David J. Kinniment

He Who Hesitates is Lost:

Decisions and free will in men and machines.

School of EECE, Newcastle University, Merz Court,
Newcastle upon Tyne, NE1 7RU, United Kingdom

Main contact for enquiries: Professor Alex Yakovlev,
Telephone: (+44) 191 2228184,
E-mail: alex.yakovlev@ncl.ac.uk

Contents:

Preface. The ability to choose

What is it that distinguishes people from machines? This question seems simple. Machines are mere inert collections of metal and plastic that simply respond to outside influences, push the accelerator, and the car goes faster, push the brake pedal, and it stops, whereas we believe that people have some internal quality that makes them individuals, and enables them to make choices for good or for evil. The concept that people have free will and consciousness, and machines do not is something that few question, but the more we try to explain what these ideas mean, the more difficult it seems to be to pin down any concrete notion that clearly distinguishes between men and machines.

Machines do not appear to decide for themselves, but people do, machines are subject to outside forces, and the laws of nature, they have no soul, they are unable to make a moral choice. But what is this soul that allows people to make moral choices but not machines? Perhaps we should start, with the Greek philosophers, whose world was one of harmony, in which rational thought and whole numbers could explain everything. The earth was at the centre of the universe, and there was a simple relationship between the diameter and the circumference of a circle. For them, the reason that the earth stayed at the centre of everything was that it had no reason to do otherwise. If there was no reason to go to the edge of the universe why should it?

In the middle ages concerns were more with the spiritual than the materialistic, God was the source of everything, but mankind was in some way special, so there was a need to understand this special relationship between God and Man. Men have immortal souls, but what about animals? Do animals have a soul? And if they do, do plants also have a soul, and so on down to the lowest form of single-celled life.

Why stop there? If a virus has a soul, why not a rock or a volcano? Volcanoes are not things you would want to annoy, they have in the past been worshipped as gods. What is animate and what is inanimate? This is a circular argument, where we just replace the words ‘having a soul’ with ‘animate’ and get no further. The debate in the middle ages centred on an apparent paradox called the paradox of Buridan’s Ass in which an Ass (an animal chosen apparently because of its obvious stupidity) is placed half way between two equally attractive bales of hay. Unfortunately it is unable to decide which one to eat first, and thus it starves to death. If its actions are purely determined by outside circumstances, like the placing of the bales of hay, their attractiveness, and so on, and past history, we can easily propose that there is nothing in these circumstances that inclines it one way or the other. So it starves, it is no better than a stone, which has no forethought for the future, no moral judgement, and no free will.

But people have free will, they are not the slaves of circumstances and history, and can see that without making a choice between two food-laden tables, they might stay half way between and die. And it does seem possible in principle that a man might find himself in such a position. If he did, wouldn't he always find some way of choosing a table to go to first rather than let himself starve to death? `If I concede that he will [starve to death],' said Spinoza (1632 - 77), `I would seem to conceive an ass, or a statue of a man, not a man. But if I deny that he will, then he will determine himself, and consequently have the faculty of going where he wills and doing what he wills.'

So what is this ‘differentiating principle’ or ‘origination’ that allows a man to choose one course or another when both seem equal in value? Is it something that originates in the soul, and allows a man to take a decision, which is in opposition to all of his history and his circumstances? Or is there no such thing, and we are all subject only to outside forces, history, our genes and the laws of physics? In short, is everything already determined?

Computers seem to be the ultimate example of machines whose outputs are determined only by their inputs and the history of those inputs. They are expected to be thoroughly reliable and deterministic, and consequently are taking over many of the decisions that previously had been made by people, such as air traffic control, but in the 1960's it became apparent that they were not as deterministic as we would like, and could make errors in an non-deterministic way. This turned out to be a fundamental problem, which cannot be avoided. It is not just associated with computers, but with the nature of decisions, and is associated with even the simplest mechanism asked to make an arbitrary choice. If both choices seem equally desirable how is the choice made? In the extreme, if the basis for a choice does not exist, either the machine cannot make it, or it is made on the basis of some internal quality innate in the machine and not dependant on outside influences.

Since an arbiter in a computer is purely mechanical the problem of choice without preference could be observed and measured without the need to take account of humanity. The religious and philosophical implications were no longer relevant, and the results of the measurements made it clear that while it was not possible to get an exact balance of preference, as the balance became more equal the time taken for the choice increased without limit. More importantly, well before the donkey starved, random fluctuations in the decision making process would determine both the outcome, and the amount of time taken for a decision.

Results like these seem difficult for people to accept. They hanker after 'yes', or 'no' rather than infinite variation in decision time, and certainty rather than the statistics of probability.

