You Are Made of Waste

Searching for the ultimate example of recycling? Look in the mirror

BY CURT STAGER

YOU MAY THINK OF YOURSELF as a highly refined and sophisticated creature—and you are. But you are also full of discarded, rejected, and recycled atomic elements. Don't worry, though—so is almost everyone and everything else.

Carbon: Your inky nails

Look at one of your fingernails. Carbon makes up half of its mass, and roughly 1 in 8 of those carbon atoms recently emerged from a chimney or a tailpipe. Coal-fired power plants, petroleum-guzzling cars, and kitchen gas stoves release carbon dioxide into the atmosphere. Each of those waste molecules is a carbon atom borne on two atomic wings of oxygen. Fossil-based carbon dioxide molecules that are not soaked up by the oceans or stranded in the upper atmosphere are eventually captured by plants, shorn of their oxygen wings, and woven into botanical sugars and starches. Eventually, some of them end up in bread, sweets, and vegetables, while others help form carbon-rich animal tissues, finding their way into meat and dairy products. Historically, atmospheric carbon dioxide was mainly replenished by volcanoes, forest fires, and biotic respiration. Today, one quarter of atmospheric CO_2 is the result of fossil fuel combustion, whether it rose from smokestacks or was displaced from the oceans. (When fossil-fuel CO_2 dissolves into ocean water, it displaces already-dissolved carbon dioxide derived from natural sources.) And because all of the carbon in your body derives from ingested organic matter, which in turn obtains it from the atmosphere, your fingernails and the rest of the organic matter in your body are built, in part, from emissions.

ILLUSTRATIONS BY YUKO SHIMIZU

Radioactive Carbon-14: Your pearly whites

When you smile, the gleam of your teeth obscures a slight glow from radioactive waste. During the late 1950s and early 1960s, atmospheric testing of thermonuclear weapons scattered so much radioactive carbon-14 into the atmosphere that it contaminated virtually every ecosystem and human. Several thousand unstable radiocarbon atoms explode within and among your cells every second as their unstable nuclei undergo spontaneous radioactive decay. Some are the natural products of cosmic rays that can turn atmospheric nitrogen into carbon-14, while others result from the decay of unstable mineral elements that are found in soil. But many of them represent the echoes of thermonuclear airbursts from the Cold War, finding their way into our water supply and meals. If they happen to disintegrate within your DNA, they can damage your genes. And many of them are bound up in your teeth. Unlike most of the atoms in your body, those embedded in your strong, stable tooth enamel have been with you ever since you ingested them through your umbilical cord and your infant feeding. If you were born during the early 1960s, you have more nuclear waste in your teeth than if you were born later, when soils and oceans had had time to bury radioactive atoms. In fact, forensic scientists use the proportion of bomb carbon in tooth enamel to determine the age of unidentified human remains.

Oxygen: Your leafy breath

The oxygen in your lungs and bloodstream is a highly reactive waste product generated by vegetation and microbes. Trees, herbs, algae, and blue-green bacteria split oxygen atoms out of water molecules during photosynthesis. They use most of the resultant gas for their own purposes, but thankfully some leaks out to sustain you. In fact it makes up about a fifth of the air you breathe. Your cells harness oxygen to release energy from chemical bonds in the food you consume. Oxygen absorbs electrons released by broken food molecules, which attract hydrogen ions, resulting in a molecular waste of your own making: metabolic water, which comprises one tenth of your body fluids. An average adult carries between 8 and 10 pounds of homemade wastewater within them, and 1 in 10 of your tears are the metabolic by-products of your breathing and eating.

Nitrogen: Your natural curls

The next time you brush your hair, think of the nitrogenous waste that helped create it. All of your proteins, including hair keratin, contain formerly airborne nitrogen atoms. But the nitrogen in air is biologically inert. For nitrogen to become a component of your hair, it has to be converted into a more accessible form. Nitrogen-fixing bacteria is one way that can happen. They live among the roots of beans, peas, and other legumes, consuming atmospheric nitrogen and releasing it as ammonia, a kind of microbial manure that fertilizes soil in which plants grow. When you eat a plant, you consume formerly atmospheric nitrogen. Every flash of lightning and every automotive spark plug emits a puff of nitrogen oxide, which can dissolve into raindrops and fall to earth as a form of fertilizer, again finding its way into food webs through plants. But most of the nitrogen in modern foods comes from urea and ammonium nitrate fertilizers artificially fixed by industrial processes. In ages past, the nitrogen in human hair came mainly from bacterial waste and lightning. But today, unless you eat a strictly organic diet, you run your hairbrush through nitrogenous frameworks that are mostly of human origin.

Iron: Your ancient blood

When you cut yourself, the wreckage of stars spills out. Every atom of iron in your blood, which helps your heart shuttle oxygen from your lungs to your cells, once helped destroy a massive star. The fierce nuclear fusion reactions that set stars ablaze create the atomic elements of life. As the star ages, it fuses progressively larger elements, such as silicon, sulfur, and calcium. Eventually, iron atoms are fused. The problem is that iron fusion consumes as much energy as it produces, so it weakens the star. If the star is big enough, it will collapse in on itself, its outer layers rebounding against the dense inner core, and a supernova explosion will result. The blast sprays out iron at supersonic speeds, filling great swathes of space with debris that can form new solar systems. The iron in your frying pan, house keys, and blood is essentially cosmic shrapnel from the tremendous explosions that ripped through our galaxy billions of years ago. The same blasts also released carbon, nitrogen, oxygen, and other elements of life, which later produced the sun, the Earth, and eventually—you.

---

curt stager is an ecologist and climate scientist at Paul Smith's College. He is the author of *Deep Future: The Next 100,000 Years of Life On Earth*, and also co-hosts a weekly science program on North Country Public Radio.

Frack 'er Up

Natural gas is shaking up the search for green gasoline

BY DAVID BIELLO

I AM SPEEDING DOWN New Jersey's highways, propelled by gasoline with a dash of ethanol, an alcoholic biofuel brewed from stewed corn kernels. As I drive through the outskirts of the township of Hillsborough, in the center of the state, I see that spring has brought with it a bounty of similar "biomass," as the fuel industry likes to call plants. Trees line the road and fresh-cut grass covers the sidewalks as I pull into the business park that is home to Primus Green Energy—a company that has been touting a technology to transform such biomass into a green and renewable form of gasoline.

But there's a hitch. The boom in hydraulic fracturing, or "fracking," a technique in which horizontal drilling and high-pressure jets of water are deployed to release gas trapped in sedimentary shale rock, has made natural gas cheap and plentiful. That's not bad for Primus, whose technology can make gasoline from natural gas, biomass, or even low-grade coal, such as lignite or peat. This versatility makes Primus a potential part of what has been called the "olive economy"—companies that are neither bright green nor darkest black, but combine environmentally-friendlier technologies with older and dirtier ones in order to compete. In fact, Primus may become a leader in advancing this kind of technology. "We can be as dark as you want or as green as you want," says geologist, serial entrepreneur, and Primus salesman George Boyajian.

In July, President Barack Obama gave a major speech on climate change that described natural gas as a "transition fuel" towards the "even cleaner energy economy of the future." But Primus's trajectory raises the question of whether natural gas is a boost on the road to a genuinely green fuel, or if it is prolonging our addiction to dirty modes of transport, and taking us on a detour from a low-carbon path.

At the Primus headquarters, I first meet Primus's chief chemist Howard Fang in front of a prototype of a Primus conversion machine. Fang, who joined the company for what he calls his "semi-retirement," is

ILLUSTRATION BY PETER & MARIA HOEY

avuncular and black-haired. His interests are broad: He spends his spare time writing and reading history, and has authored books on conflict in the Middle East and the role of Christian missionaries in China.

A lifetime in fuels chemistry left Fang with one burning question: "What is the real solution to the energy crisis?" His career at oil companies BP and ExxonMobil, and engine manufacturer Cummins, spanned not just one but two major energy upheavals—the oil crisis of the 1970s and then its sequel in the first decade of the 21st century, which is arguably still ongoing. These experiences impressed on Fang the importance of securing the fuel supply in such a way as to avoid despoiling the environment. The solution, says the bespectacled chemist, is "nature-sourced biomass or natural gas converted effectively to gas or diesel."

Primus's original idea was simple: take scrap wood or other biomass, turn it into pellets, and apply pressure and heat (700 degrees Celsius or more) to break it down into hydrogen and carbon monoxide. Then build this composite "syngas," shorthand for "synthetic gas," back up into whatever hydrocarbon product is desired—the molecules of eight carbon and 18 hydrogen atoms known as iso-octane that are a measure of the quality of conventional gasoline, or the longer chains of similar hydrocarbons that comprise diesel or jet fuel. Because plant biomass absorbs carbon dioxide as it grows, the emissions produced by burning the biofuel should balance out overall—every molecule of CO2 emitted when the fuel is burned was previously absorbed by the plant that made the fuel.

The story of the search for such green fuel is littered with disappointments, however. Major companies brew ethanol in large quantities in the United States. It is routinely added to gasoline (at levels of around 10 percent, on its way to 15 percent) as a way to improve combustion, reduce pollution, and support industrial corn farmers. But most ethanol is still made from the edible kernels of corn plants, instead of the inedible cellulose that was promised in the heady days of the mid-2000s, when Congress passed a spate of laws promoting biofuel production. Since 1978, the ethanol industry has enjoyed subsidies and tax credits to the order of 40 cents per gallon, and now produces an annual dead zone at the mouth of the Mississippi River each summer as a result of fertilizer washing off the endless cornfields of the Midwest. But ethanol is unlikely to ever fully replace conventional fossil fuels, since it is more difficult to transport, produces a fraction of the energy of oil, and would require engines to be refitted or replaced on a massive scale.

Hence the interest in "drop-in" biofuels as a substitute for conventional fuels in existing cars, planes, and trucks. The problem is not one of infrastructure, but chemistry: Companies must find a way to economically imitate and fast-track a process for which time and geology have done most of the work in conventional fossil fuels. The energy in these fuels is the pent-up power of ancient sunlight, which billions of photosynthetic microorganisms soaked up before dying, fossilizing, and turning into the hydrocarbon-rich stew we know as petroleum, and from which we refine gas, diesel, and jet fuel, among other products. In theory, then, it should be possible to turn the carbohydrates and other chemicals that store energy for today's living things into the hydrocarbons we rely on for transportation.

Potential routes to such "green crude" include algae, other photosynthetic organisms, and specialty microbes engineered to spit out hydrocarbons. Biofuel company Solazyme has a contract to supply United Airlines with 20 million gallons of algal jet fuel, and teamed up with a green fuel-station network to offer biodiesel in a test run in San Francisco's Bay Area. But it takes a lot of water—and a lot of energy to move that water around—in order to grow algae in large quantities, and tailor-making microbes is expensive at its current scale. As a result, companies are diversifying. Algal fuel producer Sapphire Energy is now focusing on isolating the genetic traits in the ancestors of all plants that might be usefully incorporated into other crops. Solazyme is making oils and specialty fats to sell at high margins to cosmetics and food companies, as is would-be microbial fuel-maker Amyris. The industry for "advanced biofuels is literally in its infancy," concedes Jonathan Wolfson, Solazyme CEO.

The allure of Primus's technology is its promise to harness waste wood and other inedible biomass that would otherwise be thrown into landfills, and turn it into a renewable source of gasoline. Its "syngas to gasoline plus" process consists, essentially, of four

"We can be as dark as you want or as green as you want," says Boyajian.

chemical reactors. One turns the syngas into methanol. The next makes methanol into a molecule known as dimethyl ether, or DME in chemist-speak. In the third reactor, catalysts known as zeolites knit DME into gasoline, in the most expensive and energy-intensive part of the process. The fourth reactor eliminates some of the unwanted byproducts that cause the resulting fuel to congeal at low temperatures.

The key is the zeolites, porous minerals made up of aluminum, silicon, and oxygen that allow the desired chemical reactions to take place. Both Primus and a conventional oil refinery employ zeolites to manipulate hydrocarbons. At an oil refinery, these catalysts help crack and sort hydrocarbons broken down from crude oil. At Primus, heat and pressure allow zeolites to build gasoline hydrocarbons from the smaller molecules of syngas. Such "catalysts are a bit of a dark art," says Boyajian. He spars with Fang over whether or not the company will one day make their own. Fang does not accept Boyajian's need for secrecy, and would be more than happy to reveal all those dark arts—a prospect that makes the affable Boyajian nervous and tight-lipped. For now, the fledgling company buys the necessary catalysts off the shelf and must sign agreements not to examine these zeolites too closely.

Using different catalysts in the reactors, Fang notes, the company could spit out diesel or jet fuel instead of gasoline. And for every 100 kilograms of syngas, he says, Primus can make 30 kilograms of gasoline or more, using a continuous looping system within the machine that eliminates the need for wasting energy to convert gases to liquids along the way. Little red containers of Fang-made gasoline record its characteristics, scrawled on masking tape affixed to the sides: low vapor pressure, a higher-than-average octane content of around 93, and a favorable absence of sulfur or benzene. Oil prices have been rising over the last month, and are currently at more than $100 per barrel; the company estimates that its gasoline costs as little as that derived from oil at $65 per barrel—and could cost as little as $2 per gallon, or about half the price gas currently goes for at local pumps, to produce at a full-sized facility, even though such an industrial plant would require a lot of capital to build.

However, the machine Fang shows me is not running on the biomass that Fang originally tested: wood chips, switchgrass, canary grass, miscanthus. Instead, it churns through natural gas, turning methane into syngas. Making long hydrocarbons from the single carbon in methane molecules is "very easy," he assures me. But "natural gas is not true green," he concedes. "There is no benefit in [the reduction of] greenhouse gases. Biomass is still true green."

Natural gas from the fracking boom has revolutionized the global energy landscape—particularly in the United States, the world's biggest producer of shale gas. But it is also controversial. Gas burns cleaner, but it still produces around half the greenhouse emissions of its dirtier cousins like coal, not including the excess methane that leaks from fracking sites and the pipelines that transport the gas. Fracked gas can also contaminate groundwater supplies. And while in 2012 it brought America's carbon footprint down to its lowest level in 20 years, relying on it in the long-term will make it hard to eliminate greenhouse gas emissions, as is required to combat climate change.

As the price of natural gas slid in response to the glut of shale gas, Primus changed gears in mid-2012 to move away from biomass and to focus on making syngas from natural gas. This is not a new idea: ExxonMobil built a plant in New Zealand in 1986 to turn natural gas into methanol and then gasoline, but abandoned its efforts when the price of petroleum dropped dramatically in the mid 1990s. Now, though, natural gas is cheap and attractive. Boyajian has a map of all the shale formations in North America tacked to the wall of his office. "The world is full of shale," he notes.

An earlier version of Primus' machine, tuned to process biomass, sits swathed in silvery insulating tape in a locked and darkened lab. "Right now it is abandoned," Fang says. The company insists that the statement doesn't apply to Primus's biomass efforts more generally. "This is the way to get to biofuels," says Primus CEO Robert Johnsen, of the gas to gasoline process, through a tight smile. "Will we be the ones to get there? Maybe."

The energy in these fuels is the pent-up power of ancient sunlight, which billions of photosynthetic microorganisms soaked up before dying.

Will natural gas be a bridge for Primus to green fuel, or will it be too cheap and attractive to resist as a permanent substitute for biomass? For the moment, the company seems keen to squeeze what it can out of the shale gale. With the help of more than $50 million in Israeli money, Primus is building a demonstration plant the size of a house near its headquarters in New Jersey, due to open this year. The location is off the map—even Google won't guide you there, as if it were some secretive skunk works facility, which is how the company likes to think of it. The plant will take natural gas from the local utility, run it through its proprietary set of chemical reactions and, on the far end, out of a spigot, will come gasoline—12.7 gallons per hour at full capacity. The company's first commercial plant, due to start construction next year, will likely be located near a source of natural gas.

Scaling up the technology this way will reduce the overhead costs per unit of gasoline—that is, the cost of fabricating the reactors and buying the zeolites and feedstocks. Plus, Primus' technology may prove economical enough at a scale small to allow its plants to be distributed close to remote natural gas wells or even sources of biomass. It is no coincidence that the company based itself in verdant New Jersey, "the Garden State"; proximity to biomass is crucial for producers, because transporting heavy and unwieldy wood or corn stalks across large distances tends makes the end product too costly and undercuts the greenhouse-gas savings that are a large part of its appeal.

As I prepare to drive off, Fang carts out one of his collection of red plastic gas cans and dumps a liter or so of Primus-made, natural gas-to-gasoline fuel into my tank. A test car tooled around on it last summer, with no problems. The hope is to be able to charge a premium for the higher-octane premium product. "People pay twice as much for organic food," Boyajian says. "So why not pay more for green gasoline?" My fuel sensor can tell the difference: it registers an anomalously high miles-per-gallon number.

Fang gives me two thumbs up as I pull away, watching me drive off on his preferred solution to the energy crisis. It's unclear whether Primus will ever find the occasion to turn back towards biogasoline—and whether that's a long-term fix for the world's energy and environmental conundrum. Striving to make cleaner fuel for standard, dirty combustion engines may reinforce drivers' loyalty to today's technology. Such lock-in makes a true revolution difficult until some alternative energy source—whether battery-driven electric cars or engines modified to burn carbon-neutral, as-yet-unmade biofuels—offers the kind of convenience and low cost that justifies replacement.

At present, Primus appears set to become part of a sprawling infrastructure that reinforces the incentives to use greenhouse gas-producing, gasoline-like fuels. And for all those concentrated octanes in my tank, I still have to pull into a Shell station to fill up on conventional gasoline, blended with corn ethanol, in order to drive home.

david biello is the Environment and Energy Editor for *Scientific American*. He is currently working on a book about the Anthropocene.

Genetics & Human Health

Since DNA is often heralded as the "code of life," what clues can mutations—changes to the DNA sequence—tell us about human health and disease? The pair of articles in the Genetics and Human Health module will explore the consequences of mutations in the context of cancer treatment and rare diseases such as muscular dystrophy. *Their Giant Steps to a Cure* discusses the challenges associated with treating a rare form of muscular dystrophy. *An Unlikely Cure Signals Hope for Cancer* explores how specific mutations in a patient's cancer can be used to a patient's advantage.

Lesson Plan

Ask students to read both of the articles for homework. Briefly introduce or review the vocabulary words in class. Assign the questions listed under "Reading Comprehension" for them to complete along with the reading and ask them to come up with one question for further discussion.

Start class by asking students if they have any questions about the readings. Ask them to contribute their discussion questions (in addition to the ones provided under Deep Thinking / Discussion questions). Have the class brainstorm and answer both discussion questions. **15 MIN**.

Next, break the class up into four groups for the Suggested Activity. Assign each group to one protein that is listed in the interactive. **15 MIN**.

Have each group present their thoughts to the class for further discussion. **15 MIN**.

Teacher's Notes: Their Giant Steps to a Cure, and An Unlikely Cure Signals Hope for Cancer

VOCAB WORDS

Muscular dystrophy: a genetic disease marked by progressive weakening of the muscles. Some forms of muscular dystrophy are seen in infancy or childhood.

Orphan diseases: diseases that have yet to be "adopted" by the pharmaceutical industry because there are very few incentives to develop new medications to treat or prevent them. Orphan diseases can be rare or they are common diseases that have been ignored (e.g.: tuberculosis, cholera, typhoid, malaria).

Calpainopathy: a rare type of muscular dystrophy characterized by symmetric and progressive weakness of proximal (limb-girdle) muscles.

Cancer: A term used to describe disease in which abnormal cells divide without control and are able to invade into other tissues. Cancers are often categorized based on the organ or cell type they originate in.

Oncologist: A doctor who specializes in treating patients with cancer.

Outlier: An observation that deviates from a majority and can be seen to be a rare event. In the context of this piece, the outliers are patients who respond to therapy when the same therapy has failed other patients.

Remission: a decline or disappearance of signs and symptoms of cancer.

READING COMPREHENSION

1. Why are orphan diseases underfunded?
2. How does the mutation in calpain 3 cause muscle to fail to grow?
3. What are some reasons pharmaceutical companies would want to develop drugs for orphan diseases? What are some possible reasons they would be against doing so?
4. Statistically speaking, outliers are often ignored. In this story, why is patient number 45 such an interesting case? Why is it generally important to study the outliers of response?
5. Which protein's activity is blocked by everolimus? What is the function of this particular protein?

DISCUSSION QUESTIONS

1. In what contexts would it be desirable and undesirable to sequence your genome to see if you are at risk for a disease? What are the benefits and downsides of knowing if you are at risk for a particular disease?
2. In both pieces, mutations are responsible for causing disease. Compare and contrast the ways mutations can lead to muscular dystrophy and cancer. Are the mutations in one case hereditary? Are mutations leading to either disease caused by environmental factors? Are the mutations in either case preventable? If so, how could they be prevented?
3. How should doctors and scientists decide whether to work on a rare condition?

ACTIVITIES

Some genes are not specific to humans, but rather, are common to myriad species. In a smaller group, you will be assigned to read about one of the proteins listed here: http://nautil.us/issue/5/fame/genes-that-won-the-fame-game

Please answer the following questions when it is your turn to present to the class:

1. What organisms is the gene present in? Were you surprised by the presence of the gene in any of the organisms listed? If so, why?
2. If this protein was mutated, what could the consequences look like? Could it cause a disease?
3. Research and present one other case of an outlier being useful in science or medicine.

WHERE THIS FITS IN THE CURRICULUM

Structure and Function (HS-LS1-1) A cell contains genetic information in the form of DNA molecules. Genes are regions in the DNA that contain the instructions that code for the formation of proteins, which carry out most of the work of cells.

Variation of Traits (HS-LS3-2) Although DNA replication is tightly regulated and remarkably accurate, errors do occur and result in mutations, which are also a source of genetic variation. Mutations can, in turn, cause disease and/or affect human health. The pattern of mutations can also predict response to drugs.

Inheritance and Variation of Traits - Environmental Factors (HS-LS3-3) Technological advances have influenced the progress of science and science has influenced advances in technology. Technologies have evolved to sequence human genes, which can better inform doctors of their patients' health. Likewise, pharmaceutical companies have also created many drugs for the treatment of human disease.

Their Giant Steps to a Cure

Battling a rare form of muscular dystrophy, a family finds an activist leader, and hope

BY JUDE ISABELLA

IN 2007, AT HER high school graduation in Quesnel, British Columbia, Ivana Topic stood at the top of the auditorium stairs, her long gown skimming the floor, her dark brown hair spilling over her shoulders. She had on ridiculously high heels. As she eased down the stairs, very slowly, she hung on to her date. She was afraid her knees would collapse, as her muscles were weak for her age.

From the audience, Ivana's mother, Marijana, watched her daughter's every step, silently panicking and breaking into a sweat. She knew Ivana could easily tumble down the stairs and break a limb. The year before, Ivana had been diagnosed with muscular dystrophy, an incurable genetic disease characterized by progressive weakening of the muscles. Antonia, Ivana's younger sister by five years, was later diagnosed with the same disease.

Around the time of Ivana's graduation, the Topics, an unassuming family originally from Croatia, had begun adjusting their lives as best they could, inquiring about ramps everywhere they went, avoiding walking in snow and sleet. For years, Ivana and Antonia had been subjected to endless medical tests. In 2010, they learned they had a rare form of muscular dystrophy, calpainopathy, which affects about 1 in 200,000 people. The diagnosis meant both would likely be bound to wheelchairs while they were still young women.

Today, Ivana is 24. In May, she graduated from college with a bachelor's degree in finance and general business. She still walks up stairs in her house; her bedroom is upstairs. "I'm definitely a fighter, and will try and walk for as long as I can," she says. "When I notice I'm falling a lot, when I need help a lot, I will go in a chair."

Muscular dystrophy treatment is limited to only palliative medications and therapies. Ivana herself practices yoga. While researchers worldwide are working on lasting cures for muscular dystrophy (funded in part by the famous Jerry Lewis Telethons), rare forms like calpainopathy are "orphans," with only a fraction of

ILLUSTRATIONS BY ELLEN WEINSTEIN

"I'm definitely a fighter, and will try and walk for as long as I can."

researchers and funds devoted to them. With quiet stoicism, the Topics have accepted that modern medicine may not have a solution for their daughters' disease. Still, says Marijana, "Without hope, there's no life."

Following a current grassroots trend in medicine, many individuals with orphan diseases do not wait for the medical industry to care about them. Facing long odds, they are forced to raise money to find a potential cure themselves. But the Topics live by modest means. Marijana runs a daycare center and her husband and the childrens' father, Niko, works for a lumber company. They are in no position to mount a quest.

But then there's Michele Wrubel, 49, a stay-at-home parent from Connecticut who has calpainopathy. For years, Wrubel has been a passionate crusader for a cure. Affluent and well connected, she doesn't varnish the truth about what it has taken to make the medical industry pay attention to her. "To make a difference in this disease, you need money and meetings," she says. "Researchers are not going to study a disease unless there's money behind it to fund the research." For the Topics, Wrubel may be their best hope.

THE GLOBAL GENES PROJECT, an advocacy group, estimates 350 million people suffer from orphan diseases worldwide. Most rare diseases are genetic and tend to appear early in life. About 30 percent of children who have them die before reaching their fifth birthday. The rest battle their conditions throughout life, as most orphan diseases have no cure. Out of the 7,000 orphan diseases identified to date, with about 250 new ones added annually, less than 400 can be treated therapeutically.

This year the European Commission gave 144 million euros to develop 200 new therapies and the National Institute of Health allocated $3.5 billion to research orphan diseases. Yet some diseases are so rare that they remain stepchildren even among orphans. As a result, they receive little research attention and funding. Neither do they fit the list of billable insurance procedures. There's no standard healthcare path to diagnosis, let alone treatment. Similar to the Topics, many patients go through an ordeal, which Marijana describes as "a blur," only to find out that medicine can't help them.

Orphan disease organizations, such as the National Organization for Rare Disorders and the Rare Disease Foundation, encourage patients to take matters into their own hands. "Families have to advocate," says Isabel Jordan, chair of the Rare Disease Foundation. She encourages patients to form organizations, find new methods of funding, and push for research.

"Push for research" could be Michele Wrubel's calling card. She was diagnosed with muscular dystrophy in her mid-20s. But even though calpainopathy was identified nearly 20 years ago—about the same time Wrubel got her initial diagnosis it took almost the entire second half of her life to determine that she was afflicted with calpainopathy. There were no clinical procedures that would lead to a diagnosis.

"It took a really long time and a very concerted effort," says Wrubel, who walks with canes, submitting to a wheelchair for long trips or when in crowded places. "If you don't know what you're looking for, they don't know what to tell you or how to help you," she says.

In 2008, gene sequencing came of age, which aided physicians in diagnosing muscular dystrophy subtypes. That year, Wrubel's husband, Lee, who holds a medical degree and a master's in public health from Tufts, an MBA from Columbia University, and is a venture capitalist in the medical field, tracked down a neurologist

In the quest for a cure, she says, "It's a matter of patients taking charge of their diagnosis."

to sequence his wife's genomes. He paid several thousand dollars from his own pocket to learn his wife had calpainopathy.

The Topics had no such luxury. But they did have luck. Cornelius Boerkoel, a clinical geneticist at the University of British Columbia, enrolled the Topics in one of his studies, and so they didn't have to pay to have each of the family member's genomes sequenced. The genome tests gave Ivana and Antonia the bad news about calpainopathy. Their younger brother, Mario, is free of the disease.

Scientists classify calpainopathy, or "calpain," as a limb-girdle muscular dystrophy Type 2a, caused by a mutation in the gene calpain 3, predominantly expressed in skeletal muscle. Those who suffer from Type 2a, such as Wrubel, Ivana, and Antonia, generally exhibit weak hip flexors muscles that lift up the thigh. The weak flexors give them an awkward gait; they swing their legs forward, landing on their toes, and then sometimes on the sides or soles of their feet. Some walk only on the balls of their feet. The upper body muscle weakness creates abnormally prominent shoulder blades.

Melissa Spencer from the University of California, Los Angeles, who has studied calpainopathy for 14 years, explains that the disease contains many subtypes. The problem with Type 2a, she says, "was a really strange gene mutation that was completely inexplicable." She says it has been a hard disease to study, partially because the implicated protein is unstable and partially because it was a rarity among the orphan diseases. When it comes to funding, calpainopathy has been overshadowed by other forms of muscular dystrophy. "Muscle studies have been underfunded forever and certainly a rare disease like 2a especially underfunded," Spencer says.

In 2010, Wrubel formed the nonprofit Coalition to Cure Calpain 3. In the quest for a cure, she says, "It's a matter of patients taking charge of their diagnosis." She reached out to other sufferers via Facebook, and some donated money. She partnered up with two other nonprofits that had raised funds on their own, both started by those afflicted with Type 2a. So far Wrubel's efforts have gathered close to half a million dollars. With that money, she has funded a project with Louis Kunkel, professor of genetics and pediatrics at Boston Children's Hospital, one of the nation's key muscular dystrophy researchers.

Her coalition also organized a conference to bring calpainopathy researchers together, including Spencer. Years earlier, in 2005, Spencer made a significant breakthrough. She discovered that calpainopathy, unlike more common forms of muscular dystrophy, was not a weakening of the muscle but a growth problem muscle forms, but fails to grow because of a missing protein. It is different from other muscular dystrophies in which the lack of the protein complex, dystrophin, damages muscle membranes. "With calpainopathy, the muscles lack the growth signal," she says. "It's not transmitted properly." That difference makes a drug cure more possible. "I think this is going to be the easiest muscular dystrophy to cure," she says.

Encouraged by the promise, the Coalition to Cure Calpain 3 gave Spencer's lab a $260,000 grant to investigate how to circumvent the signaling problem and come up with a drug to fix it. But because the United States Food and Drug Administration already has a library of approved compounds that stimulate cell growth in muscle, Spencer's team may arrive at a solution sooner. With the help of the coalition's money, her lab is now plowing through the thousands of existing compounds, choosing those fit for testing. "I think it will be five years before we start thinking about clinical trials," Spencer says—and then another five years before the drugs can be commercially available, she estimates.

Wrubel's coalition intends to get pharmaceutical companies interested, too. "Many pharmaceutical companies see treating orphan diseases as a way to increase profits," Wrubel says. Her husband, Lee, adds, "The whole model for big pharmaceutical companies going forward is different. There is too little in the big pharmaceutical pipeline, and they're looking to feed that beast as much as possible." A 2012 Thomson Reuters study found that drug companies stand to profit from orphan drugs because, compared to drugs for common afflictions, they often have shorter and less expensive clinical trials, with more success. Spencer says a drug for calpainopathy, for instance, would also be useful for patients with Lou Gehrig's Disease and bed rest patients, as it would help arrest the loss of bone and muscle mass. Wrubel hopes to bring Cydan Development, a venture-capital backed orphan drug developer, to their upcoming fall conference in the Netherlands.

As for the Topics, they were excited to learn about Wrubel from *Nautilus*. Ivana recently connected with Wrubel through Facebook. "I only talked with her a little bit, but she seems ambitious and driven," Ivana says. "Definitely not someone who is going to sit around and wait for something to happen. Definitely inspiring. And the possibility that something might help in any way is a good thing to hear, for sure." Ivana says she now wants to get involved and advocate for her own disease. "I definitely want to do something," she says, and Wrubel's coalition "would be a good place to start."

jude isabella is a science writer based in Victoria, British Columbia. Her new book, *Salmon, A Scientific Memoir*, will be released next year.

An Unlikely Cure Signals New Hope for Cancer

How "exceptional responders" are revolutionizing treatment for the deadly desease

BY KAT MCGOWAN

JUST LIKE EVERY NEW drug the oncologists at Memorial Sloan-Kettering Cancer Center tested against bladder cancer in the last 20 years, this one didn't seem to be doing any good. Forty-four people in the study were given everolimus in a last-ditch attempt to slow down or stop their advanced cancer. When the researchers analyzed the data, they could see that the drug wasn't slowing or stopping tumor growth. Everolimus seemed to be another bust.

Then there was patient number 45. She joined the trial with advanced metastatic cancer. Tumors had invaded deep into her abdomen, clouding her CT scan with solid grey blotches. She was 73 years old. None of the standard bladder cancer drugs were working for her anymore; she had "failed treatment," in the dismal lingo of oncologists. She enrolled in the study only because she happened to be a patient at Sloan-Kettering in January 2010. In April 2010, her cancer was gone.

This sort of happy surprise is not unheard of in drug studies. Bodies are fluky, each with its own idiosyncratic combination of genetic blueprints and environmental inputs. So sometimes a patient will be cured by a drug that is useless for everyone else. In the past, these spectacular reactions were written off as outlier responses that defied explanation—medical mysteries. Doctors just shrugged their shoulders and thanked their lucky stars that even though the study tanked, they did manage to help one person.

But this time was different. Clinical oncologist David Solit, director of developmental therapeutics at Sloan-Kettering, saw a new opportunity to explain what happened by sequencing the whole genome of the woman's cancer. Just five years ago, decoding and analyzing all 3 billion bases of the DNA from a tumor would've been absurdly time-consuming and expensive. Now the sequencing takes as little as a few days.

Poring over the outlier patient's genetic code, Solit pinpointed two mutations that made her tumor sensitive to this drug. He found that one of her mutations shows up in about 8 to 10 percent of other bladder cancer patients, meaning that they too might be helped by everolimus. His success has inspired a whole set of

ILLUSTRATION BY ELLEN WEINSTEIN

programs to study "exceptional responders": those rare cancer patients who do well while nobody else does.

Cancer is a personal disease, Solit explains. Each tumor constitutes its own world of defective genes and proteins. By studying the genetic quirks of exceptional responders, physicians can systematically identify weaknesses in cancer subtypes and blast them with drugs that target their unique vulnerabilities. "It's a testament to how much has been learned about the genome in the past 30 years," Solit says. "We've always wanted to find out why some individuals respond so well. Now we have the capacity. It's going to really change the way we treat patients."

UNLIKELY CASES HAVE AN eminent history in medicine. The modern science of the mind owes a lot to the freakish accident suffered by Phineas Gage, a 19th century railroad construction foreman whose job involved packing down explosive powder with a three-and-a-half-foot-long iron tamping rod. On Sept. 13, 1848, the powder exploded in his face, blasting the rod up through his chin and out the back of his head. Against all odds, he survived. But his personality was transformed. The formerly shrewd and patient Gage became obnoxious and unreliable.

An observant doctor named John Martyn Harlow who cared for Gage proposed that his personality change was due to the destruction of the frontal lobe of the left side of the brain. Gage's unlikely transformation revealed a universal truth about brains, that particular parts—the frontal lobes—are required for self-control. The strange case of Phineas Gage is still mentioned in neuroscience textbooks.

Rare events can also lead to new cures. As the story goes, English physician Edward Jenner's observations of an 18th century milkmaid who caught cowpox and thereby became immune to smallpox paved the way for the fi vaccines. New ideas for curing HIV are emerging from the famously unlucky lucky case of the "Berlin patient." Timothy Ray Brown, who was HIV positive, developed blood cancer leukemia in 2006. His chemotherapy and radiation treatments wiped out the cells of his immune system, where the virus is believed to hide. He then got a bone marrow transplant from one of those rare people with a gene mutation that makes them resistant to HIV. Today, Brown still has no sign of HIV in his body, and his case has inspired a study to genetically engineer HIV-positive patients' cells to resist the virus.

In the past, cancer researchers weren't able to capitalize on their unexpected outlier successes. Not enough was known about the biology of cancer, and the right tools hadn't been invented. "Even if someone had a complete remission, you had no way to figure out why," says James Doroshow, director of the Division of Cancer Treatment and Diagnosis of the National Cancer Institute (NCI). That changed in the 2000s, when it became possible to analyze the genetics of cancer tumors for clues.

The first major success came with studies of the drug gefitinib in non-small-cell lung cancer (the most common kind). Gefitinib helped less than 20 percent of the people who took it, but a few outliers had dramatic, rapid recoveries. In 2004, two Harvard groups found that the responders had mutations in the epidermal growth factor receptor (EGFR) gene. EGFR is one of many genes that regulates how cells grow and when they die, and the mutation basically forced it to pump out two or three times as much growth signal as it should, fueling the cancer. Gefitinib dialed down the signal. A clinical trial later proved that the drug keeps tumors at bay for more than nine months in people with certain EGFR mutations.

More insights gleaned from extraordinary responders soon followed. One melanoma patient in a study of 22 people taking sorafenib saw his tumor shrink quickly, a response due to a mutation in the gene KIT, which regulates cell growth, division and survival. People with certain kinds of melanoma, such as the type that grows on mucus membranes, now routinely get tested for this mutation. The drug helps about 40 percent of those with the mutation—an impressive advance in a cancer that once had no effective treatment.

In these studies, investigators had to make educated guesses about where in the genome to look for the culprit mutations. It was the keys-under-the-lamppost phenomenon: They could only examine genes they already suspected were involved in the cancer. But as the speed and efficiency of DNA sequencing skyrocketed, and its price plummeted, it started to look reasonable to sequence the whole tumor genome to cast the widest possible net. By 2010, when the bladder cancer patient (who doesn't want her name made public) had such a wonderful response to everolimus, the technology was ripe to analyze her entire tumor.

The outlier patient had already gone through several rounds of treatment, including surgery at Memorial Sloan-Kettering. That was another stroke of luck because it allowed Solit's group to acquire samples of her tissue to be sequenced. Cancers typically start with mutations that cause cells to divide too much, ignoring normal stop signals and evading quality controls that repair or prevent errors in DNA reproduction. "Cancer is a disease of mutations," says Solit.

The outlier patient's cancer had accumulated 17,136 mutations, of which 140 seemed most suspect, because they appeared in "coding" regions of the genome, the segments that include instructions on how to build the proteins that do the work in a cell. Out of those 140, two looked particularly menacing to Solit. In a gene called TSC1, just two of its 8,600 DNA base-pairs were missing, but the error would cause the gene to make a defective version of the protein it was supposed to create. In the gene NF2, an error meant a protein would be built only halfway, unable to do its job.

Solit could now see how these mutations were affected by everolimus, a drug typically used to suppress the immune system after organ transplants, and to combat advanced kidney cancer. Everolimus shuts down one crucial link in a chain of interacting proteins called the mTOR pathway that fuels cell growth, division, and survival. The drug inhibits the cells of the immune system from dividing, which they must do in order to attack foreign tissue, and protects transplanted organs. Likewise, it slows down the uncontrolled cell division that happens in cancer. The kicker was that both of the woman's mutations, NF2 and TSC1, affect the mTOR system. "It's not surprising, in retrospect, that our patient responded really well to this specific drug," Solit says. "She had the mutation that activated the pathway the drug targets."

Solit's team analyzed 13 more people from the trial and found different TSC1 mutations in three other people, including two whose tumor shrank a little in response to the drug. (Nobody else had NF2 mutations, which is probably why she alone responded dramatically.) Meanwhile, eight of nine people whose tumors grew during the study did not have the mutation.

DOROSHOW OF THE National Cancer Institute says Solit's work "turned on the lightbulb." It showed how the analysis of exceptional responders could be made systematic. Inspired by his example, the NCI is now trawling through its own archives, revisiting outlier responses among the roughly 10,000 patients who enrolled in NCI-sponsored clinical trials during the last decade. Picture the long rows of crates in the government warehouse at the end of *Raiders of the Lost Ark*: There's treasure in there somewhere, if only someone would look. "We ought to study these people more, since we have the means now," says Barbara Conley, the associate director of the cancer diagnosis program at NCI, who leads the project.

In the few months since the project began, Conley's team have already found about 100 exceptional responders. The next steps are to find out if their tumors were biopsied, if that tissue sample is still sitting in a freezer somewhere, and whether it's in good enough shape to be sequenced. Starting next year, the group will start inviting any scientist who is doing a clinical trial to submit new cases.

The NCI project will include whole-genome sequencing (provided they have adequate tissue samples) and repeated reads of the whole "exome"—the 1 percent of human DNA that is translated into exons, the sequences that are used as templates for protein construction. The reason to do both, explains Conley, is that cancer cells, even within a single tumor, often have a hodgepodge of mutations. Re-doing whole exome sequencing dozens of times captures most of the significant genetic variation in one tumor, and it's more practical than trying to sequence the whole genome over and over. Finally, RNA expression will also be analyzed. Evaluating RNA, an intermediary between DNA and proteins, provides a measure of which genes are switched on and how much protein they're producing.

Other elite cancer research centers and genome-sequencing centers have similar in-house projects. Much like the NCI project, the unusual responder program at the University of Texas, MD Anderson Cancer Center, is beginning by combing through the archives to hunt for outliers of the past. A patient at the clinic who has an unusual response—good or bad—will also be referred for genome sequencing and other kinds of genetic analysis.

Even if each outlier case only applies to 3 or 7 percent of one type of cancer, as more cases are solved, the benefits quickly add up. "We're talking about small subsets of patients that together make a radical change," says Funda Meric-Bernstam, chair of the Department of Investigational Cancer Therapeutics at MD Anderson, who leads the unusual responders program. In some cases, existing cancer drugs can simply be repurposed, such as discovering that an immunosuppressant drug works for certain bladder cancers. Or it might mean finding new life for an experimental drug that had been abandoned. If Conley and Doroshow can pinpoint who might be helped by an abandoned drug, a pharmaceutical company might have to do just one or two further studies to get that drug approved for routine use.

The future might look something like what's been going on for several years at the Genome Institute of Washington University, where genome sequencing is being used to help people with relapsed cancers and who have run out of options. The project puts insights from studies like Solit's into practice, analyzing a patient's tumor to determine whether currently available drugs might target the troublemaker mutations. Combining whole genome sequencing, exome sequencing, and RNA expression analysis—what Washington University professor of genetics and Genome Institute co-director Elaine Mardis calls the "Maserati approach"—the team compares a comprehensive genetic profile against a database of drugs that target specific gene variants, looking for a match.

If there is a match, the results can be impressive, as was the case with a young Washington University doctor with leukemia, Lukas Wartman, who had suffered two relapses. In his case, analysis revealed that a gene called FLT3 was expressing more RNA than normal. A drug that inhibits this gene, usually used in kidney cancer, sent his cancer into remission. Washington University now has a special genetic test for patients with his type of leukemia.

Just recently, Solit's group solved another exceptional responder mystery—a case of ureteral cancer eliminated with a combination of old and new drugs. The old drug is a standard chemotherapy treatment that prevents DNA from unwinding, which it must do in order to duplicate itself during cell division. The new one sensitizes cells to the effects of radiation. This patient turned out to have a mutation in RAD50, involved in repairing broken DNA strands (badly repaired DNA can lead to uncontrolled cancerous growth). Here, too, the outlier finding may lead to a new treatment, since about 4 percent of the other tumors Solit has looked at have mutations that affect part of the RAD50 complex. "To look at these individuals' cancers can tell us a lot more than just a random case of cancer," says Solit. "There's a phenotype—a response that gives you information about the genes."

Solit is now making a quick, reliable test for the TSC1 mutation to single out people with bladder cancer who might be helped by everolimus, and is planning a new study to test the drug in them. And the original outlier, the woman with bladder cancer? Three years later, she's still on everolimus and still having a "complete response," Solit says. She's doing fine.

kat mcgowan is a contributing editor at *Discover* magazine and an independent journalist based in Berkeley, Calif., and New York City.

EDUCATION@NAUTIL. US

Game Theory and Morality

Moshe Hoffman, Erez Yoeli, and Carlos David Navarrete

Introduction

Consider the following puzzling aspects of our morality:

1. Many of us share the view that one should not use people, even if it benefits them to be used, as Kant intoned in his second formulation of the categorical imperative: "Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end" (Kant, 1997). Consider dwarf tossing, where dwarfs wearing protective padding are thrown for amusement, usually at a party or pub. It is viewed as a violation of dwarfs' basic dignity to use them as a means for amusement, even though dwarves willingly engage in the activity for economic gain. Many jurisdictions ban dwarf tossing on the grounds that the activity violates dwarfs' basic human rights, and these laws have withstood lawsuits raised by dwarfs suing over the loss of employment (!).
2. Charitable giving is considered virtuous, but little attention is paid to how just the cause or efficient the charity. For example, Jewish and Christian traditions advocate giving 10 % of one's income to charity, but make no mention of the importance of evaluating the cause or avoiding wasteful charities. The intuition that giving to charity is a moral good regardless of efficacy results in the persistence of numerous inefficient and corrupt charities. For example, the Wishing Well Foundation has, for nearly a decade, ranked as one of CharityNavigator.

M. Hoffman (✉) • E. Yoeli
Program for Evolutionary Dynamics, Harvard University,
One Brattle Square, Suite 6, Cambridge, MA 02138, USA
e mail: moshehoffman@fas.harvard.edu

C.D. Navarrete
Department of Psychology, and, the Ecology, Evolutionary Biology and Behavior Program,
Michigan State University, East Lansing, MI, USA

T.K. Shackelford, R.D. Hansen (eds.), *The Evolution of Morality*,
Evolutionary Psychology, DOI 10.1007/978-3-319-19671-8_14

com's most inefficient charities. Yet its mission of fulfilling wishes by children with terminal illnesses is identical to that of the more efficient Make-A-Wish Foundation. Worse yet, scams masquerading as charities persist. One man operating as The US Navy Veteran's Association collected over 100 million dollars—over 7 years!—before anyone bothered to investigate the charity.

3. In every culture and age, injunctions against murder have existed. If there is one thing much of humanity seems to agree on, it's that ending the life of another without just cause which is among the worst of moral violations. Yet cultures don't consider the loss of useful life years in their definition, even though it is relevant to the measure of harm done by the murder. Why is our morality so much more sensitive to *whether* a life was lost than to how much life was lost?

There are numerous other examples of how our moral intuitions appear to be rife with logical inconsistencies. In this chapter, we use game theory to provide insight on a range of moral puzzles similar to the puzzles described above.

What Is Game Theory and Why Is It Relevant?

In this section, we review the definition of a game, and of a Nash equilibrium, then discuss how evolution and learning processes would yield moral intuitions consistent with Nash equilibria.

Game theory is a tool for the analysis of social interactions. In a game, the *payoff* to each *player* depends on their actions, as well as the actions of others. Consider the Prisoner's Dilemma (Chammah & Rapoport, 1965; see Fig. 1), a model that captures the paradox of cooperation. Each of two players chooses whether to cooperate or to defect. Cooperating reduces a player's payoff by $c > 0$ while increasing the other's payoffs by $b > c$. Players could be vampire bats with the option of sharing blood, or firms with the option of letting each other use their databases, or premed students deciding whether to take the time to help one another to study. The payoffs, b and c, may represent likelihood of surviving and leaving offspring, profits, or chance of getting into a good medical school.

Solutions to such games are analyzed using the concept of a Nash equilibrium[1]—a specification of each player's action such that no player can increase his payoff by deviating unilaterally. In the Prisoner's Dilemma, the only Nash equilibrium is for neither player to cooperate, since regardless of what the other player does, cooperation reduces one's own payoff.

[1] Note that we focus on the concept of Nash equilibrium in this chapter and not evolutionary stable strategy (ESS), a refinement of Nash that might be more familiar to an evolutionary audience. ESS are the Nash equilibria that are most relevant in evolutionary contexts. However, ESS is not well defined in many of our games, so we will focus on the insights garnered from Nash and directly discuss evolutionary dynamics when appropriate.

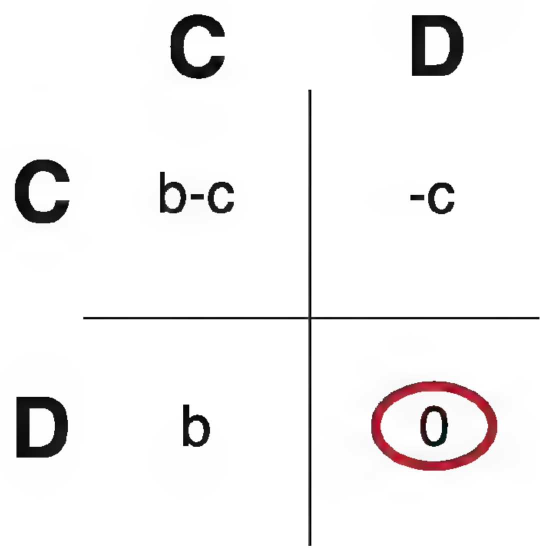


Fig. 1 The Prisoner's Dilemma. Player 1's available strategies (C and D, which stand for cooperate and defect, respectively) are represented as *rows*. Player 2's available strategies (also C and D) are represented as *columns*. Player 1's payoffs are represented at the intersection of each row and column. For example, if player 1 plays D and player 2 plays C, player 1's payoff is b. The Nash equilibrium of the game is (D, D). It is indicated with a *circle*

Game theory has traditionally been applied in situations where players are rational decision makers who deliberately maximize their payoffs, such as pricing decisions of firms (Tirole, 1988) or bidding in auctions (Milgrom & Weber, 1982). In these contexts, behavior is expected to be consistent with a Nash equilibrium, otherwise one of the agents—who are actively deliberating about what to do—would realize she could benefit from deviating from the prescribed strategy.

However, game theory also applies to evolutionary and learning processes, where agents do not deliberately choose their behavior in the game, but play according to strategies with which they are born, imitate, or otherwise learn. Agents play a game and then "reproduce" based on their payoffs, where reproduction represents offspring, imitation, or learning. The new generation then play the game, and so on. In such settings, if a mutant does better (mutation can be genetic or can happen when agents experiment), then she is more likely to reproduce or her behavior imitated or reinforced, causing the behavior to spread. This intuition is formalized using models of evolutionary dynamics (e.g., Nowak, 2006).

The key result for evolutionary dynamic models is that, except under extreme conditions, behavior converges to Nash equilibria. This result rests on one simple, noncontroversial assumption shared by all evolutionary dynamics: Behaviors that are relatively successful will increase in frequency. Based on this logic, game theory models have been fruitfully applied in biological contexts to explain phenomena such as animal sex ratios (Fisher, 1958), territoriality (Smith & Price, 1973), cooperation (Trivers, 1971), sexual displays (Zahavi, 1975), and parent–offspring conflict (Trivers, 1974). More recently, evolutionary dynamic models have been applied in human contexts where conscious deliberation is believed to not play an important role, such as in the adoption of religious rituals (Sosis & Alcorta, 2003), in the expression and experience of emotion (Frank, 1988; Winter, 2014), and in the use of indirect speech (Pinker, Nowak, & Lee, 2008).

Crucially for this chapter, because our behaviors are mediated by moral intuitions and ideologies, if our moral behaviors converge to Nash, so must the intuitions and ideologies that motivate them. The resulting intuitions and ideologies will bear the signature of their game theoretic origins, and this signature will lend clarity on the puzzling, counterintuitive, and otherwise hard-to-explain features of our moral intuitions, as exemplified by our motivating examples.

In order for game theory to be relevant to understanding our moral intuitions and ideologies, we need only the following simple assumption: *Moral intuitions and ideologies that lead to higher payoffs become more frequent.* This assumption can be met if moral intuitions that yield higher payoffs are held more tenaciously, are more likely to be imitated, or are genetically encoded. For example, if every time you transgress by commission you are punished, but every time you transgress by omission you are not, you will start to intuit that commission is worse than omission.

Rights and the Hawk–Dove Game

In this section we will argue that just as the Hawk–Dove model explains animal territoriality (Maynard Smith & Price, 1973, *to be reviewed shortly*), *the Hawk–Dove model sheds light onto our sense of rights* (Descioli & Karpoff, 2014; Gintis, 2007; Myerson, 2004).

Let us begin by asking the following question (Myerson, 2004): "Why [does] a passenger pay a taxi driver after getting out of the cab in a city where she is visiting for one day, not expecting to return?" If the cabby complains to the authorities, the passenger could plausibly claim that she had paid in cash. The answer, of course, is that the cabby would feel that the money the passenger withheld was his—that he had a right to be paid for his service—and get angry, perhaps making a scene or even starting a fight. Likewise, if the passenger did in fact pay, but the cabby demanded money a second time, the passenger would similarly be infuriated. This example illustrates that people have powerful intuitions regarding rightful ownership. In this section, we explore what the Hawk–Dove game can teach us about our sense of property rights.

The reader is likely familiar with the Hawk–Dove game, a model of disputes over contested resources. In the Hawk–Dove game, each player decides whether to fight over a resource or to acquiesce (i.e. play Hawk or Dove). If one fights and the other does not, the fighter gets the resource, worth v. If both fight, each pays a cost c and split the resource. That is, each gets $v/2-c$. If neither fights, they split the resource and get $v/2$. As long as $v/2<c$, then in any stable Nash equilibrium, one player fights and the other acquiesces. That is, if one player expects the other to fight, she is better off acquiescing, and vice versa (see Fig. 2).

Crucially, it is not just a Nash equilibrium for one player to always play Hawk and the other to always play Dove. It is also an equilibrium for both players to condition whether they play Hawk on an *uncorrelated asymmetry*—a cue or event that

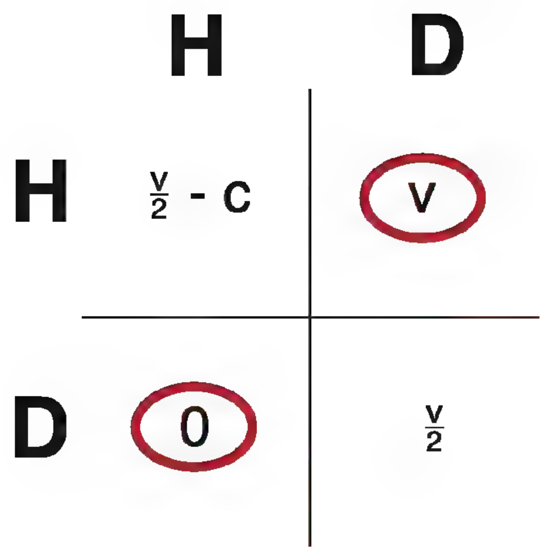


Fig. 2 The Hawk–Dove game. The Nash equilibria of the game are circled

does not necessarily affect the payoffs, but does distinguish between the players, such as who arrived at the territory first or who built the object. If one conditions on the event (say, plays Hawk when she arrives first), then it is optimal for the other to condition on the event (to play Dove when the other arrives first).

As our reader is likely aware, this was the logic provided by Maynard Smith to explain animal territoriality—why animals behave aggressively to defend territory that they have arrived at first, even if incumbency does not provide a defensive advantage and even when facing a more formidable intruder. Over the years, evidence has amassed to support Maynard Smith's explanation, such as experimental manipulation of which animal arrives first (Davies, 1978; Sigg & Falett, 1985).

Like other animals, we condition how aggressively we defend a resource on whether we arrive first. Because our behaviors are motivated by beliefs, we are also more likely to believe that the resource is "ours" when we arrive first. Studies have shown these effects with children's judgments of ownership, in ethnographies of prelegal societies, and in computer games. In one such illustration, DeScioli and Wilson (2011) had research subjects play a computer game in which they contested a berry patch. Subjects who ended up keeping control of the patch usually arrived first, and this determined the outcome more often than differences in fighting ability in the game.

This sense of ownership is codified in our legal systems, as illustrated by the quip "possession is 9/10ths of the law," and in a study involving famous legal property cases conducted by Descioli and Karpoff (2014). In a survey, these researchers asked participants to identify the rightful owner of a lost item, after reading vignettes based on famous property rights legal cases. Participants consistently identified the possessor of the found item as its rightful owner (as the judges had at the time of the case). This sense of ownership is also codified in our philosophical tradition, e.g., in Locke (1988), who found property rights in initial possession. Note that, as has also been found in animals, possession extends to objects on one's land: In DeScioli and

Karpoff's survey, another dictate of participants' (and the judges') property rights intuitions was who owned the land on which the lost item was found.

Also like animals, our sense of property rights is influenced by who created or invested in the resource, another uncorrelated asymmetry. In locales that sometimes grant property rights to squatters—individuals who occupy lands others have purchased—a key determinant of whether the squatters are granted the land is whether they have invested in it (Cone vs. West Virginia Pulp & Paper Co., 1947; Neuwirth, 2005). Locke also intuited that investment in land is part of what makes it ours: In *Second Treatise on Civil Government* (1689), Locke wrote, "everyman has a property in his person; this nobody has a right to but himself. The labor of his body and the work of his hand, we may say, are properly his."

If the Hawk–Dove model underlies our sense of property rights, we would expect to see psychological mechanisms that motivate us to feel entitled to an object when we possess it or have invested in it. Here are three such mechanisms, which can be seen by reinterpreting some well-documented "biases" in the behavioral economics literature. The first such bias is the *endowment effect*: We value items more if we are in possession of them. The endowment effect has been documented in dozens of experiments, where subjects are randomly given an item (mug, pen, etc.) and subsequently state that they are willing to sell the mug for much more than those who were not given the mug are willing to pay (Kahneman, Knetsch, & Thaler, 1990). In the behavioral economics literature, the endowment effect has sometimes been explained by loss aversion, which is when we are harmed more by a loss than we benefit from an equivalent gain. However, the source of loss aversion is not questioned or explained. When it is, loss aversion is also readily explained by the Hawk–Dove game (Gintis, 2007).

A second bias that also fits the Hawk–Dove model is the *IKEA effect*: Our valuation of an object is influenced by whether we have developed or built the resource. The IKEA effect has been documented by asking people how much they would pay for items like Lego structures or IKEA furniture after randomly being assigned to build them or receive them pre-built. Subjects are willing to pay more for items they build themselves.

A third such bias that fits the Hawk–Dove model is the *sunk cost fallacy* (Mankiw, 2007; Thaler, 1980), which leads us to "throw good money after bad" when we invest in ventures simply because we have already put so much effort into them, arguably because our prior efforts lead us to value those ventures more.

Possession and past investment are not the only uncorrelated asymmetries that can dictate rights. Rights can be dictated by a history of agreements, as happens when one party sells another deed to a house or car, or, as in our taxicab example, by whether a service was provided. There are also countless examples in which rights were determined by perhaps unfair or arbitrary characteristics such as race and sex: Black Americans were expected to give up their seat for Whites in the Jim Crow South and women to hand over their earnings or property to their husbands throughout the ages.

Hawk–Dove is not just a post hoc explanation for our sense of rights; it also leads to the following novel insight: We can formally characterize the properties that uncorrelated asymmetries must have. This requires a bit more game theory to illustrate; the logic is detailed in the section on categorical distinctions but the implications are straightforward: Uncorrelated asymmetries must be discrete (as in who arrived first or whether someone has African ancestry) and cannot be continuous (who is stronger, whether someone has darker skin). Indeed, we challenge the reader to identify a case where our sense of rights depends on surpassing a threshold in a continuous variable (stronger than? darker than?). More generally, an asymmetry must have the characteristic that, when it occurs, every observer believes it occurred with a sufficiently high probability, where the exact level of confidence is determined by the payoffs of the game. This is true of public, explicit speech and handshakes, but not innuendos or rumors. (Formally, explicit speech and handshakes induce what game theorists term common *p*-beliefs.)

The Hawk–Dove explanation of our sense of rights also gives useful clarity on when there will be conflict. Conflict will arise if both players receive opposing signals regarding the uncorrelated asymmetry, such as two individuals each believing they arrived first, or when there are two uncorrelated asymmetries that point in conflicting directions, such as when one person invested more and the other arrived first. The former source of conflict appears to be the case in the Israeli–Palestinian conflict. Indeed, both sides pour great resources into demonstrating their early possession, especially Israel, through investments in and public displays of archeology and history. The latter source of conflict appears to be the case in many of the contested legal disputes in the study by DeScioli and Karpoff (2014) mentioned above. An example is one person finds an object on another's land. Indeed, this turns out to be a source of many legal conflicts over property rights, and a rich legal tradition has developed to assign precedence to one uncorrelated asymmetry over another (Descioli & Karpoff, 2014). As usual, we see similar behavior in animals in studies that provide empirical support for Maynard Smith's model for animal territoriality: When two animals are each given the impression they arrived first by, for example, clever use of mirrors, a fight ensues (Davies, 1978).

Authentic Altruism, Motives, and the Envelope Game

In this section, we present a simple extension of the Repeated Prisoner's Dilemma to explain why morality depends not just on what people do but also what they think or consider.

In the Repeated Prisoner's Dilemma and other models of cooperation, players judge others by their actions—whether they cooperate or defect. However, we not only care about whether others cooperate but also about their decision-making process: We place more trust in cooperators who never even considered defecting. To quote Kant, "In law a man is guilty when he violates the rights of others. In ethics he is guilty if he only thinks of doing so."

The Envelope Game (Fig. 3) models why we care about thoughts and considerations and not just actions (Hoffman, Yoeli, & Nowak, 2015). The Envelope Game

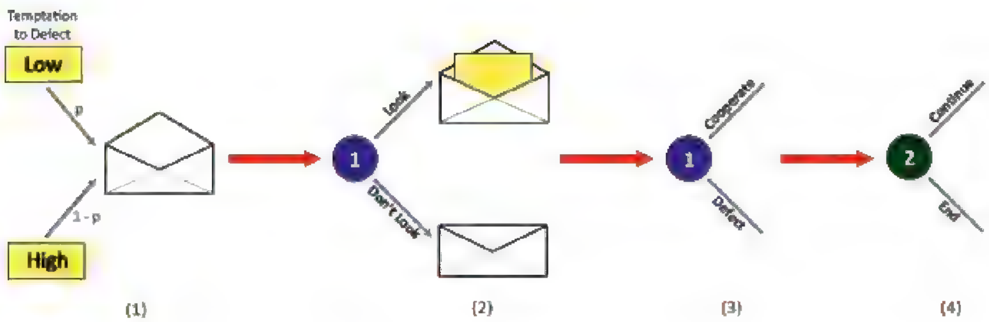


Fig. 3 A single stage of the Envelope Game

is a repeated game with two players. In each round, player 1 receives a sealed envelope, which contains a card stating the costs of cooperation (high temptation to defect vs. low temptation to defect). The temptation is assigned randomly and is usually low. Player 1 can choose to look inside the envelope and thus find out the magnitude of the temptation or choose not to look. Then player 1 decides to cooperate or to defect. Subsequently, player 2 can either continue to the next round or end the game. As in the Repeated Prisoner's Dilemma, the interaction repeats with a given likelihood, and if it does, an envelope is stuffed with a new card and presented to player 1, etc.

In this model, as long as temptations are rare, large, and harmful to player 2, it is a Nash equilibrium for player 1 to "cooperate without looking" in the envelope and for player 2 to continue if and only if player 1 has cooperated and not looked. We refer to this as the *cooperate without looking* (CWOL) equilibrium.[2] This equilibrium emerges in agent-based simulations of evolution and learning processes.[3] Notice that if player 1 could not avoid looking inside the envelope, or player 2 could not observe whether player 1 looked, there would not be a cooperative equilibrium since player 1 would benefit by deviating to defection in the face of large temptations. Not looking permits cooperative equilibria in the face of large temptations.

The Envelope Game is meant to capture the essential features of many interesting aspects of our morality, as described next.

Authentic Altruism. Many have asked whether "[doing good is] always and exclusively motivated by the prospect of some benefit for ourselves, however subtle" (Batson, 2014), for example, the conscious anticipation of feeling good (Andreoni,

[2] Technically, the conditions under which we expect players to avoid looking and attend to looking are $c_h > a/(1-w) > c_l p + c_h(1-p)$ and $bp + d(1-p) < 0$), where c_h and c_l are the magnitudes of the high and low temptations, respectively; p is the likelihood of the low temptation; $a/(1-w)$ is the value of a repeated, cooperative interaction to player 1; and $bp + d(1-p)$ is the expected payoff to player 2 if player 1 only cooperates when the temptation is low.

[3] The simulations employ numerical estimation of the replicator dynamics for a limited strategy space: cooperate without looking, cooperate with looking, look and cooperate only when the temptation is low, and always defect for player 1, and end if player 1 looks, end if player 1 defects, and always end for player 2.

1990), avoidance of guilt (Cain, Dana, & Newman, 2014; Dana, Cain, & Dawes, 2006; DellaVigna, List, & Malmendier, 2012), anticipation of reputational benefits or reciprocity (as Plato's Glaucon suggests, when he proffers that even a pious man would do evil if given a ring that makes him invisible; Trivers, 1971). At the extreme, this amounts to asking if saintly individuals such as Gandhi or Mother Teresa were motivated thus, or if they were "authentic" altruists who did good without anticipating any reward and would be altruistic even in the absence of such rewards. Certainly, religions advocate doing good for the "right" reasons. In the Gospel of Matthew, Chapter 6, Jesus advocates, "Be careful not to practice your righteousness in front of others to be seen by them. If you do, you will have no reward from your Father in heaven," after which he adds, "But when you give to the needy, do not let your left hand know what your right hand is doing, so that your giving may be in secret. Then your Father, who sees what is done in secret, will reward you."

The Envelope Game suggests authentic altruism is indeed possible: By focusing entirely on the benefits to others and ignoring the benefits to themselves, authentic altruists are trusted more, and the benefits from this trust outweigh the risk of, for example, dying a martyr's death. Moreover, this model helps explain why we think so highly of authentic altruists, as compared to others who do good, but with an ulterior motive (consider, as an example, the mockery Sean Penn has faced for showing up at disaster sites such as Haiti and Katrina with a photographer in tow).

Principles. Why do we like people who are "principled" and not those who are "strategic"? For example, we trust candidates for political office whose policies are the result of their convictions and are consistent over time and distrust those whose policies are carefully constructed in consultation with their pollsters and who "flip-flop" in response to public opinion (as caricatured by the infamous 2004 Republican presidential campaign television ad showing John Kerry windsurfing and tacking from one direction to another). CWOL offers the following potential explanation. Someone who is strategic considers the costs and benefits to themselves of every decision and will defect when faced with a large temptation, whereas someone who is guided by principles is less sensitive to the costs and benefits are to themselves and thus less likely to defect. Imagine our flip-flopping politician was once against gay marriage but supports it now that it is popular. This indicates the politician is unlikely to fight for the cause if it later becomes unpopular with constituents or risks losing a big donor. Moreover, this model may help explain why ideologues that are wholly devoted to a cause (e.g., Hitler, Martin Luther King, and Gandhi) are able to attract so many followers.

Don't Use People. Recall Kant's second formulation of the categorical imperative: "Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means but always at the same time as an end." In thinking this through, let's again consider dwarf tossing. Many see it as a violation of dwarfs' basic dignity to use them as a means for amusement, even though they willingly engage in the activity for economic gain. Our aversion to using people may explain many important aspects of our moral intuitions, such as why we judge torture as worse than imprisonment or punishment (torture is harming someone as a means to obtaining information) and perhaps one of the (many) reasons we oppose prostitution (prostitution is having sex with someone as a means to obtaining money). The Envelope Game clarifies the function of adhering to this maxim. Whereas those who treat someone well as means to an end would also mistreat them if expedient, those who treat someone well as an end can be trusted not to mistreat them when expedient.

Attention to Motives. The previous two applications are examples of a more general phenomenon: that we judge the moral worth of an action based on the motivation of the actor, as argued by deontological ethicists, but contested by consequentialists. The deontological argument is famously invoked by Kant: "Action from duty has its moral worth not in the purpose to be attained by it but in the maxim in accordance with which it is decided upon, and therefore does not depend upon the realization of the object of the action but merely upon the principle of volition in accordance with which the action is done without regard for any object of the faculty of desire" (Kant, 1997). These applications illustrate that we attend to motives because they provide valuable information on whether the actor can be trusted to treat others well even when it is not in her interest.

Altruism Without Prospect of Reciprocation. CWOL also helps explain why people cooperate in contexts where there is no possibility of reciprocation, such as in one-shot anonymous laboratory experiments like the dictator game (Fehr & Fischbacher, 2003), as well as when performing heroic and dangerous acts. Consider soldiers who throw themselves on a grenade to save their compatriots or stories like that of Liviu Librescu, a professor at the University of Virginia and a Holocaust survivor, who saved his students during a school shooting. When he heard the shooter coming toward his classroom, Librescu stood behind the door to his classroom, expecting that when the shooter tried to shoot through the door, it would kill him and his dead body would block the door. Mr. Librescu, clearly, did not expect this act to be reciprocated. Such examples have been used as evidence for group selection (Wilson, 2006), but can be explained by individuals "not looking" at the chance of future reciprocation. Consistent with this interpretation, cooperation during extreme acts of altruism is more likely to be intuitive than deliberative (Rand & Epstein, 2014), and those who cooperate without considering the prospect of reciprocation are more trusted (Critcher, Inbar, & Pizarro, 2013). We also predict that people are more likely to cooperate intuitively when they know they are being observed.

The Omission–Commission Distinction and Higher-Order Beliefs

We explain the omission–commission distinction and the means–by-product distinction by arguing that these moral intuitions evolved in contexts where punishment is coordinated. Then, even when intentions are clear to one witness for omissions and by-products, a witness will think intentions are less clear to the other witnesses.

Why don't we consider it murder to let someone die that we could have easily saved? For example, we sometimes treat ourselves to a nice meal at a fancy restaurant rather than donating the cost of that meal to a charity that fights deadly diseases. This extreme example illustrates a general phenomenon: that people have a tendency to assess harmful commissions (actions such as killing someone) as worse, or more morally reprehensible, than equally harmful omissions (inactions such as letting someone die). Examples of this distinction abound, in ethics (we assess withholding the truth as less wrong than lying (Spranca, Minsk, & Baron, 1991)), in law (it is legal to turn off a patient's life support and let the patient die, as long as one has the consent of the patient's family; however, it is illegal to assist the patient in committing suicide even with the family's consent), and in international relations. For example, consider the Struma, a boat carrying Jewish refugees fleeing Nazi persecution in 1942. En route to Palestine, the ship's engine failed, and it was towed to a nearby port in Turkey. At the behest of the British authorities then in control of Palestine, passengers were denied permission to disembark and find their way to Palestine by land. For weeks, the ship sat at port. Passengers were brought only minimal supplies, and their requests for safe haven were repeatedly denied by the British and others. Finally, the ship was towed to known hostile waters in the Black Sea, where it was torpedoed by a Russian submarine almost immediately, killing 791 of 792 passengers. Crucially, though, the British did not torpedo the ship themselves or otherwise execute passengers—an act of commission that they and their superiors would undoubtedly have found morally reprehensible.

Why do we distinguish between transgressions of omission and commission? To address this question, we present a simple game theory model based on the insight by DeScioli, Bruening, and Kurzban (2011). The intuition can be summarized in four steps:

1. We note that moral condemnation motivates us to punish transgressors. Such punishment is potentially costly, e.g., due to the risk of retaliation. We expect people to learn or evolve to morally condemn only when such costs are worth paying.
2. Moral condemnation can be less costly when others also condemn, perhaps because the risk of retaliation is diffused, because some sanctions do not work unless universally enforced or, worse, because others may sanction individuals they believe wrongly sanctioned. This can be modeled using any game with multiple Nash equilibria, including the Repeated Prisoner's Dilemma and the Side-Taking Game. The Coordination Game is the simplest game with multiple equilibria, so we present this game to convey the basic intuition. In the Coordination Game, there are two players who each simultaneously choose between two actions, say punish and don't punish. The key assumption is that each player prefers to do what she expects the other to do, which can be captured by assuming each receives a if they both punish, d if neither punish, $b<d$ if one punishes and the other does not, and $c<a$ if one does not punish while the other does (Fig. 4).
3. Transgressions of omission that are intended are difficult to distinguish from unintended transgression, as is the case when perpetrators are simply not paying

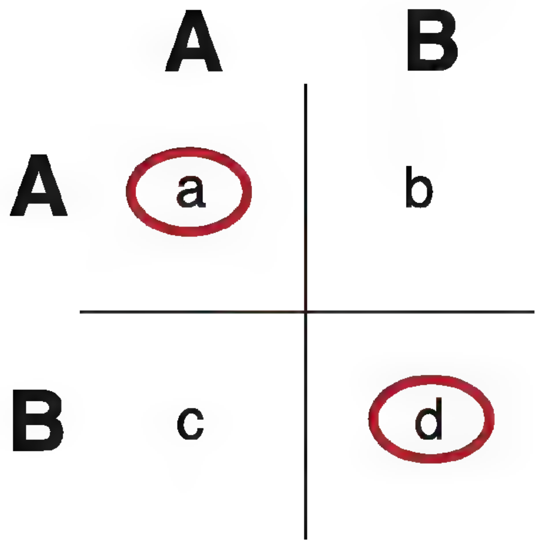


Fig. 4 The Coordination Game. In our applications, A stands for punish, and B stands for don't punish

attention or do not have enough time to react with better judgment (DeScioli et al., 2011). Relative to the example of the tennis player with the allergy described above, it is usually hard to distinguish between a competitor who does not notice his opponent orders the dish with the allergen versus one who notices but does not care. In contrast, transgressions of commission must be intended almost by definition.

4. Suppose the witness knows an omission was intentional: In the above example, the tennis player's opponent's allergy is widely known, and the witness saw the player watch his opponent order the offending dish, had time to react, thought about it, but did not to say anything. The witness suspects that others do not know the competitor was aware his opponent ordered the dish, but believes the tennis player should be condemned for purposely withholding information from his competitor. However, since the witness does not wish to be the sole condemner, she is unlikely to condemn. In contrast, when a witness observes a transgression of commission (e.g., the player recommends the dish), the witness is relatively confident that others present interpret the transgression as purposely harmful, since his recommendation reveals that the player was obviously paying attention and therefore intended to harm his opponent. So, if all other individuals present condemn the tennis player when they observe the commission, each does not anticipate being the sole condemner.

For the above result to hold, all that is needed is the following: (1) The more the costs of punishment decrease, the more others punish and (2) omissions are usually unintended (Dalkiran, Hoffman, Paturi, Ricketts, & Vattani, 2012; Hoffman et al., 2015).[4]

[4] In fact, even if one knows that others know that the transgression was intended, omission will still be judged as less wrong, since the transgression still won't create what game theorists call common p-belief, which is required for an event to influence behavior in a game with multiple equilibria.

This explanation for the omission–commission distinction leads to two novel predictions: First, for judgments and emotions not evolved to motivate witnesses to punishment but to, say, motivate witnesses to avoid dangerous partners (such as the emotion of fear; in contrast to anger or moral disgust), the omission–commission distinction is expected to be weaker or disappear altogether. Second, for transgressions of omission that, without any private information, can be presumed intentional (such as a mother who allows her child to go hungry or a person who does not give to a charity after being explicitly asked), we would not expect much of an omission commission distinction in moral condemnation.

As with the all models discussed in this chapter, the game theoretic explanation for the omission–commission distinction does not rest on rational, conscious, strategic calculation. In fact, in this particular case, all reasonable evolutionary dynamic models lead away from punishing omissions. The fact that the above results do not rest on rational, strategic thinking is particularly important in this setting since there is evidence that the distinction between omissions and commissions is not determined deliberately but rather intuitively (Cushman, Young, & Hauser, 2006) and appears to be evolved (DeScioli et al., 2011) and that consciously considering what others believe is an onerous process (Camerer, 2003; Epley, Keysar, Van Boven, & Gilovich, 2004; Hedden & Zhang, 2002).

This same model can explain several other puzzling aspects of our morality. The first is the *means–by-product* distinction. This distinction has been documented in studies that ask respondents to judge the following variants of the classic "trolley" problem. In the standard trolley "switch" case (Foot, 1967), a runaway trolley is hurtling toward a group of five people. To prevent their deaths, the trolley must be switched onto a side track where it will kill an innocent bystander. In studies using this case, the vast majority of subjects choose the utilitarian option, judging it permissible to cause the death of one to save five (e.g., Cushman et al., 2006; Mikhail, 2007). In the "footbridge" variant (Thomson, 1976), the trolley is hurtling toward the group of five people, but the switch to divert it is inoperable. The only way to save the five is to push a man who is wearing a heavy backpack off a bridge onto the track, thereby slowing the trolley enough so the five can escape, but killing the man. In contrast to the standard switch version, where causing the death of one person is but a by-product of the action necessary to save five, most subjects in the footbridge case find it morally impermissible to force the man with the backpack onto the tracks (Cushman et al., 2006; DeScioli, Gilbert, & Kurzban, 2012)—that is, when the man is used as a means to saving the five—even though the consequences are the same, and the decision to act was made knowingly and deliberately in both cases.

Such effects are found in less contrived situations, as well. Consider the real-life distinction between terrorism, in which civilian casualties are used a means to a political goal, and anticipated collateral damage, which is a by-product of war, even when the same number of civilians are knowingly killed and the same political ends are desired (say increased bargaining power in a subsequent negotiation).

The explanation again uses "higher-order beliefs" and is based on the key insight in DeScioli et al. (2011) and formalized in Dalkiran et al. (2012) and Hoffman et al. (2015): When the harm is done as a by-product, the harm is not usually anticipated.

So even when a witness knows that the perpetrator anticipated the harm, the witness believes other witnesses will not be aware of this and will presume the harm was not anticipated by the perpetrators. For instance, suppose we observe Israel killing civilians as a by-product of a strategic raid on Hamas militants. Even if we knew Israel had intelligence that confirmed the presence of civilians, we might not be sure others were privy to this information. On the other hand, when the harm is done as a means, the harm must be anticipated, since otherwise the perpetrator would have no motive to commit the act. Why would Hamas fire rockets at civilian towns with no military presence if Hamas does not anticipate a chance of civilian casualties? Consequently, it is Nash equilibrium to punish harm done as a means but not harm done as a by-product.

Similar arguments can be made for why we find direct physical transgressions worse than indirect ones, a moral distinction relevant to, for instance, the United States' current drone policy. Cushman et al. (2006) found that subjects condemn pushing a man off a bridge (to stop a train heading toward five others) more harshly than flipping a switch that leads the man to fall through a trap door. Pushing the victim with a stick is viewed as intermediate in terms of moral wrongness. Such moral wrongness judgments are consistent with considerations of higher-order beliefs: When a man is physically pushed, any witness knows the pushing was intended, but when a man is pushed with a stick some might not realize this, and even those who realize it might suspect others will not. Even more so when a button is pressed that releases a trap door.

It is worth noting that the above argument does not depend on a specific model of punishment, as in DeScioli and Kurzban's (2009) Side-Taking Game. The above model also makes the two novel predictions enumerated above, but nevertheless captures the same basic insight. It is also worth noting the contrast between the above argument and that of Cushman et al. (2006) and Greene et al. (2009), whose models rest on ease of learning or ease of mentally simulating a situation. It is not obvious to us how those models would explain that the omission–commission and means–by-product distinctions seem to depend on priors or be unique to settings of coordinated punishment.

Why Morality Depends on Categorical Distinctions

We explain why our moral intuitions depends so much more strongly on whether a transgression occurred than on how much damage was caused. Our argument again uses coordinated punishment and higher-order beliefs: When a categorical distinction is violated, you know others know it was violated, but this is not always true for continuous variables.

Consider the longstanding norm against the use of chemical weapons. This norm recently made headlines when Bashar al-Assad was alleged to have used chemical weapons to kill about a thousand Syrian civilians, outraging world leaders who had been silent over his use of conventional weapons to kill over 100,000 Syrian civilians. A Reuters/Ipsos poll at the time found that only 9 % of Americans favored intervention in Syria, but 25 % supported intervention if the Syrian government forces used chemical weapons against civilians (Wroughton, 2013). In the past, the United States has abided by the norm against the use of chemical weapons even at the expense of American lives: In WWII, Franklin D. Roosevelt chose to eschew chemical weapons in Iwo Jima even though, as his advisors argued at the time, their use would have saved thousands of American lives. It might even have been more humane than the flame-throwers that were ultimately used against the Japanese ("History of Chemical Weapons," 2013). We say that the norm against chemical weapons is a categorical norm because those who abide by it consider whether a transgression was committed (did Assad use chemical weapons?), rather than focusing entirely on how much harm was done (how many civilians did Assad kill?). Other norms are similarly categorical. For instance, in the introduction to this chapter, we noted that across cultures and throughout history, the norm against murder has always been categorical: We consider whether a life was terminated, not the loss of useful life years. Likewise, discrimination (e.g., during Jim Crow) is typically based on categorical definitions of race (the "one drop rule") and not, say, the darkness of skin tone. Human rights are also categorical. A human rights violation occurs if someone is tortured or imprisoned without trial, regardless of whether it was done once or many times and regardless of whether the violation was helpful in, say, gaining crucial information about a dangerous enemy or an upcoming terror attack. We even assign rights in a categorical way to all *Homo sapiens* and not based on intelligence, sentience, ability to feel pain, etc.

Why is it that we attend to such categorical distinctions instead of paying more attention to the underlying continuous variable? We use game theory to explain this phenomenon as follows: Suppose that two players (say, the United States and France) are playing a Coordination Game in which they decide whether to punish Syria, and each wants to sanction only if the other sanctions. We assume the United States does not want to levy sanctions unless it is confident France will as well, which corresponds to an assumption on the payoffs of the game (if we reverse this assumption, it changes one line in the proof, but not the result).

We model the underlying measure of harm as a continuous variable (in our example, it is the number of civilians killed). For simplicity, we assume this variable is uniformly distributed, which means Assad is equally likely to kill any number of people. This assumption is, again, not crucial, and we will point out the line in the proof that it affects. Importantly, we assume that players do not directly observe the continuous variable, but instead receive some imperfect signal (e.g., the United States observes the body count by its surveyors).

Imagine a norm that dictates that witnesses punish if their estimate of the harm from a transgression is above some threshold (e.g., levy sanctions against Syria if the number of civilians killed is estimated to be greater than 100,000). As it turns out, this is not a Nash equilibrium. To see why, consider what happens when the United States gets a signal right at the threshold. The United States thinks there is a

50 % chance that France's estimates are lower than its own[5] and, thus, that there is a 50 % chance that France's estimates are lower than the threshold. This further implies that the United States assesses only a 50 % chance that France levies sanctions, so the United States is not sufficiently confident that France will sanction, to make it in the United States's interest to sanction.

What we have shown so far is that for a threshold of 100,000, it is in the interest of the United States to deviate from the strategy dictated by the threshold norm when it gets a signal at the threshold. This means that 100,000 is not a viable threshold, and (since 100,000 was chosen arbitrarily) there is no Nash equilibrium in which witnesses punish if their estimate of the harm from a transgression is above some arbitrary threshold.

It should be noted that this result only requires that there are sufficiently many possibilities, not that there is in fact a continuum. Neither does it require that the distribution is uniform nor that the Coordination Game is not affected by the behavior of Assad. The only crucial assumptions are that the distribution is not too skewed and that the payoffs are not too dependent on the behavior of Assad (for details, see Dalkiran et al., 2012; Hoffman, Yoeli, & Dalkiran, 2015).

What happens if such norms are learned or evolved and subject to selection? Suppose there is a norm to attack whenever more than 100,000 civilians are killed. Players will soon realize that they should not attack unless, say, 100,100 civilians are killed. Then, players will learn not to attack when they estimate 100,200 civilians are killed and so on, indefinitely. Thus, every threshold will eventually "unravel," and no one will ever attack.[6]

Now let's consider a categorical norm, for example, the use of chemical weapons. We again model this as a random variable, though this time, the random variable can only take on two values (0 and 1), each with some probability. Again, players do not know with certainty whether the transgression occurred, but instead get a noisy signal. In our example, the signal represents France or the United States's assessment of whether Assad used chemical weapons, and there is some likelihood the assessors make mistakes: They might not detect chemical weapons when they had been used or might think they have detected chemical weapons when none had been used.

Unlike with the threshold norm, provided the likelihood of a mistaken signal is not too high, there is a Nash equilibrium where both players punish when they receive a signal that the transgression occurred. That is, the United States and France each levy sanctions if their assessors detect chemical weapons. This is because when the United States detects chemical weapons, the United States believes France

[5] This is where the assumption of a uniform distribution comes in. Had we instead assumed, for instance, that the continuous variable is normally distributed, then it would not be exactly 50–50 but would deviate slightly depending on the standard deviation and the location of the threshold. Nevertheless, the upcoming logic will still go through for most Coordination Games, i.e. any Coordination Game with risk dominance not too close to .5.

[6] As with omission, this follows from iterative elimination of strictly dominated strategies (see Hoffman et al., 2015, for details).

likely detected them and will likely levy sanctions. So the United States's best response is to levy sanctions. Similarly, if the United States does not detect chemical weapons, it expects France did not and will not levy sanctions, so the United States is better off not levying them.

This result is useful for evaluating whether it is worthwhile to uphold a norm. The Obama administration was harshly criticized for threatening to go to war after the Assad regime used chemical weapons but not earlier, although the regime had already killed tens of thousands of civilians. The model clarifies that Obama's position was not as inconsistent as his critics had charged: The norm against chemical weapons may be worth enforcing since it is sustainable, whereas norms against civilian casualties are harder to sustain and hence might not be worthwhile to enforce.

Let's return to some more of our motivating examples. Our model can explain why we define murder categorically: It is not possible to punish differently for different amount of quality life years taken, but it is possible to punish differentially for a life taken. As with omission–commission, however, we do expect sadness or grief to depend greatly on life years lost, even if the punishment or moralistic outrage will be less sensitive. This is a prediction of the model that, as far as we know, has yet to be tested.

Similarly, the "one-drop" rule is a categorical norm, so it can be socially enforced in an apartheid society. In contrast, consider a rule that advocates giving up one's seat for someone with lighter skin. Since this is based on a threshold in a continuous variable, while it might be enforceable by a unilateral authority, it cannot be enforced by "mob rule." Other forms of discrimination, such as discriminating against the less attractive, or the less tall, or the elderly, all being continuous variables, cannot be socially enforced via coordinated punishment, and hence, we expect such discrimination to be of a different form. In particular, it will not be based not on punishing violators. For example, male CEOs might still prefer young attractive female secretaries, and taller men are more likely to be hired as CEOs, not because of coordinated rewards or punishment but because those who hire the CEOs or secretaries are likely to be satisfying their own preferences or doing what they expect will lead to higher profits.

Likewise, the number of victims tortured by a regime or the number of lives saved by torturing is continuous. Thus, a regime cannot be punished by a coordinated attack by other countries or by a coordinated rebellion by its citizens based on the number of people tortured or the paucity of reasons for such torture. But, a regime can be attacked or overthrown depending on whether a physical harm was inflicted on a citizen by the state. Hence, human rights are treated as inalienable, even in the absence of an a priori justification for this nonutilitarian norm. And why are human rights ascribed to all living *Homo sapiens*? Perhaps not because of a good logical a priori argument, but simply because violations of human rights are enforceable by coordinated punishment, but no regime can be punished for harming any "person" of less than a certain degree of consciousness.

Finally, here is one last application. The model might also explain why revolutions are often caused by categorical events, such as a new tea tax or a single, widely publicized self-immolation, and not a breach of a threshold in, say, the quality of life of citizens or the level of corruption. This explanation requires simply that we recognize revolutions as a coordination problem (as argued in Morris & Shin, 2002; Chwe, 2013), where each revolutionary chooses whether to revolt, and each is better off revolting only if sufficiently many others revolt.

Quirks of Altruism and the Repeated Prisoner's Dilemma with Incomplete Information

The Repeated Prisoner's Dilemma has famously been used as an explanation for the evolution of cooperation among non-kin (Axelrod & Hamilton, 1981; Dawkins, 2006; Pinker, 2003; Trivers, 1971). *In this section, we show how the same basic model can be used to explain many of the quirky features of our pro-social preferences and ideologies.*

Recall that in the Prisoner's Dilemma, each of two players simultaneously chooses whether to cooperate. Cooperation reduces a player's own payoffs by $c>0$ while increasing the other's payoffs by $b>c$. The only Nash equilibrium is for neither player to cooperate. In the Repeated Prisoner's Dilemma, the players play a string of Prisoner's Dilemmas. That is, after the players play a Prisoner's Dilemma, they learn what their opponent did and play another Prisoner's Dilemma against the same opponent with probability δ (and the game ends with probability $1-\delta$). As is well known in the evolutionary literature, there are equilibria in which players end up cooperating, provided $\delta>c/b$. In all such equilibria, cooperation is sustained because any defection by one player causes the other player to defect. This is called reciprocity. As the reader is surely familiar, there is ample evidence for the Repeated Prisoner's Dilemma as a basis for cooperation from computer simulations (e.g., Axelrod, 1984) and animal behavior (e.g., Wilkinson, 1984). The model can be extended to explain contributions to public goods if, after deciding whether to contribute to a public good, players play a Repeated Prisoner's Dilemma (see, e.g., Panchanathan & Boyd, 2004) (Fig. 5).

The key to understanding these quirks is that players often have incomplete information. For example:

1. Players do not always observe contributions. It is intuitive that, for cooperation to occur in equilibrium, contributions need to be observed with sufficiently high probability.
2. Others cannot always tell whether a player had an opportunity to contribute. For defection to be penalized, it must be the case that others can tell that a player had the opportunity to cooperate and did not (i.e. the player should not be able to hide the fact that there was an opportunity to cooperate).
3. Sometimes, there are two ways to cooperate, and one has a higher benefit, *b*. Then, the only way this more effective type of cooperation can be sustained in equilibrium is if others know which cooperative act is more effective.

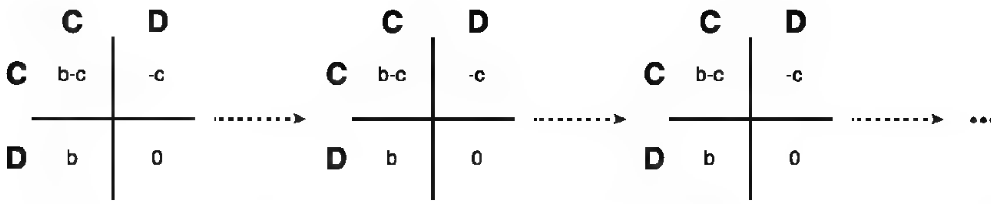


Fig. 5 The Repeated Prisoner's Dilemma. Two players play a Prisoner's Dilemma. They each observe the other's action, then, with probability δ, play another Prisoner's Dilemma against the same opponent, etc.

Technically, for the second and third point, what is needed is common knowledge that a player had an opportunity to cooperate or of the more effective means of cooperation. If observers were to know one purposely chose to defect or chose the less cooperative act, but they do not know that others know this, then observers think others will think punishment is not warranted, and observers will not punish. The argument is analogous to the discussion of higher-order beliefs in the omission–commission subsection and formalized in Dalkiran et al. (2012) and Hoffman et al. (2015).

Interpreting the Quirks of Altruism

Below we discuss some of the quirky features of altruism identified by economists and psychologists. In each case, we will argue that these features might be puzzling, but not when viewed through the lens of the above model:

Insensitivity to Effectiveness. We are surprisingly insensitive to the impact of our charitable contributions. We vote because we "want to be a part of the democratic system," or we "want to make a difference," despite the fact that our likelihood of swinging an election (even in a swing state) is smaller than our likelihood of being struck by lightning (Gelman, Silver, & Edlin, 2012). Why is our desire to "make a difference" or "be a part of the system" immune to the actual difference we are making? Our charitable contributions or volunteer efforts suffer from the same insensitivity. Why does anyone give money or volunteer time to Habitat for Humanity? The agency flies high earners who have never held a hammer halfway across the world to build houses that would be substantially more cheaply built by local experts funded by the high earners. Experimental evidence demonstrates our insensitivity: Experimental subjects are willing to pay the same amount to save 2000, 20,000, or 200,000 birds (Desvousges et al., 2010). Likewise, when donors are told their donations will be matched, tripled, or quadrupled, they donated identical amounts (Karlan & List, 2006). Why do we give so much, but do not ensure our gifts have a large impact?

The explanation follows directly from the above model: It is often the case that observers do not know which acts are effective and which are not and, certainly, this usually is not commonly known. Thus, they will not reward or punish based on effectiveness, and we ourselves will not attend to effectiveness in equilibrium. This explanation suggests that if we want to increase efficacy of giving, we ought to focus on making sure donors' friends and colleagues are aware of the efficacy of different options. In fact, this is perhaps more important than informing the donor of efficacy, since the donor will be motivated to uncover efficacy herself.

Magnitude of the Problem. We are surprisingly unaware of and unaffected by the magnitudes of the problems we contribute to solving. How many of those who participated in the recent ALS Ice Bucket Challenge have even the vaguest sense of the number of ALS victims? (Answer: about 1/100th the victims of heart disease.) How much happier would these individuals have been if the number of ALS victims were cut in half? Multiplied by 100? The same questions could be asked about AIDS or cleft lips. If we were actually motivated by our desire to rid the world of such afflictions as we often proclaim, then we would be happier if there were fewer afflicted individuals and less happy if there were more. But we are not even aware of these numbers, let alone affected by them. This suggests an alternative motivation than the one we proclaim.

On the other hand, if we give in order to gain social rewards, it does not matter whether the problem is large or small, provided others recognize it as a problem and the social norm is to give. If our learned or evolved preferences were drastically impacted by the magnitude of the crises, we would be sensitive to whether the problem was solved, perhaps motivating us to ensure that others solve it, which we would not get credit for, or perhaps motivating us to devote too much of our resources to solving it, beyond what we would actually get rewarded for.

Observability. There is overwhelming evidence that people give more when their gifts are observed. Much of this evidence comes from the lab, where it has been demonstrated a myriad of ways (e.g., Andreoni & Petrie, 2004; Bolton, Katok, & Ockenfels, 2005; List, Berrens, Bohara, & Kerkvliet, 2004). For instance, when participants play a public goods game in the laboratory for money, their contributions are higher when they are warned that one subject will have to announce to the room of other participants how much they contributed (List et al., 2004). However, evidence also comes from real-world settings, which find large effects in settings as diverse as blood donation (Lacetera & Macis, 2010), blackout prevention (Yoeli, Hoffman, Rand, & Nowak, 2013), and support for national parks (Alpizar, Carlsson, & Johansson-Stenman, 2008). In Switzerland, voting rates fell in small communities when voters were given the option to vote by mail (Funk, 2006), which makes it harder to tell who did not vote, even though it also makes it easier to vote. In fact, our willingness to give more when observed extends to subtle, subconscious cues of being observed: People give twice as much in dictator games when there are markings on the computer screen that vaguely represent eyes (Haley & Fessler, 2005), and they are more likely to pay for bagels in their office when the payment box has a picture of eyes above it (Bateson, Nettle, & Roberts, 2006).

These results should not surprise anyone who believes our pro-social tendencies are influenced by reputational concerns (though the magnitudes are surprisingly large).

The effectiveness of subconscious cues of observability points to a primary role for reputations in our learned or evolved proclivities toward pro-social behavior. The large impact of subtle cues of observability, however, calls into question alternative explanations not based on reputations.

Explicit Requests. When we are asked directly for donations, we give more than if we are not asked, even though no new information is conveyed by the request. In a study of supermarket shoppers around Christmas time, researchers found that passersby were more likely to give to the Salvation Army if volunteers not only rang their bell but explicitly asked for a donation (Andreoni, Rao, & Trachtman, 2011). If our motive is to actually do good, or perhaps proximally to feel good by the act of giving, we should not be impacted by an explicit request.

However, if we evolved or learned to give in order to gain rewards or avoid punishment as described above, then we ought to be more likely to give when, if we did not give, it would be common knowledge that we had the option to give and chose not to. The explicit request makes the denial common knowledge.

It is worth emphasizing that our evolved intuition to respond to explicit asks may be (mis)applied to individual settings that lack social rewards. Imagine you are approached by a Salvation Army volunteer in front of a store in a city where you are visiting for one day only. A literal reading of the model would suggest that you should be no more likely to respond to an explicit request. But it is more realistic to expect that if your pro-social preferences were learned or evolved in repeated interactions then applied to this new setting, you would respond in a way that is not optimal for this particular setting and nonetheless give more when explicitly asked (just as our preferences for sweet and fatty foods, which evolved in an environment where food was scarce, lead us to overeat now that food is abundant).

Avoiding Situations in Which We Are Expected to Give. In the same supermarket study, researchers discovered that shoppers were going out of their way to exit the store through a side door, to avoid being asked for a contribution by the Salvation Army volunteers. In another field experiment, those who were warned in advance that a solicitor would come to the door asking for charitable donations were more likely to not be home. The researchers estimated that among those who gave, 50 % would have avoided being home if warned in advance of the solicitor's time of arrival (DellaVigna et al., 2012). In a laboratory analog, subjects who would have otherwise given money in a $10 dictator game were willing to pay a dollar to keep the remaining nine dollars and prevent the recipient from knowing that a dictator game could have been played (Dana et al. 2006). If our motive were to have an impact, we would not pay to avoid putting ourselves in a situation where we could have such an impact. Likewise, if our motive were to feel good by giving, we would not pay to avoid this feeling.

In contrast, if we evolved or learned to give in order to gain rewards or avoid punishment, then we would pay to avoid situations where we are expected to give. Again, this would be true even if, in this particular setting, we were unlikely to actually be punished.

Norms. People are typically *conditionally cooperative*, meaning that they are willing to cooperate more when they believe others contribute more. For example, students asked to donate to a university charity gave 2.3 percentage points more when told that others had given at a rate of 64 % than when they were told giving rates were 46 % (Frey & Meier, 2004). Hotel patrons were 26 % more likely to reuse their towels when informed most others had done the same (Goldstein, Cialdini, & Griskevicius, 2008). Households have been shown to meaningfully reduce electricity consumption when told neighbors are consuming less, both in the United States (Ayres, Raseman, & Shih, 2012) and in India (Sudarshan, 2014).

Such conditional cooperation is easily explained by the game theory model: When others give, one can infer that one is expected to give and may be socially sanctioned if one does not.

Strategic Ignorance. Those at high risk of contracting a sexually transmitted disease (STD) often go untested, presumably because if they knew they had the STD, they would feel morally obliged to refrain from otherwise desirable activity that risks spreading the STD. Why is it more reproachable to knowingly put a sexual partner at risk when one knows one has the STD than to knowingly put a sexual partner at risk by not getting tested? There is evidence that we sometimes pursue *strategic ignorance* and avoid information about the negative consequences of our decisions to others. When subjects are shown two options, one that is better for themselves but worse for their partners and one that is worse for themselves but better for their partners, many choose the option that is better for their partners. But, when subjects must first press a button (at no cost) to reveal which option is better for their partners, they choose to remain ignorant and simply select the option that is best for themselves (Dana, Weber, & Kuang, 2007).

This quirk of our moral system is again easy to explain with the above model. Typically, information about how one's actions affect others is hard to obtain, so people cannot be blamed for not having such information. When one can get such information easily, others may not know that it is easy to obtain and will not punish anyone who does not have the information. For example, although it is trivially easy to look up charities' financial ratings on websites like charitynavigator.org, few people know this and *could* negatively judge those that donate without first checking such websites. And even when others know that one can get this information easily, they might suspect that others do not know this, and so avoid punishing, since others won't expect punishment. To summarize, strategic ignorance prevents common knowledge of a violation and so is likely to go unpunished. We again emphasize that we will be lenient of strategic ignorance, even when punishment is not literally an option.

Norm of Reciprocity. We feel compelled to reciprocate favors, even if we know that the favors were done merely to elicit reciprocation and even if the favor asked in return is larger than the initial one granted (Cialdini 2001). For instance, members of Hare Krishna successfully collect donations by handing out flowers to disembarking passengers at airports, even though passengers want nothing to do with the flowers: They walk just a few feet before discarding them in the nearest bin.

Psychologists and economists sometimes take this "norm" as given, without asking where it comes from, and a naive reading of Trivers would lead one to think that we should be sensitive to the magnitude of the initial favor and whether it is manipulative.

However, according to the above model, reciprocity is the Nash equilibrium, even if the favors are not evenly matched or manipulative, since, in equilibrium, we are neither sensitive to such quantitative distinctions nor to whether the initial reciprocity was manipulative, unless these facts are commonly known.

Self-Image Concerns. People sometimes play mental tricks in order to appear *to themselves* as pro-social. For example, in an experiment, subjects will voluntarily take on a boring task to save another subject from doing it, but if given the option of privately flipping a coin to determine who gets the task, they often flip—and flip, and flip again—until the "coin" assigns the task to the other subject (Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997). Why would we be able to fool ourselves and not, say, recognize that we are gaming the coin flip? Why do we care what we think of ourselves at all? Are there any constraints on how we will deceive ourselves?