Even for computers, we cannot be absolutely certain that the results of a computation will be always be the same, machines are usually reliable and deterministic, given a set of inputs, the output can be predicted. The opposite is also true, the roll of dice cannot usually be predicted, but over many rolls of the same set of dice any bias in the weighting will become apparent. One number may come up more often than others. Machines are always partly deterministic, and partly non-deterministic.

If determinism exists alongside non-determinism in computers, why not in men? Free will was once thought to be one of the defining aspects of the soul, but a soul was something that could only exist in people. The medieval idea of origination and free will, which existed only in people, appears also to exist in decision-making machines, because the machines can produce outputs independently of their inputs and circumstances. Today the existence of a soul is sometimes questioned, but regard for the will of each individual is still strong, so the debate has moved on. Instead of focussing on free will we theorize about consciousness. Instead of the internal moral agent that chooses between good and evil, there is an internal pilot who sits at the controls of the brain when we are conscious, and is the real me. The unsolved problem is how to map the brain activity observable on MRI scans, the brain-space, to the thoughts and experiences of the conscious mind, the mind-space. Susan Blakemore says "The trouble lies with subjectivity, or 'what it is like' being me, now…my experience seems to be private, fleeting, ungraspable and utterly undeniable, and this is what we mean by consciousness". We are convinced that consciousness exists, yet unable to define it. Arguments about what it is to be human have moved from examining the soul to explaining consciousness. Perhaps, as Sigmund Freud wrote "Humanity has in the course of time had to endure three great outrages upon its naïve self love: the discovery that our world is not the centre of the celestial spheres, but rather a speck in a vast universe, the discovery that we were not specially created, but instead descended from animals, and the discovery that often our conscious minds do not control how we act, but instead tell us a story about our actions". In that awakening, the problem of choice without preference has helped in a small way towards all three discoveries, and this book is an account of its history and its influence.

Chapter 1. The rabbit in the headlights.

The pilot's dilemma

Late on Monday 1 July 2002, Captain Alexander Gros was flying a Tupolev 154 of Bashkirian Airlines from Moscow to Barcelona, carrying 52 children to a two-week holiday to the Mediterranean coast. The trip was the reward for the best students of a UNESCO affiliated school in Bashkortostan, and there had already been a delay when they were mistakenly sent to Domodedovo airport rather than Sheremetyevo 1 in Moscow, and they missed their flight. The tour agency booked them into a Moscow hotel and rescheduled the flight from the Saturday to the Monday, but they had finally got on their way to the Estival Park Hotel on the Costa Dorada.

At around 11:40pm the plane was approaching the border between Switzerland and Austria, near Lake Constance, at 36,000 feet when the Swiss air traffic controller noticed two blips converging on his radar screen. The Tu-154 was on collision course with a DHL Aviation Boeing 757 flying from Bahrain to Brussels.

Accidents do not usually have a single cause, and the Swiss control centre was not as alert as it might have been. It was a quiet night, so the short-term conflict alert system, which could have warned the controllers much earlier, had been taken out of service for routine maintenance. At the same time the Zurich centre's telephone lines were being worked on so the German air traffic controllers got a busy signal when they tried to alert their Swiss colleagues to the converging blips they could see on their own radar screens. Normally there are two air traffic controllers on duty at the same time, but one had just taken a coffee break, so one man was dealing with all the traffic. Hard pressed, he did not issue a conflict avoidance instruction to captain Gros until approximately 50 seconds before the planes would hit.

Both planes were equipped with an automatic collision avoidance system called TCAS, which can deal with potential collisions when human controllers are not available. It detects any other plane with a transponder within about 40 nautical miles, and can check whether they are on collision course. There is also a Resolution Advisor, which issues advice such as “climb” or “descend” so that one plane will pass safely over the other. To avoid a collision, one plane must be told to climb and the other to descend, so the systems in both planes have to be in communication to decide which goes up and which down. TCAS is very reliable, but not everyone trusts machines as much as they trust people, so when TCAS decided that Captain Gros should *climb*, and Paul Phillips in the DHL 757 should descend Captain Gros had a problem. The Swiss ATC had told him to *descend.*

Alexander Gros was 52 years old, experienced, and had logged 12,000 hours of flying time in 31 years to destinations such as Pakistan, Saudi Arabia and the UAE. He understood perfectly well what the air traffic control and the TCAS system were telling him, if he did nothing, he was about to collide with another plane, but should he go up or down?

He hesitated, and seconds later, the agitated Swiss air traffic controller called again. This time captain Gros acknowledged the instruction to descend, but still could not decide what to do. The TCAS hardware was said to be reliable, but in Russia machines are less reliable, and human controllers carry more authority. For 14 long seconds he agonized, as the other plane drew closer, and with only a few seconds left he decided to descend. The planes hit at 35,400 ft.