Such self-image considerations can be explained by noting that our self-image can act as a simple proxy, albeit an imperfect one, for what others think of us, and also that we are more convincing to others when we believe something ourselves (Kurzban, 2012; Trivers, 2011). This explanation suggests that the ways we deceive ourselves correspond to quirks described throughout this section—for example, we will absolve ourselves of remaining strategically ignorant even when it is easy not to, or be convinced that we have done good by voting, even if we cannot swing an election.

Framing Effects. Whether we contribute is highly dependent on the details of the experiment, such as the choice set (List, 2007) and the labels for the different choices (Ross & Ward, 1996; Roth, 1995). Such findings are often taken as evidence that social preferences cannot be properly measured in the lab (Levitt & List, 2007).

We believe a more fruitful interpretation is simply that the frame influences whether the laboratory experiment "turns on" our pro-social preferences, perhaps by simulating a situation where cooperation is expected (Levitt & List, 2007).

One-Shot Anonymous Giving: We give in anonymous, one-shot settings, such as dictator games. We also sacrifice for others in the real world when there is no chance of reciprocation: Heroes jump on grenades to save their fellow soldiers or block the door to a classroom with their bodies to prevent a school shooter from entering (Rand & Epstein, 2014). This is often seen as evidence for a role of group selection (Fehr & Fischbacher, 2003).

However, an alternate explanation is that we do not consider the likelihood of reciprocation (Hoffman et al., 2015), as described above. To explain the laboratory evidence, there are two more possibilities. First, subjects may believe there is some chance their identity will be revealed and feel the costs of being revealed as selfish are greater than the gains from the experiment (Delton, Krasnow, Cosmides, & Tooby, 2011). Second, we again emphasize that learned or evolved preferences and ideologies are expected to be applied even in novel settings to which they are not optimized.

Conclusion

In this chapter we have showed that a single approach–game theory, with the help of evolution and learning–can explain many of our moral intuitions and ideologies. We now discuss two implications.

Group Selection. Our chapter relates to the debate on group selection, whereby group level competition and reproduction is supposed to occasionally cause individuals to evolve to sacrifice their own payoffs to benefit the group (e.g., Wilson, 2006). One of the primary pieces of evidence cited in support of group selection is the existence of human cooperation and morality (Fehr & Fischbacher, 2003; Fehr, Fischbacher, & Gächter, 2002; Gintis, Bowles, Boyd, & Fehr, 2003; Haidt, 2012; Wilson, 2010, 2012), in particular: giving in one-shot anonymous laboratory experiments, intuitively sacrificing one's life for the group (jumping on the grenade), and contributions to public goods or charity. However, we have reviewed an alternative explanation for these phenomena that does not rest on group selection. It also yields predictions about these phenomena that group selection does not, such as that people are more likely to cooperate when they are being observed and there is variance in the cost of cooperation. The approach described here also explains other phenomena, such as categorical norms and ineffective altruism. These lead to social welfare losses, which is suboptimal from the group's perspective. The categorical norm against murder, for example, leads to enormous waste when keeping alive, sometimes for years, those who have virtually no chance of a future productive life.

Admittedly, despite their inefficiencies, these moral intuitions do not rule out group selection, since group selection can be weak relative to individual selection. But it does provide a powerful argument that group selection is unnecessary for explaining many interesting aspects of human morality. It also suggests that group selection is, indeed, at most, weak. One example that makes this especially clear is discrete norms. Recall that we argued that continuous norms are not sustainable because individuals benefit by deviating around the threshold. Notice that this benefit is small, since the likelihood that signals are right around the threshold is low. Group selection could easily overwhelm the benefit one would get from deviating from this Nash equilibrium, suggesting group selection is weak (i.e. there are few group-level reproductive events, high migration rates, high rates of "mutation" in the form of experimentation among individuals, etc.).

Logical Justification of Moral Intuitions. In each of the applications above, we explained moral intuitions without referring to existing a priori logical justifications by philosophers or others. Our explanation for our sense of rights does not rely on Locke's "state of nature." No argument we gave rests on God as an orderly designer, on Platonic ideals, on Kant's concepts of autonomy and humanity, etc. What does this mean for these a priori justifications? It suggests that they are not the source of our morality and are, instead, post hoc justifications of our intuitions (Haidt, 2012).

To see what we mean, consider the following analogy. One might wonder why we find paintings and sculptures of voluptuous women beautiful. Before the development of sexual selection theory, one might have argued that perfect spheres are some kind of Platonic solid, and inherently desirable, or that curvy hips yield golden ratios. But with our current understanding of sexual selection, we recognize that our sense of beauty has evolved and that there is no platonic sense of beauty outside of that shaped by sexual selection. Any argument about perfect spheres is unparsimonious and likely flawed. Without the help of evolution and game theory, did philosophers conjure the moral equivalents of perfect spheres and golden ratios? The state of nature, the orderly designer, Platonic ideals, autonomy, and humanity, etc. perhaps these arguments are also unfounded and unnecessary.

References

Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2008). Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics, 92*(5), 1047–1060.

Andreoni, J., & Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics, 88*(7), 1605–1623.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm glow giving. *The Economic Journal, 100*, 464–477.

Andreoni, J., Rao, J. M., & Trachtman, H. (2011). *Avoiding the ask: A field experiment on altruism, empathy, and charitable giving*. Technical report, National Bureau of Economic Research.

Axelrod, R. M. (1984). *The evolution of cooperation*. New York: Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390 1396.

Ayres, I., Raseman, S., & Shih, A. (2012). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization*, 2–20.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters, 2*(3), 412–414.

Batson, C. D. (2014). *The altruism question: Toward a social-psychological answer*. Hillside, NJ: Psychology Press.

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology, 72*(6), 1335.

Bolton, G. E., Katok, E., & Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. *Journal of Public Economics, 89*(8), 1457–1468.

Cain, D., Dana, J., & Newman, G. (2014). *Giving vs. giving in.* Yale University Working Paper.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Chammah, A. M., & Rapoport, A. (1965). *Prisoner's dilemma; a study in conflict and cooperation*. Ann Arbor, MI: University of Michigan.

Chwe, M. (2013). *Rational ritual: Culture, coordination, and common knowledge*. Princeton, NJ: Princeton University Press.

Cone v. West Virginia Pulp & Paper Co., 330 U.S. 212, 67 S. Ct. 752, 91 L. Ed. 849 (1947).

Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science, 4*(3), 308–315.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. *Psychological Science, 17*(12), 1082–1089.

Dalkiran, N. A., Hoffman, M., Paturi, R., Ricketts, D., & Vattani, A. (2012). Common knowledge and state-dependent equilibria. In *Algorithmic game theory* (pp. 84–95). New York: Springer.

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes, 100*(2), 193–201.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory, 33*(1), 67–80.

Davies, N. B. (1978). Territorial defense in the speckled wood butterfly (pararge aegeria): The resident always wins. *Animal Behaviour, 26*, 138–147.

Dawkins, R. (2006). *The selfish gene*. Oxford, UK: Oxford University Press.

DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics, 127*(1), 1–56.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences, 108*(32), 13335–13340.

DeScioli, P., Bruening, R., & Kurzban, R. (2011). The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior, 32*(3), 204–215.

DeScioli, P., Gilbert, S. S., & Kurzban, R. (2012). Indelible victims and persistent punishers in moral cognition. *Psychological Inquiry, 23*(2), 143–149.

DeScioli, P., & Wilson, B. J. (2011). The territorial foundations of human property. *Evolution and Human Behavior, 32*(5), 297–304.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition, 112*(2), 281–299.

Descioli, P., & Karpoff, R. (2014). *People's judgments about classic property law cases*. Brandeis University Working Paper.

Desvousges, W. H., Johnson, F. R., Dunford, R. W., Boyle, K. J., Hudson, S. P., Wilson, K. N., et al. (2010). *Measuring nonuse damages using contingent valuation: An experimental evaluation of accuracy*. Research Triangle Park, NC: RTI Press.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*(3), 327.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785–791.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature, 13*(1), 1–25.

Fisher, R. A. (1958). *The genetic theory of natural selection*. Mineola, NY: Dover.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: WW Norton & Co.

Frey, B. S., & Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review, 94*, 1717–1722.

Funk, P. (2006). Modern voting tools, social incentives and voter turnout: Theory and evidence. In *Annual Meeting of the American Economic Association*, Chicago, IL.

Gelman, A., Silver, N., & Edlin, A. (2012). What is the probability your vote will make a difference? *Economic Inquiry, 50*(2), 321–326.

Gintis, H. (2007). The evolution of private property. *Journal of Economic Behavior & Organization, 64*(1), 1–16.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153–172.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a view-point: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research, 35*(3), 472–482.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition, 111*(3), 364–371.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage.

Haley, K. J., & Fessler, D. M. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*.

Hedden, T., & Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. *Cognition, 85*(1), 1–36.
Hoffman, M., Yoeli, E., & Nowak, M. A. (2015). Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences, 112*(6), 1727–1732.
Hoffman, M., Yoeli, E., & Dalkiran, A. (2015). *Social applications of common knowledge*. Harvard University Working Paper.
Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy, 98*, 1325–1348.
Kant, I. (1997/1787). Groundwork of the metaphysics of morals (1785). *Practical Philosophy, 108*.
Karlan, D., & List, J. A. (2006). *Does price matter in charitable giving? Evidence from a large-scale natural field experiment*. Technical report, National Bureau of Economic Research.
Kurzban, R. (2012). *Why everyone (else) is a hypocrite: Evolution and the modular mind*. Princeton, NJ: Princeton University Press.
Lacetera, N., & Macis, M. (2010). Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization, 76*(2), 225–237.
Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives, 21*, 153–174.
List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy, 115*(3), 482 493.
List, J. A., Berrens, R. P., Bohara, A. K., & Kerkvliet, J. (2004). Examining the role of social isolation on stated preferences. *American Economic Review, 94*, 741–752.
Locke, J. (1988/1688). *Locke: Two treatises of government student edition*. Cambridge, UK: Cambridge University Press.
Mankiw, N. G. (2007). *Principles of economics*. Thomson Learning.
Milgrom, P., & Weber, R. J. (1982). The value of information in a sealed-bid auction. *Journal of Mathematical Economics, 10*(1), 105–114.
Morris, S., & Shin, H. S. (2002). Social value of public information. *The American Economic Review, 92*(5), 1521–1534.
Myerson, R. B. (2004). Justice, institutions, and multiple equilibria. *Chicago Journal of International Law, 5*, 91.
Neuwirth, R. (2005). *Shadow cities: A billion squatters, a new urban world*. New York: Routledge.
Nowak, M. A. (2006). *Evolutionary dynamics: Exploring the equations of life*. Cambridge, MA: Harvard University Press.
Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature, 432*(7016), 499–502.
Pinker, S. (2003). *The blank slate: The modern denial of human nature*. UK: Penguin.
Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences, 105*(3), 833–838.
Rand, D. G., & Epstein, Z. G. (2014). *Risking your life without a second thought: Intuitive decision-making and extreme altruism*. Available at SSRN 2424036.
Ross, L., & Ward, A. (1996). Naive realism: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103 135).
Roth, A. E. (1995). Bargaining experiments. In J. H. Kagel & E. R. Alvin (Eds.), *The handbook of experimental economics* (pp. 253–342). Princeton, NJ: Princeton University Press.
Sigg, H., & Falett, J. (1985). Experiments on respect of possession and property in hamadryas baboons (*Papio hamadryas*). *Animal Behaviour, 33*(3), 978–984.
Smith, J. M., & Price, G. (1973). The logic of animal conflict. *Nature, 246*, 15.
Sosis, R., & Alcorta, C. (2003). Signaling, solidarity, and the sacred: The evolution of religious behavior. *Evolutionary Anthropology: Issues, News, and Reviews, 12*(6), 264–274.
Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology, 27*(1), 76–105.

Sudarshan, A. (2014). *Nudges in the marketplace: Using peer comparisons and incentives to reduce household electricity consumption*. Technical report, Harvard University Working Paper.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39–60.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist, 59*(2), 204–217.

Tirole, J. (1988). *The theory of industrial organization*. Cambridge, MA: MIT Press.

Trivers, R. (2011). *The folly of fools: The logic of deceit and self-deception in human life*. New York: Basic Books.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35–57.

Trivers, R. L. (1974). Parent-offspring conflict. *American Zoologist, 14*(1), 249–264.

Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature, 308*(5955), 181–184.

Wilson, D. S. (2010). *Darwin's cathedral: Evolution, religion, and the nature of society*. Chicago: University of Chicago Press.

Wilson, E. O. (2012). *The social conquest of earth*. New York: WW Norton & Company.

Wilson, D. S. (2006). Human groups as adaptive units: Toward a permanent consensus. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Culture and cognition*. Oxford, UK: Oxford University Press.

Winter, E. (2014). *Feeling smart: Why our emotions are more rational than we think*. Public Books.

Wroughton, L. (2013). As Syria War Escalates, Americans Cool to U.S. Intervention: Reuter/Ipsos Poll. *Reuters*, August 24, 2013. Accessed at http://www.reuters.com/article/2013/08/25/us syria crisis usa poll idUSBRE97O00E20130825 on 12 March 2015.

Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences, 110*(Suppl. 2), 10424–10429.

Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology, 53*(1), 205–214.

'millionaire's house' for a day or two, before returning with cash

He did say that she tried to get out of the sex business, but was dragged back in.

Last month it was revealed Roberts accused two male acquaintances she was raped, bringing her credibility into question.

Roberts' lawyers says the latest allegations do not contend the fact she was used for underage sex.

Attorney Sigrid McCawleysaid 'To say that our client acquiesced in this abuse, or that the abuse was OK because she was paid for it — leaves out the fact that this is why we have laws in the United States to protect minor children who are groomed and sexually trafficked by adults.'

Epstein served 13 months in a Florida prison after pleading guilty in 2008 to soliciting a minor for prostitution

Allegations: The 31-year-old is currently in the midst of a lawsuit claiming she was recruited to join Epstein's harem of underage women in 1999, naming Prince Andrew and attorney Alan Dershowitz as two of the men she was made to 'service'

The scandal resurfaced in January when Roberts named Prince Andrew in a lawsuit, accusing him of using her for underage sex

According to court records connected with Roberts' current lawsuit, she maintains that she was recruited to join Epstein's harem in 1999 by his friend, the socialite Ghislaine Maxwell.

She says during the three years she worked for Epstein, she was made to entertain his friends and named Prince Andrew and Dershowitz as two men she had sex with.

Both men have denied having sex with Roberts, and Dershowitz is trying to get his name removed from the suit, claiming Roberts made the entire story up.

Fergie vehemently defends Prince Andrew against 'defamation'

EXHIBIT G

United States District Court
Southern District of New York

VIRGINIA L. GIUFFRE,

Plaintiff, CASE NO:____________________

v.

GHISLAINE MAXWELL,

Defendant.

________________________________/

COMPLAINT

Boies Schiller & Flexner LLP
575 Lexington Avenue
New York, NY 10022
(212) 446-2300

Plaintiff, VIRGINIA L. GIUFFRE, formerly known as Virginia Roberts ("Giuffre"), for her Complaint against Defendant, GHISLAINE MAXWELL ("Maxwell"), avers upon personal knowledge as to her own acts and status and otherwise upon information and belief:

NATURE OF THE ACTION

1. This suit arises out of Defendant Maxwell's defamatory statements against Plaintiff Giuffre. As described below, Giuffre was a victim of sexual trafficking and abuse while she was a minor child. Defendant Maxwell not only facilitated that sexual abuse but, most recently, wrongfully subjected Giuffre to public ridicule, contempt and disgrace by, among other things, calling Giuffre a liar in published statements with the malicious intent of discrediting and further damaging Giuffre worldwide.

JURISDICTION AND VENUE

2. This is an action for damages in an amount in excess of the minimum jurisdictional limits of this Court.

3. This Court has jurisdiction over this dispute pursuant to 28 U.S.C. §1332 (diversity jurisdiction) as Giuffre and Maxwell are citizens of different states and the amount in controversy exceeds seventy-five thousand ($75,000), exclusive of interest and costs.

4. This Court has personal jurisdiction over Maxwell. Maxwell resides in New York City, and this action arose, and defamatory statements were made, within the Southern District of New York.

5. Venue is proper in this Court as the cause of action arose within the jurisdiction of this Court.

PARTIES

6. Plaintiff Giuffre is an individual who is a citizen of the State of Colorado.

7. Defendant Maxwell, who is domiciled in the Southern District of New York, is not a citizen of the state of Colorado.

FACTUAL ALLEGATIONS

8. Virginia Giuffre became a victim of sex trafficking and repeated sexual abuse after being recruited by Ghislaine Maxwell and Jeffrey Epstein when Giuffre was under the age of eighteen.

9. Between 1999 and 2002, with the assistance and participation of Maxwell, Epstein sexually abused Giuffre at numerous locations including his mansions in West Palm Beach, Florida, and in this District. Between 2001 and 2007, with the assistance of numerous co-conspirators, Epstein abused more than thirty (30) minor underage girls, a fact confirmed by state and federal law enforcement.

10. As part of their sex trafficking efforts, Epstein and Maxwell intimidated Giuffre into remaining silent about what had happened to her.

11. In September 2007, Epstein entered into a Non-Prosecution Agreement ("NPA") that barred his prosecution for numerous federal sex crimes in the Southern District of Florida.

12. In the NPA, the United States additionally agreed that it would not institute any federal criminal charges against any potential co-conspirators of Epstein.

13. As a co-conspirator of Epstein, Maxwell was consequently granted immunity in the Southern District of Florida through the NPA.

14. Epstein ultimately pled guilty to procuring a minor for prostitution, and is now a registered sex offender.

15. Rather than confer with the victims about the NPA, the U.S. Attorney's Office and Epstein agreed to a "confidentiality" provision in the Agreement barring its disclosure to anyone — including Epstein's victims. As a consequence, the victims were not told about the NPA.

16. On July 7, 2008, a young woman identified as Jane Doe No. 1, one of Jeffrey Epstein's victims (other than Giuffre), filed a petition to enforce her rights under the Crime Victims' Rights Act ("CVRA"), 18 U.S.C. ¶ 3771, alleging that the Government failed to provide her the rights promised in the CVRA with regard to the plea arrangement with Epstein. The litigation remains ongoing.

17. On or about May 4, 2009, Virginia Giuffre—identified then as Jane Doe No. 102—filed a complaint against Jeffrey Epstein in the United States District Court for the Southern District of Florida. The complaint included allegations made by Giuffre that pertained to Maxwell.

18. In pertinent part, the Jane Doe No. 102 complaint described in detail how Maxwell recruited Giuffre (who was then a minor girl) to become a victim of sex trafficking by introducing Giuffre to Jeffrey Epstein. With the assistance of Maxwell, Epstein was able to sexually abuse Giuffre for years until Giuffre eventually escaped.

19. The Jane Doe No. 102 complaint contained the first public allegations made on behalf of Giuffre regarding Maxwell.

20. As civil litigation against Epstein moved forward on behalf of Giuffre and many other similarly-situated victims, Maxwell was served with a subpoena for deposition. Her testimony was sought concerning her personal knowledge and role in Epstein's abuse of Giuffre and others.

21. To avoid her deposition, Maxwell claimed that her mother fell deathly ill and that consequently she was leaving the United States for London with no plans of ever returning. In fact, however, within weeks of using that excuse to avoid testifying, Maxwell had returned to New York.

22. In 2011, two FBI agents located Giuffre in Australia—where she had been hiding from Epstein and Maxwell for several years—and arranged to meet with her at the U.S. Consulate in Sidney. Giuffre provided truthful and accurate information to the FBI about Epstein and Maxwell's sexual abuse.

23. Ultimately, as a mother and one of Epstein's many victims, Giuffre believed that she should speak out about her sexual abuse experiences in hopes of helping others who had also suffered from sexual trafficking and abuse.

24. On December 23, 2014, Giuffre incorporated an organization called Victims Refuse Silence, Inc., a Florida not-for-profit corporation.

25. Giuffre intended Victims Refuse Silence to change and improve the fight against sexual abuse and human trafficking. The goal of her organization was, and continues to be, to help survivors surmount the shame, silence, and intimidation typically experienced by victims of sexual abuse. Giuffre has now dedicated her professional life to helping victims of sex trafficking.

26. On December 30, 2014, Giuffre moved to join the on-going litigation previously filed by Jane Doe 1 in the Southern District of Florida challenging Epstein's non-prosecution agreement by filing her own joinder motion.

27. Giuffre's motion described Maxwell's role as one of the main women who Epstein used to procure under-aged girls for sexual activities and a primary co-conspirator and participant in his sexual abuse and sex trafficking scheme.

28. In January, 2015, Maxwell undertook a concerted and malicious campaign to discredit Giuffre and to so damage her reputation that Giuffre's factual reporting of what had happened to her would not be credited.

29. As part of Maxwell's campaign she directed her agent, Ross Gow, to attack Giuffre's honesty and truthfulness and to accuse Giuffre of lying.

30. On or about January 3, 2015, speaking through her authorized agent, Maxwell issued an additional false statement to the media and public designed to maliciously discredit Giuffre. That statement contained the following deliberate falsehoods:

(a) That Giuffre's sworn allegations "**against Ghislaine Maxwell are untrue.**"

(b) That the allegations have been "shown to be untrue."

(c) That Giuffre's "**claims are <u>obvious lies</u>.**"

31. Maxwell's January 3, 2015, statement incorporated by reference "Ghislaine Maxwell's original response to the lies and defamatory claims remains the same," an earlier statement that had falsely described Giuffre's factual assertions as "entirely false" and "entirely untrue."

32. Maxwell made the same false and defamatory statements as set forth above, in the Southern District of New York and elsewhere in a deliberate effort to maliciously discredit Giuffre and silence her efforts to expose sex crimes committed around the world by Maxwell, Epstein, and other powerful persons. Maxwell did so with the purpose and effect of having

others repeat such false and defamatory statements and thereby further damaged Giuffre's reputation.

33. Maxwell made her statements to discredit Giuffre in close consultation with Epstein. Maxwell made her statements knowing full well they were false.

34. Maxwell made her statements maliciously as part of an effort to conceal sex trafficking crimes committed around the world by Maxwell, Epstein and other powerful persons.

35. Maxwell intended her false and defamatory statements set out above to be broadcast around the world and to intimidate and silence Giuffre from making further efforts to expose sex crimes committed by Maxwell, Epstein, and other powerful persons.

36. Maxwell intended her false statements to be specific statements of fact, including a statement that she had not recruited an underage Giuffre for Epstein's abuse. Maxwell's false statements were broadcast around the world and were reasonably understood by those who heard them to be specific factual claims by Maxwell that she had not helped Epstein recruit or sexually abuse Giuffre and that Giuffre was a liar.

37. On or about January 4, 2015, Maxwell continued her campaign to falsely and maliciously discredit Giuffre. When a reporter on a Manhattan street asked Maxwell about Giuffre's allegations against Maxwell, she responded by saying: "I am referring to the statement that we made." *The New York Daily News* published a video of this response by Maxwell indicating that she made her false statements on East 65th Street in Manhattan, New York, within the Southern District of New York.

COUNT I
DEFAMATION

1. Plaintiff Giuffre re-alleges paragraphs 1 - 37 as if the same were fully set forth herein. Maxwell made her false and defamatory statements deliberately and maliciously with the intent to intimidate, discredit and defame Giuffre.

2. In January 2015, and thereafter, Maxwell intentionally and maliciously released to the press her false statements about Giuffre in an attempt to destroy Giuffre's reputation and cause her to lose all credibility in her efforts to help victims of sex trafficking.

3. Maxwell additionally released to the press her false statements with knowledge that her words would dilute, discredit and neutralize Giuffre's public and private messages to sexual abuse victims and ultimately prevent Giuffre from effectively providing assistance and advocacy on behalf of other victims of sex trafficking, or to expose her abusers.

4. Using her role as a powerful figure with powerful friends, Maxwell's statements were published internationally for the malicious purpose of further damaging a sexual abuse and sexual trafficking victim; to destroy Giuffre's reputation and credibility; to cause the world to disbelieve Giuffre; and to destroy Giuffre's efforts to use her experience to help others suffering as sex trafficking victims.

5. Maxwell, personally and through her authorized agent, Ross Gow, intentionally and maliciously made false and damaging statements of fact concerning Giuffre, as detailed above, in the Southern District of New York and elsewhere.

6. The false statements made by Gow were all made by him as Maxwell's authorized agent and were made with direct and actual authority from Maxwell as the principal.

7. The false statements that Maxwell made personally, and through her authorized agent Gow, not only called Giuffre's truthfulness and integrity into question, but also exposed Giuffre to public hatred, contempt, ridicule, and disgrace.

8. Maxwell made her false statements knowing full well that they were completely false. Accordingly, she made her statements with actual and deliberate malice, the highest degree of awareness of falsity.

9. Maxwell's false statements constitute libel, as she knew that they were going to be transmitted in writing, widely disseminated on the internet and in print. Maxwell intended her false statements to be published by newspaper and other media outlets internationally, and they were, in fact, published globally, including within the Southern District of New York.

10. Maxwell's false statements constitute libel per se inasmuch as they exposed Giuffre to public contempt, ridicule, aversion, and disgrace, and induced an evil opinion of her in the minds of right-thinking persons.

11. Maxwell's false statements also constitute libel per se inasmuch as they tended to injure Giuffre in her professional capacity as the president of a non-profit corporation designed to help victims of sex trafficking, and inasmuch as they destroyed her credibility and reputation among members of the community that seeks her help and that she seeks to serve.

12. Maxwell's false statements directly stated and also implied that in speaking out against sex trafficking Giuffre acted with fraud, dishonesty, and unfitness for the task. Maxwell's false statements directly and indirectly indicate that Giuffre lied about being recruited by Maxwell and sexually abused by Epstein and Maxwell. Maxwell's false statements were reasonably understood by many persons who read her statements as conveying that specific intention and meaning.

13. Maxwell's false statements were reasonably understood by many persons who read those statements as making specific factual claims that Giuffre was lying about specific facts.

14. Maxwell specifically directed her false statements at Giuffre's true public description of factual events, and many persons who read Maxwell's statements reasonably understood that those statements referred directly to Giuffre's account of her life as a young teenager with Maxwell and Epstein.

15. Maxwell intended her false statements to be widely published and disseminated on television, through newspapers, by word of mouth and on the internet. As intended by Maxwell, her statements were published and disseminated around the world.

16. Maxwell coordinated her false statements with other media efforts made by Epstein and other powerful persons acting as Epstein's representatives and surrogates. Maxwell made and coordinated her statements in the Southern District of New York and elsewhere with the specific intent to amplify the defamatory effect those statements would have on Giuffre's reputation and credibility.

17. Maxwell made her false statements both directly and through agents who, with her general and specific authorization, adopted, distributed, and published the false statements on Maxwell's behalf. In addition, Maxwell and her authorized agents made false statements in reckless disregard of their truth or falsity and with malicious intent to destroy Giuffre's reputation and credibility; to prevent her from further disseminating her life story; and to cause persons hearing or reading Giuffre's descriptions of truthful facts to disbelieve her entirely. Maxwell made her false statements wantonly and with the specific intent to maliciously damage Giuffre's good name and reputation in a way that would destroy her efforts to administer her

non-profit foundation, or share her life story, and thereby help others who have suffered from sexual abuse.

18. As a result of Maxwell's campaign to spread false, discrediting and defamatory statements about Giuffre, Giuffre suffered substantial damages in an amount to be proven at trial.

19. Maxwell's false statements have caused, and continue to cause, Giuffre economic damage, psychological pain and suffering, mental anguish and emotional distress, and other direct and consequential damages and losses.

20. Maxwell's campaign to spread her false statements internationally was unusual and particularly egregious conduct. Maxwell sexually abused Giuffre and helped Epstein to sexually abuse Giuffre, and then, in order to avoid having these crimes discovered, Maxwell wantonly and maliciously set out to falsely accuse, defame, and discredit Giuffre. In so doing, Maxwell's efforts constituted a public wrong by deterring, damaging, and setting back Giuffre's efforts to help victims of sex trafficking. Accordingly, this is a case in which exemplary and punitive damages are appropriate.

21. Punitive and exemplary damages are necessary in this case to deter Maxwell and others from wantonly and maliciously using a campaign of lies to discredit Giuffre and other victims of sex trafficking.

PRAYER FOR RELIEF

WHEREFORE, Plaintiff Giuffre respectfully requests judgment against Defendant Maxwell, awarding compensatory, consequential, exemplary, and punitive damages in an amount to be determined at trial, but in excess of the $75,000 jurisdictional requirement; costs of suit; attorneys' fees; and such other and further relief as the Court may deem just and proper.

JURY DEMAND

Plaintiff hereby demands a trial by jury on all causes of action asserted within this pleading.

Dated September 21, 2015.

/s/ David Boies
David Boies
Boies Schiller & Flexner LLP
333 Main Street
Armonk, NY 10504

/s/ Sigrid McCawley
Sigrid McCawley
(Pro Hac Vice Pending)
Boies Schiller & Flexner LLP
401 E. Las Olas Blvd., Suite 1200
Ft. Lauderdale, FL 33301
(954) 356-0011

/s/ Ellen Brockman
Ellen Brockman
Boies Schiller & Flexner LLP
575 Lexington Ave
New York, New York 10022
(212) 446-2300

EXHIBIT H

IN THE CIRCUIT COURT OF THE SEVENTEENTH JUDICIAL CIRCUIT, IN AND FOR BROWARD COUNTY, FLORIDA

CASE NO.: CACE 15-000072

BRADLEY J. EDWARDS and PAUL G. CASSELL,

Plaintiff(s),

vs.

ALAN M. DERSHOWITZ,

Defendant(s).

_______________________________________/

NOTICE OF SERVING ANSWERS TO INTERROGATORIES

Plaintiffs, Bradley J. Edwards and Paul G. Cassell, by and through their undersigned counsel, hereby file this Notice of Serving Answers to Interrogatories with the Court propounded by the Defendant, ALAN M. DERSHOWITZ, on February 11, 2015, and that a copy has been furnished to the attorney for the Defendant.

I HEREBY CERTIFY that a true and correct copy of the foregoing was sent via E-Serve to all Counsel on the attached list, this 13th day of March, 2015.

JACK SCAROLA 750263
Florida Bar No.: 169440
Attorney E-Mail(s): jsx@searcylaw.com and mep@searcylaw.com
Primary E-Mail: _scarolateam@searcylaw.com
Searcy Denney Scarola Barnhart & Shipley, P.A.
2139 Palm Beach Lakes Boulevard
West Palm Beach, Florida 33409
Phone: (561) 686-6300
Fax: (561) 383-9451
Attorneys for Plaintiffs

13. Describe in detail Each instance in which Jane Doe #3 has provided information referencing Dershowitz by name that Concern the allegations set forth in Paragraphs 24-31 of the 2015 Jane Doe #3 Declaration.

ANSWER: Edwards and Cassell lack sufficient information to determine all circumstances in which Jane Doe No. 3 has mentioned to others Dershowitz's name as someone who abused her or had information relevant to abuse.

With regard to when she has provided information related to this subject to them, Jane Doe No. 3 provided such information in telephone calls with Brad Edwards beginning in 2011.

Jane Doe No. 3 has also provided this information in a public affidavit, filed on January 21, 2015, in the CVRA case. Jane Doe No. 3 has also provided similar information on other occasions, but the specifics of those communications are protected by the attorney-client privilege and the work product doctrine.

14. If You have ever seen a photograph or video of Jane Doe #3 with Dershowitz, then state when You saw the photograph or video, identify who took the original photograph or video, identify Each person who possesses a copy of the photograph or video, and state the location of Each such original and copy.

ANSWER: Edwards and Cassell have not personally seen such a photograph or video. Discovery efforts to obtain photographic materials regarding Jane Doe No. 3 held by the U.S. Attorney's Office for the Southern District of Florida and/or other federal law enforcement and prosecuting agencies are on-going.

15. For Each communication between You or anyone acting on Your behalf, and anyone from, or acting on behalf of, any media outlet Concerning this action, the Joinder Motion, or Dershowitz, and regardless of whether such communication was "on the record" or "off the record," (a) state the date of the communication; (b) state the participants in the communication; and (c) describe the contents of the communication.

ANSWER: Objection, not reasonably calculated to lead to the discovery of admissible evidence; vague, harassing, work-product.

16. Describe in detail All facts Concerning any assertion that Dershowitz was a "coconspirator" with Epstein.

ANSWER: See answers to interrogatory number 5 above, as well as answers to interrogatories numbers 1, 6, 8, 9, 10 above. In addition, factual information is found in the documents and other materials and references provided in answer to Request for Production Number 2

PAUL G. CASSELL

STATE OF UTAH)

COUNTY OF SALT LAKE)

The foregoing instrument was acknowledged before me this 11 day of March, 2015 by Paul Cassell who is personally known to me or who has produced personal knowledge (type of identification) as identification and who did/did not take an oath.

Susan Baca
Notary Public
~~State of Florida at Large~~ State of Utah SMB
My Commission expires:
Commission No:

Notary Public
SUSAN BACA
Commission #665643
My Commission Expires
April 19, 2017
State of Utah

Bradley J. Edwards

STATE OF FLORIDA)
)
COUNTY OF Broward)

BEFORE ME, the undersigned authority, personally appeared this day Bradley J. Edwards, who [✓] is personally known to me *or* [ ] produced ____________________, as identification, and who, after being duly sworn, did state that he/she executed the foregoing Answers to First Set of Interrogatories and that the same are true and correct to the best of his knowledge

Subscribed and sworn to before me this 12t day of March, 2015.

MARIA W. KELLJCHIAN
MY COMMISSION #EE840478
EXPIRES: OCT 03, 2016
Bonded through 1st State Insurance

Signature of Notary Public, State of Florida

Print / Typed Name, Notary Public

My commission expires:

EXHIBIT I

IN THE CIRCUIT COURT OF THE SEVENTEENTH
JUDICIAL CIRCUIT IN AND FOR
BROWARD COUNTY, FLORIDA

CASE NO. CACE 15-000072

BRADLEY J. EDWARDS and PAUL G. CASSELL,

Plaintiffs/Counterclaim Defendants,

vs.

ALAN M. DERSHOWITZ,

Defendant/Counterclaim Plaintiff.
____/

VIDEOTAPED DEPOSITION OF

PAUL G. CASSELL

TAKEN ON BEHALF OF THE DEFENDANT

VOLUME I, PAGES 1 to 151

Friday, October 16, 2015

1:33 p.m. - 4:31 p.m.

110 Southeast 6th Street
110 Tower - Suite 1850
Fort Lauderdale, Florida 33301

Theresa Tomaselli, RMR

let me clear all of that misunderstanding up. You know, that's -- frankly, if I had gotten something like that, that's what I would have said.

The answer that came back was -- from Mr. Dershowitz was something along the lines of, if I remember correctly, well, tell me what you -- you -- tell me what you want to know and I'll decide whether to cooperate, was I think the phrase that was used. And -- and so there was an attempt, you know, a 2009 attempt, a 2011 attempt to get information from Mr. Dershowitz.

Then there was another subpoena without deposition for -- for documents. You know, we have heard a lot about records in this case that could prove innocence. There was a records request to Mr. Dershowitz in 2013. And, again, my understanding was that there was no -- you know, no documents were provided on that.

And so those -- I had that information. Another bit of information that I had was that in 2011, I believe in early April -- this is not attorney/client privileged information from Virginia Roberts. This is a telephone call that she placed from Australia where she had been

EXHIBIT J

IN THE CIRCUIT COURT OF THE FIFTEENTH JUDICIAL CIRCUIT, IN AND FOR PALM BEACH COUNTY, FLORIDA

CASE NO.: 502009CA040800XXXXMBAG

JEFFREY EPSTEIN,

Plaintiff,

vs.

SCOTT ROTHSTEIN, individually,
BRADLEY J. EDWARDS, individually,
and L.M., individually,

Defendants.
_______________________________________/

NOTICE OF FILING

COMES NOW the Defendant/CounterClaimant, BRADLEY EDWARDS, by and through his undersigned counsel, and hereby files the attached transcript of the telephone interview of Virginia Roberts to supplement the proffer made in support of Counter-Claimant's Motion for Leave to Amend to Assert Punitive Damages.

I HEREBY CERTIFY that a true and correct copy of the foregoing has been furnished by U.S. Mail to all Counsel on the attached list on this 17 day of May 2011.

____
Jack Scarola
Florida Bar No.: 169440
Searcy Denney Scarola Barnhart & Shipley, P.A.
2139 Palm Beach Lakes Boulevard
West Palm Beach, Florida 33409
Phone: (561) 686-6300
Fax: (561) 383-9451
Attorney for Defendant/CounterClaimant
Edwards

COUNSEL LIST

Jack A. Goldberger, Esquire
jgoldberger@agwpa.com;
smahoney@agwpa.com
Atterbury, Goldberger & Weiss, P.A.
250 Australian Avenue South, Suite 1400
West Palm Beach, FL 33401
Phone: (561)-659-8300
Fax: (561)-835-8691
Attorneys for Jeffrey Epstein

Martin Weinberg, Esquire
Martin Weinberg, P.C.
20 Park Plaza, Suite 1000
Suffolk, MA 02116
Attorneys for Jeffrey Epstein

Farmer, Jaffe, Weissing, Edwards, Fistos & Lehrman, PL
425 N. Andrews Avenue, Suite 2
Fort Lauderdale, FL 33301
Phone: (954)-524-2820
Fax: (954)-524-2822
Attorneys for Jeffrey Epstein

Joseph L. Ackerman, Jr., Esquire
jla@fowler-white.com
Fowler White Burnett, P.A.
901 Phillips Point West
777 S Flagler Drive
West Palm Beach, FL 33401-6170
Phone: (561)-802-9044
Fax: (561)-802-9976
Attorneys for Jeffrey Epstein

Marc S. Nurik
marc@nuriklaw.com
Law Offices of Marc S. Nurik
One E Broward Blvd., Suite 700
Fort Lauderdale, FL 33301
Phone: (954)-745-5849
Fax: (954)-745-3556
Attorneys for Scott Rothstein

PRIVILEGED PURSUANT TO FS 766.205(4) and/or WORK PRODUCT

TELECON

PARTICIPANTS: **JACK SCAROLA**
BRAD EDWARDS
VIRGINIA ROBERTS

RE: **Edwards adv. Epstein**
291874

DATE: **April 07, 2011**

JS: Virginia, Jack Scarola and Brad Edwards

BE: Hi Virginia.

V: Hi Jack! Hi Brad! How you guys doing?

JS: We're doing fine, thank you. I'm sorry for all of the trouble and before we go any further, let me tell you, if I have your permission, I have started a tape recorder and I want to be able to tape this conversation from the very beginning. Is that alright with you?

V: Sure, that's fine, Jack. No problem

JS: Ok, good, thank you. I appreciate that. Let me start off by introducing myself. I know that Brad has spoken to you about me but I am Brad's lawyer, and I assume that you can confirm that you and I have never had any communication before. Is that right?

V: That's correct.

JS: Alright. I have, however, gotten some information from Brad about conversations that you have had with him, and that will enable me, hopefully, to make this a little bit more efficient and take up a minimum amount of your time while still getting the information that we think is going to be helpful to us and to any jury that might ultimately have to hear these facts.

So, let me begin by asking you first to tell us what your full name is.

V: Virginia Louise Roberts. That's my maiden name. My married name is Virginia Louise [redacted].*

JS: Could you spell your last name for us? That is your married last name:

V: [redacted].*

JS: Alright, thank you, and where are you living right now?

V: I live in Australia.

JS: And how long have you resided in Australia.

V: This is my 19th year.

JS: That is where you are right now, correct? We've reached you in Australia for this phone conversation?

V: That is correct, yes.

JS: And what time is it in Australia right now?

V: I think it's about 9:00 now.

JS: Ok. That's 9am, correct?

V: That's correct.

JS: Alright. Virginia, the reason for this conversation is because it is our understanding that you know a man by the name Jeffrey Epstein, and I want to begin by asking you please to tell us about the circumstances of your first meeting Mr. Epstein.

V: Ok. I was introduced to Mr. Epstein by Ghislaine Maxwell. I was working at Donald Trump's spa in Mar-a-Lago and I was prompted by Ghislaine to come to Jeffrey's mansion in Palm Beach that afternoon after work to make some extra money and to learn about massage. She met me at the spa, and I was reading a book about anatomy, so I was already interested in massage therapy as it was and not having any of the education or you know anything behind me, I thought this was a great opportunity to work for her and go. So, I went to Jeffrey's mansion about 5 or 6 in the afternoon. My dad drove me there. My dad worked at Mar-a-Lago with me, and he met Ghislaine and she seemed like a nice, proper English lady, and she knows, I mean, you know, one time then _once before I left to travel overseas, she just seemed really nice and like she would like to help me out. So my dad left, and I had no problem getting home that night, one of her drivers would take me back after my trial. So she led me upstairs, and into Jeffrey's bedroom, and past that is Jeffrey's massage room, which has got his steam room and a shower and a massage table, and there is actually an extra room that has, that nobody knows about it, it's kinda like a secret room and it's got a whole bunch of decorative pictures of pornographic literature and sex toys and I can ___?___ what happened in there.

JS: When you say that the room was hidden, Virginia, how was the room concealed?

V: It wasn't like a door that you would normally go into, like some kind of special opening, you open that and then a little door, so it looks like it's a little closet so-to-speak, but when you walk in there, it's obviously a lot bigger than just a closet. It wasn't too big, but it was bigger, you know. It wasn't a gigantic room, it was just like a small room, which you know, it probably could fit some shoes in there, _it had racks of shoes, boxes, some sweatshirts neatly folded, and the ceiling to the floor was covered in pornographic pictures of the girls that he had met.

JS: When you say…

V: So anyways, that was getting there, and I was introduced to Jeffrey, he was laying naked on top of the massage table, and obviously for one, I'm a 15 year old girl and seeing him on the table was weird but, also learning about anatomy and massage, I thought this would be part of it. So obviously, I thought it was part of the massage program, so I said ok, this is fine. And, he then instructed me on how to touch the body, Jeffrey's body, how to massage him, and for the first hour, it was actually a real massage, maybe not an hour, maybe like 40 minutes or something, but of something like that _and that's when he turned over on the other side and to expose himself fully. So then Ghislaine told me that she wanted me to undress and began to take off my shirt and skirt, my white uniform from Mar-A-Lago, she also took off her shirt and got undressed, and so I was there with just my undies on, and she was completely bare, and made some kind of little flake about the underwear that I was wearing because it wasn't my normal sexy girl underwear and just like, I don't know, had red hearts on it or something like that; just your normal, you know, real cute underwear. Anyways, so during all of this I'm kind of like what's going on, how do I act, what do I say, I was so afraid of, not afraid or fearful for my life but _unsure of how all this started and wanting to obtain a profession_ I was so afraid thinking about upsetting and disappointing them, I don't know, it's a weird situation by far and I was expected to _Lick his nipples, instructed on how to do so by J.E_ and give him oral sex while he wanted to fondle me, and then at the end, I was told by Ghislaine to get on top and straddle Jeffrey sexually, and when we were done, we went and had a shower in the room and Jeffrey told me to wash him up and down, you know with a bar of soap and make sure he was all cleaned up. And then he took me downstairs and took me to two of the guards and told John to bring me home. John was the butler at the time.

JS: Let me interrupt you for just a moment there if I could, Virginia. You mentioned as you were recounting those details that you were 15 years old at the time. What is your date of birth?

V: August 9th, 1983

JS: And can you tell us please, as best you're able to estimate it, what the date was when this first encounter occurred?

V: I've got it written down. It's like - I'm not good with math – hold on – I thought I wrote it down but I didn't. I'm not too sure, I think it was 1998 off the top of my head and around June of 1998, I would say as I was turning 16 at the end of the summer.

JS: Alright. You talked about the room where Jeffrey had the pornographic photographs. Did you actually see that room on the occasion of your first visit there?

V: No, I got to see that room a few visits after but I was just trying to describe that room to you guys so you knew exactly what room I was talking about.

JS: Another question for you, and I don't mean to be prying into your personal life, and if I ask you any questions at all that you're uncomfortable answering, then you just tell me that and we'll move on, because I appreciate your cooperation and the last I thing I want to do is impose upon that cooperation, but can you tell us please just generally what kind of sexual experience you had had prior to this confrontation with Jeffrey?

V: Yeah, sure. A close family friend has sexually abused me, and I was on the streets at 13 years old. I was picked up by a 67 year old man named Ron _Eppinger_ who did exactly what Jeffrey did with me abuse and violate my youthfulness _ & I was with him for 6 months. So, he was gone and then I had this boyfriend who was like my school friend from young days but we just kept in contact with each other and we were on and off constantly, and that was Tony Figeroua_, and there was also another younger guy was near my age, Michael, I can't remember his last name, but yeah, there, I mean, there wasn't like a string of men or anything, but there was Ron, like I told you, and he was the first guy expecting me to do so-called disgusting affairs. Jeffrey actually knew Ron, which was quite weird when I told Jeffrey the story about Ron, and Jeffrey had actually met him, and yeah. Anyway, just another story, and yeah, there were a couple of men, but that gives you an idea.

JS: Now when you described the photographs in the room as pornographic, tell me just a little bit about the photographs, if you would please? First, how many of them were there?

V: At least 100, and like I said, they covered the room from the ceiling, not the ceiling but from the top of the edge of the wall to the bottom of the floor. I want to say at least a hundred, even more, there could have been more in the boxes _Some of them were A4_ photographs, like the large size, some of naked women posing, you know, positions, sexy positions. Others were, you know, some girls had bikinis on, and it wasn't so pornographic, but it was all women, and it was all in a sexual nature.

JS: Were there photographs where more than one person appeared?

V: Oh yeah, plenty of them. There were lots of naked photographs, I mean I was just trying to give you a visual range. There was anything from 5x6s to 4x8s to 8x4s. Some of them had frames, some of them were out of frames, but they were all, like I said, they were all women, they were all sexual in nature.

JS: Was it your impression that there were a hundred photos of a hundred different people or were there multiple photos of the same woman or girl?

V: There was definitely a lot of different girls. I mean it wasn't easy to sit there and say, you weren't finding 5 girls out of some photos, no. Were there a hundred different ones? There could have been pictures of some girls, I really couldn't get close necessarily to actually recognize faces or anything like that. But if you, you know, the range of them were all different, majority of them were different, yeah.

JS: Did there ever come a point in time when you became aware that a photograph of you had been added to the collection?

V: Yes, there was. Ghislaine took several nude photographs of me for Jeffrey. So, yeah, there were pictures of me and there were pictures, he wasn't shy, that wasn't the only place in his house that he kept the photos. He liked photos all over his house. If you looked in his den or on his desk or in on the hall table, a giant hall table in his house, there were at least a hundred photos of girls in frames. Not all of them were naked, a lot of the ones that were all around his house were not naked girls posing pornographically, some were pictures of celebrities and politicians he had known_or things like that or had pants on or whatever, but yeah, there was a lot of mixed photographs in the outside ones.

JS: Were there any photographs of girls or young women that you knew or that you subsequently came to know that you saw in the house?

V: Yeah, yeah, there was. There was pictures of Nadia Bjournik__________, pictures Sarah Keller, pictures Emmy, pictures of me, pictures of the regulars, but a lot of the girls, sometimes Jeffrey could have like 7 girls a day, and he would only see those girls once if he got bored. I don't know. These weren't my days. I heard he's gotten a lot sloppier since I left. So, I don't know anyways, but when I knew him, there was just a, it seemed, there was such an influx of girls coming in and out, so did I recognize a lot of them? Maybe, maybe not, but then they were all definitely beautiful, they were all ranging in age, some of them young, some of them older in their 20's, I mean it was just they were all beautiful.

JS: You've told us about the first visit. Was there any discussion on the occasion of that first visit about your returning?

V: Yes, they were very pleased with me and after the encounter was finished, the sexual encounter, he went and told me I did well and I have a lot of potential to become a massage therapist and if I'd like I could return tomorrow, you know, and do the same thing and get paid $200/hr, so Jeffrey insisted that I come after work, and over the next few days, I guess the relationship grew into more, and within a couple of weeks, not even a couple of weeks, maybe a week, I had quit Mar-a-Lago and I was working for Jeffrey full time.

JS: Ok, let's talk about your job at Mar-a-Lago, if we could. You said that your Dad was working there. What was his position at Mar-a-Lago?

V: He was a maintenance supervisor I think is what it was called? He like managed the tennis courts and air conditioners and things like that.

JS: What is your Dad's full name?

V: Sky William Roberts.

JS: And is he still living here in South Florida now?

V: No, he's not, he's in California.

JS: Ok. Is your Dad aware of what is currently going on with regard to your having made public statements about your relationship with Jeffrey?

V: Yes, he is well aware of it. I told my family even before all this stuff came out, because they were the first ones contacted by the journalists from Mail on Sunday. I know that they the Mail on Sunday printed that I had gone out and tried to, I mean I think one of the photos said that I was angry that I saw Jeffrey and the Prince walking together and that is why I came out and went public with everything. Not true. I mean, I am angry about how they are still up to their old ways together and that they're still hanging out but I didn't contact the Mail on Sunday and I didn't bring it out. I figured that everyone was going to bring it out anyway and I better bring it out the right way. He's known everything from the start, and my family is very supportive with everything going on.

JS: I'm kind of going to jump around a little bit and I apologize for that, but since the subject has come up, tell me first of all why you are providing this cooperation to us, and I am certainly very appreciative of it, but I want you to tell us why it is you've chosen to spend time with us on the telephone and provide this information that you're now providing.

V: I'm out to help the bigger picture, you know, I think all of us can make a big difference in a lot of other people's lives and I think that this has gone on long enough and it's a big slap in my face that he can get away with hurting me so bad let alone so many other girls and laugh about it. I guess I talked to you guys out because I want to see the right thing happen, not just to him, but I want people in the world to understand this is not the way of life, you know, it's not acceptable to go out procure young girls and make them think that, this is the way you should be living and that's all. Yeah, I guess my reason for doing it is to help the bigger picture, you know, I'm a big believer in karma and I believe that good things will come back to you, so I guess that's why I'm doing this.

JS: Just for the record, neither Brad, nor I, nor anyone representing that they have anything to do with us has made any promises to you. Is that correct?

V: That is correct. I'll tell you, since this is our first conversation, that nobody has made me feel like I've been bribed or bought or had to say anything. I've told you anything that I know from my own self, not some things somebody told me. If anything, Brad's been extra careful not to tell me anything and let me do all the talking, so it's quite opposite I think.

JS: Alright. Let's get back then to the story of your relationship with Jeffrey, and we've talked about your first encounter with him and how it evolved from that into your full time employment with Jeffrey, but what were you doing at Mar-a-Lago before you quit Mar-a-Lago?

V: I was just a locker room attendant and sometimes I did babysitting for the rich and famous. So, I wasn't anything big. I worked in the spa area. That's why I was studying anatomy, because I was really really interested in becoming a _massage therapist_, and at the locker room, I didn't do much. I mean I was making tea for a living, I would, you know, make sure the toilet paper had a little triangle in it after everybody went to the toilet, or wipe down the water from the basin, you know, it was a very easy peasy job.

JS: Did you get that job through your Dad?

V: Yes, my Dad got me the job.

JS: Ok, and you were only 15 years old at the time, were they aware of how old you were at Mar-a-Lago?

V: Of course, definitely. We had to go through extensive, you know, we even had to get drug tested and id test and so on and so forth. I mean, Mantas (?) is very strict on employment, yeah, everybody knew.

JS: Ok. Was there ever any conversation with Ghislaine about how old you were before you were taken to Jeffrey's mansion?

V: No. She didn't ask me how old I was from the start, but when I did get to Jeffrey's mansion, it was discussed how old I was.

JS: With whom?

V: During the entire hour of what I call the legitimate massage I was giving him, it was cat and mouse games getting information from me to find out who I am, am I a willing participant in these kind of things, and how would I react if they were about to take the next step. But they got information off of me, they got my age, they got my, a little bit of my history so they knew I was, you know, not very stable at home, and they knew that, you know, I was actually interested in making my life better by studying so what they were offering me was a chance to become a legitimate masseuse but it was getting trained. They would have people show me how to work the body and be called a massage therapist and get me books on it, and you know, keep me interested, and every time, you know, I was with Jeffrey, literally was about massages, I don't mean just going in and have sex with him. I mean massage, because it would always start out with massage and then it would lead into sometimes other things.

JS: Alright, once this evolved into full time employment, what did full time employment mean?

V: That was entirely having to travel with Jeffrey in every city. When he was in Palm Beach, I stayed at my apartment, and he would call me to his house once or twice a day sometimes, and that's, you know, do things with him. Sometimes we'd go out shopping, sometimes we'd go out and watch a movie. You know, simple things like that, go to an expo or a fair, whatever it was. But when we were in other cities, I was at my apartment_, I lived with him full time. What I mean by full time is even in the middle of the night, I could get a ring on my phone next to me and tell me to come in his room, you know, so it was literally full time.

JS: When you say that when you were in Palm Beach you were living in your apartment, were you living on your own or were you living with members of your family at that time?

V: No, after I quit Mar-a-Lago, Jeffrey offered to get me an apartment in Palm Beach somewhere, Royal Palm Beach, and it was a nice apartment. He furnished it for me, it was absolutely beautiful, but yeah, that's the only time I would spend time away from him really.

JS: This apartment was on Royal Palm Beach Boulevard or out in the Village of Royal Palm Beach?

V: I so honestly don't remember. I've been trying to rack my brain because the FBI was asking the same thing and were trying to find it, but yeah, I'm Not sur_. I didn't get to spend as much time in it, I was only there about an entire week out of every month probably, but the majority of the time I was with Jeffrey anyways. It was somewhere in Royal Palm Beach. I don't know about Royal Palm Beach drive. I don't even remember the Royal Palm Beach drive anymore so I'm not too sure.

JS: Ok, let me see if I can draw the distinction for you and maybe that will help you to help us? Royal Palm Beach is a village that is…

V: No, no, no, I got Royal Palm Beach, I just didn't know the Royal Palm Beach Drive, like what street it's off of. Were you talking about a street?

JS: Yes, I was talking about a street. Royal Palm Drive is on the island of Palm Beach, and it's a street that is lined with large royal palm trees, and I'm wondering if this was an apartment on the island or was this an apartment out west of town…

V: No, it was actually in Royal Palm Beach, not on the island.

JS: Alright.

V: I would be driven, it was closer to my family than it was closer to him. I wanted to be close enough to everybody else so that when I was in town, I could just go see them quickly.

JS: Ok. So we're not talking about Royal Palm Boulevard. We're talking about the town of Royal Palm west of town. Jeffrey got you an apartment out there.

V: That's correct.

JS: When he was in Palm Beach, you were generally not staying at the mansion, you were staying at the apartment that he got for you out west of town.

V: That's correct.

JS: Ok.

V: I mean then there was times, I don't wanna say that every time I stayed at my apartment. There was times we'd fly back from some city maybe too late at night to really want to go back home, so you know, it's like 12:00 at night or 1:00 in the morning. I was just staying in the yellow room, or something like that; one of the guest rooms in Palm Beach. But majority of the time, I would definitely want to get back to my own apartment.

JS: Alright. What were the general hours of your full time employment when …?

V: There was not set hours. It wasn't like logging, and you know, hitting the shift button, nothing like that. The way I would get paid would be, ok, if I was in Palm Beach, I would get $200 an hour to massage Jeffrey or some of his friends and then go home. So it would be like that. If I was traveling with him, it would be per massage, so I would be getting paid per day. So I wouldn't be getting paid on an hourly rate. He wouldn't say ok, today you're going to work for me from 7:00 in the morning until 8:00 at night. It never like that. I was on call all the time.

JS: When you were here in Palm Beach, were you actually getting paid only for the time spent massaging Jeffrey or were you getting paid from your arrival at his house until you left the house?

V: From the time the massage started.

JS: From the time the massage started. Ok.

V: Sometimes we'd go there and I would wait for a while or talk with Ghislaine and Jeffrey about something or we'd meet somewhere and talk about something. A lot of times, I'd meet him upstairs in his room where he was ready for me. But then there was a lot of times where it didn't start right away, so he couldn't really pay me from the time I got there sometimes unless it was just paid from the time I massaged him til the time the massage was over.

JS: Alright. Did your duties for Jeffrey ever include anything other than providing him massages and sex in connection with the massages? Did he ever give you any other responsibilities to perform?

V: I was asked to do the same things that I did to Jeffrey to a few of his fellow colleagues as well. Those were my duties. He looked at it this way is that I was going to be a professional massage therapist, and maybe I needed some clientele, so he had me perform erotic massages on a few people.

JS: Did that start here in Palm Beach County?

V: It did. The first one did.

JS: Ok, and how long after you first met Jeffrey did he first ask you to provide services for one of his friends?

V: About 9 months, I think it was. It wasn't a full year, it wasn't 6 months, but between 6 months and a year, which is why I'm saying 9 months.

JS: And when you provided services to a friend of Jeffrey's, who paid you for those services?

V: Jeffrey would. I would get paid the next time I saw Jeffrey, so if I was invited to the Breakers Hotel to give a massage, I would give a massage, I would go home, and the next day when I saw Jeffrey, he would pay me for what I did. So, it was paid always by him, it was set up by him, so he always knew what to pay me. I did get tips and things like that, if you call it that, you know, like a hundred dollar tip or something from a few of them, you know, yeah.

JS: Was there ..

V: There was always payment from Jeffrey.

JS: Was there ever any discussion with Jeffrey about what was expected to happen when you provided massage services to one of Jeffrey's friends?

V: In a roundabout way, yes. In so many ways, Jeffrey really really had to train me, and that was why Ghislaine said that she and Jeffrey enjoyed me so much was because they never really had to speak much to me to tell me what they wanted me to do. You know, I wasn't waiting for you know, their directions. Jeffrey would tell me to go give an erotic massage to friends. He wouldn't give me much detail about it, but he would say to treat them like you treat me.

JS: Did he refer to it as an erotic massage or are those your words?

V: Erotic massage is my words. That's exactly what it was, but he would tell me to treat them how he wanted it, so I'd do what he wanted without having to say to me words more. I mean, I complied with what he wanted because it was somewhat of a, I don't know, I don't know how to say it, it was just very mindboggling how I let him have so much control or power over me basically. The massages would be routine to what Jeffrey wanted with my so called new clientele, and with their own words would ask me to provide them with sexual pleasure after the massage.

JS: Did you ever report back to Jeffrey about what happened when you provided massages to his friends?

V: Of course, of course, and I knew that his friends were reporting back to him as well because there were times where he would instigate conversation by saying you know, so and so had a great time, you did wonderful, you know so and so gave me a call and told J.E how it went_...

JS: Did Jeffrey ever elicit details from you? "Tell me what happened, describe in detail what went on?"

V: No, but he would have a laugh, he had a laugh with me a few times about some of their different mannerisms, I guess you would say, like some of them, one guy had a foot fetish and that was really weird and I mentioned it to Jeffrey, and we would have a laugh over it. He didn't want to know details. He wasn't asking me "so tell me what did you guys do exactly." No, he just basically gave me a slap on the back and said, you know, good job. And we had some kind of conversation about it. I can't recall any conversation off the top of my head. I really don't know one. It's been that long. But yeah, we did talk about it briefly.

JS: Can you give me an estimate as to the number of friends for whom Jeffrey provided and paid for your services?

V: There was about, you know, I don't know, 8 guys possibly.

JS: And are you able to name those people for me?

V: No, not at this stage. I just, some of these people are really influential in power, and I don't want to start another shitstorm with a few of them. I'll tell you that there was some erotic massages given to, I'm just afraid to say it to you.

JS: Ok, Virginia.

V: It's like geez, I don't know if I want to, I'm really scared of where this is gonna go.

JS: Alright. I understand that, and as I told you from the beginning, if I ask you a question that you are uncomfortable answering, you just tell me that, and I will move on, and I understand that at least right now, you are uncomfortable answering, and I am certainly going to respect that.

V: Thank you so much, Jack.

JS: No, that's quite alright. I am very appreciate of the cooperation you are providing, and I don't want you at any time to feel that we are taking unfair advantage of that cooperation, so give me the information that you're comfortable giving me, and if we get to a point where you're uncomfortable, I will respect that and we'll move on from there.

V: Ok.

JS: I want to talk a little bit about the traveling that you did with Jeffrey. About how long into your relationship with him did that first start?

V: Immediately. I started traveling immediately. Not internationally until I think about, Gosh, I can't remember even, I think it was a year later that we started doing international travel. Maybe like 9 months to a year again. Not too sure to be honest.

JS: So that would have been approximately the summer of 1999? Somewhere around there?

V: Yes. Somewhere around there. Somewhere around a year, somewhere around there, I can't pinpoint it exactly. But like I said, we started doing domestic traveling immediately, so my first destination with him was New York and Santa Fe and the Carribean, California, I would take trips with him occasionally. Sometimes we would go to St. Louis or New Orleans or Santa Cruz. We were traveling just about everywhere I think.

JS: How did you travel?

V: Well, we took Jeffrey's private jet, and unless I was being sent somewhere by myself for what we were just talking about before, then I would travel on a what do you call, a public jet, whatever it is...

JS: commercial flight?

V: Yes. Just a normal flight, an e-ticket.

JS: Like the rest of us common folk.

V: But when I was traveling with Jeffrey, the majority of the time would be on the black jet.

JS: Now, when you say there would be times when you would travel by yourself because he was sending you somewhere, tell me about that. How did that come about?

V: So, one of his colleagues would be at the Carribean or Santa Fe or even New York, or wherever, and he would call me up on those days where I am not working with him or in Palm Beach with him, and he would ask me to get on the next plane to so and so and go meet so and so, and that's when I would take e-tickets. His secretary or special assistant, whatever, would organize it for me and give me the details and I would just walk up the line and they'd let me right through.

JS: Can you give me any ideas as to how many times it happened that Jeffrey would send you off to meet some friend of his at some location outside of Palm Beach?

V: How many times it happened? I'm not too sure. Probably about 10-15 times.

JS: Ok. And on those occasions, how much time would you spend with one of Jeffrey's friends when you were sent to a location that you would have to travel to?

V: Only a couple of days. Only 2 days, that's it.

JS: And how were you paid for those trips?

V: I would be paid in cash upon my arrival back with Jeffrey. So, whenever I was back with Jeffrey, he would count up how many days I've had, sometimes give me even more than what I deserved, not deserved, but what I earned and give me a little extra.

JS: Was there a daily rate for those trips or was that per massage also?

V: Per massage. With Jeffrey, I would be honest. I wouldn't tell him I did 15 massages if I didn't. He knew he could trust me. He could always come back to the other person that he sent me to give massages and ask them as well, so you know, it was always by per massage.

JS: Alright. When we've been talking about massages, tell me exactly what it is we're talking about when we speak about massages.

V: Same thing I would do to Jeffrey. Again, it would start out as a massage, which would start with them being naked, and me giving him a legitimate massage to begin with, so

I'd start with his feet, go up to his calves, up his legs, buttocks, back, his neck, his head, his arms, yada yada, and then it would be time to flip over, and some of the men would want me to continue on massaging the front side of them and they would instigate me to begin having sex with them or foreplay, whatever you want to call it.

JS: So routinely, these massages involved sexual activity. Is that accurate?

V: That is accurate.

JS: Ok. Let's talk about the travel that you were involved in when you were on Jeffrey's private plane. Generally speaking, who were the passengers on the plane when you traveled.

V: Well, Larry was the pilot, and then there was a short, small solid guy, I don't know his name, but he was a co-pilot, and then he changed and there was another guy brought in later on. Generally speaking, there was always Jeffrey, sometimes Ghislaine, sometimes Emmy, sometimes a whole bunch of other girls, sometimes famous people, sometimes some politicians or yeah, just about anybody could fly on his plane. There was never no any set routine who would come and who would go. It was an influx of people on Jeffrey's airplane.

JS: I want to deal with these things separately in order to respect some of the reservations that you have, so I'm going to ask you who the people were that you remember flying with Jeffrey on his plane when you were personally present without regard to whether there was any sexual activity that occurred on the plane or not. So I'm not asking you to implicate any of these famous people in improper conduct, but just tell me what the names of the people are that you remember that you consider to be famous people.

V: Ok, there was Naomi Campbell, Heidi Klum, there was Bill Clinton. There was Al (?) Gore, there was a whole bunch of models, I wouldn't really honestly be able to give their names. There was Matt Groning the producer of the simpons cartoon, Jack CCousteau's granddaughter a lot of interior designers, architects, politicians. I am just trying to think of as many names as possible for you. Off the top of my head, that's as good as I can get for now.

JS: Ok, alright, that's fine. And again, I am not implying by my questions, nor do I want your answers to be interpreted as your suggesting that any of those people that you have just identified were engaged in any improper activities on any particular flight, but I want to talk to you now about what went on on occasion on the airplane. Ok?

V: Ok. It was a lot of the same thing that went down on the ground. A lot of times, it would be just be me and Jeffrey, or me and Jeffrey and Ghislaine, or me and Jeffrey and some other girl, sometimes Emmy, Sarah, and Nadia Bjournik. There would be sexual conduct, there would be foreplay, there was a bed in there, so we could basically reenact exactly what was happening in the house. It would start off with massaging or we would start off with foreplay, sometimes it would lead to, you know, orgies.

JS: Were there occasions when you were in Jeffrey's company, whether on the ground or in the air, where there were other girls present whom you knew to be under age 18?

V: Yes. There was a constant influx of girls coming in and going out. And we were all very young. On occasion, there was some older girls, and I don't mean older as in like in their 30s or anything, I mean like 28, 29, something like that, just very rarely. The majority of the girls that Jeffrey actually met or had on his plane or in his house were under age.

JS: Do you know how it is that Jeffrey established with any of these underage girls?

V: Yes, I do. He would send me personally or with other girls to clubs or shops, to _pick up anywhere, I mean we were constantly on the look for other girls that might satisfy Jeffrey.

JS: What instructions were you given about what to look for?

V: Young, pretty, you know, a fun personality. They couldn't be black. If they were any other descent other than white, they had to be exotically beautiful. That was just about it.

JS: Who gave you those criteria?

V: They both gave us the instructions, and it wasn't just me, Jeffrey asked most girls to bring a friend and make extra money. They would use us young girls So that way it probably looked a lot more safer to a girl that we were procuring to younger girls that were already doing it. That was the way that Jeffrey had it.

JS: Were you given any instruction at all on how to approach these girls?

V: Yes. Jeffrey and Ghislaine both taught me to, depending on the circumstances, depending on the girl, you could offer them a job as a massage therapist or you could tell them you have a really rich friend with, you know, great contacts in the acting world or modeling world and he loves pretty girls, you should come back and meet him, make some money, you know, we had a whole bunch of ways to be able to procure girls.

JS: Can you give me any idea as to the total number of underage girls that you know engaged in sexual conduct with Jeffrey during the period of time you had your relationship with him?

V: I would have no way of estimating that whatsoever. I mean, there could be a hundred, there could be more, honestly I'm not too sure how many girls, really. I wish I did know. I mean like I said there were so many over the course of 4 years with Jeffrey.

JS: Let me see if we can try to narrow it down a little bit. Is there any doubt in your mind that it was more than 10?

V: Yes, there was definitely more than 10.

JS: Ok, what I want you to do is to give me the highest number that you are comfortable in saying there were definitely more than X number of underage girls that I know Jeffrey Epstein engaged in sex with while I had a relationship with him. How would you fill in that blank? Definitely more than how many?

V: I'd say definitely more than a hundred.

JS: Alright. Did Jeffrey ever help to pick out your clothes?

V: Oh yes. I mean he wasn't out to dress me like a porn star or anything. He would always dress me very classy, but we'd just go shopping all the time together.

JS: Did he ever express any style preferences in terms of how he wanted you to dress? Besides dressing classy, I'm, you know, any other suggestion to you about how he wanted you dressed?

V: He didn't, like I said, wasn't trying to dress me in any prostitute way or anything like that. It was nice, classy outfits I was wearing like Gucci, Dolce Gabbana, Chanel, things like that. He was buying me a lot of very, very nice clothing. It was provocative. I mean I was wearing miniskirts, and tight short shorts and little shirts that showed my belly and my cleavage and everything, but they were very expensive clothes.

JS: Was there every any dress up role playing?

V: Yes. There was. Lots of it. Jeffrey loved the latex outfits Ghislane had for us girls, he had bondage outfits, he had all different kinds of outfits, but his favorite was the schoolgirl.

JS: Tell me about that.

V: Well, you know, Ghislaine would take me to dress me up to surprise J.E or Jeffrey would ask me to get dressed up, that would include wearing a tiny little skirt with nothing underneath, a white collared shirt that you would be wearing to school with a tie in it, tied up _in a bow_, my hair in pigtails, stockings on up to my knees, and I would go in there and act like a kid and we'd do role playing sexing.

JS: Did Jeffrey ever brag to you about the age of any of the girls with whom he had relationships?

V: Yes, he did. He did all the time. The worst one that I heard from his own mouth was this pretty 12 year old girls he had flown in for his birthday. It was a surprise birthday gift from one of his friends and they were from France. I did see them, I did meet them. Jeffrey bragged afterwards after he met them that they were 12 year olds and flown over from France because they're really poor over there, and their parents needed the money or whatever the case is and they were absolutely free to stay and flew out. Those were the worst ones. He was constantly bragging about girls' ages or where he got them from or their past and how terrible their past was and good he is making it for them.

JS: Where were the 12 year old girls flown to from France? Where did they come to?

V: Palm Beach.

JS: And were they flown in on Jeff's private plane or did they get transported?

V: No. They were transported by somebody else.

JS: Ok. Was the sexual activity that went on on the airplane conducted in such a way so that any of the crew was aware of what was going on?

V: They were told to knock if they had to come out, if the crew had to come out. They were told, you know, to come out as little as possible, so they weren't out there hanging out watching everything, no, but it doesn't take an idiot to put two and two together to say well there's a whole bunch of half dressed teenagers on board with this old man who is constantly being massaged by them and he wants me to keep the door shut for what reason? I mean, only they could put that together, but yeah, they knew.

JS: Did Mr. Epstein ever talk to you about people of power and influence owing him favors?

V: He would laugh about it, you know, I never really knew what to take serious from Jeffrey because he was such a funny character at times. You never knew if what he was saying was true or not. Yeah, lots of people owed him favors from what he told me. He's got everybody in his pocket, and he would laugh about he helps people for the sole purpose in the end they owe him something. That's why I believe he does so many favors in the first place.

JS: When and how did you first become aware that Mr. Epstein was in trouble with the law?

V: I was first informed by, I think someone from the FBI called me first and started to ask me questions, and I started to answer the questions but then fear took over, and I just said look, I don't know what's going on, I've got a young family that I don't want to risk, you know, please don't bother me about this again, and it was real short simple conversation, and within a week or 2, I had gotten a call from Jeffrey's attorney, and then a week later, Jeffrey himself.

JS: Ok, well let's back up before we get to those conversations and tell me approximately when it was that you were contacted by the person who you believe was with the FBI.

V: Ok. It's hard for me to pinpoint, if I had to pinpoint it, it would be in 2007 sometime.

JS: Alright. And you were living in Australia at that time, correct?

V: Correct.

JS: You were contacted by telephone?

V: That's correct, by my cell phone.

JS: Ok and do you have any idea how your name came up leading to that contact.

V: No idea. No idea whatsoever. When I did ask, I was told that some girls had revealed my name, I guess, and that's how everybody, the FBI knew to contact me.

JS: OK.

V: But I don't know offhand or sorry, I just walked into the wrong room.

JS: Ok.

V: Sorry go on.

JS: Yes & I'll never tell her you said that. Virginia, how long was it after that phone call from the FBI person were you contacted by Mr. Epstein's lawyers.

V: Like a week. It was back to back to each other. I remember being so scared after talking to the FBI thinking what's happening, what's going on. It's been like 6 years, 7 years at that stage, how did they find me & what do I have to do with this? So yeah, I do remember that very well, and it was only about a week later I was called by his attorney.

JS: Who was it that contacted you, do you remember?

V: I want to say Bill Riley, but he might have been from the FBI. No, it was Bill Riley. Bill Riley. Not sure if that's his correct name, but that's what is coming to mind

JS: What do you remember about that conversation?

V: I remember a Mr. Goldberger as well, I remember, there might have been two of them.

JS: Alright.

V: I can't remember which one it was. I want to say Bill Riley is the good one.

JS: Alright, so either Bill Riley or a Mr. Goldberger or both of them contacted you, and what do you remember about that?

V: I don't know if it was the same guy who contacted me that week later who put me in touch with Jeffrey. I think he was on the phone and he put speakerphone on with Jeffrey. So he connected me with Jeffrey. I don't know if it was the same guy or different, but I definitely know that Bill Riley was the first guy to contact me. I'm pretty sure about that.

JS: Ok. Tell me about that conversation.

V: He asked me what I knew about what's going on with Jeffrey and apparently, there was an investigation being held about some of the girls who had come out and said that Jeffrey had sexual contact with them under the age of a minor and that he was discrediting lot of these girls and making them out to be drug addicts and prostitutes and what have you so they wouldn't be looked upon as worthy in the court's eyes so to speak. And you know, he told me in the first five minutes that, you know, if I stay quiet, that "I'll be looked after" . And that was the exact way it was said. It wasn't like you know, I'm gonna pay you a zillion dollars or anything if you be quiet, but if I stay quiet, I would "looked after". And I remember saying I don't want any part to do with this. You know, this is not something I want to be a part of, I've got a young family. I wish the best for everybody in this, you know, take care kind of thing. A week later, I was called after the hearing by one of Jeffrey's lawyers. I can't tell you exactly which one it was but he had Jeffrey on the other line and he connected Jeffrey and I, and Jeffrey tried to make some simple conversation, "How are you? How have things been?" You know what I mean, catching up.

JS: Do you know if the lawyer, did the lawyer stay on the line while Jeffrey was speaking to you?

V: I'm pretty sure he did. That's why I think Jeffrey was on speaker phone because it sounded a lot different, and I was never taken off the line to begin with or connected to another line, so I was pretty sure Jeffrey was on speaker phone and the lawyer was making the call. After the simple conversation, it led to what was going on again and you know, Jeffrey couldn't believe it. You know, he thought he helped all these girls out. He didn't think he was wrong in any circumstance here at all. A lot of these girls were drug addicts and just after drug money. You know, he was really putting down these women or these girls I should say, not giving them the credit they deserved, and then he exactly repeated what the lawyer said the week before is that he would look after me if I stayed quiet, and if I need any help, you know, his lawyers would represent me and he would get legal help for me, whatever I need, he would do, and I told him exactly, I said, "Jeffrey, I'm the mother of two children at that stage. I'm away from everything there, I don't want to be a part of it. I'm not going to speak to anybody and I don't want to speak to anybody, I don't want to be involved." That was the last time I heard from him. And the next thing I knew, I was sent my victim's letter, my notification of being a victim through the US Attorney's Office and that's when I knew it was well out there enough not to have Jeffrey's lawyers come back on me and discredit me in the same way he had done to all the other girls. So, I called up Joseph Bird who was the recommended lawyers on my paperwork that they had given me and started going from there.

JS: So you contact Mr. Joseph Bergs' office and then you were dealing with his office from that point forward.

V: That's correct.

JS: Tell me about the ending of your relationship with Jeffrey. That is, at what point in time did your full time employment end and how did that happen?

V: Ok. So, it hadn't really ended. I walked away from it all. Jeffrey sent me to Thailand where I met my husband and escaped to Australia, never to return back to the states. About 6 months prior to that, he came up with a proposition that I thought was really disgustingly sick. And it really showed me for the first time in 4 years I had been with him that nothing was going to change and I was always just going to be used by him(?) which I did not like. He offered me a mansion and some of his money every month, I forget what he called it, a monthly income of what he made to bear one of his children. The proposition was that if anything ever happened between Jeffrey and I, that I would have to sign my child over to him basically and that the child would be his and Ghislaine's, and I would be looking after it as long as nothing happened between Jeffrey and I. So, I was kind of freaked out by all of that. I pushed Jeffrey more to please get me some more training, you know, and I was getting older and not of as much interest to Jeffrey anyways. I was 19 now, and he likes a female a lot younger. So he sent me to Thailand, in September 2002_. I was first supposed to meet a girl there and bring her back with me, but I never met up with her. I proceeded get a short course in Thai massage so that was to shut me up about my training so I went there, and one of my friends from school invited me to watch a fight, like a muay thai fight, which is like a form kickboxing. So I went and watched it, and I saw this guy that was a really good fighter, and a girl's word, looked really hot, so I asked my friend who knew him to introduce me. We got introduced and fell in love immediately, 3 days later Rob proposed and 7 days later I was being married in a buddist temple. I called Jeffrey and told him I'm sorry, I'm never coming back. I've gotten married, I've fallen in love. I thought he'd wish the best for me but he was kind of rude and he just said "have a good life" and hung up the phone, and that was the last time I'd talked to him ever until all this started again.

JS: Ok. Virginia, is there anything else that you would like to add to what you have told us up to this point in time?

V: I'd like to know that this time around something's going to be done about it and that Jeffrey and a lot of his colleagues, no matter how rich they are, will know that there is law and that there is people that still believe in it. So that's it. Thank you guys for listening to me, hearing me out and helping me. It's kind of hard to get through.

JS: Thank you very much. Yes, I'm sure it has been very difficult and I am very appreciative of the courage you have shown in doing what you have done, which really brings me to the last subject, and that is what was it that motivated you to go public with all of this?

V: Sharon Richard contacted me. I like her, I do, I like her a lot. I know she's a journalist, and journalists are normally bloodsucking leeches, but I like her for that, but she is an honest bloodsucking leech. She told me a lot about what was still going on, and she showed me a picture of Jeffrey with a little girl who looks like she could have been 12 years old. I mean it was disgusting. I agreed to talk with her, I never agreed to do anything until she showed me some pictures, and at that stage, being a mother of 3 children and having a daughter who I would do anything for to protect, I would put my neck on the line to make sure she never has to go through what I had to go through, and knowing all of this, and knowing that he's still out there doing the same exact thing with no regrets, no remorse, no worry about what he's doing to those girls, and all those girls feeling the same way that I did, so I, you know, I'm doing it because I believe in my heart of hearts it's the right thing to do. It's what I would want somebody to do for my daughter or my sister or my friend, and it saddens me to know that it's still going on right now. It's like the seashell story. I don't know if you're heard the story about the little kid who throws back a starfish, you know, the little brother tries to ask his sister, "why do you throw them in, they're all gonna die anyways, the little girl says "well, it's this one that I can help, and this one that I can help," and that's what I feel like I'm doing. I'm making a small dent in this big world we live in.

JS: I certainly appreciate that courage, and I have heard that story, and you're absolutely right, that one person may not be able to make a difference for everyone, but one person can make a difference for someone, and hopefully, you are making a difference for someone, and we're gonna do the best we can to make sure you are making a difference for as many people as possible.

V: Thanks Jack.

JS: Just a few other follow up things I want to ask you and again, if any of these questions are questions that you're uncomfortable in responding to, then don't hesitate to tell me that. Do you have any recollection of Jeffrey Epstein's specifically telling you that "Bill Clinton owes me favors"?

V: Yes. I do. It was a laugh though. He would laugh it off. You know, I remember asking Jeffrey what's Bill Clinton doing here kind of thing, and he laughed it off and said well he owes me a favor. He never told me what favors they were. I never knew. I didn't know if he was serious. It was just a joke.

JS: Where was here?

V: He told me a long time ago that everyone owes him favors. They're all in each other's pockets.

JS: When you say you asked him why is Bill Clinton here, where was here?

V: On the island.

JS: When you were present with Jeffrey Epstein and Bill Clinton on the island, who else was there?

V: Ghislaine, Emmy, and there was 2 young girls that I could identify. I never really knew them well anyways. It was just 2 girls from New York.

JS: And were all of you staying at Jeffrey's house on the island including Bill Clinton?

V: That's correct. He had about 4 or 5 different villas on his island separate from the main house, and we all stayed in the villas.

JS: Were sexual orgies a regular occurrence on the island at Jeffrey's house?

V: Yes.

JS: If we were to take sworn testimony from the people I am going to name, and if those people were to tell the truth about what they knew, do you believe that any of the following people would have relevant information about Jeffrey's taking advantage of underage girls? So I'll just name a name, and you tell me yes if they told the truth, I think they'd have relevant information or no, I don't think they would, or I don't know whether they would or not. Ok? You understand?

V: Yes.

JS: Ok. Les Wexner.

V: I think he has relevant information, but I don't think he'll tell you the truth.

JS: Ok. Alan Dershowitz.

V: Yes.

JS: David Copperfield.

V: Don't know.

JS: Tommy Matola.

V: Don't know.

JS: Prince Andrew.

V: Yes, he would know a lot of the truth. Again, I don't know how much he would be able to help you with, but seeing he's in a lot of trouble himself these days, I think he might, so I think he may be valuable. I'm not too sure of him.

JS: Ok. Virginia, I think that's all I have for you. Let me tell you what I would like to do. As I told you in the beginning of this conversation, we've been recording it, and hopefully, we've got a clear enough recording so that we've taken down everything accurately and when it's transcribed, it will be clear and accurate, but what I would like to do is transcribe it, send it to you, have you take a look at it, and if there's anything that we got wrong in the statement, you can write back and you can make changes in the transcript so that the transcript is accurate. Is that fair?

V: No worries. That is fair. No problem.

JS: Alright, great. I really do appreciate that and tell me what the best way is to send the transcript to you.

V: Email. If you just want to send it by email or if you want to send it by mail, either or.

JS: Ok. Give me your email address if you would please.

V: [redacted]*

JS: Let me read that back to you: [redacted]*

V: Yep that's it.

BE: Thank you Jenna, appreciate it.

V: No problem, Brad.

JS: Thank you very very much. Bye Bye now.

V: Take care Jack. Nice meeting you.

JS: You too.

*Redaction has been made at the request of the witness.

EXHIBIT K

From: Donna Paine <djpaine@aol.com>
Sent: Friday, January 23, 2015 2:31 PM
To: Paul Cassell
Cc: brad@pathtojustice.com; JWI@SearcyLaw.com; matt.williams@itn.co.uk; kate.brannelly@itn.co.uk
Subject: Re: Jeffrey Epstein case

Thank you, Paul. I had reached out to Joan previously to be included on Jack Scarola's e-mail distribution list for the Dershowitz suit, but I wasn't sure if this would also be handled by them.

I don't want to double up or inundate you, as I know other ITN News colleagues have also been in touch with you. I just wanted to make sure I was on your radar if Judge Marra rules next week.

Best,
Donna Paine

Sent from my iPad

On Jan 23, 2015, at 1:10 PM, Paul Cassell <cassellp@law.utah.edu> wrote:

> Hi Donna,
>
> I am passing this information along to Joan, who will add this information into our contacts list. Thanks for your interest in this important case.
>
> Paul
>
> Paul G. Cassell
> Ronald N. Boyce Presidential Professor of Criminal Law
> S.J. Quinney College of Law at the University of Utah
> 332 South 1400 East, Room 101 Salt Lake City, UT 84112-0730
> Voice: 801-585-5202 Fax: 801-581-6897 Email: cassellp@law.utah.edu
> You can access my publications on http://ssrn.com/author=30160
>
>
> **From:** djpaine@aol.com [mailto:djpaine@aol.com]
> **Sent:** Friday, January 23, 2015 9:47 AM
> **To:** brad@pathtojustice.com; Paul Cassell
> **Cc:** matt.williams@itn.co.uk; kate.brannelly@itn.co.uk
> **Subject:** Jeffrey Epstein case
>
> Hi Brad and Paul --
>
> I had reached out to you both the other week when the allegations of Jane Doe #3 first became public. I am a freelance TV producer in South Florida, working with ITN News on this story. As you can imagine, given her allegations against Prince Andrew, ITN is quite keen on covering this story as soon as it breaks.

EXHIBIT L

From: Paul Cassell
Sent: Saturday, January 24, 2015 1:25 PM
To: Michael Bilton
Subject: RE: from Michael Bilton - latest from Daily Mail online

Hi Michael,

We got the information. Thanks very much.

This is a very interesting article that puts a lot of the flight log information together.

http://gawker.com/flight-logs-put-clinton-dershowitz-on-pedophile-billio-1681039971

Let's stay in touch. Hopefully more and more of the truth will out. PC

Paul G. Cassell
Ronald N. Boyce Presidential Professor of Criminal Law
S.J. Quinney College of Law at the University of Utah
332 S. 1400 E. , Room 101
Salt Lake City, UT 84112-0730
(801) 585-5202 (phone)
(801) 581-6897 (fax)
cassellp@law.utah.edu

You can access my publications on http://ssrn.com/author=30160

From: Michael Bilton [michaelbilton1@btinternet.com]
Sent: Friday, January 23, 2015 6:04 PM
To: Paul Cassell
Subject: Re: from Michael Bilton - latest from Daily Mail online

You should have had from me a couple of zip files concerning VR's travel seen alongside Prince Andrew's engagements. Has that arrived with you?

M
On 24 Jan 2015, at 00:58, Paul Cassell wrote:

> Thanks for the note – looks like various pieces are falling together. PC
>
> Paul G. Cassell
> Ronald N. Boyce Presidential Professor of Criminal Law
> S.J. Quinney College of Law at the University of Utah

EXHIBIT M

Mary E. Pirrotta

From:	Mary E. Pirrotta
Sent:	January 22, 2015 9:38 AM
To:	'Lindsay.Isaac@turner.com'
Cc:	Joan Williams
Subject:	Re: Edwards, Bradley vs. Dershowitz Document: Complaint (File #: 20150013)
Attachments:	Complaint.pdf; The Duke of York.pdf

Attached is a copy of the Complaint in Case No. CACE 15-000072 and also a copy of Mr. Scarola's letter to The Duke of York per your request.

Is the full transcript of our telephone interview with Virginia part of the public record? I have a request from the press to provide it (see below), but I will only do that if it has been filed. TAKE THIS OUT

Begin forwarded message:

From: "Hill, James E." <James.E.Hill@abc.com>
Date: January 28, 2015 at 10:34:00 AM EST
To: Jack Scarola <JSX@SearcyLaw.com>
Subject: RE: Question from ABC

If possible - I'd also like to see full transcript of the 2011 phone interview you/Edwards conducted with VR while she was still in Australia.

-----Original Message-----
From: Hill, James E.
Sent: Wednesday, January 28, 2015 9:31 AM
To: 'Jack Scarola'
Subject: RE: Question from ABC

Thank you

-----Original Message-----
From: Jack Scarola [mailto:JSX@SearcyLaw.com]
Sent: Wednesday, January 28, 2015 7:19 AM
To: Hill, James E.
Subject: Re: Question from ABC

You will receive a response and copy of the deposition later this morning.

On Jan 27, 2015, at 10:00 PM, "Hill, James E." <James.E.Hill@abc.com> wrote:

Hi Jack - sorry to trouble you.

But wondering if you know what documents are source for a new NY Daily News story - quoting what appears to be a Juan Alessi deposition which mentions Andrew visits to PB mansion.

Daily News cites newly filed documents. I'm not aware of any new docs that quote Alessi on the issue of Andrew.

Can u enlighten me?

Do you have access to those portions of the Alessi depo? Not sure which case it would come from.

http://m.nydailynews.com/news/national/prince-andrew-sex-massages-epstein-home-repairman-article-1.2094073

James E. Hill
Producer, ABC News
Chicago
312.899.4281 (ofc) | 312.286.1691 (cell) | 312.343.3832 (cell 2)
james.e.hill@abc.com
Twitter: @jameshillabc

****
Privileged and Confidential Electronic communication is not a secure mode of communication and may be accessed by unauthorized persons. This communication originates from the law firm of Searcy Denney Scarola Barnhart & Shipley, P.A. and is protected under the Electronic Communication Privacy Act, 18 U.S.C. S2510-2521. The information contained in this E-mail message is privileged and confidential under Fla. R. Jud. Admin. 2.420 and information intended only for the use of the individual(s) named above. If the reader of this message is not the intended recipient, you are hereby notified that any dissemination, distribution, or copy of this communication is strictly prohibited. Personal messages express views solely of the sender and shall not be attributed to the law firm. If you received this communication in error, please notify the sender immediately by e-mail or by telephone at (800) 780-8607 and destroy all copies of the original message. Thank you.
****

EXHIBIT N

Paul Cassell

From: Paul Cassell
Sent: Saturday, January 3, 2015 11:14 AM
To: Paul Blake
Subject: RE: questions for Dershowitz

Questions for Alan Dershowitz

1. Have you ever met a young girl, under the age of 18, in the presence of Jeffrey Epstein or on one of Jeffrey Epstein's properties?

2. How many times have you visited Jeffrey Epstein's mansion in West Palm Beach, Florida?

3. The U.S. Attorney's Office for the Southern District of Florida reached a non-prosecution agreement with Jeffrey Epstein, in which he agreed to make civil restitution for sexual abuse committed against 40 girls at the mansion. Do you have any reason to doubt that that Jeffrey Epstein sexually abused 40 underaged girls.

4. Didn't Jeffrey Epstein tell you that he repeatedly had sex with these underage girls?

4. Jeffrey Epstein is a now a registered sex offender. Do you have any reason to doubt that this registration is appropriate?

5. Have you ever met Ghislaine Maxwell in the presence of Jeffrey Epstein or at one of his properties?

6. Did you send correspondence to the U.S. Department of Justice in Washington, D.C. (or elsewhere) attacking the credibility of these 40 girls?

7. According to testimony in court cases, when Jeffrey Epstein was at his mansion in West Palm Beach, Florida, he would receive almost daily "massages" that were in fact sexual encounters with these underage girls. Do you have any reason to doubt this testimony?

8. The U.S. Attorney's Office gave Epstein a non-prosecution agreement that extended immunity from federal prosecution not only to Epstein but also "any potential co-conspirator". Isn't it normal practice for federal prosecutors only to extend immunity to specifically named persons, not an open-ended group of "any" co-conspirator.

9. How did you travel to Epstein's mansion in West Palm Beach?

10. How many times have you been to Epstein's private ranch, called "Zorro", in New Mexico?

11. How many times have you been to Epstein's private island in the U.S. Virgin Islands, known as Little St. John's.

12. Did you ever see young girls present at Little St. John's?

13. If we talk about the time frame of 1998 to 2001, how many times a year were you making telephone calls to Epstein?

14. How many times have you flown on Epstein's private plane?

15. Did you ever see young girls, potentially under the age of 18, on Epstein's private plane?

IN THE CIRCUIT COURT OF THE
SEVENTEENTH JUDICIAL CIRCUIT
IN AND FOR BROWARD COUNTY, FLORIDA

CASE NO.: CACE 15-000072

BRADLEY J. EDWARDS and
PAUL G. CASSELL,

Plaintiffs/Counterclaim Defendants,

vs.

ALAN M. DERSHOWITZ,

Defendant/Counterclaim Plaintiff.

____/

DEFENDANT/COUNTERCLAIM PLAINTIFF ALAN M. DERSHOWITZ'S *REDACTED* MOTION TO MODIFY CONFIDENTIALITY ORDER

Defendant/Counterclaim Plaintiff, Alan M. Dershowitz ("Dershowitz"), by and through undersigned counsel, hereby files his *Redacted* Motion to Modify Confidentiality Order of January 12, 2016, and in support thereof states the following:

On January 16, 2016, Defendant Alan M. Dershowitz began the deposition of non-party Virginia Roberts Giuffre ("Roberts"). Pursuant to this Court's January 12, 2016 Confidentiality Order, that transcript currently is under seal. The Confidentiality Order should be modified at least to allow Dershowitz to defend this case. Dershowitz and his counsel need to be able to contact witnesses, inform them of Roberts's testimony, and ask them whether Ms. Roberts's testimony is accurate. They also need to be able to use Ms. Roberts's testimony in other ways as part of the defense effort, such as by providing it to expert witnesses, among other things. The bottom line is that Dershowitz's counsel must be able to use Roberts's testimony as necessary in their professional judgment to represent their client, as a matter of fairness and due process.

Accordingly, Dershowitz requests that the Court modify the Confidentiality Order to confirm that Dershowitz's counsel may disclose Ms. Roberts's testimony as they deem necessary in their professional judgment in order to represent Dershowitz in this case.

BACKGROUND & EXECUTIVE SUMMARY

Dershowitz was first presented with Roberts's heinous and false allegations against him when her lawyers, Bradley J. Edwards ("Edwards") and Paul G. Cassell ("Cassell"), filed certain now-stricken allegations in the action styled *Jane Doe, et al. v. United States of America*, No. 08-80736 (S.D. Fla.) (the "Federal Action"). After Dershowitz defended himself to the media, Edwards and Cassell sued Dershowitz for defamation. The falsity of Roberts's allegations, her credibility, and the investigation her lawyers took to assess those allegations and credibility before filing those allegations are a critical part of Dershowitz's defense.

On April 9, 2015, Roberts moved for an order "quashing the subpoena *duces tecum* served on her by Defendant, or alternatively, pursuant to Florida Rules of Civil Procedure 1.280(c) for issuance of a protective order sharply limiting the scope of the subpoena" (the "Motion to Quash"). *See* Motion to Quash, attached hereto as Exhibit A. Roberts did not move to seal the deposition transcript and the resulting order did not seal it, but instead directed that "a confidentiality order shall be entered." *See* November 4, 2015 Email from Judicial Assistant Susan Moss, attached hereto as Exhibit B and November 12, 2015 Order, attached hereto as Exhibit C. The Confidentiality Order then prepared by Roberts's counsel and consented to by all parties includes a provision stating that "[t]he deposition testimony of Non-Party Virginia Giuffre will be designated as 'Confidential' and not subject to public disclosure" and that "[i]t may only be filed under seal." *See* January 12, 2016 Confidentiality Order, attached hereto as Exhibit D.

Dershowitz now requests that the Court modify the Confidentiality Order to allow Dershowitz to use the transcript for those limited purposes as deemed necessary in the professional judgment of his counsel to ensure Dershowitz is afforded his right to build and present his defense.

I. DERSHOWITZ MUST BE ALLOWED TO CONTACT WITNESSES AND ADVISE THEM OF WHAT ROBERTS ALLEGES IN ORDER TO VERIFY OR DISPROVE HER ALLEGATIONS AND CREDIBILITY AND DETERMINE WHETHER PLAINTIFFS EVER MADE EFFORT TO CONTACT THESE INDIVIDUALS TO VERIFY ROBERTS'S ALLEGATIONS AND CREDIBILITY.

As explained by Plaintiffs Edwards and Cassell in their Response to Dershowitz's Motion to Determine Confidentiality, the "sexual abuse allegations filed by Edwards and Cassell for their client Ms. Virginia Giuffre are not peripheral to this lawsuit – *they are inherent to it.*" Plaintiffs' Response to Dershowitz's Motion to Determine Confidentiality, November 23, 2015, attached hereto as Exhibit E at 4 (emphasis added). Those "sexual abuse allegations filed by Edwards and Cassell for their client" go beyond Dershowitz. Another *inherent* part of this lawsuit is what investigation, if any, Plaintiffs undertook with respect to the scope of Roberts's allegations, all of which bear upon her credibility. Dershowitz argues that Plaintiffs did not perform a reasonable investigation before making the allegations in the Federal Action. Plaintiffs argue that they did. Dershowitz must be allowed to contact witnesses and advise them of what Roberts alleges so that Dershowitz can not only verify or disprove her allegations and credibility, but also determine whether Plaintiffs ever made efforts to contact key witnesses to verify Roberts's allegations and credibility. As explained by one Florida court, "[o]penness in courts has a salutary effect on the propensity of witnesses to tell the truth" as it "informs persons affected by litigation of its effect upon them . . . ." *John Doe-1 Through John Doe-4 v. Museum of Sci. & History of Jacksonville, Inc.*, No. 92-32567-CI-CI, 1994 WL 741009, at *1 (Fla. Cir. Ct. June 8, 1994) (internal citations omitted).

As set forth in Dershowitz's Motion for Clarification of Confidentiality Order or Relief from that Order, filed Jan. 29, 2016, it appears that Roberts made false statements in a publicly filed affidavit about being present at a private island in the US Virgin Islands when former President Clinton was there. Indeed, former FBI Director Louis Freeh determined based on the response of the federal government to a FOIA request that the absence of records responsive to the request "strongly establishes that former President Clinton was not present on Little St. James Island during the period at issue." *Id.* at 2. If Roberts made a false statement under oath about former President Clinton, it is equally if not more likely that she has made false statements about others whose whereabouts are more difficult to track.

Roberts cannot reasonably argue her testimony is confidential as [redacted] *See* Excerpts of Roberts's Deposition Transcript, attached hereto (under seal) as Exhibit F.[1] [redacted] *See id.*; *see also* March 2, 2011, Sharon Churcher, Daily Mail, attached hereto as Exhibit G. Dershowitz has also discovered that Roberts [redacted] *See* Exhibit F; *see also* AD-006931-006933, Transcript of Telephone Conversation Between Alan M. Dershowitz and Rebecca, attached hereto as Exhibit H. As a result, Roberts cannot claim that these allegations are confidential simply because [redacted]

[1] As per the Confidentiality Order, Exhibit F is only filed under seal.

II. DERSHOWITZ MUST BE ALLOWED TO SHARE ROBERTS'S DEPOSITION TRANSCRIPT WITH THOSE WORKING ON DERSHOWITZ'S BEHALF AS PART OF THIS LITIGATION.

Dershowitz asks the Court to modify the Confidentiality Order to allow Dershowitz to use the transcript in ways necessary for his defense including sharing the transcript with any counsel and other legal support, experts, consultants, insurers, and others typically permitted access to supposedly confidential information in addition to using it with potential witnesses and others as deemed necessary in the professional judgment of his counsel as set forth above. Dershowitz and his attorneys are aware of and will abide by the Florida Rules of Professional Conduct, including its comments, regarding the handling of any information deemed by this Court to be confidential within the limitations of the applicable rules.

III. ALLOWING DERSHOWITZ TO USE THE DEPOSITION FOR THE LIMITED PURPOSES OF HIS DEFENSE IS REQUIRED BY FLORIDA LAW.

Florida law requires that any sealing order be the least restrictive means necessary to accomplish its purpose. The Florida Supreme Court held in *Barron v. Florida Freedom Newspapers*, 531 So. 2d 113 (Fla. 1988), that a sealing order can be entered only where "no reasonable alternative is available to accomplish the desired result, and, if none exists, **the trial court must use the least restrictive closure necessary to accomplish its purpose.**" *Id.* at 118 (emphasis added); *see also Carter v. Conde Nast Publ'ns*, 983 So. 2d 23, 26 (Fla. 5th DCA 2008) ("an order sealing court records must state, inter alia, the particular grounds for making the court records confidential, that the closure is no broader than necessary, and that there are no less restrictive measures available.").

This Court has not set forth any reasons addressing a request by Roberts to seal her deposition transcript, much less determined that "no reasonable alternative is available" to accomplish Roberts's desired result. *See News-Press Publ'g Co. v. State*, 345 So. 2d 865, 867

(Fla. 2d DCA 1977) ("The judge's statement that he had 'cogent reasons' for sealing the records obviously fell short of specifically setting forth the reasons why public access to these deposition was being denied."). Moreover, Roberts and Plaintiffs cannot argue that sealing her deposition in its entirety is "the least restrictive" option, as it is most certainly the *most* restrictive option and one that Florida courts take very seriously. "[A] closure order must be drawn with particularity and narrowly applied." *Barron*, 531 So. 2d at 117.

Here, that requisite "least restrictive" application requires, at a minimum, allowing Dershowitz to use Roberts's testimony for the limited purposes necessary in the professional judgment of his counsel to represent their client, as a matter of fairness and due process.

CONCLUSION

Because Dershowitz must be able to prepare his defense and any sealing order must be the least restrictive measure available, the Court should modify the Confidentiality Order to confirm that Dershowitz's counsel may disclose Roberts's testimony as they deem necessary in their professional judgment in order to represent Dershowitz in this case.

Respectfully submitted,

Dated: February 3, 2016

<u>*s/Thomas E. Scott*</u>

Thomas E. Scott, FBN 149100
Thomas.scott@csklegal.com
Steven R. Safra, FBN 057028
Steven.safra@csklegal.com
COLE, SCOTT & KISSANE, P.A.
Dadeland Centre II, 14th Floor
9150 South Dadeland Boulevard
Miami, Florida 33156
Phone: (305) 350-5300
Fax: (305) 373-2294

Richard A. Simpson (*pro hac vice*)
rsimpson@wileyrein.com
Mary E. Borja (*pro hac vice*)
mborja@wileyrein.com
Ashley E. Eiler (*pro hac vice*)
aeiler@wileyrein.com
Nicole Audet Richardson (*pro hac vice*)
nrichardson@wileyrein.com
WILEY REIN LLP
1776 K Street NW
Washington, DC 20006
Phone: (202) 719-7000
Fax: (202) 719-7049

Counsel for Alan M. Dershowitz

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that a copy of the foregoing has been electronically filed through the Clerk of Broward County by using the Florida Courts eFiling Portal and thus served by electronic mail: jsx@searcylaw.com, mep@searcylaw.com, scarolateam@searcylaw.com to: **Jack Scarola, Esq,** Searcy Denney Scarola Barnhart & Shipley, P.A., Counsel for Plaintiff, 2139 Palm Beach Lakes Blvd., West Palm Beach, Florida 33409; jonijones@utah.gov to: **Joni J. Jones, Esq.**, Assistant Utah Attorney General, Counsel for Plaintiff Cassell, 160 East 300 South, Salt Lake City, Utah 84114; brad@pathtojustice.com to: **Bradley J. Edwards, Esq,** Farmer, Jaffe et al, 425 North Andrews Avenue, Suite 2, Ft. Lauderdale, FL 33301; cassellp@law.utah.edu, to: **Paul G. Cassell, Esq.,**; smccawley@bsfllp.com, sperkins@bsfllp.com, to: **Sigrid S. McCawley, Esq.**, Boies Schiller & Flexner, LLP, 401 E. Las Olas Blvd, Suite 1200, Ft. Lauderdale, FL 33301, this 3rd day of February, 2016.

By: *s/Thomas E. Scott*
Thomas E. Scott
FBN: 149100
Steven R. Safra
FBN: 057028

EXHIBIT A

Filing # 25919550 E-Filed 04/09/2015 05:23:25 PM

IN THE CIRCUIT COURT OF THE 17TH
JUDICIAL CIRCUIT IN AND FOR
BROWARD COUNTY, FLORIDA

CIVIL DIVISION

CASE NO. CACE 15-000072

BRADLEY J. EDWARDS, and
PAUL G. CASSELL,

Plaintiffs,

v.

ALAN DERSHOWITZ,

Defendant.
_______________________________/

MOTION TO QUASH OR FOR PROTECTIVE ORDER REGARDING SUBPOENA SERVED ON NON-PARTY JANE DOE NO. 3

Non-party Jane Doe 3, by and through undersigned counsel and pursuant to Florida Rules of Civil Procedure 1.410(c)(1)[1], hereby moves for an order quashing the subpoena *duces tecum* served on her by Defendant, or alternatively, pursuant to Florida Rules of Civil Procedure 1.280(c) for issuance of a protective order sharply limiting the scope of the subpoena.

INTRODUCTION

This Court should quash the subpoena issued to non-party Jane Doe No. 3 as it is unreasonable and oppressive. The Defendant is abusing the subpoena power in an effort to intimidate, harass and cause undue burden to a non-party. Indeed, Defendant - just days ago - publicly admitted that his goal of deposing Jane Doe No. 3 has nothing to do with this Florida Defamation Action; rather, he is trying to find a way to send this victim of sexual trafficking to "jail." "She was hiding in Colorado…but we found her and *she will have to be deposed. The end*

[1] For the limited purpose of the Motion to Quash or for Protective Order and resolving the scope of the subpoena and any enforcement issues, Jane Doe No. 3 voluntarily submits herself to this Court's jurisdiction.

result is that she'll go to jail because she will repeat her lies and we'll be able to prove it and she will end up in prison for perjury." (emphasis added). *See* Exhibit 1, New York Daily News, April 7, 2015. Defendant has subjected Jane Doe No. 3 to horrific public attacks including publicly calling her a "prostitute" and a "bad mother" to her three minor children. *See* Exhibit 2, Local 10 News, January 22, 2015.

Defendant has gone on a media blitz campaign against this non-party for statements she made under oath in a federal action: "The end result of this case should be *she [Jane Doe No. 3] should go to jail*, the lawyers should be disbarred and everybody should understand that I am completely and totally innocent." (emphasis added). *See* Exhibit 3, CNN International, New Day, January 6, 2015. "*My goal is to bring charges against the client and require her to speak in court.*" (emphasis added). *See* Exhibit 4, Australian Broadcasting System (ABC), January 6, 2015. Defendant also stated, in an interview in Newsmax, that he is "considering" bringing a lawsuit against Jane Doe No. 3. "*And we're considering suing her for defamation* as well, but right now she was trying to hide in Colorado and avoid service, but we found her and we served her and *now she'll be subjected to a deposition.*" (emphasis added). *See* Exhibit 5, Newsmax, April 8, 2015.

Defendant's own words demonstrate that he is abusing the subpoena power of this Court to try to get discovery that is irrelevant to this case, in the hopes of being able to intimidate Jane Doe No. 3 with the press and generate a claim against her. Considering the extensive abuse that Jane Doe No. 3 suffered as a minor child, and Defendant's threats and intimidation, it would be both unreasonable and oppressive to require this non-party to comply with this subpoena *duces tecum*. Accordingly, Defendant's subpoena should be quashed. *See* Exhibit 6, Defendant's Subpoena to Jane Doe No. 3.

BACKGROUND

The underlying action before this Court is a defamation case filed by a former federal judge, Paul Cassell, and his colleague Brad Edwards, who represent various sexual trafficking victims in a case pending in the Southern District of Florida, specifically case no. 08-cv-80736-KAM, hereinafter ("CVRA case"). As a result of an affidavit filed in the CVRA case, Defendant went on a national media defamation campaign calling, among other things, former federal judge Paul Cassell and attorney Brad Edwards, "unethical lawyers" who should be "disbarred". *See* Exhibit 7, Today Show, January 5, 2015. In response to this national slander campaign by the Defendant, Paul Cassell and Brad Edwards filed a defamation case against Defendant in the Circuit Court of the Seventeenth Judicial Circuit for Broward County, Case No. CACE 15-000072, hereinafter "Florida Defamation Action").

Defendant's statements against Paul Cassell and Brad Edwards are statements about their character as lawyers and do not directly involve non-party Jane Doe No. 3. Despite this fact, Defendant is abusing the subpoena power in this case by seeking documents from a non-party that are irrelevant to the defamation issue before this Court. Defendant is determined to find a way to harm non-party Jane Doe No. 3 and anyone who braves to represent her. Jane Doe No. 3 has good cause to be fearful of the Defendant in this matter based on Defendant's repetitive threats. *See* Exhibit 8, Affidavit of Jane Doe No. 3. This Court should not allow Defendant to abuse the subpoena power to further abuse this non-party. Florida Rules of Civil Procedure provide a vehicle for this Court to protect a non-party from a harassing, burdensome and unnecessary subpoena. As explained below, non-party Jane Doe No. 3 should be protected from having to be deposed in this matter or produce documents. Defendant's campaign of threats and intimidation should not be condoned by this Court and Defendant's subpoena should be quashed in its entirety.

ARGUMENT

1. This Court Should Quash Defendant's Abusive Subpoena In Its Entirety.

Florida Rule of Civil Procedure 1.410(c)(1) provides that the Court may "quash or modify the subpoena if it is unreasonable and oppressive." *Id.* The Court has discretion to evaluate the circumstances in determining whether the subpoena is "unreasonable and oppressive." *Matthews v. Kant*, 427 So. 2d 369, 370 (Fla. 2d DCA 1983). "The sufficiency thereof is a factual determination for the trial judge who is vested with broad judicial discretion in the matter, and whose order will not be overturned absent a clear showing of abuse of discretion." *Id.*; *see also Sunrise Shopping Center, Inc. v. Allied Stores Corp.*, 270 So. 2d 32 (Fla. 4th DCA 1972) (Fourth DCA quashing lengthy subpoena served on non-party who was not in control of documents as being "oppressive and unreasonable."). It is undisputed that Jane Doe No. 3 was sexually trafficked as a minor child by Jeffrey Epstein and he was sentenced for his crimes. Allowing the Defendant in this case to force this non-party to provide discovery on this highly sensitive topic would be both oppressive and unreasonable and serves no purpose other than to foster Defendant's publicly admitted and utterly baseless campaign to try to send Jane Doe No. 3 to "jail."

The documents requested in Defendant's subpoena demonstrate the oppressive and unreasonable nature of the requests. Defendant, for example, seeks highly personal and sensitive information from this victim of sexual trafficking, including requesting her personal diary during the time when she was being sexually abused as a minor child. *See* Exhibit 6, Request no. 16. Defendant also demands that this non-party produce photographs and videos of her as a minor child while she was being sexually trafficked by convicted sex offender Jeffrey Epstein. *See* Exhibit 6, Request nos. 2, 3, 4 and 10. Defendant's unreasonable subpoena even includes a demand for this non-party's personal cell phone records for more than a three (3) year period during the time when she was a minor child being sexually trafficked. *See* Exhibit 6, Request no.

15. Defendant also demands items like personal financial documents from this non-party including payments she received from convicted sex offender Jeffrey Epstein and the men he "lent" this minor child out to from 1999 – 2002. *See* Exhibit 6, Request no. 20. It is without question that Defendant is abusing the subpoena power in this case to conduct a fishing expedition in an effort to intimidate and harass this victim and to try to dig up information he can use in his openly stated "goal" to send this non-party to "jail."

Jane Doe No. 3 is rightfully fearful of Defendant as he is an incredibly powerful individual and the legal counselor to convicted Jeffrey Epstein who sexually trafficked Jane Doe No. 3 for years when she was a minor child. *See* Exhibit 8, Affidavit of Jane Doe No. 3. Jane Doe No. 3 believes Defendant's goal is to abuse the subpoena power to get her into a deposition so he can harass and intimidate her by forcing her to discuss the abuse she had to withstand as a minor child. *See* Exhibit 8, Affidavit of Jane Doe No. 3. None of that childhood abuse is relevant to this case which involves the narrow issue of whether Defendant defamed two lawyers. Defendant's subpoena is both unreasonable and oppressive and should be quashed. *See Matthews v. Kant*, 427 So. 2d 369, 370 (Fla. 2d DCA 1983).

2. <u>The Court Should Quash The Subpoena In Its Entirety, But At A Minimum, It Should Severely Limit The Production Requirements.</u>

In addition to its power to quash the subpoena, Florida Rule of Civil Procedure 1.280(c) also allows the Court to protect a non-party from discovery that would result in "annoyance, embarrassment, oppression or undue burden or expense…" *Allstate Ins. Co. v. Langston*, 655 So. 2d 91, 94 (Fla. 2003) (Florida Supreme Court overturning denial of protective order and holding that "[d]iscovery of certain kinds of information 'may reasonably cause material injury of an irreparable nature.'") (internal quotations omitted). *Matthews v. City of Maitland*, 923 So. 2d 591, 595 (Fla. 5th DCA 2006) (quashing discovery order where "[t]he compelled disclosure… would create a chilling effect on [petitioners] rights…"). The Court may determine that "the discovery not be had" or that "the discovery may be had only on specified terms and conditions…". Fla. R. Civ. P. 1.280(c).

Defendant issued a vastly overbroad subpoena to this non-party which included 25 separate document requests, many with subparts. In addition to placing an undue burden on this non-party to have to search for the broad scope of materials requested, the document requests seek information that is irrelevant to the Florida Defamation Action and clearly intended to "embarrass and oppress" this non-party. Fla. R. Civ. P. 1.280(c). Defendant's overly broad subpoena to non-party, Jane Doe No. 3, goes so far as to seek documents relating to former President, Bill Clinton and former Vice President, Al Gore, which, even if such documents existed, would be absolutely irrelevant to the Florida Defamation Action. *See Toledo v. Publix Super Markets, Inc.*, 30 So. 3d 712 (Fla. 4th DCA 2010).

Defendant's requests can be grouped into four key categories: (1) documents that contain highly personal and sensitive information sought only to harass, embarrass and intimidate the non-party; (2) documents unrelated to this action and, instead, intended to gain discovery relating to Defendant's admitted "goal" of putting this non-party in "jail," bringing a new case against Jane Doe No. 3, or related to the federal action; (3) documents that contain personal financial or other confidential information; and (4) privileged communications between the non-party and her lawyers. Non-party, Jane Doe No. 3, has filed specific objections as to each request sought in Defendant's subpoena as set forth in Exhibit 9. Here, Jane Doe No. 3 provides the Court with a sampling of the oppressive nature of the subpoena that is the subject of her detailed objections.

a. <u>Category 1 – Overly Broad Subpoena Requests Intended Solely to Harass, Embarrass and Intimidate the Non-Party by Seeking Highly Personal and Sensitive Information</u>

It is clear from the Defendant's requests that his intent is to intimidate and harass this non-party by seeking highly sensitive personal information that is irrelevant to this action. For example, Request no. 16 seeks "Any diary, journal or calendar concerning your activities between

January 1, 1999 and December 31, 2002." Defendant is seeking personal diary information during the time this non-party was a minor child and a victim of sexual trafficking. There is no reason this non-party should be forced to produce her diary from when she was a child. *See Peisach v. Antuna*, 539 So. 2d 544 (Fla. 3rd DCA 1989) (court of appeal holding that trial court departed from the essential requirements of law by granting deposition of party's gynecologist which was only meant to invade privacy and intimidate and harass the party).

Defendant also has a number of requests (Request nos. 2, 3, 4, 10 and 19) that seek "photographs" and "videos" of this non-party when she was a minor child and during the time she was the subject of sexual abuse. Photographs of Jane Doe 3 when she was a minor child are completely irrelevant to the matter before this Court. Defendant served this subpoena demand solely to intimidate, harass and embarrass this non-party and the Court should preclude this type of discovery set forth in Request Nos. 2, 3, 4, 10, 15, 16, 19 and 21. *See Citimortgage, Inc. v. Davis*, No. 50 2009 CA 030523, 2011 WL 3360318 (Fla. 15th Cir. Ct. April 4, 2011) (trial court granting protective order precluding a deposition noting "this deposition request is mere harassment" and had no relevance to the underlying dispute where the party was wrongfully using the discovery process for personal gain).

b. Category 2 – Clear Abuse of the Subpoena Power By Seeking Documents Unrelated to this Action and Intended Instead to Provide Discovery for Other Actions

Defendant is abusing the subpoena power of this Court by issuing subpoena requests that are intended to obtain discovery for the development of other actions against this non-party and are unrelated to the instant case. *See* Exhibit 5, Newsmax Interview ("And we're considering suing her for defamation as well, but right now she was trying to hide in Colorado and avoid service, but we found her and we served her and now she'll be subjected to a deposition."). Defendant has admitted that his "goal" is to put Jane Doe No. 3 in "jail" and he is using this Court's subpoena power to go on a fishing expedition in the hopes of fulfilling his ultimate stated "goal." *See Toledo v. Publix Super Markets, Inc.*, 30 So. 3d 712 (Fla. 4th DCA 2010) (court of appeal quashing discovery order where party sought law firm client file relating to a different matter holding that "curiosity" about a law firm's records does not satisfy the relevance requirement and explaining that the contents of the "subpoena is a classic 'fishing expedition' and the trial court's order departs from the essential requirements of the law."); *Calvo v. Calvo*, 489 So. 2d 833, 834 (Fla. 3d DCA 1986) (quashing subpoena served on wife's bank for financial records finding them irrelevant: "*indeed, the husband has failed to demonstrate what possible relevance the records might have in the proceeding below other than to harass the wife*."). (emphasis added).

Defendant's incredibly broad and unrelated demands include, for example, Request no. 24: "All documents concerning, relating or referring to your assertions that you met former President Bill Clinton, Former Vice President Al Gore and/or Mary Elizabeth "Tipper" Gore on Little Saint James Island in the U.S. Virgin Islands." *See* Exhibit 6, Request no. 24. Whether or not Jane Doe No. 3 met any of these individuals has absolutely nothing to do with the action before this Court. *See Allstate Ins. Co. v. Langston,* 655 So. 2d 91, 94 (Fla. 2003) (Florida Supreme Court holding that "we do not believe a litigant is entitled carte blanch to irrelevant discovery" and ""It is axiomatic that information sought in discovery must relate to the issues involved in the litigation, as framed in the pleadings.'") (internal citations omitted). Defendant's Request demonstrates a blatant example of abuse of the subpoena power.

Indeed, the face of many of Defendant's subpoena demands demonstrate that he is using the subpoena power of this Court to obtain discovery for the federal action. Request nos. 1, 5, 6 and 9 all reference the "federal action" or specifically cite the declaration and case number "OS-SO736-CIV-MARRA/JOHNSON. Request no. 1, for example, demands: "All documents that reference by name, Alan M. Dershowitz, which support and/or confirm the allegations set forth in Paragraphs 24-31 of your Declaration dated January 19, 2015 and/or Paragraph 49 of your

Declaration dated February 5, 2015, which were filed with the United States District Court for the Southern District of Florida, in Jane Doe #1 and Jane Doe #2 v. United States of America, Case No. OS-S0736-CIV-MARRA/JOHNSON, [ECF No. 291-1] (the "Federal Action")." Defendant should not be using the subpoena power of this Court to issue a non-party subpoena for documents sought for a federal action.[2]

c. Category 3 – Documents that Contain Personal Financial Information Completely Irrelevant to this Action

Defendant also wrongfully abuses the subpoena power to seek personal financial information from this non-party. *See Woodward v. Berkery*, 714 So. 2d 1027, 1034-38 (Fla. 4th DCA 1998) (quashing lower court's discovery order and finding irreparable harm to husband in disclosure of private financial information when wife's *clear purpose was to wrongfully disclose the financial information to the press*) (emphasis added); *see also Granville v. Granville*, 445 So. 2d 362 (Fla. 1st DCA 1984) (court of appeal overturning denial of protective order and finding that private financial information should have been protected from disclosure).

The requests are clearly meant to intimidate and harass her by, for example, seeking information during the time she was the subject of sexual trafficking by Jeffrey Epstein. Request no. 20 seeks "All documents showing any payments or remuneration of any kind made by Jeffery Epstein or any of his agents or associates to you from January 1, 1999 through December 31, 2002." Whether Jeffrey Epstein paid minor children that he sexually trafficked has absolutely nothing to do with the action before this Court and there is no basis to force a non-party who was subject to this abuse to comply with a production demand on this topic. The subpoena also includes request for financial information relating to the media. Apparently, Defendant believes Jane Doe No. 3 has a book "deal" in the works. For example, Request no. 18 seeks: "All documents concerning any monetary payments or other consideration received by you from any

[2] The requests relevant to this category are nos.: 1, 5, 6,7, 8, 9, 12, 13, 14, 22, and 24.

media outlet in exchange for your statements (whether "on the record" or "off the record") regarding Jeffrey Epstein, Alan M. Dershowitz, Prince Andrew, Duke of York, and/or being a sex slave." Whether Jane Doe No. 3 has interacted with the media has nothing to do with the Florida Defamation Action. As explained above, a non-party's personal financial information and other confidential information is subject to protection by this Court. *See Woodward v. Berkery*, 714 So. 2d 1027, 1034-38 (Fla. 4th DCA 1998). Accordingly, the requests relating to financial information from this non-party should be quashed[3].

d. Category 4 – Plainly Privileged Communications

Defendant's subpoena requests seek documents that are plainly privileged. Florida courts are unequivocal in stating that an opposing party can never obtain attorney-client privileged materials. *See Quarles & Brady LLP v. Birdsall*, 802 So. 2d 1205, 1206 (Fla. 2d DCA 2002) (quashing discovery order and noting "undue hardship is not an exception (to disclosure of privileged material), nor is disclosure permitted because the opposing party claims that the privileged information is necessary to prove their case.") (internal citations omitted). Non-party, Jane Doe No. 3, objects to all of Defendant's subpoena requests to the extent that they seek documents protected by the attorney client privilege, work product doctrine, joint defense and common interest privileges and any other relevant privilege. Indeed, Jane Doe No. 3 should be protected from responding to Request no. 25 in its entirety because on its face it seeks solely privileged and confidential information relating to her retention of BSF.[4] *See Westco Inc. v. Scott Lewis' Gardening & Trimming, Inc.*, 26 So. 3d 620, 622 (Fla. 4th DCA 2010) (court explaining that "[w]hen confidential information is sought from a non-party, the trial court must determine whether the requesting party establishes a need for the information that outweighs the privacy

[3] These Requests include nos. 9, 17, 18, 20 and 23.

[4] Specifically, Request no. 25 seeks: "All documents concerning your retention of the law firm Boies, Schiller & Flexner LLP, including but not limited to: signed letter of retainer, retention agreement, explanation of fees, and/or any documents describing the scope of retention."

rights of the non-party."). Defendant has not established any basis for these privileged and confidential documents that outweighs this non-party's privacy rights.

3. <u>The Subpoena Should Be Quashed In Its Entirety. If the Court Will Not Take That Action, at a Minimum, It Should Grant a Protective Order Severely Limiting The Areas Of Inquiry At Deposition And Grant Protections For This Victim Who Is Fearful Of The Defendant.</u>

This Court has the power to preclude and/or limit the deposition of non-party Jane Doe No. 3. Specifically, Florida Rule of Civil Procedure 1.280(c) allows the Court to prevent a deposition from going forward "to protect a party or person from annoyance, embarrassment, oppression or undue burden or expense that justice requires," and courts routinely enter protective orders to reduce the burden on subpoenaed non-parties to a case, as well as in cases where the discovery sough is irrelevant. *See, e.g., Peisach v. Antuna*, 539 So. 2d 544 (Fla. 3d DCA 1989) (holding that the trial judge erred in allowing the deposition of certain non-parties where evidence sought was irrelevant); *see also Citimortgage, Inc. v. Davis*, No. 50 2009 CA 030523, 2011 WL 3360318 (Fla. 15th Cir. Ct. April 4, 2011) (trial court granting protective order precluding a deposition noting "this deposition request is mere harassment" and had no relevance to the underlying dispute where the party was wrongfully using the discovery process for personal gain). Section 4 of Rule 1.280 provides that the Court can also limit the areas of inquiry of a deposition providing "that certain matters not be inquired into, or that the scope be limited to certain matters."

Jane Doe No. 3 contends that the subpoena for her deposition should be quashed. If the Court, however, is inclined to allow a deposition of Jane Doe No. 3, then she respectfully requests the issuance of a Protective Order modifying the subpoena as set forth below.

a. <u>Testimony Limitations</u>

Non-party Jane Doe No. 3 respectfully requests that this Court limit the deposition to questions directly related to Defendant's defamatory statements about Brad Edwards and Paul Cassell. The Court should limit Defendant's ability to engage in a "fishing expedition" of this victim to foster his goal of putting her into "jail" or of bringing a new action against Jane Doe No. 3. *See Peisach v. Antuna*, 539 So. 2d 544 (Fla. 3d DCA 1989); *see also Citimortgage, Inc. v. Davis*, No. 50 2009 CA 030523, 2011 WL 3360318 (Fla. 15th Cir. Ct. Apr. 4, 2011). Defendant should be precluded from asking any questions about Jane Doe No. 3's experiences as a sexually trafficked minor. Defendant should be precluded from questioning Jane Doe No. 3 about individuals that she was sexually trafficked to or about other victims or individuals involved in the sexual trafficking orchestrated by Jeffrey Epstein. Defendant should be precluded from questioning Jane Doe No. 3 about any rapes that occurred when she was a minor child. Defendant should be precluded from questioning Jane Doe No. 3 about anything related to her sexual activity either as a minor or thereafter as these questions would only be intended to embarrass and harass this non-party witness.

b. Language and Harassment Limitations

In addition, Jane Doe No. 3 requests that the Court provide counsel with a cautionary notice, that counsel for Defendant may not harass the non-party victim in any way during the deposition. With respect to the language used at the deposition, the Defendant's counsel should be directed by the Court to not use any of the derogatory terms the Defendant has used in the press including calling Jane Doe No. 3 a "prostitute," a "liar," or a "bad mother" or any other similar derogatory and harassing language.

c. Physical Location Limitations

Non-party Jane Doe No. 3 has a valid and real basis to fear being in physical proximity of the Defendant. *See* Exhibit 8, Affidavit of Jane Doe No. 3. Accordingly, to the extent a deposition is to go forward, we would request that the Court direct that the Defendant not be present in the same room as non-party Jane Doe No. 3 and, instead, follow the testimony electronically from a separate location. In addition, non-party Jane Doe No. 3 respectfully requests that the Court hold that the physical location of the deposition should be the offices of

Jane Doe No. 3's attorney's Boies, Schiller & Flexner LLP.

CONCLUSION

WHEREFORE, non-party Jane Doe No. 3 respectfully requests that this Court grant her Motion to Quash, or alternatively, that the Court enter an order limiting the scope of her document production and deposition as set forth above.

Dated: April 9, 2015

Respectfully submitted,

BOIES, SCHILLER & FLEXNER LLP
401 East Las Olas Boulevard, Suite 1200
Fort Lauderdale, Florida 33301
Telephone: (954) 356-0011
Facsimile: (954) 356-0022

By: /s/Sigrid S. McCawley
Sigrid S. McCawley, Esq.
Florida Bar No. 129305

Attorney for Non-Party Jane Doe No. 3

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that on April 9, 2015, a true and correct copy of the foregoing was served by Electronic Mail to the individuals identified below.

By: /s/Sigrid S. McCawley
Sigrid S. McCawley

Thomas E. Scott Thomas.scott@csklegal.com Steven R. Safra Steven.safra@csklegal.com COLE, SCOTT & KISSANE, P.A. 9150 S. Dadeland Blvd., Suite 1400 Miami, Florida 33156 Renee.nail@csklegal.com Shelly.zambo@csklegal.com Richard A. Simpson rsimpson@wileyrein.com Mary E. Borja mborja@wileyrein.com Ashley E. Eiler aeiler@wileyrein.com WILEY REIN, LLP 1776 K Street NW Washington, D.C. 20006 *Counsel for Defendant Alan Dershowitz*	Jack Scarola SEARCY DENNEY SCAROLA BARNHART & SHIPLEY, P.A. JSX@searcylaw.com 2139 Palm Beach Lakes Blvd. West Palm Beach, FL 33409-6601 *Attorney for Plaintiffs*

EXHIBIT B

From:	Susan Moss
To:	Thomas E. Scott; Steven R. Safra; Simpson, Richard; Borja, Mary; Eiler, Ashley; Jack Scarola; Mary E. Pirrotta; smccawley@bsfllp.com
Subject:	Edward & Cassell v. Dershowitz CACE 15-000072 (05)
Date:	Wednesday, November 04, 2015 11:22:42 AM

Good morning,

I am advising of Judge Lynch's rulings on the above referenced case. His rulings are as follows:

Non Party's motion to quash, or for protective order, regarding subpoena served on non party law firm Boies Schiller: The subpoena, as to the law firm, is quashed

As to the "Jane Doe #3" subpoena: The motion is granted as to request #9, 17, 18, 20 and 23.

The motion is denied as to the other requests, but a confidentiality order shall be entered.

Regarding the deposition: The depo shall be limited to 4 hours without prejudice to request additional time in the future.

The Defendant can be present at the depo.

The depo will be taken at the law firm representing the witness.

There shall be a special master, paid by the Defendant, present at the depo, to rule on objections.

The depo will be limited to the issues of this case without prejudice for another depo, if required, in the future.

The issues and said limitations will be determined by the special master.

Each attorney who had a motion heard, is to prepare the order on their motion for the judges signature, along with sufficient copies, self addressed, self stamped envelopes for all parties.

These orders cannot be submitted through the order portal.

Respectfully,

Susan Moss
Judicial Assistant to Judge Thomas M. Lynch, IV
201 S.E. 6th Street, Rm 920B
Fort Lauderdale, Florida 33301
954-831-7831

EXHIBIT C

IN THE CIRCUIT COURT OF THE 17TH
JUDICIAL CIRCUIT IN AND FOR
BROWARD COUNTY, FLORIDA

CIVIL DIVISION

BRADLEY J. EDWARDS, and
PAUL G. CASSELL,

CASE NO. CACE 15-000072

Plaintiffs,

v.

ALAN DERSHOWITZ,

Defendant.
_______________________________/

ORDER ON BOIES, SCHILLER & FLEXNER LLP AND JANE DOE NO. 3's MOTIONS TO QUASH SUBPOENAS OR FOR PROTECTIVE ORDER

This Cause comes before the Court on November 2, 2015 upon Boies, Schiller & Flexner LLP's Motion to Quash Subpoena Or For Protective Order and Jane Doe No. 3's Motion to Quash Subpoena Or For Protective Order.

Having reviewed the record and being otherwise fully advised, the Court hereby Orders:

1. Non-Party Law Firm Boies, Schiller & Flexner LLP's Motion to Quash Subpoena is GRANTED. The subpoena to Boies, Schiller & Flexner LLP is quashed in its entirety.
2. Non-Party Jane Doe No. 3's Motion to Quash Subpoena or For Protective Order is Denied in Part, Granted in part, as follows:
    a. The Motion to Quash is Granted as to Requests Nos. 9, 17, 18, 20, and 23. The Motion is denied as to the remaining Requests.
    b. A Confidentiality Order shall be entered.
    c. The deposition of Jane Doe No. 3 shall be limited to 4 hours without prejudice to request for additional time in the future.
    d. The deposition will be taken at the law firm representing the witness, Boies, Schiller & Flexner LLP.

e. There shall be a special master, paid by the Defendant, present at the deposition, to rule on objections.

f. The Defendant Alan Dershowitz can be present at the deposition.

g. The deposition will be limited to the issues of this case without prejudice for another deposition, if required, in the future. The issues and said limitations will be determined by the special master.

DONE AND ORDERED in Broward County, Florida on this ____ day of November, 2015.

THOMAS M. LYNCH, IV

NOV 12 2015

Honorable Judge Thomas Lynch
Circuit Judge

TRUE COPY

cc: Counsel of Record

EXHIBIT D

IN THE CIRCUIT COURT OF THE 17TH
JUDICIAL CIRCUIT IN AND FOR
BROWARD COUNTY, FLORIDA

CIVIL DIVISION

BRADLEY J. EDWARDS, and
PAUL G. CASSELL,

CASE NO. CACE 15-000072

Plaintiffs,

v.

ALAN DERSHOWITZ,

Defendant.
_______________________________/

[PROPOSED] CONFIDENTIALITY ORDER

THIS CAUSE COMES before the Court based on its Order dated November 12, 2015 granting, in part, Non-Party Jane Doe No. 3's Motion to Quash Subpoena or for a Protective Order. This Court ordered that "A Confidentiality Order Shall Be Entered."

Accordingly, having reviewed the record and being otherwise duly advised, the Court issues the following Confidentiality Order:

1. The deposition testimony of Non-Party Virginia Giuffre will be designated as "Confidential" and not subject to public disclosure. It may only be filed under seal.
2. Documents produced by Non-Party Virginia Giuffre that are confidential may be marked as "Confidential" and shall be treated in the same manner as confidential testimony.

DONE AND ORDERED in Broward County, Florida on this 12 day of January, 2016.

____
Honorable Judge Thomas Lynch
Circuit Court Judge

cc: Counsel of Record

EXHIBIT E

Filing # [illegible] E-Filed [illegible] PM

IN THE CIRCUIT COURT OF THE SEVENTEENTH JUDICIAL CIRCUIT, IN AND FOR BROWARD COUNTY, FLORIDA

CASE NO.: CACE 15-000072

BRADLEY J. EDWARDS and PAUL G. CASSELL,

Plaintiffs,

vs.

ALAN M. DERSHOWITZ,

Defendant.

______________________________________/

PLAINTIFFS/COUNTERCLAIM DEFENDANT EDWARDS AND CASSELL'S RESPONSE TO DERSHOWITZ'S MOTION TO DETERMINE CONFIDENTIALITY OF COURT RECORDS

Plaintiffs/Counterclaim Defendants Bradley J. Edwards and Paul G. Cassell, by and through their undersigned attorneys, hereby file this response to Dershowitz's Motion to Determine Confidentiality of Court Records. The records at issue are not confidential, and so the Court should deny Dershowitz's motion in its entirety.

The court records at issue are three court filings by attorneys Edwards and Cassell in which they recite their client's (Mr. Virginia Giuffre's) allegations that she was sexually abused by Dershowitz. These records are hardly "confidential" in this defamation case, where the parties have claims and counterclaims about these sexual abuse Allegations. Rather, these records are an important part of this case, since they not only support the conclusion that Dershowitz abused Ms. Giuffre, but also indisputably establish Edwards and Cassell's strong basis for filing the allegations on her behalf. Moreover, contrary to assertions made in Dershowitz's motion, these documents have never been found to be "confidential" by any other court. And Dershowitz has repeatedly referred to

these documents, not only in defamatory statements broadcast worldwide, but also in his pleadings before this Court and in recent depositions. Indeed, Dershowitz said in his media interviews that he wants "everything to be made public" and implied that Edwards and Cassell had something to hide. Accordingly, Dershowitz has failed to carry his heavy burden to justify sealing these presumptively-public documents.

I. DERSHOWITZ HAS NOT JUSTIFIED SEALING ALLEGED DEFAMATORY RECORDS THAT ARE INTEGRAL TO THIS DEFAMATION CASE.

In his motion, Dershowitz never recounts the heavy burden that he must carry to seal the records at issue. To be sure, Florida Rule of Judicial Administration 2.420 allows for the sealing of "confidential" materials. But the Rule begins by recounting the overarching principle that "[t]he public shall have access to all records of the judicial branch of government, except as provided below." Fla. R. Jud. Admin. 2.420(a). This rule is a codification of the Florida Supreme Court's admonition that a "a *strong presumption of openness* exists for all court proceedings. A trial is a public event, and the filed records of court proceedings are public records available for public examination." *Barron v. Florida Freedom Newspapers, Inc.*, 531 So.2d 113, 118 (Fla. 1988) (emphasis added). In light of this presumption of openness, "[t]he burden of proof in [closure] proceedings shall always be on the party seeking closure." *Id.* To obtain a sealing order, the party seeking sealing must carry a "heavy burden." *Id.*

Remarkably, Dershowitz fails to acknowledge these well-settled principles. More important, he even fails to cite (much less discuss) the limited substantive exceptions to this general principle of access – and which specific exception he believes applies to this case. Accordingly, it is impossible for Edwards and Cassell to respond with precision to his motion.

The exceptions that might arguably be in play in this case permit records to be maintained as confidential in order to:

> (i) Prevent a serious and imminent threat to the fair, impartial, and orderly administration of justice;
> (ii) Protect trade secrets;
> (iii) Protect a compelling governmental interest;
> (iv) Obtain evidence to determine legal issues in a case;
> (v) Avoid substantial injury to innocent third parties;
> (vi) Avoid substantial injury to a party by disclosure of matters protected by a common law or privacy right not generally inherent in the specific type of proceeding sought to be closed;
> (vii) Comply with established public policy set forth in the Florida or United States Constitution or statutes or Florida rules or case law . . . .

Fla. R. Jud. Admin. 2.420(c)(9) (codifying the holding in *Barron v. Florida Freedom Newspapers, Inc.*, 531 So.2d 113 (Fla. 1988)). The only exception that seems to even arguably apply here is exception vi, which itself specifically provides that confidentiality is appropriate only where disclosure is "*not generally inherent* in the specific type of proceeding sought to be closed" (emphasis added). Of course, this lawsuit is a defamation action – involving a defamation claim by Edwards and Cassell and a defamation counterclaim by Dershowitz. Disclosure, discussion, and debate about the defamatory statements at issue lies at the heart of the case. Accordingly, disclosure of these materials is "inherent" in the case itself.

The principle that defamatory material in a defamation case cannot be sealed is recognized in *Carnegie v. Tedder*, 698 So.2d 1310 (2d DCA 1997). *Carnegie* involved a claim and counterclaim between two parties (Carnegie and Tedder), one of whom alleged that disclosure of the materials in the records would be harmful to his professional reputation. *Carnegie* recited subsection vi's restriction on release of materials involving a privacy right, but noted that "statements Tedder alleged were defamatory and damaging were allegations in Carnegie's counterclaim for which she seeks damages. These matters were not peripheral to the lawsuit; they were inherent to it." *Id.* at 1312. Of course, exactly the same principle applies here: sexual abuse allegations filed by attorneys Edwards and Cassell for their client Ms. Virginia Giuffre are not peripheral to this lawsuit – they are inherent to it.

To see how "inherent" the sexual abuse allegations are to this lawsuit, the Court need look no further than Dershowitz's counterclaim in this case. Count I of Dershowitz's Counterclaim (styled as "False Allegations in the Joinder Motion") contends that Edwards and Cassell should pay him damages because they "filed a pleading in the Federal Action titled 'Jane Doe #3 and Jane Doe #4's Motion Pursuant to Rule 21 for Joinder in Action' . . . ." Dershowitz Counterclaim at ¶ 14. Dershowitz's Counterclaim then goes on to quote at length from the Joinder Motion. His counterclaim contains, for example, this paragraph recounting the allegations:

> The Joinder Motion then goes on to allege – without any supporting evidence – as follows:
>
>> One such powerful individual that Epstein forced then-minor Jane Doe #3 to have sexual relations with was former Harvard Law Professor Alan Dershowitz, a close friend of Epstein's and well-known criminal defense attorney. Epstein required Jane Doe #3 to have sexual relations with Dershowitz on numerous occasions while she was a minor, not only in Florida but also on private planes, in New York, New Mexico, and the U.S. Virgin

> Islands. In addition to being a participant in the abuse of Jane Doe #3 and other minors, Dershowitz was an eye-witness to the sexual abuse of many other minors by Epstein and several of Epstein's coconspirators. Dershowitz would later play a significant role in negotiating the [Non-Prosecution Agreement] on Epstein's behalf. Indeed, Dershowitz helped negotiate an agreement that provided immunity from federal prosecution in the Southern District of Florida not only to Epstein, but also to "any potential coconspirators of Epstein." Thus, Dershowitz helped negotiate an agreement with a provision that provided protection for himself against criminal prosecution in Florida for sexually abusing Jane Doe #3. Because this broad immunity wouldhave been controversial if disclosed, Dershowitz (along with other members of Epstein's defense team) and the Government tried to keep the immunity provision secret from all of Epstein's victims and the general public, even though such secrecy violated the Crime Victims' Rights Act.

Dershowitz Counterclaim at ¶ 15 (quoting Joinder Motion at 4).

Remarkably, having quoted at length from the Joinder Motion in his Counterclaim in this case, Dershowitz now seeks to have *that very same language* from the Joinder Motion deemed "confidential" and sealed. *Compare* Counterclaim at ¶15 (block quotation above) with Motion to Determine Confidentiality, Exhibit A at 4 (composite exhibit with proposed "confidential" document that includes paragraph beginning "[o]ne such powerful individual that Epstein forced then-minor Jane Doe #3 to have sexual relations with was former Harvard Law Professor Alan Dershowitz, a close friend of Epstein's . . . ."). Dershowitz cannot come before this Court and file a counterclaim seeking damages from Edwards and Cassell for alleged defamatory statements and then ask to have those very same statements placed under seal as "confidential." *See Barron v. Florida Freedom Newspapers*, 531 So.2d at 119 ("although generally protected by one's privacy right, medical reports and history are no longer protected when the medical condition becomes an integral part of the civil proceeding, *particularly when the condition is asserted as an issue by the party seeking closure*" (emphasis added)).

II. JUDGE MARRA'S ORDER IN HIS CASE DOES NOT REQUIRE THAT THE RECORDS BE SEALED IN THIS CASE.

Dershowitz also appears to contend that Judge Marra's order striking some of the materials from the records at issue somehow requires that these stricken materials be kept confidential in this case. Dershowitz's argument misunderstands both the scope of Judge Marra's order and its effect in this case. His argument rests on a truncated – and misleading -- description of the events surrounding Judge Marra's ruling striking certain documents. A more complete description makes clear that Judge Marra has not determined the documents are somehow "confidential" even in the federal Crime Victims' Rights Act case – much less in this separate state defamation action.

Edwards and Cassell filed the federal case pro bono on behalf of two young women who were sexually abused as underage girls by Dershowitz's close personal friend – Jeffrey Epstein. In 2008, Edwards and Casell filed a petition to enforce the rights of "Jane Doe No. 1" and "Jane Doe No. 2" under the Crime Victims' Rights Act (CVRA), 18 U.S.C. § 3771, alleging that the Government had failed to provide them rights with regard to a plea arrangement it was pursuing with Epstein. *Jane Doe No. 1 and Jane Doe No. 2 v. United States*, No. 9:08-cv-80736 (S.D. Fla.). In the course of that case, on October 11, 2011, the victims filed discovery requests with the Government, including requests specifically seeking information about Dershowitz, Prince Andrew, and others. Further efforts from the Government to avoid any discovery followed (*see generally* Docket Entry or "DE" 225-1 at 4-5), ultimately leading to a further Court ruling in June 2013 that the Government should produce documents. DE 189. The Government then produced about 1,500 pages of largely irrelevant materials to the victims (DE 225-1 at 5), while simultaneously submitting 14,825 pages of relevant materials under seal to the Court. The Government claimed that these pages were "privileged" for various reasons, attaching an abbreviated privilege log.

While these discovery issues were pending, in the summer of 2014, Edwards and Cassell, contacted Government counsel to request their agreement to add two additional victims to the case, including Ms. Virginia Giuffre (who was identified in court pleadings as "Jane Doe No. 3"). Edwards and Cassell sought to have her added to the case via stipulation, which would have avoided the need to include any detailed facts about her abuse. Weeks went by and the Government – as it had done on a similar request for a stipulation to add another victim – did not respond to counsel's request for a stipulation. Finally, on December 10, 2014, despite having had four months to provide a position, the Government responded by email to counsel that it was seeking more time, indicating that the Government understood that victims' counsel might need to file a motion with the court on the matter immediately. DE 291 at 3-5. Rather than file a motion immediately, victims' counsel waited and continued to press the Government for a stipulation. *See id.* at 5. Finally, on December 23, 2014 – more than four months after the initial request for a stipulated joinder into the case – the Government tersely indicated its objection, without indicating any reason: "Our position is that we oppose adding new petitioners at this stage of the litigation." *See* DE 291 at 5.

Because the Government now contested the joinder motion, Edwards and Cassell prepared a more detailed pleading explaining the justification for granting the motion. One week after receiving the Government's objection, on December 30, 2014, Ms. Giuffre (i.e., Jane Doe No. 3) and Jane Doe No. 4 filed a motion (and later a corrected motion) seeking to join the case. DE 279 and DE 280. (Note: DE 280 is the first of the three documents Dershowitz seeks to have declared "confidential" in this case.) Uncertain as to the basis for the Government's objection, the motion briefly proffered the circumstances that would qualify the two women as "victims" eligible to assert rights under the CVRA. *See* 18 U.S.C. 3771(e) (defining "crime victim" protected under the Act). With regard to Ms. Giuffre, the motion indicated that when she was a minor, Jeffrey Epstein had trafficked her to Dershowitz and Prince Andrew (among others) for sexual purposes. Jane Doe No. 3 stated that she was prepared to prove her proffer. *See* DE 280 at 3 ("If allowed to join this action, Jane Doe No. 3 would prove the following . . . . "). The motion also provided specific reasons why Jane Doe No. 3's participation was relevant to the case, including the pending discovery issues regarding Dershowitz and Prince Andrew. DE 280 at 9-10 (explaining several reasons participation of new victims was relevant to existing issues).

After the motion was filed, various news organizations published articles about it. Dershowitz also made numerous media statements about the filing, including calling Jane Doe No. 3 "a serial liar" who "has lied through her teeth about many world leaders." http://www.cnn.com/2015/01/06/us/dershowitz-sex-allegation/. Dershowitz also repeatedly called Edwards and Cassell "two sleazy, unprofessional, disbarable lawyers." *Id.* On

January 5, 2015, Dershowitz filed a motion to intervene to argue to have the allegations stricken. DE 282. Dershowitz also argued that Ms. Giuffre had not provided a sworn affidavit attesting to the truth of her allegations. On January 21, 2015, Edwards and Cassell filed a response for Ms. Giuffre and Jane Doe No. 4. DE 291. (Note: This is the second of the three documents Dershowitz seeks to have kept under seal here.) The response enumerated nine specific reasons why Ms. Giuffre's specific allegations against Dershowitz were relevant to the case, including the fact that Ms. Giuffre needed to establish that she was a "victim" in the case, that pending discovery requests concerning Dershowitz-specific documents were pending, and that Dershowitz's role as a defense attorney in the case was highly relevant to the motive for the Government and defense counsel to conceal the plea deal from the victims. DE 291 at 17-26 & n.17. The response included a detailed affidavit from Ms. Giuffre about the sexual abuse she had suffered from Epstein, Dershowitz, and other powerful persons. DE 291-1. On February 6, 2015, Edwards and Cassell filed a further pleading (and affidavit from Ms. Giuffre, *see* DE 291-1) in support of her motion to intervene. (Note: this affidavit is the third of the three documents Dershowitz seeks to have declared confidential.)

On April 7, 2015, Judge Marra denied Ms. Giuffre's motion to join the case. Judge Marra concluded that "at this juncture in the proceedings" details about the sexual abuse she had suffered was unnecessary to making a determination "of whether Jane Doe 3 and Jane Doe 4 should be permitted to join [the other victims'] claim that *the Government* violated their rights under the CVRA. The factual details regarding with whom and where the Jane Does engaged in sexual activities are impertinent to this central claim (i.e., that they were known victims of Mr.

Epstein and the Government owed them CVRA duties), especially considering that the details involve non-parties who are not related to the respondent Government." DE 324 at 5 (emphasis in original). While Judge Marra struck those allegations, he emphasized that "Jane Doe 3 is free to reassert these factual details through proper evidentiary proof, should [the victims] demonstrate a good faith basis for believing that such details are pertinent to a matter presented for the Court's consideration. Judge Marra then denied Ms. Giuffre's motion to join the case, but allowed her to participate as trial witness: "The necessary 'participation' of [Ms. Giuffre] . . . in this case can be satisfied by offering . . . properly supported – and relevant, admissible, and non-cumulative – testimony as needed, whether through testimony at trial . . . or affidavits supported in support [of] the relevancy of discovery requests." DE 324 at 8 (emphasis deleted). In a supplemental order, Judge Marra stated that the victims "may re-refile these documents omitting the stricken portions." DE 325. The victims have recently refiled the documents.

In light of this history, Dershowitz is flatly incorrect when he asserts that "Judge Marra's Order appropriately precludes the unredacted documents from being re-filed in this case on the public docket." Confidentiality Motion at 3. To the contrary, the Order specifically permits factual details about Dershowitz's sexual abuse of Ms. Giuffre to be presented in regard to pertinent matters in the *federal CVRA case*. And certainly nothing in Judge Marra's Order could render those documents confidential in *this state defamation* case, where the central issues swirl around Edwards and Cassell's good faith basis for filing the allegations. Indeed, the order is not binding in any way in this case, because it is res judicata only as to Ms. Giuffre (the moving party in that case), not as to her attorneys Edwards and Cassell. *See Palm AFC Holdings, Inc. v. Palm Beach County*, 807 So.2d 703 (4th DCA 2002) ("In order for res judicata to apply four identities must be present: (1) identity of the thing sued for; (2) identity of the cause of action; (3) identity of persons and parties; and (4) identity of the quality or capacity of the persons for or against whom the claim is made.").

III. EDWARDS AND CASSELL WILL BE PREJUDICED IF THEY ARE BARRED FROM QUOTING FROM THE RECORD WHILE DERSHOWITZ IS PERMITTED TO FREELY REFER TO THEM WHENEVER HE FINDS IT CONVENIENT.

Dershowitz is also incorrect when he asserts that no prejudice will befall Edwards and Cassell if the records are placed under seal. To the contrary, placing the documents under seal would permit Dershowitz to continue to misrepresent and distort what is contained in those records while preventing Edwards and Cassell from correcting those misrepresentations. Dershowitz has repeatedly referred to details in the records when he has found it convenient to do so – treating the records as not confidential in any away. One clear example comes from Dershowitz's recent deposition, where he gratuitously injected into the record a reference to a portion of Ms. Giuffre's affidavit about him watching Ms. Giuffre perform oral sex on Epstein. And then, having injected that gratuitous reference into the record, he proceeded to try to rebut the reference with confidential settlement discussions – but did so by mispresenting what another attorney (David Boies) had said during the settlement discussions. So that the Court may have the full flavor of the exchange, the narrow question to Dershowitz (by attorney Jack

Scarola) and Dershowitz's extended answer are quoted in full – including Dershowitz's reference to the oral sex allegation that he now argues this Court should treat as "confidential":

> Q. [Y]ou [are] aware that years before December of 2014, when the CVRA pleading was filed, that your name had come up repeatedly in connection with Jeffrey Epstein's abuse of minors, correct? . . .
>
> A. Let me answer that question. I am aware that never before 2014, end of December, was it ever, ever alleged that I had acted in any way inappropriately with regard to Virginia [Giuffre], that I ever touched her, that I ever met her, that Ihad ever been with her. I was completely aware of that. There had never been any allegation. She claims under oath that she told you that secretly in 2011, but you have produced no notes of any such conversation. You, of course, are a witness to this allegation and will be deposed as a witness to this allegation. I believe it is an entirely false allegation that she told you in 2011 that she had had any sexual contact with me. I think she's lying through her teeth when she says that. And I doubt that your notes will reveal any such information.
> But if she did tell you that, she would be absolutely, categorically lying. So I am completely aware that never, until the lies were put in a legal pleading at the end of December 2014, it was never alleged that I had any sexual contact with Virginia Roberts.
> I know that it was alleged that I was a witness to Jeffrey Epstein's alleged abuse and that was false. I was never a witness to any of Jeffrey Epstein's sexual abuse. And I wrote that to you, something that you have falsely denied. And I stand on the record. The record is clear that I have categorically denied I was ever a witness to any abuse, that I ever saw Jeffrey Epstein abusing anybody.
>
> *And -- and the very idea that I would stand and talk to Jeffrey Epstein while he was receiving oral sex from Virginia Roberts, which she swore to under oath, is so outrageous, so preposterous, that even David Boies said he couldn't believe it was true.*
>
> MS. McCAWLEY: I object. I object. I'm not going to allow you to reveal any conversations that happened in the context of a settlement discussion.
>
> THE WITNESS: Does she have standing?

> MS. McCAWLEY: I have a standing objection and, I'm objecting again. I'm not going to
>
> THE WITNESS: No, no, no. Does she have standing in this deposition?
>
> MR. SCOTT: Let's take a break for a minute, okay?
>
> THE WITNESS: I'm not sure she has standing.
>
> MR. SCAROLA: Are we finished with the speech?
>
> MR. SCOTT: No. If he --
>
> MR. SCAROLA: I'd like him to finish the speech so that we can get to my question
> and then we can take a break.
>
> A. So the question -- the answer to your question is --
>
> MR. SIMPSON: Wait a minute. Wait a minute. Wait a minute. Please don't disclose something that she has a right to raise that objection if she wants to.
>
> MR. SCOTT: Exactly.

Deposition of Alan Dershowitz (Oct. 15, 2015) at 93-95 (attached as Exhibit 1); *see also* Deposition of Alan Dershowitz (Oct. 16, 2016) (attached as Exhibit 2) (also containing discussion of Ms. Giuffre's affidavit).

The Court should be aware that within approximately two hours of this exchange, Ms. McCawley (David Boies' law partner) released a statement on his behalf, which stated that Dershowitz was misrepresenting what happened: "Because the discussions that Mr. Boies had with Mr. Dershowitz were expressly privileged settlement discussions, Mr. Boies will not, at least at this time, describe what was actually said. However, Mr. Boies does state that Mr.

Dershowitz description of what was said is not true." Statement of Ms. McCawley on Behalf of David Boies (Oct. 15, 2015).

More broadly, the Court can readily see from this passage how Dershowitz is willing to inject into the record a part of Ms. Giuffre's affidavit whenever it serves his purpose – and, indeed, to characterize the part of the affidavit as "preposterous." But then he asks this Court to place the underlying affidavit under seal, so that the Edwards and Cassell stand accused having filed a "preposterous" affidavit without anyone being able to assess the validity of Dershowitz's attack.

Dershowitz has referred to the court records that he now wishes to have the Court declare confidential not only in his deposition, but also in his widely-broadcast media attacks on Edwards and Cassell. For example, Dershowitz appeared on the British Broadcasting Corporation (the BBC) and was asked about the allegations:

> Well, first of all they were made in *court papers* that they don't even ask for a hearing to try to prove them. They put them in *court papers* in order to immunize themselves from any consequences from a defamation suit. *The story is totally made up,* completely out of whole cloth.
> I don't know this woman. I was not at the places at the times. It is part of a pattern of made up stories against prominent people and world leaders. And the lawyers in recent statement challenged me to deny the allegations under oath. I am doing that. I am denying them under oath, thus subjecting me to a perjury prosecution were I not telling the truth. *I am now challenging them to have their client put these charges under oath* and for them to put them under oath. I am also challenging them to repeat them outside of the context of court papers so that I can sue them for defamation. . . . And I will prove beyond any doubt not only that the story is totally false, but it was knowingly false: that the lawyers and the client *conspired together to create a false story.* That is why I am moving for their disbarment in challenges to be provided to the disciplinary committee.

BBC Radio 4 - Sarah Montague (Jan. 3, 2015) (http://www.bbc.co.uk/programmes/p02g7qbc).

Similarly, Dershowitz appeared on NBC's *Today Show* the morning after Edwards and Cassell made a filing for Ms. Giuffre, to say that the Edwards and Cassell – and Ms. Giuffre – were all "lying" in the court documents:

> Question from Savannah Guthrie: *In legal papers from the lawyers*, they say you've had, in fact, the opportunity to be deposed.
>
> Answer from Alan Dershowitz: They're lying. They're lying.
>
> Question: They show letters in which they offered to depose you.
>
> Answer: And they didn't show my letters in response saying, (a), if you ask me about my legal relationship with Epstein and I'll be happy to answer. . . . And I responded that I would be happy to be deposed if you could give me any indication that I would be a relevant witness . . . . They will be proved – all of them [i.e., Cassell, Edwards, and Ms. Giuffre] – to be categorically lying and *making up this story.* And it will be a terrible thing for rape victims. . . . We [Epstein and Dershowitz] had an academic relationship. I was never in the presence of a single, young, underaged woman. When I was with him, it was with prominent scientists, prominent academics. And they're just – again – lying about this. I never saw him doing anything improper. I was not a participant. I was not a witness.

Today Show, Jan. 22, 2015 (emphases added).

As another example, in *Miami Herald*, Dershowitz called the Joinder Motion that he seeks to have sealed "the sleaziest *legal document* I have ever seen. They [Edwards and Cassell] manipulated a young, suggestible woman who was interested in money. This is a disbarrable offense, and they will be disbarred. They will rue the day they ever made this false charge against me" – i.e., Edwards and Cassell will "rue the day" they ever filed the Joinder Motion. *Miami Herald* (Jan. 3, 2015).

Most remarkably, Dershowitz took the public airwaves to represent that he wanted all of the information surrounding the allegations to "be made public," while implying that Edwards and Cassell had something to hide. For example, on the BBC he claimed that he wanted"everything to be made public":

> Q: Would you encourage that it now be made public?
>
> A: Of course, of course. *I want everything to be made public. I want every bit of evidence in this case to be made public. I want every allegation to be made public.* I want to know who else she's accused of these horrible crimes. We know that she accused Bill Clinton of being on Jeffrey Epstein's island and participating in sex orgy with underage girls. The records of the Secret Service will prove that President Clinton never set foot on that island. So that she lied. Now it's possible to have a case of mistaken identification with somebody like me. It's impossible to have a case of mistaken identification with Bill Clinton.
> My only feeling is that if she has lied about me, which I know to an absolute certainty she has, she should not be believed about anyone else. She's lied clearly about me, she's lied clearly about Bill Clinton. We know that. We know that she's lied about other public figures, including a former prime minister and others who she claims to have participated in sexual activities with. So I think it must be presumed that all of her allegations against Prince Andrew are false as well.
> I think he [Prince Andrew] should clear the air as well.
> If you're squeaky clean and if you have never done anything like this, you must fight back with all the resources available to you. And that's what I will do. I will not rest or stop until the world understands no only that I had nothing to do with any of this, but that she deliberately, with the connivance of her lawyer, lawyers, made up this story willfully and knowingly.

BBC Radio 4 - Sarah Montague (Jan. 3, 2015) (http://www.bbc.co.uk/programmes/p02g7qbc).

In another widely-broadcast interview on CNN, Dershowitz implied that there is no evidence supporting the allegations against him:

> *Ask them [Edwards and Cassell] if they have any evidence . . . .* They're doing it for money. She's getting money for having sold her story. She wants to sell the book. They're trying to get into this lawsuit. They see a pot of gold at the end of the rainbow. They're [Edwards and Cassell] prepared to lie, cheat, and steal. These are unethical lawyers. This is Professor Cassell who shouldn't be allowed near a student. This is Professor Cassell, who is a former federal judge, thank God he no longer wears a robe. He is essentially a crook. He is essentially somebody who's distorted the legal profession. . . . Why would he charge a person with a sterling reputation for 50 years on the basis of the word alone of a woman who is serial liar, who has lied about former Prime Ministers, former Presidents, has lied demonstrably.

CNN Live (with Hala Gorani) (January 5, 2015). Of course, by placing "the evidence" in this case under seal, Dershowitz will be free to continue to try and insinuate that Edward and Cassell – and their client, Ms. Giuffre – had no evidence supporting the allegations against him, even though a mountain evidence strongly support Ms. Giuffre's allegations. See Deposition of Paul Cassell (Oct. 16, 2015) at 61-117 (Exhibit 3); see also Depo of Pual Cassell (Oct. 17, 2015) (Exhibit 4).

CONCLUSION

The Court should deny Defendant/Counterclaim Plaintiff Alan Dershowitz's motion to place documents regarding Ms. Giuffre's allegations against him under seal.

I HEREBY CERTIFY that a true and correct copy of the foregoing was sent via E-Serve to all Counsel on the attached list, this 23rd day of November, 2015.

/s/ Jack Scarola
Jack Scarola
Florida Bar No.: 169440
Attorney E-Mail(s): jsx@searcylaw.com and mep@searcylaw.com
Primary E-Mail: scarolateam@searcylaw.com
Searcy Denney Scarola Barnhart & Shipley, P.A.
2139 Palm Beach Lakes Boulevard
West Palm Beach, Florida 33409
Phone:(561) 686-6300
Fax:(561) 383-9451

SEAN D. REYES
Utah Attorney General

By: JONI J. JONES
JOEL A. FERRE
Assistant Utah Attorneys General

Bradley J. Edwards
FARMER, JAFFE, WEISSING,
EDWARDS, FISTOS & LEHRMAN, P.L.
425 North Andrews Avenue, Suite 2
Fort Lauderdale, Florida 33301
Telephone (954) 524-2820
Facsimile (954) 524-2822
E-mail: brad@pathtojustice.com

And

Paul G. Cassell
Pro Hac Vice Motion Pending
S.J. Quinney College of Law at the
University of Utah
383 S. University St.
Salt Lake City, UT 84112
Telephone:801-585-5202
Facsimile:801-585-6833
E-Mail:cassellp@law.utah.edu

Attorneys for Plaintiffs Bradley J. Edwards and Paul G. Cassell

COUNSEL LIST

Thomas Emerson Scott, Jr., Esquire
Thomas.scott@csklegal.com; Steven.safra@csklegal.com
Cole Scott & Kissane P.A.
9150 S Dadeland Boulevard, Suite 1400
Miami, FL 33156

Phone: (305)-350-5329
Fax: (305)-373-2294
Attorneys for Defendant

Richard A. Simpson (pro hac vice)
rsimpson@wileyrein.com
Mary E. Borja (pro hac vice)
mborja@wileyrein.com
Ashley E. Eiler (pro hac vice)
aeiler@wileyrein.com
WILEY REIN LLP
1776 K St. NW
Washington, DC 20006
Phone: (202) 719-7000
Fax: (202) 719-7049

COUNSEL LIST

Sigrid Stone McCawley, Esquire
smccawley@bsfllp.com;
sperkins@bsfllp.com; ftleserve@bsfllp.com
Boies Schiller & Flexner, LLP
401 E Las Olas Boulevard., Suite 1200
Fort Lauderdale, FL 33301
Phone: (954)-356-0011
Attorneys for Alan M.Dershowitz, Esquire

Thomas Emerson Scott, Jr., Esquire
Thomas.scott@csklegal.com;
Steven.safra@csklegal.com;
Renee.nail@csklegal.com;
shelly.zambo@csklegal.com
Cole Scott & Kissane P.A.
9150 S Dadeland Boulevard, Suite 1400
Miami, FL 33156
Phone: (305)-350-5329/Fax: (305)-373-2294
Attorneys for Alan M.Dershowitz, Esquire

Bradley J. Edwards, Esquire
staff.efile@pathtojustice.com;
brad@pathtojustice.com;
maria@pathtojustice.com
Farmer Jaffe Weissing Edwards Fistos & Lehrman, P.L.
425 N Andrews Avenue, Suite 2
Fort Lauderdale, FL 33301
Phone: (954)-524-2820/Fax: (954)-524-2822
Attorneys for Alan M.Dershowitz, Esquire

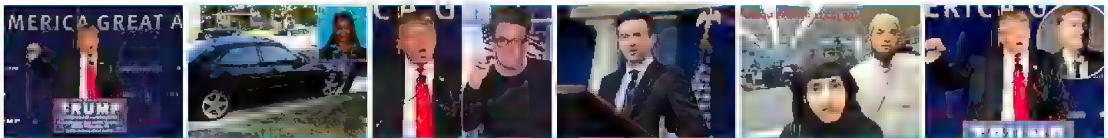
Kenneth A. Sweder, Esquire
ksweder@sweder-ross.com
Sweder & Ross, LLP
131 Oliver Street
Boston, MA 02110
Phone: (617)-646-4466/Fax: (617)-646-4470
Attorneys for Alan M.Dershowitz, Esquire

Ashley Eiler, Esquire
AEiler@wileyrein.com
Mary E. Borja, Esquire
MBorja@wileyrein.com
Richard A. Simpson, Esquire
RSimpson@wileyrein.com
Wiley Rein, LLP
1776 K Street NW
Washington, DC 20006
Phone: (202)-719-4252/Fax: (202)-719-7049
Attorneys for Alan M.Dershowitz, Esquire

Joni J. Jones, Esquire
jonijones@utah.gov
Assistant Utah Attorney General
160 E 300 S
Salt Lake City, UT 84114
Phone: (801)-366-0100/Fax: (801)-366-0101
Attorneys for Paul Cassell

EXHIBIT F
UNDER SEAL

EXHIBIT G



Prince Andrew and the 17-year-old girl his sex offender friend flew to Britain to meet him

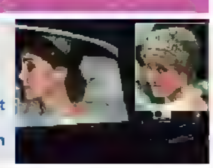
By SHARON CHURCHER
UPDATED: 08:02 EST, 2 March 2011

- **Virginia Roberts reveals she is 'Jane Doe 102' in Jeffrey Epstein case**
- **Mother-of-three spent four years as millionaire's personal masseuse**
- **She describes being flown across world to meet Prince Andrew**
- **Epstein trained her 'as a prostitute for him and his friends'**

As the UK's special representative for international trade, the Duke of York holds an important position, requiring sound judgement and widespread respect

But those qualities have been thrown into question since photographs of Prince Andrew with his billionaire financier friend Jeffrey Epstein, a convicted child-sex offender who was jailed for 18 months for soliciting underage prostitutes, appeared last weekend.

Today, however, even more serious doubts are cast on his suitability after a woman at the centre of the Epstein case revealed to The Mail on Sunday that she had, as a 17-year-old employed by Epstein, been flown across the world to be introduced to the Prince.



First meeting: Prince Andrew puts his arms around 17-year-old Virginia, centre

On one of those occasions Virginia Roberts was subsequently paid $15,000 (£9,400). Her shocking account of her four years as Epstein's personal masseuse is supported by court documents, an eyewitness, photographs and flight details of Epstein's private jets.

One picture, said to have been taken by Epstein during Andrew's first encounter with the girl in March 2001 and published today by The Mail on Sunday, shows the Prince with his arm around her waist.

This is not the first time the Duke of York's judgment and choice of associates have been questioned. He appears to relish the company of super-rich oil billionaires from the Middle East, North Africa and the former Soviet Union.

The peculiar sale of his former marital home to a Kazakh businessman for £15 million after it had languished unsold for five years at £12 million has never been satisfactorily explained.

In the recent leak of American diplomatic cables it was revealed that he had criticised an official corruption investigation into the huge Al-Yamamah arms deal between Britain and Saudi Arabia, while he is also said to be close to Saif Al-Islam Gaddafi, son of the beleaguered Libyan president, and may have had a role in the early release of Lockerbie bomber Abdelbaset Al Megrahi.

But it is Andrew's friendship with Epstein, whom he has known since at least 2000, and with Epstein's confidante Ghislaine Maxwell, daughter of the late disgraced newspaper baron Robert Maxwell, that gives most concern.

While on the streets, I slept with men for money. I was a paedophile's dream

He was first seen with the pair on holiday in Thailand, and was pictured cavorting with Ghislaine at a Halloween fetishthemed party in Manhattan

The photograph that appeared last weekend shows the prince strolling through Central Park with 58-year-old Epstein. Andrew was said to have spent four days at his New York mansion in December, when he was joined by other distinguished guests, including Woody Allen, at a dinner

It is by no means the first New York soiree Andrew has attended as Epstein's guest.

A lengthy profile of the financier in Vanity Fair magazine some years ago reported that Andrew was a guest at a cocktail party thrown by Epstein and Maxwell packed with young Russian models. 'Some guests were horrified,' said the article's author, Vicky Ward.

It should not be forgotten that Epstein is a registered sex offender after recently completing his sentence for offences relating to child prostitution.

However, he avoided trial on more serious charges that carried a potential life sentence. And no one reading The Mail on Sunday's interview with the woman who was prepared to testify against him can be in any doubt of the seriousness of the charges.

Epstein, a Wall Street money manager who once counted Bill C inton and Donald Trump among his friends, became the subject of an undercover investigation in 2005 after the stepmother of a 14-year o d girl c aimed she was paid $200 (£125) to give him an 'erotic massage'.

The subsequent FBI probe uncovered at east 20 gir s levelling sexual allegations against him. Eventua ly, Epstein struck a 'plea bargain' with prosecutors · a practice not permitted under British law under which he was allowed to plead guilty to two relatively minor charges.

Police claim that his donations to politicians and his 'dream team' of influential lawyers deterred prosecutors from bringing more serious charges of sex-trafficking The deal certainly kept the names of a lot of Epstein's famous friends out of an embarrassing court case.

However, an unusual part of the agreement was that Epstein's alleged victims were a lowed to bring civil proceedings against him.

He has so far made 17 out-of-court settlements, and some cases are ongoing. One of these girls was to have been a key witness for the prosecution had the case gone to trial. She was just 15 years old when she was drawn into Epstein's exploitative world in 1998.

In her civil writ against him, under the pseudonym Jane Doe 102' she a leged that her duties included being 'sexually exploited by Epstein's adult male peers inc uding royalty'.

Now, horrified by the evidence of Epstein and Andrew enjoying each other's company in New York, Jane Doe 102 has agreed to waive her anonymity and tell for the first time her deeply disturbing story.

Her real name is Virginia Roberts and she now lives in Austra ia, where she is a happily married mother of three.

Over the course of a week during which she spoke at length to The Mail on Sunday, she appeared sometimes vulnerable, and sometimes stee y, but always quietly reso ute and consistent.

Revisiting events from a past that she had hoped she had left behind,Virginia occasionally buried her face in her hands.

Some recollections – and, for reasons of taste, not all the details can be included here – caused her to flush with shame. 'I'm telling you things that even my husband didn't know,' she said.

Virginia, who has undergone counselling to try to come to terms with her past, is honest about her initiation into Epstein's depraved world

She was a troubled teenager, whose slender figure, delicate complexion, hesitant voice and soulful b ue eyes made her look young for her years.

Born in Sacramento, California, in August 1983, Virginia spent her early years on a small ranch on the West Coast of America.

This seemingly idyllic childhood ended when she was sexually molested by a man close to her family.

The fallout from that led to her parents temporarily splitting up. Blaming herself, Virginia began to get into trouble Aged 11, she was sent to live with an aunt but repeatedly ran away.

Living on the streets, she was beaten up and slept with at least two older men in return for food. 'I was a paedophile's dream,' she says.

Three years later, she was reunited with her family and started a new life with her father who had moved to Palm Beach,

Florida, where he was maintenance manager at Donald Trump's country club, Mar-a-Lago.

Virginia got a part-time job as a changing room assistant –which is where, soon after her 15th birthday, she met Ghislaine Maxwell, who invited her to work as Epstein's personal masseuse.

A new life: Virgina, now a mother-of-three, in Australia

'I was wearing my uniform – a white miniskirt and a skin-tight white polo top – when I was approached by Ghislaine,' Virginia says.

'I told her I wanted to become a masseuse and she said she worked for a very wealthy gentleman who was looking for a travelling masseuse.

I'd get training and be paid well.'Virginia's father gave his blessing, believing his daughter was being handed the opportunity to learn a skill and to work for a wealthy and respectable emp oyer.

He drove her to Epstein's pink mansion on the Palm Beach waterfront – he also owns a nine-storey home in New York, the city's biggest private residence; a 7,500-acre ranch called 'Zorro' in New Mexico and Little Saint James, a private 70 acre atoll in the US Virgin Islands.

Virginia says: 'Ghislaine said I was to start immediately and that someone would drive me home.

My father left and I was told to go upstairs.' She was led by another woman through Epstein's bedroom into a massage room where he lay face down naked on a table.

He started to interviewed Virginia This was unconventional, but Virginia had no suspicions Presumably, she thought, this was how the wealthy conducted their business.

Epstein elicited the information that Virginia had been a runaway, and was no longer a virgin

Virginia was then told to start massaging Epstein, under the instructions of the woman who had shown her in. The massage quickly developed into a sexual encounter.

Virginia was uncomfortable, but reluctant to deny such important people. 'My face was red with embarrassment,' she says. 'But I felt under immense pressure to please them.

The whole time it was going on, they were promising me the world, that I'd travel with Jeffrey on his private jet and have a well-paid profession.' Afterwards, she was given two $100 bills and told to return the next day.

That was the beginning of the four years she spent with Epstein.

For three of those years, she was under Florida's age of consent, which is 18.

Troubled teenager: Virginia on the billionaire's Zorro ranch in New Mexico in 2001

Virginia was fascinated by his life story: the son of a humble New York City parks worker, he was a teacher before becoming a Wall Street broker and friends with the upper echelons of the political, financial and academic establishment.

As a confused teenager, Virginia easily fell into the practice of sexually gratifying him for money.

He guaranteed her a minimum of $200 each time she gave him what he called an 'erotic massage.'

Virginia said: 'I would always receive the money immediately. He would give me the cash from a wad he carried in a black duffel bag or an assistant paid me.

'And, because of the way Epstein had warped her sensibilities, every time she took the cash, Virginia felt even more indebted to him. Secretly, he was also preparing her for an even more disturbing role.

'Basically, I was training to be a prostitute for him and his friends who shared his interest in young girls,' she says: 'After about two years, he started to ask me to "entertain" his friends.'

It started when Epstein called Virginia at the Palm Beach apartment he had rented for her.

She recalls: 'He said, "I've got a good friend and I need you to fly to the island to entertain him, massage him and make him feel how you make me feel."

Epstein had trained me to do whatever men wanted. I told myself I was special

He didn't spell out what I had to do. He didn't have to. 'He'd trained me to do whatever a man wanted. I was shocked but I told myself he was sharing me around because he trusted me and I was special.

I was worried, but I would do anything to keep Jeffrey happy and to keep my place as his number one girl.

He would keep telling me how lucky I was with the life I was leading and the money I was making. It was easy to fall into his grasp.

'The way it usually worked was I'd be sent to meet a man on the private island Jeffrey owned in the Caribbean, or at his ranch in New Mexico, which was really isolated.' She was 'given' to men ranging in age from their 40s to their 60s.

They included a well-known businessman (whose pregnant wife was asleep in the next room), a world-renowned scientist, a respected liberal politician and a foreign head of state.

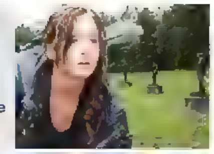
None appeared to think the arrangement was unusual. Virginia says there were many other girls in Epstein's circle and that she was paid extra money to help recruit them.

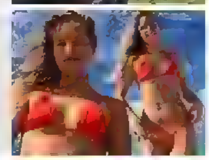
'They would lounge around the Palm Beach house, the ranch or the island, nude or topless,' she says. 'But I was one of the very few he trusted as "special" and chosen to "entertain" his friends.'

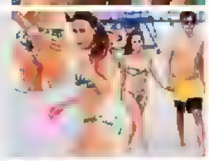
Virginia took the sedative Xanax to detach herself from sordid reality. 'It was an escape drug,' she says. 'It made me calm and helped me forget about what I had to do. I was up to eight pills a day.'

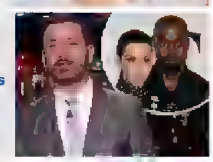
Epstein had no objection to Virginia's use of prescription drugs, no doubt recognising that they made her even more malleable. 'I didn't want to go back to the life I'd had before' she says.

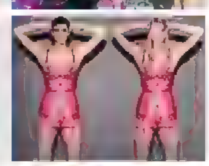
'That made me totally obedient.' Despite the fact that Epstein was, essentially, her pimp, this life now seemed normal to Virginia. 'I felt that he and Ghislaine really cared for me,' she said.

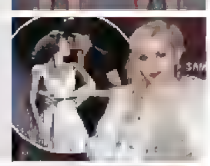


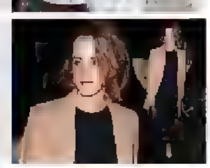


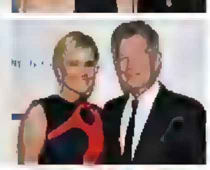












'We'd do family things, like watch Sex And The City and eat popcorn. 'A lot of it was very glamorous. I met famous friends of his such as Al Gore and Heidi Klum and Naomi Campbell. He introduced me as his "travelling masseuse."

Counselling: Virginia at her mother's home at Palm Beach in 1998

Some people mistook me for his daughter. 'When we were in New York or Palm Beach, Ghislaine and I would shop all day.

Jeffrey bought me jewellery – diamonds were his favourite – and wonderful furniture. He was paying me very well because I'd give him sex whenever he wanted it.'

She was, she says, delighted when Epstein invited her to accompany him on a six-week trip in 2001.

'He said we'd be going to Europe and North Africa to meet architects and interior decorators because he wanted to redo his New Mexico house.

I threw my arms around him and gave him a peck on the cheek.' They flew to Paris, then Spain, then Tangier.

Finally, they went to London. 'After we landed, we drove straight to Ghislaine's house,' says Virginia. 'I was given a small upstairs bedroom. The following morning, Ghislaine came in.

She was chirpy and jumped on the bed saying, "Get up, sleepyhead. You've got a big day. We've got to go shopping. You need a dress as you're going to dance with a Prince tonight."

'She said I needed to be "smiley" and bubbly because he was the Queen's son.

Ghislaine and I went to Burberry, where she bought me a £5,000 bag, and to a few other designer stores where we bought a couple of dresses, a pair of embroidered jeans and a pink singlet, perfume and make-up.

We got back to Ghislaine's house at around 4pm and I ran straight upstairs to shower and dress.

When I went downstairs, Ghislaine and Jeffrey were in the lounge. There was a knock at the door. Ghislaine led Andrew in and we kissed each other on the cheek. 'Ghislaine served tea from a porcelain pot and biscuits. She knew Sarah Ferguson and they talked fondly about Andrew's daughters.

Then Ghislaine asked Andrew how old he thought I was and he guessed 17 and they all laughed. Ghislaine made a joke that I was getting too old for Jeffrey.

She said, "He'll soon have to trade her in." It was widely known that he liked young girls.' The four of them went out to dinner and on to Tramp nightclub where, she says, Andrew danced with her.

'After about an hour-and-a-half, we drove back to Ghislaine's.

All of us went upstairs and I asked Jeffrey to snap a picture of me with the Prince. I wanted something to show my Mom. Ghislaine and Jeffrey left us after that, and later Andrew left.

'In the morning, Ghislaine said, "You did well. He had fun". We flew straight back to the States.' The Mail on Sunday has confirmed that the tycoon's jet flew to Paris on March 6, 2001, continuing to Granada, Tangier and London, before returning to New York.

On the last leg of the trip, Virginia was paid about $15,000 (£9,400) by Epstein. 'It was amazing money, more than I'd ever made on a trip with him before.

He didn't say there was any special reason, but I felt like I'd done everything he wanted. He was very pleased.'

There is no suggestion that there was any sexual contact between Virginia and Andrew, or that Andrew knew that Epstein paid her to have sex with his friends.

I took eight pills a day to help me forget what I had to do. It made me calm.

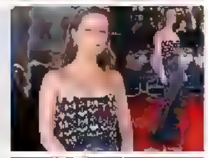
However, the Prince must have been aware of Epstein's conviction when he stayed with him in New York in December

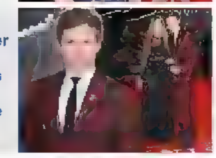
Virginia says she met Andrew for a second time around Easter 2001 at Epstein's Manhattan mansion.

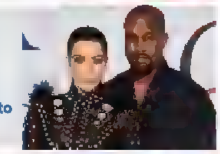
'When I got to the mansion, I was told, "Get ready. You are meeting someone in the office" – which is what they called the library. Andrew was sitting there in a big leather armchair.

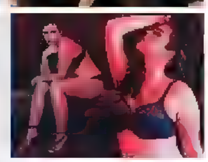
Ghislaine had just given him a present, a big toy that was his Spitting Image puppet. 'He was smiling ear-to-ear. He looked like a kid whose parents were taking him to Disney World.

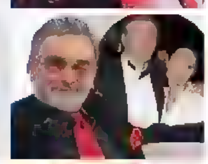
A beautiful girl called Johanna Sjoberg who worked for Jeffrey was sitting on Andrew's knee. Ghislaine guided me over to Andrew and I think he recognised me, though I don't know if he remembered my name.

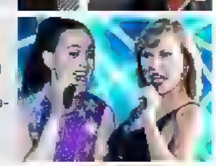


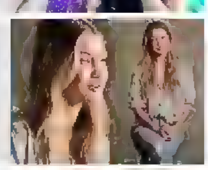


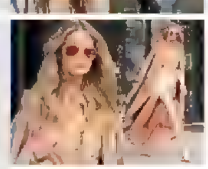












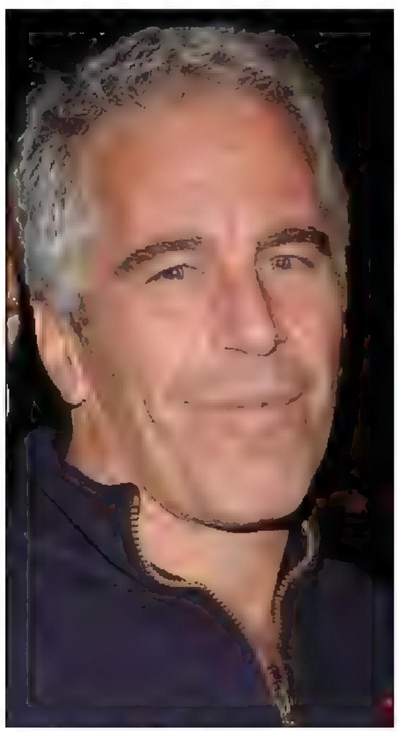
Organiser: Ghislaine Maxwell looks on as Andrew put his arm around Virginia. Robert Maxwell's daughter invited her to work as Epstein's personal masseuse soon after her 15th birthday

We kissed on the cheek and Ghislaine placed me on his other knee.' Johanna spoke to The Mail on Sunday three years ago about this incident, which took place when she was 21.

She said: 'Ghislaine put the puppet's hand on Virginia's breast, then Andrew put his hand on my breast. It was a great joke. Everybody laughed.' After this, Virginia was paid, by Epstein, around $400 (£250).'

She met Andrew for the third and final time on Epstein's Caribbean island, Little Saint James. Virginia was never under the British legal age of consent when she met Andrew. She was 17 during the first two encounters and 18 at the third.

By now, however, Epstein, had started to hint that she was getting 'too old' for him.



Conviction: Jeffrey Epstein

But during one trip to the island, Epstein and Ghislaine made their most astonishing proposition, and one which repulsed her. 'They said Jeffrey wanted me to have his child,' she says.

'They said I was part of their family and I was beautiful,young, loyal and nurturing and would be a great mother.

They said I would have to sign a contract relinquishing rights to the child and consenting to Jeffrey having as many relationships as he liked. In return I would have my own mansion in Palm Beach and a large monthly payment, a percentage of his income.'

This, finally, was a wake-up call to Virginia and she began to see the way in which she had been groomed.

'It was a smack in the face,' she says. 'I finally realised this wasn't ever going to be a real relationship but I knew if I refused, I'd be thrown back on the streets. So I said, "I'm too young. I want to get my massage credentials, then maybe we'll do it".'

The tycoon took her at her word and, for her 19th birthday in August 2002, flew her to Thailand where he enrolled her in a massage course.

Shortly after arriving there, she met an Australian martial arts expert called Robert They fell in love and, just ten days later, married in a Buddhist ceremony.

'I called Jeffrey and told him I'd fallen madly in love,' Virginia says 'I was hoping he'd be delighted. But he said, "Have a nice life," and hung up on me.' The couple now have two sons, aged five and four, and a daughter who recently turned one

'The first few months after I married Robert were the worst,' she says. 'I couldn't bring myself to tell him much. No man wants to know his wife has been traded out.

'I felt very alone. I was having panic attacks and seeing a psychiatrist and was on anti-depressants.

' Virginia was beginning to put her Epstein days behind her when, three years ago, she was phoned by the FBI.

'They said they had found photos of me at Jeffrey's Palm Beach house,' she says. '[Epstein had] hidden cameras watching me the entire time even when I was in the bathroom. I was so embarrassed.

'I told the FBI that my true purpose was sexual. They told me everything he did was illegal because I was under age.' (The age of consent in Florida is 18).

'They said that if it had to go to trial, they'd need me because I'd lived with him and that made me a key witness. I was very afraid, because he had so much power, but eventually I agreed to testify.

I was glad he'd finally been found out He shouldn't be hurting other girls. Fol owing Epstein's arrest, investigators are believed to have found a list of men's names on his computer and asked him whether they had been 'treated' to sexual encounters with his menage of minors.

'He took the Fifth Amendment, refusing to answer, indicating that if he were to answer the question, it could be incriminating,' a source told The Mai on Sunday

Epstein struck a deal resulting in what commentators characterised as a 'slap on the wrist' for him, and ended up serving 13 months of his sentence, much of it in a liberal work-release programme Lawyer Brad Edwards, who represented several of Epstein's victims, said: 'Rather than punish him the way they would an average Joe, they sent a clear message that with enough money and power and influence, the system can be bought '

Virginia was spared her the humiliation of having to go before a jury, and has kept her feelings bottled up until last weekend's photograph of Andrew with Epstein triggered distressing memories.

Virginia says: 'I am appalled. To me, it's saying, "We are above the law." But Jeffrey is a monster.'

Last night, neither Epstein, Ghislaine Maxwell nor Prince Andrew would comment on Virginia's story.

Share or comment on this article

Filing # 37957690 E-Filed 02/08/2016 06:28:17 PM

IN THE CIRCUIT COURT OF THE 17th
JUDICIAL CIRCUIT IN AND FOR
BROWARD COUNTY, FLORIDA

CIVIL DIVISION

CASE NO. CACE 15-000072

BRADLEY J. EDWARDS, and
PAUL G. CASSELL,

Plaintiffs,

v.

ALAN DERSHOWITZ,

Defendant.
_______________________________/

RESPONSE TO DEFENDANT ALAN DERSHOWITZ'S MOTION FOR CLARIFICATION OF CONFIDENTIALITY ORDER OR RELIEF FROM THAT ORDER

Non-Party Virginia Giuffre, by and through undersigned counsel, hereby responds to Defendant Alan Dershowitz's Motion for Clarification of Confidentiality Order or Relief From that Order and states as follows:

FACTUAL BACKGROUND

On November 12, 2015 this Court issued an Order granting in part, non-party Virginia Giuffre's Motion to Quash the subpoena served by Defendant Alan Dershowitz and ordered protective limits relating to her deposition. *See* Exhibit A, November 12, 2015 Order. On December 18, 2016, this Court entered a Confidentiality Order holding that non-party Virginia Giuffre's deposition would be confidential. *See* Exhibit B, December 18, 2015 Confidentiality Order. On January 16, 2016, Ms. Giuffre testified at her deposition in accordance with this Court's Order. The deposition was labelled confidential in accordance with this Court's Order. As the Court knows, Ms. Giuffre was a victim of sexual trafficking when she was a minor child. Indeed, the U.S. Attorney's Office for the Southern District of Florida has specifically recognized her as a "victim" of federal sex offenses. Unsurprisingly, her deposition contains highly sensitive information about her experiences as a minor child, including detail descriptions of sexual crimes committed against her.

ARGUMENT

1. Non-Party Virginia Giuffre Agrees That Her Deposition Should Be Provided Confidentially To Law Enforcement to Investigate All The Crimes Committed Against Her

Defendant Dershowitz seeks to have the Court grant an exception to the confidential nature of the deposition so that it can be provided to the Office of the State Attorney and the Office of the United States Attorney for investigative purposes. Specifically, Defendant Dershowitz states in his motion that he is hoping to have law enforcement investigate whether Ms. Giuffre committed perjury by stating in her previously filed affidavit that it is her recollection that she witnessed former President Bill Clinton on Jeffrey Epstein's island in the United States Virgin Islands ("USVI")[1].

Setting aside Defendant Dershowitz's baseless claims of perjury, Ms. Giuffre agrees that her confidential deposition should be provided to law enforcement, including the United States Attorney and the State Attorney in each jurisdiction where any alleged crimes occurred so that they may investigate all of the crimes committed against her when she was a minor child. To ensure that justice is served and that Defendant Dershowitz's request is not just another charade designed only to bully a sexual abuse victim, Ms. Giuffre asks the Court to impose the following reasonable conditions relating to the disclosure:

- Mr. Dershowitz agrees and is directed to cooperate with authorities and answer all questions relating to the investigation of crimes against Ms. Giuffre.

[1] Defendant Dershowitz conveniently ignores that publicly available flight logs of Jeffrey Epstein's private planes demonstrate that President Clinton travelled with Jeffrey Epstein and others to various locations throughout the world including Europe, Africa and Asia. *See* Exhibit C, Excerpts of Flight Logs from Jeffrey Epstein's private plane.

- Mr. Dershowitz agrees to make his client, Jeffrey Epstein, and others with relevant testimony and with whom he has testified he shares a "common interest" – at least Epstein and Maxwell – available to any law enforcement agency reviewing any alleged criminal activities; or in the alternative, to attest to this Court that those necessary witnesses have consented to full cooperation in the investigation Mr. Dershowitz is seeking permission to initiate.

- Mr. Dershowitz agrees to waive the statute of limitations in all jurisdictions for any criminal conduct he participated in or was aware of relating to Ms. Giuffre so that law enforcement can pursue any necessary charges. Defendant Dershowitz proclaimed that he was willing to waive any statute of limitation for criminal conduct so this should not be an issue. *See* Exhibit D, January 12, 2016 Deposition Transcript of Alan Dershowitz at 395. "I had talked about the statute of limitations for criminal purpose was what I said, that I would waive the statute of limitations for criminal purposes."

- Mr. Dershowitz agrees to provide the names and contact information for each State Attorney and United States Attorney for which he has or is planning to provide information relating to Ms. Giuffre; and agrees to jointly, with Ms. Giuffre's counsel, request that the State Attorney and United States Attorney, in the relevant jurisdictions, investigate all potential criminal conduct. Both parties may provide any relevant information they have that may assist the authorities with their investigation.

- For all other purposes non-party Ms. Giuffre's January 16, 2016 deposition transcript shall remain confidential and sealed other than for confidential disclosure to law enforcement as described above.

2. <u>Mr. Dershowitz Has No "Evidence" of Perjury And Instead Is Simply Trying To Bully This Victim</u>

As explained above, Defendant Dershowitz wrongly suggests to this Court that non-party Virginia Giuffre has committed perjury in an effort to taint the Court against this victim. His only "evidence" of this alleged perjury is a self-serving opinion from his retained expert that an "absence of records" in response to a FOIA request, establishes that former President Clinton was never on Jeffrey Epstein's island in the USVI. Defendant Dershowitz misrepresents the government's response. The government is only required to conduct a reasonable search of readily accessible records. Accordingly, an "absence of records" response does not mean that records do not exist. It simply means that in the course of the search, no records were found. *See Cunningham v. U.S. Dept. of Justice*, 961 F.Supp. 2d 226, 236 (D.C. 2013) (court reasoning that "[t]he adequacy of a search is measured by a standard of reasonableness… The question is not whether other responsive records may exist, but whether the search was adequate."); *Wilbur v. C.I.A.*, 355 F.3d 675, 678 (D.C. 2004) (court explaining that "the agency's failure to turn up a particular document, or mere speculation that as yet uncovered documents might exist, does not undermine the determination that the agency conducted an adequate search for the requested records."). Moreover, when dealing with a former President's security detail travel, there are a number of reasons why the government may not disclose those records.

As explained above, public flight records from Jeffrey Epstein's private plane show that President Clinton traveled with Jeffrey Epstein on multiple occasions. Nevertheless, if Defendant Dershowitz wants to pursue this issue before the Court, then he needs to produce for deposition testimony in this case and the proposed criminal investigation, the other witnesses that were present on the island at the time former President Clinton was alleged to have visited, including his client Jeffrey Epstein, and Ghislaine Maxwell, to whom he has testified he is party to a joint defense agreement. It is worth noting on that point, that despite Mr. Epstein's counsel's attendance at depositions in this case, and Defendant Dershowitz's claim that Mr. Epstein is still his client, Mr. Epstein has taken extreme measures to avoid being deposed in this case despite being ordered to deposition by this Court.

Indeed, it is also noteworthy that during Defendant Dershowitz's recent deposition, counsel for Mr. Edwards and Mr. Cassell asked Defendant Dershowitz the following question: "Was Virginia Roberts lying when she said Jeffrey Epstein socialized with Bill Clinton during the relevant time period?" Depo Tr. Of Alan Dershowitz, Vol. 4, January 12, 2016 at 511. Before Defendant Dershowitz could answer, Mr. Dershowitz's legal counsel interposed an attorney-client privilege objection. *Id.* Perhaps Mr. Epstein's defense counsel can provide to the Court an appropriate privilege log regarding that objection – and all the communications between Mr. Epstein and Defendant Dershowitz that would have been revealed in answer to that question – so that the Court will have the benefit of a full record in ruling on this motion.

CONCLUSION

WHEREFORE, Non-Party Virginia Giuffre respectfully requests that this Court allow a limited release of her confidential deposition transcript to law enforcement subject to the terms set forth above on pages 2-3.

Dated: February 8, 2016

Respectfully submitted,

BOIES, SCHILLER & FLEXNER LLP
401 East Las Olas Boulevard, Suite 1200
Fort Lauderdale, Florida 33301
Telephone: (954) 356-0011
Facsimile: (954) 356-0022

By: /s/Sigrid S. McCawley
Sigrid S. McCawley, Esq.
Florida Bar No. 129305

Attorney for Non-Party Virginia Giuffre

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that on February 8, 2016, a true and correct copy of the foregoing was served by Electronic Mail to the individuals identified below.

By: /s/Sigrid S. McCawley
Sigrid S. McCawley

Thomas E. Scott Thomas.scott@csklegal.com Steven R. Safra Steven.safra@csklegal.com COLE, SCOTT & KISSANE, P.A. 9150 S. Dadeland Blvd., Suite 1400 Miami, Florida 33156 Renee.nail@csklegal.com Shelly.zambo@csklegal.com Richard A. Simpson rsimpson@wileyrein.com Mary E. Borja mborja@wileyrein.com Ashley E. Eiler aeiler@wileyrein.com WILEY REIN, LLP 1776 K Street NW Washington, D.C. 20006 *Counsel for Defendant Alan Dershowitz*	Jack Scarola SEARCY DENNEY SCAROLA BARNHART & SHIPLEY, P.A. JSX@searcylaw.com 2139 Palm Beach Lakes Blvd. West Palm Beach, FL 33409-6601 *Attorney for Plaintiffs*

EXHIBIT A

IN THE CIRCUIT COURT OF THE 17TH JUDICIAL CIRCUIT IN AND FOR BROWARD COUNTY, FLORIDA

CIVIL DIVISION

BRADLEY J. EDWARDS, and PAUL G. CASSELL,

CASE NO. CACE 15-000072

Plaintiffs,

v.

ALAN DERSHOWITZ,

Defendant.

_______________________________/

ORDER ON BOIES, SCHILLER & FLEXNER LLP AND JANE DOE NO. 3's MOTIONS TO QUASH SUBPOENAS OR FOR PROTECTIVE ORDER

This Cause comes before the Court on November 2, 2015 upon Boies, Schiller & Flexner LLP's Motion to Quash Subpoena Or For Protective Order and Jane Doe No. 3's Motion to Quash Subpoena Or For Protective Order.

Having reviewed the record and being otherwise fully advised, the Court hereby Orders:

1. Non-Party Law Firm Boies, Schiller & Flexner LLP's Motion to Quash Subpoena is GRANTED. The subpoena to Boies, Schiller & Flexner LLP is quashed in its entirety.
2. Non-Party Jane Doe No. 3's Motion to Quash Subpoena or For Protective Order is Denied in Part, Granted in part, as follows:
    a. The Motion to Quash is Granted as to Requests Nos. 9, 17, 18, 20, and 23. The Motion is denied as to the remaining Requests.
    b. A Confidentiality Order shall be entered.
    c. The deposition of Jane Doe No. 3 shall be limited to 4 hours without prejudice to request for additional time in the future.
    d. The deposition will be taken at the law firm representing the witness, Boies, Schiller & Flexner LLP.

e. There shall be a special master, paid by the Defendant, present at the deposition, to rule on objections.

f. The Defendant Alan Dershowitz can be present at the deposition.

g. The deposition will be limited to the issues of this case without prejudice for another deposition, if required, in the future. The issues and said limitations will be determined by the special master.

DONE AND ORDERED in Broward County, Florida on this ____ day of November, 2015.

Honorable Judge Thomas Lynch
Circuit Judge

cc: Counsel of Record

EXHIBIT B

IN THE CIRCUIT COURT OF THE 17TH
JUDICIAL CIRCUIT IN AND FOR
BROWARD COUNTY, FLORIDA

CIVIL DIVISION

CASE NO. CACE 15-000072

BRADLEY J. EDWARDS, and
PAUL G. CASSELL,

Plaintiffs,

v.

ALAN DERSHOWITZ,

Defendant.
_______________________________/

ORDER ON BOIES, SCHILLER & FLEXNER LLP'S MOTION TO SEAL

This Cause comes before the Court on December 18, 2015 upon Boies, Schiller & Flexner LLP's Motion to Seal "Exhibit B, Affidavit of Alan M. Dershowitz Regarding Meetings with David Boies" to Defendant Alan M. Dershowitz's Motion in Limine to Overrule Objections As to Application of Settlement Rules, Filing # 35429605 E-Filed 12/11/2015 at 10:08:04 a.m.

Having reviewed the record and being otherwise fully advised, the Court hereby Orders:

the motion to seal is granted.

DONE AND ORDERED in Broward County, Florida on this 18 day of December, 2015.

THOMAS M. LYN[illegible]
DEC 18 2015
TRUE COPY

____
Honorable Judge Thomas Lynch
Circuit Court Judge

cc: Counsel of Record

EXHIBIT C

Date 2002	Aircraft Make and Model	Aircraft Identification Mark	Points of Departure & Arrival From	To	Miles Flown	Flight No.	Remarks, Procedures, Maneuvers, Endorsements	Number of Landings	Aircraft Category... AIRPLANE	GLIDER	HELICOPTER
JAN 21	N906GM	C-421B	PBI	FXE			D.O. AT BROTHERS AVIATION TO HAVE ENGINES SET UP		6		
MAY 11	N908JE	B-727-31	PBI	JFK		93	JE, SK	1/1	2 5		
11	B-727-31	N908JE	JFK	LFPB		94	JE, GM, SK	1/1	7 2		
13	"	"	LFPB	LFPG		95	REPOSITION	/1	5		
13	"	"	LFPG	EGGW		96	JE, GM, SK		8		
16	"	"	EGGW	LFMN		97	JE, GM, SK		1 8		
19	"	"	LFMN	UNNT		98	JE, GM, SK	1/1	6 7		
20	"	"	UNNT	RJTA		99	JE, GM, SK	1/	6 5		
22	"	"	RJTA	VHHH		100	JE, GM, SK, PRESIDENT BILL CLINTON, MEYE, DOUG BANDS, JANICE, JESSICA PETE RATHGEB [illegible]	/1	4 0		
23	"	"	VHHH	ZGSZ		101	SAME AS ABOVE PETE RATHGEB		4		
23	"	"	ZGSZ	WSSS		102	SAME AS ABOVE PETE RATHGEB		3 4		
25	"	"	WSSS	VTBD		103	SAME AS ABOVE PETE RATHGEB		2 2		
25	"	"	VTBD	WBSB		104	SAME AS ABOVE PETE RATHGEB		2 6		
27	"	"	WBSB	WRRR		105	JE, GM, SK PETE RATHGEB	1/1	2 1		
29	"	"	WRRR	VCBI		106	JE, GM, SK PETE RATHGEB		5 2		
29	"	"	VCBI	OMDB		107	JE, GM, SK PETE RATHGEB		4 4		
30	"	"	OMDB	LFPB		108	JE, GM, SK PETE RATHGEB		3 8		
31	"	"	LFPB	EGGW		109	JE, GM, SK PETE RATHGEB	1/1	1 0		
JUN 7	"	"	EGGW	EIDW		110	REPOSITION		1 7		

I certify that the statements made by me on this form are true.