Witnesses said they heard a noise like thunder, and a fireball erupt in the sky over lake Constance. Then pieces of wreckage fell in and around the lake, and scattered fires were started. There were no casualties on the ground, but 71 people died in the two planes that collided

A police officer and his dog search the debris of a cargo Boeing 757 near Ueberlingen,

Captain Gros had decided to obey the man rather than the machine, and this cost him and his passengers their lives. It was the wrong decision, but he had had to do something. For both captains to ignore the warnings, and do nothing would have also resulted in disaster. Because they were on collision course and planes can't stop, there were only seconds to make the decision, but there was not enough information to make the decision. Should he obey the air traffic controller or the TCAS system? The machine is faster than the man, but the man may know things that the machine cannot know, for example, he may direct one plane to descend 1000 feet and the other to descend 2000 feet to avoid turbulence at a higher level. On the other hand humans are much more prone to error, so in weighing the evidence the decision maker can easily see a close balance between the alternatives, where it is not obvious which is the best.

What's the problem?

All decisions have to be taken by weighing the evidence, and then coming to a conclusion. That takes time, and it takes longer if there is very little to choose between the alternatives. The real problem comes when the choice has four characteristics:

1. The decision *has* to be made. If you don't do anything you'll end up in a worse state.
2. There are two (or more) distinct alternatives. You've got to choose one of them.
3. Each alternative looks just as good (or bad) as the other.
4. There is a limited amount of time in which to make the choice.

It's very possible that if you can't decide, and run out of time, you'll up in a worse position than if you had chosen any of the alternatives.

This is not a rare occurrence, it can also happen in a much less dramatic way in real life. Most people have had the experience of hurrying to get through a busy public place, where people are all trying to go about their own business. Sooner or later you find yourself heading straight for someone coming in the opposite direction. How to get past? To the left or to the right, both are equally possible. One person moves to the left in the hope that the other will move to their left, and the two can pass. This may happen, but equally often the second person moves to their right, and they are still on collision course. Both simultaneously realise their mistake, and move to the other side. But then they are still in each other's way and getting closer all the time. Most of the time this situation is resolved with only a little embarrassment, but occasionally they bump into each other and apologise. Not much damage is done if they do, and there is no need for them to bump anyway, because both can both slow down, or even stop while they sort out how to pass each other.

Because you can slow down and stop while you decide, this is slightly different from the two aircraft on collision course, but though you can avoid the chance of a collision, it's uncertain how long it will take to pass, and there's even the possibility that it might take minutes while you both dodge back and forth. It's only a problem if you *have* to decide. The two possibilities of a crash or a long wait can both be avoided if you don't have to decide. Someone else can decide for you. The rule of the road says that you must pass on the left in the UK, and on the right in the US. Everyone knows which side of the road to go to, so there's no decision, and no problem.

Aircraft can't stop, and can't even slow down much because they would stall and fall out of the sky, so the time available for making the decision of where to pass is limited, probably to a few seconds. The choices are climb, or descend, so a Rule of the Sky, that says always climb, doesn't help. Both planes climbing has the same result as neither. Taking the decision out of the hands of the captain does help, but only because it can be given to a computer which can take decisions faster than people.

But computers still have to make decisions and they still take time. So suppose the planes have two identical TCAS systems manufactured by the same company each with the same software. Both systems notice they are about to collide at the same time, and have to negotiate which is going up and which down. Both of them will go through exactly the same decision making process, and unless there is something different about the two systems, they both will make the same decision, say, to climb. Then each informs the other of the decision they have made, and because they both want to go up, there will still be a collision. The decisions clearly must be revised, but if both revise their own decision to “descend” they will still collide. This is actually worse than the human pilot, because now the collision is certain. There needs to be some small difference either in the systems, or the data supplied to the systems to arrive at diverging paths for the two planes. This different data might be the height of the two planes. If the height of both is not exactly the same, the higher one could be directed up, and the lower one down. Using an external measure like height difference to determine the outcome reduces the chance of collision but can never reduce it to zero, because there is always a chance that differences between the data supplied to both systems are very small, and it may take a lot of processing to find a difference.

Can it happen to me?