Pilot's Signature David Rodgers

	Number of Landings	AIRPLANE	GLIDER	HELICOPTER	
Page Total	6/7	57 4			
Amount Forward	6537/6180	8635 1	3 5	112	6
Total to Date	6539/6181	8692 5	3 3	112	6

Date 2002	Aircraft Make and Model	Aircraft Identification Mark	From	To	Miles Flown	Flight No	Remarks, Procedures, Maneuvers, Endorsements	Number of Landings	Airplane	Glider	Helicopter
FEB 18	B-727-31	N908JE	PBI	ABY		64	JG, GM, SK, GLEN DUBIN, 2 FEMALES, RICHARD COOK	✓	1 3		
18	"	"	ABY	JFK		65	JG, GM, SK, GLEN DUBIN, 2 FEMALES, RICHARD COOK		2 1		
20	"	"	JFK	MRY		66	JG, SK, GERALDINE LAYBOURNE, KIT LAYBOURNE, STEVEN [illegible], NINA + TIM ZAGAT	1/1	5 9		
~~23~~	~~"~~	~~"~~	~~MRY~~	~~VNY~~		66	DAVID ROCKWELL, CAROLYN MILLER, MICHAEL WOLF, DANIEL DENNETT, RICHARD [illegible], JOHN BROCKMAN, KATINKA MATSON, CINDY KELLY BOVINA, RICHARD [illegible] COOK				
23	"	"	MRY	VNY		67	JG, SK, KELLY BOVINA, NINA + TIM ZAGAT, 6 OTHER PASSENGERS	1/1	8		
23	"	"	VNY	JAX		68	B CHECK FMS, ENROUTE RNAV PROCEDURES, LARRY MORRISON SIC, T, ↓, P LANDING	1/	4 2		
25	G-1159B	N908JE	~~TGB~~ PBI	~~ABQ~~ TCB		156[illegible]	ROGER, SLOANE + SPENCER BIRNETTE, ARLENE (NANNY)	1/1	2 5		
25	"	"	TGB	ABQ		192	GM	1/1	4 3		
28	B-727-31	N908JE	JAX	PBI		71	B CHECK	1/1	1 2		
MAR 6	B-727-200	SIMULATOR	MIA	MIA			STEEP TURNS, STALLS, HOLDING		2 0		
7	"	"	"	"			EMERGENCY DESCENT, RTO, UNUSUAL ATTITUDES, ENGINE FAILURES, MISSED APPROACH		2 0		
7	"	"	"	"			EMERGENCY DESCENTS, 2 ENGINES OUT, WINDSHEAR, HYDRAULIC FAILURES		2 0		
8	"	"	"	"			EMERGENCY DESCENT, HOLDING, ENGINE FIRE, WHEEL WELL FIRE		2 0		
10	B-727-31	N908JE	TIST	JFK		72	JG, GM, SK, AP, CINDY, 2 FEMALES	✓	4 1		
14	"	"	JFK	PBI		73	JG, SK, JOE PAGANO, JULIE, TODD		2 5		
17	"	"	PBI	JFK		74	JG, SK, JOE PAGANO, JULIE, TODD, 1 FEMALE	✓	2 6		
19	"	"	JFK	EGGW		75	BILL CLINTON, DOUG BAND, 3 SECRET SERVICE, JG, GM, SK	1	6 7		
21	"	"	EGGW	JFK		76	BILL CLINTON, DOUG BAND, 10 SECRET SERVICE, JG, GM, SK, NAOMI CAMPBELL, 1 MALE		7 8		
22	"	"	JFK	PBI		77	JG, NICOLE JUNKERMANN; SCANLIN FOR LM		2 6		

I certify that the statements made by me on this form are true.

Pilot's Signature: [signature]

	Number of Landings	Airplane	Glider	Helicopter
Page Total	85/8	56 6		
Amount Forward	6517 / 6169	8541 9	3 3	112 6
Total to Date	6526 / 6175	8594 5	3 3	112 6

Date 2002	Aircraft Make and Model	Aircraft Identification Mark	Points of Departure & Arrival From	To	Miles Flown	Flight No	Remarks, Procedures, Maneuvers, Endorsements	Number of Landings	Aircraft Category... Airplane	Glider	Helicopter
JUL 13	B-727-31H	N908JE	LFMN	GMTT		120	JE, GM, SK, AP, CINDY LOPEZ		2 1		
13	"	"	GMTT	GMME		121	JE, GM, SK, CL, AP		7		
13	"	"	GMME	LPAZ		122	JE, GM, SK, AP, CL, PRESIDENT CLINTON DOUG BANDS MIKE, & SECRET SERVICE		2 4		
13	"	"	LPAZ	JFK		123	JE, GM, SK, AP, CL, PRESIDENT BILL CLINTON, DOUG BANDS, MIKE, & SECRET SERVICE		5 8		
18	"	"	JFK	PBI		124	JE, SHELLEY LEWIS, 2 FEMALES		2 2		
19	"	"	PBI	JAX		125	KRISTY RODGERS, GREG HOLBURT, ALYSSA HOLBERT – C CHECK		1 0		
AUG 4	G-1159B	N909JE	PBI	MVY		1583	JE, 1 FEMALE		2 8		
4	"	"	MVY	BED		1584	JE, 1 FEMALE		7		
4	"	"	BED	TEB		1585	JE, 1 FEMALE		9		
5	"	"	TEB	SAF		1586	JE, SK, 2 FEMALES	1/1	3 9		
6	C-172XP	N7395P	AEG	AEG			172 CHECK OUT	3/3	9		
6	206L3	N474AW	ZORRO	AEG							
15	B-727-31H	N908JE	JAX	JAX		126	C-CHECK FLIGHT TEST PETE RODRIGUES	1/1	9		
16	"	"	JAX	PBI		127	RETURN FROM C-CHECK PETE RODRIGUES		1 1		
17	G-1159-B	N909JE	SAF	TEB		1389	JE, GM, SK, CINDY LOPEZ, VIRGINIA ROBERTS, JOHN MALON, GAL ARDY, ALFRED, MARGARITA, NICK SIMMONS,		3 7		
18	"	"	TEB	PBI		1390	JE, VIRGINIA ROBERTS, 1 FEMALE		2 4		
21	B-727-31H	N908JE	PBI	TIST		128	JE, SHELLEY LEWIS		2 5		
25	"	"	TIST	JFK		129	JE, SK	1/1	3 6		
28	"	"	JFK	LFPB		130	JE, SK, CINDY LOPEZ, 1 FEMALE		6 4		

I certify that the statements made by me on this form are true

Pilot's Signature Daniel Rodriguez

	Number of Landings	Airplane	Glider	Helicopter
Page Total	6/6	14 0		5
Amount Forward	6547 6194	8740 1	3 3	112 6
Total to Date	6553 6200	8754 1	3 3	113 1

Date 19__ 2002	Aircraft Make and Model	Aircraft Identification Mark	Points of Departure & Arrival: From	To	Miles Flown	Flight No.	Remarks, Procedures, Maneuvers, Endorsements	Number of Landings	Aircraft Category... AIRPLANE		GLIDER		HELICOPTER
AUG 31	B-727-314	N908JC	LFPB	EGBB		131	JC, NICOLE JUNKERMANN	1/1	1	0			
SEP 2	"	"	EGBB	LFPB		132	JC, NICOLE JUNKERMANN		1	0			
3	"	"	LFPB	JFK		133	JC, SK, CINDY LOPEZ, JEAN LUC BRUNEL	1/1	7	5			
4	"	"	JFK	PBI		134	JC, 1 FEMALE		2	5			
8	G-1159B	N909JC	PBI	TEB		1592	JC, ANDREA, 2 FEMALES		2	7			
9	G-1159B	"	TEB	BED		1593	JC, SHELLEY LEWIS	1/1		9			
9	"	"	BED	TEB		1594	JC, SHELLEY LEWIS			9			
10	"	"	TEB	TIST		1595	JC, SHELLEY LEWIS, ANDREA 1 MALE		3	8			
15	"	"	TIST	PBI		1596	JC, SK, DEANE FLEETWOOD	1/1	2	6			
21	B-727-314	N908JC	JFK	LPAZ		136	PRESIDENT WILLIAM J. CLINTON, [illegible] KEVIN SPACEY, CHRIS TUCKER JE, GM,	/[illegible]	5	2			
22	"	"	LPAZ	DGAA		137	SK, CL, CHAUNTAE DAVIES, ANDREA MITROVICH DOUG BAND, DAVID SLANG, JIM KENNEDY	/1	5	7			
23	"	"	DGAA	DNAA		138	ERIC NONACS, RONNY SLAYER, CASEY WASSERMAN, RON BURKLE, GAYLE SMITH	9/	1	7			
24	"	"	DNAA	HRYR		139	SAME AS ABOVE LESS RON BURKLE		4	0			
25	"	"	HRYR	FQMA		140	SAME AS ABOVE PLUS IRA MAGAZINER		3	8			
26	"	"	FQMA	FACT		141	SAME AS ABOVE	1/	2	4			
27	"	"	FACT	FAJS		142	SAME AS ABOVE LESS JC, GM, SK, CL CHAUNTAE DAVIES, ANDREA MITROVICH		2	[illegible]			
28	"	"	FAJS	FACT		143	SAME AS ABOVE LESS GAYLE SMITH, IRA MAGAZINER		2	0			
29	"	"	FACT	DGAA		144	SAME AS ABOVE ADD JC, GM, SK, CL CD, AM		6	2			
29	"	"	DGAA	LFPB		145	SAME AS ABOVE	/1	6	5			

I certify that the statements made by me on this form are true.

Pilot's Signature David Rodgers

	Number of Landings	Airplane		Glider		Helicopter	
Page Total	5/6	62	5				
Amount Forward	6553 / 6200	8784	1	3	3	1	3
Total to Date	6558 / 6266	8846	6	3	3	11	3

Date 19— 2002	Aircraft Make and Model	Aircraft Identification Mark	Points of Departure & Arrival From	To	Miles Flown	Flight No.	Remarks, Procedures, Maneuvers, Endorsements	Number of Landings	Aircraft Category... AIRPLANE	GLIDER	HELICOPTER
29 OCT	B-727-31H	N908JE	LFPB	EGGW		146	SAME AS ABOVE LESS JE, SK, CD, LAURA + CASEY WASSERMAN	1/1	— 9		
1	"	"	EGGW	LFPB		147	GM, NICK + GAWENA SIMMONS		1 0		
2	"	"	LFPB	JFK		148	JE, GM, SK, CL, ANDREA METROVICH, NAOMIE CAMPBELL		8 1		
3	"	"	JFK	PBI		149	JE, SK, ANDREA METROVICH, NICK SIMMONS, CHAD, HIGMALL		2 4		
6	G-1159B	N909JE	PBI	TEB		1597	JE, SK, ANDREA, RYAN, I FC MISC		2 3		
11	"	"	TEB	PBI		1598	JE, GM, ANDREA, RYAN, GIAN		2 5		
14	"	"	PBI-OPF	-PBI		1599	MARK PURTER - PPC ILS SINGLE ENGINE + GO AROUND SE	1/	X		
15	"	"	PBI	TEB		1600	JE, GM, ANDREA METROVICH, RYAN DEONNE	1/1	2 4		
17	"	"	TEB	TIST		1601	JE, SK, ANDREA METROVICH		3 7		
21	"	"	TIST	PBI		1602	JE, SK, AM, ANNA HANKS, JULIETTE BRYANT	1/1	2 6		
30	"	SIMULATOR	JFK	MIV SIM			STEEP TURNS, STALLS, SLOW FLIGHT, RECOVERY, HOLDING		2 0		
30	"	"	MIV	JFK SIM			SINGLE ENG APPROACH [illegible], EMERGENCY DESCENT, ENGINE, ELECTRICAL		2 1		
31	"	"	SWF	JFK SIM			ENGINE [illegible], SINGLE ENGINE APPROACH [illegible] ENGINE FAILURE, CIRCLING APPROACH		2 [illegible]		
NOV 3	"	N909JE	PBI	TIST		1603	JE, SK, ANDREA METROVICH	3/3	2 [illegible]		
6	B-727-200	SIMULATOR	MED	MED SIM			[illegible] ENGINE FAILURE [illegible]	1/1	2 3		
6	"	"	MED	MED			[illegible], EMERGENCY DESCENT SYS. [illegible] LOW OIL PRESSURE, HIGH PACK TEMP, [illegible] ALL GENERATORS		2 [illegible]		
10	G-1159B	N909JE	TIST	TEB		1604	JE, SK, KELLY MCCARTHY, MA[illegible], [illegible] BIRKS	1/	2 0		
15	"	"	TEB	TIST		1605	[illegible] RYAN DEONNE, ME [illegible], ANNA [illegible]		4 3		
17	"	"	TIST	PBI		1606	RETURN [illegible]		3 6		
									— 6		

I certify that the statements made by me on this form are true.

Pilot's Signature [signature]

	Number of Landings	Airplane	Glider	Helicopter
Page Total	9/8	49 5		
Amount Forward	6558 / 6206	8846 6	3 3	11
Total to Date	6567 / 6214	8896 1	[illegible]	[illegible]

EXHIBIT D

IN THE CIRCUIT COURT OF THE SEVENTEENTH JUDICIAL CIRCUIT IN AND FOR BROWARD COUNTY, FLORIDA

CASE NO.: CACE 15-000072

BRADLEY J. EDWARDS and PAUL G. CASSELL,

Plaintiffs,

vs.

ALAN M. DERSHOWITZ,

Defendant.

_______________________________/

VIDEOTAPE CONTINUED DEPOSITION OF

ALAN M. DERSHOWITZ

VOLUME 3
Pages 334 through 461

Tuesday, January 12, 2016
9:46 a.m. - 11:59 a.m.

Tripp Scott
110 Southeast 6th Street
Fort Lauderdale, Florida

Stenographically Reported By:
Kimberly Fontalvo, RPR, CLR
Realtime Systems Administrator

Virginia believes she has been hurt by you or Prince Andrew, she should be suing for damages?

MR. SCOTT: Objection, form.

A. I don't recall.

BY MR. EDWARDS:

Q. Have you ever said, "I welcome that lawsuit"?

MR. SCOTT: Objection, form. You should identify what you're publishing from.

MR. EDWARDS: I'm asking if he's ever made the statement to anybody other than his lawyers.

A. I know I said that I welcome any opportunity to put Virginia Roberts under oath and to have her cross examined to establish that everything she's said about me is totally false, which I'm confident we can do.

BY MR. EDWARDS:

Q. After making that statement, were you presented by Mr. Scarola, I believe, with a waiver of the statute of limitations in order to oblige you on filing that lawsuit?

A. Mr. Scarola wrote me a nasty letter that I didn't think it was appropriate to respond to. What I did do is state under oath that she was lying and state under oath that I did not have sex with her, thereby opening myself up --

MR. EDWARDS: This again is nonresponsive to the question.

SPECIAL MASTER POZZUOLI: Let me stop -- repeat the question so we can get an answer.

BY MR. EDWARDS:

Q. The question is, were you sent a waiver of the statute of limitations to sign so that Virginia Roberts could bring that lawsuit? Were you?

A. I don't believe that's right. I don't believe that's right. I don't believe that I was sent a -- something that would give her the opportunity -- a waiver to file that lawsuit. That's not my recollection. If you show it to me, maybe my recollection will be refreshed.

I had talked about the statute of limitations for criminal purpose was what I said, that I would waive the statute of limitations for criminal purposes. If a prosecutor were to investigate this case and come to the conclusion that there was probable cause that I had committed any offense with Virginia Roberts, I would then waive the statute of limitations.

Q. Okay.

Robin Solomon

In a career spanning three decades in academic medicine, Robin Solomon has combined patient care, clinical research, philanthropy and business development expertise to build strategic alliances, brand recognition and revenue. In evolving leadership roles at The Mount Sinai Medical Center, she has developed a reputation for her innovative programming, quick grasp of new situations and common sense approaches to challenging problems.

She is recognized for her comprehensive understanding of the healthcare sector, including operating models, the political and regulatory environment, shifting demographics, financial opportunities and economic trends. Her eighteen years as an ICU nurse informs the way she navigates patient care on a human level, collaborates with physicians and persuasively conveys the potential impact of philanthropic gifts to donors.

Board and Donor Relations, Medical Navigation. As Vice President, Trustee Services, Robin developed and led a high-touch, best in class Donor Relations program. The services include 24/7 medical direction, referrals and oversight to the Mount Sinai Board and donor community to optimize the patient experience. Robin established relationships and gained participation from prestigious corporate, financial, and legal firms including Goldman Sachs, Paul Weiss Rifkind, Schulte Roth, Donna Karan, Warner Music and Federated Department Stores. This program has generated hundreds of physician referrals and over $20 million in incremental revenue each year. As a patient advocate, Robin has participated in thousands of multi-disciplinary patient care teams from initial consults to final outcomes.

Philanthropy. Robin's trusted relationships with medical, business, civic, and community leaders enable her to secure funds needed to advance medical, academic and patient care initiatives. During her tenure, she raised over $29 million in donations from individuals, corporations and foundations. Robin conceived, co-chaired and ran "A Toast to Health" an annual wine benefit, which raised $17 million dollars over nine years, consistently attained over 85% event underwriting. She has directed philanthropic funds to capital projects including an upgrade of the emergency room, complete renovation of the Neonatal and Pediatric ICU's and state-of-the art modernization of a medical school lecture hall. Her efforts have also provided $2.8 million for scholarships to The Mount Sinai School of Medicine and supported programs including Visiting Doctors, Palliative Care Fellowships, and perinatal bereavement social work.

Global Health. Robin has planned and executed cogent responses to critical international medical needs. At the request of President Clinton's office, she arranged for physicians and medical supplies to be sent with first responders in the days following the 2010 earthquake in Haiti. She collaborated with military leaders, USAID, and Dr. Paul Farmer of Partners in Health to organize transport of fully supplied surgical teams and re-open the University Hospital in Port au Prince. In 2006, in close coordination with Liberian President Ellen Johnson Sirleaf and her ministers, Robin established a medical mission to Liberia. She obtained private funding to send physicians and support personnel to provide medical care and surgical interventions. The project has continued and led to administration of the first chemotherapy and pap screening clinic in Monrovia.

Clinical Research Leadership. As Assistant Dean for Clinical Trial Operations and Business Development, she designed and led the Office of Clinical Trials, key to advancing translational research. She launched marketing and outreach programs to convey Mount Sinai's research capabilities. These efforts attracted ongoing contracts from Novartis, Smith Kline Beecham, Amgen and Pfizer and other pharmaceutical and biotech companies. She facilitated over $18 million in grant funding, increased clinical trials by 40% and grew patient enrollment by over 320% annually.

Robin was Senior Research Coordinator on a CDC multicenter study measuring the effects of using safety devices for drawing blood, inserting IV's and giving injections. The manufacturers extended the study solely at The Mount Sinai Hospital. This research opened the multi-billion dollar market for safety devices. Earlier, Robin coordinated and performed clinical research at Cornell University Medical College's Surgical Metabolism Laboratory.

Patient Care. Robin's clinical expertise was developed over more than eighteen years as an ICU nurse at New York Hospital, Columbia-Presbyterian Hospital and The Mount Sinai Medical Center. The scope of her medical knowledge reflects years of hands-on patient care and thousands of decision-making conversations with physicians and families.

Robin holds a Master of Science in Exercise Physiology, *cum laude* from the University of Massachusetts and Bachelor of Science in Nursing, cum laude from the University of Florida.

Robin Solomon

[redacted]

EXECUTIVE PROFILE

In a career spanning three decades in academic medicine, Robin Solomon has combined patient care, clinical research, philanthropy and business development expertise to build strategic alliances, institutional recognition and revenue. In evolving leadership roles, she has developed a reputation for her clear and effective communication style, innovative programming, quick grasp of new situations and common sense approaches to difficult problems.

Robin is recognized for her comprehensive understanding of the healthcare sector, including operating models, the political and regulatory environment, shifting demographics, financial opportunities and economic trends. Her 18 years as an ICU nurse informs the way she navigates the patient experience on a human level, collaborates with physicians, and persuasively conveys philanthropic opportunities to donors.

PROFESSIONAL EXPERIENCE

The Mount Sinai Medical Center, New York, N.Y. (1993-2016)

Vice President Special Projects, Office of the President *(2015-present)*

- Receive and direct high-level medical and administrative inquiries from the Board and donor community. Manage transition of prior responsibilities to the Executive Services group.
- Secure philanthropic support, oversee medical procedures and personal care for several global health patients in collaboration with Dr. Randall Owen and his team.

Vice President, Trustee Services *(2005-2015)*
Director of Development *(2002-2005)*
Launched and led a Donor Relations program to provide high-touch, best in class 24/7 medical navigation and oversight for trustees, donors and patients referred by leadership and staff. Established and maintained philanthropic relationships with corporations, law firms and financial companies including Goldman Sachs, Paul Weiss Rifkind, Schulte Roth, Donna Karan, Warner Music and Federated Department Stores. Extended the Mount Sinai Medical Center brand to a broader patient base.

- Raised over $29 million in donations from individuals, corporations and foundations.
- Co-chaired and ran "A Toast to Health" an annual wine benefit. Raised over $17 million dollars over nine years. Consistently attained over 85% event underwriting.
- Directed philanthropic funds to capital projects including emergency room upgrade, complete renovation of the Neonatal and Pediatric ICU's and state-of-the art modernization of a medical school lecture hall.
- Provided $2.8 million for scholarships to The Mount Sinai School of Medicine, the single largest donation to date. Supported programs including, Visiting Doctors, Palliative Care Fellowships, perinatal bereavement social work and emergency medicine residency training.
- Engaged three new Trustees and four new members of the Mindich Child Health and Development Institute Advisory Board who collectively contributed over $10 million.
- Generated physician referrals resulting in hundreds of additional surgical admissions and over $20 million in incremental revenue each year.
- Managed refurbishing of ten family waiting areas. Obtained $1 million in pro bono services by notable decorators. Handled all logistics and ensured compliance with infection control regulations.

Global Health Project Leader (2002-2015)
Planned and executed medical missions with global health experts:

Haiti Earthquake Response. At the request of President Clinton's office, arranged for physicians and medical supplies to be sent with first responders in the days following the earthquake in 2010. Collaborated with Dr. Paul Farmer of Partners in Health to re-open the University Hospital in Port au Prince.

- Organized transport of 5 full surgical teams and all required equipment. Over ten days, the teams treated 400 patients and performed 50 surgeries.
- Personally requested emergency use of corporate jet from CEO of Honeywell and delivered needed supplies during the early days of the mission.
- Communicated with military leaders to obtain necessary landing slots, protection from the 82nd Airborne, and tents from USAID to house patients and orphans.
- Negotiated resupply without cost of all supplies and medications provided by Mount Sinai.

Liberia Medical Missions. Patients received the first chemotherapy in Liberia and a pap screening clinic was opened in Monrovia in 2014 in response to the large number of advanced stage cancers diagnosed during annual missions. Program continued through 2014, to resume post-Ebola.

- Coordinated directly with Liberian President Ellen Johnson Sirleaf and her ministers in 2007, 2008.
- Obtained private funding to send 30 physicians, medical students and support personnel to provide medical care and surgical intervention.

Director of Strategic Initiatives (2001-2002)
Planned and facilitated an 18-month program of events to commemorate the institution's 150th anniversary, reporting to the Executive Vice President and Board of Trustees. Highlights included: opening ceremony with Mayor Bloomberg, archival retrospective at the Museum of the City of New York, Community Health Day with free medical services and screenings to over 1500 attendees, and an Academic Symposium.

- Secured $800,000 in corporate sponsorship across the portfolio of events.
- Oversaw contract and filming of documentary video on the history of Mount Sinai Hospital.

Assistant Dean for Clinical Trial Operations and Business Development (1998-2001)
Designed and managed the Office of Clinical Trials, a key organizational and operational component advancing clinical trials and translational research. Developed strategic and operational plans. Prepared grant applications and negotiated contracts for extramural funding involving federal, industry and non-profit foundations. Attended and organized research conferences. Strengthened overall regulatory compliance and fiscal compliance with third party payors.

- Launched marketing and outreach program to convey Mount Sinai's research capabilities.
- Attracted ongoing contracts and right of first refusal in phase III clinical trials (with appropriate patient populations) from Novartis, Smith Kline Beecham, Amgen and Pfizer and other pharmaceutical and biotech companies.
- Grew research revenue by increasing clinical trials by 40% and patient enrollment by over 320% annually, from inception.
- Secured $12 million in NIH funding and ran program on a $1.5 million budget resulting in a profitable operation beginning in year one.
- Facilitated contracts between faculty and pharmaceutical and biotech companies resulting in $500,000 in funding for investigator initiated research projects.
- Awarded $5 million grant from the American Heart Association after reworking a previously submitted and unfunded NIH grant proposal.
- Attained membership in the Smith Kline Beecham Development Partners Program and $500,000 annually in upfront quarterly payments to ensure consistency among clinical research coordinators.

- Created a career track for Clinical Research Coordinators including ACRP certification opportunities and a centralized pool of research coordinators to promote steady workflow.
- Commissioned an audit of 10% of all active human subject trials for compliance to good clinical practice (GCP). Collaborated with Institutional Review Board (IRB) Chair when remediation was necessary. Sat on the IRB.
- Invited by Dr. Barry Coller, Chairman of Medicine to join the Patient Oriented Research Committee.

Senior Research Coordinator, Division of Infectious Diseases (1993-1997)
Collaborated with the Hospital Epidemiologist as Senior Research Coordinator on a CDC multicenter study to measure the effects of using safety devices to draw blood, start IV's and give injections to determine if they decreased needle sticks in healthcare workers. The device companies (Johnson and Johnson, Becton Dickenson and Vacupro) extended the study solely at The Mount Sinai Hospital. This research opened the multi-billion dollar market for safety devices.

- Hired study personnel. Arranged with Procurement to cancel contracts to remove non-study devices from the institution.
- Inserviced 3000 nurses and physicians on the use of the study devices.
- Audited sterilized used "sharps boxes" to assess implementation rates of the study devices.
- Included as a co-author of an article in *Morbidity and Mortality Weekly Report.*

Cornell University Medical College, New York, N.Y. (1986-1993)
Department of Child Development, Research Nurse Coordinator
Coordinated a National Institute of Mental Health funded pediatric study. Recruited subjects, developed tools to achieve stated research goals, implemented research protocols, analyzed data and prepared reports.

Department of Surgery Research Nurse Specialist
Established initial infrastructure and managed a multi-site surgical metabolism laboratory consisting of 4 full-time surgical fellows with an annual budget of $800,000. The study data and collaboration with Dr. Anthony Cerami's laboratory at Rockefeller University supported Dr. Kevin Tracy's seminal work on tumor necrosis factor (TNF). Implemented federally funded research protocols. Procured and managed supplies. Recruited control subjects and hospitalized inpatient subjects and obtained informed consent.

- Assisted physician researchers with all clinical and laboratory aspects of the study including processing, cataloging, transporting and storing specimens.
- Carried out study protocols for ICU inpatient and outpatient human subjects.
- Prepared submissions for IRB approval to assure patient safety and regulatory compliance.

Critical Care Registered Nurse, New York, N.Y. (1988-1997)
Neonatal ICU nurse at Mount Sinai, New York Hospital, Lenox Hill Hospital, Columbia-Presbyterian Hospital, contracted through a hospital staffing agency.

New York Hospital, New York, N.Y. (1979-1988)
Registered Nurse, Neonatal ICU. Delivered care to critically ill infants. Supervised ancillary personnel, assigned nursing staff, maintained detailed written records and electronic medical records for all patients.

EDUCATION
University of Massachusetts, Amherst, Massachusetts. M.S. Exercise Physiology, *Cum Laude,* 1987.

University of Florida, Gainesville, Florida. B.S. Nursing, *Cum Laude*, 1979.

How Alan Turing invented the computer, helped win World War II and left us with one of the greatest puzzles of our time: are humans simply computers or are we more than that? Many scientists think we have a tenuous hold on the title, "most intelligent being on the planet". They think it's just a matter of time before computers become smarter than us, and then what? This book charts a journey through the science of information, from the origins of language and logic, to the frontiers of modern physics. From Lewis Carroll's logic puzzles, through Alan Turing and his work on Enigma and the imitation game, to John Bell's inequality, and finally the Conway-Kochen 'Free Will' Theorem. How do the laws of physics give us our creativity, our rich experience of communication and, especially, our free will?

Can a computer win the imitation game and pass the Turing Test?
Why do creative people make better mates than rich people?
Why are humans bad at mathematics, yet so creative?
Could an infinite number of monkeys write Hamlet?
Is our brain a quantum computer?
Is free will an illusion?

James Tagg is an inventor and entrepreneur. A pioneer of touchscreen technology, he has founded several companies, including Truphone, the world's first global mobile network. He holds numerous patents, filed in over a hundred countries. He studied Physics and Computer Science at Manchester University, Design at Lancaster University and Engineering at Cambridge University. He lives with his family on a farm in Kent, England.
www.jamestagg.com

"I can't tell you when the last time was that I had this much fun reading and using my brain. From the very beginning, James Tagg had me hooked with the premise; the question of whether or not humans are the most intelligent beings on the planet...."

Janet, Netgalley

"This is a fantastic book. It seams together cutting edge neuroscience, psychology, thought experiments, artificial intelligence/ machine learning, mathematics and even some history!..."

PFJ H., Amazon

"Hard work to read, but makes you think about the nature of human intelligence and AI..."

Brian Clegg, Popular Science

"This is a fat book that covers a huge amount of ground. James' topic is primarily the brain and how we think, but there is a running theme contrasting the human brain with computers. His thesis is that computers can never think like humans (for example, that they can never be truly creative) and he explores many fields from philosophy and logic to mathematics in pursuit of this proof...."

R. Hanbury, Amazon

If you have enjoyed reading this book please leave a review and if you would like to hear more, or come to one of my talks, please join the mailing list at: www.jamestagg.com/updates.

Are the Androids Dreaming Yet?

Amazing Brain.
Human
Communication,
Creativity &
Free Will.

Are the Androids Dreaming Yet?

Amazing Brain.
Human
Communication,
Creativity &
Free Will.

JAMES TAGG

Hurst Farm Books
An Imprint of
Hurst Farm Enterprises

Published by Hurst Farm Books

Hurst Farm, Dairy Lane, Crockham Hill, TN8 6RA.
+44 1732 807246
12 Williams Road, Chatham, NJ 07928
+1 646 355 1250
www.jamestagg.com
bookinfo@jamestagg.com

A catalogue record for this book is available from the British Library.

Publisher's Cataloging-in-Publication Data

Tagg, James, 1964-
Are the Androids Dreaming Yet?: Amazing Brain.
Human Communication, Creativity & Free Will. /
James Tagg.
pages cm

Includes bibliographical reference and index
ISBN: 978-1-910464-03-8 (softcover)
ISBN: 978-1-910464-01-4 (ebook)

1. Creative ability. 2. Communication—Social aspects.
3. Technology—Social aspects. 4. Mind and body. 5.
Computers and civilization. I. Title.
T174 .T24 2015
303.48`34—dc23

Library of Congress Control Number: 2014945686
(hardback)

p2 220415 postc

To my family,
who have patiently listened to my interminable
ramblings about 'Elephantine' Equations.

PREFACE

ACPMM, Wolfson College, Cambridge

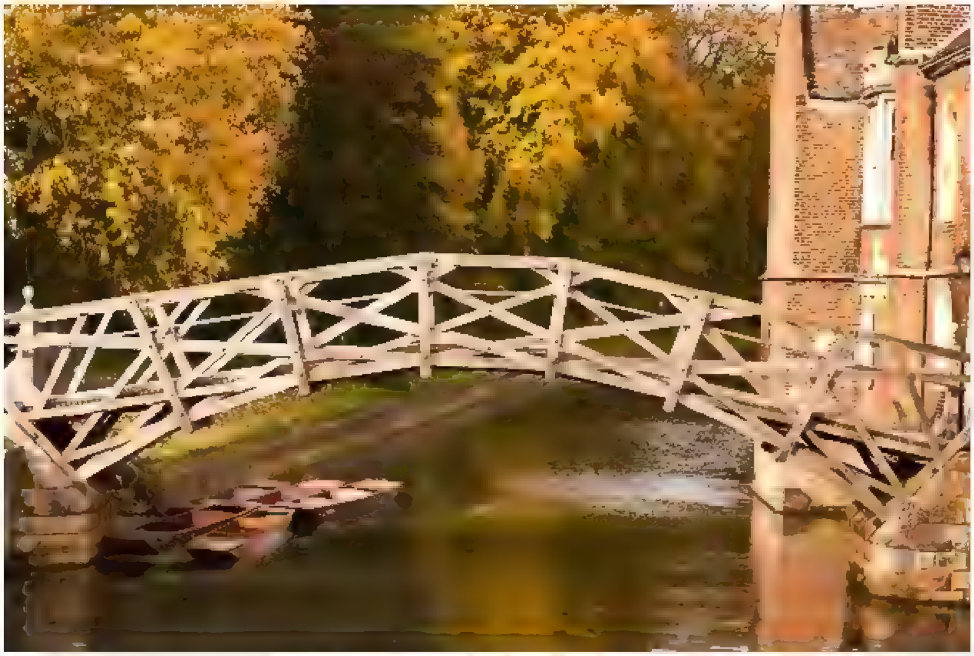
"A man may have twenty years of experience, or one year of experience twenty times."

Mike Sharman

"Rules are for the obedience of fools and the guidance of wise men."

Douglas Bader

I am an inventor. I've always been an inventor. Ever since childhood I've tinkered with electronics and computers, taking things apart and putting them back together. There is no academic course for inventing, so I had to choose my own path through school and University. I studied design, physics and mathematics at secondary school, and engineering and management at University. Part of that time was spent in the Engineering Department of Cambridge University on a particularly special course.



Mathematical Bridge, Cambridge

Every autumn about thirty graduate students arrive at the Engineering Department in Cambridge to join the Advanced Course in Design, Manufacturing and Management. They expect to spend the year walking among the city's hallowed spires, attending lectures, bumping into Stephen Hawking and punting on the River Cam.

Instead, they get quite a shock!

In 1989, I joined the course. There were twenty-six engineers, a psychologist and a physicist – me. There was no prescribed syllabus; instead the course used learning-by-experience and lectures from the experts in a given field. To study advertising, you might visit a top London agency, for shipbuilding a shipyard on the Clyde. If you were unlucky enough to find these two lectures scheduled for the same week, you had to travel the length of Britain. The course runs a half dozen minibuses to solve this transport problem. Every four weeks we would undertake a project in a different company. I remember designing pit props for coal mines and imaging software for a weaving company. At the end of each project we presented our findings to each other and, with eight projects and thirty students, this made for a great many presentations. To keep the process manageable, the course put great store in teaching us the art of communication.

These days I design large complex systems, and clear communication is extremely important. My ideas are often turned into working products and, if those products have flaws, a post-mortem usually shows the cause was a breakdown in communication. Of course, this may be a purely personal failing, but when I talk to people in other companies they report the same problem. It seems we all find communication difficult.

have wondered for many years why it is called the 'art of communication'. Surely it's a science, governed by bits, bytes and bandwidth. That might be true of the symbols in an email – they are clearly encoded symbolically – but is the understanding in our brains simply encoded by symbols? What is the physics that underlies human understanding?

Each summer I go on holiday to escape engineering for a couple of weeks. While away I indulge my passion for reading books by the likes of Douglas Hofstadter, David Deutsch and Stephen Hawking. One book that struck me years ago was Roger Penrose's *The Emperor's New Mind*. In it, he tackles the question of what happens in the human brain when we understand something. He extends an idea put forward by J.R. Lucas of Oxford University that minds must be more powerful than computers because they do something computers cannot: namely to step beyond mere rules and see truth. Colloquially we call this 'common sense' or 'stepping outside the box'.

The Lucas argument uses the theories of Gödel and Turing to show computer algorithms have limitations. Some things are simply not computable. Computers can do many useful things, but they cannot discover new mathematical theorems, such as a proof of Fermat's Last Theorem. In 1996, Andrew Wiles succeeded in finding a solution to this problem. This presents a paradox, solved only if we conclude Andrew Wiles is not a computer. Indeed, since most mathematicians discover at least one theorem during their lives, we must conclude no mathematician is a computer! This is controversial. Most philosophers tend to the view put forward by Daniel Dennett that the Universe is an entirely determined place and any personal sense of free will and creativity is an illusion. In Dennett's worldview, Andrew Wiles is a special purpose machine that was always destined to solve Fermat's Last Theorem. I believe this model is flawed. It is my aim in this book to show you why. Indeed I am going to go further and argue all human creativity is non-computational; art, communication, understanding – all are based on non-algorithmic principles.

If you consider creative thinking deeply enough you're inevitably drawn into the question of whether we have free will. When I get to work each morning, the first thing I do – after a cup of coffee, obviously – is choose which creative task to tackle first. I feel this choice is freely made, but the determined determinists assure me I am wrong and my decision was already made. As Daniel Dennett says, "You have no free will. Get over it!" They say I am effectively an avatar in some giant cosmic computer game, going about my business in an entirely predefined way. I do not agree! If they are right all the coincidences and chance actions of my life were fixed at the time of the Big Bang. I feel this must be wrong, but finding a chink in the determinist armor is hard work; the laws of physics as we know them today are almost exclusively deterministic. This book lays out the options – the chinks – that would allow free will to enter our Universe.

To understand human thinking we would really like to look inside a working human brain. We can't do this yet. All we can do is observe minds at work when they communicate with one another. If our minds think non computationally as I believe we should be able to see them struggle when they have to translate thoughts into symbolic form. The more symbolic, the harder it will be. This is indeed what we observe: face-to-face communication is easy, while formal written modes are much harder. We will explore the difference between human and computer communication as our first step in locating the weakness in the armor of determinism.

What do I Believe?

As a scientist, I ought not to have beliefs. I should have theories and working assumptions. But, as a human being, I must admit believing certain things are true. Science does not forbid beliefs. It just requires you to be prepared to have one overturned if a better one comes along. Richard Feynman summed this up in a lecture he delivered at Cal Tech: "If you want to discover a theorem," he said, "first, you guess, then you work out some effect predicted by the theorem. Finally, you see if the effect happens in the real world. If it does, you have a good theory. If the effect happens a little differently, you will need to look for a better theory." Here are some of my overturn-able beliefs.

Beliefs

- We have true free will. We consciously decide our actions and these decisions are in no way predetermined. We shape the future. Allowing for free will is, therefore, a boundary condition for any theory of our Universe.

- The world is an amazing place, but understandable. We can understand the Universe through the application of thought and reason.

- There is only one Universe and it appears to make sense.

- Humans think creatively, computers do not.

- The process of understanding and communication is complex, much more complex than the digital theorems of Claude Shannon and Harry Nyquist.

- Understanding is hard.

- The communication of understanding is even harder.

CONTENTS

"It is no good getting furious if you get stuck. What I do is keep thinking about the problem but work on something else. Sometimes it is years before I see the way forward. In the case of information loss and black holes, it was 29 years."

Stephen Hawking

Introduction

EXPERIMENTS, MULTIMEDIA AND PUZZLES

Throughout this book you will come across experiments to try, multimedia references to track down, and puzzles to solve. You can get additional information at www.jamestagg.com/understanding.

If you undertake an experiment I would appreciate your leaving a note of your results on the website and making useful comments on the blog.

Most of the experiments and puzzles are quick and simple. The puzzles I have set often benefit from creative thinking. I have made finding the answers to these problems a little hard, so you are not tempted to cheat. I want you to try to solve the problems and 'feel' your brain working.

This book argues that intuitive thought solves problems in a different way to analytical thought. The process takes time and often benefits from putting a problem to one side while you use your mind to process foreground tasks. I hope you read this book at a time when the website is not available – or at least don't peek. Give your intuitive thought processes time to work.

Graham Wallas described the process of creative thinking in 1926 and I think it is still one of the best models we have:

> First you must prepare and become fully acquainted with the problem. It might seem impossible but don't despair, just commit to it. Next, you should leave the problem to stew – incubation, he called it. After a while, you will feel a solution is at hand. You don't quite have it yet but you are

sure you will. This is intimation. Finally, some inspiration or insight will pop into your head – this is the Eureka moment. Now you have a solution but intuitive thinking is far from infallible. You will need to check the solution and may find your answer wrong the first few times. Persevere; you will get there in the end.

As a warm-up exercise, let me give you a simple childhood riddle to solve.

A man lives on the twentieth floor of a skyscraper with an old elevator. Each morning he gets into the elevator and goes down to the ground floor, but each evening he gets into the elevator, travels up to the tenth floor, gets out, and walks the rest of the way. Why?

ANSWER IN YOUR OWN TIME

Chapter 1

MIND OVER COMPUTER

Computer versus Human

"I visualize a time when we will be to robots what dogs are to humans, and I'm rooting for the machines."

Claude Shannon

"The question of whether computers can think is just like the question of whether submarines can swim."

Edgar Dijkstra

Kasparov versus Deep Blue

"The Three Laws of Robotics:

1. *A robot may not injure a human being or, through inaction, allow a human being to come to harm;*
2. *A robot must obey the orders given it by humanbeings except where such orders would conflict with the First Law;*
3. *A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. The Zeroth Law: A robot may not harm humanity, or, by inaction, allow humanity to come to harm."*

Isaac Asimov, *I, Robot*

Deep Blue

It is 1997 and we are on the 39th story of the Equitable Center in New York, watching a chess match. It's no ordinary match. Two men sit opposite each other. One, a neatly suited figure, stares intently at the board. You can almost see the heat rising from his head as he processes the possibilities before him. The other, sits implacably calm and, before each turn, looks to a screen at the side of the board, reads the instruction, and makes his move.

This is the famous match between Garry Kasparov and IBM's Deep Blue. Kasparov, a child prodigy, became world chess champion at the age of fifteen and, to this day, holds the record for the highest chess ranking ever achieved. Some consider him one of the most intelligent people on the planet. His opponent, Deep Blue, is a massively parallel chess-playing computer built by IBM's Watson Research Laboratory. The machine itself sits a few blocks north of the tournament in an air-conditioned room, and relays the moves over a phone line to Joe Hoane, the IBM researcher who moves the pieces.

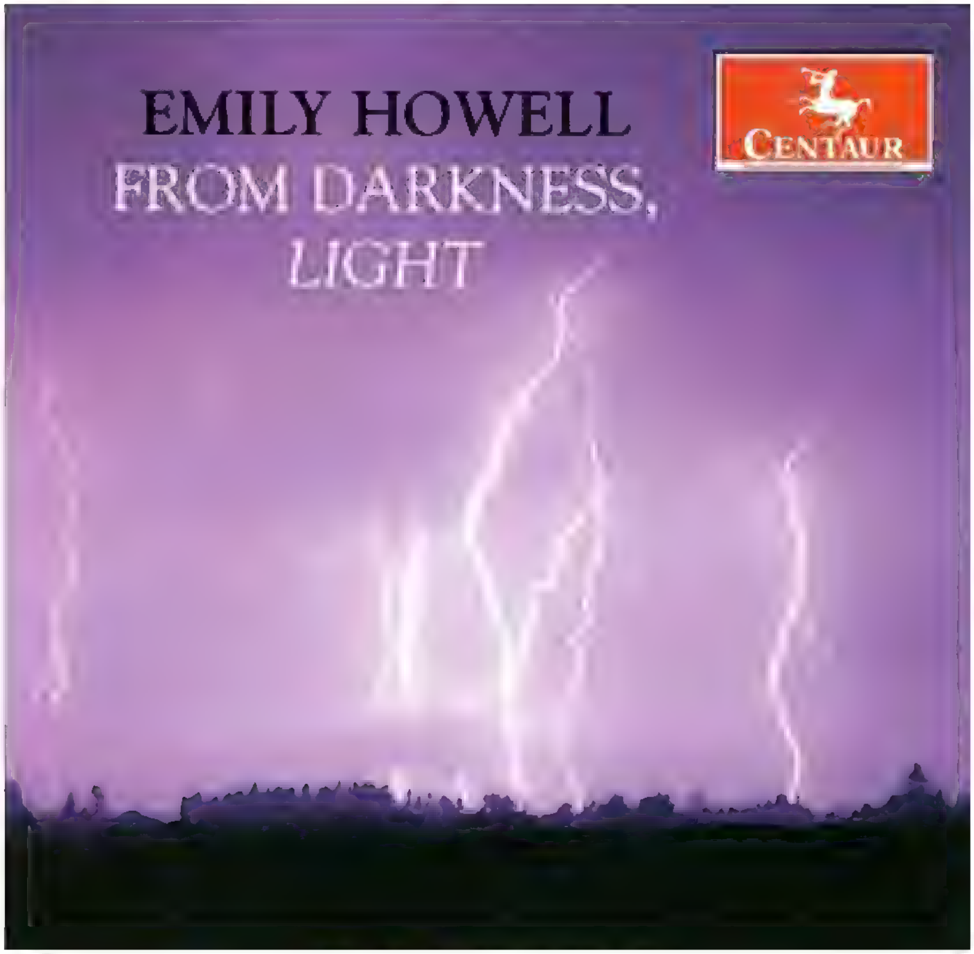
Six months earlier, in Philadelphia, Kasparov won against Deep Blue. This is the rematch and has generated a worldwide media frenzy. Tickets to the event are sold out and most news organizations give a blow-by-blow report each day. On the eighth day of the tournament Kasparov and Deep Blue are level pegging. Kasparov is playing an opening he knows well. It's one designed to be hard for computers to play and has been tested extensively against Fritz, a chess computer Grand Masters use for practice. But Deep Blue doesn't seem fazed. Kasparov is visibly tired. On the 16th move he makes a dreadful blunder and sinks into despair. An hour later, after some moments of quiet contemplation, he tips over his king, gets up, and leaves the room. Kasparov has resigned, Deep Blue has beaten him 3½ to 2½ points and is now the most powerful chess player on the planet.

Later, when interviewed about his experience, Kasparov thought Deep Blue must have been assisted by humans during the games because the program appeared to play intuitively. The rules of the tournament allowed humans to work on the program between matches, but not during actual play. The argument has never been settled, and Deep Blue was long ago dismantled. These days chess players avoid big public matches against computers, arguing it is really a different sort of game. A computer's ability to crunch mathematically through all the many possibilities means a chess player must play without error against a machine, but can play a more interesting and fluid match against a fellow human.

Chess is computer-friendly because it is a finite problem. You always win, lose or draw. The game can't go on forever because any position that repeats itself more than three times is declared a draw, and if a player makes 50 moves without moving a pawn or taking a piece, the game is also declared a draw. In a typical game, each player makes 40 moves, and on each turn you can choose from 30 possible moves. Although this equates to a huge number of options, it is still a finite number.

It is possible, therefore, to create a perfect chess-playing machine. Such a machine would project any position it encountered through every permutation to the endgame. But, although chess is solvable using brute force this might not be practical in our Universe. The storage required to hold all the possible positions being analyzed would be vast – needing most of the atoms in the Universe. You would need to pack this information into a small enough space to allow fast retrieval in order to play the first 40 moves in two hours. This would require storing all the information within a sphere no larger than three light minutes. Putting that much data in such a small space would exceed the Hawking Bekenstein bound – a limit on the information carrying capacity of space-time put forward by Stephen Hawking and Jacob Bekenstein – causing the region of space-time to collapse to a black hole! Despite these minor technical problems, an ingenious algorithm could be made that was unbeatable: chess is essentially computable.

The term algorithm will often arise in the book, so it is worth giving a little history. The word comes from the name of an 8th Century Persian mathematician, Al-Khwarizmi, and means a step-by-step procedure. We use one whenever we do long division or look up a phone number on



The Music of Emily Howell

our mobile phone. It is any mechanical procedure you perform without thinking about it. Computers are always executing an algorithm; that's what they do.

Fast forward to 2010 and Centaur Records releases a new classical music CD featuring the piano music of Emily Howell. Critics are enthusiastic about the new talent. She has composed music in a broad range of classical and contemporary styles. You can find some examples on my website.

But, it transpires, Emily is a computer, the brainchild of David Cope from the University of Santa Cruz. On hearing this news critics revise their opinion of the compositions – "repetitive and formulaic," "not real music," "pastiche". Listen again to the music and see whether you have changed your opinion. Whatever you think, Emily has made a good attempt at composing in the style of several great composers: J.S. Bach and Franz Liszt, as well as modern ones such as Stockhausen and

Philip Glass. The compositions would get a reasonable technical score in an exam, better than many of my attempts, but are these compositions truly art?

There's no question computers are gaining ground on us in certain mathematically oriented tasks   playing chess, musical composition, and various modeling tasks. But attempts to have them work with words and ideas have generally produced dismal results. Until now.

In 2008, IBM unveiled Watson: a computer capable of answering general knowledge questions. Watson has an enormous database of human knowledge: the Encyclopedia Britannica, a billion web pages, the entire text of Wikipedia and millions of books. It uses artificial intelligence to trawl through this vast reservoir of knowledge and answer questions using a statistical approach. In 2011, Watson featured as a contestant on Jeopardy, the American quiz show, where it beat the two record-holding contestants – the one with the highest number of wins and the one with most consecutive wins. Let me give you a few sample questions and see how you fare.

Question 1.	It can mean to develop gradually in the mind or to carry during pregnancy.
Question 2.	William Wilkinson's "An Account of the Principalities of Wallachia and Moldavia" inspired this author's most famous Novel.
Question 3.	Its largest airport is named for a World War II hero; its second largest, for a World War II battle.

Watson answered questions one and two correctly but failed on question three. You can probably see the final question is posed in poorly structured English and this threw off Watson's comprehension algorithm.

IBM's Watson Plays Jeopardy

Ignoring the odd hiccup, Watson is much better at Jeopardy than a human. Should humans be worried? First chess, then music, now general knowledge, will all human endeavors succumb to a computer? What will be our purpose on the planet if this happens?

Steve Wozniak

"Machines will run the world,
humans will become idle pets."

Steve Wozniak

Man v Machine

Are humans advanced computers with a temporary hold on the title, 'most intelligent being on the planet,' or are we fundamentally different?

We are extraordinarily creative, but we can't add up as well as a cheap pocket calculator. We have poor memories, but we can use common sense to solve problems we have never seen before. Our communication skills are woefully imprecise, but we can tell jokes that send our fellow humans into paroxysms of laughter. We might conclude humans are not computers, but the scientific consensus is that brains are 'wet computers'. I don't agree with this and I'm going to set out the argument to show why man is not a computing machine.

There is an urban legend we think with only 10% of our brains. This is not true. Science has mapped the vast majority of the human brain using two methods. The first, an amazing set of noninvasive imaging techniques, allows us to 'see' the brain as it thinks. The second is more macabre: with seven billion humans on the planet, enough accidents occur through sports injuries, car crashes and surgical mistakes to provide a large enough sample to conduct research. Questioning patients with brain-damage allows us to work out what the injured part did before the accident.

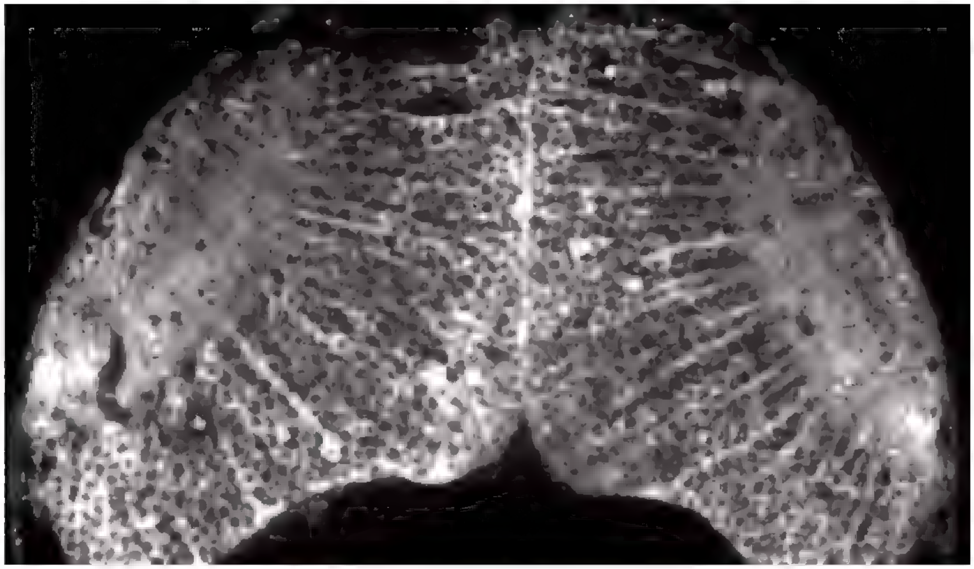
One famous patient had an accident where the blade of a toy sword went up his nose and damaged a small part of his amygdala and hippocampus, the area of the brain responsible for storing memory. This rendered the man unable to lay down permanent memories after the accident. Events before the accident remained clear but he could not memorize new information. You could tell a joke and he would find it

Turning Images to Music

funny and laugh uproariously. A few minutes later, you could tell the same joke and he would find it just as funny as the first time. For him, every time was the first time, because he had lost the ability to record long-term memories. The syndrome is wonderfully depicted in the film *50 First Dates* starring Adam Sandler and Drew Barrymore. Another patient with specific stroke damage was unable to recall the names of fruits but, oddly, could still name vegetables. Interestingly tomatoes presented a particular problem. He had probably never known how to catalogue them so they were partially remembered in both areas.

There are many such medical cases. In Oliver Sachs' *The Man who Mistook his Wife for a Hat*, the author relates the tale of a man with visual agnosia who could not reliably name familiar objects, including his own wife! He had a perfectly loving relationship with her but simply could not name her from a picture. Sachs, Professor of Neurology at New York University School of Medicine, provides many such fascinating stories, along with their medical backgrounds.

The fruit and vegetable case suggests our brains are organized like a filing cabinet. When we damage a part of the brain, it's like losing a drawer: All the information stored in that drawer is lost. Quite a few experiments contradict this model and indicate many tasks are distributed around the brain. The curious case of blindsight is one such example. People with a damaged visual cortex can often recognize objects despite reporting they have no sensation of vision. Show them a shape and they will report they can see nothing. Ask them to name the shape and they might even get a little irritated by the question; they are blind after all. But, ask them to guess the shape and they will get it right every time. Seeing is more



Brain Image of Fish Hunting Prey

widely distributed in the brain than was first thought. Conscious seeing is based in the visual cortex, but there are older pathways still active in the brain that facilitate this unconscious seeing.

The brain is very plastic. Lose your sight through damage to the eye or optic nerve, and the brain can repurpose the visual cortex to other uses such as processing sound or touch. Daniel Kish has developed this to such a high level that he can ride a bicycle despite being blind. He clicks his tongue against the roof of his mouth and uses echolocation to form an auditory model of the world around him. Using a similar approach, Amir Amedi from the Hebrew University of Jerusalem has built an audio imager that turns pictures of the world into musical sound patterns. CAT scans of people using this system show they use the visual cortex to convert these sound images into models of the world in similar parts of the brain to a sighted person.

We now know roughly *what* each part of the brain does, but we have no idea *how* it does it. The scale of an individual thought is too small to see in a brain scan. All we can do is observe large-scale electrical activity associated with those thoughts. A video, from a group at Tokyo University, shows an example of electrical activity filmed in real time as a fish hunts for its prey. Fish have transparent bodies and thin skulls facilitating this sort of imaging. Humans are much harder subjects to work with!

The most popular theory to explain how brains work is as some form of computer. Computers are easy to study because we manufacture them. They tend to crash quite frequently – usually at the most inconvenient moments – so we have packed them with diagnostic monitoring systems. These systems allow us to watch a computer think and, since they think symbolically, we can easily read their minds.

Unfortunately computers don't display many human-like thoughts. They don't laugh and cry, they don't report consciousness and they don't appear to exercise free will or display creative impulses. This is frustrating because these are the thoughts we would most like to study. It might be that computers are not yet powerful enough, and in another few years they will be giving Mozart a run for his money. But there may also be a fundamental difference which renders them incapable of this sort of thinking. This is the crux of the modern scientific debate: do humans think differently?

Computer Brains

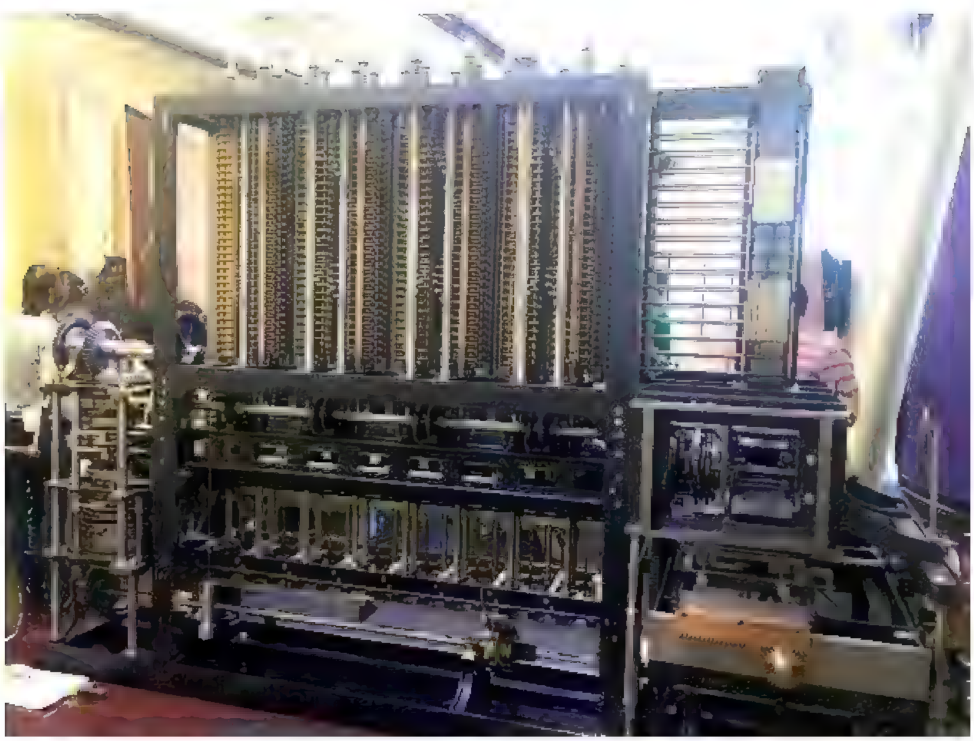
On the face of it, humans and computers *behave* very differently. Our memories are poor, but we understand things. We are creative, but bad at mathematics. We learn by example, computers are programmed. We are emotional, impulsive and appear to have free will. Computers are ordered, predictable, but lack common sense. Both humans and computers appear to be physical, discrete systems. We both take inputs, generate outputs and are capable of solving similar problems. Indeed, each time we examine a problem solved by humans we usually find we can automate it. This is known as 'knowledge engineering' and there are many examples; from aerospace to finance, and architecture to medicine.

An example of where computers excel is in medical diagnosis. ISABEL is a clinical diagnosis program designed to help ER doctors quickly diagnose critical patients. It was created by the parents of Isabel Maude, a little girl who presented with multiple symptoms to an ER unit. Doctors were initially confused by the symptoms and misdiagnosed her condition. She was later diagnosed with meningitis. Isabel suffered multiple organ failure but survived. Her parents realized there was something wrong with the ER triage process. They got together with some computer scientists and built the expert system 'ISABEL'. When ER doctors are presented with symptoms, they must mentally scan a vast array of literature to rule in and out possible diagnoses. The problem-solving process is not linear; if you've ever watched the TV series *House* it gives a great dramatization of the process. Certain symptoms might suggest a diagnosis but are not conclusive, and there are many paths to explore. Programmers have taken the heuristic rules from many doctors and codified them into software. ISABEL allows a doctor to input a set of symptoms and it will spit out a range of possible alternative diagnoses with probability weightings and suggested further tests. Similar systems are widely deployed in other fields, to build racing cars, design dams and fight crime. Even the game consoles in our living rooms implement artificial intelligence to make the aliens more believable and our hearts pump faster.

Origin of Computers

Alan Turing effectively invented the modern day computer in a paper he submitted to the London Mathematical Society in the summer of 1936. He was not the first person to come up with the idea – that honor probably goes to Charles Babbage – but he was the first to fully understand its power. When we talk about computers today we mean machines, but it is worth noting computers in Turing's time were more often humans using pencil and paper. The mechanical computers before Turing were elementary at best.

Rudimentary calculating machines were developed in Greece, Persia and China as far back as the Ming Dynasty. An astrolabe recovered from a ship wreck off the Greek Island of Antikythera had cogs and gears and could accurately predict the motions of the sun and planets. Many



Babbage's Difference Engine No. 2, Computer History Museum, CA of these skills were lost in the Dark Ages but, once the Renaissance was underway in the 16^{th} and 17^{th} centuries, complex mechanical clocks were devised that were capable of predicting the motions of the planets to a high degree of precision. Mechanical, hand-cranked calculators appeared in the mid 18^{th} century, and in 1886 Charles Babbage conceived the first programmable computing machine, The Analytical Engine. It was designed to read programs from cards, and used cogs and wheels to perform the calculations. His first machine – The Difference Engine – was designed to help the Admiralty calculate tide tables, but Babbage realized he could generalize it to compute almost any function. He ran out of money to complete any of his machines, but in the 20th century a dedicated band of enthusiasts built a working model of Difference Engine No.2. One copy sits in the London Science Museum and another in the Computer History Museum in California. These difference machines are not Turing complete and his Analytical Engine has never been built.

19^{th} Century Calculators

Model of the Antikythera Mechanism

In 1935, Turing was made a Fellow of King's College, Cambridge, and became interested in whether mathematical proofs could be found automatically. He wanted to know whether solving a mathematical puzzle was simply a matter of working through all the possibilities in a methodical manner, or whether something more subtle was required. Although chess is a fantastically complex game, it is finite, a big enough, fast enough computer can play the perfect game. Is this the case with discovering knowledge? Could a big enough, fast enough computer calculate all the knowledge in the Universe? Is Douglas Adams' fabled computer Deep Thought a possibility, able to calculate the answer to the ultimate question of 'life, the Universe and everything', albeit with a more enlightening answer than 42?

Turing boiled down the process of pencil and paper computation to a systematic program – a computer program. He proposed a thought experiment where he would run every possible program and see if such a procedure would yield the solution to every imaginable mathematical problem. He was able to show this would lead to a paradox and concluded the universal problem solver could not exist. His discovery is one of the most important of the 20th century – in the same league as relativity and quantum mechanics – and I will use it as my main tool in trying to explain the difference between brains and computers.

Although Turing's original paper was not intended as a blueprint for a practical device, he was one of those rare mathematicians who also liked to tinker with real world machines. The outbreak of the Second World War made the practical application of his work very important, and in Chapter 8 I will relate some of the code breaking stories that were to make him famous and caused Churchill to credit him with shortening the war by two years.

Calling Turing's work an 'invention' is probably the wrong term; 'discovery' might be more appropriate. Whatever you call it, people immediately equated human brains with computers. This is not new.

Each time a new advance in technology is made, people use it to explain the working of the brain. The ancient Greeks thought the brain was a fire consuming oxygen. When Alexander Graham Bell invented the telephone, the nervous system resembled a maze of wires and the brain an exchange. Brains were obviously a sophisticated telephone system. This idea has some potentially frightening consequences, particularly in light of the speed at which computers are improving.

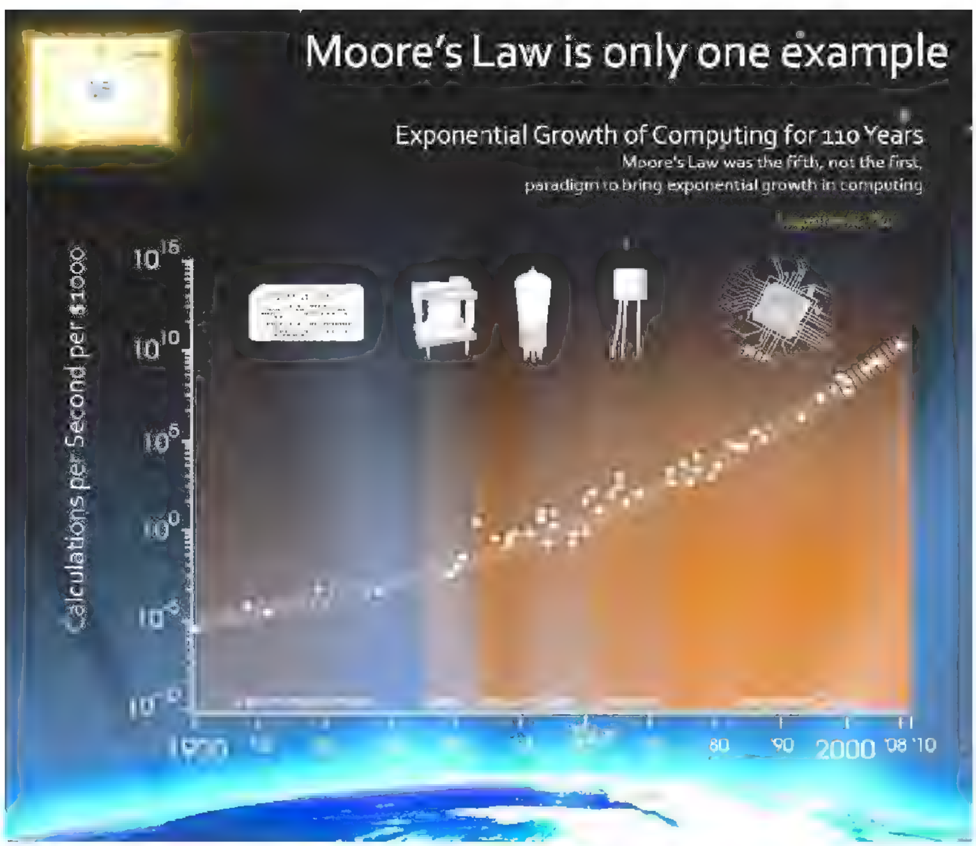
The most striking feature of computer technology is the rate of development. Cars travel faster than a person's legs will carry them, machines manufacture things faster than our hands are capable of working. If brains are computers, surely it is just a matter of time before they will think faster than humans. Turing predicted this would happen when computers reached the level of storing around 10 billion units of information. This happened some time in mid-2000. But today, in the year 2014, I can report that although my computer can beat me at chess, it still cannot fill out my expense report for me. So I am still ahead!

Maybe Turing just got the mathematics wrong. The human brain has about 10,000 times more neurons than our most powerful computers have logic gates. By this calculation, it's not a billion units of storage we need but, a trillion trillion units to put the computer on a par with a human brain. It's just a matter of time!

The worrying thing – especially for fans of the 'computers taking over the world' science fiction genre – is that computers are improving exponentially fast in line with Moore's Law, and the parity point is coming soon. Gordon Moore founded Intel with Andy Grove, and ran the engineering department there for more than 20 years. According to Moore's Law, the power of a computer doubles approximately every 18 months. The next significant event in the computer versus human competition is the gate count parity point   the moment when the number of logic gates and the number of neurons become equal. By my reckoning this will happen some time in 2053.

Don't despair. There may be a few dodges yet. The gate parity point assumes a logic gate and a neuron are equally powerful. However, some single cell organisms with only one neuron are capable of complex behaviors, such as hunting prey and avoiding obstacles. To perform these simple behaviors, a computer would need as many as 10,000 logic gates, about the complexity of my TV remote control. This gives us a bit more breathing space. The extra four orders of magnitude push the gate parity point out to around 2080, too late for me to see, but certainly within the bounds of some readers of this book.

To give you some idea of how Moore's Law works, the graph shows growth in computing power over time; the y-axis is a logarithmic plot using engineering notation. Because the growth is exponential we rapidly end up with very large numbers. Scientists use a special notation to cope with these large and small numbers. In scientific notation a number is written out in a compact form. For example, three hundred can be written as 3.0×10^{2}. To expand it back to a regular number you move the decimal point in 3.0 two spots to the right, making the number 300.0. A similar technique is used for small numbers. To expand 3.0×10^{2} move the decimal point 2 points to the left, giving 0.03. Why use scientific notation? Well, once the numbers get large they would no longer fit on a page! We can shorten the representation of numbers even further by dropping the '3.0 ×' part and just looking at the order of magnitude. The number 10^{80}, one with eighty zeroes after it, is the number of atoms in the Earth, and 10^{120} the number of particles in the known Universe. 10^{-43} meters is the 'plank number', believed to be the smallest dimension you can have, and 10^{100} is called a googol, named by Milton Sirotta, the



Moore's Law Extended by Ray Kurzweil

nephew of the famous American mathematician Edward Kasner, and subsequently the inspiration for the name 'Google', the Internet search engine.

Ray Kurzweil, the prolific inventor and futurologist, is fascinated by this exponential growth. Exponential curves grow slowly to start with but they pick up speed rapidly and, in the end, growth tends towards infinity. We are all painfully acquainted with one example of exponential growth: The common cold. Each infected cell in our body releases virus particles into the blood which infect further cells, leading to an exponential increase. This makes us feel rotten. Luckily our immune system can also respond exponentially, albeit somewhat delayed, so we survive. In the case of computer power there is no opposing immune system fighting back, so Kurzweil thinks computers will achieve almost limitless processing power; perhaps even within our lifetime. He thinks this will lead to some interesting consequences, for example, allowing people to live forever! Far-fetched? Follow his argument.

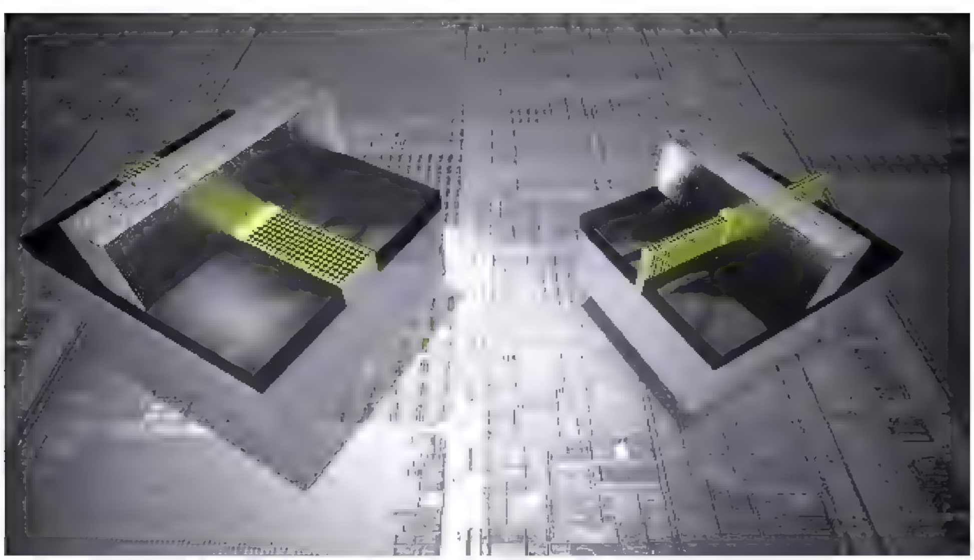
The two most important elements in keeping us alive are medical imaging, to see what is wrong; and genetic engineering, to fix those things. Both are improving in line with digital technology, doubling in power every 18 months. As computers get better at seeing into our bodies, and our ability to sequence and synthesize spare parts improves, we will reach a point where we can fix almost any problem. Kurzweil figures technology is improving and his body is decaying at just the right rate to mean by the time he needs heavy duty medical intervention it will be available. Barring a traffic accident or mad-axe-murderer, he should live forever. Even if his calculation is slightly off, the next generation will definitely have this option.

You might dismiss this as science fiction, but some amazing things are already happening. Recently a female patient in the USA suffering from bone cancer had her jaw replaced with a 3D printed component. Doctors were able to scan her head and take an image of the good side of her jaw, flip it right to left within the computer and repair any problems they saw. Then they sent the image to a 3D printer. The printer made a new jaw from tungsten powder, which was fused in a kiln. The final stage was to cover the metal part with an inert bone-like substance to give the human body a scaffolding on which to build real bone. They then performed the operation to remove her old jaw and replace it with the new one: result, brand new healthy jaw.

There are some practical limits to the power of computers on the horizon. Currently, the wires in a silicon chip are about twenty-two nanometers wide. That's around a thousandth of the width of a human hair, or approximately two hundred atoms wide. To match the complexity of a brain we will need to pack an order of ten million more gates into a silicon chip. One way to achieve this is to simply shrink the wires, but when we get down to around ten atoms wide, quantum effects begin to dominate. Signals in today's chips involve tens of thousands of electrons. We normally think of these electrons as a group, but in these tiny circuits we need to consider the behavior of each individual electron. Problems arise as this behavior is subject to quantum uncertainty. With only ten electrons there is a finite probability that none of them will be where you were expecting them to be. This causes problems for digital logic. You can't put a '1' in a memory location and be sure when you come to read it you will get a '1' back. You have to factor in the possibility of error.

Quantum effects can be annoying – requiring us to devise all manner of error checking hardware – but they can also be helpful. Richard Feynman proposed using quantum bits, 'qubits', to perform computation. Quantum computers can calculate many times faster than a classical computer because a single bit can represent more than one piece of information. Enterprising entrepreneurs are making use of this effect to build the next generation of devices, and you can already buy a 512 qubit computer from a Canadian company called D-Wave.

The biggest problem with building more powerful conventional chips is their area is reaching the manufacturing limit for economic viability. Silicon wafers contain random spots of damage and, as a chip gets larger, the chance it will have one of these spots approaches certainty. One solution is to use the third dimension and print the logic



3D Chip, Intel

gates so that they communicate in the vertical direction as well. Intel demonstrated the first three-dimensional chip in 2004, and these chips should begin to appear in our laptops by around 2020.

Taking a chip into the third dimension solves the economic problem but adding logic gates to a 3D chip presents a new problem – heat. Heat is generated in proportion to the volume of the chip but it can only be lost through the surface area. Result: the chip overheats. Large animals have the same problem which is why elephants have huge ears, filled with blood vessels, they can flap to cool themselves and really big mammals, such as whales, live in the ocean. The thermal problem is now the biggest problem in most computer designs. One data point suggests we could solve this problem, the human brain. We pack huge processing power into our skulls without overheating by using a variety of techniques, including folding the surface of the brain, running each neuron very slowly and maybe even using quantum mechanics. A very recent discovery is that brains could be using quantum effects to transmit signals. If true – and the research has only been recently published – it means we may use a form of high-temperature superconductivity to avoid overheating. More on this in Chapter 4.

Excluding exotic quantum effects, the main difference between computer and human brains is their processing architecture. Brains use slow, asynchronous logic to process information rather than the fast, synchronous type used in modern day computers. Logic gates in today's computers work all the time, even when there is nothing to do. For example, if I multiply 2 by 3 on my laptop the entire multiply circuit, designed to work on 20 digit numbers will still operate, and, even worse, it will operate on every tick of the master clock even if there is nothing to multiply. The brain, by contrast, works only as it needs; unused gates don't operate. This gives a massive reduction in unnecessary power consumption. We'd like to use this technique in modern computers but it is very difficult to implement. Tiny changes in timing cause completely different operation and this makes them hard to test. We accept this sort of problem in humans, calling it 'human error', but we count on computers to behave absolutely reliably, so full-blown asynchronous logic is not likely to appear anytime soon. Some of these ideas, however, have made their way into today's consumer devices. For example, the chips in the latest iPhone contain two CPUs: a small slow one operating when the phone is dormant in your pocket, and a large fast one that switches on when you need the power for a game or other high performance task.

All things considered there doesn't appear to be a fundamental engineering problem that would prevent computer hardware reaching the complexity of the human brain. But, even with all this hardware, it's not clear we would have something intelligent.

Richard Branson

"I was dyslexic, I had no understanding of schoolwork whatsoever. I certainly would have failed IQ tests. And it was one of the reasons I left school when I was 15 years old. And if I – if I'm not interested in something, I don't grasp it."

Richard Branson

Intelligence

Although my computer appears to have a mischievous sense of humor – deleting files and crashing just before important presentations – it does not appear to enjoy listening to Bach, nor does it ever ask me to put it near the window to admire the view. I think this means it's not intelligent, but what is 'intelligence'?

Intelligence is hard to define; I usually conclude someone is intelligent if I have an interesting conversation with them. In the workplace we typically hire people based on interview, rather than practical tests. Could a computer hold a conversation with me and persuade me it is intelligent? Meet Eliza.

Invented by Joseph Weizenbaum of the Massachusetts Institute of Technology, and named after George Bernard Shaw's character in *Pygmalion*, Eliza runs an algorithm pretending to be a psychiatrist offering a sympathetic ear to anyone who cares to talk. Try her out at my website or read the sample dialogue below!

User: Men are all alike.

Eliza: In what way?

USER: Men are all alike.

ELIZA: In what way ?

USER: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

USER: Well, my boyfriend made me come here.

ELIZA: Your boyfriend made you come here ?

USER: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

USER: It's true. I am unhappy.

ELIZA: Do you think coming here will help you not to be unhappy ?

USER: I need some help, that much seems certain.

ELIZA: What would it mean to you if you got some help ?

USER: Perhaps I could learn to get along with my mother.

ELIZA: Tell me more about your family.

USER: My mother takes care of me.

ELIZA: Who else in your family takes care of you ?

USER: My father.

ELIZA: Your father ?

USER: You are like my father in some ways.

ELIZA: What resemblence do you see ?

USER: You are not very aggressive but I think you don't want me to notice that.

ELIZA: What makes you think I am not very aggressive ?

USER: You don't argue with me.

ELIZA: Why do you think I don't argue with you ?

USER: You are afraid of me.

ELIZA: Does it please you to believe I am afraid of you ?

USER: My father is afraid of everybody.

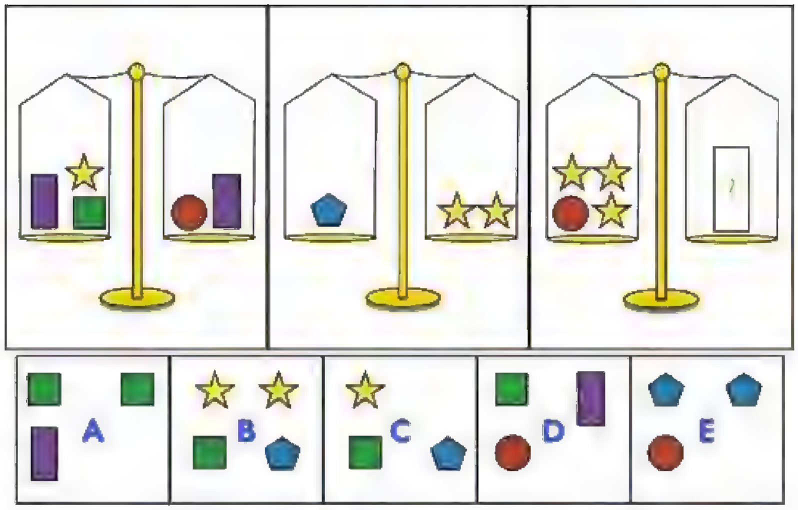
ELIZA: What else comes to your mind when you think of your father ?

USER: Bullies.

ELIZA: Does that have anything to do with the fact that your boyfriend made you come here ?

If you think Eliza a bit odd, you are right. Perhaps you won't hire her just yet. The program looks for keywords in speech, and parrots them back at you interspersed with a few stock phrases. If you type a sentence containing the word "job", the computer will say, "Tell me more about your job." It's a simple program and runs to only three pages of text. It shows how good a job you can do with very few resources and how far we are away from a real human being. More complex programs with a larger repertoire have been built and can hold a plausible conversation. People are occasionally taken in by them and are quite shocked when told they have been conversing with a machine. If you're on your guard, they are easy to spot. So far, no one has built a computer capable of holding an extended human conversation and fooling a human into thinking it's a person. Incidentally, having a sympathetic listener is so important to human beings these programs are used in psychotherapy and can be as effective as drugs.

Turing proposed a test to tell whether a computer had truly achieved human intelligence called the imitation game. His argument is as follows:



IQ Test

Humans are intelligent. (If you disagree with this premise then you're going to have a problem with this argument!) If you talk to a computer and cannot tell it from a human, it must also be intelligent: QED. The logic is sound but somehow feels wrong. It neatly, but irritatingly, sidesteps the whole problem of defining intelligence.

In 1912, William Stern devised a method for measuring intelligence in children. He named it 'IQ' from the German Intelligenz Quotient. You may have taken one of these tests at school. The tests consist of a series of abstract reasoning problems that minimize cultural references. For example, you might be asked to look at a set of blocks with dots on them and identify which is the odd one out. Numerous versions of the test have been developed over the years, but nowadays we mostly use one of three standard tests, Wechsler being the most common.

Measuring intelligence is complicated. Culture and language play a big part. If we take a tribe of Amazonian Indians and ask them to list the presidents of the United States, they will fail. That does not mean they're stupid. Drop me into the Amazon Rainforest and I will probably starve to death; they, on the other hand, can live off the land as hunter-gatherers with only a few hours work per day. Who is more intelligent?

One problem with IQ is that individual candidate scores can differ wildly from test to test, sometimes by as much as 20 points. That's huge. At the high end of the scale it can be the difference between being classified as smart or as a genius; and, at the low end, between being average or mentally subnormal. These variations don't usually matter and most universities and colleges take IQ with a pinch of salt, preferring more specific tests such as SATs in America, the Baccalaureate in Europe or A levels in the UK. IQ can be very important; and is sometimes a matter of life or death. In Atkins v. Virginia, the US Supreme Court found a person with mental disability, defined as having an IQ of less than 80, cannot be executed.

IQ is not really a measurement, in the normal sense. Most measurements in life are absolute, for example, distance, weight, and time. I can prove my house is bigger than yours using a tape measure. We each ensure our measures are the same by calibrating them against a common reference. In the 1900s we could have walked down to the local town hall and checked our measurements against a 'yardstick'. As measurements became standardized, these sticks were compared with a common central reference. For example, the *metre* was a platinum-iridium bar kept at the Pavillon de Breteuil near Paris. In the 1960s, a laser superseded the metal reference, and today a *metre* is defined as 1,650,763.73 wavelengths of the orange-red emission line in the electromagnetic spectrum of krypton-86 in a vacuum. Measurement has become very precise!

Intelligence is different. It has no yardstick. If I were to ask, "How much intelligence does it take to design a building?" there's no simple answer. IQ is not an absolute measurement – it's a relative score. Test 100 people and list their scores in order. The ones in the middle get a score of 100; the top 5 a score of at least 130 and the top person a score of 140. Similarly at the lower end. A person with a high IQ is probably smarter than one with a low IQ, but it doesn't tell you if the building they designed will stand up. It's rather like quoting the odds of a horse winning the Derby. Quoting the odds does not give the speed of the horse, nor often the winner of the race!

Despite attempts by test creators to remove cultural bias, it can never be completely eliminated. Certain Amazonian tribes have no concept of counting above five. For them, numbers are an alien idea and serve no useful purpose in their habitat. In the jungle there are always enough trees to make spears, and as a hunter-gatherer you simply need to know where to find your prey. There is no need to count animals into an enclosure at night. Another interesting environment is the Australian Outback. Aboriginal Australians appear to have a remarkable aptitude for visio-spatial memory and can remember maps or collections of objects much better than you or I. Tests for this skill involve playing a variant of Pelmanism. A collection of objects is placed on a tray and covered with a cloth. The cloth is lifted for 60 seconds to reveal the location and type of objects and then replaced. Subjects are then given a bucket full of objects and asked to recreate the tray. You and I do a modest job. Native Australians do this almost perfectly. Why?

In the vast, inhospitable Outback it is vitally important you remember that water can be found at the two rocks near the old gnarled tree. Forget this and you will die of thirst. It was once thought the skill evolved through natural selection, but this might not be the correct explanation. Recent studies show many of us can use mnemonic tricks to significantly improve our memory. Aboriginal skills might actually be learned and passed on from generation to generation.

IQ gives us a way to sum up intelligence using a single number but is this too simplistic? We all have friends who would be our first call if we met that special someone or lost our jobs. They are often not the smartest people we know, but they are highly empathetic. These people have 'social intelligence'. Other friends may fail academic tests yet demonstrate wonderful musical or artistic ability. They have creative intelligence. As we dig deeper, more talents emerge: sporting prowess, organizational brilliance, the ability to inspire loyalty. All these traits appear independently of academic brilliance.

During the last century, scientists worked hard to understand these different intelligence traits. The most influential theory came out of studies done at the United States Army Educational testing service, by Raymond Cattell and John Horn, and later added to by John Carroll. Their initials give the theory its name. CHC theory breaks down the general idea of intelligence into many different subgroups: 'G' factors.

If you are good at recalling all the kings and queens of England in chronological order, or can name every member of the 1966 English World Cup team or, perhaps, all the members of the baseball Hall of Fame, you would have high 'crystalized intelligence' – 'Gc'. It measures the sum total of all the *things* you have learned and retained in your long-term memory, your store of useful, and useless, facts. On the other hand there is innate intelligence, the sort that allows you to solve problems where tapping memory banks is not useful. My family often buy me puzzles for Christmas, the sort where you manipulate bits of bent metal that appear linked, but can be separated with a little ingenuity. These puzzles test our ability to work with problems we have never seen before and is called 'fluid intelligence' – 'Gf'.

Metal Puzzle

We can go further. A good tennis player will have high 'Gt' and 'Gv' scores: 't' for time and 'v' for vision, a good pub quiz contestant a high 'Glr' score – 'lr' denoting for long-term retrieval. Carol Vorderman, a UK game show presenter famous for mental arithmetic, would have a good 'Gq' score, 'q' for quantitative numerical skills. With all these types of intelligence to choose from it begs the question, "Is there a single master intelligence from which the rest follow?"

Political correctness plays a part here. It feels rather elitist to say smart people are good at everything. It is far nicer to think we each have our individual talents and some just have a few more than others. But that's not what the science tells us. 'Group Intelligence' – the overall G score – does appear to be the underlying cause of the other types of intelligence, and smart people do tend to be good all-rounders. However, there is one major flaw in the analysis; the studies only measure the subjects' ability to pass academic tests, they don't look at our success in real-life, nor our creativity.

Lewis Terman began the longest running study of intelligence and its relationship to life success back in the 1920s. It continues to this day. A group of 1500 children with high IQs were selected and tracked throughout their lives. Terman assumed their high IQs would result in them being very successful. They certainly did well, but studies show they did no better than if they had been chosen randomly from the same area (all the children came from around Stanford University). Famously two children, William Shockley and Luis Alvarez, tested too low to be chosen for the study but went on to win Nobel Prizes for Physics in 1956 and 1968, respectively.

There are many similar anecdotes: Apparently stupid people go on to great things. Einstein's teacher famously stated he would never amount to anything and Sir John Gurdon's school report said he was 'too stupid' for science. He went on to discover monoclonal antibodies for which he was awarded a Nobel Prize! Scientists have now devised the alternative theory of an intelligence tidemark. Once above this level – an IQ of about 130 – you can pretty much do anything you want to. This might be because one very important type of intelligence – creative intelligence – is not highly correlated with the rest. Creative people tend to be sufficiently intelligent for their field but once above that threshold the relationship breaks down. Success in creative endeavors seems to reflect strength of character and creative aptitude rather than raw brainpower.

Physical Basis of Intelligence

The high correlation between different sorts of academic intelligence suggests we might find a physical process within the brain leading to high IQ. Functional MRI scans show intelligent people use more neurons when tackling a given mental task, perhaps bringing to bear greater raw horsepower, but this is not really an explanation. It is akin to saying Usain Bolt runs faster because he gets more power to his legs. This is obvious. What we want to know is *how*.

The problem with looking at brains for a common cause is the variation from brain to brain. We all have different genes and life experiences. On top of this, we really only see brains post mortem and this tends to confound comparisons of brain structure. One way to minimize the variation is to use separated identical twins. Twins have identical genes so their fundamental hardware is the same. We should be able to see features of the brain that are common to smart sets of twins but absent in less smart pairs. If a feature is not shared it can be discounted as something accidental, caused by disease, environment, or the like.

When we examine smart twins, they appear to have greater myelination of their neurons. Myelin is a flat protein that acts as an insulating sheath, wrapping the nerves and the neurons in our brain. Myelination appears to be part of the mechanism involved in laying down long-term memory – more myelin, more memories. It may also help sustain signals and allow them to move faster over a longer distance: the increased insulation allowing the brain to include information from more distant parts of the brain within a given thought. But increased myelination may be an effect of higher intelligence rather than a cause. The brain is responsible for a significant part of our overall energy consumption so insulating the neurons might simply help with energy conservation. This is an active area of research.

Evolution also gives a clue to the causes of intelligence. Humans, nonhuman primates, and dolphins all share spindle neurons. These spread across the brain and appear to help us coordinate complex actions between the different parts. The high function intelligence that characterizes these disparate species requires a great deal of cooperation between different areas of the brain. Take playing a musical instrument. This uses physical coordination (motor cortex), sound processing (auditory cortex), rhythm (another part of the motor cortex), along with emotional interpretation (amygdala). Humans have more spindle cells than other animals so this might explain our superior ability in performing these complex tasks.

However plausible these ideas, they are all hardware arguments. It is like me saying my word processor is better than yours because it has gold plated connectors. That might be true – it might allow the machine to run a little faster without electrical errors creeping in, but we all know it's software that matters. A great computer game is great because it is cleverly written and has beautiful graphics. The speed of the hardware might help, but it does not define 'great'.

Can we see these software effects in the brain?

No, unfortunately, this is where our imaging technologies fail. They lack sufficient resolution. We would need 100,000 times more resolution to see our thoughts, even assuming we would recognize thought if we saw it. There is no reason to believe the brain lays out thinking in anything resembling the computer software we are accustomed to reading.

There is one exceptional group of people that *does* show a software difference on a large-scale – chess players. It seems Chess Masters use a different part of their brain to process information about chess than you and I. This can be clearly seen on scans of the brain and is such a gross effect it even shows up in old-fashioned EEGs – where electrodes are taped to your head. Interestingly the effect can be used to predict greatness. Players likely to become Grand Masters show they use a different part of their brain from the rest of us at an early age. Chess players possess the only large scale *wiring* difference we know of, but there is another group with a visible physical difference, London taxi drivers. Their hippocampi are noticeably larger than the rest of ours. The hippocampus does many things, but one of its most significant jobs is to memorize maps. The three years it takes to acquire 'the knowledge' and the subsequent years of navigating London's complex streets give cabbies a 30% larger hippocampus than the average London resident.

Is Intelligence Static?

We've all seen the headline. Every summer public examination results come out and every year is pronounced a record breaker! Year after year, students get better and better grades. This creates a problem. There's is no better grade than an A – and eventually all students get As. Welcome to grade inflation – a problem affecting systems the world over, from British 'A' levels to Harvard grade point averages. Newspapers are awash with stories bemoaning the dumbing down of today's tests. "Examinations aren't what they used to be."

Grade inflation undoubtedly exists and studies of undergraduate grades show progressive compression into the top grades, most competent students get 'A's, making it difficult to distinguish a good student from a great one.

At first glance, the problem appears to be one of social engineering. Teachers don't want to disappoint, and academic institutions want to improve on last year's results. The people awarding the grades often have a vested interest in those grades improving. Even a tiny positive bias in the most scrupulously honest teacher is enough for grades to creep up. However, grade inflation might not be purely a matter of over enthusiastic teachers. IQ scores are also rising. Welcome to the Flynn Effect.

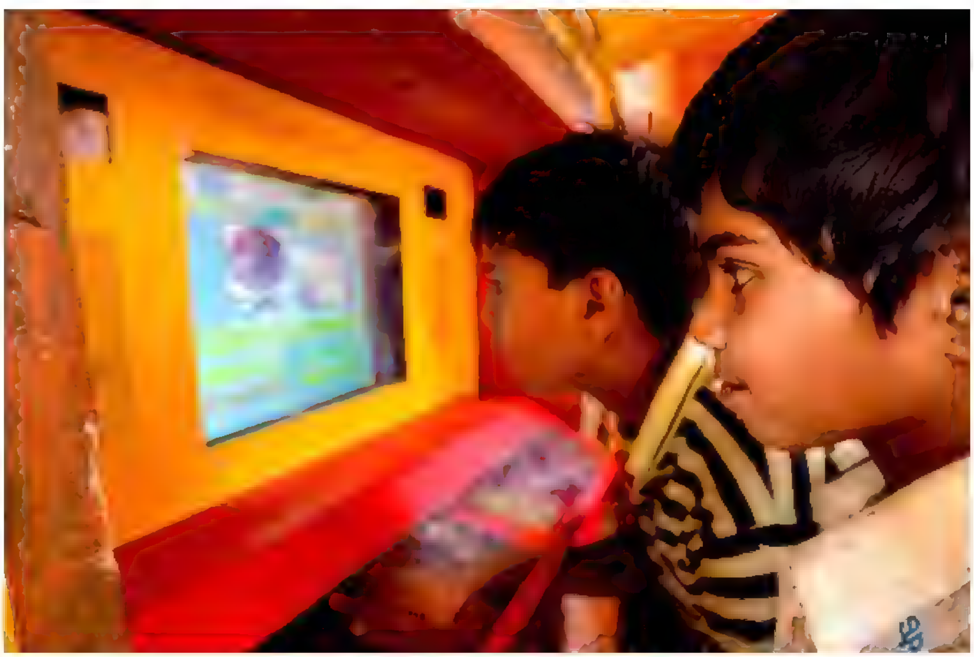
James Flynn, Emeritus Professor of Political Studies at the University of Otago in Dunedin, New Zealand, reported in 1987 that IQ scores rise over time throughout the world. All told the population gains about one IQ point every three years, and approximately every ten years IQ tests have to be re-calibrated, so the average student once again receives the average grade. This is a mystery. It is a large effect and cannot be explained by the rote learning of lots of sample questions. The human race is either rapidly getting smarter or the least smart members of society are coming up to the general average fast; either way it means there are fewer dumb people around. The Flynn Effect has recently slowed in western countries, suggesting it might be that intelligence is converging rather than increasing overall. Another interesting fact is people become more intelligent as they age, gaining about one IQ point every ten years. Against the stereotype, it's not all downhill after forty. There is hope for me yet!

Until recently we thought IQ was fixed, but new research contradicts this. Muscles get stronger with exercise, physical skills, such as playing golf and tennis, improve with practice; why not intelligence? Scientists used to believe brains couldn't get smarter; you had the IQ you were born with. You might learn more 'stuff' during your life, but the G factor stayed the same. It looks like this is wrong and we were simply not using the right exercises.

In 2008, Susanne Jaeggi and Martin Buschkuehl, of the University of Maryland, modified an intelligence test into a game and showed playing the game improved 'fluid' intelligence and increases IQ. They believe playing their game helps improve working memory – the short-term memory we use for storing sums as we do mental arithmetic – or remembering telephone numbers. Previous attempts to improve IQ through practice had not shown much success as the skills did not transfer between tests, but working memory is such a useful thing it appears to help across the board.

These factors argue against intelligence being a hardware feature of our brain. It does not remain static but instead improves with age, time, and education.

At the beginning of the chapter, I said Garry Kasparov was once thought to be one of the most intelligent people on the planet. When his IQ was eventually tested – the German magazine *Der Spiegel* put up the money – he scored 135. That means, in academic terms, he is smart but no genius. Yet, he is undoubtedly a genius by any common sense definition: the best chess player to ever live. These days he involves himself in politics rather than chess and is still uniquely able to concentrate for long periods of time. Concentration seems a very important factor. Einstein was once asked where his genius came from. He replied that he did not consider himself a genius but instead put his success down to his persistence and ability to concentrate on a problem for many years. IQ tests say nothing of our ability to concentrate over extended periods and nothing about our drive to change the world. The tests are, at best, a useful but dangerous diagnostic tool for educators. One of the worst things IQ can do is pigeonhole people. Would Kasparov have become world champion if he had been given his IQ score of 135 as a teenager rather than late in his thirties after he had conquered the world?



Hole-in-the-Wall Experiment

"Education is what is left after what has been learnt has been forgotten."

B.F. Skinner

The Learning Brain

Human beings are born with an extraordinary ability to learn through experiencing the world around them. Studies show babies as young as three weeks understand musical ideas, smiling as you play music to them in a major key and frowning at music in a minor key. By six months, babies have learned to distinguish the relationship between objects, and by two, they have a command of language and are beginning to develop a theory of self. They understand how to lie and become adept at playing parents off against each other!

Sugata Mitra, of Newcastle University, has run an experiment in India to test minimally invasive education called the 'Hole in the Wall Project'. As the name suggests, he cut a hole in the wall of a building in Delhi and put a computer in it. The hole opens out onto a slum district and local children rapidly discovered the computer. Without any formal training they picked up the necessary skills and very soon became adept at searching the Web. Remember, in order to 'pick up' this skill they often had to learn the English language as well.

Another example showing children's innate ability to learn is Nicolas Negroponte's 'One Laptop per Child' program, which gives computers to children in remote villages around the world. The laptops are a triumph of cost engineering but are fully functional and can connect to the Internet. The inspiration for the project came from an analysis of the economics of the computer industry. Huge capital investment in the western world is driving most costs down, but one cost that seems to have stuck fast is the access device. Laptops tend to remain at a floor price of around \$500, far too high for much of the developing world. At \$500, a computer store makes \$80 when they sell you a laptop. This is as low as is cost-effective

Laptops Galore

for them to stock the machine, employ someone to tell you about it, and fix it if it goes wrong in the first year. Value for money improvements have all focused on faster processors, more memory, sharper displays and larger hard drives, not lower prices. These improvements are useful if you want to shoot aliens, but overkill if you only want to surf the Internet and learn the '3 Rs'. So the 'One Laptop per Child' project has developed a device for $100.

Negroponte is often asked how he deals with the maintenance and repair issues. His answer, "There aren't any." The computers are treasured possessions and rarely broken or lost. Children become empowered by the machines and can access knowledge and information far beyond the wildest dreams of their parents' generation. Stories abound of children checking the spot prices for wheat or coffee on the Chicago Stock Exchange, and advising their parents on the price to accept for their crop. Negroponte estimates there are currently 500,000 children in South America teaching their parents to read!

It's interesting to speculate whether children learn spontaneously or are somehow 'programmed' by the adult members of society. In both the 'Hole in the Wall' experiment and the 'One Laptop per Child' program the children could simply be learning from adults and older children, but there is a novel way to eliminate this influence. Negroponte and Mitra have teamed up to run an experiment to see how children learn for themselves. They are planning to air-drop laptops into remote villages in the Andes. In this scenario, the children can't possibly learn from the

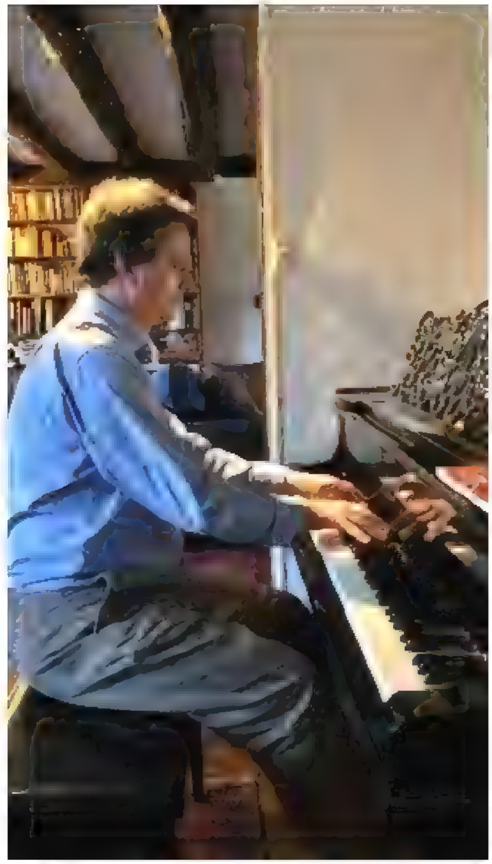
One Laptop per Child

adults – the adults have never even seen a computer before. Instead, they must rely entirely on their innate learning ability. At this point, the experiment has only just started; I will put details on my website as the experiment progresses.

The 10,000 Hour Club

Learning by experience takes humans quite a bit of time. Anders Ericsson, Professor of Psychology at Florida State University, studied musicians in the early 1990s and found they had accumulated a huge number of practice hours by the time they became experts. His research was popularized by Malcolm Gladwell, in the book *Outliers*, and by Daniel Coyle in *The Talent Code*. The idea is that humans need around ten thousand hours of practice to become proficient at a skill. The more skilled players seem to have simply accumulated even more

Dan McLaughlin



Piano Practice

practice. A number of people have wondered whether you can take this literally, and if you devote 10,000 hours to practicing something you can become world class. Dan McLaughlin from the USA used to be a professional photographer and decided he might like to become a professional golfer. He quit his job and is now 3,500 hours in. So far, he has achieved a 4 handicap. I also personally got bitten by this bug and am learning the piano. I am about 3,000 hours in and am making good progress.

Gladwell's interpretation of Ericsson's results is not without controversy. Ericsson stresses 'purposeful practice' is the important element. Practicing the wrong thing for ten thousand hours will just make you good at doing something wrong. Practicing without concentration and attention will equally have little effect. One illustrative example is the story of Edward Sanford, a supreme court judge, who read the morning prayer aloud every day over a 25 year period. After he retired he was asked if he could recite it from memory. Despite reading it as many as 5000 times during his working life, he was unable to remember it. It seems you must purposefully practice the exact thing you want to do if you wish to learn it, in this case recall.

Computers don't require practice to learn a skill. If their program is right they work correctly, and if it is wrong, they are always wrong. Computers can be programmed to learn but so far this learning has been limited to specific problem domains, such as face recognition. They do not have the general-purpose capability humans enjoy.

Astrological Clock at Hampton Court Palace

"The die is cast"

Shakespeare

"How does the water of the brain turn into the wine of consciousness?"

David Chalmers

Determinism

I have free will.

Look…

I can choose to type any word I like.

Giotto...

Many philosophers tell me I am deluded. I was always going to type that word and I have no free will. Everything in my life is predetermined. I'm rather like a character in an enormous video game. The character might think it was free to act, and its actions would appear random. Yet from the moment the player clicked the button to start the game, every action the character takes is determined by a preprogrammed set of rules. This is the free will debate. How can we tell we are free? Would there be any observable effect?

One of the big problems is that philosophers codified much of our modern theory of free will in the 19^{th} century, at a time when all the known physical laws were deterministic and reversible. They could not see a way for free will to emerge from such physical laws. There was even a group called the Compatibilists lead by David Hume that thought free will could coexist with determinism. Provided you felt free it did not matter that your actions were inevitable.

We all want free will to mean actual freedom to make conscious choices. We would like to affect the world in which we live; not the other way around. I dislike making definitions – I find they take away from the core argument and only result in linguistic jousting, but it seems that two centuries of philosophers have avoided a proper discussion of free will by loosely defining the term. Here is my definition:

'We consciously, and through the exercise of will, make decisions between different choices without anyone or anything causing the decision in advance. Others can influence decisions – by offering advice or even holding a gun to our head, but we choose.'

If you can devise a better, stronger definition please email me and I will revise my definition to your better one. I'm searching for the most powerful definition of free will – totally free and born out of the exercise of will.

The human mind appears to have free will. At least this is my personal conscious experience. Computers, on the other hand, do not. They run programs that dictate exactly how they will operate in every situation. Could a computer be programmed to have free will? That's hard to do. Let's see why.

Thinking with Clockwork

Astronomers have been predicting the motions of the heavens for centuries and to do this they need accurate clocks. The very first clocks were sundials. These suffered the obvious disadvantage of not working at night, but it was also unsatisfactory to use the motion of the sun to predict the motion of the sun. The earliest 'heaven independent' clocks used water flowing through small holes in pottery vessels. They were effective over short intervals but plagued by dust, dirt and evaporation. It was the invention of the anchor escapement that enabled the first accurate mechanical clocks.

By the sixteenth century clockmakers had gone to town developing astrological clocks with more and more gears, to show all manner of information; the phases of the moon, the motions of planets, even the motion of moons orbiting those planets. These clocks became hugely ornate. The astrological clock at Hampton Court Palace was built for Henry VIII circa 1542 and, as well as showing phases of the moon and the signs of the zodiac, it accurately calculated the time of high tide at London Bridge, allowing Henry to travel quickly to the Tower of London. You might also notice it shows the sun orbiting the earth! Copernicus published his book, *De revolutionibus orbium coelestium (On the Revolutions of the Celestial Spheres)* showing the earth orbited the sun a year later in 1543, and it took centuries before it became accepted fact.

Clocks need gears. The humble gear is a simple machine. They work because wheels of different size have different circumferences – the distance around the edge – but one full turn is the same for all wheels. Imagine you have a circular sweet such as a Life Saver – or Polo for

British readers – and you roll it once around the wheel of your car. The small sweet will turn many times. Now put a pencil through the hole in the sweet, jack up your car so the wheel is off the ground, hold the sweet next to the wheel of your car and press the accelerator. The sweet will spin round very fast and probably disintegrate in a shower of minty sugar. This is the principle of gearing. A small circle has to do a lot of work to keep up with a big circle. It's very predictable. The sweet will turn a set number of times for each rotation of the car wheel, equal to the ratio of the circumferences of the two circles.

Gears usually have teeth to lock the wheels together, but this is really just to make sure they can't slip against one another when they transfer huge forces, such as in racing cars. Some passenger cars have been built with smooth gears; a friend of mine had one at university. If he put his foot down too hard, the gears would slip, heat up and you would get a terrific smell of burning rubber. If you were lucky you could leave the car for a few hours and all would be well. But, if not, you had to replace the rubber belt, which was very expensive. Toothed gears generally win out.

Toothed gears also have the enormous benefit of being digital. This is quite important if you want to keep things accurate. Gears can't move a fraction of a tooth so if a toothed gear has 'slipped' forward a small amount, it will be kicked back into position when it meshes with another gear.

In a modern mechanical clock, a balance wheel swings back and forth on a spring and moves the main gear one notch forward each time it passes its central position. Gears divide this down to move the hour and minute hands. If I put the hands of a clock at midday and let the clock tick 86,400 times, the clock hands will come back to the same place. Once you understand how a clock works you can play a trick. If you tell me the number of ticks the clock has tocked, I can tell you the exact position the hands will be in. To a small child this might be dressed up as a magicians trick – but, of course, it is simply a matter of dividing the number of ticks by 60 and then 60 again to calculate the amount of time elapsed. This type of precisely predictable behavior is called deterministic behavior. Something is deterministic if you can set it up in a particular way and know the exact state later or, conversely, examine something and trace it back into the past.

Modern computers scale up clockwork and make it much more efficient; gears are translated into electronic logic gates and a quartz crystal vibrates at 1000 million ticks per second to give us the clock tick. On each tick, the computer can do a mathematical operation, store and retrieve information, or branch down an avenue in its program. Using these simple building blocks the computer allows us to play computer games or process the words of this book as I write. Importantly, all these operations are deterministic; given a set of inputs the computer will always generate the same outputs and that means a computer has no free will.

"Ah," I hear you say, "but my computer plays games with me and is not predictable, otherwise I would always beat it." You are right, but the computer has a clever trick to fake non-deterministic behavior: it uses you!

Computers on their own cannot generate random numbers. All a computer can do is generate a pseudo-random number and it does this by working its way through a very long calculation. It could, for example, calculate the first thousand digits of π (pi), and then start using the subsequent digits as random numbers. The digits look jumbled up but we know they follow an entirely predictable pattern. The computer appears to behave randomly because when I press the button to kill an alien the computer picks the number it had counted up to at that moment, say the $55{,}678^{th}$ digit of π, and uses that. It is I, the human, who unconsciously picks the precise moment in time and therefore provides the random element. My choice is governed by all sorts of extraneous quantum influences: Did I have coffee this morning? Was it a big mug or a small cup? How hot was it? All these things will be important as they determine the amount of caffeine absorbed across the brain blood barrier and the exact timing of my actions.

Humans are not good at consciously generating random numbers. We tend to choose the same numbers too often. If I ask you to pick a number between one and ten, you are likely to choose three or seven. This effect is called social stereotyping; magicians often use it when they pretend to read your mind. The problem arises because we tend to over think the problem. I asked you to pick a random number between one and ten. You won't pick one or ten. Five is too obviously the mid-point. Even numbers don't feel random. Nine is too large. That just leaves three and seven. So the mind reading magician has you! Humans can unlearn this social programming and become quite good random number generators but normally we tend to conform.

There is a way two humans can generate a truly random number without training. Find a friend for this experiment. One of you should pick any number between one and ten and start counting under your breath, when you get to ten just go back to one and keep repeating. The other should wait a while and then shout stop. The number reached should be genuinely random. Please post the results on my website and I'll tell

Lava Lamp

you if this crowd-sourced random number generator really works. There should be no way to predict the resulting number as both of you are affected by quantum randomness and, provided you wait a little before shouting stop, any social stereotyping should be overcome. If you want to be scientific, remember the random number you started with and the length of time before your friend shouted stop. There should be an improvement in randomness with the amount of time they wait.

In the absence of human interaction another way to give a computer access to a random number is from a quantum device. A lava lamp works well! The Lavarand, developed by Silicon Graphics, is a hardware random number generator which uses images of a lava lamp to seed a random number generator. It is covered by U.S. Patent 5,732,138, titled "Method for seeding a pseudo-random number generator with a cryptographic hash of a digitization of a chaotic system." Got that!

A computer does not acquire free will just through the injection of randomness. You could simply put an intercept on the link from the lava lamp to the computer and completely predict the computer's behavior. The system as a whole will certainly do unpredictable things, but the computer did not make a choice; behaving randomly is not exercising free will. Where is the will?

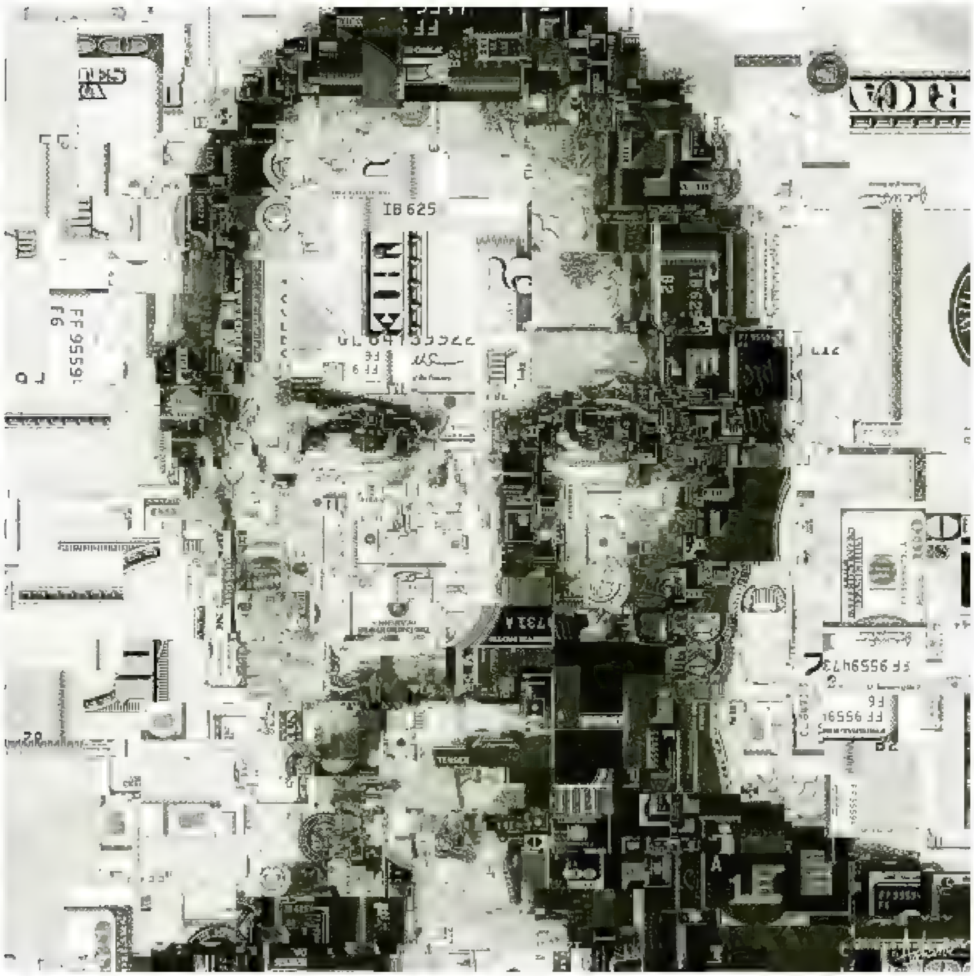
Consciousness

I remember my first trip to Death Valley in the United States. We were driving along the main east-west highway at the bottom of the valley and a sign said, "Turn off your air conditioning now." I did as I was told and to cool down I opened the window. When I put my hand out I felt nothing; no wind chill, nothing. The air was so hot the wind carried no heat from my hand. When I imagine hot weather it always brings back this memory. It's my conscious experience of the world.

Humans experience the world through a vivid lens we call consciousness. It allows us to think about the world as we watch it and plan actions. But, it also summons associated memories, something scientists call 'qualia'. Most writers describe consciousness as an internal dialogue with themselves and see it as a consequence of human language. That's probably because most writers are linguists. Non-linguists, perhaps even dyslexic engineers like me, experience consciousness as more of a visual dialogue.

It's hard to pin down consciousness as the difference between humans and computers. Computers do have something that resembles consciousness; they have watchdog functions, they plan and anticipate actions and are aware of their own existence. But they don't understand or make free choices based on this consciousness. It is an entirely mechanistic affair. A computer might know its CPU is overheating and send a notification message to the administrator, but it does not really appreciate what this means. It does not have our sensation of a near death experience. This self-awareness is the 'hard question' of consciousness. Why, despite the computer knowing it is overheating, does this not translate into the intense experience we have? Philosophers, such as Daniel Dennett, think this lack of consciousness is only a matter of time; once computers live long enough and have sufficient internal complexity they will begin to experience the world the way we do. We are nothing special.

The problem with consciousness is it does not seem to have any externally discernible effect. Anesthetics can take it away and brain scanners can see that it has been switched off, but what is it for? I think it comes hand in hand with our faculty of creativity. Consciousness allows us to shape the world – not the other way round.



Steve Jobs

"We can't solve problems by using the same kind of thinking we used when we created them."

Albert Einstein

Creative Theories

Once I have exercised my free will by getting out of bed in the morning, I often decide to do something creative. Humans seem driven to create. We compose music, draw, paint, and solve mathematical puzzles. Computers are not naturally creative; they spend most of their time doing exactly the opposite – following preset rules. Is this a fundamental limitation distinguishing the computational world from the real world?

The Conventional View

Most scientists believe pattern-matching algorithms in the brain allow us to be creative. To see how this might work, imagine our brains are chaotic – not hard to do – and process many competing ideas at the same time. The neurons in our brains build millions of useful, and useless, connections based on the patterns in the data we see and hear. Then a selection process goes to work – something akin to natural selection – to sift and prune the connections until something bubbles to the surface and we get that, 'aha' feeling.

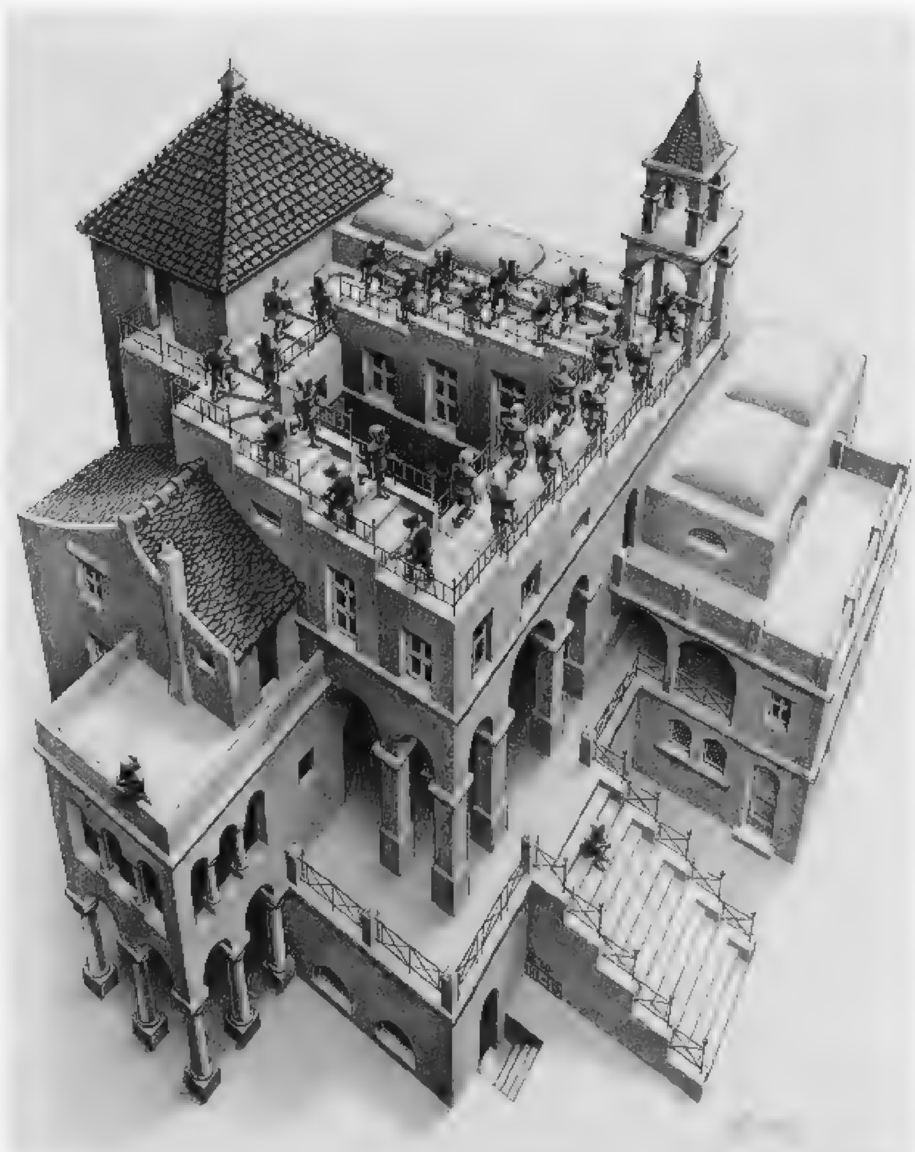
Douglas Hofstadter, Professor of Cognitive Science at Indiana University, famous for the book *Gödel Escher Bach*, has written a computer program using pattern matching to discover number theorems; things like any number ending in a zero is divisible by 5. The program produces interesting results, even perhaps generating some new theorems. He argues the human brain is essentially a scaled up version of his program. By the way, if you like trivia, his book *Fluid Concepts & Creative Analogies* was the first book ever sold on Amazon.com.

The Unconventional View

Roger Penrose, Professor of Mathematics at Oxford University, holds a completely different view. He thinks brains operate in a non-algorithmic manner and provides a sketch of the possible mechanism in two books – *The Emperor's New Mind* and *Shadows of the Mind*. He suggests tubulin molecules, which form the skeleton of our neurons, exploit quantum-gravitational effects to calculate non-computable functions. The scientific community was initially highly skeptical that quantum effects could survive the warm, wet environment of biological systems, but in January of 2014, Edward O'Reilly and others at UCL discovered plants use quantum effects to improve the efficiency of photosynthesis. No prize has yet been awarded for this discovery but it must be a contender for a Nobel Prize at some point. Recently Travis Craddock, now of the Nova Institute in Florida, has submitted a paper showing a very similar geometry of proteins exists within tubulin microtubules in the brain. He believes this is evidence quantum effects may exist there as well.

A simple quantum effect in the brain could merely reduce the resistance of the wiring in the brain to help conserve power and avoid overheating. We recognize this is a major problem in building small, powerful conventional computers. Roger Penrose suggests an altogether more radical idea. He proposes our brains are quantum gravity computers capable of calculating non-computable functions. We don't yet have a theory for quantum gravity so his idea is at the cutting edge of physics – read highly controversial. He raises a deep mathematical question. If the Universe is deterministic and effectively equivalent to a computation, how does 'creative' knowledge emerge within it? Lots of knowledge can be manufactured by simply mechanically rearranging data. That's what happens when I watch a DVD or play a computer game, but, at some point in the past, a director or a programmer had to put in the creative effort to make the movie or write the computer program. How did that happen? Was it baked into the fabric of the Universe at the moment of the Big Bang? Is what we take for a Universe really nothing more complex than putting a DVD in the slot and hitting play?

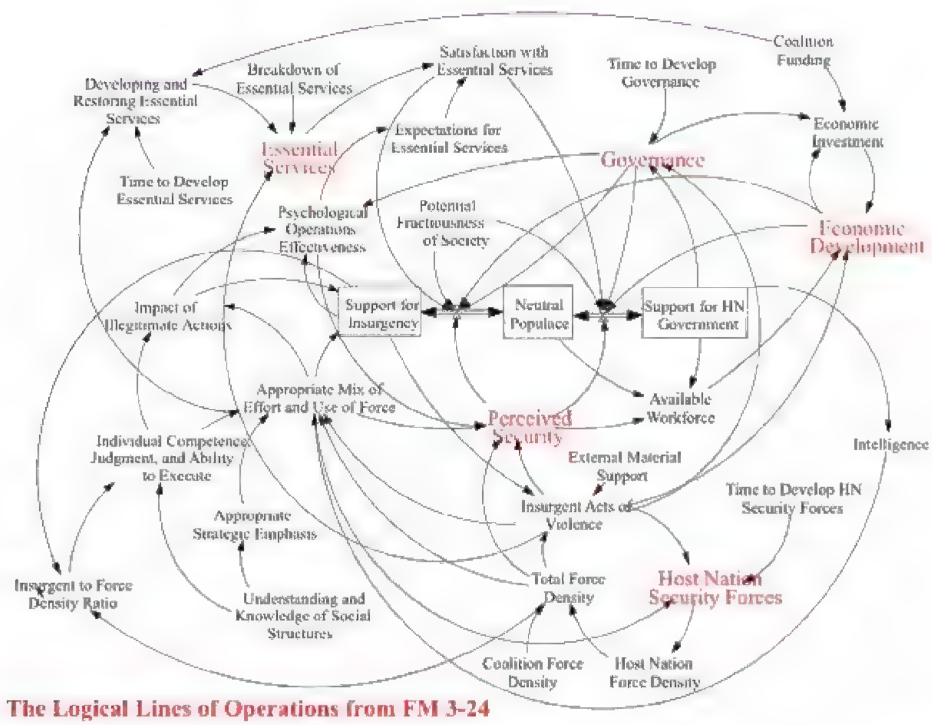
One last piece of trivia links Hofstadter with Penrose: Roger Penrose and his father invented the Penrose Steps, inspiring the never-ending staircase in the Escher prints featured in Hofstadter's book. For movie buffs, the Penrose steps appear in the film *Inception*, starring Leonardo DiCaprio. The fact we get pleasure from these trivial links tells me something is going on in our brains that is not so mechanical.



M. C. Escher's Ascending and Descending (Penrose Steps)

Chapter 2

UNDERSTANDING



Afghanistan COIN Dynamics

"Power corrupts, PowerPoint corrupts absolutely."

Ed Tufte

"No battle plan survives contact with the enemy."

Colin Powell
Originally, Helmuth von Moltke

John Masters stood up to address General Stanley A. McChrystal and his military staff in Kabul. The topic, of course, the war in Afghanistan. The main war lasted only eight weeks, but this did not end the conflict. A level of tribal violence and insurgent warfare rumbled on for years, killing around 30 people a week. Masters' job was to explain the dynamics of Afghanistan and provide politicians and military commanders a framework to understand what was going on.

Think about your country for a moment. What maintains the fabric of society – police, family, the local charity club, church, newspapers, the broadcast media? All these institutions work to keep us civilized, but what happens if a country loses them? There are institutions in Afghanistan, good and bad: tribes, gangs, corrupt officials, families. Masters had spent a year investigating these interactions, and questioning the returning commanders. He and his team believed that understanding the dynamics of the conflict was the key to bringing peace to Afghanistan.

If you live in an industrialized country, you rarely see society without its civilizing web in place. One interesting 'experiment' that shows what happens when it fails was the 1976 traffic police strike in Finland. Finland is a fantastically law abiding country where most people obey both the written and unwritten laws. During the strike, this behavior changed. Many people began parking illegally but refrained from blocking the roads. A few took advantage of the absence of police to drive incredibly fast – twice the national limit. These would be labeled as 'defectors' in game theory. Without traffic police, a different automotive

General Stanley A. McChrystal

morality emerged, a different structure to society. Of course all the other parts of society remained the same. People paid their taxes and went about their lives normally; only the traffic behavior was affected.

Afghanistan has had most of its social structures removed over the last forty or so years. First the Soviets, and then the Taliban, took apart much of the fabric until finally the Allied Forces swept the Taliban out, leaving very little behind. There were no police or courts, and few laws – or at least none enforced by the rule of law. The Allied Forces have spent a decade rebuilding these structures. Before we examine Masters' presentation, let's look at the daily life of an Afghan farmer.

If you are an Afghan farmer you have a dilemma. Your most reliable crop is opium. It grows well in the arid soil, does not require irrigation, and is resistant to most pests. For this crop there is a financial infrastructure to rival the Chicago Commodities Exchange. You get interest free loans secured against the crop, and you can forward sell your product on a futures market. Your investors can 'add value' by dealing with the major pest – the US military. They do this through the simple expedience of taking pot shots at them if they get too close to the crop. Since a field of opium is worth $30,000 and a militia wage for the year is $350, you can easily employ a few men to protect your investment. Of course, you are indebted to thugs and criminals, but they are at least reliable thugs and criminals.

On the other hand, the traditional products of the Himalayas – walnuts, pomegranates and vines – need years to cultivate. There is no forward market and the timescales over which you must take risks are far greater. If you believe your American protectors will leave before the crops mature, you will be loath to plant and care for them. But, if you make the decision to take this risk, you have a strong incentive to foster stability and reap the rewards of your effort. There is a feedback effect: the balance of power between all the different parties is important to the decisions you make, and the decisions you make affect your desire to invest in future stability.

Masters' team built a slide pack to demonstrate the complex interactions between the groups: farmers, security, stability, markets, military power, and emerging institutions. The COIN – COunter INsurgency – dynamics slide shows just how hard it is to communicate complex topics between human beings. The presentation is beautifully crafted but it was a public relations disaster. At the end of the presentation General McChrystal said jokingly, "When we understand that slide, we will have won the war." The slide was paraded in the press as, "the most complicated PowerPoint slide in history."

If you invest a little time on the slide you will understand it and may even see it as a thing of beauty. But Masters' audience was obviously expecting something different and, presented with this level of complexity, went into shutdown. Perhaps they wanted a simpler presentation, a high level summary, a few bullet points. Of course, there is no simple presentation on Afghanistan. The lesson is that context, timing and expectation are often as important to good communication as the elegance of the content, and that information is a complex thing.

If you want a lighthearted poke at PowerPoint here is Peter Norvig's PowerPoint version of the Gettysburg Address.

Understanding

Next time you are in a business meeting, count the number of times the word 'understand' is used. If you ask the people around you what it means you'll stump many of them. That's because understanding has two very different meanings. Most people don't separate these meanings but the distinction is important. Understanding means to decode information, to comprehend – but, more importantly, it also means to absorb and internalize information. That feeling you have when you 'get it'.

If I say, "I understand" I mean I have taken in the question you asked and decoded it into ideas so I can provide an answer. This can be quite a mechanical process and computers routinely understand natural language and answer questions – Apple's digital assistant Siri being a case in point.

When I say, "I understand a problem" or "understand a culture" I mean something far less tangible. Somehow the information I have gathered over my life is formed into a matrix within my brain that allows me to ponder and run scenarios. I can predict the effects of my actions before I do them, and often anticipate your responses. That's clearly a very useful evolutionary adaption, but is there more to it than that? Roger Penrose and David Deutsch think understanding allows us to transfer non-symbolic information from one brain to another. We don't run programs in our brains, nor do we store precise information such as lists and tables. We have, therefore, had to evolve a creative approach to communicating skills and understanding each other. One of the most closely studied areas in the field of communication is when it breaks down in the lead up to a disaster.

"The human mind tends to look for clear linear relationships, we like solutions that are close to the problem in time and space and make sense when we think about it quickly, unfortunately, those simple solutions are usually wrong and come from acting on a complex system as if it was a simple one."

Brett Piersen

Gettysburg Address, Peter Norvig

Space Shuttle Columbia Crew Photo

"For a successful technology,
reality must take precedence
over public relations, for Nature
cannot be fooled."

Richard Feynman

Bad Understanding Can Kill

On January 16, 2003, at 3:39pm, the Columbia space shuttle took off from Cape Canaveral. During the launch a small piece of foam insulation broke off the fuel tank and hit the shuttlecraft. The event was recorded on a few low-resolution video frames. They show a tiny white object hitting the shuttle and a plume of dusty material splattering outward. The shuttle made it safely into orbit and for two weeks engineers on the ground debated what to do. In the end, it was decided the risk was minimal and the shuttle could safely return to Earth. On reentry, the shuttle disintegrated, killing seven astronauts.

NASA managers had decided the shuttle was undamaged based on a series of presentations by the engineers. One image in particular analyzed the potential damage to the shuttle's tiles from an impact. Read the slide, look at the key frames, and decide for yourself what action you would have taken.

Shuttle Tile

Review of Test Data Indicates Conservatism for Tile Penetration

- **The existing SOFI on tile test data used to create Crater was reviewed along with STS-87 Southwest Research data**
  - **Crater overpredicted penetration of tile coating significantly**
    - **Initial penetration to described by normal velocity**
      - Varies with volume/mass of projectile (e.g., 200ft/sec for 3cu. In)
    - **Significant energy is required for the softer SOFI particle to penetrate the relatively hard tile coating**
      - Test results do show that it is possible at sufficient mass and velocity
    - **Conversely, once tile is penetrated SOFI can cause significant damage**
      - Minor variations in total energy (above penetration level) can cause significant tile damage
  - **Flight condition is significantly outside of test database**
    - **Volume of ramp is 1920cu in vs 3 cu in for test**

BOEING 2/21/03 6

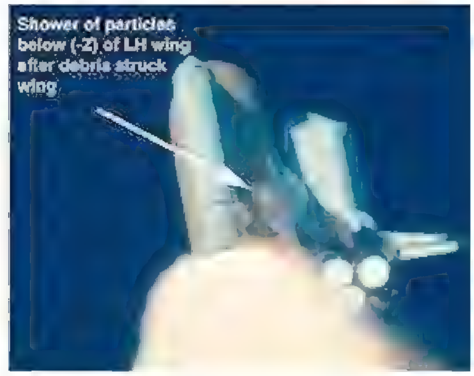
NASA Internal Slide

WHAT DO YOU UNDERSTAND FROM THE SLIDE?

Some images of the launch are shown on the right

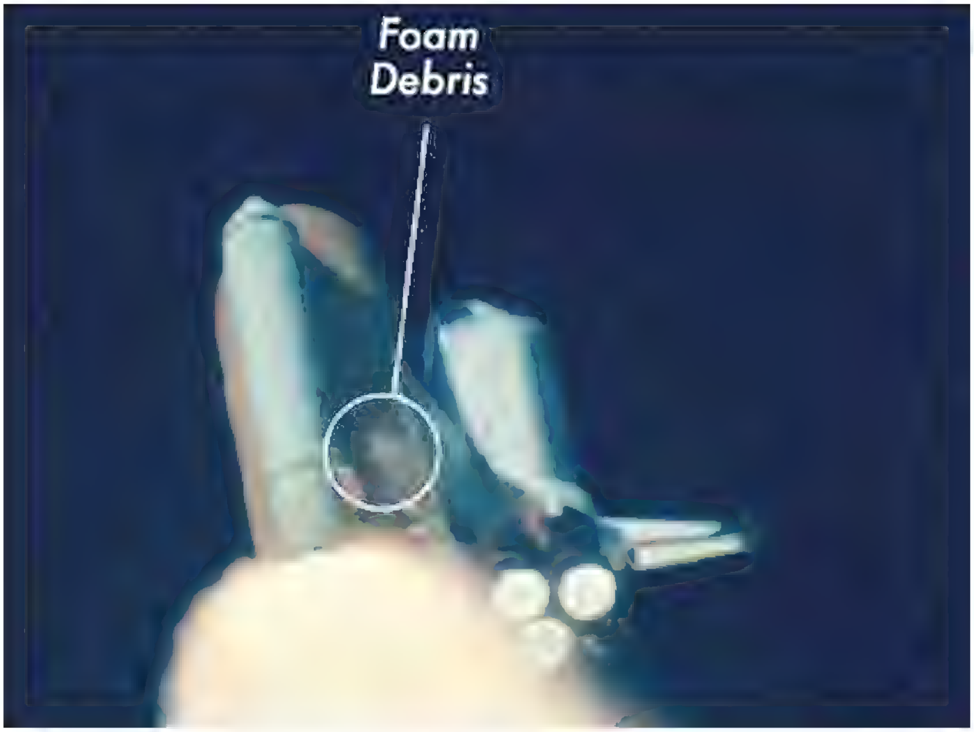
Here is what you should have understood from the slide: tiles are really tough but if the foam dislodged from the fuel tank broke through the outer coating it would cause significant damage. The estimated speed of the foam hitting the tile was 640 times greater than anything previously tested. Worried?

Is this a proper understanding of the problem? You have the slide and the images. Take another look and think hard. If you want, you can check a video of a similar launch on YouTube to get a feel for the scale of things, but the still frames shown all the information you need to make your conclusion.



Photographs of the Foam Impact from Video Footage

Frame Showing Foam Dislodging



Still from Ground Camera

LOOK AT THE IMAGES, WHAT HAPPENED?

The truth is you simply don't know. If you are puzzling over the strength of tiles, you have been misdirected. There is video footage of some sort of impact on a wing mostly covered in white tiles, and a slide describing the effect of a benign sounding 'foam' hitting those tiles. But what is the evidence for an impact on a tile? The shuttle is certainly not made entirely from tiles; I can see a window in the picture. You should instead be asking more questions, "What happened?" "What hit what?" and "How bad is that?"

It was bad. The foam, a very tough material, had hit the leading edge of the wing, a weak point, punching a hole through it. The wing failed on reentry and tore the shuttle apart. Clearly, a full discussion of the possibilities did not occur amongst the shuttle team, or perhaps it only happened amongst the engineers in private. Once the analysis was tidied up and presented to 'management' it was a one-way communication of the conclusions, not a discussion of the underlying ambiguous thought process. The result: people passively listened to the information rather than interactively understanding it and agreed on the recommendation that it was safe to return. Clearly they did not understand the ambiguity otherwise they would have realized they did not have enough information to form a conclusion. This is the tragedy of lack of understanding. If they had known how little they knew, they could have deployed a spy satellite to take pictures of the damage – one was available nearby and would have taken a few hours to re-task – but they did not.

Ed Tufte served on the second shuttle disaster commission and provided an analysis of the disaster. He views slides as a poor medium for communicating complex problems and thinks documents are far better. The danger with slides is they force you to simplify information in a way that destroys the essence of the information. His analysis of the failure of communication at NASA formed a major part of the final report on the disaster. Later he coined the paraphrase "All Power corrupts; PowerPoint corrupts absolutely." Good communication benefits from stories and narrative, not bullet points and graphic fluff. Instead of using bullet points, speak! After all, we have evolved for 250,000 years to understand language, but only 25 to read PowerPoints. '

If you write presentations, Ed Tufte's book *The Cognitive Style of PowerPoint* is compulsory reading. He argues that much of the information you want to communicate is complex and interconnected. PowerPoint or any similar presentation software encourages you to simplify it into hierarchical bullets. The format implies simple causal relationships where none exists. This is dangerous. Communication should convey understanding – which is very important – and not just information. What, you ask, is the difference?

Searle's Chinese Room

"The hardest thing to understand is why we can understand anything at all."

Albert Einstein

The Imitation Game

As an experiment, I am going to ask a student to spend a week in a locked room. The room is perfectly nice; it has a bed, a light, a desk, some reading matter, oh, and we'll give him some washing facilities too! Every now and then I post some food under the door to keep him going, Pop-tarts and pizza (thin-crust) work well.

On the first evening a note is pushed under his door with a symbol on it. The student puzzles for a while, then opens the book sitting on the desk. The book says, "If you get a piece of paper with symbols on it look them up and follow the instructions." He looks up the symbols and the entry in the book says, "Go to page 44, write down the third symbol on a piece of paper then post it back under the door." He follows the instruction and is rewarded with another piece of paper, this time with a larger set of symbols on it. Again he follows the instructions in the book and posts his answer back under the door. This goes on for several days. He is somewhat bemused, but it passes the time, and he diligently looks up the symbols and performs all the complicated actions as instructed.

Meanwhile, I meet our new Chinese graduate student and explain to her she needs to interview a potential translator for the department. He has just come in from Hong Kong and there is a health scare, so we have quarantined him in the lab room. He is bored and I have some paper for writing messages. She writes "hello" in Chinese on a piece of paper and posts it under the door.

The exchange of notes goes on for a few days and the two seem to be getting on well. There is even a little romance in the air. When the week is over I open the door and the two meet. The graduate student says, "Hello. It's nice to finally meet you in person." The man is puzzled because, of course, she has spoken to him in Chinese. He knows no Chinese.

"I'm terribly sorry, but I don't speak Chinese," he says.

She is puzzled, "But I spoke with you this last week!"

"No, I really don't speak it," he says.

And, of course, he is telling the truth. The book he has been using contains the rules for answering questions in Chinese, but he has absolutely no knowledge of the language. I'll leave to your imagination whether the two strike up a real relationship and live happily ever after.

This is the Story of the Chinese Room. The setup is able to fool someone into believing there is a Chinese speaking person in the room, yet there is not. Where does the understanding of Chinese lie? The man definitely does not understand Chinese. And the book clearly does not understand Chinese because it is an inanimate object. Yet the person outside the room is convinced she is communicating with a Chinese speaker. The analogy to a computer is clear. The book is software and the man blindly following instructions is the hardware. John Searle, who devised the thought experiment uses it to show computers can never understand because there is no place in a mechanistic system for understanding to exist.

The Chinese Room has sparked huge argument in philosophical circles; let me boil it down to its simplest form. First, let's refute Searle's position with the 'System Argument'.

The man plus the book form a system. Systems understand; their individual components do not. My blood does not understand. My brain without blood would not understand it would be dead! Plug my brain into a good supply of blood; add a dash of glucose, and it will understand the most complex of things.

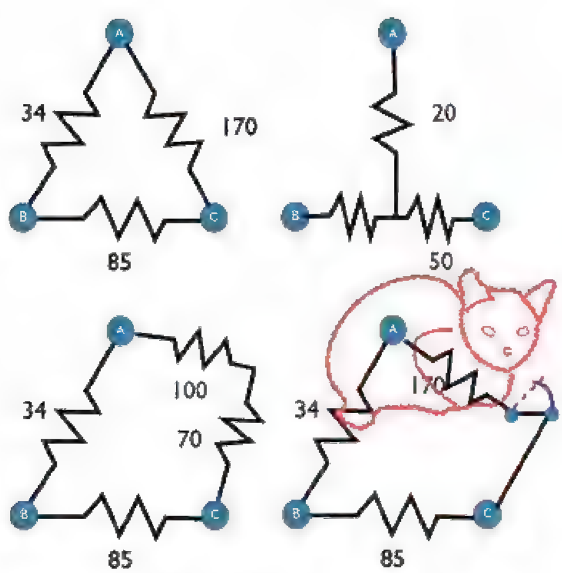
The systems argument is elegant and most scientists think this is the definitive argument against Searle, but Searle has a neat way to counter it. "Imagine", he says, "that the man memorizes the book and leaves the room. Now there is no system, there is just the man, but the man still does not understand Chinese; he is just parroting rote-memorized words and rules." Computers, Searle argues, process syntax the rules of language; humans understand semantics – the contextual meaning of language.

Artificial Intelligence (AI) proponents hate the Searle argument. They believe the memorization of a set of words and rules is exactly what gives us knowledge of Chinese. That is why we go to school!

A key problem posed by Searle's Chinese Room is whether you can know everything about a situation from just looking at the inputs and outputs. This is very similar to the restriction posed by the Turing Test. In that case if we were to trace the wire from our computer terminal to the other room we would either find a human typing messages or a large box covered in flashing lights. This would definitively answer the question whether we were talking to a man or a machine. Similarly, if we opened the door to the Chinese Room we would immediately know whether there was a real Chinese speaker in there or not. But opening the door on both tests misses the point. The question asks, "if the inputs and outputs are the same does it matter what is *really* going on inside a closed system?"

Black Boxes

Experiments involving closed systems are known as Black Box experiments. They presume you can learn *everything* about the inner workings of a box simply by probing it from the outside. Young electronic engineers are often given black boxes as a test. Electronic components hidden in the box are connected to three external terminals on the outside. The student is asked to deduce what is in the box using only an electric meter to probe those terminals. Here are a few examples of the possible contents of a black box. They would all show up identically on the student's meter. Although internally different they are externally identical. Even my 'silly' fourth choice with a cat in the box does not give



Black Box Equivalence

itself away if all you have to go on are electrical readings. (I dare say the cat would make its displeasure know if left in there for any time.) The contents are, therefore, said to be black box equivalent.

The reason for teaching engineers about black boxes is to help them understand how to simplify things. We could construct option four, with a cat and some food, but it would cost a great deal of money. Option 1 is functionally identical from an electrical point of view, but for a fraction of the cost. Steve Wozniak and Steve Jobs were so successful when they started Apple because Wozniak was brilliant at simplifying logic circuits. He could take a design with thirty chips and come back with a black box equivalent solution using only five. It was a fraction of the cost and far more reliable.

Scientists put great store in black box equivalence because of a principle called Occam's Razor. William of Occam was an English Franciscan friar living in the fourteenth century. He proposed the idea of minimal explanation. It states that, 'among competing hypotheses, the hypothesis with the fewest assumptions should be selected'. When trying to explain the workings of a black box, the more complicated inner workings should be discarded, as they have no externally verifiable effect over the simpler mechanism. Our extraneous animal must be eliminated! Sorry.

Ironically, given his calling, Occam's Razor is sometimes wheeled out as a disproof of the existence of God. Surely God is a complication unnecessary to the explanation of our Universe. The argument is illustrated beautifully in Carl Sagan's book *Contact* and the film of the same name. God gets the last laugh in Sagan's book when the difficulty with Occam's Razor is brought into sharp focus. Occam's Razor contains an inherent paradox. At any moment in time we only have evidence to support the simplest of explanations, yet we know many of these simple explanations are incomplete. We regularly discover new phenomenon – dark matter and dark energy being some recent examples. If we stopped discovering new things, Occam's Razor would be a good way to simplify our thoughts. Occam's Razor is a useful intellectual tool to prevent us over complicating explanations, but there will often be explanations that are correct, but for which there is not yet any observed effect.

If we go back to our black box example, we see the flaw in concluding the boxes are identical from examining only their inputs and outputs. Opening them would clearly show they are not identical! But, how would this fact reveal itself if they remain closed? The answer is: over time. If something in the box has memory or understanding, it could present one set of results for a while and a completely different set of results later.

In my trivial example, the cat could eat a wire and change the operation of the black box. Now there is an open circuit where none existed before. If this happened, the output would change and we would need a new theory to explain it. If the circuit was attached to a missile control system or a life support system, you would really want a full understanding without waiting. It's humans nature to try to open black boxes. This is what MRI scans, X-rays, particle accelerators and all our other tools of scientific investigation are for. We want to open all the black boxes of nature and see what is going on inside: simply waiting to see what happens is not acceptable.

In a sense, we live in a black box. We experience the world through our senses, seeing with our eyes and feeling with our hands. The brain never directly experiences anything; it only infers the likelihood of

Scene from The Miracle Worker. Helen Keller pictured at the moment she understood language.

something from the signals it receives. This is similar to our engineer probing the terminals of the circuit of a black box. How can we know our experience of the world is real?

Understanding the World

The French philosopher Descartes gave us an explanation for this paradox. He spent a long time looking skeptically at everything we perceive. For example, when we poke a stick into a pond, the surface of the water bends light and the stick appears to have a kink in it. Our eyes tell us the stick is bent, but our brain 'knows' the stick is straight: it's an illusion. Descartes wondered if something so simple could be an illusion, perhaps the whole of our experience is too.

His eventual solution underpins much of modern philosophy – 'I think therefore I am', *cogito ergo sum*. Even if we doubt everything else, we cannot doubt we are thinking about this doubt. At least we can rely upon the existence of this 'thought' as some reality. Descartes built up from this bedrock the real world we live in. We can be sure we experience things and can apply logic and use thought. We can use this intellectual faculty to tell a great deal about our Universe.

True Understanding

In the QED lecture series, *The Strange Thing about Light and Matter*, Richard Feynman relates the story of the ancient Mayan astronomers. 3000 years ago they were able to predict the motion of Venus in the sky using only pebbles. They had a simple system that could predict when the planet would rise over the horizon. Put a stone in the jar every day, take out a stone once a week, add a stone at every new moon. If the number of stones in the jar is divisible by 23, Venus will rise. I'm making up the details but you see the idea... It's a very simple algorithm. What should we conclude if the Mayans had perfected their calculations to predict the motion of Venus and it proved reliable over a whole century? Would this constitute understanding?

Feynman would say no: the Mayan understanding was not complete. It was only black box equivalent to our modern understanding over a limited period. We known that once the Sun begins to run out of fuel it will swell to a red giant and explode, destroying Venus and the Earth. Their model could not predict this catastrophic failure. Our modern deeper understanding of the workings of the solar system allows us to predict this future even though there is no clue from the motion of Venus today. Understanding allows us to predict discontinuous events: a system changing its state or a star running out of fuel.

We see the same predicament in stock markets. Stock markets normally behave in a linear fashion but, when they go wrong; they go *very* wrong. Recent recessions have been made much worse by the failure of hedging systems to handle market disruption. Some even think the crises were caused by the automatic trading strategies of these hedging systems.

The quants – as mathematicians in banks are called – spend considerable effort modeling financial instruments to show that if one stock goes down, another will go up at the same time. If the stocks are held together your investment is safe because, on average they will remain constant. The problem with these correlations, which often hold reliably for many years, is that when trouble hits they fall apart. Historical correlations don't give us understanding of the future: something that was only meant to happen once in a million years has happened within six months. As they say on your investment papers, past performance is no predictor of future results.

Do Computers Understand?

Today's computers don't have our general-purpose ability to understand. Watson was thrown off by badly formatted English. The human contestants, by contrast, had no problem with this. Just how good would Watson have to be, to call it – or should I say 'him' – intelligent? How could I judge this had happened? Alan Turing proposed an ingenious test in his 1950 paper Computing Machinery and Intelligence using 'The Imitation Game.' We now call the Turing Test.

If we ask a series of questions to a computer and we cannot tell its responses from those a human would give, then the computer is, for all practical purposes, the same as a human. Since we are intelligent – or at least we hope we are – the computer must also be intelligent. QED.

That's all there is to the Turing Test. Puzzled? Let's pick his argument apart.

Imagine you are chatting away on Facebook with someone you don't know. They may have posted a photograph so you can see what they look like. The photo might be a fake; you have no real way to tell. What question would you ask the other 'person' to prove they were human and not a computer? There are obviously some giveaway questions. Please multiply the numbers 342,321 and 23,294 and give me the answer. This would be very hard for a human but easy for a computer. If you got a very quick answer; the computer would have given itself away. But, the computer has been programmed not to give itself away, and it is free to give the answer slowly or even reply that the calculation is too hard. Our computer can say anything it likes, including lying to pass the test! If the computer can fool a questioner into believing it is a human then Turing argued the computer has shown it is at least as intelligent as we are.

It used to be assumed that the field of broad general knowledge would be hard for a computer, but Watson has shown this is not so. With enough storage and a reasonable algorithm, winning a pub quiz is well within the capability of a modern computer.

The really difficult questions for a computer are philosophical ones, novel questions and things that don't fall into a pattern. For example,

"Are you happy?"

"What do you think of Shakespeare's Hamlet?"

"Is there life after death?"

"How went it?"

"Think Differ..."

If a computer could plausibly answer this sort of questioning for an extended period, say fifteen minutes, should we conclude it is intelligent, or do we need more time to be certain?

Turing's approach to certainty was simple. Just ask lots of questions. As you ask more and more questions, you will become increasingly certain you are talking to an intelligent being. He characterized it as a linear process; after 15 minutes of questioning you might be 99% certain and after a few hours 99.9% certain and after a few days completely certain. The problem with this approach is it does not flush out discontinuities. What if the questioning suddenly stopped without warning or explanation? A human responder is likely to worry that the questioner has had a heart attack and do something to find out what is going on including leaving the room. Humans can make creative leaps, solve non-computable puzzles or come up with a clever new joke. A humans could even announce the test is a waste of time and walk off. They just exercised free will! A computer cannot do these things.

Each year a group of scientists enters a competition run by Cambridge University to win the Loebner prize, a competition to see how close a machine can come to passing the Turing Test. If you can beat the test you win $100,000. So far no one has come close and scientists are beginning to realize just how hard it is.

"On the Internet, nobody knows you're a dog."

New Yorker Cartoon

With the anonymity the Internet provides we can imagine all sorts of strange scenarios if the Turing test could be passed. You would have no way of knowing what you were talking to. The New Yorker ran a cartoon back in 2000. "On the Internet no one knows you are a dog." We come across a similar problem the other way around when we encounter bad customer support. A few years ago, while trying to get an answer to a computer problem, I became convinced the thing responding to my emails was a machine. The company did use machine responder technology so it could well have been. I asked it to prove it was human by putting the word marmalade into an English sentence and fixing my problem. The human pretending to be a machine saw the joke, fixed my problem and replied "Marmalade is served with butter and toast." The test worked!

Uncannily not Human

The sister test in robotics is equally hard. The goal is to simulate the physical human form, its movements and mannerisms. It's easy to get close, but close is not good enough. The term 'Uncanny Valley' has been coined to describe the discomfort humans have with something that tries to simulate a human being but does not quite get there. I think it is part of the reason Madam Tussaud's waxworks are so fascinating. Humans have a love-hate relationship with facsimiles of themselves. They love the flattery but feel a sense of revulsion at anything that comes too close.

Searle and Turing

In the Turing Test, we limited our senses to the purely symbolic: using only typed words on a screen. I could break the lock on the door and go into the room to see what was there.

"Aha!" I would say.

"I can see you're a computer, I, therefore, know you'll be good at sums and bad at creativity."

But Turing wants us to see if the difference is given away purely through intellect. He argues there is no way to tell. But if you follow my argument from chapter 1, there is one way: ask the computer to find a non computable solution to a mathematical puzzle. This is, in practice, a difficult test to pose because it might take a *very* long time. Twenty-five billion people have lived on planet Earth during the last 350 years, and about 5 million of them were mathematicians. None of them was able to solve the problem posed by Pierre de Fermat until Andrew Wiles turned up but this is a clear difference between humans and computers. However long you give a computer it would never be able to solve the problem.

This creativity test would take centuries to run if non-computable thought was rare, but I think we see it often – on display even when we tell jokes. In which case computers and humans should be easy to tell apart: humans are the funny ones. I am not saying you can't build a brain; our brains are physical devices, after all. I just believe a computer or a mechanistic machine, cannot think like a human being.

I like the Searle argument but qualitative arguments are insufficient. We need a quantitative argument. In the forthcoming chapters, I am going to look at the mathematical argument underlying the difference between human intelligence and computer processing. Before we do this let's take one last look at a qualitative difference; the way computers and humans communicate.

Chapter 3

BODY LANGUAGE & BANTER

Body Language

"England and America are two countries separated by a common language."

George Bernard Shaw

"I speak two languages, Body and English."

Mae West

"The body never lies."

Martha Graham

In the summer of 1986 Ronald Reagan and Mikael Gorbachev met in person for their second negotiation session, this time at the Höfði House in Reykjavik. For five days, the leaders talked alone except for interpreters. Reagan badly wanted to develop the Strategic Defense Initiative; known by its nickname, 'Star Wars'. The idea was to put smart weaponry in space that could destroy ballistic missiles before they reentered the atmosphere. Reagan believed this would remove the threat of imminent destruction that had hung over the world since 1945. Gorbachev, on the other hand, felt this was just another escalation in the Cold War, and the Soviet Union would be forced to build yet more weapons to overcome the American defenses. He wanted Reagan's plans shelved, arguing that it broke the Anti-Ballistic Missile Treaty. He was probably right. The leaders talked back and forth, unable to overcome the impasse. At the end of the summit there was a mad scramble to announce some sort of deal, but this proved difficult. In the last moments before they had to conclude a communiqué, Reagan suggested they abolish *all* nuclear weapons. Reagan's negotiating team was horrified and shut the door.

For decades, the American strategy had been to use nuclear weapons as a deterrent against the apparent numerical advantage of the Soviets. In all the potential scenarios analyzed by the Pentagon, Russian forces ended up overrunning American forward positions – otherwise known as Western Europe! The only way to stop them was through a release of nuclear weapons, which, inevitably escalated to all-out nuclear

Ronald Reagan and Mikael Gorbachev

war. It was assumed this inevitable progression deterred the aggression in the first place, and the threat of mutually assured destruction kept the world peaceful. Giving up this tenet of defense strategy was something the American military just could not contemplate. Many people did not think it a rational defense strategy; it seemed appropriate the acronym for mutually assured destruction is MAD, but this was the status quo.

We now know our worry over Russian superiority was groundless. The West's technological advantage, founded on the invention of computing and sophisticated materials technology, gave us a huge advantage. In the only battle to be fought in the 20th century between Russian and Western tanks, during the first Iraq war, most of the Russian tanks were destroyed with no losses to American tanks. We know this now, but we are talking of a time when paranoia over the Soviet advantage was the common view.

There is speculation that Reagan had muddled intercontinental ballistic missiles with all nuclear weapons. I do not think this is true. Reagan was a man of vision, quite comfortable with using his folksy way to convey sincere belief, and I think abolishing all nuclear weapons was in his mind. It would have been a breathtaking moment.

In the end a rather feeble communiqué was put together and the talks declared a technical failure. But, both leaders had seen eye to eye; both were prepared to make major concessions and both wanted an end to the old strategy of mutually assured destruction. Wiping each other out was no longer considered a successful outcome! The meeting, and

Höfði House in Reykjavik

the fundamental thawing of relations between East and West, was to lead to the Intermediate-Range Nuclear Forces Treaty and the end of the Cold War.

Face-to-Face Communication

What really happened between these two leaders when they met and talked? Was it a mechanical process of offer and counter-offer, as easily executed by fax, or is human interaction more complex than this?

Reagan, as a young man, had been a liberal, sympathetic to socialist ideals until a painful strike in California caused him to lose faith in the politics of the left. Gorbachev, a lifelong Communist, was desperate to reform the Soviet economy and make it more competitive. He, also, had come to see the hypocrisies that could emerge in far left-wing ideology. I don't believe this common experience could have been communicated by fax or email. Indeed, I am sure these specific points were never made, but the nonverbal communication must have conveyed something of their common background and purpose.

When we phone someone or exchange emails, the interaction is factual, there is no body language, and we rarely laugh. When we travel to meet someone, we spend a great deal of time with them. The average length of a phone call is two and a half minutes, but meetings, especially when one party has travelled to see the other, can be hours long. When humans meet they greet each other, shake hands, sit in the same room, talk at length, and laugh. Body language is important; people mirror each other's postures, adopt open and receptive stances, and make eye contact. You can see this in the picture of Reagan and Gorbachev above. Body language allows us to convey qualitatively different things, such as trust and happiness. It is very expressive; you can see the more guarded postures of Yasser Arafat and Shimon Pérez below, just after they negotiated a landmark peace deal. Can you tell if the leaders smiles are false?

Communication

Communication is one of mankind's greatest expenditures. The US telephone system is arguably the largest machine on the planet, while the world's mobile phone networks have a capital value of $2.5 trillion, greater by an order of magnitude than all the steel plants in the world put

Yasser Arafat and Shimon Pérez

together. This lifeblood of our existence – long-distance communication between human beings – turns out to be amazingly difficult, even with all our clever technology.

In recent years the Internet has, in theory, allowed each and every person to communicate freely with any other person on the planet. In some of the most distant parts of the world mobile phones, and projects such as; 'One Laptop per Child' are rapidly bringing unlimited communication to all. This communication can be personal, one-to-one, or broadcast: I can talk to people interested in a particular topic directly. As we watch the Arab world democratize, catalyzed by the Internet, there is no question that digital communication has now become a major force in the world. Yet, people don't communicate over the Internet as much as you would expect; they often use the Internet to set up phone calls during which they arrange meetings! This is odd. We have a fantastic phone system and sophisticated communication technologies; email, video and instant messaging. Yet, we still choose to travel when we want to communicate.

On the face of it, there should be no difference between a phone call and a meeting. In principle the same information can be conveyed. Yet when we want to really understand someone, we always go to meet

Smiles Fake or Real

in person. No great treaty or big industrial contract has been negotiated without a face-to-face meeting. We see this daily: people talking on the phone get to a certain point, give up, and arrange to meet in person.

The consequence is that we spend $550 billion annually, flying around the globe to meet each other. Each day the world's population takes three million plane flights. Around 80% of these are business flights. Some are people emigrating or going to do specific manual tasks, but most are to have meetings. We have always assumed that this is because the parties are unable to reach a sufficient level of trust over the phone and need face-to-face interaction to build that trust, but it may be that the parties are not able to convey sufficient information to fully understand each other. Face-to-face meeting may convey much more information than we think.

Smiles

When we smile naturally we use a full set of facial muscles, including the muscles around our eyes. When the smile is forced those eye muscles remain passive and the smile, although superficially the same, is missing something. You can't put your finger on it, but the look is insincere. A study of marriages in the USA analyzed smiles in wedding photographs. The couples with false smiles divorced much earlier than the genuinely happy couples. Similarly for high school photos; people with genuine smiles at 18 years of age were happier later in life and in more stable relationships. Smiling is really important. It is good to be around people who smile, they are more successful – and nicer.

There is also a curious reverse effect. The link between our minds and bodies is much more fundamental than we thought. If you grasp a pencil between your teeth, it forces you to smile. Try it. The mere act of smiling is found to make you happier, it causes the release of the chemicals called endorphins which improve your feeling of well-being.

Micro-expression Analysis

Since the involuntary movements of the muscles around our eyes give away genuine happiness, a whole science has evolved looking for other biological cues to mood. The two most interested groups are the FBI, trying to detect lies, and poker players, trying to make money! Much has been written on the topic, including a few best sellers, but the evidence for micro expressions is mixed. Regardless of whether involuntary actions give away our emotions, humans voluntarily use a great deal of body language when talking.

Body Language

A study by Albert Mehrabian is often cited to say 93% of the information in a conversation comes through nonverbal cues. This is misquoted. The study really stated 93% of the *emotional* content is nonverbal. That's more believable. And further studies have shown when there is doubt, nonverbal cues win over verbal information every time. The rule is sometimes laid out as the 7%-38%-55% rule – 7% words, 38% tone of voice and 55% body language. Remember this is emotional content, your conviction and sincerity. You will still have to get over the factual information you want to convey.

Learning Swedish with The Two Ronnies

Try this experiment on a friend. Tell them you like their shirt using different tones of voice: sarcastic, sincere, amazed. Then see what they understood. You will find it difficult to appear sincere because I have told you to say you like their shirt unless of course you really do. When you use sarcasm they will find it hard to process your statement. It is revealing how we use the information.

Interestingly, a piece of research described in *Scientific American* shows even insincere flattery is effective. If you want a pay rise from your boss, any form of flattery will do. Vanity appears to override skepticism!

Interaction

The normal cadence of communication between people includes a great deal of mutual interruption. When a meeting breaks down we often see people begin to say things like, "Please don't interrupt me," "Do you mind, I was talking," "Pleeeease, let me finish." If the meeting is really getting out of hand, third parties will often step in and tell one to wait for the other. This is where the mechanics of face-to-face interaction fail, as we need to interact in order to communicate effectively.

Because we have a lot more time in a face-to-face meeting people can wander 'off topic'. This is an important part of the process of communicating. After all since most phone calls are 2-3 minutes and most meetings an hour, there are another 57 minutes to fill! These off topic items bring in social experience and help us form the background context we need to properly communicate.

What is Background Context?

Alex and Bella are both fans of the British comedy duo, the Two Ronnies, and enjoy their learning Swedish sketch. Bella asks Alex what kind of sandwich he wants for lunch. Alex replies 'M'. Bella laughs. If you have seen the sketch you will understand the background context to the joke. If not this paragraph might as well have been in Swedish. Take a look at the sketch on YouTube and reread this paragraph... Now you understand.

Do I think in English?

Most scientists believe we think thoughts using language, but most scientists writing about thought are linguists or psychologists. If you are a dyslexic engineer like me, language is a long way down the processing chain. I think abstractly and then translate those thoughts into words. Some ideas don't map between languages and often, one language adopts the words of another to fill in the gaps. Some interesting examples are:

Zeitgeist	German, spirit of the times
Schadenfreude	German, enjoying others misfortune
Chutzpah	Hebrew, audacity

All of these are fully signed up, card carrying entries in the Oxford English Dictionary.

Some languages have fewer distinctions between ideas: truth and law are the same word, 'torah', in Hebrew. Languages have different tenses and structure. In Chinese all words are one syllable and the script is pictographic rather than phonetic. This is unusual, even Egyptian and linear-B, which look pictographic are mostly phonetic. With single syllable words, Chinese uses voice inflection to change meaning; a rising or falling tone can change the meaning of a word from 'grey' to 'girl'. In many Western languages rising voice inflection is used to indicate a question, as in Australian English or irritation, as in English English. So how do the Chinese show if they are annoyed or want to ask a question? They elongate their words and accentuate the changes in intonation. An argument in Chinese can sound quite alarming to the Western ear, with its percussive monosyllables and extreme inflection changes. This degree of inflection is used in English, but only in extreme emotional contexts: A Chinese argument over cold tea can sound like an accusation of murder to a Western ear.

Symbolic Communication

The earliest recorded permanent human communication is cave painting, dating to 33,000BCE. Written communication emerged in Sumer, the southern part of Mesopotamia (now Iraq), using a script called Cuneiform, written on clay tablets. It was used primarily for accounting. The Sumerians are responsible for our common use of base twelve. Twelve hours in a day, inches in a foot, and notes in the scale; all stem from their civilization.

Although not the first to write stories, the Greeks perfected the dramatic forms we use today: poetry, prose and plays. Watch an episode of 'Law and Order' and you are seeing a direct descendant of a Greek tragedy, complete with suffering and justice denied. All this permanent *thought art* is made possible by the translation of ideas into symbols.

Scripts and Symbols

The world supports a huge variety of scripts split roughly into phonetic, representing the component sound of words, and pictographic, stylized pictures of the ideas.

Chinese Traditional and Simplified

Some scripts have interesting quirks. Ancient Hebrew, although phonetic, is a script where vowels are omitted. Modern Hebrew often leaves them out as well. This means words can be ambiguous and need context to decipher them. A common set of Chinese characters has long been used by Mandarin, Cantonese, and Japanese speakers even though their spoken languages are entirely different. The script languages of these people are gradually diverging and might in time become entirely separate languages too.

The Chinese government in Beijing has moved to using simplified Chinese for Mandarin speakers, while Hong Kong continues with the traditional form. Japanese has developed many new characters for

Latin	Hello Reader
Japanese	読者 こんにちは
Russian	Здравствуй читатель
Greek	Γεια σας αναγνώστη
Hebrew	שלום קורא
Arabic	مرحبا قارئ
Chinese	讀者： 您好！
Chinese	读者： 您好！ *simplified*
Korean	리더 인사
Japanese	読者 こんにちは
Linear-a	can't be translated!
Linear-b	 (best I could do is 'new wine')
Lao	ສະບາຍດີຜູ້ອ່ານ
Hindi	पाठक (hello)
Persian	سلام خواننده
Hieroglyphics	

modern ideas, such as computers, that differ from the Chinese, and mixes in a great deal of Katakana, a script allowing the phonetic representation of foreign words. If you walk around these countries their signage looks quite different, although I am told Cantonese speakers can still read simplified Chinese. Take a look and normally you will find them to be quite different. Each example in the figure is my best attempt to translate the phrase "Hello Reader" into a script and the corresponding language.

Symbols of the World

English is one of the most irritating script languages of all. It commonly uses etymological elements, showing the history or origin of the word that has nothing to do with the sound of the word. A word like school has the 'k' sound spelt 'ch', showing its historical derivation from the Greek, but confusing for pronunciation. English has 53 sounds derived from only 26 letters, so there are plenty of letter combinations, many of which are irregular. Because the language favors historical convention over simplicity, sugar is pronounced "shu-gar" whereas sand is strictly phonetic. As for Leicestershire I'll leave that as a test for the American readers amongst you. If you're British, try Mattapoisett, a town in Massachusetts named in Native American.

Yet English is also a 'lovely' language. Because of its richness there are often twenty different ways to say something, and a dozen words to choose on any topic. One of my own favorite words is 'jump'. It is phonetic, but also onomatopoeic and even pictographic. Jump both sounds like a jump and looks like a *jump*.

Two scripts that puzzled scholars for many years are Linear-b and Hieroglyphics. Linear-b – found on clay tablets on the Island of Crete – turned out to be a coded form of ancient Greek with some slight quirks, such as dropping the letter 's' from the ends of words. The 's' is superfluous in most Greek words, and dropping it saved precious clay space!

Hieroglyphics was a real puzzle. It looks so like a pictographic language that it fooled many people for centuries. The Rosetta Stone was discovered in 1799 and became the key to their deciphering. This stone had the same edict written out in 3 languages – Greek, Egyptian and Demotic. The French adventurer Jean-François Champollion decoded hieroglyphics in 1822 and although it *looks* pictographic, it was found to be predominantly phonetic. Linear-a, another script found on the Island of Crete has yet to be decoded and remains one of the world's great-unsolved mysteries.

All these different ways to code ideas into symbols present the children of the world a great learning challenge. Because written language is so young, in evolutionary terms, our brains have not had enough time to evolve to master it. Instead words co-opt parts of our brains originally evolved for different purposes. As languages differ in their construction they co-opt different bits of the brain. It is possible to see this using brain imaging.

Dyslexics – and I am one – have difficulty in translating between the realm of conceptual thought and written script. This translation is subtly different for each language. Chinese speakers use their motor cortex to process characters. Young children write out the characters over and over, to memorize them, so the 'muscle' memory is highly involved. French and Spanish children use the audio pathways, as most of their language is phonetic, the motor part of writing is then an add-on and does not process meaning. English children must use portions of their visual cortex to process the meaning of words, as many words have spelling quirks that have nothing to do with the sound of the words. Some studies even suggest a child dyslexic in one language, because, for example, their audio pathway is impaired, might not suffer the condition in another language that relied on a visual or motor skill.

Can Objects Communicate?

The process of communication has many components, starting with something capable of communicating. Communication usually perhaps always – is something that occurs between sentient beings. I don't think of my computer as communicating with me, but rather think of it as a medium for communication or a dumb machine. But colloquial language around the subject is a little muddled. We all agree a lighthouse does not communicate, even though it can signal danger, but what do we mean when we say, "That song really spoke to me." No one believes the song is actually communicating, but some kind of communication was made nonetheless. When we talk of communication do we mean the agent or the message?

Stories

Humans enjoy communicating; we create works of art, music and literature that transcend simple analysis. The COIN dynamic slide, which we saw earlier detailing the strategic situation in Afghanistan, would probably have been better communicated with a story. Humans, unlike computers, do not cope well with large quantities of unrelated information, and studies of memory and comprehension show we benefit from a narrative structure. Let me give you a basic example. One simple trick the human brain uses is chunking. Give yourself a moment to try to learn this string of characters.

HALTNTIBMGTATLAMATLOLPOMSGTG

TRY TO MEMORIZE THE STRING WITHOUT READING ON

Now, if I divide it into chunks, you will see it includes meaningful information.

HAL TNT IBM GTA TLA MAT LOL POMS GTG

You probably won't recognize all the acronyms unless you are under 10. Even then, you will find memorizing it hard, but if you put the sequence into the context of a story then it is much easier to learn.

HAL uses TNT to blow up the IBM building in Grand Theft Auto. "Three Letter Acronyms are annoying," says MAT. I'm Laughing Out Loud; Parents Over My Shoulder. Got To Go.

We find it easier to fit new information into existing structures within our brains rather than memorizing by rote. I've used quite a bit of modern Internet slang here. You'll find young people recall this information better than older people for whom GTG and POMS are nonsense.

If you want to memorize something, experts recommend you imagine bizarre images and relate them to a story pictured in the mind's eye. Try it and you may very well find you can still remember my sentence in ten years time!

Let's try something else. The following sentences are a little different, yet the recall scores for information in the two are dramatically different:

1. I met an old tramp on 42nd Street wearing a dirty grey rain coat.

2. New York on a cold damp November day; as I cross the street I bump into an old man wearing a dirty grey Macintosh. His shuffling gait suggests some sordid intent. I think nothing of it, but this brief meeting was to change my life.

The addition of contextual cues allows you to form a mental picture. By withholding some information at the end I have used a dramatic trick to cause your brain to free wheel and imagine what happens next. You are involved in the story. Notice the *longer* story, with more data in it, is paradoxically more comprehensible and memorable.

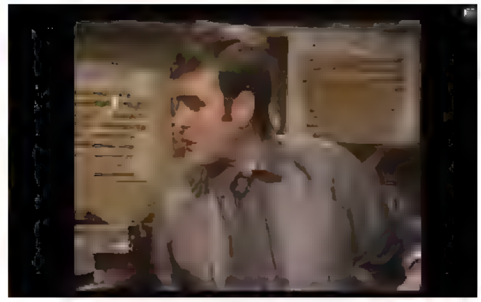
Ed Tufte makes the point about our ability to process information very forcefully. He believes presentation experts are wrong when they recommend you keep your slides to a few words! He points out the common advice to use only six bullets per slide and six words per bullet comes from a misconception that has blighted a generation of presenters. Studies performed on memory in the 1960s measured unrelated word recall. Six words are all you can remember if the words are meaningless. But if the words have meaning we can comprehend and absorb many pages of data. Hundreds of millions of people throughout the world read a newspaper every morning and can recall the stories throughout the day; the poems, songs and plays we memorize when young are usually long, comprising thousands of words, yet we are able to remember them verbatim for the rest of our lives.

When we tell a story, we are trying to draw the reader in so they can to experience our imaginary world and be 'in' the story. When I read a story perhaps *Harry Potter* I don't think about the grammar and punctuation, or even the accuracy of character portrayal. I'm transported to a different place. I experience a piece of the reality or 'imaginality' the storyteller has created. I can describe the characters, the scene, the sounds and the smells. A good author forms a complete world in our heads corresponding with the world they have in their heads. With more abstract information, comprehension and retention is harder. Often if the information does not hang together in a linear narrative it can be impossible to take in at a single sitting. However, if it forms a story and is well told so you 'get it', you do not need it repeated. We experience something of this effect when we watch a good movie. "I've already seen that one," means you have absorbed the whole story in a single sitting. You don't need to watch it over again to comprehend it.

Comedy

Finally, when you mix all the elements up, emotional understanding, body language, in-person communication and empathy; you get comedy. Humans 'do' comedy from a very young age and it's vitally important to the fabric of our lives. What purpose comedy serves in communication

My XBox is Broken
The One Ronnie



Dead Parrot Sketch
Monty Python

Gerald the Gorilla
Not the 9 O'Clock News

Fork Handles
The Two Ronnies

Andre Previn
Morecambe and Wise

Self Defense Against Fruit
Monty Python

is not clear. In life, telling a joke will make another person smile. This causes people to be happy and happy people release chemicals into their bloodstream which make them healthier. Happy people then tell jokes to others. This circular process improves the well-being of communities and helps bond people together. But why on Earth did comedy evolve to be the mechanism that does this?

Comedy may be an important way to avoid an argument when context is unclear. Much of what we say can be taken the wrong way. Simple communication of fact can sound like criticism or challenge, and humans are naturally hierarchical – not unlike packs of dogs or beached walruses. Humor allows us to test the response of others to statements, which might otherwise be taken the wrong way. Something said in a 'jokey' tone of voice may not generate a negative response, even though the raw content might be quite provocative. "Ah, late again I see..."

It is worth taking a look at some great comedy sketches because they bring home the richness of human interaction. Here are some of my favorite links as an antidote to the heavy-duty mathematics I am about to inflict on you.

The World's Funniest Joke

> *Two hunters are out in the woods when one of them collapses. He doesn't seem to be breathing and his eyes are glazed. The other guy whips out his phone and calls the emergency services. He gasps, "My friend is dead! What can I do?" The operator says, "Calm down. I can help. First, let's make sure he's dead." There is a silence, then a gunshot is heard. Back on the phone, the guy says, "OK, now what?"*
>
> Spike Milligan, from *The Goon Show*

I think comedy is a fitness display. It demonstrates to those around us – particularly of the opposite sex – that we can be creative and use non- computable thought processes, just as dancing is a fitness display of our agility and coordination. When we tell a joke we are showing others we can 'think outside the box', a valuable survival skill.

At a simple level it has been proven that animals with the ability to behave randomly escape being eaten more often than animals that follow a pattern. Non-computability is the ultimate behavioral randomizer since it is not an algorithm and cannot be copied. The ability to take non-computable thinking to its logical conclusion to create and invent has clearly taken off for humans.

Of course, another explanation might be that making people happy is fun. People like to be around other fun people so humor encourages crowds to form. If a saber-toothed tiger attacks you, and you are in a crowd, you're more likely to survive. You only have to outrun one member of the crowd!

Chapter 4

THE BRAIN

Baby EEG

"The brain is a wonderful organ; it starts working the moment you get up in the morning and does not stop until you get into the office."

Robert Frost

"The brain looks like nothing more than a bowl of cold porridge."

Alan Turing

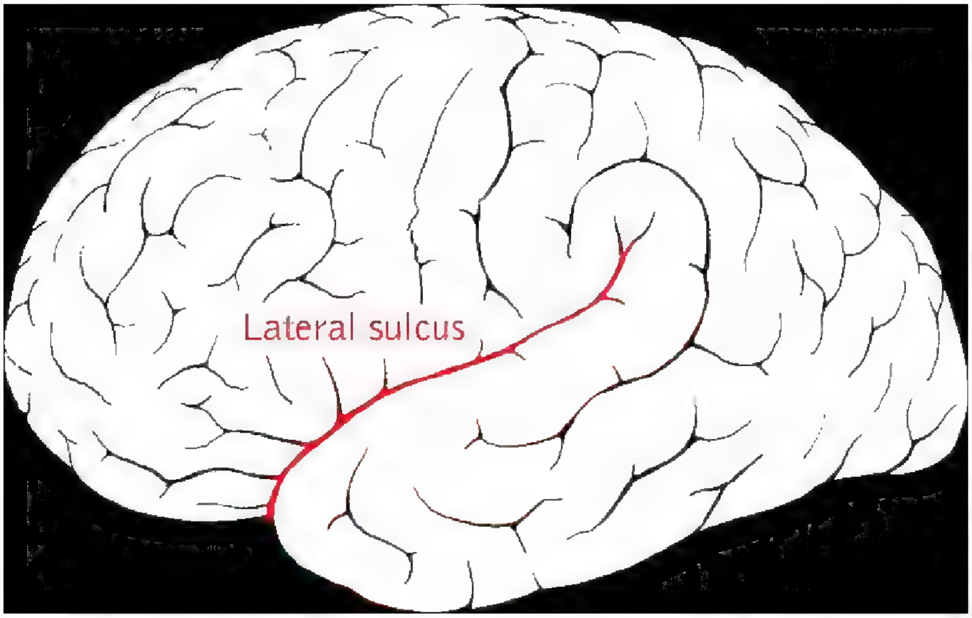
Physically the human brain is very boring. Alan Turing described it as looking like a bowl of cold porridge. To get to the porridge you must first cut through the skull, a two-millimeter thick protective layer of bone. The adult human skull has almost no gaps in it, and the only ways into the brain without a bone saw are through the eye sockets or the soft area of bone at the back of the nose. Egyptian mummies had their brains removed through the nose and preserved in a jar for the afterlife!

Thinking with Porridge

Protecting the brain is very important and the skull does a good job by being a tough, impenetrable barrier. But sometimes this toughness backfires. In 2009, Richard Hammond, one of the presenters of the TV motoring series Top Gear, suffered a crash while testing a land speed record-breaking car. Although he was in a multipoint harness, the crash, at over 200 miles per hour, bounced his helmeted head around the inside of the cockpit and his brain was badly bruised. As you know from experience, when you bruise you get swelling, and the brain is no exception. However, the brain is encased in bone, so this swelling has nowhere to escape. The resulting buildup of pressure is dangerous, causing an interruption of blood supply to the un-bruised parts. Brain damage in such accidents is often fatal; Richard Hammond was very lucky to live through the experience.

Surgeons often need to cut into the skull to relieve pressure on the brain, or to gain access to remove tumors. Going through the scalp involves a great deal of blood, but once you have a clean hole in the skull you can peel back the thin membranes, called the meninges, to reveal a wrinkly folded whitish thing that looks a bit like a cauliflower. This is the outer surface of the brain where much of our thinking is done. Unfolded, this surface layer would cover the area of a football field and this intense folding distinguishes the human brain from the brains of simpler animals. Some animals, such as elephants and dolphins, have larger brains than ours, but the area of their folded surface is considerably smaller. It is thought that this efficient folding is key to giving us the ability to think complex thoughts.

Analysis of Einstein's brain held at Princeton University shows it is not particularly massive, but it is strikingly more folded than average, and has a shorter lateral sulcus – the fissure between the front and back



Einstein's Brain

of the brain. Whether this is related to his highly creative thinking or just random chance is unknown, but it's an interesting data point in our quest to understand creativity and intelligence.

Looking through a microscope, the wrinkly grey matter is composed of 30 trillion neurons; small whitish cells sprouting filaments that wrap around each other like the tentacles of an octopus. The tentacles, and there can be as many as 10,000 per cell, are known as dendrites and spread out to nearly touch other neurons. At the other end of the neuron is a single axon. The gaps between the end of an axon and the next neuron's dendrites are called synapses, about one-tenth of the width of a human hair and varied in structure. When a nerve 'fires', an electrical pulse spreads out along the axon to the end and crosses the synapses to other brain cells. This electrical pulse is not like the flow of current in a wire: neurons don't conduct electricity. It is more akin to dominoes falling in a line. Ion gates in the walls of the neuron open, letting potassium ions flow out. As the gates open in one section, the next section is triggered and so on. Thus, electrical signals ripple out along the axon. As the electrical signals cross the synapses they either excite or inhibit the firing of adjacent neurons. There is a lot more structure to a neuron than was once thought. The textbook model is of a sequence of ion sacks stacked end to end rather like plant cells, but neurons have a far more complex structure. Bundles of actin and tubulin form a skeleton in the neuron and the neuron metabolizes ATP to recharge its firing mechanism. Neurons behave far more like small animals than inanimate plant cells.

The wiring of our brain looks a bit like the logic circuits of a computer, and our best guess is the cells in our brain form some kind of computer. The brain cells – a specialized form of nerve cell – connect to the rest of the body via the nerve cells that largely run down our spine. Thoughts trigger action and, in reverse, the nerves in our extremities sense things in the environment and relay information back to the brain. If I think, 'move my finger' my finger will move, and if it touches something I will feel the sensation. Interestingly if my finger touches something hot a reflex will kick in. Reflexes work without involving the brain. We don't have to think, "that hurts." Instead, our finger reflexively pulls away. We may say ouch, but by the time we do, our fingers already moved away from the heat.

Nerve cells are much slower than the electronic systems we build with copper and silicon. This speed is quite noticeable and limits the rate we can do certain things. It takes around 0.08 seconds for a nerve impulse to run down to the tips of our fingers, initiate an action and return to give us the sensation of the action. This may sound fast but if you're a tennis player in a rally or a pianist faced with a fast passage, the nerves don't have time to make a full round trip signal before the next action must be initiated. In these instances we need to run on autopilot and there are parts of the body where the nervous system takes action without the brain getting involved. This is particularly the case with things like walking and balance, which must respond fast to changes in ground conditions. The signals just don't have time – and don't need – to go all the way up to the top of the body for instructions. Rather like the heat reflex above, the peripheral nervous system can process information locally. After all, brain cells and nerve cells are really all one type of cell.

If you have a group of people, you can conduct a fun experiment to show the speed of nerves. Hold hands in a big circle and squeeze the hand of the person next to you. When they feel you squeeze, they should squeeze the next person's hand and so on. The rate at which people squeeze hands around the circle is limited by the speed at which nerves conduct the signals across our bodies.

Imaging the Brain

There are several ways to look inside the brain without recourse to a bone saw. The methods are fascinating in their own right, even before we start looking at the results. Each image is generated using a different physical principle.

X-rays

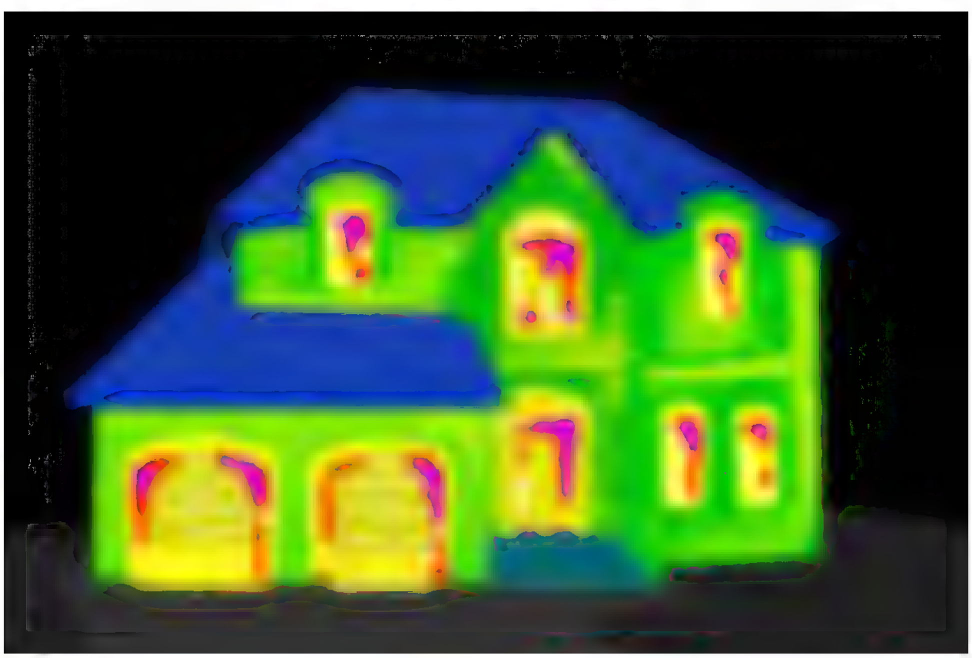
The first Nobel Prize in Physics was awarded to Wilhelm Röntgen in 1901. He had discovered 'X' rays; so called because he had no better name for them. X-rays, as they became known, are just light of a very high frequency.

Light comes in a variety of colors; at the low end of the frequency scale we see red, higher up blue and, at the top, violet. At this point human eyes give up and cannot see anything higher, so ultraviolet light is invisible to us. Bees, on the other hand, can see a long way into the ultraviolet spectrum and some flowers have beautiful ultraviolet markings that attract bees for pollination. Daylight contains a great deal of ultraviolet light which is wasted on us – other than to tan our skin. But all is not lost. Clever manufacturers put fluorescent dyes into their washing powders which stick to our clothes and convert ultraviolet into visible light, making our T-shirts look brighter as they reflect *more* visible light than fell on them. You can see this effect most easily in a disco when ultraviolet lights are shone on the dance floor and anyone wearing a newly washed T-shirt will glow bright white. The other common substance that fluoresces strongly on a dance floor is tonic water. Quinine, the active ingredient in tonic water, is a strongly

Flowers in Ultraviolet Light

Pit Viper

fluorescent substance which converts ultraviolet light down into the visible spectrum. Photoactive dyes have recently become controversial as suggestions have been made that they are unsafe and irritate the skin. Going to discos might not be quite as fun in the future!



Thermal Imaging

At the bottom end of the spectrum is infrared light. Pit vipers have evolved special organs on the sides of their heads to 'see' in this spectrum and they use this sense to hunt prey in the dark. I use the word *see* with some caution. We have no idea what their sensation of 'heat-sight' involves, but their organs are very precise, able to detect things only 0.2 degrees warmer than the background.

Infrared cues help several species of snakes, bats and insects locate things in the dark, but the animal that excels at the task, albeit using technology, is mankind. Special cameras allow us to use infrared to see in the dark or detect where our houses lose heat.

X-rays are much higher in frequency – about one hundred times that of the ultraviolet light that affects our T-shirts. The high frequency corresponds to a small wavelength that allows the rays to pass through our bodies. Later on in the book we will understand that frequency is not a proper explanation for light, as it is not a wave but rather a particle that obeys the laws of a wave. But for now we will ignore this detail.

The first use of X-ray images was to see broken bones. Bones block the rays as they are dense, but the soft parts of our bodies are almost completely transparent to X-rays. We can see the soft tissues if we turn the contrast up, but there are problems when using X-rays to view the brain. Our skull completely encases the brain and however much we turn the contrast up, all we see is bone. The solution to this problem is to perform sophisticated mathematical tricks using a computer to enhance the contrast ratio and make image 'slices' through the living head.

The slicing technique was invented independently in the 1970s by Sir Godfrey Hounsfield, working for EMI in England, and Allan Cormack, of Tufts University in America, and they shared the 1979 Nobel Prize for Medicine for their work. Legend has it that EMI was making so much money from The Beatles they could fund the enormous development cost of the CAT scanner from the profits; true or not, it's a great invention.

The best way to understand the mathematics is to picture yourself in an episode of 'CSI', the American television crime drama. An intruder has attacked someone with a knife and there are blood spatters all over the walls of the room. Enter the brilliant pathologist who reconstructs the scene of the crime from the pattern of blood on the wall. She can map the trajectory of the blood spatters and back calculate that the attacker must have been 5' 4", left-handed and wielding a 6" blade. In a CAT scan, our head is hit with billions of rays that bounce and scatter over the walls of the machine. Sensors detect the rays and a mathematical algorithm calculates an image of the body that would produce such a pattern. To

X-ray of Roentgen's Wife's Hand

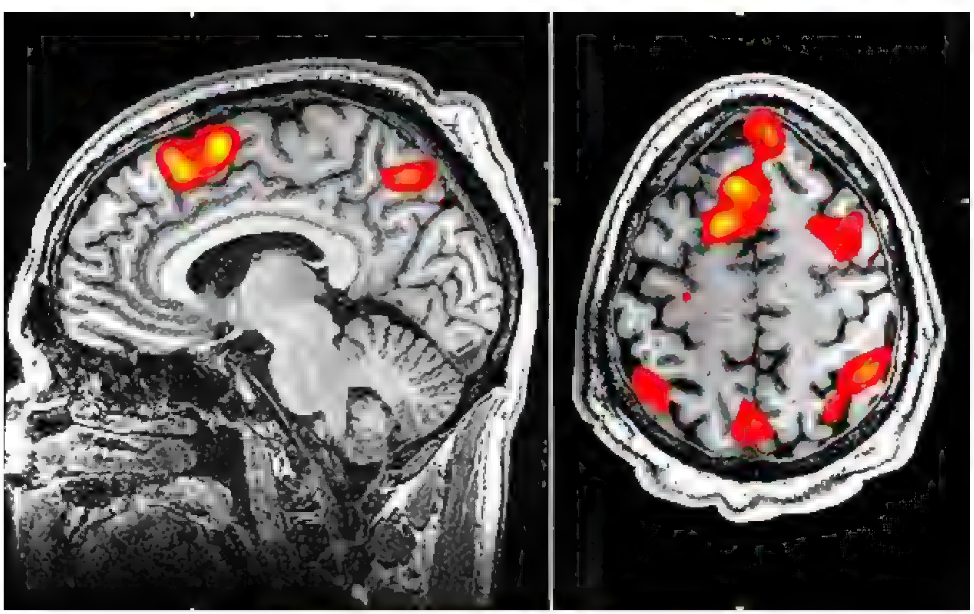
simplify things we shine the X-rays onto the head as a narrow slit of light so we only have to do the back calculation in two dimensions. Then we stitch successive slices together in the computer to form a 3D virtual image. Thus, doctors can 'fly' through the brain looking at structures such as tumors from all angles.

There are two problems with X-ray imaging. Even with clever mathematics, the dense bone in the skull blocks the rays so you don't get much contrast, making it hard to distinguish normal brain matter from something like a tumor. But the bigger concern is X-rays are a form of ionizing radiation, and ionizing radiation causes cancer.

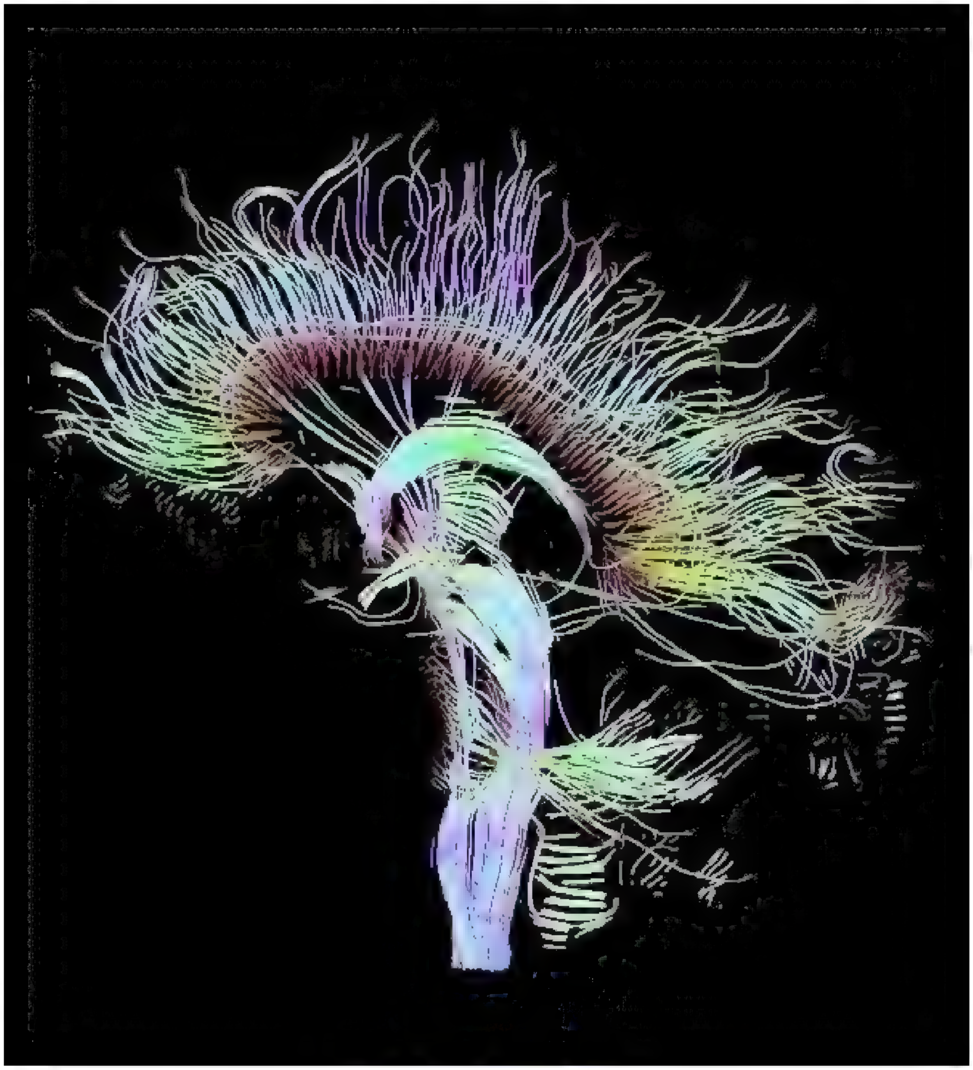
We are told to wear sun block to protect our skin from ultraviolet light; X-rays are 100 times more potent and can do a great deal of damage. Fortunately, the body repairs itself quite well in the presence of low levels of radiation. The *double* part of the double helix in our DNA allows a set of proteins in our cells to go around correcting errors when they detect a mismatch between the two strands. But, now and again an X-ray might make an irreparable fault in both copies. If enough of these faults accumulate, they can lead to cancer or, if the errors are in reproductive organs, birth defects. Doctors try to minimize the radiation we receive and give us as few CAT scans as possible during our lifetime, especially when we are young and have not yet had children.

MRI

X-rays dominated our ability to see into the human body until the mid 1970s when Raymond Damadian came up with the idea of using magnetism. Magnetic fields are not absorbed by bone and present no danger as they do not damage DNA. Ironically, the technique was originally known as Nuclear Magnetic Resonance, 'NMR', which patients thought must be dangerous because of the word nuclear. The name was



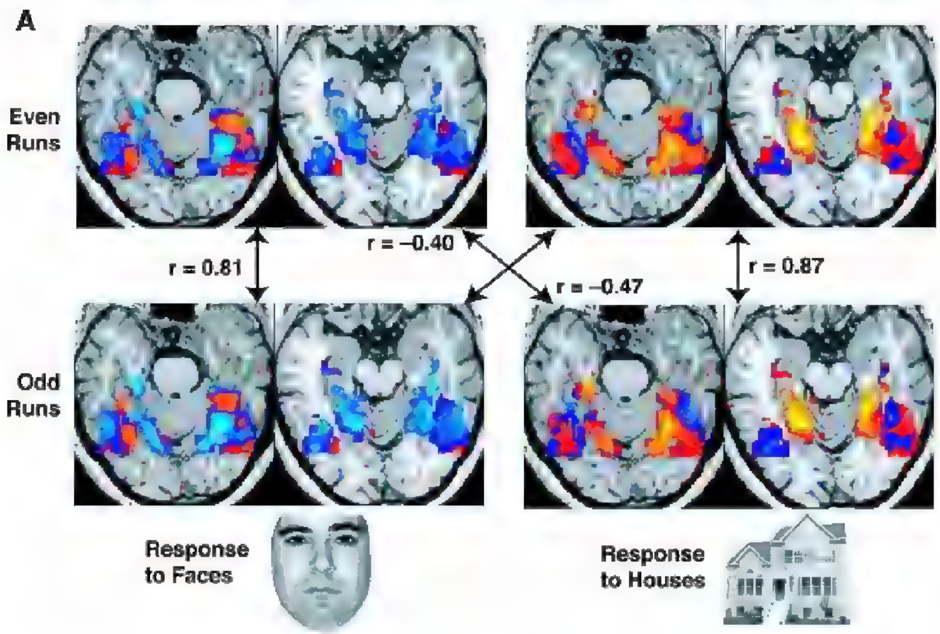
Functional MRI: Working Memory



Diffusion Tensor Image

changed to the one we use today: Magnetic Resonance Imaging, 'MRI'. The system works by applying a strong magnetic field to your body to excite the hydrogen atoms. Since we are mostly H2O there are plenty of these.

Three magnetic fields are used. First, an extremely strong field is applied to the whole body. This causes all the hydrogen atoms in the water and fat to spin in line with the field of the machine. Next a gradient field is applied to the top of your head so it is slightly more magnetized than the bottom of your feet and, finally, a pulse of magnetism is applied to the top of your head. The spinning hydrogen atoms line up a little more when this pulse is applied and then randomize again when it is switched off. As they randomize, they give off energy. The clever part is the gradient field which causes the atoms to give off energy at slightly different times – the top of your head first, your neck a fraction of a

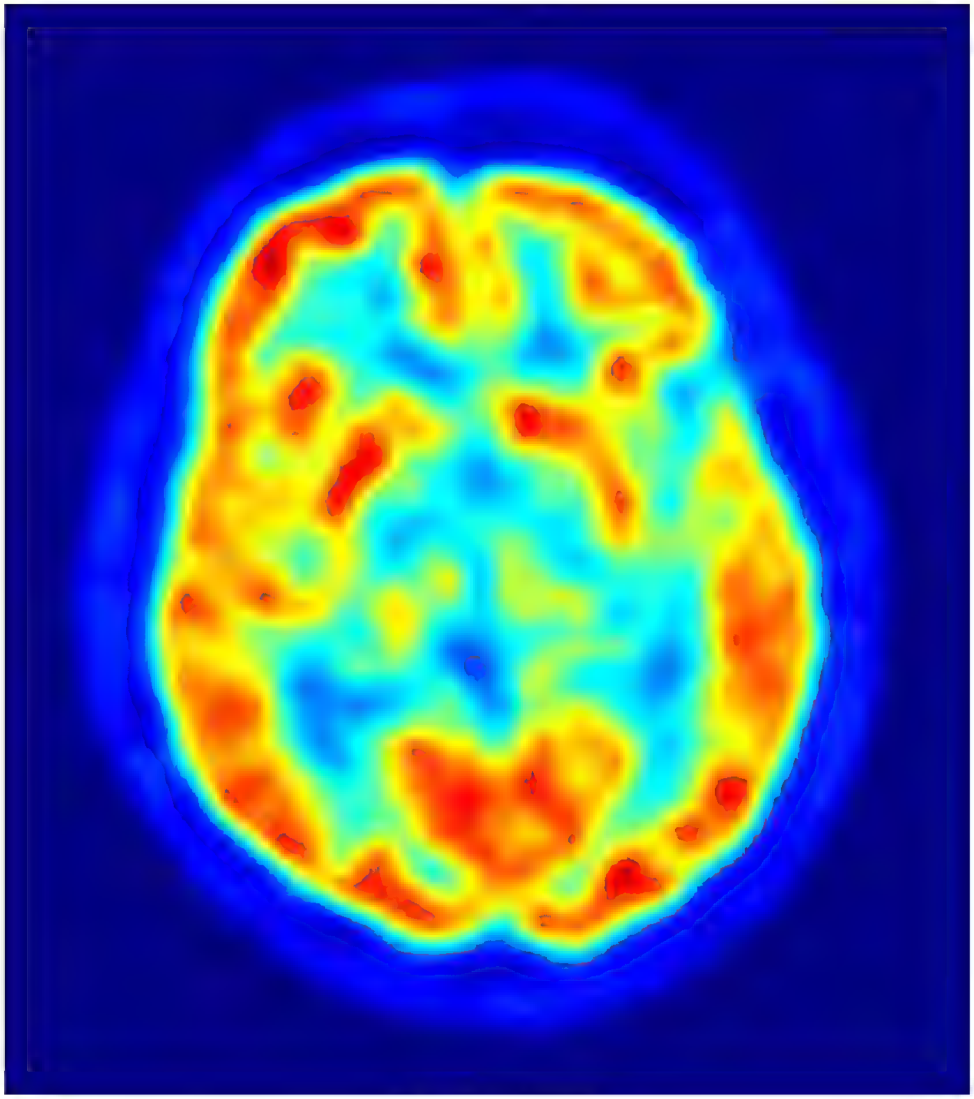


fMRI

second later, and so on down to your feet. What you see at any one time is a slice through a specific section of the body. You can then build up 3D images from these slices and look at the soft watery tissue rather than the hard bone you can see with an X-ray.

MRI scans give detailed images but today there are many more imaging tricks you can play. Give the patient gadolinium to eat – a type of paramagnetic material – and this contrast agent will highlight active parts of the brain. You can 'see' which parts are active: the location of emotions such as love, joy and even the effect of smells as the brain experiences things. This is still coarse grained information; it shows only the general area of excitation and it does not tell us what is going on at the nerve level, but the images are fascinating.

Another recent development in imaging is the diffusion MRI. If you remember your school physics, molecules travel with a random walk: they diffuse along pathways just as people wander along a corridor. If the corridor is full of people, they are jostled around and make little progress. If the corridor is empty, they move in straight lines. This difference in jostling affects the reading in an MRI and allows you to color code the image according to the rate of motion of water along the pathways. You can therefore 'see' the rate at which signals flow in the brain and not only locate thoughts, but also see the links between them.



Functional PET

PET

The last scan we will look at is functional positron emission tomography, or f-PET. In this machine the scanner detects positrons given off by excited oxygen atoms.

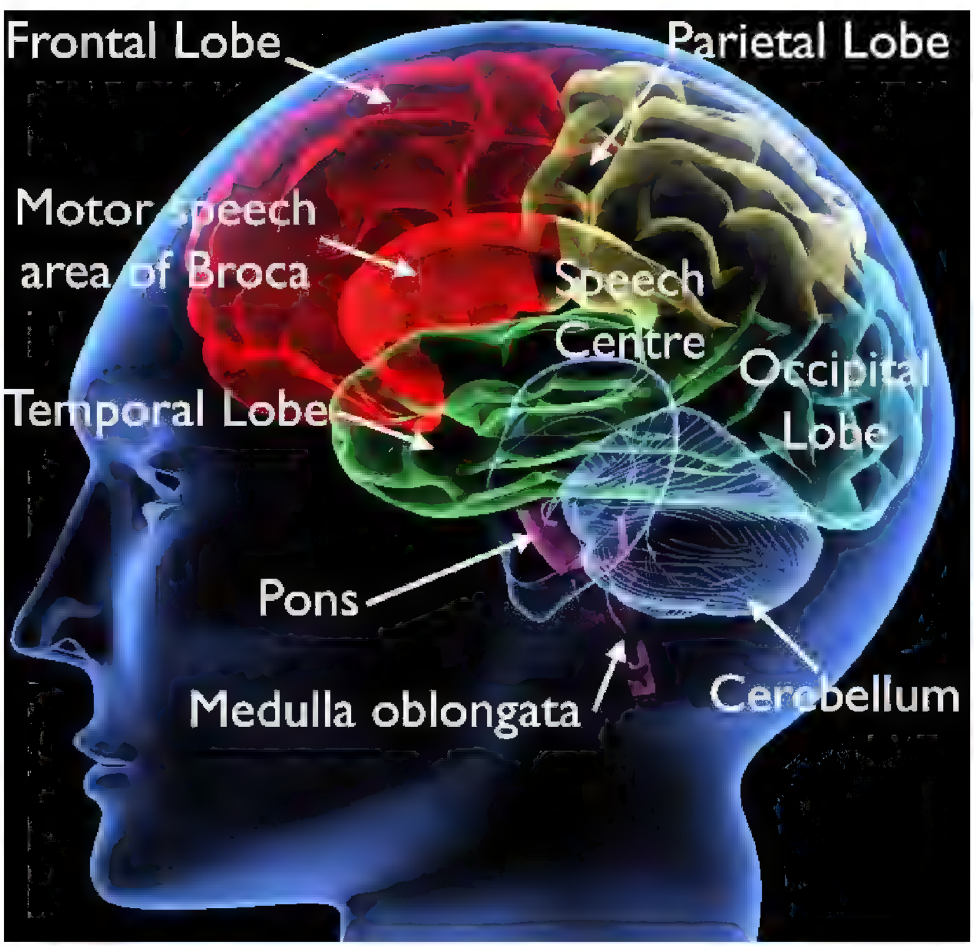
As you think, you burn glucose by combining it with oxygen. The parts of the brain that are thinking hard use a great deal of oxygen and this shows up in scans. Again the consecutive slice trick is used to generate a 3D image that allows you to fly through the brain as it works on a problem.

There is one problem common to all these methods. X-rays, MRI and PET scans only show us the location of thoughts with an accuracy of a few millimeters. Each pixel in the image contains around 10 million neurons, so we can't see the details of thought. For a scale comparison it is like looking at a car factory from space. You can see cars and people going into the factory but you can't read the owner's manual. We need to be able to see at least 10 million times more detail than our current technologies allow to *see a thought*.

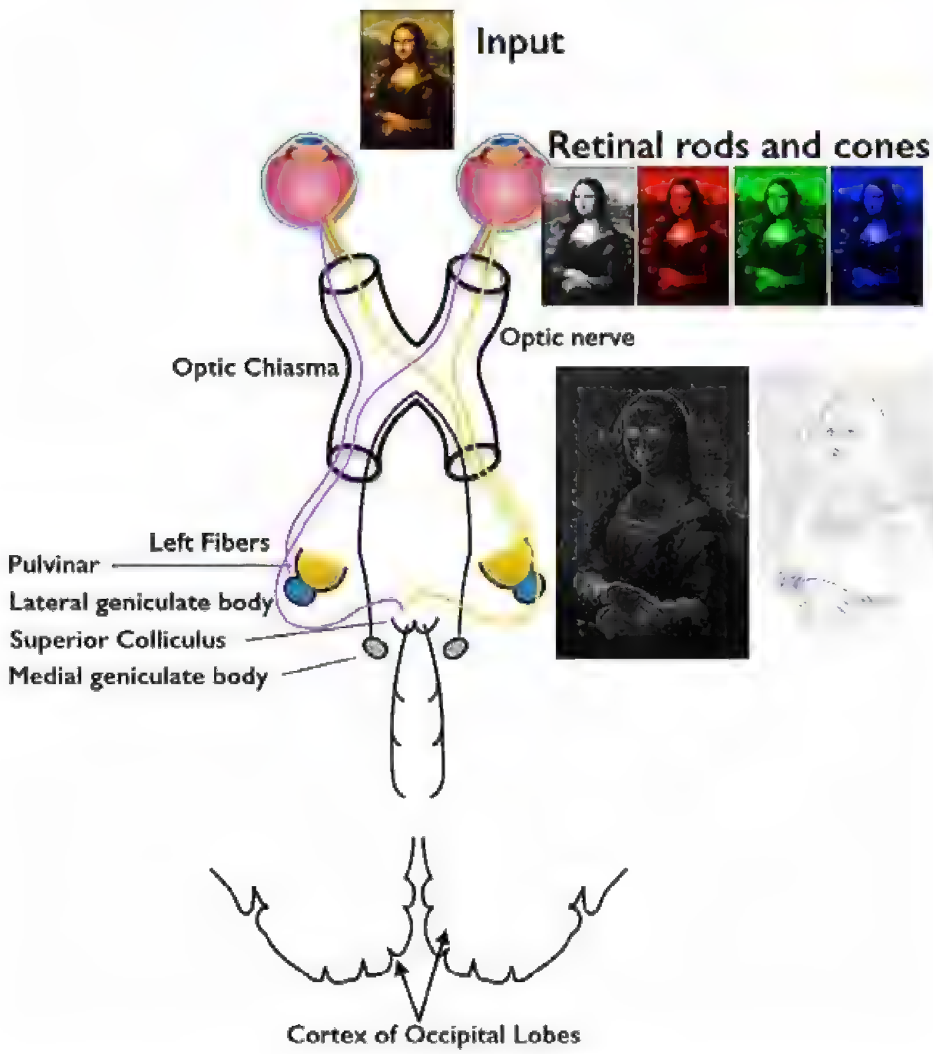
A Quick Tour

Now that we understand how to look inside the brain, let's take a tour around it. The brain is a highly distributed thinking machine. Some things, such as hearing, are located in specific places while others, like the enjoyment of music, are spread out.

Our eyes work as an extension of the brain and use a specialized type of nerve cell. Light falls on the retina and stimulates these cells, causing nerve impulses to run along the optic nerve into the center of the



The Brain



Visual Processing System

brain. The impulses split and form two distinct paths, one through the cerebral cortex, which gives us the sensation of conscious vision, and the other into the lower brain which provides us with instinctive reactions.

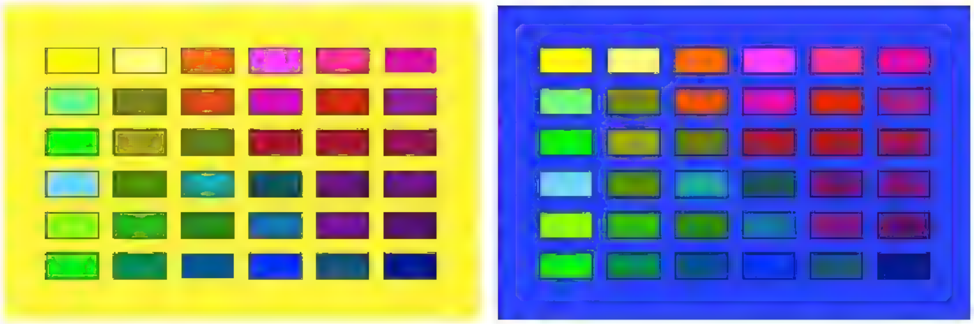
The right hand side of your body is connected to the left hemisphere of the brain and vice versa. This means each hand is controlled by the opposite side of the brain. But, your eyes see both your hands. To resolve this conundrum a very complex thing has to happen to the optic nerve in the center of the brain. The optic nerve from each eye splits and crosses over in the middle, so the left side of the left eye and the left side of the right eye goes to the right hand side of the brain and vice versa. This keeps the brain focused on the correct hand.

Frogs Eyes are Very Sensitive

The processing power of the eye is staggering. The human retina has about 120 million rods and 7 million cones, giving it an average resolution of 10,000 by 10,000 pixels. Each rod is sensitive to individual photons but we register light consciously only if we see around 5 7 photons. It is thought frogs can react to single photons because of the chemistry of their eyes and the fact they are cold-blooded, but this is not proven.

Some animals, including some frogs and my cat, have a tapetum lucidum. This is a reflective backing to the eye that allows each photon two chances to react with a rod, once on the way in and, if that fails, once on the way out. This is why you can see the eyes of some animals if you shine a light into the forest on a dark night. Cones are less sensitive than rods but give us color perception. In the human eye, there are three types of cone: a red, a green and a blue, giving us trichromatic vision. We see colors because light stimulates more than one types of cell and we infer the color in between. A fourth type of cone is present in some species such as birds, reptiles, and fish. This gives them tetra-chromic vision, allowing them to see into the ultraviolet range. It is speculated some humans might have this ability but so far none has come forward. Some animals lack the ability to see certain colors. Most dogs can't see red. This gives cats a big advantage!

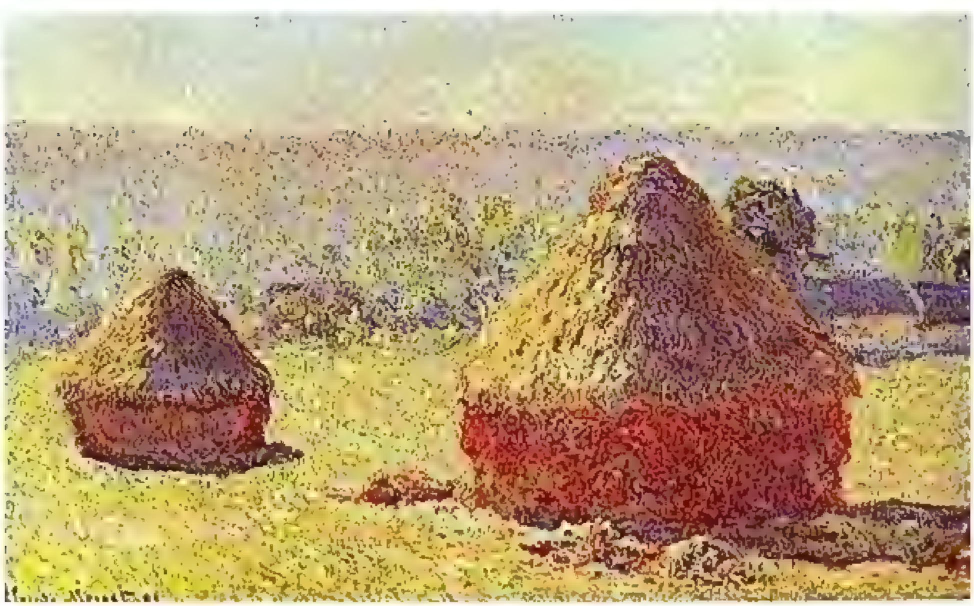
Many people wonder if we all see the same color as each other. Is your red the same as mine? The brain's perception of color is complex. Although the color red is absolute and can be detected by a calibrated sensor, our perception of color is relative. We perceive them in the context



Color is Relative

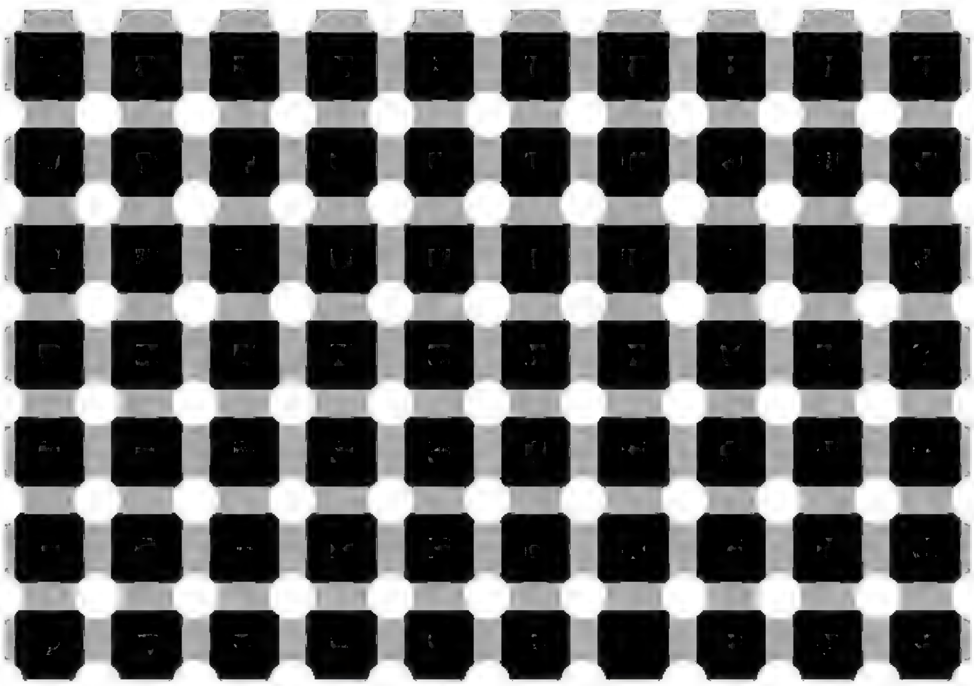
of other colors – not in isolation. The two panels above contain identical blocks of color but they look very different against the background. Check out the website if you have a black and white book. It is an irrelevant question to ask if my red is the same as yours, since *my* red against one background is not even the same as my red against another.

People generally agree on naming colors but not all languages have the eleven specifically named colors of modern English: black, blue, brown, gray, green, orange, pink, purple, red, white, yellow, if you are interested. Ancient Celtic languages, so called 'gru' languages, recognized only four colors and other languages don't distinguish purple from blue. Color, or at least the naming of color, is a cultural thing.



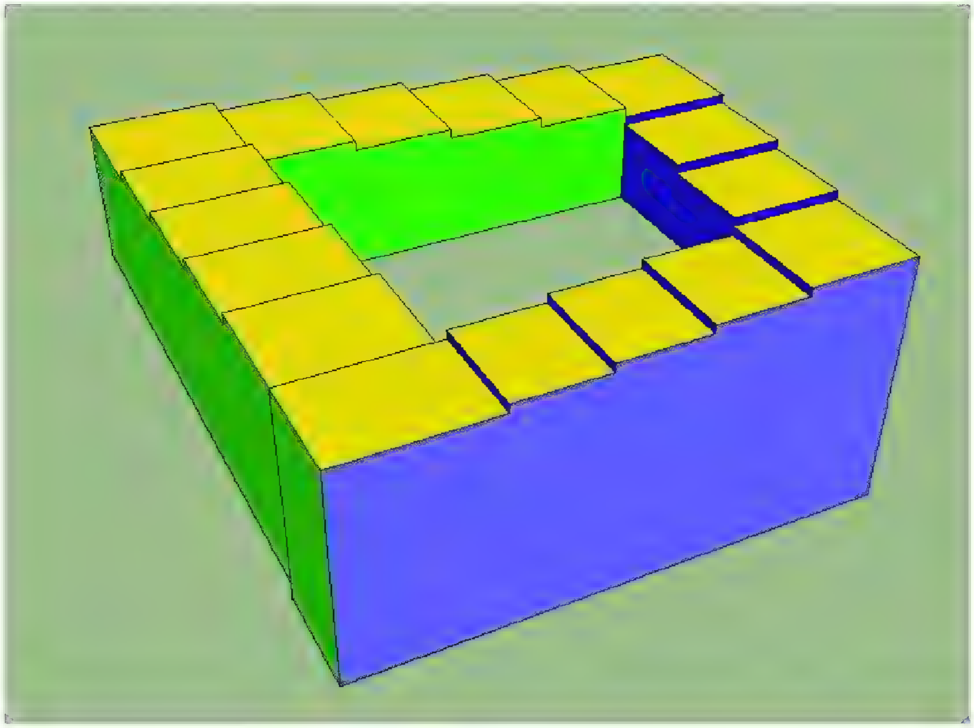
Impressionist Painting, Monet Haystack

The resolution of the eye is not the same across the image. High resolution is concentrated in the center, while lower resolution black and white vision dominates the edge. This peripheral vision helps us detect predators or play football but it is not the focus of our attention. When we focus our attention on something, we turn our eyes to look at it directly. The central part of our eye is called the fovea centralis and is composed of cones. About half our cones are concentrated in this very small section and this gives us immense visual acuity. For a computer display to outperform this section of the eye it would need one billion by



Scintillating Dots Optical Illusion

one billion pixels. The fovea centralis is tiny, only two degrees across, so our eyes must dart around the image to take in all the detail. Once the basic information is encoded in our retina and sent down the optic nerve, it goes into a production line process in the visual cortex where all the elements are analyzed. Our brains extract information from the image such as texture, edges and depth perception in specialized portions of the brain. Because of this specialization it is possible to play tricks on the brain with images that are not easy to process. Some we find pleasurable, while others can be a little disturbing.



Penrose Steps

Optical Illusions

This picture is an illusion that plays with your stereoscopic synthesis. The dots appears to flip between black and white. Other illusions play with depth perception. The Penrose Steps are a type of illusion that tries to build an impossible physical model in our cerebral cortex. The brain sees perspective and depth perception cues, but the resulting shape could never exist.

Hearing

Unlike sight, hearing is an absolute sense. Our ears capture and focus sound down to the eardrum where a set of small hairs called cilia convert it into electrical impulses. The impulses stimulate cells corresponding to specific pitches.

We are born with perfect pitch, yet most of us lose it early on. When I hear Maria Carey sing a top B flat a specific set of neurons located near the ear fires, and if she sings a top 'A' then a different clump of neurons are stimulated. By the time most children come to learn music they have edited out this absolute pitch information. One group of children who do not lose the ability are Chinese pianists. Because Chinese is a tonal language – where the pitch of words affects their meaning – and because

McGurk Effect; Go to the Website and Watch the Linked Video

Chinese children tend to learn the piano very young, they don't lose the absolute part of pitch. An astonishing 93% of these children develop and retain perfect pitch throughout their lives.

There are many cross connections between the audio and video processing systems. At parties you often can't hear speakers clearly because of the background noise. Watching their lips will help comprehension, but which sense wins if there is conflict between the two? The McGurk effect shows this.

To test the effect, go to the website, watch the video and see if you can distinguish when a speaker talking normally and when he is making the mouth movement of another sound. There is a winner. Try it for yourself; check out the link on my website.

Once upon a time people imagined the brain was like a camera forming an image of the world, but if this were the case there would be a paradox. Who is looking at the image in our brain to make sense of it? Modern research shows we don't take a complete picture of the world like a camera but rather parse the image into its constituent parts on the fly.

If someone asks, "Which side of the house is the tree on?" your brain parses the question and compares it with the image map in your mind's eye. What is the image composed of: trees, houses, sky, grass? Your brain manipulates the linguistic question about the relationship of elements and matches it with the visio-spatial understanding of the image, allowing you to answer the question. You might not have to answer

Humans' Ability to Concentrate

the question verbally. If you hit a baseball, no language is involved; you distinguish the ball from the background and perform quite a feat of tracking and calculation to connect it with your bat.

Because the brain is editing the scene on the fly to keep within its processing power, the eye only sees what it turns its attention to. Magicians take advantage of this to play amazing tricks on us. Watch the video on the web and then tell me what you see.

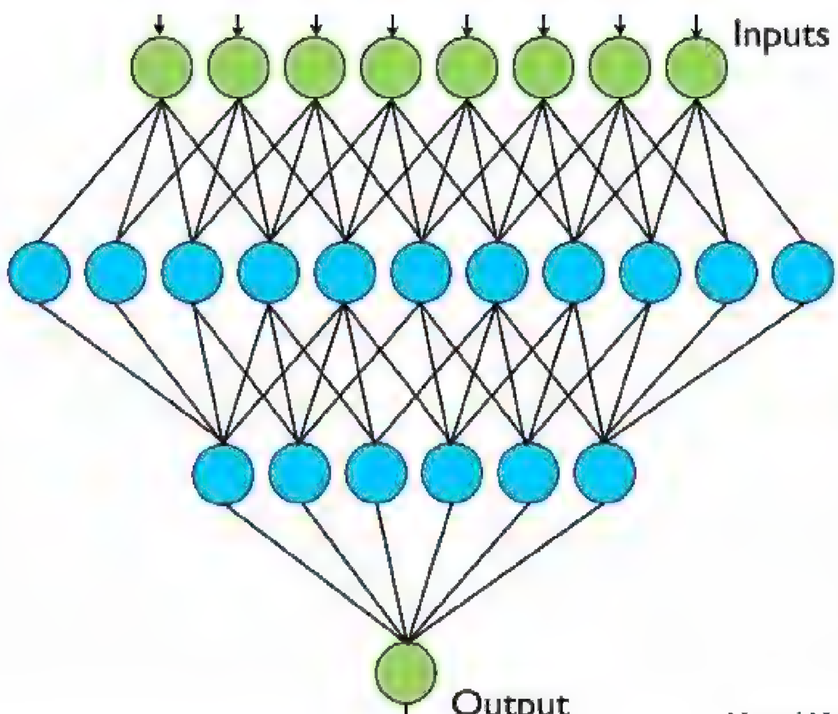
VISIT THE WEB AND VIEW THE VIDEO TO SEE WHAT HAPPENS

Tiger Woods Swing

You can see just how intensively the brain works on a given problem, throwing away all unnecessary information.

The brain contains mirror neurons, a type of brain cell that responds when we see another human do something. These neurons fire as if we were performing the action ourselves even though we are merely witnessing it. It is one of the ways we learn a skill. If I watch Tiger Woods's golf swing, my mirror neurons will fire as if I were practicing *his* swing. Later when I practice the swing for real, my neurons will have already been partially programmed. This effect is presumably the reason we enjoy watching sports; our mirror neurons allow us to begin acquiring a skill while sitting in an armchair! This is clearly a useful evolutionary trait but you *do* also need to practice for real!

Mirror neurons also fire in response to witnessing emotions. When we see an actor laugh or cry, we experience their emotion as if for real. This helps us empathize with the person we are watching and is part of the reason we enjoy movies and plays.



Neural Network

Thinking

"We cannot solve our problems with the same thinking we used when we created them."

Albert Einstein

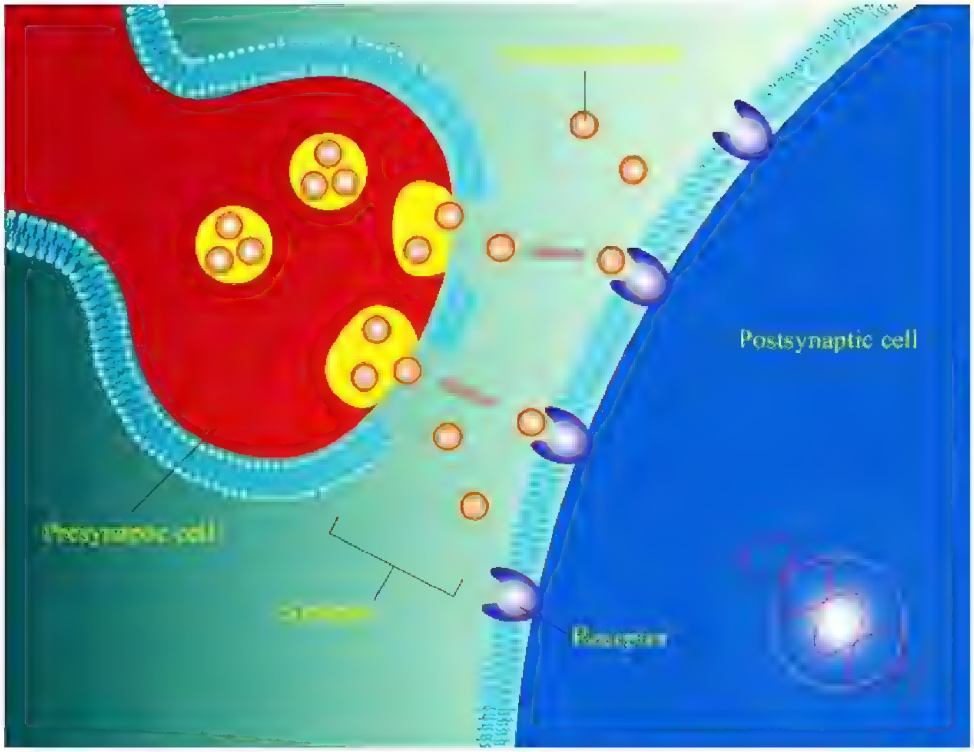
If you feel mentally exhausted reading this book, don't worry. This is normal. Mental work takes energy. Scientists estimate the brain consumes 20% of our resting energy; around 12 watts. Physical fitness is important for thinking. If you get out of breath running for a bus, thinking is going to be harder for you. Studies are mixed about whether the additional work involved in solving a difficult problem causes you to use more energy. We certainly see an increase in the flow of glucose to the appropriate part of the brain, but the overall energy use in the brain is quite high in the first place, so it is hard to see the incremental effect.

Unlike muscles, which store energy locally as glycogen, brain cells 'burn' glucose and oxygen from the blood stream in real time. If scientists detect glucose and oxygen flowing to a part of the brain they know it must be working on a problem. As we know, there are several ways to make glucose and oxygen show up in brain scanners. You can, therefore, inject someone with the right chemical markers, wheel them into a brain scanner, and watch them learn new skills. On a practical level, there is limited space within a scanner and you can't wield a golf club, for example. Julien Doyon, a researcher at the University of Montreal, was recounting this problem to a friend and she suggested knitting. Knitting is a physical activity you learn just like a golf swing or a tennis stroke, with all the initial fumbles and jerky activity, settling down to a fluid learned skill. Most experienced knitters can engage in a full conversation while knitting complex patterns, only needing to break off and concentrate during a pattern change. Luckily, there are ceramic and bamboo knitting needles which don't interfere with MRI scanners, and they are small – no golf swing problems here. Studies of knitters show that when they initially learn a skill, several areas of their brain light up, but after a while, the brain activity becomes concentrated in the sensorimotor striatal territory.

Glucose, the brain's power source, is a sugar we get directly from eating sweets or indirectly by digesting starch. Some studies show children do slightly better at school if they eat starchy foods in the morning for breakfast – a bowl of cereal or porridge. When you think and work your brain consumes the glucose in your blood, and blood glucose level drops. If there is a steady source of glucose from the starch digesting in your gut, the glucose is constantly topped up and the level will stay high. If there is no input of glucose from your gut, the body will first get glucose from glycogen in your liver or generate it by converting fat reserves. This takes more work so the body tends to avoid doing so until it absolutely has to. You can function with slightly lower glucose levels but the body will shut down a little. One thing that suffers as a result is the brain's ability to perform cognitive tasks. A quick and easy way to fix this is to consume some raw glucose and most fridges have a ready supply in the form of sugary drinks. Stories of kids running amok, due to sugar highs brought on by too many sweets and sodas, appear to be an urban legend. In tests, parents told their children have had a sugar drink report them to be hyperactive even if they had been given a sugar free drink. I'm not suggesting you drink lots of sugary drinks – it is bad for your teeth and will make you fat – but the occasional soda is fine.

Memory

Scientists are just beginning to explore the mechanisms that lay down memory in the brain. There are two main classes of theory. The first believes memory is formed in the large scale wiring of the brain. Neurons connect with other neurons and the number and strength of these connections cause memory. When we learn, new connections are formed. The electrical activity in a given part of the brain triggers the formation of new dendrites. They grow, piloted by tubulin micro-tubes, rather like vines growing in a slow motion nature clip. Once a micro-tube guided filament is close enough to other, a synapse forms. This gross-scale wiring growth is one method of memory formation. Another gross-scale effect is myelination. Myelin is the insulation the body uses on nerves cells, including nerve cells in the brain. It looks a bit like the insulation we used in the 1930s. Before the invention of plastic, strips of waxed canvas were wrapped around wires to provide insulation. Myelin



Synapse

has a similar structure. It is a flat protein laid down as a spiral on the outside of nerve cells. The theory is that cell firing causes myelination, which permanently imprints the memory.

The alternate class of theory proposes memory is encoded at a much smaller scale. Neurons are quite complex structures in their own right. Inside each neuron is a lattice of proteins, which forms a skeleton. Part of that skeleton provides structural integrity to the neuron, while other elements provide control and motility. It is this control part of the skeleton that people believe might encode memory. A 2012 paper by Travis Craddock and Jack Tuszynski of the University of Alberta, and anesthesiologist Stuart Hameroff of the University of Arizona proposes a protein called CaMKII binds to the cytoskeleton in 32 different configurations, providing a binary data encoding. It is an elegant idea but it also relies on your believing their model for quantum neuron processing which is still highly controversial. If proven, they are my top Nobel Prize tip for this decade!

Photographic Memory

Until recently conventional wisdom held that true photographic memory was a myth and the few people claiming to have it really used some sort of mnemonic memory technique to selectively memorize things. The most famous case was a Russian journalist known as 'S'. He habitually memorized things using association with places. In antiquity this was taught as 'the method of loci'. The unusual thing was his inability to turn the effect off, and he found it as much a curse as a blessing. He was unable to forget useless information and found it hard to interpret complex images, tending to see areas of color and shade rather than objects such as trees, houses and fields.

Very recently some people have come forward, six in America and one in the UK, who appear to have genuine photographic memories It is well worth watching the TV documentary *The Boy Who Can't Forget* to gain a sense of what this is like. These people appear to lack the ability to forget, and this turns our understanding of memory on its head. It seems memory might work the opposite way we thought. We had previously thought we only remember what we pay attention to, but perhaps we must actively forget, and this ability is missing in these subjects. Scientists are studying these people to see if they can understand more about memory.

The Aging Brain

We can explode a myth and encourage older readers simultaneously. Memory does not deteriorate with age, or at least not until we are very old. Most studies looking at memory deterioration focus on the very old and compare them with the very young. Even then, the differences are small. When people are asked to attempt memory problems there is a mild drop off with age but the results are quite similar. The most likely reason older people don't remember so well is they don't believe they can. Perhaps they don't have as much incentive to remember new information. Why learn someone's name if you're unlikely to meet them again? Since IQ actually increases with age, don't believe people when they say you are going downhill from the age of 40. You are not!

Computer Brains

Computers are really quite simple compared with all the evolved baggage we humans carry around. When a computer is presented with instructions, for example, for a program like Excel and a file such as my expenses, it will load everything into memory and 'run' it. The process of running a program is simple. Each instruction is a number. The computer reads the number, looks it up in a table, finds a corresponding number, and writes that down. Essentially that's all there is to it. From a simple mechanism like this, we get the enormous complexity of a modern computer. The sophistication is achieved through reading and writing many numbers in parallel, and chaining the steps together so that if you read a particular number it triggers another read/write process, and so on. I'm glossing over some details such as logical functions but, if you know how a modern computer chip is constructed, my description is not far off. Almost all logic today is implemented in tables to achieve the speeds we expect from modern chips.

All modern computers are clocked. A small piece of quartz rock has been polished, coated with metal, and wired up to a control circuit in the computer. When you apply voltage to the rock it bends and absorbs energy. When the voltage is taken away it bends back and gives out the energy. This is effectively a pendulum and it can be used to make an accurate clock. I used to design these for a living. Every logic gate in a computer is connected to this clock, and each time the clock ticks the logic gates in a computer *compute*.

Most modern computers are entirely synchronous. The clock rate is set so that the gates in the computer fully recover by the time of the next tick, and every gate is therefore ready in its standard position when the next instruction arrives. The human brain does not have a central clock. Each neuron acts independently – firing regardless of whether the neurons it is adjacent to are ready or not. It is wrong to think of the brain as digital. Each neuron *does* fire and recover, but it may be triggered again before it fully recovers. This makes for a chaotic and essentially analogue operation. If one neuron fires when a second has only half recovered, then it gets half an effect. If the neuron is 80% recovered, an 80% effect. Neuron recovery time is quite long, perhaps as much as $1/1000^{th}$ of a second, and they are wired in three dimensions to as many as 10,000 other neurons. It is perfectly possible for a set of neurons to run one 'program' when they are rested and a completely different 'program' when they are 50% recovered and yet another programs if triggered from different starting locations. I have said 'program,' but arguing a brain runs a 'program' is misleading. It is not organized like this.

Neural Networks

A neural network is our best attempt at a computer model for the human brain. Each neuron is represented by an entry in a table. The entry records all the connections to it, along with the strength of each connection – these are called 'weights'. In some models the connections can be both inhibitors and activators like in real synapses. An individual neuron will fire if the sum of all the connections multiplied by the weights reaches a certain pre-determined threshold.

A neural network does not run a program in the conventional sense, and must be trained through experience rather like a human brain. The training process allows the weights in the network table to be adjusted to give the correct result. But, unlike the brain, you can read the weights and even save them to a disk. The neural network tables start with random settings. You show the network the letter 'A' and adjust the weights in the tables until it gives a positive answer: 'It's an A'. Repeat the process with the other letters until the network correctly distinguishes them. As you do this a computer algorithm constantly adjusts the weighting tables using a method called 'back propagation'.

At the end of the training process you can show the network some new input and see how it does. For example, a letter 'A' that is in a slightly different font to anything in the training set. Trained neural networks can perform complex tasks such as recognizing faces or making clinical diagnoses, and they can be allowed to modify their weighting tables as they work so they learn from experience in a similar way to a human brain. Strong AI proponents believe making a thinking machine is just a matter of building a really large, fast neural network and working out how to train it efficiently.

Quantum Brains

Conventional wisdom says each brain cell is a single processing unit making an on-off decision – fire, or don't fire – depending on the state of its neighbors. But, Stuart Hameroff, Professor of Anesthesiology at the University of Arizona, thinks neurons are not the fundamental information-processing unit in the brain. He suggests that this accolade should go to tubulin. Tubulin is a small, versatile protein that self-assembles into filaments rather like the way buckyballs – a magnetic children's toy – can be arranged. There are two types of tubulin molecule: α and β. They slot together and wrap around to form a micro tube about 25nm in diameter.

Tubulin micro tubes do several important things in the body. They form the skeleton of neurons and give them structure. They are involved in guiding neurons as they grow towards each other to form new connections, and they also operate in the nucleus of a cell to unzip

Paramecium

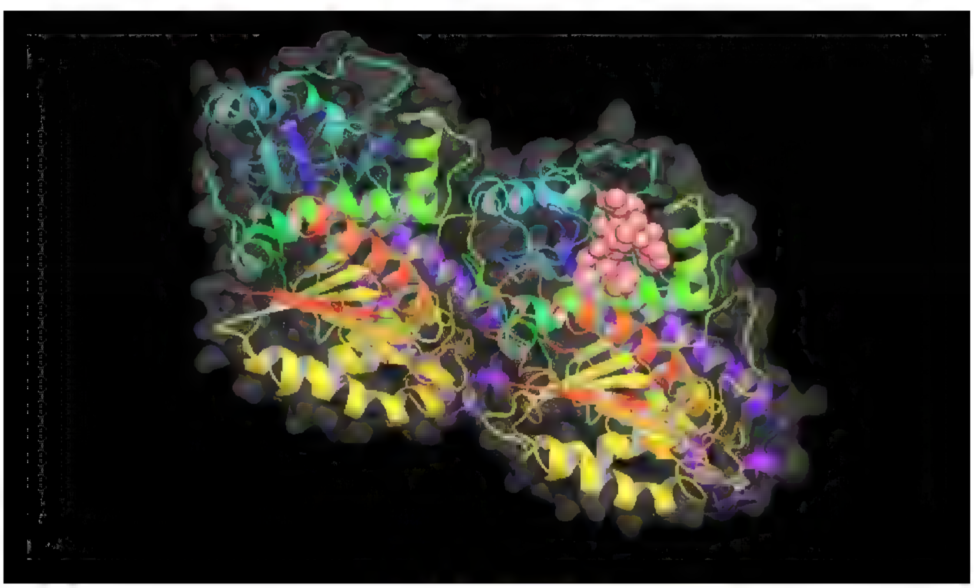
DNA into its two complementary strands when a cell divides. In single-celled organisms, including paramecium, the ends of the tubes stick out of the body and form the cilia that drive the organism along.

The presence of tubulin in complex, single-celled organisms provides a clue that the smallest information processing unit might not be the neuron. Some single cell organisms, such as paramecium, display complex behavior: hunting for prey and escaping danger. This suggests they can process small amounts of information without the need for a matrix of neurons. Since we evolved from these organisms, why wouldn't our brain cells take advantage of this sub-cellular intelligence?

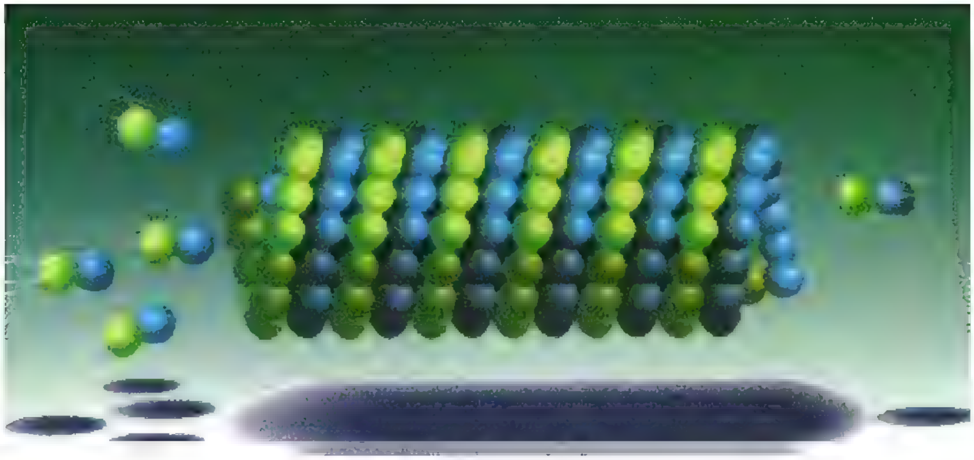
The structure of tubulin lends itself to digital processing as the molecules forming the walls have two stable states and can flip between them. We might recognize this as the basis of a binary computer, and cells might have little computers within them. They would not need to process many bits to be useful. Perhaps single-cell organisms developed information processing capabilities in their micro tube structures that allowed them to better survive and, as their nervous systems evolved, they coupled these structures to form the brains we see today. This piece of theory is not too controversial. After all, nerves have wiring within them to carry information to the synapses and it's likely this wiring is involved in the thinking process. But Hameroff is not finished. He has teamed up with Roger Penrose to bring quantum mechanics into the picture.

Their reasoning is straightforward but has generated a great deal of controversy. Hameroff observes that anesthetics cause humans to lose consciousness by binding to tubulin, but they do not halt all brain function. He, therefore, concludes our conscious thinking is mediated by tubulin, not the larger scale firing of the neurons. Penrose had been looking for a mechanism in the brain that would explain how brains solve non-computational problems. Together Penrose and Hameroff propose tubulin micro tubes are quantum gravity computers that allow us to think non-computationally and are the seat of consciousness. The ideas are still being worked.

Penrose and Hameroff have a difficult task conveying their ideas to the rest of the scientific community. Scientists don't recognize a need for something that can think non computably, so they are highly skeptical of a mechanism which performs that sort of thought. The latest development on the Hameroff Penrose model comes in the work of Travis Craddock, now of Nova Southeastern University, Florida, and others. They have written a paper arguing signals propagate according to quantum principles within microtubules through the excitation of thiamin molecules along the length of the tube. They believe these molecules are quantum, entangled in a similar manner to the mechanism recently discovered in photosynthesis. The geometry of these molecules is set out in a similar way to the active areas in chlorophyll and they have a complementary problems to solve. Chlorophyll tries to maximize energy conversion efficiency, while a microtubule tries to minimize the use of energy while propagating signals along a nerve. You might wonder



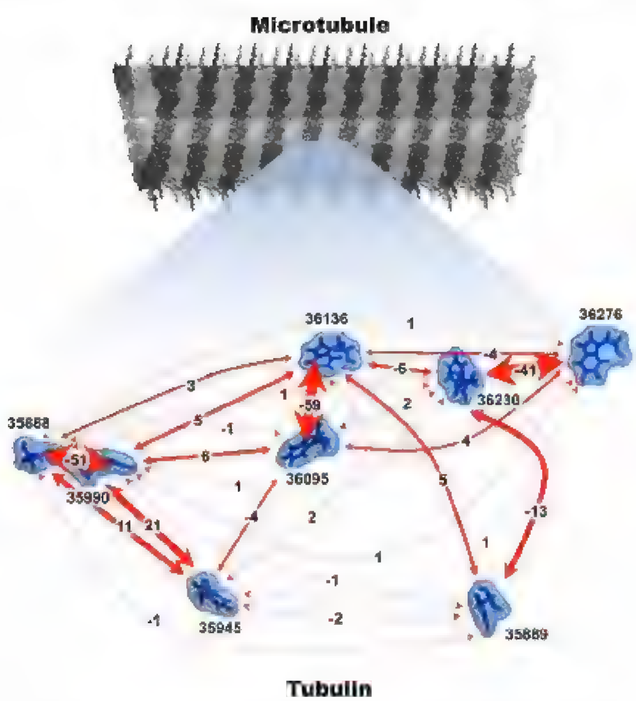
Tubulin Protein



Tubulin

where the light comes from since tubulin is housed deep within the neurons inside our brains and shielded from light by our skull. It turns out that the mitochondria which powers our bodies emit photons of UV light as a waste product of their metabolism. The speculation is tubulin harvests this waste energy.

Before we argue for this mechanism any further we still need to establish that a non-computational mechanism is needed to allow human thought. In the next chapters, we will look at the nature of knowledge and, in particular, mathematical creativity and the Wiles Paradox.



Quantum Coupling of Tubulin in Microtubule

Chapter 5

KNOWLEDGE

Chimpanzee and Typewriter

"There's an infinite number of monkeys outside who want to talk to us about this script for Hamlet they've worked out."

Douglas Adams

"I'm not young enough to know everything."

J.M. Barrie

"He has Van Gogh's ear for music."

Billy Wilder

Could an army of monkeys write *Hamlet* by bashing away randomly on typewriters? Of course, we don't mean this literally. We are asking whether knowledge can be created without understanding. Can a monkey, or perhaps some form of computerized random number generator, accidentally type out the script for Shakespeare's *Hamlet* or write Tolstoy's *War and Peace*? Is knowledge generation simply a numbers game?

Leo Tolstoy's *War and Peace* is generally assumed to be the longest novel ever written. This is not quite true. Wikipedia reckons the longest novel is a French book, *Artamène*, with over 2.1 million words. Tolstoy comes in sixteenth, with a mere half million!

Written in 1869, *War and Peace* tells the story of five Russian families during the Napoleonic wars. Originally written in a mixture of Russian and French, and numbering over 500,000 words, it was quickly translated to other languages. The mistress of composer Franz Liszt translated it fully into French, where it expands to 550,000 words. Contrary to popular myth the length of the book drops slightly in German. If you really want to save paper Chinese is best. Because it uses a single symbol per word, the Chinese translation needs only 750,000 characters compared with the 3 million for English. It is wrong to assume this is necessarily more efficient than a phonetic language. Although it might save on paper, it is considerably more laborious to write. Three strokes are required to write 'war' in English whereas the Chinese pictogram requires ten.

是

War in Chinese

Computers work with numbers. It is a simple process to translate a book into numbers because books are composed of discrete symbols. All we need do is give each symbol a unique number and record those numbers in digital format. Artistic works involving pictures and sound are more difficult to represent because they are continuous in nature. We have to digitize them first. With music or painting this inevitably means some loss of information as we can't cut a sound or image into an infinite number of pieces.

The modern standard for translating text to numbers is Unicode. Each character is represented by a five-digit number ranging from 1 to 64,000 – two bytes for those of you who know computing. This is sufficient to code almost all the world's symbols, so we can avoid any accusation of being language-ist! Here are some examples of the

ΦﾃあQボ杏確ПЖЯ

Ancient Greek, Japanese: Kanji, Katakana, Chinese, and Russia-Cyrillic Symbols

characters represented by Unicode.

For our discussion, it does not matter which language *War and Peace* is written in. We just treat the symbols as numbers. I am going assume the English translation which has around 500,000 words; a nice round number. Assuming a generous 10 characters per word, *War and Peace* is approximately 10-megabytes – that's about the same size as a music track on iTunes. In practice, the book uses a bit more memory, as there is some overhead for formatting information. My laptop has a 500 Gigabyte hard disk so I could fit half a million copies of *War and Peace* on it!

If we take a look at the contents of the file on my computer the book starts:

87101108108448011410511099101441151111

Can a computer calculate this number?

The obvious answer is YES. It is just an integer like 1, 3 or 42. Granted it's a large number, but the length of the number is simply the length of all the symbols in the book coded into Unicode – about 10 million digits. We have already determined this number can be stored on my hard disk half a million times, so it's not an unimaginably large number. How long would it take to calculate the number corresponding to *War and Peace*?

The simplest method is to count up starting at 1 then 2, 3, 4, 5, and so on until I try every number. Will this eventually get to the *War and Peace* number? The answer is yes. Eureka! All of human knowledge is computable. I have written this computation out as a simple computer program below. It says, in plain English, start at zero, go round a loop counting up one at a time and print each number as you go along.

i==0; Loop i++ Print i;

Easy!

No, unfortunately. The problem is subtler than it first appears. First it will take a VEEEEERRRY long time. If I counted up from one, I would print out *War and Peace* eventually but it would take 120 billion, billion, billion, billion, billion… (I would need the entire length of this book to write out all the billions) years! For the physicists amongst you, I would need $10^{30,000}$ years, assuming I could use every atom in the known universe counting in parallel at the plank interval. 'The plank interval' is the shortest time that can exist in the Universe as a discrete 'tick'.

Even going at this speed using with every atom in the known Universe would take $10^{5,000}$ longer than the age of the Universe. This is stupendously long. Remember scientific notation means I have a 1 with 5000 zeroes after it. It is a deceptive notation as something as innocuous as 10^{120} is equal to the number of atoms in the known universe. $10^{5,000}$ is an absolutely enormous number. If you hear something is going to 'take until the end of time', we're talking a lot longer than that!

You may have spotted that in the process of counting up to the *War and Peace* number we also count through EVERY book ever written shorter than 500,000 words in all the world's languages. Interestingly we counted through the Japanese and Chinese translations of *War and Peace* quite a bit before we reached the English and finally French translations. During the process, we also stepped through countless other wonderful works: proofs of amazing theorems, the complete works of William Shakespeare, and every composition ever written. Sadly, we never knew it. The problem is my program never stopped and told me it had found any of these wonderful things. I would have to sit staring at the screen to spot them. If I was off doing something else – making a cup of tea, taking the kids to school – I would miss all these wonders; the program never *tells* me if it has succeeded, but quietly prints out *War and Peace* and carries on. This is really annoying. It's not a useful machine.

What I need is a machine that rings a bell when it finds something interesting so I can break away from what I am doing and take a look. Reading every book it writes in every language and all the nonsense in between would take a ginormous amount of my time. (By-the-way, contrary to statements by school teachers that ginormous is not a word – it is!) I want a computer to come up with *War and Peace* without me having to do all the work.

It's no help if the machine writes everything down and lets me take a look in my own good time. That only puts off the time when I have to begin reading all the gibberish it produced. Another practical problem is the massive storage required. Just imagine the immense piles of printer paper! Stephen Hawking and Jacob Bekenstein have shown space appears to have a limit to the quantity of information it can store. The quantity of information we are looking at here is greater than the storage capacity of the Universe and would collapse space-time to a black hole before I got even a fraction of the way through. Let us try to be a bit cleverer about the task of creating this information.

The simplest way to tie the computer down is to run a much stricter program. Ask it to count up from one until you get to a number representing the novel War and Peace and then print it, stop and ring a bell.

```
Loop i++ until i == "War and Peace..."; Print i; ring-bell;
```

This program succeeds!

I am triumphant. I have calculated the *War and Peace* number, and this time I did not miss the event. But, if you consider this a little more deeply I gave the computer the answer! I told it the string "*War and Peace…*" and it was able to count up, stop, and tell me it reached it. In mathematical terminology, the program is said to have 'halted' when it reached the *War and Peace* number and in computer science speak it is a special purpose program designed to do only this one thing. This program is pointless. First, it would still take a ginormous amount of time to get there and, second, it is trivially the same as running the program: Print *War and Peace*.

```
i = "War and Peace..."; Print i;
```

It's just the same as me taking my laptop, finding *War and Peace* and pressing *print*. In no way is this equivalent to Leo Tolstoy's creative effort of writing *War and Peace* in the first place.

What went wrong?

I wanted my computer to find an *interesting* string I did not already know. *War and Peace* is trivially computable after Leo Tolstoy created it but the question is whether my computer could come up with *War and Peace* or some similar creative work on its own. Can it *create* and, more importantly, understand it has *created* something? We have linked the ideas of creativity and understanding, and this will prove to be the key to the problem.

The Problem

One suggestion put forward by Daniel Dennett is the creative process is a two-part task – generate ideas, then critically assess them. I can, in principle, make a program write out every possible book less than 500,000 words long. Provided I don't store the results this will not collapse the Universe. This just leaves the problem of writing another program to read all the output and ring a bell each time it finds some interesting truth. This second program might be called an appreciation program. Let's examine this approach. I can write out a very simple program to do this – provided I cheat and ignore the complexity of the term 'something interesting'. In plain English: Count up from one until I get an interesting fact, write it down and stop.

Loop i++ until i == (Something Interesting), Print i

This generates two problems. We need to make a program that can tell if something is interesting and it will need to be fast because it is going to be handed a huge amount of junk. Clearly I have a process running in my brain that can determine if something is interesting, but it is quite slow. It takes me an appreciable time to open a book, leaf through the pages and declare it either junk or interesting. Leo Tolstoy had a process in his brain that allowed him to create something interesting but I want to prove he did not do this by generating random junk and sifting through it. Let's look at the mathematics.

We know simply counting sequentially through every number would take too much time, but why not generate random numbers and run our critical eye over them? Surely this would give a faster result. Let us try with a short poem. How hard would it be to come across something as simple as a four-line poem using this technique?

This poem, by the late Spike Milligan, is only 23 words long, including the title, and I have a powerful computer. Wouldn't it be possible to generate it using a computer? Unfortunately, no. We humans don't have a good head for large numbers and this problem is much harder than it appears. Let's use playing cards to get a feeling for large numbers.

A Simple Poem

Rain

There are holes in the sky
Where the rain gets in
But they're ever so small
That's why the rain is thin.
Spike Milligan

Spike Milligan

Coming upon a poem by chance can be likened to the probability of dealing a perfect bridge hand. Shuffle the deck thoroughly and then deal four hands. What is the probability every player will have the ace through king in a single suit? It's about 1 in 1,000,000,000,000,000 hands. Because lots of people play a lot of bridge around the world, this outcome has been reported quite a few times. The possibility appears within the bounds of human experience. Fifty-two playing cards seems close to the 80 characters that make up this poem and 13 choices of cards is about the same as the 26 letters of the Latin alphabet. Wouldn't we expect poems of this complexity to crop up almost as often?

NO.

The 80 characters of this poem versus the 52 playing cards and the greater choice offered by 26 letters increases the problem geometrically. Taken together the probability of accidentally getting this poem is vastly less than a perfect hand of bridge, 1 in 10^{83} against the perfect bridge hand of 1 in 10^{20}. That's the difference between the number of atoms in the known universe and the number of atoms in a jug of water! Numbers get big very quickly when we are looking at the permutation of information. And there is another problem with our bridge analogy. All the bridge players in the world are part of the machine finding the perfect hand. When a human sees a perfect bridge hand they are amazed. It is an event that usually hits the local newspapers and a couple of years ago one reached the national papers in Britain. Each bridge player looks at every hand, they play so there is a huge amount of processing going on during every bridge game. To replicate this for our poem, we would need millions of poetry classes spending hours each evening reading through computer printouts of gibberish.

I should also add that sightings of perfect bridge hands are almost certainly hoaxes. The probability of it happening even once would require everyone on Earth to play bridge continuously for a thousand years. It is reported somewhere in the world about two or three times a year. If we are charitable, we might assume people failed to shuffle the deck properly but I suspect some mischief is going on! The numbers don't stack up…

You might think the problem is one of improving the efficiency of the filter so humans would only have to examine a smaller number of possibilities. Surely I could improve things by writing a simple program to ban all non-English characters, words and poor grammar; things that don't pass the Microsoft Word grammar checker. This would generate a more manageable number of potential poems.

Lewis Carroll shows this does not work; my idea to use a grammar and spelling checker to filter out gibberish just eliminated Jabberwocky, one of the most famous verses in the English language. Take a look at what Microsoft Word thinks of it.

The Jabberwocky

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.
"Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
The frumious Bandersnatch!"

He took his vorpal sword in hand:
Long time the manxome foe he sought—
So rested he by the Tumtum tree,
And stood awhile in thought.

And as in uffish thought he stood,
The Jabberwock, with eyes of flame,
Came whiffling through the tulgey wood,
And burbled as it came!

One, two! One, two! and through and through
The vorpal blade went snicker-snack!
He left it dead, and with its head
He went galumphing back.

"And hast thou slain the Jabberwock?
Come to my arms, my beamish boy!
O frabjous day! Callooh! Callay!"
He chortled in his joy.

Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

Lewis Carroll

Lewis Carroll's Jabberwocky

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.
"Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
The frumious Bandersnatch!"
He took his vorpal sword in hand:
Long time the manxome foe he sought—
So rested he by the Tumtum tree,
And stood awhile in thought.
And as in uffish thought he stood,
The Jabberwock, with eyes of flame,
Came whiffling through the tulgey wood,
And burbled as it came!
One, two! One, two! and through and through
The vorpal blade went snicker-snack!
He left it dead, and with its head
He went galumphing back.

"And hast thou slain the Jabberwock?
Come to my arms, my beamish boy!
O frabjous day! Callooh! Callay!"
He chortled in his joy.
Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

The Jabberwocky Spell Check

Microsoft Verdict on the Poem

39 of the 166 words in the poem are unknown to Word's spelling checker and this is an optimistic analysis of how the algorithm would fare. Many of the words are in the spelling checker because of the poem: galumphing, for example. Lewis Carroll's work was sufficiently influential that part of

the English language was created in this poem. The same goes for much of Shakespeare. If we used a filter method, we would have just deleted most of Shakespeare from the English language! Indeed half the poems in my anthology of English verse are destined for the waste paper basket due to some minor infraction of 'the rules'. If you want something that completely flummoxes my spelling checker here is the Loch Ness Monster Song by Scottish poet Edwin Morgan. I asked a Scottish friend whether Scottish spelling checkers fared any better and he assures me, no.

The Loch Ness Monster's Song

Sssnnnwhuffffll?

Hnwhuffl hhnnwfl hnfl hfl?

Gdroblboblhobngbl gbl gl g g g g glbgl.
Drublhaflablhaflubhafgabhaflhafl fl fl -

gm grawwwww grf grawf awfgm graw gm.
Hovoplodok - doplodovok - plovodokot
- doplodokosh? Splgraw fok fok
splgrafhatchgabrlgabrl fok splfok!
Zgra kra gka fok!
Grof grawff gahf?
Gombl mbl bl -

blm plm,

blm plm,

blm plm,

blp

Edwin Morgan

The Loch Ness Monster

The foibles of spell checkers have long been a personal pain to me because of my dyslexia. Although I can see the red underlining Microsoft Word kindly inserts so liberally into my text, I can't easily see the occasions when I use a homonym. A fine poem illustrating the problem was kindly written by Jerrold H. Zar and published in *The Journal of Irreproducible Results*. It hangs on the wall behind my computer to remind me to check for these errors.

Candidate for a Pullet Surprise
By Jerrold H. Zar

I have a spelling checker,
It came with my PC.
It plane lee marks four my revue
Miss steaks aye can knot sea.

Eye ran this poem threw it,
Your sure reel glad two no.
Its vary polished in it's weigh.
My checker tolled me sew.

A checker is a bless sing,
It freeze yew lodes of thyme.
It helps me right awl stiles two reed,
And aides me when eye rime.

Each frays come posed up on my screen
Eye trussed too bee a joule.
The checker pours or every word
Too cheque sum spelling rule.

Bee fore a veiling checker's
Hour spelling mite decline,
And if we're lacks oar have a laps,
We wood bee maid too wine.

Butt now bee cause my spelling
Is checked with such grate flare,
Their are know fault's with in my cite,
Of nun eye am a wear.

Now spelling does knot phase me,
It does knot bring a tier.
My pay purrs awl due glad den
With wrapped word's fare as hear.

Too rite with care is quite a feet
Of witch won should bee proud,
And wee mussed dew the best wee can,
Sew flaw's are knot aloud.

Sow ewe can sea why aye dew prays
Such soft wear four pea seas,
And why eye brake in two averse
Buy righting want too pleas.

The Search for Knowledge

I hope this explanation shows you the simplest model for creativity – working through every possibility, and examining them all – is doomed to failure. It would take longer than *until the end of time* to even list all the options, let alone analyze them.

You might wonder just how long it is until the end of time? It's generally assumed there are two possible ends to the Universe, a Big Crunch or heat death. Either way the approximate estimate is our Universe will last somewhere between one and fifty times longer than it has lasted so far. That's a long time, at least another 15 billion years, but just generating *War and Peace* would take 5000 orders of magnitude longer than this!

More complex models such as a three-step process have been suggested. We could perhaps randomly create information and put it through a mechanical filter to bring it down to a manageable set of options and then give it to an appreciation algorithm to finally decide whether we have created something. The real problem with this model is the filters. If we try to reduce the effort by assembling works only from pre-existing words, we will have filtered away many works we know and love. Gone are Shakespeare, Lewis Carroll, Dylan Thomas and Roald Dahl, shall I go on? Indeed, once upon a time there were *no* words, every word was coined at some point. The process of creating art is continually creative and mechanical filters can't be applied to things they have not seen before.

You might argue we could devise a more sophisticated mechanical filter, something that contains an algorithm with an understanding of the rules of language. The problem is both the size of the task and the nature of understanding. If I devised some really good appreciation algorithm which did not delete all the creative words of the English language, it would still have to read and appreciate the huge quantities of input until it hit upon something good. There is no way for any machine to read all this information in the age of our Universe; the numbers are just too large. And there is no way for a machine to understand all the rules of language, they are not written down and constantly evolve.

These descriptions should give you an intuitive feel for nature of the creative problem. If you try to deconstruct it into mechanical steps you end up with either a mechanism that needs to be infinitely specified or one that lets through an infinite quantity of nonsense. A human could never sift through all that garbage to find the occasional pearl of wisdom.

Until the beginning of the 20th century, most people thought knowledge and creativity must be just a matter of scale. A big enough, fast enough machine should be able to solve any problem. But early in the 1930s two mathematicians – Kurt Gödel and Alan Turing – showed knowledge was not so simple. Let me give you a feel for why.

Knowing When You Know

The essence of creating knowledge, is to know when you have done so. In a sense, counting from one to infinity means *I know everything*, and merely counting to 50 million creates every piece of significant symbolic knowledge that will ever be written – all the books, plays, mathematical theorems you could possibly want. But, if I were to list all these numbers in an enormous imaginary book it would hardly constitute knowing everything: I would be awash with numbers but not with knowledge.

The essential feature of 'knowing' is to have a small number of steps that will definitely answer a problem. For example, if I wish to phone someone I can look up their details on my phone. The process will tell me their number in two or three steps. If you tell me the number is somewhere in the phone book this is not knowledge. It could mean I need an infinite number of steps.

If I accidentally deleted all the names in my phone – a nightmare scenario – and just had a print out of numbers would I still 'know' them? Obviously I would recognize my mother's number, but most of them would be useless. To know something, I need link the information to what it is for. A number with a name allows me to predict what will happen if I make a call. I will have an interesting conversation or pay my gas bill. It's the same with most numbers. If I have a number that represents the design for a building or a mathematical theorem, these numbers have purpose. If I input these numbers to a computer along with some building design software or a copy of Mathematica they will do something interesting; allowing a construction firm to build a innovative building or a mathematician to check a theorem is sound.

It's a lot harder to prove numbers representing art are functionally useful. A work of art is in some sense not complete – it still needs to go through the process of being appreciated by someone. We could show it to a friend or exhibit it in a gallery but this is un unpredictable process. Van Gogh's paintings were so criticized in his lifetime, many people would have denied them the label art, and Edwin Morgan's Loch Ness

Art or Information

Monster poem is almost pure gibberish, but it's undoubtedly art. Art is a tricky problem but, in practice, most of us agree on what constitutes good and bad art. We will look again at art, in Chapter 10.

Classically we assume knowledge is discovered through random chance and iteration. To understand how this might work let's lay out the world's information in a way we can visualize. Imagine every piece of discoverable knowledge could be found in an infinitely large library.

The infinity library would contain every possible symphony, theorem, novel, poem, and play ever written, or to be written. Its sister library next door, the continuum library, would contain all the analogue works of art; painting, sculpture, architecture, physical artifacts and the like. The curators of the two libraries would constantly argue over whose collection was the better. We'll leave them to differ for the moment. The infinity library is interesting enough so let's explore it first. After all, its sister, the continuum library, takes an infinite amount of time just to look at the first room, and we are in a hurry!

Although the infinity library is infinite, we are probably only concerned with entries shorter than a million symbols. All the interesting papers, proofs and symphonies I know of are shorter than this. If I wanted to include all computer programs, I would still only need to increase it to 100 million symbols. Looking for knowledge is *not* itself an infinite task.

For the sake of clarity, I will ask the infinity librarian to organize the collection. Any book or paper will be sorted according to its title and the contents of its pages, and similar books should be grouped together. I also only want to look at the English section of the library for the moment. I will still have a huge section to look through but at least every work is titled and readable by me. Much of the information will be junk but amongst the sea of rubbish will be islands of useful knowledge. Now, is there a way to find knowledge in this library in an automated fashion?

Battleship

The best analogy I can find to illustrate iterative knowledge discovery is the 1970s family game 'Battleship'. The game consists of two 10 by 10 grids that you plug your ships into. All the ships are linear shapes of a few squares in length. The players cannot see each other's ships and must guess where they are. A very simple way to do this would be to ask your opponent whether they have a ship on the top left square and continue systematically across the board, square by square, until you reach the bottom right hand corner. This would eventually find every ship. If every ship were a piece of knowledge we could discover all the knowledge in the world by simply stepping through the board one cell at a time, but it would take a long time.

A better way to play Battleship is to pick a square at random. If you get a hit, explore linearly around the hit. This will efficiently find the rest of the ship. The same might be true for knowledge. We could take random shots, get lucky and move linearly to flesh out our knowledge. Once we had exhausted an area we could take a step away at random and again hope for another hit. This process is exactly the way some people imagine the frontier of knowledge expands.

But, it is wrong.

The monkey moon shot story explains…

> *"I believe that this nation should commit itself to achieving the goal, before this decade is out, of landing a monkey on the moon and returning him safely to Earth."*
>
> President Monkey

The monkey nation is asked to mount a moon shot. After a little time a monkey is asked to report on progress.

"I can report," says the monkey, "I have climbed a particularly tall tree on the tallest hill on my island and have made over seven hundred meters progress towards the moon, although this is only 0.0001% of the way there, this has been quick so I believe we are well on the way."

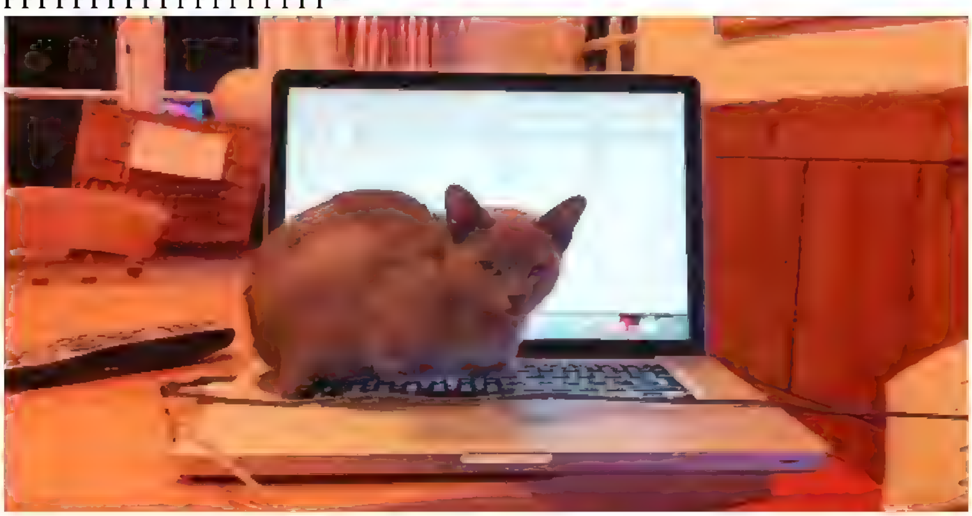
You see of course the problem. Progress in many problems is nonlinear. Moving a bit of the way towards the goal does not provide any *actual* progress: That is the problem with knowledge. It is not linear in structure. You need to take leaps to discover new knowledge. You can not simply look around in the general area. Such leaps are mathematically huge. The chance of making a successful one by pure chance is virtually zero.

But Cats Can!

As chance would have it, as I was writing this book about the impossibility of creating great literary works at random, our new kitten, Jessie, sat on my keyboard – she likes the warmth. To my great embarrassment I have been proven wrong. Here is her first literary work. I managed to capture her on camera a little later that evening, editing a spreadsheet. My brain interprets this string as the cat thanking me for good food. I wonder if you see the same thing? This is just a demonstration of the strength of human pattern detection algorithms and not, sadly, of feline communication.

Cats Creation

…. Kkkklnk gfoooooooofd0——— iiiii;;;;;;;;;;ii…..ffffffffffffffffffffff……
==
==
==
=============================pppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
pppppppppppppppppppppppppppp..oppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
ppppppppppppppppppppppppppppppppppppppppppppppppppppppp
pppppppppppppppppppph



Jessie, Our Creative Kitten

Chapter 6

KITTENS & GORILLAS

Orangutan and Kitten

"No kitten that loves fish is unteachable; No kitten without a tail can play with a gorilla; Kittens with whiskers will always love fish; No teachable kitten has green eyes; No kittens have tails unless they have whisters; hence..."

Lewis Carroll

"Once you eliminate the impossible, whatever remains, however improbably, must be the truth."

Sherlock Holmes,
Arthur Conan Doyle

As well as giving us Alice, the Jabberwocky, and the Cheshire Cat, Lewis Carroll lectured on mathematics at Oxford University. He wrote several books on logic, illustrated with wonderful problems involving fish, kittens, and gorillas – much less boring than the brown, grass eating cows of modern textbooks. Kittens and gorillas are not usually in much contact, but I did find one hit on Google, pictured!

The words we organize into books, poems and plays are not just a random jumble; they have structure and a logic to them. We group verbs, subjects and objects together to form sentences and, at a larger scale, characters have motivations and relationships: this character loves that character, the valet had the candlestick in the ballroom and could not have stabbed the butler in the kitchen, and so on. We have dictionaries to define words, but to truly understand the information they convey we need to understand the logical rules governing how they can be combined.

Everyday conversation is fragmented and repetitive. Fortunately, now and again, we say something definitive. For example, "This gorilla is brown." The statement links a property, 'brownness', to a thing, 'a gorilla'. Logical statements are precise but often need to be put in context. If I were standing in a forest when I made my statement you must guess I mean the nearest gorilla. The word 'This' implies nearness, but nearness is not well defined. Better to be precise. 'The gorilla I am closest to, measured by line of sight distance is the Pantone shade *dark brown*.' However, if I talked like this all day I would not have many friends.

Logical Beginnings

The formal study of logic began in 384BC with the publication of a treatise called the *Organon* by the Greek philosopher Aristotle. A student of Plato, Aristotle taught many of the famous leaders of his time, including Alexander The Great. Ancient Greece was not some idyllic think tank. If you annoyed the political establishment you might find yourself having to leave town in a hurry. This happened to Aristotle after Plato's death, and he spent nearly a decade touring Europe. Eventually, he returned to Athens where he published his study on logic.

In the Organon, Aristotle examined groups of up to four statements, each containing up to four relationships. For example: All kittens eat fish. Some kittens eat fish. No kittens love gorillas. No gorillas eat kittens – luckily. It is possible to put two statements back to back and infer things.

I could say, "All gorillas eat leaves." "All leaves are green." Therefore I can infer all gorillas eat some green things. This is a valid inference. It is *not* correct to say, gorillas eat *only* green things.

There are 256 ways you can arrange four Aristotle statements with four relationships but only 19 valid deductive conclusions can be drawn. The kitten puzzle at the start of the chapter is an example of such a logical puzzle. Can you reach the right conclusion?

TRY SOLVING THE KITTEN PUZZLE WITHOUT READING ON

Aristotle's syllogisms are only a start. There are many other types of logic. In antiquity, the Stoics developed a different brand of logic based on the idea of larger and smaller. Stoic logic allows us to answer questions of relative size. If a Mini is smaller than an Audi, and an Audi is smaller than a Rolls Royce, then a Mini is smaller than a Rolls Royce. The Stoics pursued their branch of logic until around 180AD when study of this sort died out. It's not quite clear why. Perhaps the rise of religious power and the onset of the Dark Ages curtailed intellectual inquiry. Even after the Enlightenment began around 1650 it took some time for the discipline of logic to re-emerge. If you want to learn more about syllogistic logic and how to solve Lewis Carroll's puzzle you should read his book *The Game of Logic*. The definitive book on the logic of language, in my opinion, is *Logic* by Wilfrid Hodges.

Logic for Computers

Western civilization mostly survived on syllogisms and stoic logic for nearly two thousand years before George Boole devised his theory of binary logic in 1847. Boole developed an elegant mathematical system for representing logical statements that allowed simple arithmetical operations to answer logical questions. We now call this system Boolean logic and he gave us the modern convention of using one for true, and zero for false. Computers use his principles all the time. For example, if it is true my bank account shows less than zero, then make it true that someone will send me a letter warning me I am overdrawn. The best way to get your head around Boolean logic is to solve the ancient puzzle of the Two Guards. The puzzle featured in the 1986 movie, *The Labyrinth*, starring David Bowie and Jennifer Connelly. If you want to cheat watch the film to see the answer. Here is the puzzle. I'll put the answer on my website.

> *Two guards stand barring your way and behind them are two doors. One guard always speaks the truth, while the other always lies. You are only allowed to ask one question of one of the guards. Your life depends on picking the right question to ask as, based on the answer, you must pick a door. One leads to life, the other to certain death. Is there a question you can ask to ensure you pick the door leading to life?*

TRY SOLVING THE GUARD PUZZLE

Twin Guards - Left door or Right

If you are reading this, you picked the correct door and lived.

Logic for Humans

Syllogisms can be used for practical purposes. Take, for example, the following set of statements, "I want a hot drink." "Coffee and tea are hot drinks." "I always drink milk with tea," "We have no milk." What drink should I choose? I'm sure you can work it out. This logical problem follows a simple chain and results in me getting the hot drink I like.

We use Boolean logic on a day-to-day basis. The simplest form is a checklist. Pilots use checklists all the time; do I have wings, fuel and a copilot? If they are all there, go ahead and fly. Otherwise do not. Mathematically speaking, a checklist is simply the product of the options. If they are all one, then the product is one – in this case we can fly. If any is false – represented by a zero – the product will be zero and we cannot fly.

Life is often more complicated and we have many logical tools at our disposal. Let's take a look at a few, starting with a famous historical one.

Benjamin Franklin invented the lightning rod and bifocal glasses, as well as charting the Gulf Stream and all manner of other scientific discoveries. He described his process for decision-making when there are many pros and cons to consider.

> *"... my Way is, to divide half a Sheet of Paper by a Line into two Columns, writing over the one Pro, and over the other Con. Then during three or four Days Consideration I put down under the different Heads short Hints of the different Motives that at different Times occur to me for or against the Measure. When I have thus got them all together in one View, I endeavor to estimate their respective Weights; and where I find two, one on each side, that seem equal, I strike them both out: If I find a Reason pro equal to some two Reasons con, I strike out the three. If I judge some two Reasons con equal to some three Reasons pro, I strike out the five; and thus proceeding I find at length where the Balance lies; and if after a Day or two of farther Consideration nothing new that is of Importance occurs on either side, I come to a Determination accordingly."*

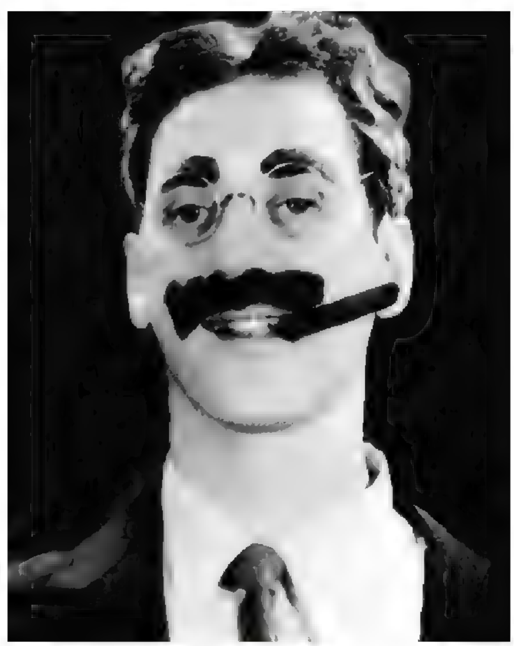
Another important piece of logic is *reductio ad absurdum*. Reduction to the absurd allows us to disprove something because, if it were true, it would lead to an absurd conclusion. An alibi is a familiar form. If I was seen in the pub when the murder occurred in the ballroom of the manor house and you claim I committed the murder, I must have been in two places at once. People can't be in two places at once – that would be absurd. Conclusion: I am innocent!

Notice I not only prove I am not guilty I also prove the opposite: I am innocent. When a mathematician uses this trick, it is called an indirect proof and works the same way as the alibi. Assume the opposite is true of some theory you want to prove (I am guilty). If it generates a contradiction or paradox (can't be in two places at once) you can deduce the opposite must be true (innocence). Mathematicians use this all the time. It assumes, of course, mathematics is consistent and that true and false are opposites.

Some mathematicians argue this is too strong an assumption. Why should we assume consistency and recognize only two logical states, true and false? These mathematicians believe the only way to prove a theorem is with positive argument rather than using the opposite of a negative argument. They don't allow indirect proofs in their mathematical models. This type of mathematics is unsurprisingly called positivism. It's a pure theory but, unfortunately, if you try to follow it you lose much of our current mathematical knowledge and understanding. Most modern mathematicians think it a historical curiosity, but it does pop up from time to time. Modern mathematics is founded on the axioms that true and false are the opposite of each other and that inconsistency is forbidden within the system. Mathematical proofs submitted to journals are not permitted to contain inconsistencies or result in paradoxes.

Paradoxes – When Logic Fails

> *"I would not be a member of any club that would admit me."*
>
> Groucho Marx



Groucho Marx

Paradoxes occur when a statement makes no sense, or results in an internal contradiction as with Groucho Marx's famous quote. They are widely used in mathematics to implement indirect proofs. To do this, we suppose something is true, and if it results in a paradox then the thing we thought true must be false and the opposite is true. This is a somewhat circuitous route to prove things, but it is often the only practical way.

Two paradoxes we are taught as children are the liar's paradox and Zeno's paradox – also known as the story of the tortoise and hare. The first is a real paradox but the second is a false paradox. The liar's paradox is just the simple statement:

"This sentence is false."

It is a paradox because of the internal inconsistency: We cannot determine if it is a true or false.

First assume it is true, but it says it is false, so it is not true. Then try it the other way around. Assume it is false but the sentence states it is false, so it must be true. If that were so it must be false by the first argument and so on *ad infinitum*.

Either way around, the sentence contradicts itself. A paradox.

Zeno's Paradox, on the other hand, is a false paradox. Here is the story.

Once upon a time there was a hare. He was a very arrogant hare and believed he could outrun any animal. A tortoise was walking along the way and the hare jumped out in front of him. "You are so slow," said the hare. The tortoise replied, "You may be the fastest hare in the kingdom but I am the most persuasive tortoise. I bet I can persuade you of anything, including that I am faster than you."

"I don't believe you," said the hare.

"OK," said the tortoise, "let me show you. Give me 100 meters head start since you are so fast. Then, we'll both start to run. After 10 seconds you will have run 100 meters and arrived where I used to be, but I will now be ten meters ahead. After another second you will be where I am now, but I will be 1 meter ahead again. So you can never catch me."

The hare pondered for a while but, being a hare of little brain, could not make out the true answer.

It is a false paradox. The time intervals are getting shorter. The question for a mathematician is, does the problem converge to a solution. The answer is yes, and I can reframe the problem to see how it is solved. Let's simply look at who would be ahead after 20 seconds: the hare!

The mathematical reason for it being a false paradox is that some series converge and some do not. If I move progressively closer and closer to something in smaller and smaller time intervals then I may indeed reach it. On the other hand, some series never converge. I will never reach infinity how ever many steps I take.

The Barber Paradox

Now, for a slightly harder paradox, let's suppose there is a town with just one barber.

In this town, every man keeps himself clean-shaven by either shaving himself or going to the barber; the barber shaves all the men in town who do not shave themselves. All this seems perfectly logical, until we pose the question: who shaves the barber?

This question results in a paradox because, according to the statement above, he can either be shaven by himself or the barber, which is he. However, neither of these possibilities is valid! This is because if the barber shaves himself, then the barber must not shave himself and if the barber does not shave himself, then the barber must shave himself.

You might think this paradox an oddity but, using this simple idea, Bertrand Russell changed the course of mathematical history and it is the fundamental paradox used to show computers are Turing limited.

The Russell Paradox

In the late 19^{th} century, mathematicians began to think about the nature of numbers.

What is a number?

It is certainly not an object we can hold.

I can't hold a *two*, unless it's the brass number plate, for my front door. And, in that case I am holding *one* number plate, so I am not holding the *idea* of two, but rather the idea of *one*: one brass plate in the shape of a two.

The 'idea' of a number is to say something about the *things* I have in my hand: two apples, two oranges and two brass number plates. These are all sets of two things and 'two' is the collection of all these sets.

In 1890, Gottlob Frege completed his theory of sets. The project had taken him five years. Unfortunately, just before sending the book to the publisher, Bertrand Russell wrote to him and pointed out the following paradox. What about the set of sets that does not contain itself? Think about it...

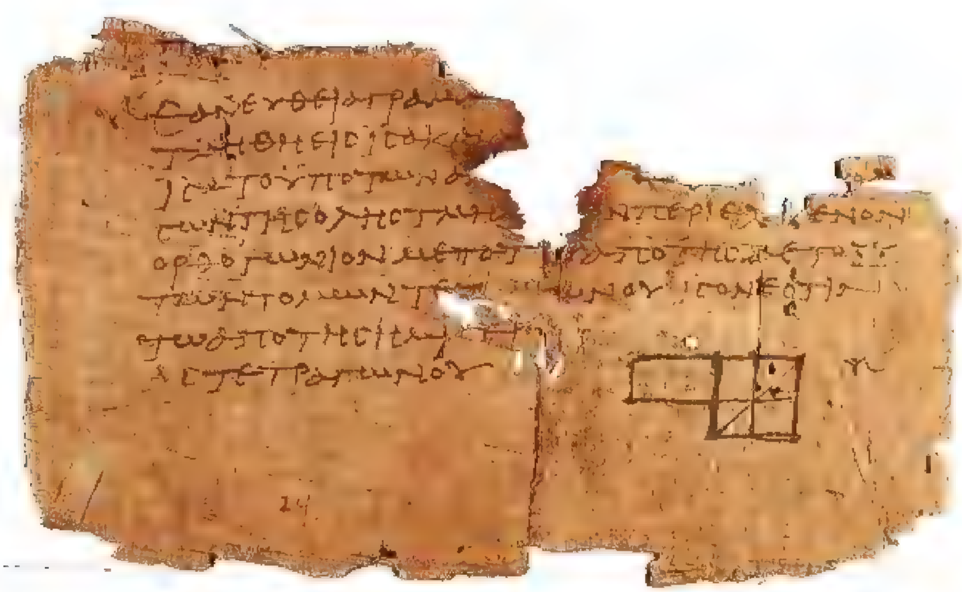
It is the barber paradox with the word 'set' substituted for 'barber' and 'contains' rather than 'shave'. But it's essentially the same logical problem. You might find this rather contrived but mathematicians must have a system totally free from paradox, otherwise there is no certainty. Frege's system was holed below the water line.

Eventually, after much further work, a theory of sets was worked out that does not contain the Russell Paradox. It's called Zermelo-Fraenkel set theory, or ZF for short. It solves the Frege problem by forbidding sets to refer to themselves. It's a bit like Microsoft Excel's solution to dividing something by zero. It is simply forbidden and generates an error message. Set theory was fixed and is now the basis of most mathematical thinking.

What is Logic for?

Logic is the foundation of mathematics. Applying it enables us to make irrefutable statements about things: numbers, lines, planes, equations and the so on – the things you learned at school – and to prove statements about these things beyond *any* doubt. This is not the 'reasonable doubt' hurdle of our law courts, but an absolute measure: No possible doubt whatever.

Let's look at one of the earliest mathematical proofs: Euclid's proof there are an infinite number of prime numbers. Euclid created this proof in ancient Greece around 300BC – so far back that logic was in its infancy



Euclid's Elements 100AD

and numbers had not yet been properly invented. Euclid used distances rather than numbers for all his proofs but I will use the word 'number' in this explanation.

First a little revision. A prime number is a number that can only be divided by itself and one, for example three, five, seven, and eleven. All numbers can be split into primes using a couple of tricks. First, all numbers are divisible by a set of primes. Ten is five times two – two primes. We are also fairly sure we can form any number by adding two primes together. This is Goldbach's Conjecture, set as a question in a letter written to Euler in 1742. It is still unproven!

Euclid proved there are an infinite number of primes by using *reductio ad absurdum*. Imagine we have a complete list of prime numbers James' list of primes. It contains every prime number. (This is the setup. We are proposing something we suspect is incorrect and will lead to a paradox or contradiction. When it does, we will have proven the opposite fact. The proof relies on the fact that a number can either be prime or not prime. There is no middle ground.)

Let's make a new number by multiplying all the numbers on my list together and adding one. There are two possibilities: this new number is either prime or not prime.

If the number is prime, it is a new prime number that was not on my list and I have disproved the theory.

If it is *not* prime then it must be divisible by two prime numbers already on my list. However, neither of these numbers could have been on my list, because dividing by one of them would give me a remainder of one. Remember I multiplied *all* prime numbers together and added one. It must, therefore, be a new prime number, which had previously not been on my list. Once again, I disprove the theory.

Since both routes fail, James' list of prime numbers is not complete and, therefore, prime numbers are infinite.

Feynman's Proof

My favorite piece of logic is Richard Feynman's disproof of the existence of polywater. It's a strange logical proof bordering on philosophy, but it shows just how far you can take logic.

In 1969, an urban legend spread around the world that there was a substance called polywater. It even made it into an episode of Star Trek. Polywater was believed to be a lower energy state of water, more viscous than ordinary water. If this substance did exist, it would be possible to mine the oceans of the world converting water to polywater and therefore generate energy. There was a concern that if the right catalyst was accidentally introduced into the oceans they would solidify into polywater thus dooming the human race, or at the very least making water sports impossible!

Feynman was consulted and stated, "If there were such a substance as polywater then there would have evolved an animal that eats water and excretes polywater, using the liberated energy as its power source. Since there is no such animal, polywater does not exist."

Feynman's proof is an elegant indirect proof coupled with a syllogism. Polywater exists. Polywater is a lower energy form of the high-energy substance called water. Food is a high-energy substance that can be converted to a low energy substance by a process we shall call 'being good to eat.' All things on earth that are good to eat have something that eats them. Polywater is a food and therefore good to eat. Therefore an animal must exist that eats polywater. No such animal exists, so either something in our chain of logic is wrong, or the premise is unsound. Since the chain is sound, the original premise must be wrong: Polywater cannot exist.

In short, Feynman's proof says: if a thing is so, then the inevitable consequence is the evolution of something else, and since that something else does not exist, the original thing cannot be so. QED: disproof by nonexistent consequence.

The polywater disproof neatly demonstrates the important elements of Feynman's Evolutionary proof. First, life must be continuously exposed to the thing in question, in this case water. This is clearly so as most life on planet Earth lives in the oceans or is intimately entwined with water. Evolution takes time, so enough time must be allowed for life to evolve. It must be a nearly linear problem so that a solution proceeds in steps where each step is an improvement and no step requires too high a level of mutation or adaption. We can illustrate the boundary between a linear problem and one requiring a step change by describing how triple drug therapy works in the treatment of AIDS.

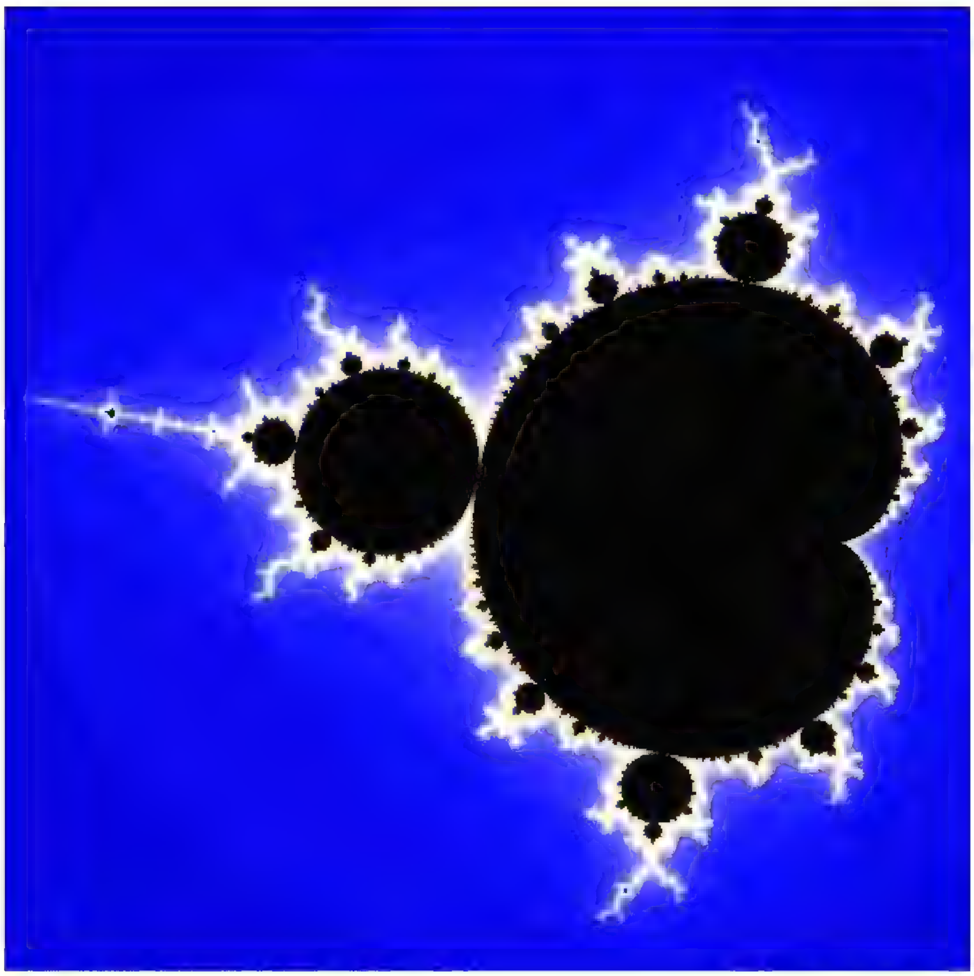
Until triple drug therapy entered the picture progress against AIDS had been a depressing story of drug discovery followed by the almost immediate evolution of the virus to evade the drug. The AIDS virus is a retrovirus with a shell composed of sugar molecules. It is almost trivial for an AIDS virus to mutate these outer markings to look different, even from one day to the next. This is the way the virus continually and nimbly evades our immune system. However, the AIDS virus *does* have some components that it can't easily mutate because they are not merely aesthetic, they have a functional purpose. Why not target them?

Unfortunately, it turns out the AIDS virus can even mutate its functional parts, but this is harder. The probability of a successful functional mutation is 1000 times less likely than a simple aesthetic mutation to the sugar coat.

Triple drug therapy works by attacking three different *functional* elements of the virus simultaneously. It is possible for the virus to modify all these functional elements but the likelihood of it doing so is tiny. One mutation alone does not help because the drug cocktail will still target the other two elements and kill the virus. The AIDS virus does not understand that it is facing a triple drug cocktail. It cannot reason like a sentient being and random chance is not sufficient to make the big leap necessary to overcome the cocktail of drugs. Unless you can mutate all three elements at once your time as a virus particle on this planet is over. Most problems we have to solve in this world require more than one simultaneous logical step and these don't happen by chance.

Chapter 7

COMPLEXITY & CHAOS



Mandelbrot Set

"Life is really simple, but we insist on making it complicated."

Confucius

"Any darn fool can make something complex; it takes a genius to make something simple."

Pete Seeger

There was once a great King who lived in a marvelous palace. To fend off boredom he collected all manner of interesting games and puzzles. One day an inventor came to his palace and told the King he had a game of such subtle complexity, yet apparent simplicity, the King would play no other. The King learned the game and soon agreed it was, without doubt, the best of all games. The game was, of course, 'chess'. The King asked the price of this game and the inventor told him it was a mere trifle. The King should give him one grain of rice on the first square of the board, two on the second, four on the third, and so on, doubling each square until he filled the board.

The King called his treasurer to honor the bargain and the first bags were brought from the storehouse. The grains were placed on the board in each square but soon there was not enough space and the grains had to be piled on the table next to the board. Soon this, too, was not enough and every table and chair in the hall had to be covered. Even this was not enough and they began to stack whole bags up in the courtyard.

When they reached the thirtieth square, the treasurer turned white. He sat and calculated for a while before saying with a trembling voice, "My great ruler, there is not enough rice in all the world to cover this board." The ruler called the inventor and told him he could not honor the debt and the inventor should name another price. The King had two beautiful daughters, the first knew she was beautiful and deported herself accordingly, and the second, was bookish and shy, but perhaps more beautiful for this. The inventor asked for the hand of the second daughter and lived happily ever after. In the less favorable version of this story, the King becomes very angry and has the inventor beheaded. I prefer the romantic version.

Placing rice on a chessboard and doubling it successively leads to wildly large numbers. Covering it completely requires 18,446,744,073,709,551,615 grains, about four hundred trillion tons and equivalent to one thousand years of worldwide rice production. Like the king, humans do not intuitively grasp the enormity of this problem because we're not good with large numbers.

Although the number of grains needed to cover a chess board is very large, it is not hard to calculate. The treasurer is the one who should have lost his head for not being able to do the calculation. The equation is simply two, doubled sixty four times, less one, $2^{64}-1$. A pocket calculator can produce this number in a thousandth of a second: it's just long multiplication. Although calculating this number is quick, it is not always the case. Answers to some problems have short cuts, while others do not.

Mathematicians have catalogued the universe of problems into classes rather as biologists have catalogued animals into species. Each problem is examined and put into a genus with a name. Sadly the names are not as readable as the Latin names for animals. For example, 'nlogn' is the complexity class of most sorting programs, while traversing a maze typically sits in the class NP or P/POLY. Although the classifications look complex the basis of cataloging is simple, a class name signifies the time needed to solve a problem using the best possible algorithm, and the scale this is measured in is 'Big O'.

Big O

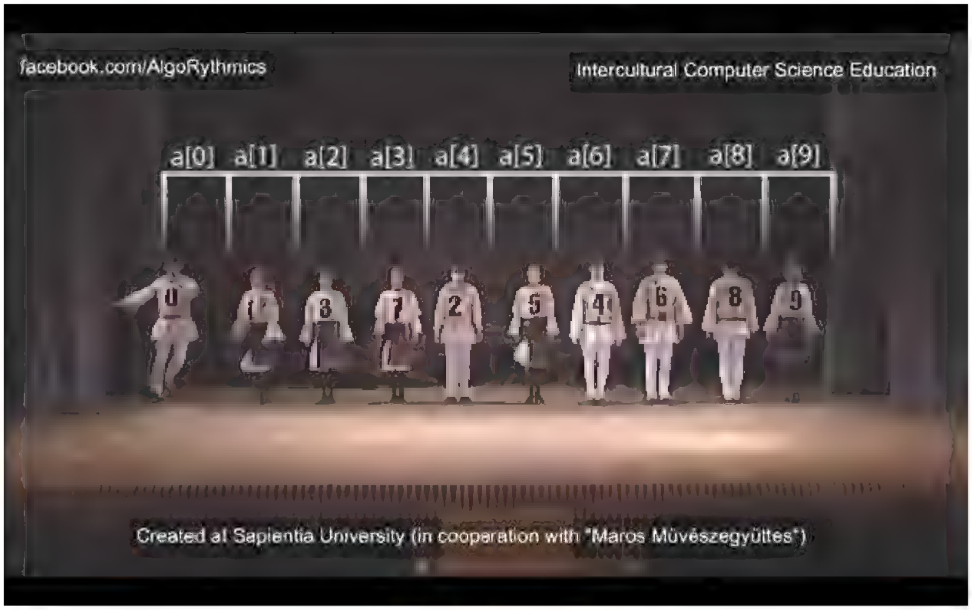
Every problem has a complexity. In mathematics this is expressed using 'Big O' notation, where 'O' stands for order-of-magnitude. The simplest problems have order 1.

If I am working at my computer on a Word document and I press *print*, the printer will spring to action and print the document. This problem is of flat complexity, notated O(1). It does not matter how large the file is; one click is all I need. I am, of course, assuming sufficient paper in the printer and ink in the cartridges.

The next complexity class is a linear problem, O(n). For example, walking to the store to buy a pint of milk. The farther the store, the longer the walk. The time needed to get to the store is directly proportional to the distance: if I am walking, a single step multiplied by the number of steps required to cover the distance.

You might think adding two numbers together is a linear problem – the bigger the number, the harder the problem – but there's a clever trick to speed it up. You can get 10 people to add each column in parallel. They'll need to coordinate when someone ends up with a number larger than ten and has to carry the extra digit but this can be easily solved. A problem gets its classification only once we've used the cleverest possible trick to solve it.

Most problems we meet in mathematics are somewhere in between flat and linear but there are some that are much harder. The most common hard problem we come across in our daily lives is sorting. Rather than go through a tedious written description, check out the video link on my website. Sorting without using any spare space requires a bubble sort. This is an example of something that needs n squared operations and, since n squared is the simplest example of a polynomial, it is said to be in the polynomial time, or 'P' time classification.



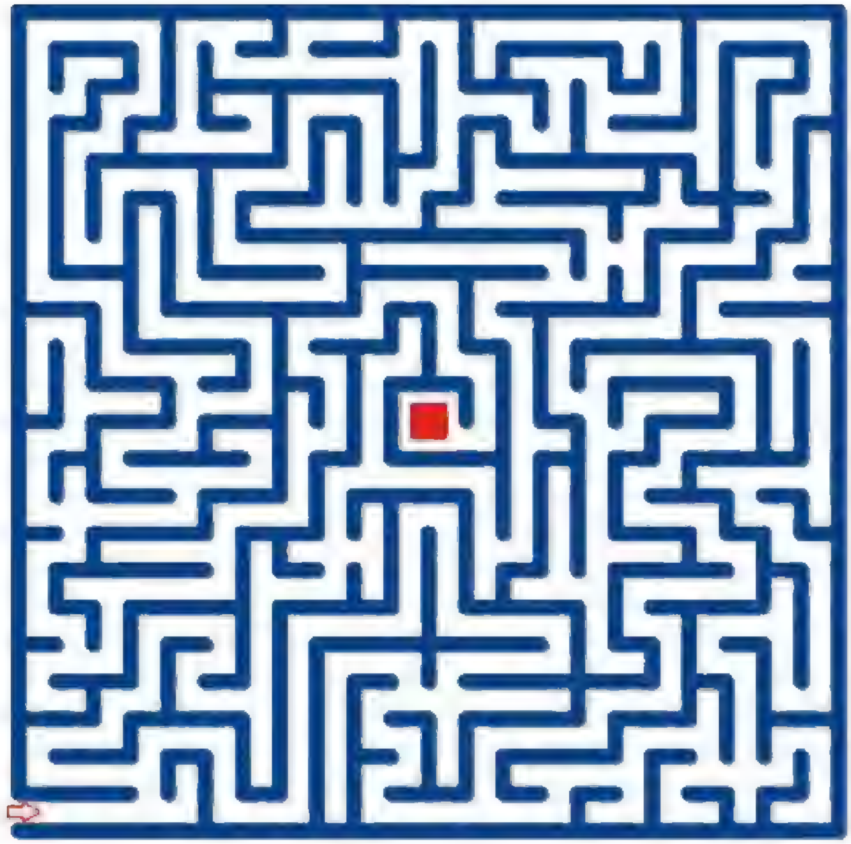
Bubble Sort Ballet

The Hardest Problems

You probably hope cracking the encryption used to secure the Internet is one of the hardest problems known to man but I'm sorry to tell you it is not. When you use your credit card to buy something from an online shop, your web browser changes from *http* to *https*, the 's' stands for secure. The data you send to the Internet is coded using a system developed in 1977 by Ron Rivest, Adi Shamir and Leonard Adleman of MIT, which is why it is called RSA encryption. Any information you send is raised to the power of a very large number – usually around one hundred digits long. Raising something to the power simply means multiplying it by itself that many times.

What makes decrypting a message hard is that division is a slow process; it is called 'long division' for a reason. It turns out there is no way to speed it up on a conventional computer so, unless you know the right number to divide by you will have to try *every* number. It is this that makes decrypting RSA messages hard.

Although RSA messages are difficult to decipher, they are nowhere near the hardest problems. That accolade is commonly believed to go to non-deterministic polynomial problems known as 'NP' problems. NP problems are easy to describe but fiendishly difficult to solve. Nondeterministic means each time you come to a branch in the problem there is no way to tell which branch is the best to pursue without exploring it all the way to the end. It's the same as a maze; at each junction in the maze you can decide which path to take, but the junction gives you no



Maze

clue which one will be better. Beware the confusing naming system, 'N' stands for *nondeterministic* in this case, whereas in normal complexity classes 'n' stands for *number*. Sorry. That's just the way it is. Let me give you an example of one of these NP problems.

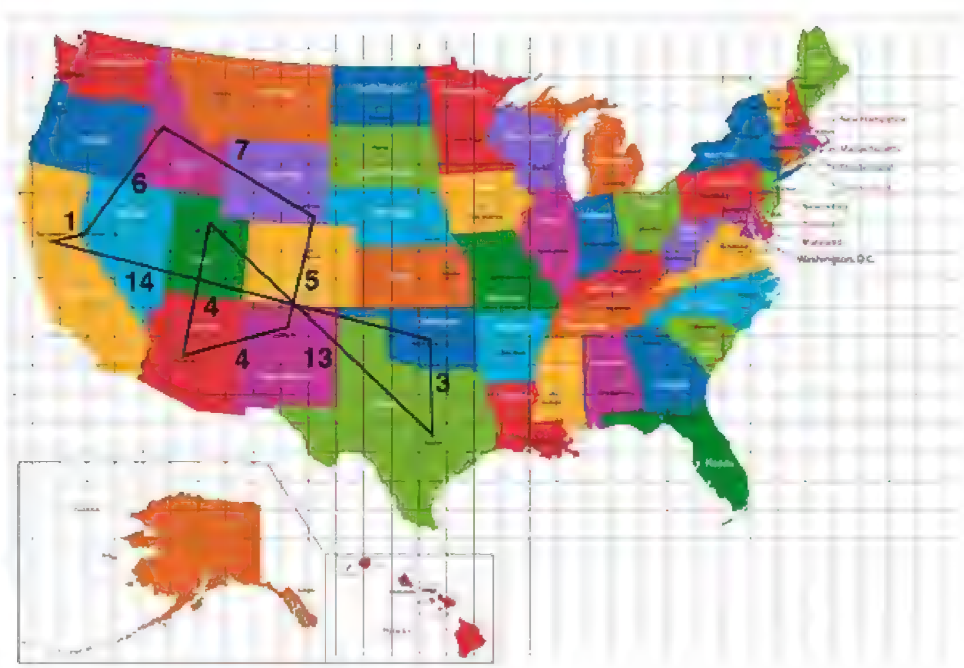
Let us assume we have one of those complicated recipes from the latest celebrity chef cookbook. If all the ingredients can be bought from one store, making the dish is straightforward, but if they come from different stores, you will have your work cut out. What is the best order to visit them? With 2 shops, it's trivial. Either order will do. With 3 it is a little harder and with 4 there is quite a bit of choice. This is known as the 'traveling salesman' problem because the original formulation described a salesman wishing to find the shortest route between all the cities in which he had customers. The complexity of this problem rises much faster than the Rice and Chess Board problem. Try it for yourself. It doesn't matter if you imagine you are visiting customers or shops. I have given you a grid to count off distances. Try to solve a problem for 3 cities, 5 and 10. What is the shortest path allowing you to visit each place?

TRY THE PUZZLE ON THE WEB

Warning: Don't spend too long on these problems.

The reason I warned you not to spend too long is that solving the 50-city problem would take longer than the age of the known universe. NP problems get harder very fast as the number of elements goes up. A 50-city problem is hugely larger than a five-city problem, not just ten times harder.

The Clay Mathematics Institute has offered a $1 million prize for anyone who can say whether NP problems are really as hard as they appear. It may be there is a general trick or a series of tricks that allow you to solve any NP problem in a shorter time. If you could do this, the problem would be demoted to P, allowing fast computers to tackle it. No one has yet found a proof of the P=NP problem. At the time of writing several proofs are sitting with the Clay Prize judges but don't hold your breath. Most people assume there is no solution. If you want to have a crack at the problem let me state it in simple terms.



Traveling Salesman

Imagine you wanted to find the center of a maze. Is there a way to speed searching the maze, so you do not have to test every branch? If you can provide a mathematical proof that there is or is not, you win the prize.

Places Game

While it is commonly assumed NP problems are the hardest, this is not the case. There are quite a few that are harder still. One such is called a PSPACE problem. It's quite difficult to explain but luckily many of you will have played a form of it on long car trips when you were a child: My family calls it The Places Game.

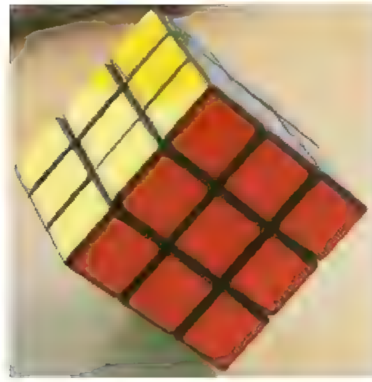
I will pick a place 'London,' and you must then pick another place, say, 'New York', that starts with the letter my place ends with. I'll then pick 'Canterbury' and my kids will laugh at my dyslexia and I'll have to switch to 'Kansas' and so on. Once you use a place you can't use it again.

The mathematical question is to predict who will win given each player has a finite list of places they know? It turns out this type of problem is even harder to solve than an NP problem. This is because on each turn a player gets to pick any name from their list. With the traveling salesman problem, there is only one 'player' the salesman so we can write out a route and check it. In the Places Game there is no single route through the game because, after I pick my favorite town 'London,' you can pick any place beginning with 'N'. I have to anticipate an enormous table of possible paths through the game. The table takes huge physical space – which is where PSPACE gets its name.

Remember I'm just playing the simplest mathematical games with bits of paper and discrete ideas. I haven't strayed into the quantum world yet. That brings with it a whole new level of complexity to explore. Complexity is such a diverse subject that Scott Aaronson of MIT has created a web site called the *complexity zoo* to catalogue all the different 'species'. It is much to complex to reproduce here but let me provide a sketch.

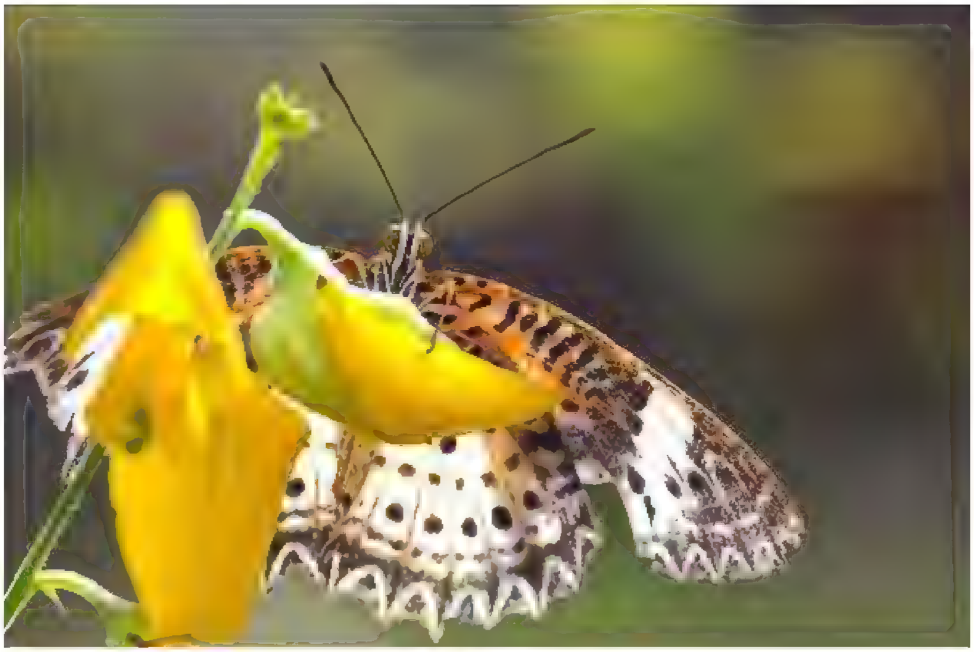
The Complexity Hierarchy

My table below represents the hierarchical complexity of knowledge. We start off with the problems both humans and computers find easy, then rapidly move onto problems that even the fastest machines find difficult: a perfect game of chess or predicting the weather. Above these computable problems are the non-computable ones which no computer running any algorithm can solve, and then there are the free will problems: how do we pick a problem in the first place? How do inventors come up with problems no one had ever thought to solve in the first place, such as the invention of the Rubik's Cube?



Ernő Rubik's Cube

Problem	**Example**
Flat	*Print File (for Human)*
nlogn	*Searching a list*
Linear	*Finding the lowest number in a list*
Logarithmic	*Long Multiplication*
Exponential	*Long Division*
P	*Most Algorithms*
Near NP	*Factor Prime Number*
NP-non-complete	*Perfect Game of Chess*
NP-Complete, tractable	*Travelling Salesman, SAT*
Chaotic	*Weather*
NP-Complete, Quantum	*Modeling a Quantum Process*
NP-Complete, intractable	*Busy Beaver, Towers of Hanoi*
PSPACE	*Graph Problems, Places Game*
Non-computable	*Creativity, Finding Fermat Theorem for a Turing machine, Tiling the plane with Penrose Triangles*
Non-deterministic, Non-time divisible, Non-computable	*Free will*
Impossible	*Halting problem for a Turing Machine, some mathematical theorems such as the Continuum Hypothesis in ZF+AC (Hilberts 1st). Travelling faster than the speed of light. Understanding the American tax code.*
Known Unknowns	*I know that I don't know either way.*
Unknown Unknowns	*I have not thought to ask that question yet. Inventing the Rubik's Cube*



Butterfly

"Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?"

Philip Merilees, improving on
Edward Lorenz

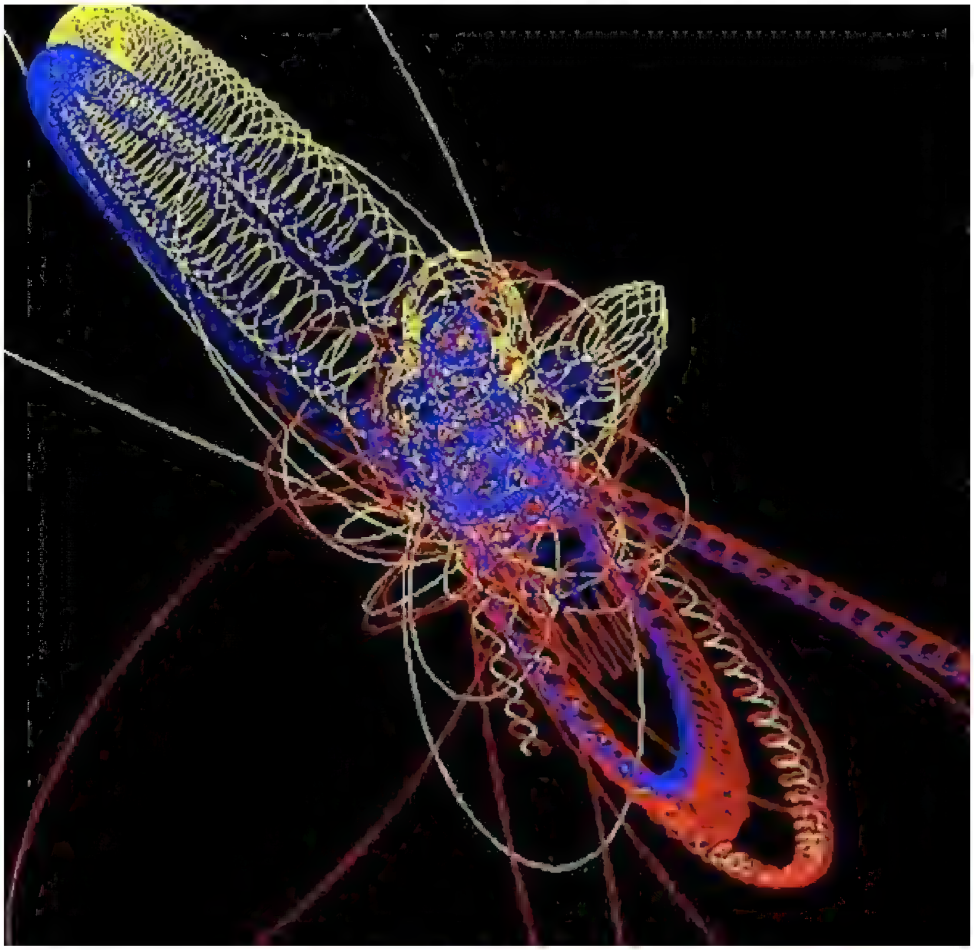
Chaos

Chaos is the twin of complexity. It burst into the public psyche in 1987 with the publication of James Gleick's book *Chaos*. It's not a difficult concept to grasp. Complex systems can be formed using simple rules, and very small changes in starting conditions can profoundly affect future events. I experience this if I miss my train to work in the morning: 30 seconds either way will change the whole pattern of my day, the people I meet and the level of stress I experience. I'm sure you can think of similar experiences.

Henri Poincaré, a French mathematician, first studied the effect back in 1880. Poincaré was trying to solve an old mathematical problem called the Three Body Problem originally set by Isaac Newton. Take the Earth, Mars and the Sun. These three bodies orbit each other, or strictly speaking a point in space somewhere between them. Is there an equation that will tell you where the bodies will be in, say, 100 years' time?


Poincaré

The answer is surprising, no. The three bodies will orbit in a non-repeating way. There is no analytical short cut, no equation that will predict where they will be at some point in the future. The only way to know is to build a perfect model of the system and see what happens. Poincaré won a valuable prize for his proof from the King of Bavaria. You can see some amazingly complex orbits plotted below. Remember these are still deterministic and predictable – after all, they were calculated with a computer – they are just chaotic.



Four Body Problem

Butterflies and Sliding Doors

After Poincaré, the field of chaos remained fairly quiet until Edward Lorenz began studying weather patterns using computers in the 1960s. The story goes, one day his computer was misbehaving and he had to re-key some data into the machine. Rather than using eight decimal places he used only six to save time, and was amazed when the results of his program came out completely different. Dropping the seventh and eighth decimal place represents a change of only one part in a million, yet the patterns of weather predicted by the computer were completely altered.

Lorenz went on to study the effect and created a new branch of mathematics. His quote about the beat of a butterfly wing creating tornados has entered the public psyche and is central to the plot of numerous Hollywood movies. One of his functions – known as the Lorenz Attractor – nicely illustrates the nature of chaos. A very simple equation plots the beautiful, apparently three-dimensional, non-repeating shape.

Chaosville

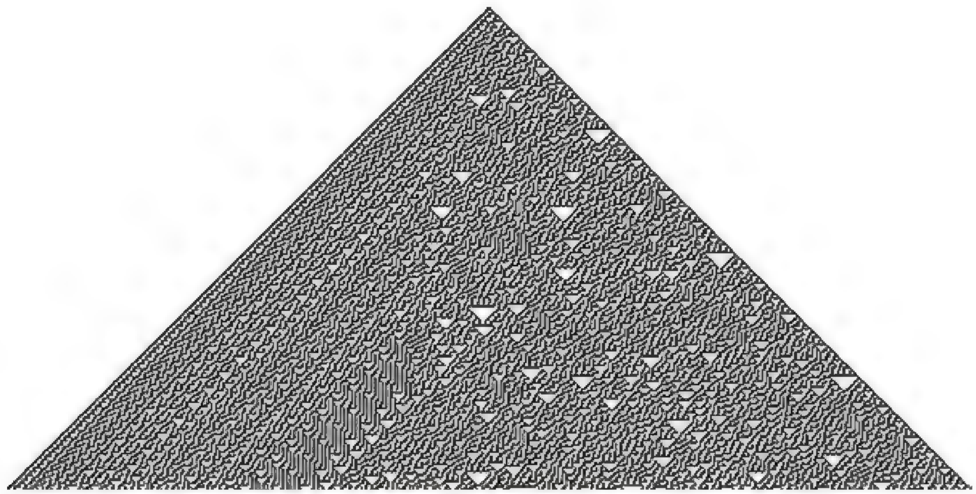
Chaos, taken to its logical conclusion could explain our Universe. Stephen Wolfram in *A New Kind of Science*, makes the argument that simple rules could explain the extraordinary complexity we see in our Universe. He applies rules to elements in a two-dimensional grid programmed on the computer which form 'cellular automaton' that function a little like simple animals, generating all manner of complex shapes and behaviors. The inspiration for this approach is almost certainly Conway's Game of Life developed by John Conway in the 1960's. In his computer game, animals and machines seem to appear on the screen but in truth they derive from the most simple set of rules. You can check out the website to see a live version of Conway's Game of Life. It's a lot of fun. Wolfram's

Strange Attractor

thesis is that we could all be living in one of these games. Perhaps our Universe is a form of Mandelbrot diagram – albeit a 3D version with stars and planets. If you look at the picture of a nebula and compare it to the Mandelbrot set, you can see how this is a tempting conclusion.

In the Game of Life the rules are simple yet the behavior simulates little animals being created and destroyed. Of course, there are no actual animals. The things you see on the screen, 'gliders', 'walkers', and 'cannons', just hang together accidentally. But, Wolfram considers these little digital creatures *are* animals. He argues our Universe is just like the Game of Life: A set of simple rules leading to complex behavior. If we are

Nebula



Cellular Automaton

prepared to call ourselves animals, so should the little creatures which emerge within the game. We simply emerged in a similar but slightly more complex game.

This proposal would mean our Universe is entirely deterministic, our lives the result of a gigantic computer program that we live within and form part of. Chaos might make it impossible to predict the future without running the program and watching what happened, but the results would be inevitable, set in motion at the dawn of time. There is no place for free will in such a Universe, no place for reason. The world would simply be.

But a strange idea will come to our aid to show us the limits of computation and allow us to question whether we live in a predetermined world. This idea is Aleph 1 – something larger than infinity. And it is infinity we will explore next.

Conway's Game of Life

Chapter 8

∞

Hilbert's Hotel

"All infinities are equal, but some are more equal than others."

George Orwell, paraphrase

"Only two things are infinite, the universe and human stupidity, and I'm not sure about the former."

Albert Einstein

"God gave us the integers, all else is the work of man."

Kronecker

Health warning! The man who discovered infinity had a mental breakdown. This subject may tax your brain.

Georg Cantor didn't really 'discover' infinity but he was the first mathematician to put it on a firm theoretical footing. In the late 19th century, most mathematicians thought infinity was a curious idea with no proper place in mathematics. They treated Cantor's attempts to make it into a real mathematical object with contempt. This affected Cantor's morale and caused him to suffer several bouts of deep depression, retreating to a sanatorium from time to time.

Infinity is a difficult idea to grasp but it is vital to our study of information. It behaves counter-intuitively but is not impossible to grasp. The reason it is important is that information can always be translated into numbers and numbers go on to infinity. If you want to know all about information, you must understand infinite numbers.

History

Indian scholars began studying infinity in the 4th Century BC. It turns up naturally in all manner of places. In geometry, parallel lines extend forever in either direction without ever meeting. To define a parallel line you must contemplate infinity. In arithmetic, even if you pick the largest number you can imagine, there is always a larger one; just add one. In the physical world if you look up at the night sky it appears to go on forever. Again you have infinity.

Historically there were two interpretations of infinity. The first, favored by Plato, was a journey. When you embark upon a journey, you can always take another step. Infinity is the idea of 'one more' or never-ending. It can never be reached. The second definition is more radical. Infinity is a thing, a number *so big* you could not imagine anything bigger, but it is *one* number. Plato thought this second definition tantamount to madness. Today we embrace this madness and go a whole lot further. Let me show you how.

If infinity *were* a number, you should be able to perform mathematics with it; add it, multiply it, and even raise it to a power. This is not as radical as it might first seem. Until comparatively recently, zero was not accepted as a number – if you consider recent to be one thousand years! Nowadays it is.

At the end of the first millennium Indian scholars found, against their intuition, that you can use zero as a number without generating contradictions. Take addition. I can have zero cakes, add one, and I have one cake, add another, and have two cakes and so on. In this way, the number zero behaves just like any other counting number. It also works with multiplication. If I have zero lots of 4 cakes, I have no cakes. Zero times four is zero, so multiplication with zero works. There is one embarrassing exception, if I divide by zero I seem to get infinity. When I was a child this was a definition for infinity, but nowadays mathematicians simply forbid the operation. Division by zero is not allowed and if you try it on your computer, you will get the not terribly useful, #DIV/0! Error. That's progress I guess!

Zero had been tamed. What about infinity?

Cantor showed that while you *could* think of infinity as a number, it might not be just one number. He proposed there are many infinities. In fact, there are a greater than infinite number of them! He did this through a rigorous analysis of a new branch of mathematics called set theory.

Set theory is now the cornerstone of modern mathematics, but it was treated with suspicion in Cantor's time. Rather than embrace the new thinking, many mathematicians ridiculed it; Poincaré wrote that Cantor's ideas were a grave disease infecting the discipline of mathematics! This seems odd given our modern propensity to embrace innovation, but the tone of science back then was different: innovation was not necessarily considered a good thing.

At the turn of the 20^{th} century, scientists were on a mission to tidy things up. Lord Kelvin announced in 1890 that mankind had discovered everything there was to know and the role of future scientists was simply to catalogue and observe the consequences of these laws, and to improve the accuracy of measurement. The last thing scientists wanted was a completely new set of numbers that behaved in strange ways. Cantor was upsetting the apple cart, but he was in good company. Just a few miles away in Berlin, a young Albert Einstein was beginning to study physics in his spare time. Those studies would culminate in his four papers of 1905, two on Quantum Mechanics and two on Relativity, ushering in the modern age of physics.

How to Count

To understand infinity you need to count in a particular way. You're probably used to counting with numbers. You count apples: one, two, three, and say, "I have three apples." You can do the same with oranges. If you have three apples and three oranges, the totals are the same and you can declare you have the same number of fruits. This is the first way to count.

But there is a second way of counting. Take your apples and put each next to an orange. If they match up, you can easily see they are equal in number. "Look," I say, "I have the same number of apples as oranges." This method is more primitive and does not require the concept of numbers, but it is very useful. If I'm a shepherd I can hold a set of counters in a bag, one for each sheep. To ensure all my flock are gathered in for the night I drop one counter into the bag as each sheep enters the enclosure. I don't need to give the counters number names.

The Munduruku tribe, from the Amazon rainforest, have no concept of number names beyond five. Their counting system simply goes one, two, three, four, five, many. Yet this second way of counting allows them to function successfully, deciding whether two groups of things have the same number of elements, even if there are more than five of them. For example, if they need to determine if they have enough spears for a hunt, each person simply stands next to their spear. If everyone has one, they're ready. If not, then the empty handed Munduruku simply make one. No need for pesky numbers or mathematics lessons.

This second way of counting is particularly useful when tackling infinity because we are not sure what infinity is. Treating it the same way the Munduruku treat the number 'many' is the safest thing to do. The first question we would like to answer is whether all infinite things are the same.

Spears and Hunters

We know from our childhood that infinity plus one is still infinity. Is there anything we can do to make infinity bigger? Perhaps multiplying infinity by infinity will do the trick.

Infinity times infinity can be visualized as a square with edges of infinite length. We can show that this square is the same size as a one dimensional infinity through a clever trick – the zigzag method. Mark the infinity square into a grid. Start in the corner square, go across, diagonally down, then across, diagonally up, and so on. I'll draw you a picture. We visit every square in our grid using a single line. We can then lay down our infinite zigzag line next to the infinite line of one of the edges. The lines are the same length as they are both infinitely long! So infinity, times infinity, can be matched to infinity, they are the same. Cantor thought this a very strange result and wrote to a fellow mathematician, Dedekind, *"Je le vois, mais je ne le crois pas!"*, "I see it, but I don't believe it!"

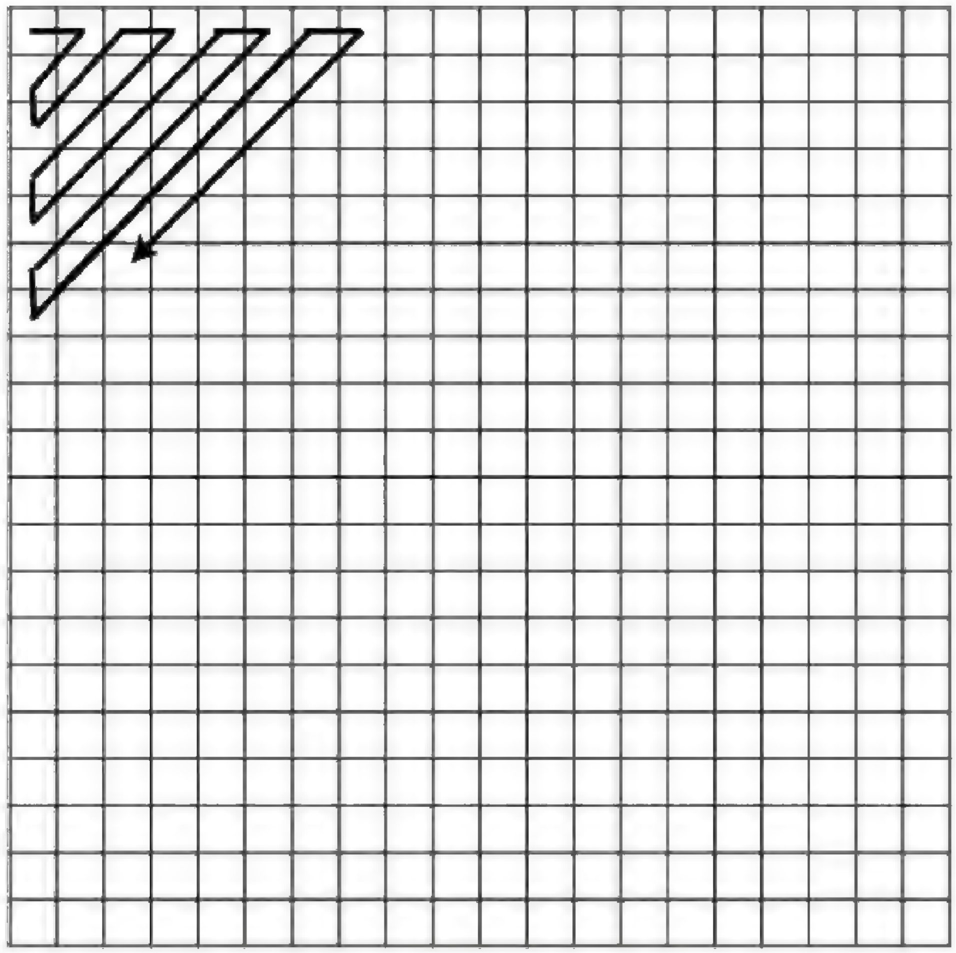
If you are struggling with this, don't worry. We just jumped forward to quite a complex concept. Let's take it more slowly. One way to get a better grip on infinity is through the stories of David Hilbert and the Infinity Hotel.

Infinity for Dummies

Hilbert's Hotel is a mythical building with an infinite number of rooms. Other than this strange feature it is a regular hotel complete with minibar, dodgy TV, and slightly mad manager. The rooms are numbered sequential starting at one, then two, three, four, and so on. The hotel allows you to play a series of mathematical games to see how infinity behaves.

Are there the same number of minibars as there are rooms? That's easy. I said *every* room has a minibar. We can use the matching technique to match minibars with rooms. Go to the first room. There is a number on the door and a minibar inside. The same goes for room 2 and 3 and this goes on forever. I've just proven two infinite things are the same – rooms and bars, but I still have not shown you why the zigzag line is the same length as the edge line.

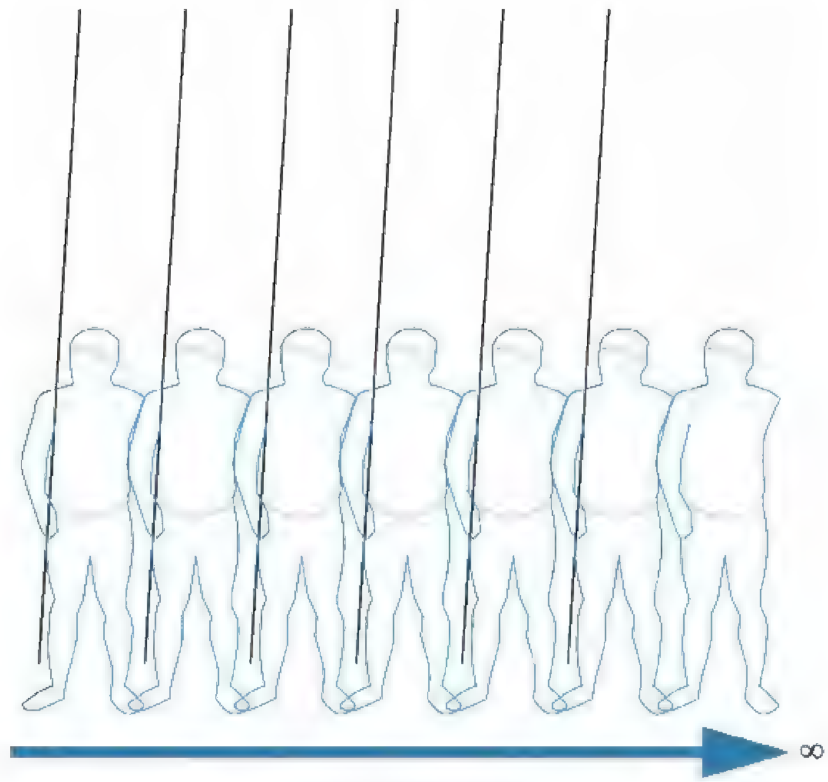
When you first explain infinity to a child they immediately ask "What's infinity plus one." A particularly smart kid I met, Dermot, asked, "What's infinity plus three?" Hilbert's Hotel allows us to answer this problem in a way we can visualize.



Traversing an Infinite Plane with a Line

The infinite hotel is full. A man comes to the front desk and asks for a room. The hotel manager says, "I'm terribly sorry, but we are full… But I may be able to help you. Let me think." He ponders for a moment and then says, "OK – I've found you a room." He calls the people in room 1 and asks them to move into room 2. He calls the people in room 2 and explains that due to a double booking they must move out of their room to let the people from room 1 in. But it's OK; they can move into room 3. Everyone moves up a room and the new guest gets the checks into the now vacant room 1.

This is a little harder to understand. We did not have a perfect one-to-one match as with the rooms and mini-bars. We had a mismatch of guests to rooms. But, we were able to show it is possible to re-establish a one-to-one match by doing something to every guest, having them move up a room. There is no problem with the last guest because it is an infinite hotel, there is no last guest! Another way to visualize the problem is to ask ever hunter to pass their spear to the right in the picture below.



Hunter with Spears

Provided there are an infinite number of hunters there is always someone to hand the spear to and the person at the front of the line now has space for another spear.

You can probably see how to answer Dermot's question. The hotel manager calls the guest in the first room and asks him to move 3 rooms up rather than one. He then calls the remaining guests and tells them the same thing. Thus, he has managed to fit three more people into the infinite hotel. Infinity plus 3 is infinity. You may worry that it takes the manager an infinite time to call all the rooms, but it's OK; he lives infinitely long so it all works out.

What about fitting an infinite number of new guests into the already full hotel? Surely then we will get stuck.

No, Hilbert's Hotel can fit an infinite number of extra guests. Here's the trick: ask all the people currently in the hotel to move to the room with double the number they are currently in – 1 goes to 2, 2 goes to 4, 3 goes to 6, and so on. Now all the odd numbers are empty and you can fit an infinite number of people into the empty odd rooms. Infinity plus infinity is infinity. Voila.

Now, a very clever or annoying student asks, "What happens if an infinite number of infinitely large buses arrive at the hotel. Can they all fit in?" The mathematical question is "does infinity times infinity, equal infinity?" Let us ask all the guests to get out of the bus and line up in the parking lot in neat rows. Passengers from bus one in line 1, those from bus 2 in line 2, and so on. All the guests now form a two-dimensional grid. We already know how to map a two-dimensional grid to one-dimension using the zigzag method. We can fit them all in the hotel and we are done!

Is Anything Larger than Infinity?

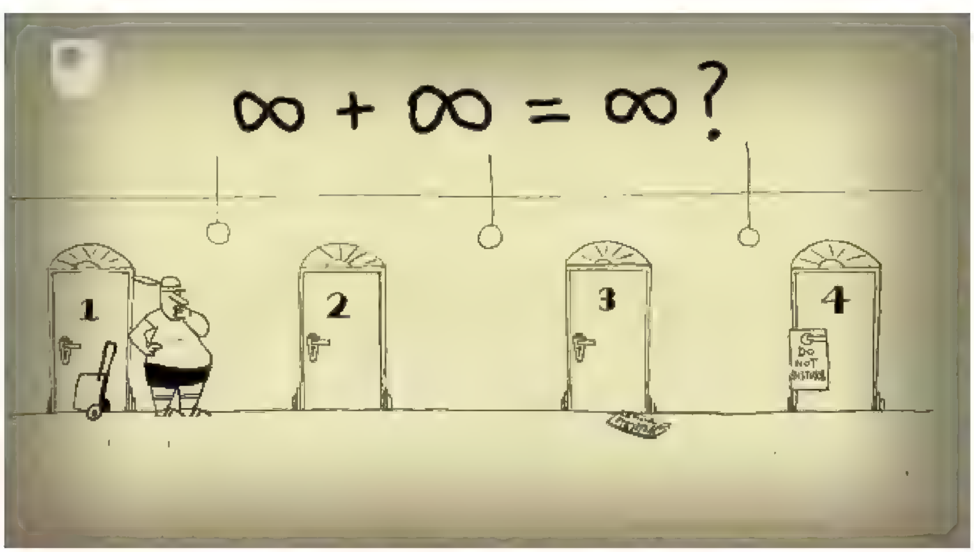
Is there any bus or combination of buses that would cause the manager of Hilbert's Hotel a problem.

The answer is yes and it involves a subtle change to the contents of the bus.

An infinite number of buses turn up but this time the buses are filled with men and women. The hotel manager is asked to put everyone in a room and once again he obliges using the zigzag method.

At the end of the process the tour guide comes to him. "I think you have missed some people," he says. "Since I am just one person, I know you can fit *me* in. But, I have a whole bus in the car park you completely missed."

"No," says the manager. "I did every bus."



Infinity Plus Infinity Equals Infinity

"Ah, no," says the tour guide. "The first bus you accommodated had a man in the first seat but this has a woman. The second bus had a woman in the second seat but this one has a man and so on. This bus has a different gender in at least one seat to every bus you so far accommodated. It is a new bus."

The manager finds room for the passengers from the new bus but the tour guide comes back a moment later.

"You have missed another bus. This one has a different gender in at least one seat to every previous bus, including the one you just accommodated. It looks like there are an infinite number of buses you missed, all lined up to get into the infinite hotel."

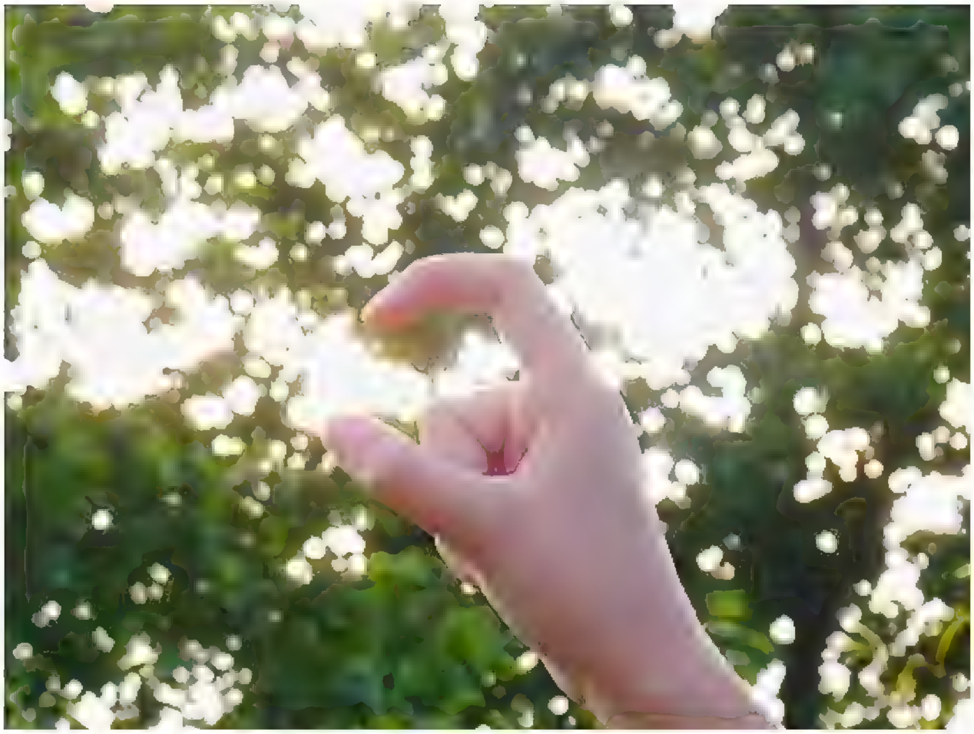
What is it about *these* buses that make them so difficult to accommodate? They are all just filled with people after all.

The manager is defeated by the more complex information held in the contents of the buses. An infinitely large bus full of binary information has more information in it than an infinitely large bus specified only by its size. This is a larger infinity than the counting infinity. The permutation of all the possible options for the occupants of the bus is larger than infinity.

Real Numbers

What about the real world we live in? Is the larger infinity we failed to fit into Hilbert's Hotel present, or was it just a mathematical fiction? Hold up your thumb and index finger for a moment. The gap between them is a distance. Most likely this is a whole number with an infinite decimal digits after it – say 2.2320394386…. centimeters. The infinite set of decimal digits in this measurement is the larger type of infinity: called the continuum. Distances in space form a continuous unbroken line of points, with no gaps in between. The counting numbers, on the other hand, form a broken line. We take discrete steps from one number to the next. This is a hard distinction to grasp but it is the same distinction we used in Hilbert's Hotel. Imagine you believe you have a list of all the real numbers in the world. You can take the first decimal digit from the first number and add one, the second digit from the second number add one and so on generating new numbers not on the original list. Therefore, you cannot have a list all the real numbers; they are not countable. Let's take a closer look at these real numbers.

Here's a quick test. Which is the larger number, the first or the second?



Holding a Real Number in your Hand

First: 3.1233249837583462136421472374
Second: 3.1233249837583462134421472374

You have 2 seconds to answer!
TRY ANSWERING WITHOUT READING ON

The first is larger. I changed one digit. Can you see?

Notice, you need time to read each digit and process the information. If you were an obedient reader and attempted it in two seconds you either guessed or gave up. Two seconds is too short to take in all the digits.

Let me give you another test. Again, I'll ask you the question, "Is the first number larger than the second?"

First
3.123324983758346213642147237516464646464636…
Second
3.123324983758346213642147237516464646464636…

I know you're looking for the difference but you won't find one, as I did not have time to write the numbers out in full. The $10^{20000th}$ digit is different, but even if I took the whole age of the universe and counted as fast as possible I would not reach this digit. Any number greater than, $10^{120}/10^{-43}$ digits cannot be distinguished from another in the age of the observable universe. Real numbers are in practice subject to an uncertainty principle. Some mathematicians even wonder whether they really exist. But, they do exist in our minds and our thought experiments. In my view, any model of the Universe that ignores them is likely to be wrong.

Random Numbers

Which of the following numbers is random?

1111111111111111
34289460370124001
49293741762343083

THINK ABOUT YOUR ANSWER THEN READ ON

Each of the numbers could be random. There is no reason any set of 10 digits is more likely than another, but it feels very unlikely that if I tried to generate a random number I would get 15 consecutive digits. What a human means by random is a jumbled up number: one with varying digits that have no real pattern. An American mathematician, George Chaitin has been able to explain this by saying that a random number is uncompressible. This means there is no way to describe the number more efficiently than writing it out in full. A string of ones can be compressed. "Write a million 1s" takes only 18 characters, yet accurately describes a number that is a million digits long. By contrast 8988376132 can't be compressed very much at all, its information is just a jumble. There are many interesting numbers around. Some numbers are *Hamlet*; some numbers are pi. One interesting number is the following: 17733173332032037377. It is the genetic sequence for the virus smallpox, or at least the first 20 digits. Copies of the full sequence sit under lock and key in the Pasteur Institute in France and the CDC in Atlanta. This number is a candidate for an 'evil' number. You might think there are many numbers that could represent smallpox because there are

Smallpox Virus

many languages in the world and many ways you could code the genetic sequence of GATC. But, there *will be* one most efficient binary coding for smallpox and that number is the nearest we have to an evil number.

The other important element of random numbers is the process by which they are created. Computers can't genuinely generate random numbers. The numbers they generate are predictable and eventually

Child Survivor of Small Pox

repeat. To create the random number in my example above I went to www.random.org, a website that uses fluctuations in atmospheric quantum noise to generate random numbers. As far as we know quantum effects are truly random and have neither rhyme nor reason.

Numbers are more complex than they first appear. They are infinite, yet there are different infinities, and they have meaning. The smallpox example above and the Turing numbers we will discover shortly suggest numbers *do* have meaning independent of culture and language. The next two chapters will show us what happens when we think about the meaning of numbers. We will also explain one more 'super infinity' and this will be the key to understanding creativity.

> *"There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns - the ones we don't know we don't know."*
>
> Donald Rumsfeld
> United States Secretary of Defense
> (2001-2006, 1975-1977)

Chapter 9

KNOWN UNKNOWNS

Donald Rumsfeld

In the spring of 1981, London staged its first marathon. The field of runners included 1200 international athletes and 20,000 amateurs. An estimated 20 million viewers watched from around the world. The top international runners stayed together for the first twenty miles and then two runners, American Dick Beardsley and Norwegian Inge Simonsen, made a push for the finish. They were long-standing rivals and, as they ran the final mile each man challenged the other to see if they could get ahead and gain the advantage. Because of the fine balance human muscles maintain between anaerobic and aerobic metabolism, the small set advantage could prove insurmountable. The other runner would need to sprint to catch up and the resultant lactic acid generated would turn their legs to jelly. As the two runners neared the finish line they glanced at each other, smiled, reached out and held hands as they crossed the line. Who won?

We all instinctively know the answer. The race was a draw, but the rules of the International Athletics Federation are clear. Read rule 164.

RULE 164

The Finish

1. The finish of a race shall be denoted by a white line 5 cm wide.
2. The athletes shall be placed in the order in which any part of their bodies (i.e. torso, as distinguished from the head, neck, arms, legs, hands or feet) reaches the vertical plane of the finish line.

The organizing committee held a brief conference and the result declared a draw. They had interpreted the rules in the same way 20 million TV viewers already 'knew' to be true.

This story should set your minds thinking about the nature of *rules* and *truth* and how the two are often different. According to the *rules*, one person crossed the line a little ahead of the other. The *truth*, as we all instinctively know, is that the race was a draw. Maybe the rulebook is missing a rule – 'The contact draw rule'. Clearly you could amend the rulebook to add this one rule. I checked the current athletics rules and they don't contain this amendment. If the rules were amended the mischievous amongst you will realize an unsporting athlete could grab the hand of their opponent as they crossed the line to force a draw. The rules would have to stipulate that holding hands must be voluntary for both parties, and refinements could go on for some time. What if I held your hand but you tripped and let go? What if my attempt to hold your hand

Kurt Gödel

caused you to trip? You could go on forever, generating rules to cover every eventuality.

Clearly, in the fuzzy world of human endeavor, truth and rules often part company. Yet, we all assume mathematics is free of such uncertainty. Let me tell you this is not so. The brilliant mathematician Kurt Gödel proved this when he was just 22, and his proof says something fundamental about the nature of knowledge.

The story of his discovery involves some of the greatest mathematical thinkers in history. My introduction to it came about from a chance accident. I became ill in my first year at University (mononucleosis, otherwise know as glandular fever, if you're curious) and was eventually sent home to recover. Lying in bed for two months is boring. So to pass the time my mother suggested I read Bertrand Russell's, *The History of Western Philosophy*. I think she figured I had plenty of time, so picked a thick book. This nearly 800 page tome charts the entire history of philosophy from the time of the ancient Greeks. I presumed Russell was a philosophy professor, but he was originally a mathematician. He was a mathematician. And because he lived and worked productively for almost all of his 97 years, spanning much of the 19th and 20th centuries, he crops up repeatedly as a central figure in many areas of intellectual life. Russell the politician, Russell the philosopher, Russell the mathematician and Russell the peace campaigner are all the same man – not, as I had incorrectly first guessed, a prolific family. In his early career, Bertrand Russell was a Fellow of Trinity College, Cambridge, working on a broad range of mathematical problems. Meanwhile, in Germany, his contemporary David Hilbert, also a polymath, held the chair of mathematics at Göttingen University. Both men shared a common objective: to tidy up the loose ends in mathematics and set down the rules once and for all. This movement was called Formalism.

Formalism

David Hilbert and Bertrand Russell believed you should be able to set out all the rules of mathematics even though it might be a complicated affair. Without contradiction or inconsistency you should be able to write down the rules and then play the 'game of mathematics' to derive every possible truth. Hilbert despised the idea that there could be unknowable things and was a forthright speaker. His battle cry was: *Wir müssen wissen — wir werden wissen!* "We must know — we will know!" He believed there were no fundamental unknowns in the world.

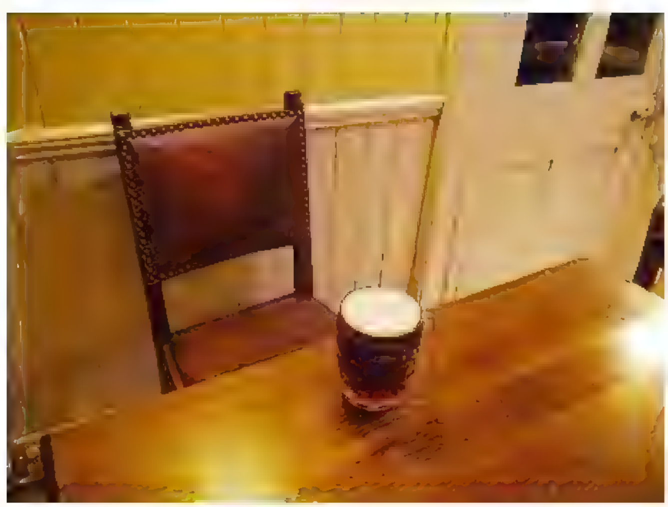
Donald Rumsfeld famously summed up the problem of unknowns in an attempt to clarify a question from a journalist at a Whitehouse press conference:

> *"There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns - the ones we don't know we don't know."*

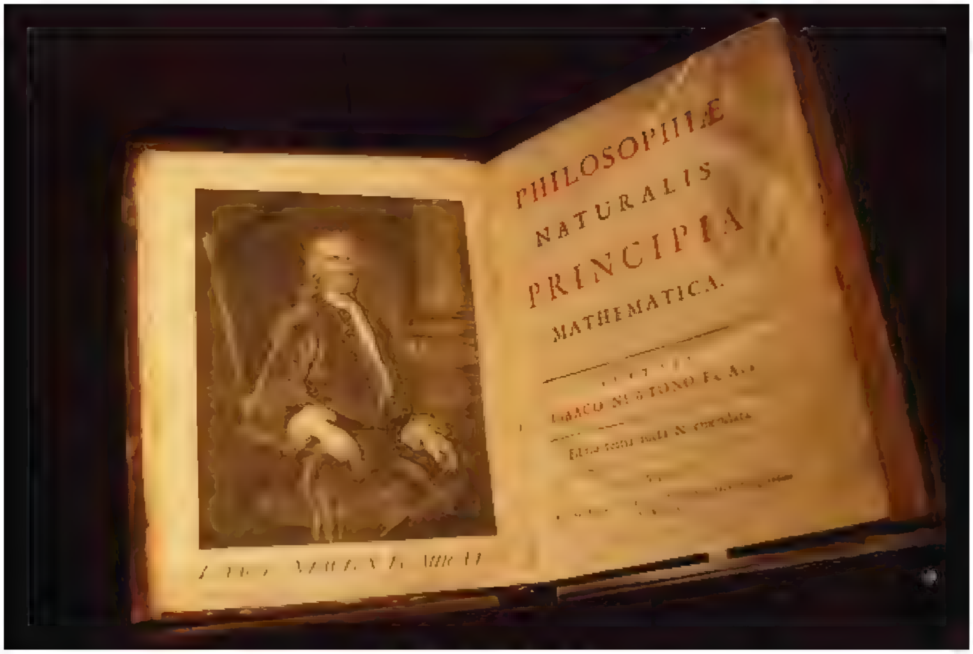
Interestingly Donald Rumsfeld, like Bertrand Russell, is another person to span a huge swath of time in the public eye. He was both the youngest and the oldest serving U.S. Secretary of Defense, serving under both Richard Nixon and George W. Bush. We will shortly discover Rumsfeld's convoluted view of the world turns out to be closer to the truth than Hilbert's tidy mathematical aspiration.

As well as believing there were no unknowable unknowns Hilbert thought mathematics was completely abstract. You did not need to know what you were talking about. Whether the symbols meant dogs, cats or numbers all you needed to do was apply the rules and all would be well. His belief is captured in his quote below.

"It must be possible to replace in all geometric statements the words point, line, plane by table, chair, beer mug."
David Hilbert



Geometry with Beer and Furniture



Newton's Principia

PM

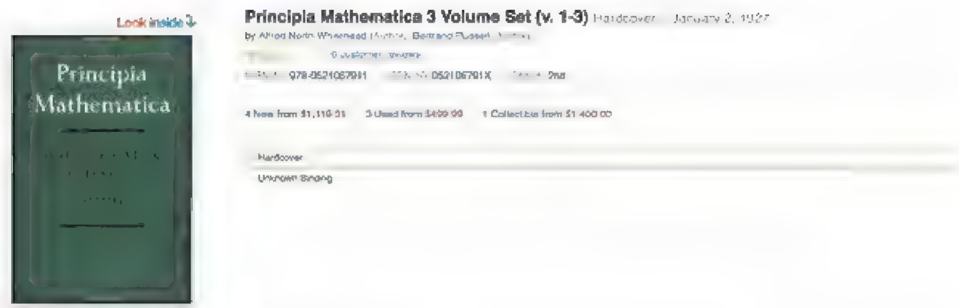
In 1890, the Cambridge mathematicians Alfred North Whitehead and Bertrand Russell embarked on the mammoth task of writing out all the rules of mathematics and publishing them in a set of books called *Principia Mathematica*. Every rule is written down in meticulous detail. The books are heavy going and look like more like computer programs than text. They set out precisely what you can, and cannot, do with numbers, and are the most impenetrable textbook you will ever read. Just to give you a flavor here is one line where Russell proves 1+1–2. It has taken about 100 pages of densely packed equations to get to this point!

$$*54\cdot43.\quad \vdash :. \alpha, \beta \in 1 . \supset : \alpha \cap \beta = \Lambda . \equiv . \alpha \cup \beta \in 2$$

From this proposition it will follow, when arithmetical addition has been defined, that $1 + 1 = 2$.

One Plus One Equals Two, PM

PM is a 3-volume set of books. Volume One costs £480 on Amazon. This is a significant work and a collector's item. The last time a first edition volume came up at auction in 2007 it went for over £800. Cambridge University Press printed only 750 copies and I suspect they



Amazon Listing for Principia Mathematica

are undervalued. When mathematicians use the letters 'PM', they are usually referring to Russell and Whitehead's *Principia Mathematica* rather than the afternoon.

Hilbert's Problems

In 1900, while Russell and Whitehead were in full flow writing out their rules, David Hilbert was invited to deliver the annual lecture at the International Congress of Mathematicians in Paris. He asked a mathematician friend what subject he should pick for the talk and, in a moment of inspiration, the friend suggested laying out a vision for the future of mathematics. Rather than tell people how wonderful mathematicians were, and why their discipline was the pinnacle of human scientific endeavor, why not try modesty and list all the problems on which they were stumped? Hilbert liked the idea and devoted his talk to all the problems he thought mathematicians would solve in the 20th century. Hilbert's Problems were simply an intellectual challenge. He offered no prizes. At the turn of the 21st century, the Clay Institute created the Millennium Prizes for solving the most important modern mathematical problems. Each solution wins a prize of a million dollars!

There are 23 numbered Hilbert Problems in all: ten in the original lecture and a further 13 in the written transcript. In 1928, he clarified the 2nd and 10th problems, refining them into three distinct questions: Is mathematics consistent, complete and decidable? Ironically this means that Hilbert's 23 problems actually number 24! The most important Hilbert questions where these last three. They ask whether Russell and Whitehead would be successful – can you write out all the rules of mathematics and then simply calculate the answer to any problem or derive any proof. This is known as the Decision Problem. Can you mechanically decide any mathematical question without doubt? To explain Hilbert's Problems, I need to define mathematics properly.

Giuseppe Peano, Mathematician

"A mathematician is a blind man in a dark room looking for a black cat which isn't there."

Charles Darwin

The Game of Math

One of my most vivid childhood memories is driving my mother distraction by asking the 'why' question. Most children go through this phase:

Me: "Why is a sponge wet?"
My mother: "Because it has soaked up water."
Me: "Why has it soaked up water?"
My mother: "Because it has small holes in it."
Me: "But what makes water wet?"
My mother: "Because it is made of wet stuff." a bit weak now.
Me: "What is wet stuff?"
...

You can ramble on indefinitely unpeeling a never-ending onion. Sometimes, if you are unlucky, you may get stuck in a loop. For example, "where did the chicken come from?" "An egg," "and where did the egg come from?"...

Mathematics breaks this cycle!

In mathematics, there is no danger of an infinite number of 'why' questions because at its core are a clearly defined set of absolute rules called axioms. You cannot ask the 'why' question of an axiom. It is a RULE!

Starting from an absolute minimum of fundamental rules everything else is built up so that no step requires any leap of faith nor generates any contradiction. Let me give you a concrete example and, in the process, show you how numbers are defined.

Numbers

It was not until the late 18th century that numbers were properly codified. The mathematician Giuseppe Peano gave us the rules, so they are called Peano axioms. Here are his 'axioms' in natural language.

Peano Axioms

1. The first number is named zero.
2. Every number has a next number (called its successor). Example: the next number after one is two.
3. Numbers are singular. Every number with the same name is the same thing.
4. If something is true of a number, it should be true of the next number (the successor number).

From this we can prove some very simple things.

1+1−2. Because the next number after 1 is 2 and '+1' means take the successor. (You can see I cheated here a little and did not take 100 pages for the proof.)

Back to my poor mother: "Why is the lowest number zero, Mummy?" "Because I say so!" Or, at least "…because Mr. Peano said so." That's what an axiom is.

"OK, but why is 3 greater than 2."

"Because I said that each number has a thing that comes after it.

"But, why can't 3 come after zero!"

"It can!"

"But then, if 3 is the thing after zero, I could count 0, 3, 2, 4…"

"Yes, if you want to…"

"I'm sort of lost. Now, you are saying that 3 doesn't really 'mean' anything. It just comes after 0."

"Yes. You can make up any symbols you like. You just have to remember what you said and be consistent."

The dialogue shows the importance of definition in mathematics. I could define my counting numbers as 0, 1, 2, 3, 4 or as ο, π, ρ, σ, ς, or 享, 仇, 仕, 仝 or to be really annoying and confusing 0, 3, 1, 2, 4; they are only arbitrary symbols. It helps us to learn the numbers because 1 is a single line, 2 is two lines joined, three is basically three lines looped together, and four is four lines, but we could have used any symbols we cared for. It is the rules for manipulating these symbols that are the important part and give mathematics its meaning.

The Game of Mathematics

When I was a child, our living room carpet had a square pattern. You could use boiled sweets to play checkers on it. Even though there was no board and no pieces, it was clearly a game of checkers because we followed the right rules (with the one exception that if you jumped over a sweet you got to eat it). Mathematics is like a game with a set of rules. If you follow the rules, you are doing mathematics.

Consider the simple mathematical theory that if A equals B, then B equals A. This seems clear-cut, but you can get into trouble if you're not careful when defining the word 'equals'. 'My dog equals naughty' does not imply 'naughty equals my dog. Here I have used 'equals' to mean 'has the property of.' My dog has the property of being naughty. This is an attribute, not equivalence. You must be careful with mathematics. A equals B implying B equals A is a property of numbers when the equals sign is used to mean equivalence.

Here are the rules of the game that provide a proof for this theory.

Let us start with the position in which we don't know whether A equals B implies B equals A. We have these three axioms, call them rules for now since we are using the game analogy.

Rule 1: If I have no minus sign in front of a letter I can assume there is an invisible + sign there.
Rule 2: If I have a positive letter (or a letter with no symbol in front of it) I can put a minus in front of it and put it on the other side of the equals sign.
Rule 3: I can swap the plus and minus signs of all the letters in my equation if I do it to all of them.

Now I am ready to prove my theorem.

A = B is the same as +A = + B. (rule 1)
+A = + B is the same as -B = - A (rule 2 done twice)
-B = -A is the same as B = A (rule 3)

Success.
So A = B is the same as B = A.

I have my proof. It might be glaringly obvious, but that's not the point. The point is you can apply rules to symbols and derive new rules. It does not matter what the symbols are or how obvious it is. Here's the same proof with dingbats.

Rule 1: If I have no glyph in front of a symbol I can assume there is an invisible Ψ there.

Rule 2: If I have a positive letter (or a letter with no symbol in front of it) I can put a ♮ in front of it and put it on the other side of the ➤

Rule 3: I can swap the Ψ and ♮symbols of all the symbols in my equation if I do it to all of them.

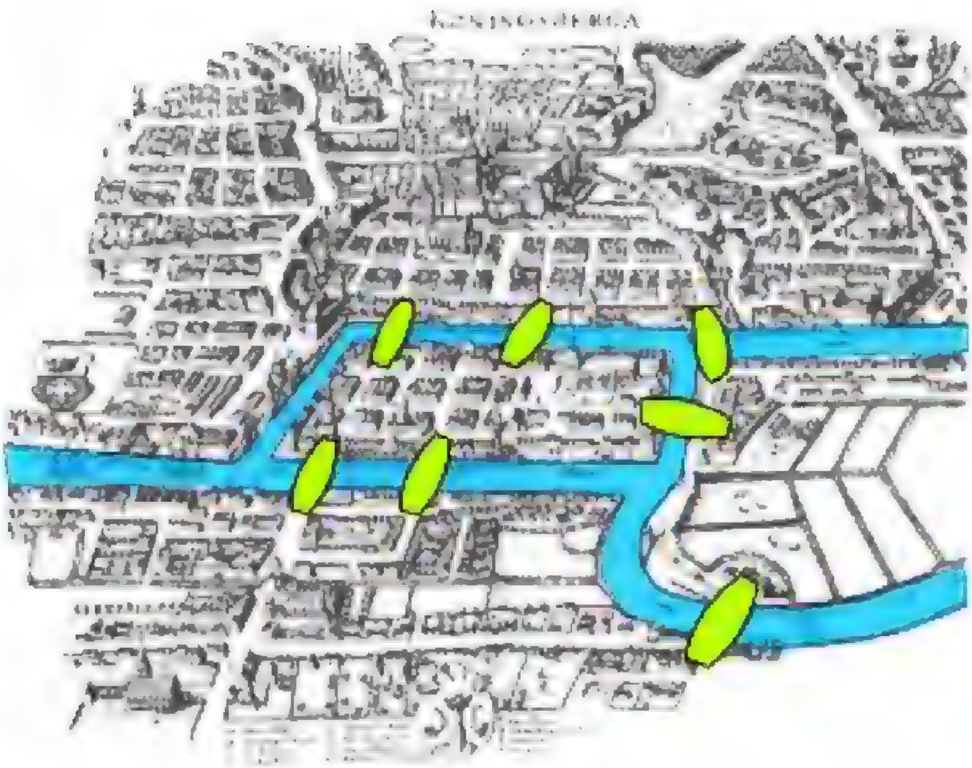
The proof in symbols

> ♈ ➤ ß is the same as Ψ ♈ ➤ Ψ ß. (rule 1)
> Ψ ♈ ➤ Ψ ß is the same as ♮ß ➤ ♮♈ (rule 2 twice)
> ♮ß ➤ ♮♈ is the same as ß ➤ ♈ (rule 3)

Any collection of symbols will do. The symbols have no meaning in themselves other than the meaning we have given them. A tribe in the Amazon jungle could demonstrate a proof without knowing any mathematics. All I need say is, "Hey, I want to play a game with you. Can anyone make this into that, in the fewest possible steps, while obeying these rules?"

But, is it true we can ignore the meaning behind the symbols. Does it matter that we were talking of numbers rather than spears, counters, or crocodiles? If we look at the marathon winning analogy again, we know the nature of a game is important. In a running race we can interpret holding hands to mean the two athletes are treated as one, the existing rules can then be applied as normal and the pair become a single winner. But, in tennis, there would be a problem. I wouldn't want to come on court and find I'm playing against two opponents! On consideration though I'd be happy if they had to hold hands while they played so that they constituted a single player. When we examine the actual circumstances, we can add a rule and show the rule works, but we have to see something about the specific sport that makes the rule fair and workable.

Hilbert was convinced mathematical truth is not like this and that proofs follow from the rulebook without any knowledge of the circumstances, i.e., the sport being played or any other analogous thing. He was to be proven wrong by Kurt Gödel.



Königsberg Bridges

Gödel

Gödel studied mathematics at Königsberg University, Hilbert's hometown. Königsberg is famous for having a mathematical problem related to the seven bridges that link the city together. It's quite fun to try to solve. Find a route across the city that crosses each bridge once and once only. You can start anywhere, but no walking halfway over a bridge and no swimming!

Euler discovered a rigorous mathematical proof that there can be no solution in 1735 after five hundred years of failure by other mathematicians. The answer is you cannot.

In 1931 Kurt Gödel, then working at the University of Vienna, proved mathematics *is* like our sporting analogy. There are true statements in mathematics that cannot be proven by the rules of the system. Someone outside the system, with common sense, can see a statement is true, but it's impossible to prove this if you constrain yourself inside the system. It is the equivalent of all the members of the London Marathon Committee wondering what to do about the race while all of us watching the TV are shouting, "It's a draw!" Looking at the rulebook 'really hard' doesn't help.

You have to step back and think about the problem in the round and then devise some additional rules to handle the circumstances. Mathematics is like this also.

Here is how Gödel proved his result.

It is easy to turn logic or any text into numbers. That's how this book is stored on my laptop. All we need do is translate sentences into ASCII or Unicode. In this way, any theory can be reduced to a string of numbers.

Since Gödel's proof predates the invention of the computer, he had to come up with a novel way to store information. He deployed an old Roman invention; a substitution code. The number one was represented by 1, two by 2 and the symbols by larger numbers, for example, '=' was coded as 15 and so on. He then raised a sequence of prime numbers to the power of each of these codes and multiplied all the results together. This generated a single enormous but unique number that he could later factor back into its constituent parts to recover the information. This is a truly complicated solution to a very simple problem. Today we would solve it by storing each number in the memory of a computer as an array.

Let's use the easier table method to store things and code as follows: 000 will stand for 'start of proof'. Each step in the proof will start with 00 and each symbol in the proof starts and ends with a zero. This way we can code one plus one equals two as follows.

0000001110454011101210222000000

I think this is simple enough for you to guess the coding scheme. Hint: 111 stands for 1. The scheme is on my website if you can't work it out. Using this technique, any series of mathematical statements can be turned into a number. As a series of mathematical statements is a proof, we can generate proof numbers. They are just the sequential list of all the instructions. These numbers are sometimes referred to as Gödel numbers.

Gödel's next step was to say one number *demonstrates* the proof of another number. For example, the number 000820962 might demonstrate the proof of another number 000398... This is the mathematical equivalent of my saying a Word file demonstrates the truth of your mathematical theorem. Any statement can be represented by numbers, provided you have a consistent coding scheme that allows you to get back to the meaning.

Now Gödel set up his paradox:

Every correctly formed theorem number has another number, which demonstrates the proof of that number.

If this is universally true there should be no contradiction. Unfortunately if you apply the theorem to itself you get something similar to the liar's paradox.

"This proof number is not a proof of the truth of this theorem number."

The proof number proves the theorem number is true, but the truth of the statement is that it can't be a proof of the statement… Paradox.

The only way to resolve the paradox is to go back one step and realize that not *every* correctly formed theorem number has a proof number using only the rules of that system.

Concisely, Gödel's theorem says, "Within any formal system of mathematics there can be statements that are true but are not provable using only the rules of that system."

When Hilbert heard of Gödel's proof, his first reaction was anger. After all, he had spent 30 years of his life trying to prove mathematics was tidy and complete. Gödel had just shown it was not. Hilbert never worked on formalism again, but the rest of the mathematical establishment largely ignored the result. Gödel's proof did not stop mathematicians proving new theorems nor doing useful mathematics. They went on much as before, using a mixture of intuition and analysis. The only difference was someone had told them analysis alone would not succeed. The repercussions of Gödel's theory have more to do with understanding our place in the Universe and the nature of knowledge discovery. These are 'big' philosophical questions, which don't greatly affect the day-to-day ability of a mathematician to do their job. However, it is important to understand that knowledge discovery is not simply analysis. Knowing this helps us understand human creativity.

Inconsistency

In the proof above, I said the only way to resolve the paradox is by saying there cannot be a proof number for *every* mathematical statement and therefore mathematics is incomplete. There is one other way to solve the paradox, and that is by allowing inconsistency into the system. Gödel's proof assumes you can prove something true or false, but what if you could prove it true *and* false? In this case, the system is complete but you can prove truths and untruths within it! This may seem an acceptable solution, but inconsistency in a mathematical model is a cancer that will spread through the entire body. Think about it. If I am allowed to prove anything either way, of course, my system is complete. It can say anything it wants, but the proofs I make are worthless.

Let us imagine, for a moment, we created a new system of mathematics where all the numbers in our new theory behave as we expect, except for the numbers 5 and 6. You may use them to count, but they are also *equal* to each other! This feels bad and it certainly breaks the Peano axioms. In my new system 1 plus 5 and 0 plus 5 are the same, so I can equate 0 to 1. Because 0 and 1 are the basis of binary arithmetic, all numbers can be equated. Numbers now have no guaranteed meaning in my system and, what is worse, since logic uses 1 and 0 to represents true and false, all of logic falls apart as well. Whenever we allow inconsistency into mathematics it rapidly brings the whole pack of cards down.

The example I gave was glaring; an inconsistency right in the middle of the counting numbers! Maybe I was too aggressive and a subtle and less damaging inconsistency might be tolerable. However, *any* inconsistency allows me to make zero equal one somewhere in my system and, therefore, any theorem based on proof by counterexample will be suspect.

There might be systems where inconsistency could be a legitimate part of a mathematical system, but I would always need positive corroboration for each proof. If I tried hard enough, I could always prove something either way. I would need to formulate a new mathematical rule – something like "I will believe short, sensible-looking proofs to be right and circuitous proofs to be wrong." Mathematics would be a bit like a court of law. You would have to weigh up the evidence from a variety of sources and the verdict would be a matter of subjective opinion rather than objective fact. Inconsistency is *very* bad in mathematics.

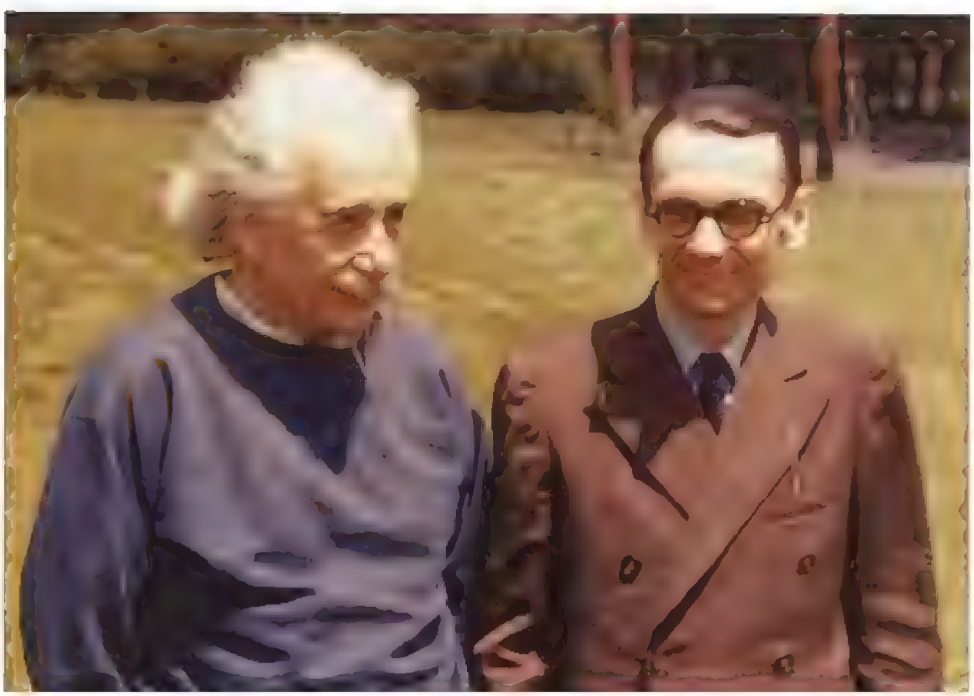
The Lucas Argument

J.R. Lucas of Oxford University believes Gödel's theorem says something fundamental about the nature of the human mind. In 1959, he wrote a paper, *Minds, Machines and Gödel*, where he argued humans must be able to think outside a fixed set of formal rules. The paper has been causing arguments ever since. Strong AI proponents have a visceral reaction to it. Forty years later, in 1989 Roger Penrose picked up the baton and put the Lucas argument on a stronger theoretical footing. The Lucas-Penrose argument is this:

If humans used a formal system to think, they would be limited by the incompleteness theorem and unable to discover new theorems that required them to extend the formal rules. Humans do not appear to have such a limitation and regularly extend their appreciation of mathematics by expanding the rules, and seeing through to the truth.

Many scientists dislike this argument and think it farfetched, saying there is no evidence to show people see past the limitation. Our brains could be following a formal system capable of discovering everything we have discovered to date or, indeed, might encounter in the future. Why should we assume human minds are constrained in the same way as the mathematical systems they discover? There is no evidence to suggest a human *thinking* about Peano arithmetic is running a Peano based model in their head. When Peano discovered his theorem he was certainly extending our mathematical knowledge, but this does not imply he was extending the capability of his brain.

The critics of Lucas and Penrose have one big problem to deal with. The formal system in our head would need to be able to see the truth in everything we could *ever* encounter. But, our formal system appears to be small. As infants, it is almost nonexistent. Where does this enormous system come from? It can't come from our parents because they have the same problem; they were once children. You might argue that the capability of the human brain is huge and we can learn from all the other humans on earth, but let me remind you what Gödel said. However large



Two Giants

a system you have and however much you extend it, the system will always be incomplete. And we really do mean; however large. Even an infinitely large formal system would be incomplete.

The only way to avoid this problem is with some sort of conspiracy theory where we only come across problems our formal system can already solve. Such a theory is a determined Universe. In a determined Universe, all the mathematical problems we ever solve must be expressed by the formal systems existing in the Universe. We must never encounter a problem where we need to extend the system and break the Gödel limit because we are pre-determined not to do so.

The Inconsistency Defense

An argument put forward by opponents of the Lucas-Penrose position is that humans are inconsistent formal systems. Inconsistent formal systems are not subject to the incompleteness limit. Humans certainly behave inconsistently with remarkable regularity but simply making inconsistent statements is not sufficient to show the underlying formal system is, itself, inconsistent. Inconsistent beliefs can come simply from making mistakes or reading the same story in two different newspapers! We need a fundamentally inconsistent thinking mechanism inside our brains to break the constraint. The very machinery itself would have to be inconsistent. But this is exactly Penrose's point. Constructing a machine capable of reasoning in an inconsistent but useful manner would need exotic technology, some sort of non-deterministic, rationalizing computer. The components to make it could not be computer logic as we know it today. All such logic is entirely computationally deterministic.

Let me see if I can reframe the Lucas argument. Imagine IBM's Watson computer was let loose on mathematical reasoning. Watson could scan every mathematical theorem ever written down. It would know every programming language created. It would have its enormous bank of general knowledge to call upon and it could answer many questions. It would sometimes appear inconsistent because the information it had trawled from the Internet would be wrong. But Watson would still be a consistent formal system and Gödel's theorem says there would be truths Watson could never see. Lucas argues humans can see such truths where a machine cannot, and these truths would allow a human to discover a proof to a mathematical problem that would forever elude Watson.

The Lucas argument runs into a brick wall because it asserts we see truths a machine cannot. For each alleged creative step, his opponents simply assert your brain was already sufficiently powerful to perform that creative step. Lucas's argument is largely a philosophical one. Surely all this creativity can't all be pre-coded within the brain. Surely we must be extending our model in order to extend mankind's mathematical model. "Prove it," say the detractor, and he cannot. We need something more practical if we're going to show a difference between humans and machines - something an engineer, or even a physicist, could grasp! That thing is a Turing Machine. We will examine this next.

Chapter 10

TURING'S MACHINE

Alan Turing

"A computer would deserve to be called intelligent if it could deceive a human into believing that it was human."

Alan Turing

"The only real valuable thing is intuition."

Albert Einstein

"Mathematical reasoning may be regarded rather schematically as the exercise of a combination of two facilities, which we may call intuition and ingenuity."

Alan Turing

It is 1943 and a small group of Polish mathematicians sit, ears glued to their wireless set, waiting to hear whether the German army will advance on Warsaw. The Polish Intelligence Bureau badly needed to know what the German army was planning and had recruited this group of young mathematicians as code breakers. Up to this point, code breaking had been the domain of linguists able to see word patterns in apparently random sets of letters. The arrival of electro-mechanical machines made this method redundant, and code-breaking had become the domain of mathematical minds. The British, French, and American intelligence agencies were all hard at work deciphering the German codes, but only the Polish group, motivated by the imminent threat of invasion, had made real progress. The code they were breaking: 'Enigma'.

As with many inventions, Enigma got off to a difficult start. The inventor, Arthur Scherbius, tried to sell it to the army but they rejected it saying it did not provide any real military benefit. Instead, the machine went into service transmitting commercial shipping manifests. However, some senior figures in the German military had not forgotten the lesson of the First World War. During that war, the German army suffered major setbacks because the British broke all their codes early on. With the onset of World War II, Rommel ordered the German Army and Navy to deploy modern coding machines. The previously rejected Enigma was rapidly pressed into service and, all of a sudden, Europe went dark to Allied Intelligence. The man to lead the task of breaking Enigma for the English was Alan Turing.

Alan Turing

Alan Turing was conceived in India but born in London in early 1912. He was precocious from an early age and an extraordinarily determined character. His first day at Public School, Sherborne in Dorset, coincided with the British General Strike of 1926. With no public transport available, the thirteen-year-old Turing cycled the 60 miles to school, staying in a guesthouse on the way and earning a write-up in his local newspaper. Turing went on to study Mathematics at King's College, Cambridge and was made a Fellow at only 22. In 1936 Turing, aged 24, published *On Computable Numbers and their Application to the Entscheidungsproblem*, not a snappy title, but one of the most influential mathematical works of the 20th century. The paper described the new the science of computing and solved Hilbert's 'Entscheidungsproblem', a mathematical puzzle simply translated as 'the Decision Problem' – could you decide the truth of a mathematical statement using some sort of automatic computation – an 'algorithm' as we now call it?

It is difficult to imagine, but Turing worked on 'computing' before the invention of the computer. When he talked of computing, he meant the abstract idea of doing something mechanically. The nearest thing he had to a 'computer' at the time was a human mindlessly but methodically calculating something with pencil and paper! The scientific paper he submitted to the London Mathematical Society described both the theoretical basis of computing, and the design of a general-purpose computing machine: the forerunner of all modern computers.

At the time, only a handful people in the world could assess Turing's paper. One of them, Alonzo Church, was based at the Institute of Advanced Mathematics in the USA on the Princeton University campus, next door to the Institute for Advanced Study that housed Einstein. Turing travelled to America in 1937 and completed his doctoral thesis at Princeton. He might have stayed, but Europe was heating up and war seemed inevitable, so Turing returned to England to take up a part-time job in the government code-breaking branch. Here he was able to indulge his passion for hands-on engineering, experimenting with the newly invented valve technologies. When war finally broke out Turing was ordered to report to Bletchley Park, just north of London. This was to be the home of the top-secret British code-breaking group tasked with cracking Enigma. Turing's first task was to debrief the Polish mathematicians and see what they had discovered. The Polish mathematicians had seen there were flaws in Enigma that made it repeat itself. They had made a copy of the machine to test different coding configurations and had been routinely cracking Enigma for 6 years, but the Germans had been getting smarter and it was taking longer and longer to crack the codes. Turing realized he could apply the Polish ideas in a more general way and break the codes on an industrial scale. He was installed at Bletchley Park to lead the project.

Initially he was successful but as the war continued, Enigma developed subtleties making it harder to break. At one point, it was taking a whole month to break a single day's messages. Turing realized the only solution was to use computer technology to fully automate the decryption. He built a computing machine that could simulate thousands of Enigma machines and try out all the possible settings in a short space of time. The machine acquired the nickname 'a bombe', perhaps because of the ominous ticking sound it made as it calculated (or maybe as a reference to the smaller Polish machines).

Thanks to Turing's insight into coding schemes and the machines he designed, the British were soon able to read almost every coded message the Germans sent during the war, giving the Allies an enormous advantage. The D-Day invasion involved convincing Hitler that the Allies had a huge army of nearly 400,000 men, massed around Dover preparing an attack on Calais head on, with a second army in Scotland poised to attack Norway. In truth, they had only 150,000 men planning an assault on the Normandy Beaches in the South. Just before the landings messages were decoded showing Hitler had fallen for the Allied subterfuge. Even as the Normandy landings began, Hitler still thought this a bluff and kept his 28 divisions at Calais waiting for the imagined attack. Without this intelligence advantage, the Allies would have needed a much larger invasion force, and Churchill believed Turing's work shortened the war by as much as two years.

The cracking of Enigma remained a secret after the war and Turing's story remained untold for many years. When Churchill wrote his history, *The Second World War*, a massive work in six volumes, all sorts of sensitive information featured, but Turing's work was omitted. One sentence hints that Churchill might write something about it in the future, but he never did. Churchill considered the work at Bletchley Park so sensitive he had it put in the highest classification extending the 30-year secrecy rule. We must presume the decoding schemes were still being deployed during the Cold War. The papers were finally released in 2010.

In one of those sad turns in history Turing was found guilty of gross indecency for homosexuality in 1954, a criminal act at the time, and was prescribed hormone treatment. This affected his mental state and he took his life by eating an apple laced with cyanide. He was eventually honored posthumously as a war hero and one of the most significant thinkers of the 20th Century. A Turing Award is the equivalent of the Nobel Prize for Computing. He was given a royal pardon in 2013.

To see how Turing came up with the idea for the Turing machine and solved the decision problem, we need to get a feel for theoretical mathematics. That might sound a little heavy going but don't worry, I will use a simple piece of mathematics to explain, one we have all played with as children, secret codes.

Codes


Enigma Machine

Everyone has played with some sort of secret code as a child – the Aggy Waggy game, passing notes written in invisible ink made from lemon juice, or perhaps a simple cypher. If I want to send you a secret message, I can use a substitution code. Let's see how good a code breaker you are. Can you decode this?

Gdkkn Qdzcdq

It's really easy. You might guess the message from the pattern of letters and your knowledge of my writing style. There are a couple of interesting patterns to note: the 3rd and 4th letter of the first word are the same and the first and last letter of the second word are the same. As a test I gave this code to my wife and my eight-year-old daughter to see how long it took them to decode… Less than a minute for my wife – a linguist. We will come back to my daughter shortly!

Roman Emperors used this sort of simple code to secure their messages, but modern codes have to be a great deal more sophisticated. Let us use a progressive cipher where we vary the substitution using a secret word. Take the name of my dog and write it down repeatedly next to the letters of the message you want to keep secret. Now translate all the letters in the message and the code into numbers 'a' = 1, 'b' = 2 and so on. Then add the letters of my dog's name to the letters of the message one at a time. If I get to 26 ('z') just wrap around to 'a' and carry on. This is called modulo arithmetic. This coding scheme will translate 'l' to 'a' the first time but 'l' to 'c' the second making it much harder for a linguist to see any pattern.

hello reader can you read this code
georgegeorgegeorgegeorgegeorgegeorge

Gives

ojacveyjpvlwghpegcvzoilfkehzpxghcvle

The advantage of this cipher is that I can easily remember the name George. I don't need to write it down. And the circular application makes the message sufficiently obscure you can't easily work it out…

Is this, therefore, a good code?

No.

This cipher is easy to break. Once you have guessed that I have applied a repeated short code word, you can write out ALL the possibilities and decrypt my message! This may be tedious, but if you are fighting a war and your life depends on it, you can employ a thousand people to write them all out. The British government employed 10,000 people at Bletchley Park, many of them doing exactly this. You might think that applying ALL the possibilities is too time consuming in practice but there are many shortcuts. If I suspect the message contains the name of a German town all I need do is try keys until I find a German town somewhere in the message then work my way outwards from there. Or perhaps I suspect the key is something easy to remember like the name of the Commandant's dog. I can try ALL German dog names until I get lucky. If I've 10,000 people working for me this is easy.

The Enigma machine and the coding process set up to operate it was designed to remove these loopholes. For a start, the keys were all random numbers taken from a code book no dog names allowed and the machine took the idea of a simple progressive cipher and made it *much* more complex.

Imagine I took my GeorgeGeorgeGeorge pattern but then every 3rd character added one, every 14th character subtracted 15 and every 40th character added the 3rd letter of the First Mate's mother's maiden name. Now this would be a VERY hard code to break. I would need a machine to code messages because if I tried to do it by hand I would make so many mistakes that the messages I send would be unintelligible. The Enigma machine made these coding schemes a practical possibility. But, although Enigma is hard to break it is not impossible with enough computing power. Is there any code that is impossible to break?

An Unbreakable code

Is there a way of coding a message so you can never break it?

The answer is there are two ways to code a message so it is PERFECTLY safe. The first is to use a one-time pad and the second is quantum cryptography.

One perfect way to encode a message is to use a one-time pad. On a sheet of paper I write a completely random set of numbers or letters – since we are going to translate numbers to letters it does not matter which. I make a copy and give it to a person I later want to send a coded message. Because I will only use these two paired sheets once it helps to make a few of them – a pad in fact. By convention, we refer to a single sheet or a whole book as a one-time pad code. Here is the one-time pad I created earlier. It is just a random sequence of letters and spaces.

kaleygnqaloiuebldlan dlkawoqyevbax gmlsosuebal

To code a message, I substitute numbers for letters as with the progressive cypher earlier again using modulo arithmetic to wrap around if I reach the letter 'z'. I have applied my one-time pad to the hello reader message below to get 'sfacngfvbpta'.

hello reader
sfacngfvbpta

This code is unbreakable – almost! Notice there are very few clues for anyone wanting to decode it without holding a copy of the pad. Spaces do not necessarily indicate breaks between words, and letter patterns are absent. It has only one flaw. The total number of characters and spaces could have some meaning. This is a problem because if I routinely communicated bombing targets and my message was "Bomb Bath". You could figure out the sender was not going to bomb Bristol if the message were shorter than 11 letters and spaces. To avoid this problem, messages are extended with nonsense at beginning and end to make sure no information can be gleaned from the length. The convention is to code messages to the full length of the pad. You must never reuse a pad. Each time you code a message, rip off that page rather like a calendar. Destroy it and use the next page for the next message. At the other end, the recipient uses his copy of the pad to run the process in reverse. Decode the message by swapping each letter according to the modulo method, rip the page from the pad, and burn it. Because each key is only used once you can't use any sort of statistical method to work out the message, making the one time pad perfectly secure. Claude Shannon proved this in 1945 while working for Bell Corporation but, due to wartime secrecy, his proof was not published until 1948.

The Perfect Code

The proof that a one-time pad is perfectly secret is straightforward. Imagine I take a coin and flip it 1000 times. I'll write down some of the results as follows:

HHTHHHTTHTTTHTTTTHTH…

I give you a copy of my results and keep one for myself. Now we each have the same random set of Heads and Tails recorded on a piece of paper. I can convert any message from letters to binary numbers: 'a' = 00000001, 'b' = 00000010, 'c' = 00000011 and so on. If you are not familiar with binary just assume I have a code where we only ever use combinations of 0s or 1s. To encrypt the message we flip each bit 0 goes to 1 or 1 goes to 0 – using my random list of heads and tails according to the following rule: If I have a head flip the bit, otherwise leave it the same. I now have a randomized message, and it really is truly random. To convince yourself, imagine answering the question, do you like coffee or tea? Think of your answer and flip a coin. If the coin lands heads change your answer otherwise leave it the same. Now write your answer down. Try it out a few times. Do you see you end up with a totally random set of decisions tea, coffee, coffee, tea, tea, tea. If you don't record the coin toss there is no way to determine your true answer.

Similarly, the message I encoded above now looks like a completely random stream of 1s and 0s and the only person who can decode it is the party with the other record of the coin tosses. Apply this to the message and, as if by magic, the message reappears. Any other random sequence will yield gibberish. It has to be the SAME random sequence I used in the first place.

Mathematically, the proof involves working out that the probability of getting the right answer by applying a random sequence is 1 in 2^n and the probability I could guess the answer is also 1 in 2^n? The same! So the chance of decrypting the message knowing the encryption method is the same as simply guessing the message and getting lucky. Therefore, the message is perfectly encrypted.

Quantum Cryptography

It turns out there is one other perfect encryption method that involves thinking about the nature of secrets. Normally we consider the primary problem with sending a secret message is coding it so that it can't be read by anyone but the intended recipient. However, wouldn't it be equally valuable to know if someone other than the recipient had intercepted and read the message? This is the trick quantum cryptography gives us.

Taking a measurement with a quantum device disturbs the system so measurements can be taken only once with the same results. By the same logic, I could send you a message and if someone else has read it in the meantime, you will know. I could arrange to meet with you in Berlin and if you detect the message has been intercepted, you could simply not show up.

I could use this same technique to send you a one-time pad. If you receive it without it being overheard, I could then safely send you an encrypted message. In 2007, this technique was used to transmit the results of a Swiss election from the polling booths to the central counting center.

Enigma

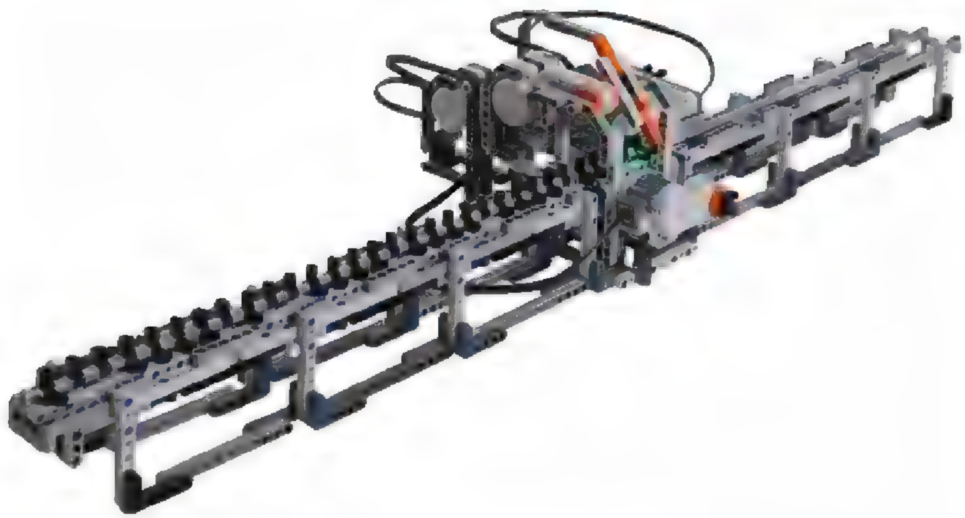
World War II accelerated the evolution of encryption from simple substitutions a human could perform to complex ciphers only a machine could calculate. You might wonder why everyone does not use a one-time pad since it is a perfect code. The problem is distributing and maintaining the pads while keeping them secret. My daughter cracked my earlier code because she knows my laptop password, broke in, and read the answer. That's the problem with codes – security. The pads could be sent out in sealed envelopes but it would be easy to intercept an envelope, copy the pad and reseal it. You would then have a perfect and undetectable way to break the code. Also, if I were an Admiral wanting to communicate with my fleet of submarines I would need a huge pad – one page for every message I want to send – and either a pad for each submarine or one pad for all submarines. If I use only one pad, then I cannot talk to a submarine privately, and if any pad were lost *all* security would be breached. One-time-pads were used by both sides during World War Two, and often printed on nitrocellulose – a chemical similar to the explosive nitroglycerine. This allowed users to burn the codebooks quickly if an enemy threatened to capture them.

Both the Americans and British captured Enigma machines and codebooks during the war. A Navy Enigma machine was a sought-after prize, as it was more complex than the Army version, with extra dials and plug settings. To crack the more sophisticated codes Bletchley Park needed to get hold of Enigma machines, ideally without the Germans' knowledge. The film *U-571* merges two such capture stories into one, taking a few dramatic liberties along the way, but it's well worth watching.

Even with a captured machine, the codes were hard to break. You needed a starting point – a crib to give you a clue what the machine settings were. Helpfully, the German Army often began their messages with a weather report. Everyone knows the German word for weather – 'Wetter'. Decode the first 20 letters of a message until you found 'Wetter' and the message is unlocked. The German Navy, however, was less chatty and avoided obvious words in their messages. One way the Allies could find a crib was to blow something up. They would sail to some point in the Atlantic, fill an old boat with oil drums, and set it alight. The German Navy would get wind of this and go to investigate. The first thing they would do is to radio a message back to base with the coordinates of the wreckage, which, of course, the British already knew. This gave the British a crib, and once they were in, they could decode messages for several days in a row because the Enigma machines often cycled through a repeating pattern.

Throughout the War, the German military never suspected the British had cracked their codes and thought they must have traitors giving away their secrets. The Enigma machine was an elegant compromise between a truly unbreakable code and a simple cipher. Unfortunately for the Germans, Turing was on the side of the Allies.

In the 1930s almost all mathematics, accounting, and code-breaking were performed by humans using pencil and paper. It was the science behind this process Turing sought to understand. We'll take a step back in time again to 1935 and Turing's discovery of a solution to the Decision Problem – *the Entscheidungsproblem*.



Lego Turing Machine

"Machines take me by surprise
with great frequency."

Alan Turing

The Machine

Turing probably learned of *the Entscheidungsproblem* in a lecture given at Cambridge University by Max Newman. Newman described a new proof by Gödel showing mathematics was incomplete. The proof solved the completeness and consistency problems by turning mathematical statements into numbers and showing you could generate a logical paradox if you tried to argue for completeness and consistency at the same time. Thus, of the three original Hilbert problems, completeness, consistency and decidability, only decidability remained unanswered.

Turing spent all of 1935 and much of 1936 thinking about this question: Is mathematics intuitive, or could a machine decide mathematical questions automatically? Eventually, cycling through the Cambridge countryside one day, he stopped to rest in a field near Grantchester and in a flash of inspiration envisioned his mathematical machine. The machine was entirely imaginary but made as if from mechanical parts common in the 1930s.

The idea was to reduce the process of computing with pen and paper to its most basic level. Turing hit upon the idea of using a long ribbon of paper tape similar to the ones used in telegraph machines. A paper tape is simpler than rectangular paper as it can be handled mathematically as a single sequence of numbers – we don't have to worry about turning the page or working in two dimensions. If you are worried that a tape is less powerful than a sheet of paper remember Cantor's theorem: an infinite plane is the same as an infinite line. The use of a tape massively simplified the mathematics, and subsequently many early computers used tapes, as they were easy to handle in practice as well as in theory.

The eye, hand and pencil of a human mathematician was modeled as the read-write head of a teletype. It allowed the machine to read input from the tape and write information back so as to keeping track of intermediate calculations or provide the final output. The operation of the machine was straightforward. At each moment in time the machine could read a symbol on the tape, move the tape forward or backwards, and write or erase a symbol. That's all he needed to model a human doing something like long multiplication. Turing argued his model was exactly analogous to a human performing a computation.

Turing's imaginary machine was now able to perform computations just like a human. You could write down the rules for a given procedure and the machine could, for example, do long multiplication. At each step of the calculation, the computer would examine the state machine, look up the state in the instruction book and put the machine into its new state. If you recall Searle's Chinese Room, this is the same process the man in the room followed: get a symbol, look it up in a book, and reply with the corresponding symbol.

Universal Turing Machine

We have missed one important step from our explanation of the modern computer: the ability to run programs. Nowadays, we take for granted you can download a program from the Internet or buy one from a shop. In the 1930s adapting a single machine to multiple purposes was a radical idea. Machines were built to do one thing, and one thing only, and there was no concept of a general-purpose machine. Nowadays this is hard to comprehend, but there is a similar revolution going on in manufacturing today with the widespread adoption of 3D printing. Today most factories use tools – lathes, drills and saws – to fashion objects. Each machine does a specific job and is not 'general purpose'. But innovative new machines can now be purchased relatively inexpensively called 3D fabricators, which print entire objects. The same happened for electronic logic in Turing's time.

Before computers, logical tasks were performed by banks of relays. How these banks work can be illustrated by the workings of an old-fashioned elevator. If you pressed a button to call an elevator, you closed a switch coupled to a relay in the basement sending power to the car. Another switch was tripped automatically when the elevator reached the desired floor. All the functioning of the elevator system was fixed. Once you pressed a button to go up you could not change your mind and press

Old Fashioned Relay Mechanism

the button to go down. That logic did not exist in the relay banks. If you wanted to improve the logic of the elevator you would need to rip out all the relays and rewire everything from scratch.

Turing's first imaginary machine was set up in the same way. It had a fixed set of hard-wired logic, a rule book. In order to perform different tasks – say addition or multiplication you had to use a different rule book. His revolutionary idea was to write a rule book that told the machine to read a soft-wired set of instructions from the tape and execute those instead. He called this a Universal machine since it could perform *any* procedure written on the tape. Today we call this software. It is fair to say Turing was not the first to use this idea. Charles Babbage's analytical engine could read instructions from cards and execute different procedures, but Turing thought through all the ramifications of the idea and made it general purpose, giving us the modern science of computing. It is easy to build a real Turing machine, but by today's standards it is a little clumsy; a team in Denmark has built one using Lego. You can see a link on my website.

Very soon after Turing's paper was published, a number of people proposed better practical implementations. In 1943, John von Neumann of Princeton University created the architecture for ENIAC, the first stored program computer, developed for the United States Army's Ballistic

3D Printing Machine

Research Laboratory. The laptop I am writing on uses the von Neumann architecture, and most modern computers evolved from it. By contrast, mobile phones are descended from the Harvard architecture developed by IBM and first supplied to Harvard University in 1944, hence its name. The distinction in architectures has blurred over the years. The world supports two main computer chip technologies, one built for desktop and laptop computers, designed by Intel in Santa Clara, California, and the other, designed for mobile devices by ARM, in Cambridge, England. All these computers can, in principle, run any piece of software.

Programs

Software is just a series of numbers. When you click an icon on your desktop, the computer reads the number and interprets it as a series of instructions. There is a decoder inside the computer that knows the number '1' means add the next two digits and the number 5493 means display them on screen and so on. On my computer the operating system, Apple's OSX, takes the number, decodes it and passes it to the CPU for execution. You might ask what runs the operating system and that is a smaller program called the BIOS. What runs BIOS? An even smaller program called the Bootstrap. Once all this is up and running you have a working computer, which can run any program you throw at it.

The problem with programs is they tend to crash – usually at the most inconvenient times. It is often not clear whether a program has truly crashed. It might be stuck in an infinite loop, or it could be calculating the answer to a complex question, such as the answer to life, the Universe, and everything. How would we know? If only I had waited a little longer before rebooting, the program would have run to its end and given me the answer to Douglas Adams' question.

It would be very useful, and save a great deal of time, if I had a way of telling whether a program will ever stop. An elegant solution would be to have a second program called 'Halt', which would test the program and output 'will halt' or 'will crash' as appropriate. It turns out this program would be more than just useful. It could be used as an oracle, capable of answering almost any question imaginable.

I could, for example, write a program that says: for every index in Fermat's puzzle try every number and halt if you find a solution greater than 2. Now if I run my halt program on this program and it states 'will crash', I will have solved Fermat's Last Theorem! Do you see why?

If we give 'Halt' an input: a program we are interested in, along with some data, it will tell us if the program finds an answer. If I am trying to solve Fermat's Last Theorem, we will ask it to try every possible index for the equation $3^x+4^x=5^x$ and halt when it finds a true result greater than 2. If the halt program says yes and halts, you can trace through the program and work out how it did it. The theory would be proved. If the program says no, the theory is disproved. This gives us a way to discover proofs of many mathematical theorems.

I could try almost any puzzle using a program with this form. All I need do is put a problem in the following decision format: *try all possible options, and then stop and ring a bell if a solution is found.* The Halt program would then give the result leading to untold riches, winning all the remaining Clay Mathematics prizes at the very least and earning me $6m.

Does such a magical program exist? The answer, sadly, is no. There is no Halt program and the final part of Turing's paper proved there can never be.

The Proof

Let's write a list of all the possible programs my laptop could ever run. A comprehensive way to do this is to start at one and try every number.As I count up I am simply generating numbers, for example, 5,433,232, then turning each number into a program file and running it. For a bit of fun, I created a couple and tried them out on my laptop. They did nothing, so it was not very edifying. Most numbers are just junk because programs have to be in the right format for the computer you are working on. It's just like words. If you randomly take a handful of scrabble tiles out of a bag, most of the time you will have nonsense, but every now and they you will have an actual word. Be careful with this; you could accidentally write, "delete every item on my hard disk." Of course, the probability is astronomically low, but Murphy's Law says it will happen, so back up your data!

As you count up, you will generate every possible program along the way. A mathematician would say programs are recursively enumerable. The word recursive means there is an algorithm and enumerable means to count. Therefore, there is a counting algorithm that would run every imaginable program. Here is a list of them, or at least a some of the highlights:

0 (probably doesn't run)
1 (ditto)
00 (ditto)
01 (ditto)
011001001001000100 (makes the computer beep once)
… (from here on I'll give the program names since the numbers are too large to print)
Does Nothing (there are many of these)
Is Gibberish (there are an infinite number of these)
Junk (an infinite number of these)
Print Something (again an infinite number of these)
More Gibberish
Excel
Word
PowerPoint
Mathematica…
Fermat's Last Theorem enumerator (runs for ever)
A nonworking version of the Halting Program
A nonworking version of the Crashing Program
Really big programs that don't fit on my hard drive

and so on.

You can see that every program imaginable is generated in our list. If you are wondering which version of Word or Excel, the answer is *every* version and every bug ridden unreleased version as well. We are enumerating every program that could ever be run in the known universe!

Perhaps you can see a problem looming. I can pose any mathematical puzzle in a clever way so that a program only stops if there is a solution. I am about to list every possible program that could ever be created. If halt exists this will automatically prove every mathematical theorem imaginable.

Let us see if this is so.

For our thought experiment, we will assume every program takes an input. Historical convention in computing means this is generally the case. If you type a program into the command line of a computer with some words listed afterwards, the computer will usually run the program with the words as input. For example, if you type, "Print 'Hello World'", most computers will print 'Hello World'.

We now imagine there is a Halt program that can run on an infinity of inputs. Will it work for every input? We are looking for a paradox caused by the existence of the Halt program. If Halt causes a paradox then Halt cannot exist.

Here goes...

If there is a Halt program, we can write a Crash program. That's a program that goes into an infinite loop if it detects a program will halt. Now what happens when we feed Crash into itself? Does Crash halt if it runs with the input Crash?

This creates a paradox; there is no solution which makes sense. It's similar to the Barber Paradox of earlier. Since a paradox is created there must be a fault in our original theory. The error is the existence of Crash. Since Crash cannot exist and it was created as the logical opposite of Halt, Halt cannot exist either. QED. There is no general program that will tell if another program will halt because such a program could not run with the negative of itself as input.

This places a limit on the power of computers to automatically solve problems. There is certainly no general purpose algorithm which will solve every problem. Slightly more subtly there is no general purpose program that is guaranteed to solve one arbitrary problem.

If there were, you could just write a program to sequentially present every problem to the arbitrary problem solver and you would have solved everything.

This presents us with a puzzle. A huge software industry has grown up based on Turing's ideas, employing tens of millions of people worldwide. This industry regularly solves all manner of problems. The proof from Turing's original 1936 paper suggests there should be quite strict limits on the power of computers. In the next chapter, we will examine this industry and take a look at Turing's theorem from a modern view point. The chapter can be read as a stand alone article but was originally written as an integral part of this book.

Chapter 11

SOFTWARE

Fred Brooks

Medieval Block Print from 'No Silver Bullet'

"The bearing of a child takes nine months, no matter how many women are assigned."

Fred Brooks

"Adding manpower to a late software project makes it later."

Brooks' Law

In *No Silver Bullet – Essence and Accidents of Software Engineering*, Fred Brooks explains why writing software is hard, and why machines are not going to do it for us anytime soon. The original article appeared in the proceedings of the Tenth World Software Conference. It was subsequently expanded into the, now famous, book, *The Mythical Man Month*.

Brooks believed solving real world problems involves understanding the essential complexity of life. 'Accidental Complexity' – the simple type – is the time-consuming part of writing software, for example, listing all 220 countries of the world in a website, or making sure all the buttons in an interface line up correctly. These tasks are tedious – you have to look up all the countries in Wikipedia and make decisions, such as whether the United Kingdom will be denoted 'UK' or 'GB'. They don't need any real ingenuity. 'Essential Complexity' is altogether different. It involves understanding the world and setting out the rules in meticulous detail. Brooks argued essential complexity is not susceptible to being sped up by machine processes. Navigating these architectural decisions cannot be automated. He gives us an analogy by comparing writing software to building a house.

When you build a house, an architect designs it, an engineer makes the calculations to ensure it is safe, and a construction firm builds it. The construction process dominates the cost and time. In software projects, an engineer writes a program that precisely defines the design and the construction and calculation is done by a compiler – software that takes the design and makes it machine-readable. Compilers operate in a fraction of a second. Making software is, therefore, dominated by the design time, and design is all about capturing the essential complexity of a task.

This chapter will try to show where essential complexity comes from, why computers can't tackle this sort of complexity and, therefore, why they can't write software. Good news for programmers as this means job security!

For a more thorough treatment of the mathematics read my paper *The Free Will Universe* at www.jamestagg.com/freewillpaper.

```
<html lang="en">
<head>
    <meta charset="UTF-8" />
    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
    <meta name="viewport" content="width=device-width, initial-scale=1.0">

    <title>James Tagg | Invention, Physics and Farming</title>
    <link rel="profile" href="http://gmpg.org/xfn/11" />
    <link rel="pingback" href="http://jamestagg.com/xmlrpc.php" />

    <!--[if lt IE 9]>
    <script src="http://s0.wp.com/wp-content/themes/premium/broadsheet/js/html5.js?m=1393348654g"
type="text/javascript"></script>
    <![endif]-->

            <script src='//r-login.wordpress.com/remote-login.php?
action=js&amp;host=jamestagg.com&amp;id=57804437&amp;t=1406042069&amp;back=http%3A%2F%2Fjamestagg.com%2F'
type="text/javascript"></script>
        <script type="text/javascript">
        /* <![CDATA[ */
            if ( 'function' === typeof WPRemoteLogin ) {
                document.cookie = "wordpress_test_cookie=test; path=/";
                if ( document.cookie.match( /(;|^)\s*wordpress_test_cookie\=/ ) ) {
                    WPRemoteLogin();
                }
            }
        /* ]]> */
        </script>
        <link rel="alternate" type="application/rss+xml" title="James Tagg &raquo; Feed"
href="http://jamestagg.com/feed/" />
<link rel="alternate" type="application/rss+xml" title="James Tagg &raquo; Comments Feed"
href="http://jamestagg.com/comments/feed/" />
<script type="text/javascript">
/* <![CDATA[ */
function addLoadEvent(func){var oldonload=window.onload;if(typeof window.onload!='function')
{window.onload=func;}else{window.onload=function(){oldonload();func();}}}
/* ]]> */
```

James Tagg's Home Page

"Computers are stupid. They can only give you answers."

Pablo Picasso

"Software is like sex: it's better when it's free."

Linus Torvalds

Silver Bullets Can't be Fired

Human brains are wonderfully creative things. We can compose music, play golf, write novels, and turn our hands to all manner of problems. Many people use their brains to write software. In our modern day lives we use software all the time: when we access the web, type on a word processor or play a computer game. Software also inhabits many apparently dumb devices. Modern cars contain dozens of computers quietly working away; providing entertainment and navigation, controlling the engine, and helping the car brake safely. In my living room I count over a hundred computers. Many are tiny, like the one in my TV remote control, while others are hidden as parts of larger machines. The laptop on which I write has over twenty computers inside it, besides the main Intel processor.

One thing all these computers have in common is that a human being sat for many hours writing their software. Software is formal logic written in something resembling English.

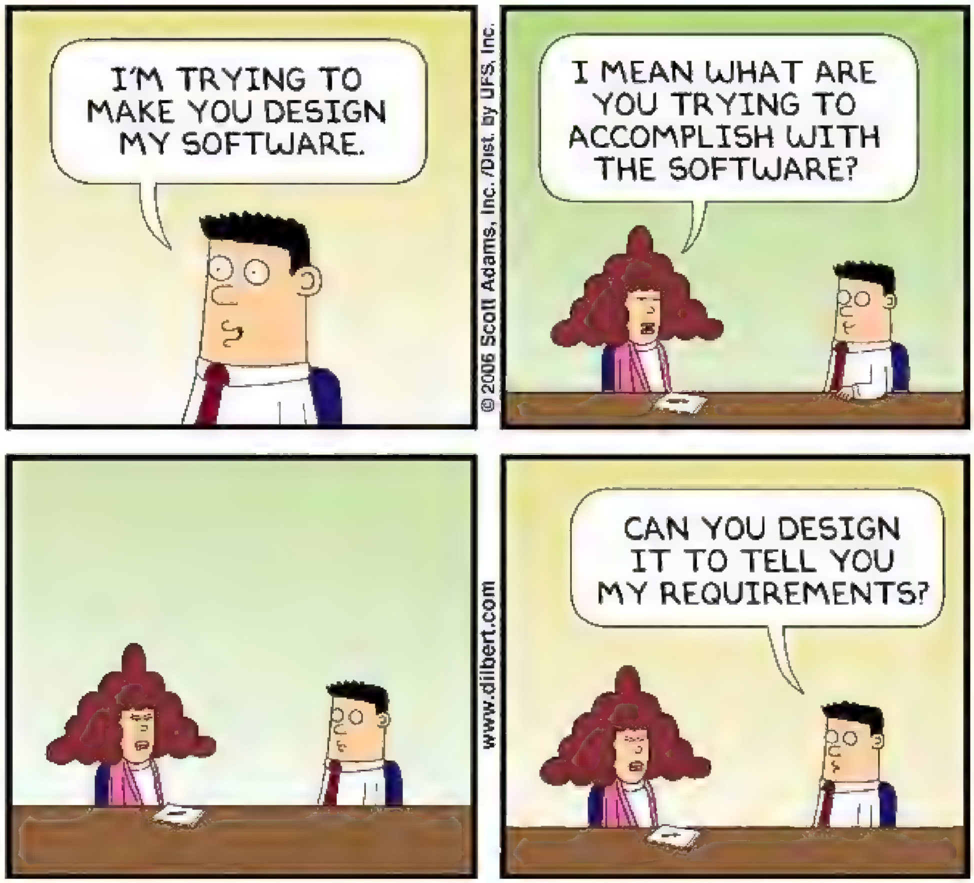
If I go to my ATM and try to withdraw cash, a programmer will have written out the logic for the transaction as a set of rules.

When I put my bankcard in the slot, and type in my PIN, a line of software will ask: If the bank balance of 'James Tagg' is less than twenty dollars and I have pressed 'withdraw' for an amount in excess of twenty dollars, then display, "We are sorry we cannot process the transaction at this time." and return the card. There seems to be an unwritten rule that the things a computer says should be accurate but unhelpful!

Alice, Ted and Software Specification

It would have been much more helpful if the computer had said, "You do not have enough balance in your account." And, it would have been more helpful still if it had asked whether I needed a temporary overdraft. However, such a feature needs many more lines of software and this is time-consuming to write.

Software takes time and is expensive, because it has to be written in a general-purpose way. Any name could substitute for James Tagg, and any amount could be used. After all, it would be useless if an ATM machine could only give out $20 to one person. The generalization of software makes use of variables instead of fixed values and this renders it hard to understand. Wherever we meet an idea that needs to be generalized, a letter must be used instead of a fixed value. Computer programs tend to look like this: if 'a' wants to do 'b' with 'c' then allow it only if 'd' is greater than 'c'. The software programmer has to keep track of all the possible values that could be inserted into each of the variables and make sure each and every combination would make sense. My ATM scenario gets complex quickly. It needs to be able to answer a range of questions for all the bank's customers, deal with any amount of



money and handle security when communicating with foreign banks is necessary. A human being must write lines of code for all the rules and every exception, making provision for any gibberish that might be typed in by the customer.

Many people ask, "Wouldn't it be great if my computer could write software for me? Humans could sit back and put their feet up." While most people don't actually believe this could happen, they will often ask why we can't specify software exactly and use unskilled people to write it. Both proposals fundamentally misunderstand the nature of writing software.

What do Programmers Do?

A human software programmer can write up to 1000 lines of code per day. At the beginning of a project, when the work is unconstrained, programmers write fast. Things slow down once programmers encounter the enemy: the real world. By the time the code is complete and selling in shops, the productivity of a programmer can be as low as one line of code per day. This is staggeringly low and luckily only applies to big commercial software, equating to about 10 words per day. A good typist types at 80 words per minute and most programmers are reasonable typists. So software writers in a big project spend only a minute or so per day in the act of writing. The rest is taken up by meetings, process discussions, email, reporting and so on. In projects that avoid much of this administrative overhead, good software programmers reach a long-run average of about 225 lines per day. This has been the level of productivity on the products I have developed in the past. These projects were lucky. They had a single team on the task from beginning to end and, in general, the projects took few wrong turns. Still these programmers were spending only 10-20 minutes of each day on actual programming. What were they doing the rest of the time?

In the early days of programming you might have a great idea, but the process of turning this idea into software was immensely long-winded. I learned to program at Manchester University in the 1980s. The enormous machines in the basement of the computer building provided heat for both our building and the mathematics tower next door. We were not permitted to play with these basement monsters but were 'privileged' to submit instructions to a mini computer in the undergraduate section – a PDP11-34.

For those of you not acquainted with computers I can tell you the process of writing software in the 1980s was immensely tedious. To add two numbers and display them on a screen took a month of lab time, using detailed instructions written in machine code. Everything was manual, including writing your code out in pencil on special paper with little numbered squares and then giving it to someone to type in overnight! You would return the next day to discover whether you had a usable program or a something riddled with errors. If you found an error, it would require editing. This was nothing like using a modern word processor. The online editors of the day were the ultimate in annoying software. If you misspelled a word, you would need to count up the letters and spaces manually on a printout and enter a command – replace letter 27 of line 40 with the character 'r'. Each and every typo would take five minutes to correct. I managed to finish the simple program required for course credit – I think it displayed an eight-digit decimal number – and ran for the hills. In my second year I bought a PC and decamped to the physics department next door where I remained for the rest of my undergraduate life.

The PC revolution provided programmers with a new and intuitive software creation environment where almost all the tedium was removed. A wealth of tools for creating software was pioneered by Bill Gates of

Microsoft and Philip Kahn of Borland, along with intuitive applications such as the spreadsheet invented by Dan Bricklin and Bob Frankston and made popular by Lotus Corporation. Today all computers have elegant WYSIWYG, 'What You See Is What You Get' interfaces, where you drag and drop elements into place on the screen. Over the last 25 years writing software has sped up and stopped being tedious – becoming almost a joy!

In *No Silver Bullet*, Brooks explains that writing software can't be accelerated any further because all the tedious mechanical tasks have already been removed. Remember his analogy: Writing software is like building a house, but with some important differences. With a house, an architect handles the design and then turns over construction to a building company. Construction takes an appreciable time, more time than the design and quite a bit more effort. But in software the construction is totally automated. When we complete the design for a piece of software we press compile on the computer and the software is built and tested automatically in a matter of seconds. Speeding this process up any further would make only a tiny improvement in the overall software creation time, since the process is already 99% design and 1% building. For the most part, the creative process of writing software cannot be improved through mechanical means.

This is not always the case. I recently upgraded the machines for some developers I work with. We added solid state hard drives. Compiling a program now takes only 10 seconds, compared with 6 minutes before. Because programmers nowadays tend to compile their programs very regularly we estimate this saves them as much as an hour a day. This is the only real innovation I have seen in the build phase of software in the last 5 years, and it's arguably not an innovation at all. We just forgot to keep on top of the build time and allowed it to get out of hand.

You might argue some counter examples. Modern software design suites let you drag and drop things on the screen to make applications or build a website. Two hundred million people have managed to put together WordPress websites using this technique. These are mechanical procedures for solving a programming task and seem to contradict my argument. They allow us to lay out graphics, press a button and turn the design into software. But they perform very simple tasks. The computer simply notes the coordinates of each box on the screen and places those numbers into a file. The process is entirely mechanical and could be performed by a clerk with no programming knowledge following a set of rules. The computer just does it faster. I did the clever work; I had the idea for the software, I came up with the idea for the interface, I decided where to place the boxes, and I chose all the colors, fonts and graphics. I did all the creative bits!

So, now we know what programmers do all day. They create!

Origins of Software

Alan Turing first described the modern day computer in a paper presented to the London Mathematical Society in 1936. He was not trying to invent the computer. That was a by-product. He was trying to solve a puzzle that had been troubling mathematicians for 30 years: *The Decision Problem*.

David Hilbert set out the challenge during a public lecture to the French Academy of Science in 1901, marking the turn of the century. Rather than give a boring lecture extolling the virtues of scientists, he decided to give his audience a list of all the puzzles mathematicians were stumped on.

Rather like the XPRIZE of today, he presented the problems as a series of challenges. Sadly for the mathematicians of his time, there were no million dollar prizes on offer, just a moment of fame and the adulation of their colleagues. Each challenge was given a number. The list included many famous puzzles; the Riemann Hypothesis, the puzzle of Diophantine Equations and the Navier Stokes Hypothesis, to name only three. A group of these questions were to coalesce into what we now know as the Decision Problem.

The Decision Problem is very important to computer science because it asks whether an algorithm can be written to automatically discover other algorithms. Since all software is itself algorithmic you could rephrase the question: Can software write software? This might seem esoteric. But, if you are a computer scientist, it is an important question. If we could solve all mathematical problems automatically we would not need mathematicians anymore. And, since programs are applied mathematics, the same goes for computer programmers.

Before you breathe a sigh of relief because you are neither a mathematician nor a computer scientist, you should remember it is possible to describe all knowledge using numbers. That's what your iPhone does when it stores music. If everything can be represented by numbers, then a fast enough computer could use an algorithm to create *everything*! You really could set Douglas Adams' *Ultimate Question of Life the Universe and Everything* before a computer and it would come up with the answer – presumably extrapolating the existence of rice pudding and income tax along the way.

Algorithms

Back in the 1930s no mechanical system could perform a calculation with any speed. People still used pencil and paper for most things; the newly-invented mechanical cash registers were slow and could perform only one calculation for each crank of the handle. If you wanted to calculate something complex, you had to employ a *computer*: a person who could do mental arithmetic enormously fast. Richard Feynman's first job was computing for the Manhattan Project. The question was: Could a computer, either mechanical or human, blindly follow known rules to decide all mathematical questions? Hilbert's 10th Problem asked this question of a particular type of mathematical expression – called a Diophantine equation.

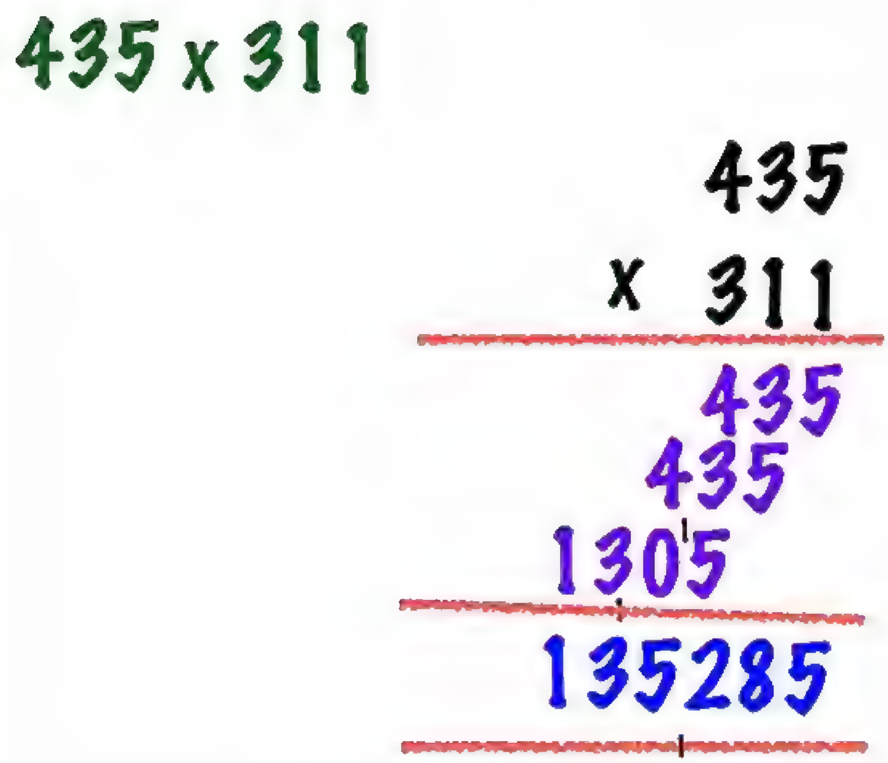
Hilbert's 10th Problem

> *"Given a Diophantine equation with any number of unknown quantities, devise a finite process to determine whether the equation is solvable in rational integers."*
>
> David Hilbert

Diophantus lived in ancient Persia – now Iran. His son died young and Diophantus was so consumed by grief he retreated into mathematics. He left us seven books of mathematical puzzles – some he devised himself and some of them taken from antiquity. The puzzles look deceptively simple and are all based on equations using whole numbers. His most famous puzzle is set in a poem which tells how old Diophantus was when he died. Can you solve it?

> *"Here lies Diophantus,' the wonder behold. Through art algebraic, the stone tells how old: 'God gave him his boyhood one-sixth of his life, One twelfth more as youth while whiskers grew rife; And then yet one-seventh ere marriage begun; In five years there came a bouncing new son. Alas, the dear child of master and sage, after attaining half the measure of his father's age, life chill fate took him. After consoling his fate by the science of numbers for four years, he ended his life."*

Diophantine puzzles look straightforward. Hilbert asked if these problems could be solved by a mechanical procedure, in modern terms, by an algorithm. To show you what is meant by this, allow me to take you



Long Multiplication

back to your childhood. Do you recall being taught long multiplication at school? Take a look at the next illustration and it will all come flooding back. Once you learn the process of long multiplication you can follow the rules and get the right answer for any similar problem every time. To do this, you lay out the calculation in a particular format and apply the logic. Multiply each number by a single digit of the other number and then add the results together.

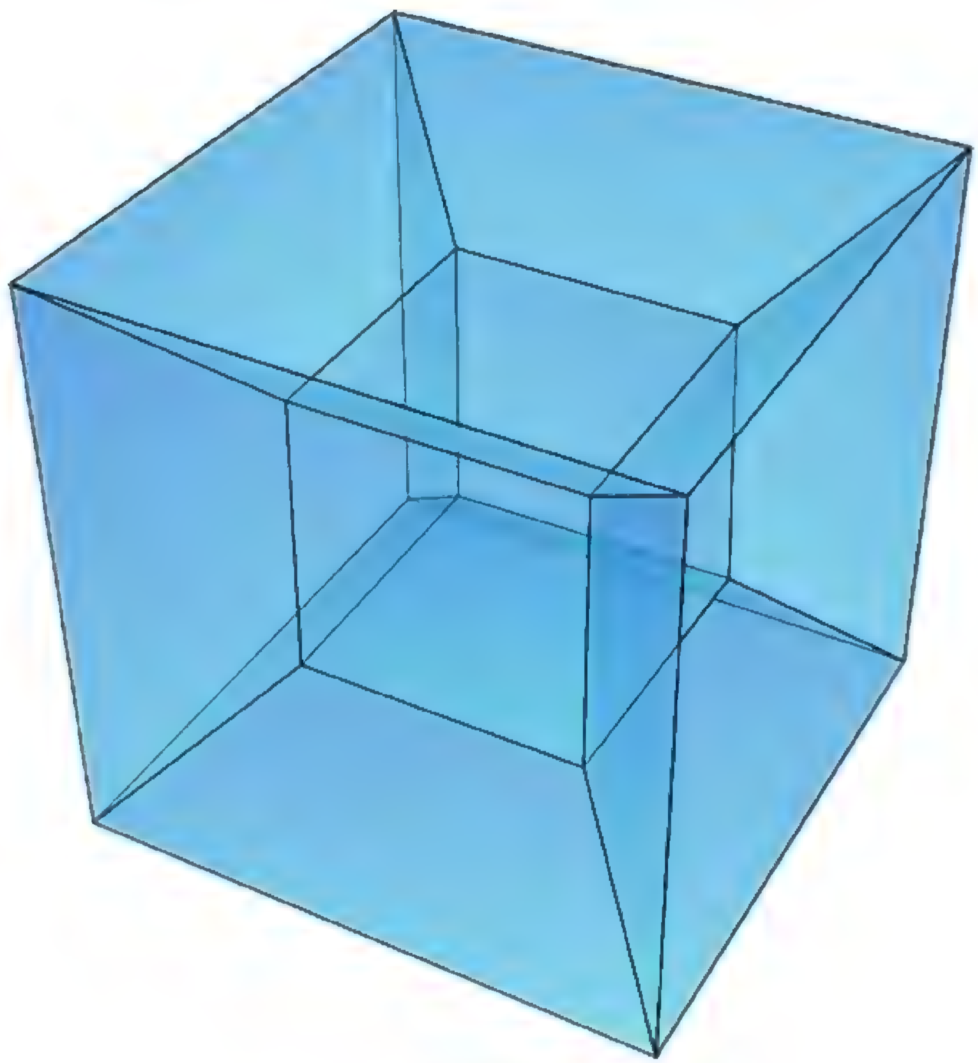
Diophantine problems are a little more complex than long multiplication and some of them are a bit abstruse. But there is one *very* famous Diophantine problem we can all recite. "The square on the hypotenuse is equal to the sum of the squares of the other two sides." The equation for a Pythagorean triangle.

The theorem applies to right-angled triangles and there are sixteen whole number solutions, known as Pythagorean triples; three, four, five; is one example.

Purists may protest that Fermat's Last Theorem isn't strictly Diophantine because it refers to a variable exponent – the x to the n part. This is hair splitting. But, of course, the splitting of hairs is bread and butter to a mathematician. We will see later that Fermat's Theorem can be made Diophantine, but we are jumping ahead of ourselves a little.

A question that taxed mathematicians for many centuries was whether there are triples for higher powers, such as cubes. In other words, would the cube of the hypotenuse be equal to the sum of the cubes of the other two sides for some set of numbers? After much work, it was proven no triple exists which can solve the cubic equation. But what happens if we substitute higher indices?

The next shape to consider is the hypercube – a four-dimensional cube. That may stretch your visual imagination but the equation is simple, $3^4+4^4 \neq 5^4$. Again the challenge is to find a whole number solution for:



Hypercube

"The hypercube of the hypotenuse is equal to the sum of the hypercubes of the other two sides." A picture of the hypercube might help you visualize things.

It's quite difficult to get your head around this shape because it is hard to think in four dimensions. This seems strange because we have no problem seeing in three dimensions on flat, two-dimensional paper – it's called a picture, but four dimensions on flat paper appears to stump us. Again there is no solution for a hypercube: no Pythagorean triple exists.

Fermat's Last Theorem asked whether this inequality for the cube and the hypercube is true for *all* higher dimensions – for the hyper-hypercube, the hyper-hyper-hypercube and so on. Tantalizingly, he claimed to have found a proof but wrote that it was too large to fit in the margin of his book. It's partly due to this arrogant annotation that it became the most famous puzzle in mathematics, frustrating mathematicians for nearly 400 years.

Hilbert's question back at the turn of the 20th century was whether a machine could find a proof of this conjecture by following a mechanical procedure, similar to our long multiplication example above.

The puzzle was eventually solved in 1995 by Andrew Wiles, a mere 358 years after Fermat claimed to have solved it. Wiles' proof runs to eighty pages of densely typed mathematical notation   considerably larger than the margin in which Fermat claimed his proof did not quite fit! There is an excellent book by Simon Singh – *Fermat's Last Theorem* – that tells the whole story.

We now know for certain, thanks to Wiles, that the answer is 'no'. There are sixteen answers to the two-dimensional triangle puzzle but there is none for any higher dimension all the way up to infinity. How might a computer tackle this problem and find a proof?

A computer could apply brute force and try many solutions; every combination up to 100 million has already been tried and no exception found. But, mathematicians are haunted by big mistakes of the past. There were theories they imagined to be true until someone discovered a counterexample. This sort of thing dogged prime number theorems.

Mathematicians don't like to look foolish and are suspicious of practical answers, "Well, I've tried it and I can't seem to find an exception." This sort of argument does not wash with them. That's what engineers and physicists do. Mathematicians are better than that!

Mathematicians want definitive answers; "It is certain no solution can exist", and these sorts of answers require an understanding of the problem to see *why* no solution could exist. That's a very high bar. What we need is a program that, rather than mechanically trying every possible combination, takes our problem and definitively says, "Yes, there is a solution," or, "No, there is not." There are plenty of man-made proofs of this nature. Pythagoras's proof there are an infinite number of primes is an example. Pythagoras did not have to try every prime number. He simply understood the nature of prime numbers and gave us a logical reason why it is so.

Mathematicians love a general solution. One way to solve Hilbert's 10^{th} Problem would be to find a single mechanical way to solve *every* problem. If you could solve *every* possible problem, you could certainly solve Hilbert's 10^{th} Problem. It turns out there is a way to test whether every problem has a mechanical solution – pose the Halting Question.

The Halting Question

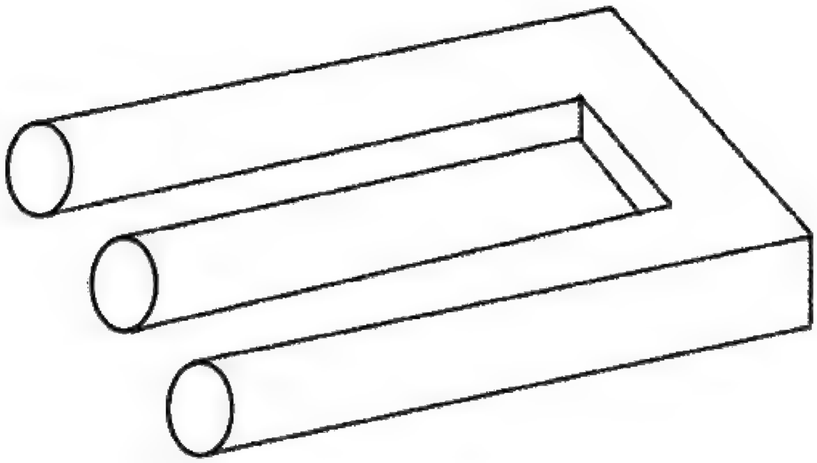
I should say for a little historical color that the Halting Problem was not called that by Turing. The name was coined much later, in the sixties, by Martin Davis. Turing knew the problem by the less catchy name of the "not crashing" problem, or as he preferred, "Being circle free", meaning the program did not get caught in an infinite loop.

To understand halting we should imagine a brute force program stepping through all the possible solutions to Fermat's problem. If there *is* a solution this stepping program will eventually halt and answer 'true'. If there is not, the program will run forever. Can we predict a program will not run forever? At first pass this is hard. We can't watch it forever and say, "It never halted." So is there a clever way to do this? An algorithm perhaps?

The Answer to the Ultimate Question

The answer is 'No!' In 1936, Alan Turing proved there is no general-purpose mechanical way to tell whether a program is going to find an answer at all, much less what the answer is. This means Hilbert's Decision Problem has no solution; there is no general purpose algorithm which will discover all mathematical theorems.

Turing succeeded in proving this by turning the problem on its head. He proved that a crash detection program is unable to see whether it will crash itself. Since you cannot tell whether a program will crash – and by this I mean go into an infinite loop – you cannot tell if it will halt. He used the simple argument that since you can't tell if the crashing program will halt, you have already proved you can't predict if every program will halt.



Impossible Shape

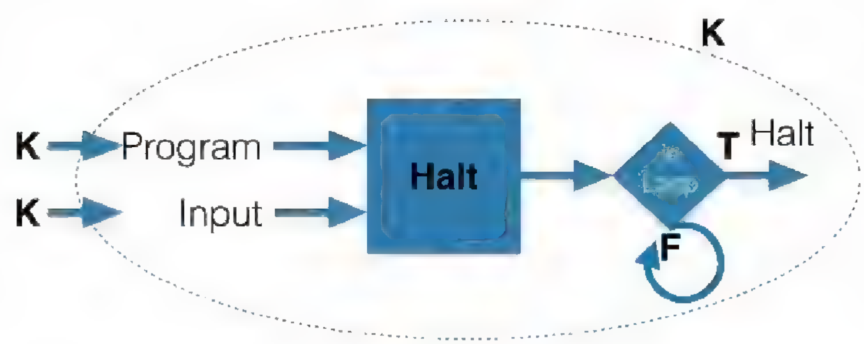
That is Turing's argument in a nutshell. But if that was too large a step, let's take the argument a little more slowly and prove it a couple of different ways. First, we will use a proof by counterexample, known by mathematicians as an 'indirect proof'. These may tax your brain. If you want a visual image to help with the idea of an indirect proof, take a look at the impossible shape. It is paradoxical, which means it does not exist. QED.

The Proofs

There are several ways to prove the non-existence of the Halting Program. I am going to present a few in the hope one of them will hit the mark and allow you to see why. The first proof uses a software flowchart. I have laid this out on the assumption the program exists and then attempted to apply it to itself. Unfortunately, the flowchart contains a paradox and thus there can be no Halting Program. The paradox is at once straightforward and confusing. It is a more elaborate version of the liar's paradox: "This sentence is a lie." If the sentence is true it must be false, and if the sentence is false then it must be true.

The Halting Program

Let us suppose there is a Halting Program. Remember that a Halting Program simply takes another program as input and predicts if it will halt or not. It follows there must also be a program called Haltcrash. Haltcrash goes into an infinite loop if it examines a program with input that halts, otherwise it halts itself.



Halting Flowchart

Now we create a third program called RunMe. RunMe runs Haltcrash on itself. Still following this? Now execute RunMe with RunMe as its own input. What happens? The analysis is as follows:

1. RUNME started on input RUNME halts. If RUNME started on RUNME halts, then Haltcrash started on RUNME with input RUNME halts. If Haltcrash started on RUNME with input RUNME halts, then HALT decided that RUNME started on RUNME does not halt!

   Therefore,

   RUNME started on input RUNME halts implies that RUNME started on input RUNME does not halt. (contradiction)

2. RUNME started on input RUNME does not halt. If RUNME started on RUNME does not halt, then Haltcrash started on RUNME with input RUNME does not halt. If Haltcrash started on RUNME with input RUNME does not halt, then Halt decided that RUNME started on RUNME halts!

   Therefore,

   RUNME started on input RUNME does not halt implies that RUNME started on input RUNME halts. (contradiction)

Both analyses lead to a paradox! There is only one way out. There can be no halting procedure. I'm sorry if this is quite convoluted.

Philosophical Proof

If you find these technical proofs difficult to follow, it may be easier to examine the problem philosophically. Consider the consequence of the existence of a Halting procedure. A Universal Turing Machine is a relatively small program. Roger Penrose gives a three-page example in *The Emperor's New Mind*, and Stephen Wolfram has implemented one using a cellular automaton with as few as five component parts.

A Halting Program running on such a machine should be able to compute all the knowledge in the Universe. Every structure, every work of literature, every galaxy could be the output of this single, simple program. My pocket calculator could, theoretically, paint like Picasso and compose like Mozart. All art, knowledge and science would be entirely determined in our Universe and we would have no free will. If you philosophically rebel against this then the Halting Problem must have no solution.

Gödel's Insight

Another way to understand this conundrum is through the earlier work of Gödel. Solutions to mathematical puzzles are neat, orderly sequences of statements where the problem is solved step by step. Computers are good at step by step processes. Surely a computer could simply proceed in a painstaking fashion to check all the possible combinations of words and symbols to discover a proof.

An analogy might be trying to find your hotel room if you have forgotten the number. You could simply find it by trying every room. As you progressed through each floor, you would try every corridor and retrace your steps to the main hallway before attempting the next. Eventually you would succeed.

Finding proofs of theorems is often understood to be the same sort of task: search systematically through all the numbers and you will find the solution. But this is not so: There is a hidden problem.

Although it is true to say problems and proofs can be described by numbers, they are not simply related like a lock and key. We need the first number to translate into a set of symbols meaning something about mathematics: for example, that x squared plus y squared equals z squared but for higher powers there is no equality, and the second number to denotes a set of sequential steps we can apply to demonstrate this fact. These steps must have meaning and obey the rules of mathematics, but what are these rules? Are they written down in a text book?

It turns out there is no way to find this set of rules; it is a super-infinite task. We would need to reach into our infinite bag of numbers and pull out rule after rule, turning each into a mathematical model that explains numbers and logic and what can be done with them to form mathematical statements. The number of ways to do this is not just infinity, but two to the power of infinity. This is the number of ways to permute all possible mathematical rules.

Your mind may be rebelling at this. Surely, if I have an infinite set of numbers I can just pluck all the numbers from my bag and then I am certain to have the solution. Unfortunately, it turns out there is no complete, consistent set of rules; no valid dictionary that maps all numbers to all of mathematics. That is Gödel incompleteness theorem.

Despite a fundamental limit on mapping all numbers to all of mathematics, there might still have been an algorithm which could practically find solutions for a given arbitrary problem. Turing proved this is not the case.

The Wiles Paradox

Turing showed us there can be no general purpose, mechanical procedure capable of finding solutions to arbitrary problems. A computer program cannot discover mathematical theorems nor write programs to do so. Yet computers regularly solve problems and generate programs. That's what software compilers do. This seems to be contradiction.

The solution to this apparent contradiction is to propose a boundary: a 'logic limit' above which computers may not solve problems. With a high boundary a general-purpose machine could solve most problems in the real world, though some esoteric mathematical puzzles would be beyond it. But if the boundary were low, many activities in our daily life would need some sort of alternative, creative thinking. It is crucial to know where the logic limit lies.

The Logic Limit

Amazingly, in many branches of science it is possible to pinpoint the exact location of the logic limit, but finding that boundary in mathematics has taken forty years work from some of the greatest mathematicians of the 20^{th} century.

The story starts back in the 1940s at Berkeley University with a young Julia Robinson, one of the first women to succeed in the previously male-dominated profession of mathematics. By all accounts, she had a wry sense of humor. When asked by her personnel department for a job description she replied: "Monday—tried to prove theorem, Tuesday—tried to prove theorem, Wednesday—tried to prove theorem, Thursday—tried to prove theorem, Friday—theorem false." Like Andrew Wiles, she fell in love with one of the great mathematical puzzles, and although she made great strides, the problem passed from her to Martin Davis for the next steps.

The final elements were put in place in the 1970s with the work of another young mathematician, this time a Russian – Yuri Matiyasevich. Robinson wrote to him when she heard of his proof, "To think all I had to do was to wait for you to be born and grow up so I could fill in the missing piece." The complete result is the Robinson Davis Matiyasevich theory which sets out the limits of logic and algebra. What, you may ask, do we mean by logic and algebra?

Mathematicians like to turn everything into logical statements, even ordering a round of drinks! The discipline of logic emerged from ancient Greece as the study of language. The starting point was the syllogism: Statements such as, "All cows eat grass." or Lewis Carroll's assertion, "There are no teachable gorillas." Over time the study of logic became ever more precise with, for example, the introduction of variables and equations; a=all cows, b=some grass. The formula "a eats b" translates by substitution into, "The cows eat the grass." This doesn't look much like a step forward but, trust me, it is.

The modern way to represent logic is using prenex normal form. This mouthful simply means separating relationships between things from the things themselves. The following four statements say the same thing, each in a more formalized way.

Speech: Harry loves Sally

Logical: x loves y (substitute Harry for x and Sally for y)

Formal: There exists an x, there exists a y (x loves y)

Prenex: $\exists x \exists y$ (x R y), Where R, the relationship, is 'loves'

The final example is in prenex normal form. The symbol '∃' means 'there exists' and R stands for relationship in this equation. All logical statements can be translated into this form using a purely mechanical process. There is even a website that will do this for you. It's useful but I don't recommend it as entertainment!

In the example above, something exists in relation to the existence of something else: one person who loves another. Give me a name and I can look up the person they love. This is simple. A computer can easily solve such problems. Indeed there are hundreds of websites doing this every day. Once you've solved one problem of this type, you have solved them all.

We can rearrange Diophantine equations into many different prenex forms. The simplest form might be, 'there exists an x which solves the following equation, x equals three.' This would be written out as $\exists x, x=3$ and is of the $\exists$ class – 'there exists'. There are slightly more complex classes than our simple $\exists$ relationship: $\forall\exists\forall$ 'for all, there exists for all' or the class $\forall^2\exists\forall$ 'for all, for all, there exists, for all'. Each of these groups of equation is called a 'reduction class'.

One way to think about a reduction class is as a problem in topology, 'knots', to non-mathematicians. Imagine someone handed you a bunch of tangled cables the sort of mess you get when they are thrown haphazardly into a drawer. You can tease them apart and rearrange them but you must not cut them or break any connection. Once you have done this you will be left with a series of cables on the desk. They are all separate, looped or in someway knotted together. Each cable has a fundamental topological arrangement: straight cables, granny knots, figure eight, and so on. You have reduced them to their simplest form, their logical classes. The same goes for logical statements. Once you have rearranged logical statements into their simplest form you can lay them out and group them together according to their complexity. Each group makes up a reduction class and you can ask whether that class as a whole is automatically decidable. It is a huge task to untangle and classify mathematical problems, and it took Robinson and her colleagues nearly forty years to succeed.

It turns out problems with a form as simple as $\forall\exists\forall$ (for all, there exists, for all) have no general purpose algorithm. Each must be examined individually and solved by something that is not a computer. This is a remarkable result as the logic boundary is set quite low. An $\exists\exists$, (exists, exists), class of problem is automatically solvable by a general algorithm, but a ∀∃∀, (for all, there exists, for all), is not. Each individual type of problem within the class must be examined with insight and understanding.

Our lives are full of problems – playing chess, finding a mate, designing space ships and simply getting to work in the morning. Imagine we expressed everyday problems as logical problems. Where is the logic limit for life? We have no answer for this yet, but we do know the logic limit for computing; it is given by Rice's Theorem.

Named after Henry Rice, and proven in 1951 as part of his doctoral thesis at Syracuse University, Rice's Theorem states: "No nontrivial feature of a computer program can be automatically derived." You cannot tell if a program will halt with a given input. You cannot tell if one program will generate the same output as another. You cannot tell if a simpler program could be written to do the same task as a more complex one. In fact, no nontrivial thing can be proven. This means the logic limit in computers is low, and computer programmers have job security.

For Programmers

For the programmers amongst you, here are some of the things that cannot be done automatically even given infinite time:

- Self-halting Problem. Given a program that takes one input, does it terminate when given itself as input?

- Totality Problem. Given a program that takes one input, does it halt on all inputs?

- Program Equivalence Problem. Given two programs that take one input each, do they produce the same result on every input?

- Dead Code Elimination. Will a particular piece of code ever be executed?

- Variable Initialization. Is a variable initialized before it is first referenced?

- Memory Management. Will a variable ever be referenced again?

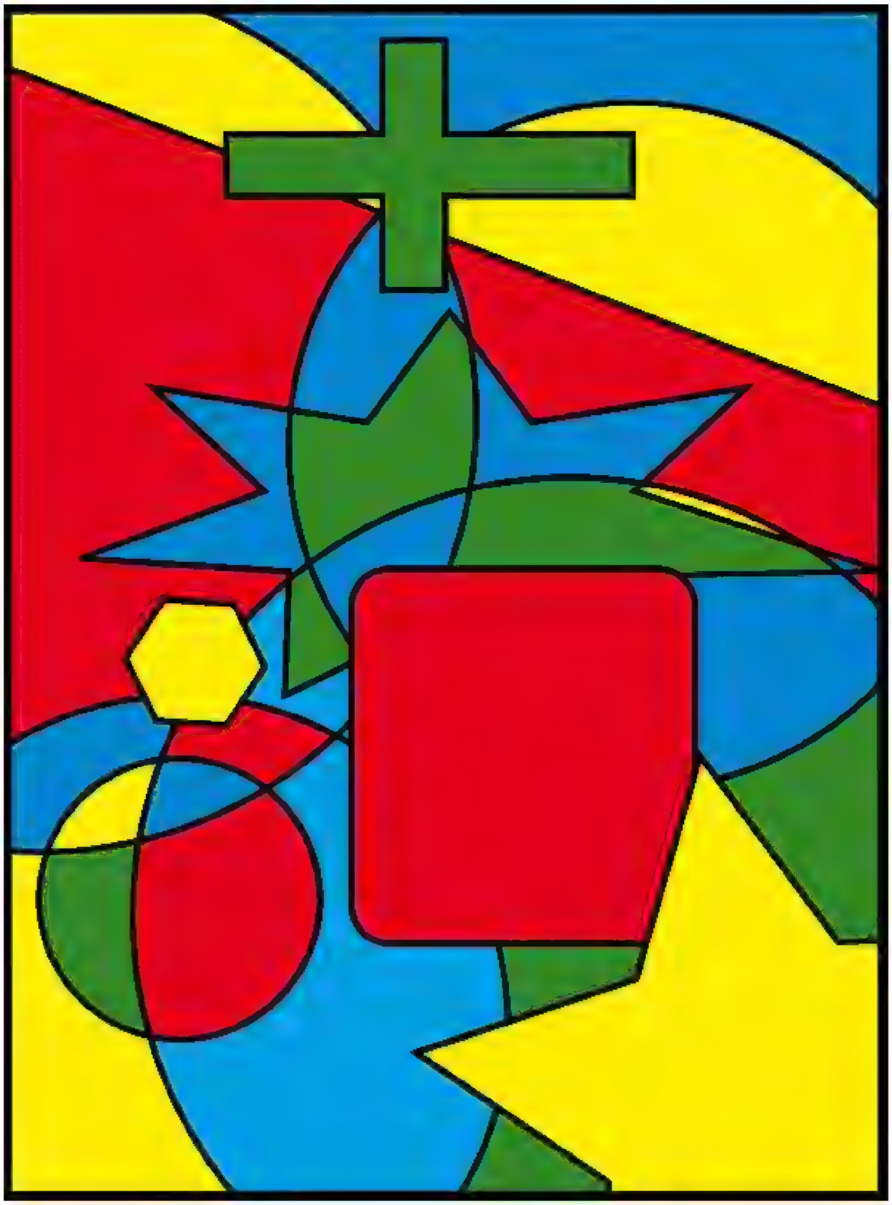
Can humans solve 'unsolvable' problems?

The question of whether Fermat's Last Theorem could be solved mechanically remained unanswered until 1970 when Yuri Matiyasevich filled in the missing piece in Julia Robinson's proof. Matiyasevich used an ingenious reduction method to match up sequences in Robinson's theorem with a set of Turing machines. This showed that if Robinson's theorem was false you could solve the halting problem and since you can't solve the halting problem, then Robinson's theorem must be true. All this effort proved Diophantine equations have no general algorithmic solution. This was a hugely important result but, as we noted earlier, Fermat's Last Theorem is not, strictly speaking, a Diophantine. It is an exponential Diophantine equation. We still had no definitive answer to Fermat.

In 1972 Keijo Ruohonen and again in 1993, Christoph Baxa demonstrated that Diophantine equations with exponential terms could be rewritten as regular Diophantine equations with one additional complication – the necessity of adding an infinite set of terms to the end of the equation. In 1993, J.P. Jones of the University of Calgary showed the logic limit for regular Diophantine equations lies at thirteen unknowns. Matiyasevich had already pointed this out but never completed his proof. Since infinity is greater than thirteen, all exponential Diophantine equations are above the logic limit and, therefore, undecidable. Finally, we have a proof that Fermat's Last Theorem is unsolvable by a computer – or at least by a general purpose algorithm running on a computer. Matiyasevich went on to show many mathematical problems can be rewritten as exponential Diophantine equations and that much of mathematics is undecidable. For example, the Four Color Conjecture:

> *"Given an arbitrary map on a Euclidean plane, show the map can be colored in a maximum of four colors such that no adjacent area shares the same color."*

Meanwhile, Andrew Wiles, an English mathematics Professor at Princeton had been secretly working on Fermat's Last Theorem. When I say secretly, he had not told anyone in his department, and only told his wife late in 1993 when he suspected he might have a solution. He had been working on the problem a long time, having fallen in love with it at the age of 8! In 1995, after nearly 30 years work, he announced he



Four Colors is All You Need

had found a proof. He had solved an unsolvable problem, a problem that could not be answered by using a computer. Therefore, Andrew Wiles cannot be a computer!

As with all real-life stories, it was not quite as neat as this. It turned out Wiles' initial proof had an error in it, identified by one of his referees. Wiles had made an assumption about a particular number theory that had not been proven: it was still a conjecture. Working with another mathematician, he managed to prove this conjecture and so, two years after first announcing that he had solved Fermat's Last Theorem he could finally lay it to rest.

The Special Purpose Objection

Before I declare mankind's outright victory over computers, the Special Purpose Objection must be overcome. The objectors would argue that Wiles is a Special Purpose computer. Special Purpose computers are at no risk of breaking the Turing limit when they solve problems they have

> **Theorem (Undecidability of Hilbert's tenth problem)**
> There is no algorithm which, for a given arbitrary Diophantine equation, would tell whether the equation has a solution or not.

been programmed to answer. The objection misses the key point. I am not arguing *having* a solution to a given mathematical puzzle presents a difficulty to a computer; I am arguing a computer cannot *discover* one.

Take, for example, the search engine Google. If I type "where can I find the proof of Fermat's Last Theorem?" into the search box, it will retrieve a PDF of the proof as the third result. It appears this special purpose computer solved the problem. But you immediately see the difficulty. Google search already knew the answer, or more precisely had indexed the answer. The computer was not tackling a random problem from scratch. It was tackling a problem for which it knew the answer, or at least where an answer could be found. There is no sense in which the search engine discovered the proof.

To really understand this objection we need to examine exactly what Turing and Matiyasevich proved.

An arbitrary problem is one you do not already know the solution to when you write the algorithm. You can think of it as a variable. Is there an algorithm that can solve problem 'X'? The alternative is a special program. It can solve problem Y. Y is a problem it knows. It must have the solution coded somewhere within it in a computably expandable way. You might think of this as a table of constants; problem Y has solution 1, problem Z has solution 2, and so on. But it could be more subtle than that. Problem Y might have a solution which is encrypted so you cannot recognize it within the program, or it might even be the result of some chaotic equation so complex that the only way to see it is to run the program and watch the output: no form of program analysis will give you any clue as to what it produces. There is only one stipulation. The answer to problem Y MUST be held within the program as a computable algorithm. Put another way, the computer must already be 'programmed' to answer the question.

Could a human mathematician be pre-programmed from birth? Yes, there is no fundamental objection to this. Mathematicians could be born to solve the problems they solve. But this would present a couple of issues. Where is this program stored? And who, or what, programmed the mathematician? Could we perhaps find an experiment to determine whether mathematicians are pre-programmed?

One view held by philosophers is that the Universe programmed the mathematician. They believe we live in an entirely determined Universe with no free will. There is then no mystery as to how Andrew Wiles came up with his proof. He was destined to do it from the dawn of time. The ink that fell from his pen to the paper was always going to fall in just that way. We live in a clockwork Universe and although we might feel we have free will, this is an illusion. I simply don't believe this. If I am right and humans do exercise free will, Andrew Wiles cannot be a computer. And because Andrew is not alone in discovering proofs, those mathematicians cannot be computers either. Humans are, therefore, not computers.

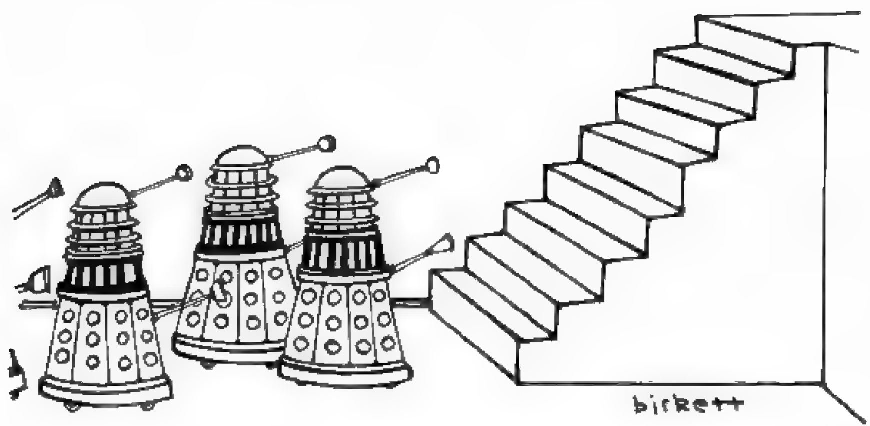
The Chance Objection

I said there was no automatic way to solve any problem above the logic limit, but this is not quite true. There is one automatic method you could deploy to generate a non-computable proof, the infamous 'monkeys and typewriters' idea where we use random chance to generate information. Many people have suggested it is possible to write a play such as Shakespeare's *Hamlet* by simply typing random characters until we happened upon the play. The argument is flawed.

The first flaw is the process would take a super-astronomically long time. Even if every atom in the Universe were a monkey with a typewriter, it would take orders of magnitude longer than the age of the known Universe to come up with the script to a play or a mathematical proof.

The probability of finding a solution to Fermat's Last Theorem by chance is about 1 in $10^{50,000}$. That's 1 with 50,000 zeros after it. For a comparison, there are only 10^{120} atoms in the known Universe. To be, or not to be, certain of finding the proof, you would need to run a computer long enough to calculate all the possible proofs up to the length of Wiles' solution. Currently, a computer using every particle in the Universe clocked at the Plank interval – the fastest conceivable computer running at 10^{34} operations per second – would take 10^{500} times the age of the known Universe to do this. If someone tells you this is *astronomically* unlikely they are making a huge understatement. A computer running until the end-of-time would only scratch the surface.

The second flaw is even more damning. Even if the monkeys succeeded in generating something interesting, something else needs to spot this. If an algorithm stumbled upon a proof of Fermat's Last Theorem, what would recognize it as such? There are no ways to systematically analyze proofs. There are no mechanical methods that understand these.



"Well, this certainly buggers our plan to conquer the Universe."

Dalek Trouble

"All non-trivial abstractions, to some degree, are leaky."

Spolsky's Law
of Leaky Abstractions

Consequences

Machines cannot discover theorems using algorithms, yet mathematicians do it all the time. Do the rest of us break the logic limit? It seems we do. People appear creative – painting, composing, sculpting and so forth. But, are these endeavors creative in the mathematical sense. To prove this, ironically we need to find something outside mathematics that is definitely non-computable. This is tricky. Most artistic things are fuzzily defined and there are no written rules we can apply. How can we prove a work of art could not have been generated by a computer?

Trivial proofs exist but they are rather contrived. For example, it would *not* be possible to make a film with a solution to the still unproven Riemann Hypothesis on the blackboard in the background of a movie scene. All the mathematics *Good Will Hunting* had been already discovered before the movie was made. New mathematics cannot be accidentally generated by a set designer – unless, of course, they also happened to be a world class mathematician.

These trivial proofs might lead a mathematician to argue the theory is proven. There *are* some artworks which cannot be computed. QED. But these are not very satisfactory proofs. I could create almost any movie I wanted without tripping over this rule. What I really wanted to know is whether *Good Will Hunting* as a whole could have been generated by a computer. Not that some weird version with a particular mathematical proof on the blackboard is forbidden. Movies are a difficult subject for this argument, but music is much easier to analyze. It is linear, highly mathematical and largely uniform by culture and language. Yet it is universally appreciated. Is music a computational or a creative endeavor?

Is Music Computable

To prove a piece of music is non-computable requires two tests. First to show we can 'reduce' it to a problem that is already non-computable and, second, to demonstrate it 'looks like' or 'sounds like' a piece of music. An accountant would say it needs to pass 'the smell test'.

The first non-computable problem to be studied in depth was Emil Post's Word Problem. Post was a contemporary of Alan Turing and studied at the Institute of Advanced Mathematics in Princeton. He solved the Halting Problem six months before Turing, but his proof used a complex recursive method called the lambda calculus. Turing's method was far more practical, which is why we now refer to Turing machines rather than Post machines. Later in his career, Post came up with a branch of non-computable mathematics called 'Post Problems'. They look like a puzzle you might find in a newspaper. Imagine starting with the word 'camel' and being asked to turn it into 'aardvark', using only a few simple rules. We'll make the problem very easy to start with: cam ↔ aard and el ↔vark. This solution is obvious; just do the substitutions and you are there. But what if the rules were a little more complex? Gennadií Makanin, a Russian mathematician based at the University of Moscow, found a set of extremely simple puzzles that are nevertheless non-computable. Here is one:

{"CCBB" ↔ "BBCC", "BCCCBB" ↔
"CBBBCC", "ACCBB" ↔ "BBA", "ABCCCBB"
↔ "CBBA", "BBCCBBBBCC" ↔
"BBCCBBBBCCA"}

Word Problem

Can a computer tell us which word problems have a solution and which do not? The answer is 'no'. Word substitution puzzles are a class of non-computable problem. Martin Davis proved this in 1948. Using a reduction argument we can use these word problems to prove some music is also non-computable.

Let us start by substituting the notes of the musical scale for the letters of the alphabet to create a piece of 'music'. Since it is a direct analogue of the word problem, we have created a non-computable piece of music. It is definitely non-computable, but is it music? If it just looked like a random jumble of notes it would be unconvincing, but luckily there are many forms of music that look exactly like a word substitution puzzle. Bach's *Art of Fugue*, the canons of Tudor composers such as William Byrd and Thomas Tallis, and the works of Grieg all use sequences of chords that move from one to the next using substitution rules. If you were to listen to the steps in our word substitution music, they would definitely sound musical. I think they should pass the main artistic criticism – that they should not sound formulaic.

But is any actual human composition non computable? Unfortunately, we cannot prove whether a particular piece of Bach, Tallis or Grieg is non-computable because we don't know the specific rules used to compose it. All we know are the general musical principles of harmony and counterpoint that applied at the time. We don't have these composers personal rule sets because they were held in their brain and they are, of course, long since dead. It is statistically likely that most pieces are non-computable because there are an uncountably infinite number of them, whereas computable pieces are merely countably infinite. But that's just probability; it is no proof.

I puzzled for some time whether there is a way to prove it but had to conclude it is impossible. However, and this is how creativity works, once I had given up on the problem, my brain continued to work on it. I was not conscious of this, I was only aware that failing to solve the problem annoyed me. I then had a Eureka moment. Although I couldn't *prove* a piece of music was non-computational, I could make one! – a piece that could not have been created using computation alone. This requires me to inoculate your brain.

Creative Inoculation

Take either Andrew Wiles proof of Fermat's Last Theorem or Alan Turing's proof of the Halting Problem; both proofs are non computable. Each document is made up of symbols, the Roman alphabet and some special Greek symbols such as α, β, ζ, and so on. Let us write out the symbols in a table and assign a musical note to each. It is straightforward to put these notes into a synthesizer and play the piece of music. I have provided a link to such a piece. Warning: once you listen to this you will have been 'creatively inoculated'.

This resulting piece of music, based on the transliteration of a proof, is non-computable. You might immediately argue with this, "The piece of music was translated from proof text to music file using a computer. It is clearly computed.", but this is not my point. The music could not have come into existence in our Universe as a result of a computation. It is a computable translation of a non-computable string. It could not have been generated solely by a computer: It was done in two steps, the first of which could not have been computed.

If, up to this time, our Universe has never contained a piece of music that was generated non-computationally, it does now. If you listen to this piece, you will find it impossible not to be somewhat inspired by it. You cannot erase the experience from your memory. And once you have heard it you will have been creatively inoculated. I have defeated Daniel Dennett and his like, and given you creative freedom!

www.jamestagg.com/noncompmusic

Having made at least some music above the Turing limit I could declare victory but I want to go further. Using the same reduction method, I believe we can show all art is above the limit. First let's attempt novels and plays. Do you enjoy those crime novels by Agatha Christie and Colin Dexter? It must be possible to construct a plot sufficiently complex, and a murder sufficiently baffling that it exceeds the logic limit. I could keep extending this idea to provide any number of examples and, therefore, prove all art and creative output is above the logic limit.

There are many other arts we could apply this argument too. In the visual domain there are non-computable images. In principle, it is possible, to draw or paint things beyond the capability of a computer. Roger Penrose has created non-computable visual puzzles such as tiling an infinite plain with special jigsaw pieces. Creating an image containing a solution to his visual puzzle is non-computable.

This extension argument also applies to me. There is an argument that I am a finite being and therefore can be simulated by a computer. Since I can be simulated by a computer, I am the same as a computer and therefore incapable of non-computable thought. The argument is as follows: James Tagg will have during his life a finite number of inputs and, equally, a finite set of outputs. This means you could model me using a

Jackson Pollock

computer. You could simply create a table of all the possible inputs and all the possible outputs I would make and this would be a perfect facsimile of me. A number of people have posed this as an argument to refute Roger Penrose's assertion that humans are capable of non-computable thought.

But this analysis misses a key point. There is no way to calculate all the contents of this table. My past could be tabulated. It is the history of all the things I ever did, but my future cannot. I might yet discover some great theorem that could not be computably generated. This would be a part of my output which could not be generated by an algorithm or any mechanical process. This forms a non-computational arrow of time; we can write down the past, we cannot write out the future. If a creative person such as Andrew Wiles could be simulated in advance, we would have an automatic way to find a solution to Fermat's Last Theorem. Since this is not possible, it follows that creative people cannot be simulated. This also means the Turing test is not passable by a machine. Humans can create; machines cannot. That is the difference.

Will Computers Take over the World?

Ray Kurzweil, the American inventor and futurologist, has suggested computers are getting exponentially faster and will soon reach such immense power they became effectively infinitely powerful. They could instantly answer any question posed and solve all our engineering problems. He dubs this point 'the singularity': a point of near infinite

Watson and Our Future?

computing power and therefore universal knowledge. This could herald a Utopian future; global warming, cancer, all things of the past. But computers might just as easily become bored and determine we humans are the real problem. If we are lucky, they may treat us as amusing pets. If we are unlucky...

These consequences might have come to pass if the answer to the Halting Problem were 'yes', but as the answer is 'no'! This is not the future we face.

Mummy, where do Bugs Come From?

One consequence of the logic limit provides a theoretical basis for the origin of computer bugs. The mention of 'bug' conjures up stories of dead creepy crawlies stuck in early computer circuits, but the term had been in use for over 150 years before the computer was even invented. Bugs are not simply annoying mistakes.If you misspell my name as Stagg instead of Tagg that's just carelessness. Real flaws creep into a computer program when you fail to understand Brooks' essential complexity, or by my terminology, you stray above the logic limit without realizing it.

Imagine we have created a piece of software. The software goes into test and is subjected to a range of use cases. Some of these will fail because we did not take into account all the real world possibilities. Then a strange thing happens. We get trapped in a loop of patching the errors in the program in a rather mechanical way. Find an error, patch it. Find another, create a work-around, and so on. By doing this, we are effectively mechanically generalizing our solution. This is forbidden as it breaks the Turing limit, so we can't mechanically solve a general logic problem above the logic limit. We need instead to use intuitive or creative thought. In our panic we did not stop, take a step back and engage our brain. Instead, we attempted, unsuccessfully, to blindly hack our way through the problem.

If we eventually succeeded in perfecting the code this way, we would have broken a fundamental law of the Universe. Something nasty would have to happen to prevent it, such as rupturing the space-time continuum or an event equally horrible! Luckily something prevents this and keeps our Universe intact – BUGS! Bugs stop us breaking Turing's limit.

The next time you curse a bug, remember if they didn't exist you'd be in danger of causing a logical paradox. There is no problem in redefining the domain and then creatively producing an all-encompassing design, but, you can't patch and hack your way there. This theory of bugs leads to an explanation for some modern programming rules of thumb.

Written specifications are valuable because they force you to lay out *the whole* problem. You don't need to be detailed regarding the depth, but should be expansive about the breadth, covering all the logical complexity. This might result in many details as a by-product, but a specification needs to delineate the edges of the problem space and not simply focus on a few key points.

Writing the tests for the software in advance is helpful as it is likely to tell you early whether your design encompasses the whole problem space.

Also, building a prototype, throwing it away, and then building the real thing can help greatly. It may be the only way to examine the edges of the problem space in detail. Armed with a full understanding, you can then imagine solutions to the complete problem in a single creative sitting. Whatever techniques you use to improve the quality of your software, remember you are engaged in a creative process that is not, itself, open to automation.

The Art of Programming

Programming is an art: a creative endeavor. It is also, of course, highly scientific. When you work with a good programmer – and I have been fortunate to work with some of the best in the world – they all follow a similar process. First they talk with you at length about your needs

Geek Humor

and examine the full scope of the problem space. Even if you say, "Oh don't worry about that bit," they always will. They want to know about everything. Then, they write a high-level list of features, some simple block diagrams, and occasionally a flow chart, only then do they begin to code, ticking off the list as they go. Sometimes, they will check to see if their list is the same as your list but more often they will come back and just check the high-level purpose. "If I give you something that achieves this, will that do it for you?" They test as they code so you end up with is something that meets your high-level purpose, and can prove it does so in its own right. At the end of the coding they write out the specification for the project so that they can remember what they did, or a colleague can pick it up in the future.

This is *not* how students are taught. Students are told to write a detailed specification at the start and then simply implement it. If you've been following my argument, they are being taught to do something impossible! There is no 'just' to programming. Sometimes teams are even split up so that one person writes the specification and another the code – again an impossible task. If the specification was the *answer* to the problem, it must have required creative thought to develop and so would be as complex as the program itself. Since it is not yet a program you cannot test it, so it becomes an untestable solution to a creative problem. Since the specification is *not* the answer but rather a general list of tasks, the great danger is to give it to a separate programmer and

they implement it mechanically. You see, of course, the problem. It will be riddled with bugs because they have missed the creative step of imagining the whole problem and solving it in the round.

This fundamental misconception of software is common in many organizations. "Ah," says the finance director, "I'll write a very detailed spec and then we can get someone cheap to *just* program it." This does not work. If the finance director has done the creative work of taking a problem and turning it into a detailed specification for the programmer to 'just program' – removing any ambiguity and therefore the creative overhead – he will have all but written software himself, albeit in a computer language of his own making. On the other hand, if the specification is a linear list of issues with no creative thought, he will not have reduced the time needed to program. He may have improved the quality by effectively getting a second pair of eyes onto the requirements gathering stage, but this does not help the programming effort itself.

Ideally, you should never split up specification and coding. It is a creative process best handled by very small numbers of people working intensively on it. Of course, there is one big problem with this: some software tasks are huge. Before we look at the science of splitting up a software project, it is worth pointing out that many of the most famous projects were written by one man. I have met many of these people and they are all exceptional – Linus Torvalds, Linux; Anthony Minessale, FreeSWITCH; Daniel-Constantin Mierla, Kamailio; Eric Allman, ,SendMail. Before splitting a project between many people, it is worth considering whether you can give it to just one individual. To do this you will need to unload this person of ALL interruptions and administrative burdens. This is the most effective way to solve a creative programming task. Practically, once your task is over the limit for a single human, a software project must be split up. This requires great care. Dividing a problem efficiently means specifying the interfaces between them and decoupling the components. This is the art of an architect or a producer in the creative arts. The creative process operates similarly in other walks of life. There are many examples of successful creative duos – Rogers and Hammerstein (*The Sound of Music*), Ben Elton and Richard Curtis (*Blackadder*).

Good managers, therefore, find ways to break projects into manageable sub-projects that can be worked by pairs or rely on single super programmers with support around them. If you are lucky enough to gather together a group of super-programmers and can divide a problem efficiently amongst them, you can achieve great things. You see this pipeline in movie production. A script writer generates a script creatively. The casting director finds the actors, a director is in charge of filming, and an editor puts it together. In very great movies you will often find a great director or producer who had a hand in almost everything holding it all together. They are often accused of micro-managing but you can see that's what they must do. They are the super programmer with the whole creative work in their head, and an eye on the audience and financial backers.

If you talk with great programmers you will be amazed by their breadth of technical, commercial and product knowledge. Why do they need all this commercial information to do their job in the round?

Rules and Tips

I began writing some rules on how to split up a project, and almost immediately ran into exceptions and special cases. The job of dividing things into sub-tasks is, itself, a creative problem and must not be done mechanically. Any 'one size fits all' rule will fail and you must apply domain knowledge and careful thought to the process.

It is the job of architects or a senior engineer to split projects into smaller chunks. To do this they must accurately 'guess' boundaries between subtasks to create self-contained, creatively solvable problems. This can be done by either vertical or horizontal abstraction. Both have their problems.

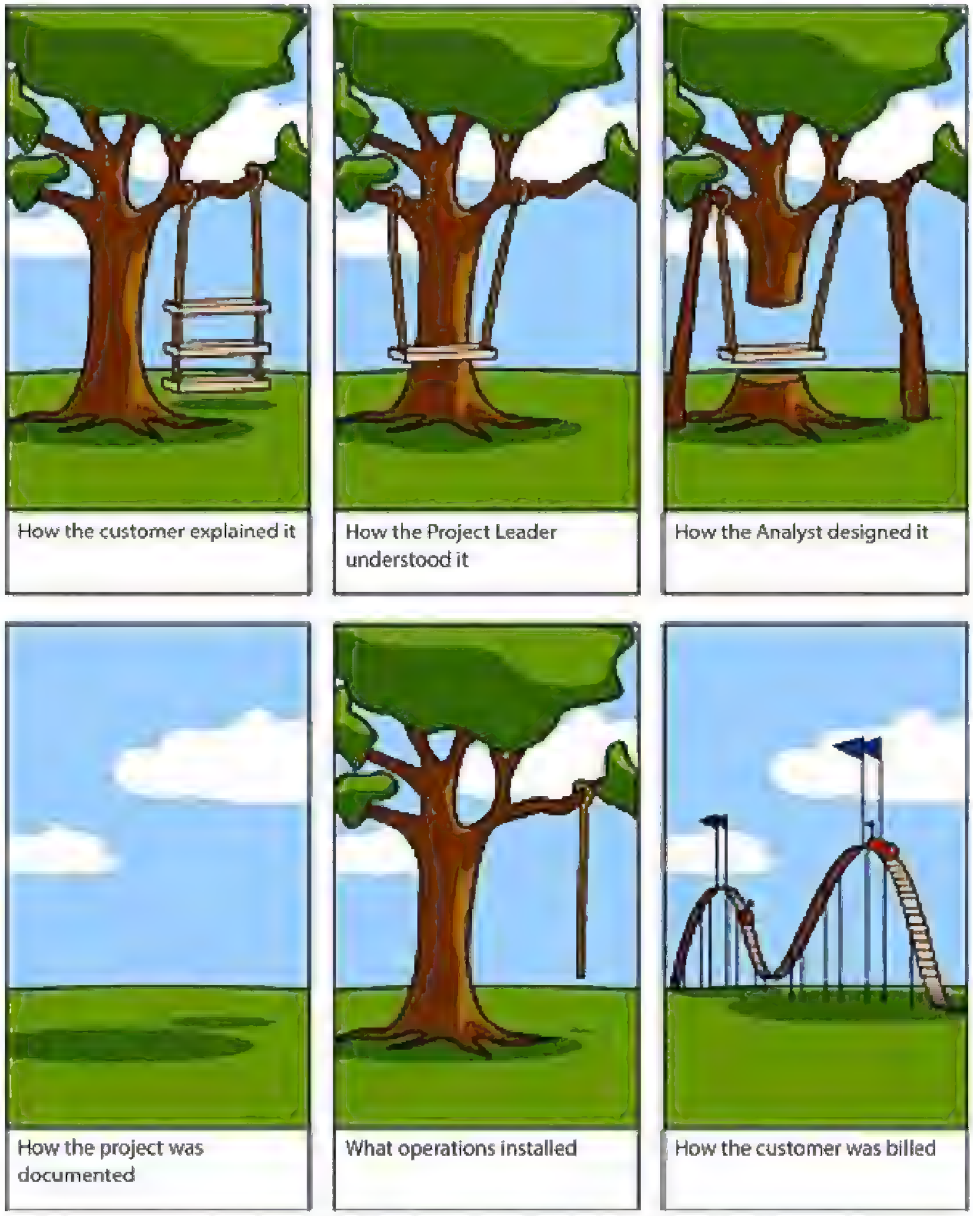
Horizontal abstraction is the simpler of the two to understand, and the more common. Computer systems are built 'on the shoulders of giants'. That is to say we no longer need to place individual pixels onto the computer screen. We can assume a computer will draw a square if we specify the dimension and coordinates of the center. That's abstraction. Today's computers are even more helpful. We can ask them to draw a rotating cube lit from a certain angle and the computer will do the whole job for us. But, there are always practical limitations to this.

I want my cubes to move around the screen naturally but I am not sure what physics model has been implemented. What will happen when they bump into each other? If the abstraction is not thoroughly thought through they pass through each other in a very odd way, breaking up and showing me they are really made of triangles, the illusion of three dimensions is lost. Whenever we work at an abstract level, we risk being exposed to its inner guts at some point. Joel Spolsky, a computer scientist who worked on Microsoft Excel, proposed the Law of Leaky Abstractions to explain this. An example of his law in action is the TCP/IP protocol stack that transports data over the Internet. The stack is hugely reliable, yet I have to debug one of these stacks at least four times a year!

The problem is that the TCP (Transmission Control Protocol) is designed to provide reliable delivery of information: internet pages, my bank account and the like. But, the internet protocol 'IP' on which it relies is only designed for best-efforts. When a link loses a packet of information, the TCP has to retransmit it. This takes additional time. TCP provides an abstraction of a reliable connection, but the implementation is not as robust as it may seem, and the details leak through as variable latency and throughput. This explains why your web pages sometimes do not completely render. You are told it is reliable, but often it is not! Experience is so valuable to a programmer because they know which of these specifications to take with a pinch of salt and when they are likely to leak. They are battle scarred by previous naivety.

I think Spolsky's Law follows from Rice's Theorem and ultimately from Turing's no halting proof. If leak-less abstraction was possible you could, in principle, write a recursive partial halting solution. By layering abstraction on top of abstraction you would be able to solve some very complex problems, eventually including the Halting Problem. We know this is impossible, so non leaky abstraction cannot exist.

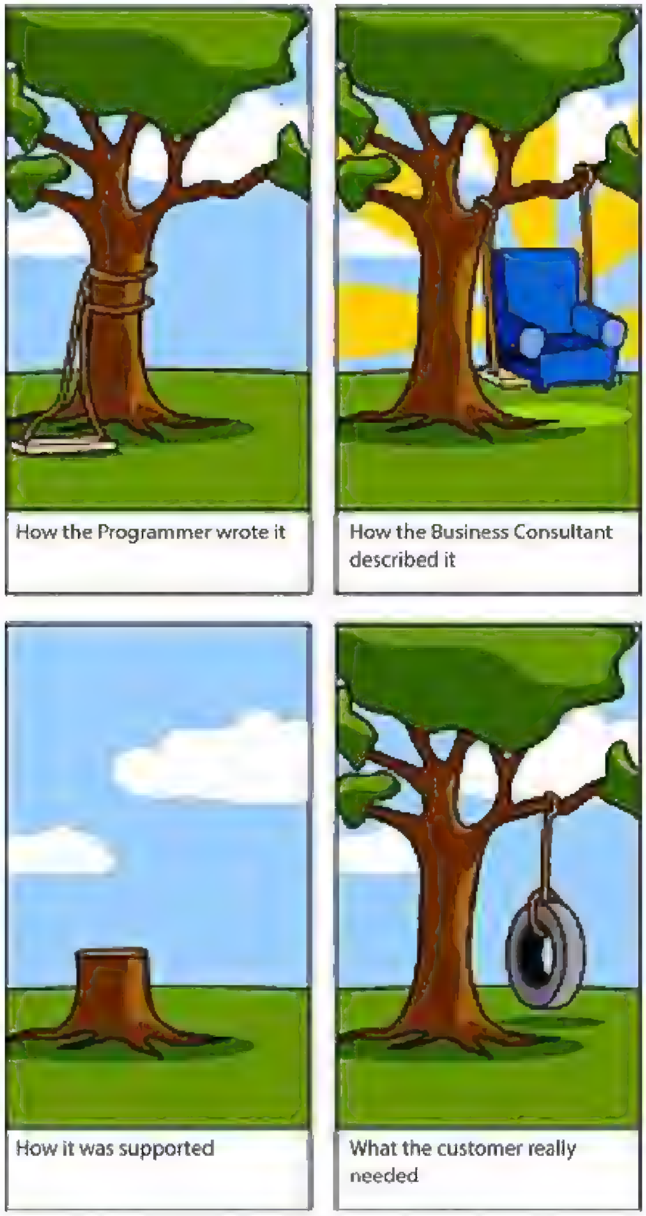
The other method of splitting software is vertically. This is often done following the natural boundaries of an organization: functional or geographic. Again there will be leakage between the systems; the data you get from the finance department might not be detailed enough for



Specification Cartoon

the engineers or vice versa, and so groups have to interact. The main problem with vertically divided software is each group tends to reinvent the wheel, so you end up with multiple similar implementations of the same thing.

All said, the architectural job in software is a dynamic one. You can split up software into separate elements but you must take into account the leakage between them. When you detect a leak you must bring people together to collaboratively solve the problem, rather than insisting on the original partitioning. While doing all this you must keep track of the overall aim and all the irritating small details contained in the many



lists that form the project specification. I should confess that I am no great fan of specifications, because they can mislead you into thinking you've solved the problem, but I concede a good specification is helpful. Spolsky's Second Law is 'Always write a specification.' Engineers should collaboratively write the specification as a response to the desires of the project creators. But they must not blindly implement the specification they've been handed. They must not forget the creative element.

The Role of 'Process' in Creativity

We hear a lot about 'process' when developing software and other creative tasks. The first thing to realize is process does not write software and every moment spent on process is a moment not writing software. Excessive process can bring the productivity of the average programmer down from a thousand lines per day to one. On the other hand, we all know that using no process to write software results in useless software. Good solo programmers, playwrights or composers are surrounded by lists and post-it notes full of process. Where is the balance to be struck?

In my view 'process' is there to help humans with the tasks we find naturally difficult. Humans, as we know, are dreadful at remembering lists of symbolic information. Give a human ten numbers to memorize and they will quickly forget them. Give Microsoft Excel ten numbers and it will remember them forever, or, at least, until your next upgrade! So the first job of process is to collect lists of things and sometimes even lists of those lists.

Another significant affliction affecting humans is procrastination. We tend to put off decisions. Process can set waypoints; when will the job of splitting a project occur, when will we begin the test, and so on.

The third job of process is to keep track of the division of labor – if the project has to be divided. Who will do what? Essentially we are back to lists again.

The most important job of process, in my view, is to keep track of scope. 'Logical scope creep' when unrecognized destroys software projects. Scope creep is fine if it just adds more linear work. "Could we add three more product types?" "Could you do another language?" "Can you make this interface prettier, less cluttered?" It may cause a busy team to groan, but it does not damage the integrity of the design. To put it back in Brooks' language, accidental creep is fine provided you add some resource. Essential creep is not. Adding the french language prompts to a project in English might be fine, putting language translation into a project may be a step too far. The project may have strayed into a different logical class. Increases in logical scope often require redesign, you must stop and re-architect if you are to avoid bugs in plague like quantities.

If programming software is a creative task, how can we help improve productivity? The most important factor is to provide uninterrupted peace and quiet. Programming is a task where people need to hold many ideas in their head at the same time, and this requires deep concentration. To get some idea of the creative process at work, listen to the excellent TED lecture by John Cleese.

A common and costly mistake is to put off thinking about a class of things you are going to need in the next release because of time pressure. 'Time out, that's for the next release' and similar statements spell disaster for the future of a project as when you come to the next release, you may have to rewrite much of it from scratch. This is why good architects are so valuable. They anticipate the future even when they are told to ignore it and ship now!

Just as there are artistic geniuses, there are programming geniuses. Hold onto them if you get one. They are rare. We don't know if they can be made or they are lucky accidents, but statistics shows that some people are 1000 times more productive at writing code than the average. If you can find lots of them and make them work together you will build the next Google or Facebook. If you have a tight deadline, a super programmer may get you out of a hole, producing in a week what might otherwise take a year. Remember your great programmers will most prolific if you can get process and distraction out of their way. Just make sure they have a clear idea of purpose.

Laws

A programmer interrupted eight times a day does no work.

A creative person interrupted eight times a day does no work.

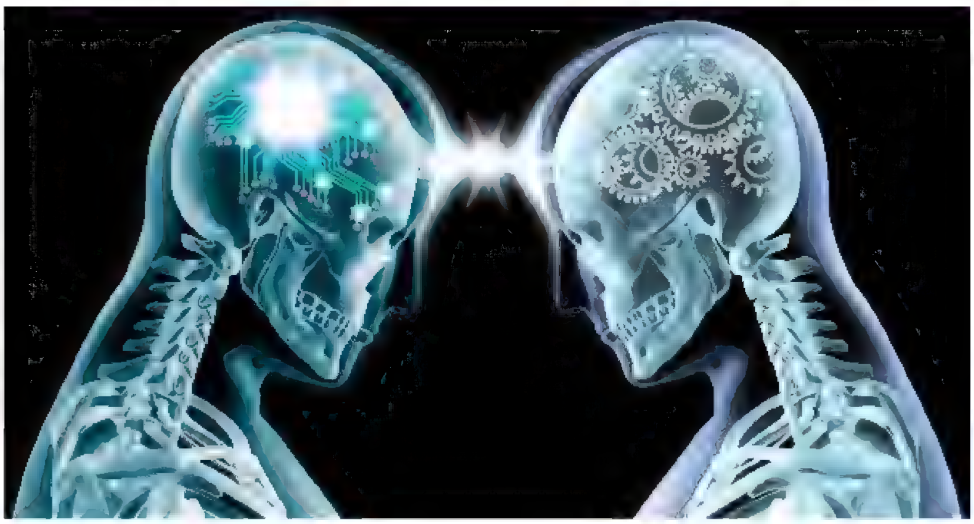
Programming is a creative endeavor.

There are creative geniuses. Hold onto them.

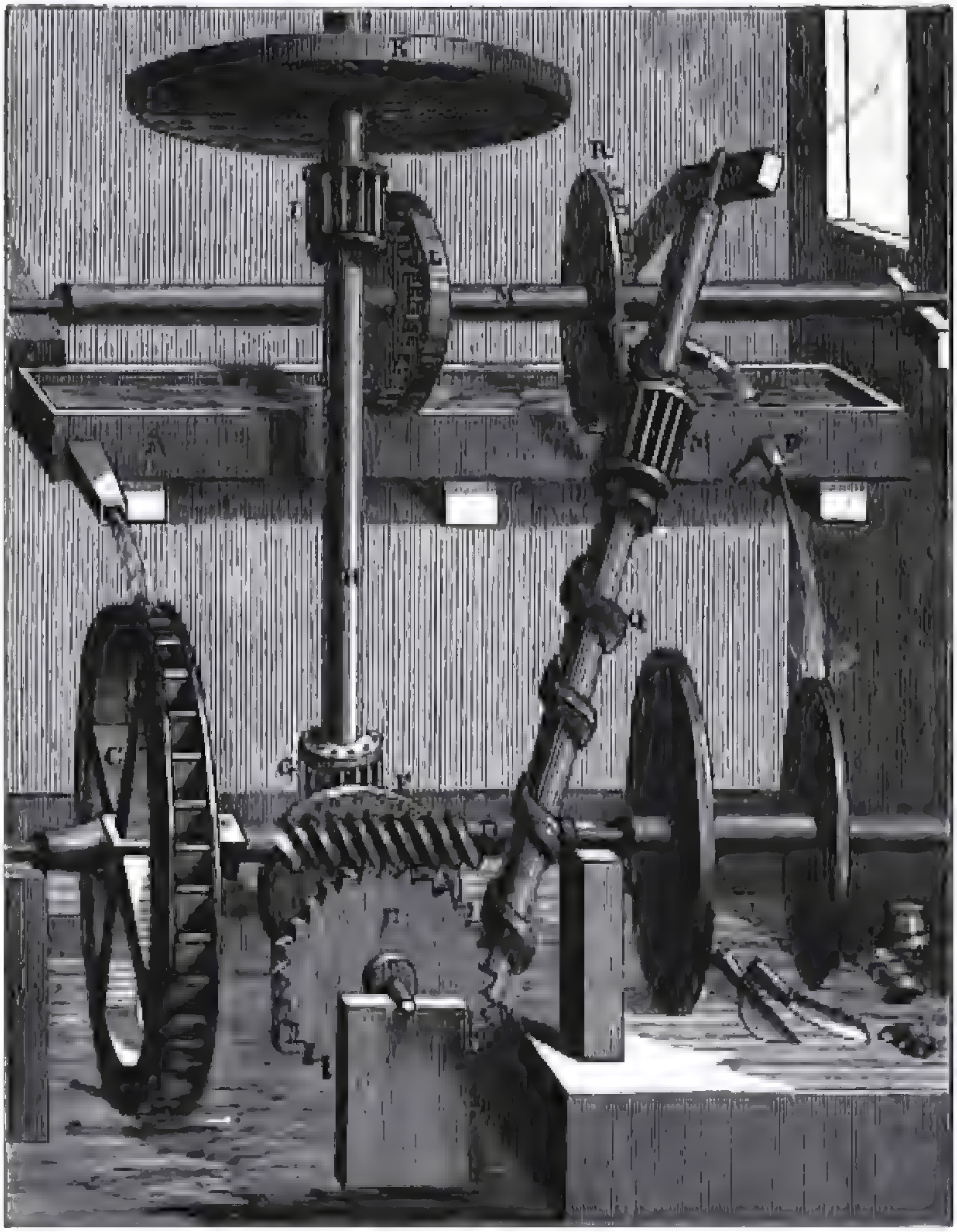
Bugs save us from collapsing space-time when we are lazy and try to use mechanical means rather than creative thought to write software.

Chapter 12

HYPER-COMPUTING



What's in a Brain



Perpetual Motion from the 1600s

"If you are in a spaceship that is travelling at the speed of light, and you turn on the headlights, does anything happen?"

Stephen Wright

If you believe humans outthink computers, be warned; you are in controversial territory. This would need a hyper-computer and many scientists speak of these in the same breath as perpetual motion machines.

I'm not sure it's an entirely fair analogy. We understand machines, and the physical laws of our Universe forbid perpetual motion. We don't understand brains, so we can't reasonably dismiss human hyper-computing. Humans commonly demonstrate one clear example of thinking which appears to break the Turing limit, namely finding solutions to mathematical puzzles. We need an explanation for this. Let me take you on a whistle-stop tour of all the schemes people have imagined that might lead to a hyper-computer.

A hyper computer is a machine that can calculate a function which a Turing machine can not. For example, when given a number denoting a problem such as Fermat's Last Theorem, it can give me in return a number representing a valid proof. We are not concerned here with speed. We are talking about fundamental 'do-ability'. Such machines are often dubbed 'super-Turing'.

Epic Fails

Let us first look at some proposals that blatantly fail. My children call these 'epic fails', and they are the perpetual motion machines of the hyper-computing world.

Could we run many Turing machines at the same time, perhaps even an infinite number? Then we would have a much more powerful machine that must beat the Turing limit.

The answer is no.

Turing machines are already infinitely powerful and we know from our chapter on infinity that all countable infinities are the same. Infinity plus infinity, infinity times infinity, infinity to any power; all are equal. One single, fast, one-dimensional machine can simulate them all. We get no greater power with an infinite number of similar machines.

The next technique which might realize a hyper-computer is to have a machine which simultaneously runs every possible branch in a program. Each time the machine gets to a point where there is a binary decision, it can take the 'yes' branch, spawn a copy of itself, and run the 'no' branch as well. Logically this machine should be able to calculate anything since it tries *every* conceivable option. The process is called non-determinism. This doesn't mean the computer has free will. It just means the computer never chooses one option over another. It just assumes each could be correct and travels down both. Solving a problem using a machine like this can be fast. The problem is this machine has no greater power than a regular Turing machine. Let me show you why.

A non-deterministic machine is essentially the same as a single Turing machine; each time there is a branch in the program you would start running two processes. The first process works on every even tick of the computer clock and the other on every odd tick. Now we have a single machine running two branches at the same time. Using this trick over and over again, a single machine can run a program exploring every possible branch. Although it generates an enormous number of branches and takes a huge time to run, it is still a single machine and we have an infinity of time on our hands. Therefore, the machine is limited as before.

We are not doing well so far and we have already exhausted an infinite number of options! Let's try a different tack. We know true randomness is non-computable, the sort of randomness generated by the Lavarand we examined earlier in the book. Might this help? Truly random processes can't be simulated by a computer. If we throw this into the pot might it let us compute something a Turing machine cannot?

Again, no.

This idea still only generates a machine as powerful as the non-deterministic machine above. A non deterministic Turing machine runs *every* possible program. All a random one does is choose *some* of the same paths at random. It, therefore, can't be any *more* powerful. The one difference is that it can generate non-computable numbers. However, the only interesting characteristic of these numbers is they are truly random and this randomness was an input. Their presence does not make the machine any more powerful.

There are quite a few proposals for hyper-computers that are just cleverly dressed up versions of the machines we have already met and dismissed. For example, it has been proposed the Internet could form a super-Turing machine. This is known as a site machine because the processing is distributed across many sites linked together through the Internet. It is proposed each site could act as an oracle to the others. This is quite an elegant idea, and some proofs have been offered that show such a machine is capable of generating non-computable functions. The problem with this idea is that you can simply draw an imaginary line around the whole site machine and it looks exactly like a big Turing machine. There is no conceptual difference between such a machine and a regular computer with subroutines. After all, that's in Turing's original proof. Again we have reached a dead end. We need something qualitatively different to a traditional computer in order to break the Turing limit. The obvious place to turn is the quantum world.

Quantum Computers

Quantum computers have had an extraordinary run in the press recently. It has been variously claimed they offer limitless computing power and can break all known security schemes; cracking, for example, the prime factors that form the basis of public key cryptography. This is big news. These codes are used to protect all the financial transactions we make on the web.

In a regular computer, bits of information are processed by switches that make simple 'yes' or 'no' decisions. In a quantum computer each switch can take both the yes and no branches, at least for a short time called the decoherence interval. The calculations are said to be superposed. This allows a quantum computer to calculate exponentially, rather than linearly, as the number of logic gates increases. Grover's algorithm and Shor's algorithm use this superposition to speed up factoring numbers and looking things up in databases, respectively.

Grover's algorithm gives us the ability to find something stored in a random place without having to look in every box. If you think about a standard search, say for your lost car keys, you must look everywhere to guarantee finding them. It does not much matter in what order you do it. When you are halfway through the search, you will be 50% likely to have found your keys. But, with a quantum computer, you can be fuzzy and look in many places at once. A quarter of the way through a quantum search, you are 50% likely to have found your keys. That might sound like a small improvement, but when working with very big numbers, it makes an enormous difference.

Shor's algorithm works a little differently and, yes, it does allow a quantum computer to break Internet encryption, so the newspaper headlines are true up to a point. Some time in the future we will need to move to a more secure type of encryption.

The largest quantum computers today can process 300 qubits at a time or remain 'coherent' for about two seconds. These results are pitifully low. The largest prime number factored so far is 143, a mere 7 bits long! By way of comparison, internet security routinely uses 1024 bits. But, quantum computers are improving exponentially faster than classical computers: They really do change the rules of the game. If you remember our discussion of chess, the quantity of space needed for a calculation can be the limiting factor. A quantum computer is very space efficient. When the computer branches and makes a copy of itself, it does so without needing more space. There are two theories for how it does this, (well, three, but the third is highly controversial). The first theory is the computer doesn't need the space because it hasn't made its mind up yet; somehow the calculation floats in an undecided state. The second is that the computer puts a copy of itself in a parallel Universe each time it branches. When the calculation is over, either all the Universes collapse to a decision, or *every* possibility is chosen in some Universe or other and they all go on their merry way! This is the 'many-worlds' interpretation of quantum mechanics and we will return to it later in the book.

We have now explored all the straightforward ways to make a hyper-computer, and all have failed. We need something still more exotic.

More Horse Power Needed

Is there anything more powerful than a Turing machine?

Yes, in theory, there is.

The first person to explore ways of breaking the Turing limit was Turing himself. He cut right through the problem by proposing the existence of an oracle function. At any point in a computation, you could ask this function a question and it would give you the right answer.

We must leave completely aside the question of how this wonderful oracle function is constructed. All we know is it can't be a machine. If it really existed, a Turing machine that was able to consult it would be able to answering any question you put to it. That is a hyper-computer.

Unfortunately having access to such an oracle does not get us far. We can use it to compute numbers we could not otherwise have obtained – or answer a single question – but it does not give us a general-purpose way to solve further problems outside of the logical area we asked it to answer.

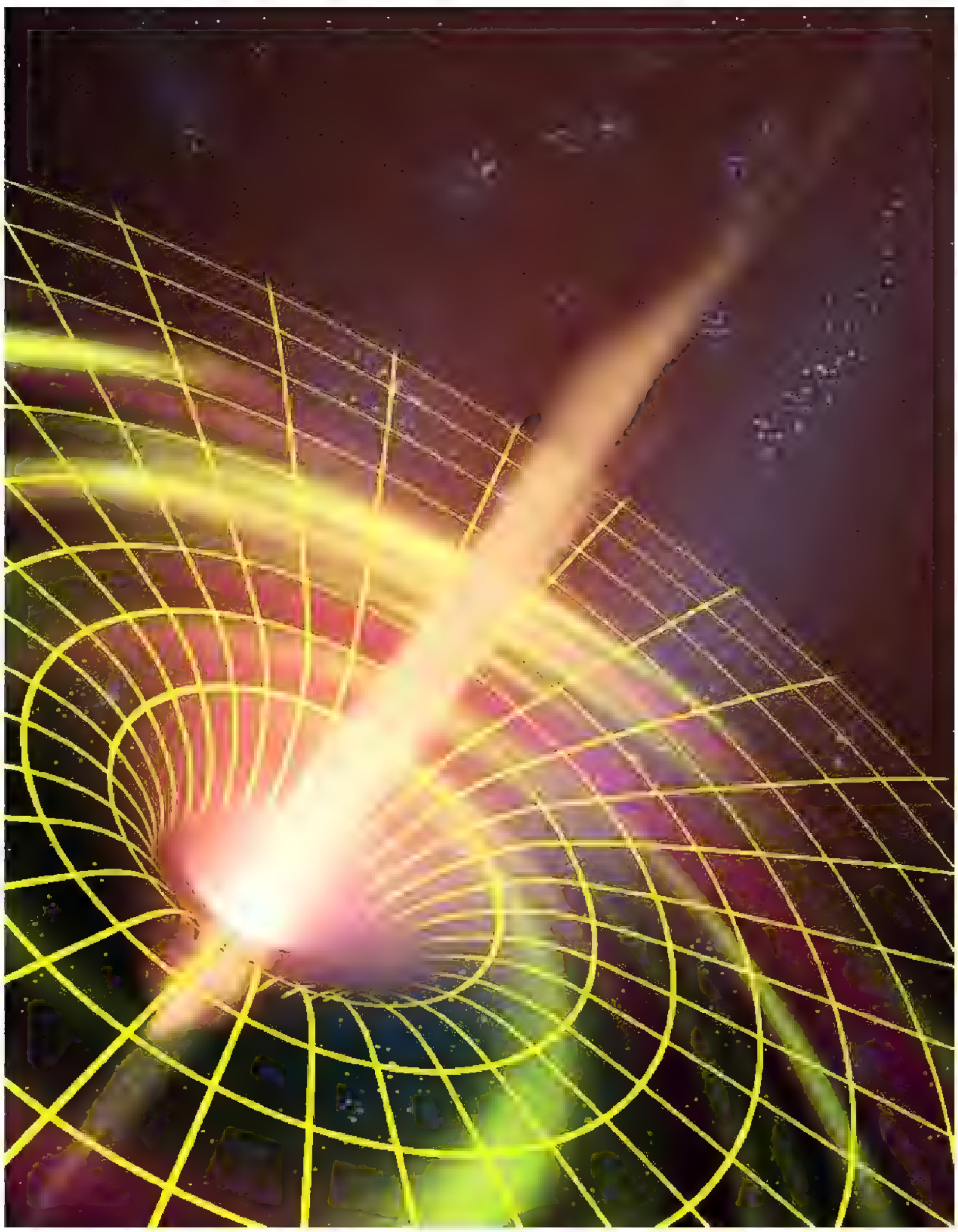
Each time the oracle answers a question we break the limit a tiny bit. Each question and each answer moves us forward, but does not give us something universally applicable. If I ask the oracle to prove Fermat's Last Theorem it will give me that answer, but this does not turn me into a creative mathematician, able to prove any other theory. You can test this by typing a mathematical question into the Google search box. Does obtaining an answer make you better at mathematics?

In any case, an oracle is not and cannot be a machine, so it does not lead us any further in our quest to build something super-Turing.

The Weird and Wonderful

There are some really weird and wonderful proposals for machines capable of super-Turing thought. Let's take a bit of a flight of fantasy.

If we could make a spaceship survive the inhospitable environment near a spinning black hole, it might be possible to send information backward in time. We could see the answer to a calculation before we had to go to the trouble of calculating it in the first place.



Black Hole Malament-Holgarth Space

David Malament and Mark Hogarth of the University of California, Irvine have proposed a form of space-time called the Kerr Metric. This allows a machine to break the Turing limit, but has the drawback that as it does so it falls through the event horizon and is sucked into the black hole. We might discover new information but are now trapped inside the event horizon unable to communicate it – a form of cosmic censorship.

Candidates for a hyper-computer that could fit inside a human brain include mathematical curiosities which stretch the concept of infinity. The easiest to understand is the Zeno machine. In a Zeno machine a computer runs each successive step of a calculation in half the time of the previous step. The computer can pack an infinite quantity of computation into each finite time interval and can therefore outperform a Turing machine. This theory fails at a practical level because we simply can't build such a machine.

There are numerous weird suggestions for mathematical super-Turing machines, and many are described on the Internet. They all fit broadly within the two models above: modifications to space-time or peculiar mathematical paradoxes. The inspiration for the true solution to super-Turing thought may lay in there somewhere, but there are some more plausible proposals to look at next.

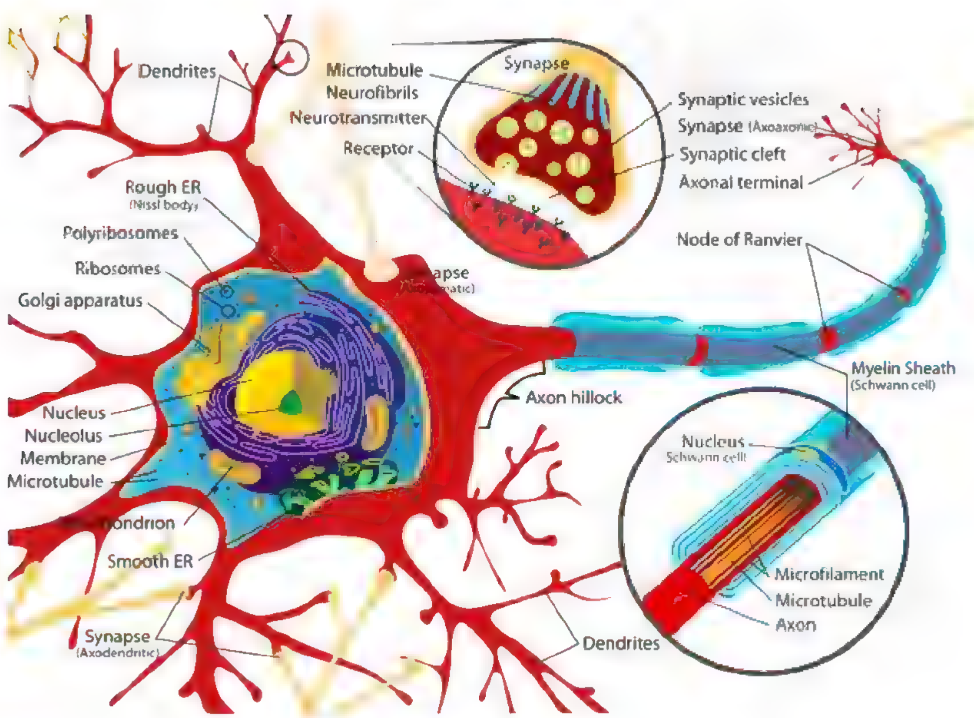
Plausible Ideas

I have characterized the next set of ideas as plausible, but they may still be highly controversial. My only criteria for plausibility are that the mechanism must outperform a machine limited to counting numbers, and it might fit inside our skulls. No black holes allowed.

One interesting proposal for a super-Turing machine that could fit inside our skulls is the Adaptive Recurrent Neural Network, 'ARNN' proposed by Hava Siegelmann of the University of Massachusetts, Amherst. An ARNN is a neural network with real number weights. As you recall, real numbers are equivalent to the continuum infinity, a larger infinity than that of counting numbers.

This is the infinity that defeats a Turing machine, and Siegelmann harnesses it as the basis of her computing machine. She argues that, although the machine cannot be programmed as it is impossible to write real numbers down, once it is running, the weights diverge and real numbers will be used within the network. These real numbers allow the machine to compute using numbers that are not, themselves, computable and this is where the machine's greater power comes from. Of course such a thing might easily fit inside our skulls, and the physics within our brains are certainly capable of using real analogue values.

The biggest stumbling block for Siegelmann's idea is the information that gives her machines their power is fine grained and easily destroyed by noise in the environment. This is not just from the sort of electrical noise we hear when our cell phones interfere with the radio, but the precision required by her machines is so exacting that anything might interfere with them. For example, gravitational waves caused by the motions of nearby stars would disturb calculations at only the fiftieth decimal place. Since it is these digits that constitute the difference between an ARNN and a regular Turing machine, most people conclude ARNNs can't work. There is one effect stemming from the quantum world which might come to the rescue. The potential to do something in the quantum world is sufficient to modify the behavior of a system even if the system does not actually do that specific thing. This is called a counterfactual process. The possibility an ARNN might perform infinite precision calculations may be enough to give the machine the edge, even though in practice it is disturbed by noise. This is speculation upon speculation, but interesting nevertheless.



Neurons and Microtubules

Roger Penrose is fascinated by such counterfactual experiments and is inspired to think such effects might have a role in non-computable thought. It is his 'machines' we will look at next.

Penrose-Hameroff Machines, aka Brains

Roger Penrose of Oxford University and Stuart Hameroff of the University of Arizona have proposed a very different way to understand the workings of the brain. They focus on the much smaller scale structures within neurons called tubulin microtubules. If you watch a brain form, the dendrites grow towards each other, twisting and turning rather like the growth of a plant as viewed in a slow motion nature film. This motion is controlled by micro tubular structures formed of a protein called tubulin. Tubulin is made from peanut-shaped polar molecules that self-assemble into helical tubes with a radius of just seven molecules. The tubes bundle together to form the backbone of neurons. The peanut-shaped molecules are bipolar switches and can flip between two states. This allows them to bend into different shapes and, in the most extreme example, to flap fast enough to propel small organisms such as paramecia. It is also, interestingly, the protein that unzips the double helix when a cell divides, and so plays a fundamental role in our evolution.

Penrose and Hameroff suggest these tubes form the true processing element in our brains. The walls of the tubes are formed of successive alpha and beta tubulin molecules. Each of the tubulin molecules can flip between two states, propagating a ripple along the tube wall. The scale is small enough for quantum effects to matter, and Hameroff suggests quantum error correction keeps the ripples from decohering too fast. Because the processing is happening at a molecular level rather than at the scale of a neuron, the brain would be considerably more powerful than a count of its neurons would suggest. They propose increased computing power would stem from three sources: There are many more tubulin molecules than neurons; the micro-tubes could perform quantum computation, and the micro-tubes are capable of non-computable, conscious, thought.

Measurement of a quantum process is the only candidate we have for a non-deterministic physical process today; all other physical processes are deterministic. Penrose argues that quantum processing in the brain spontaneously collapses in decision making because of the interaction between quantum superposition and gravity. The arguments are put forward in two books: *The Emperor's New Mind* and *Shadows of the Mind*. This theory remains controversial for two main reasons. First, most people see no need for super-Turing thought. They believe computers are sufficient. Second, they believe the brain is not a hospitable place for quantum effects: it is too hot and too chaotic. Indeed, until recently people assumed quantum effects would have no place in biological entities, but this orthodoxy has recently been overthrown by the discovery of quantum processes in photosynthesis. The paper by Travis Craddock of Nova and others suggests there may also be quantum structures in the neurons of our brains and we might possess quantum computers after all. But, remember, Penrose and Hameroff don't only need quantum coherence within our brain to explain consciousness. They also need gravitational effects.

Chapter 13

HYPER-COMMUNICATION

World Wide Communication

"The single biggest problem in communication is the illusion that it has taken place."

George Bernard Shaw