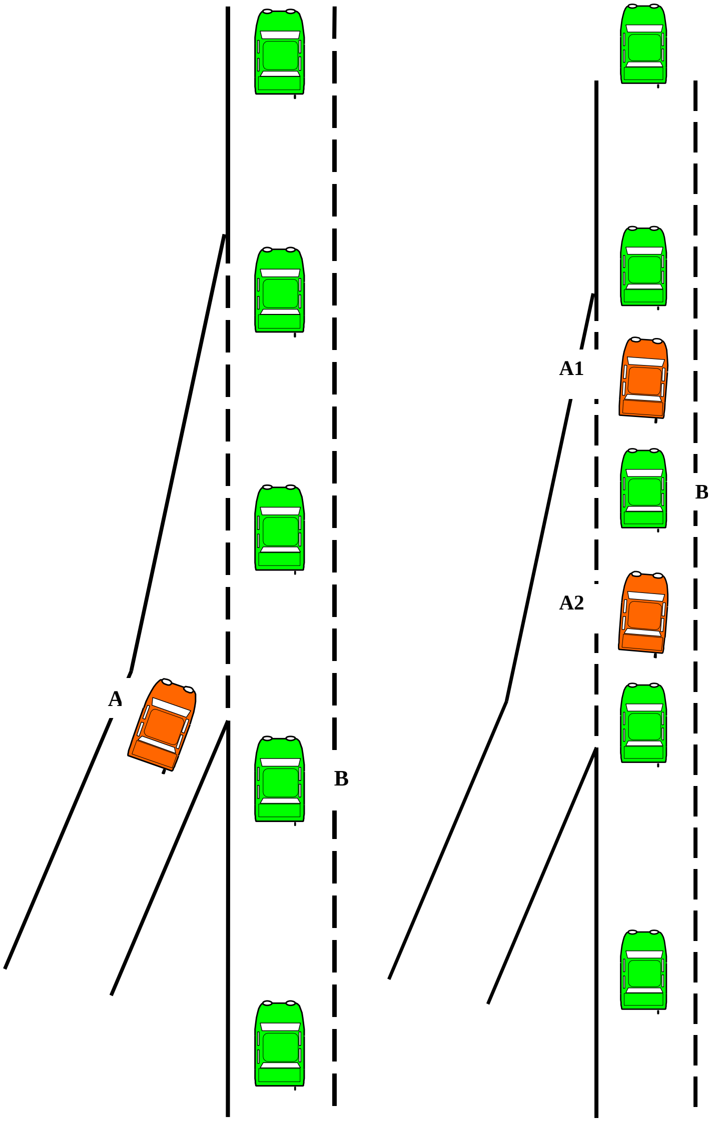
A variation of the two people passing in a crowd can be found in the story of the two polite friends. They can’t get both go through a door together, so one opens the door for the other. “After you Claude”, “No, after you, Cecil”…. If they are both polite enough neither of them gets through the door.

Perhaps it’s rude to push past a stranger in a doorway, but it’s not usually dangerous. A more worrying decision is the one you may have to make on the slip road to a major highway. It’s rush hour, the highway is busy, and cars are moving along at a constant speed. You come in on a slip road, and need to get on to the main road. The road designers have made sure that there is a short section of slip road where you run, more or less, parallel to the main road while you decide when to get on. You can't stop, or even slow down much, because otherwise you'd never get into one of the gaps between the traffic, and that means you'll run out of slip road after a while.

Now you have a decision to make. Do you speed up and go in the gap in front of the next car (A1), or slow down and go in the gap behind it (A2)?

You've got to decide. In this case there is more that one possibility, but that doesn't matter, you have to choose one of them. Any of them will do, but if you don't decide quickly enough, you run out of slip road.

What really matters here is the amount of time you can spend on the slip road making your decision. If the parallel section is very short, say only 50m, there will probably be at least one crash at the location every day. If it's over 250m, the probability of a crash is much lower, and there might not be any crashes for more than a year. The interesting thing is that there is *no* completely safe length that reduces the chance of an accident to zero.



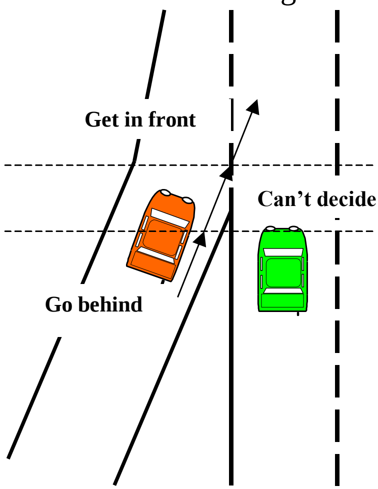
The slip road. Before or after?

Another dangerous arrangement is a Paternoster lift. This is a continuously moving chain of lift cabins intended to move large numbers of people in tall buildings. There is no lift door, just a cabin open at the front, which moves upwards, and is immediately followed by another, like the beads used to count off prayers. The chain of cabins moves fairly slowly, and you are supposed to jump on to the next one that appears at your floor, and then jump off when it has reached the next floor. At a university in the UK in the 1960's one was installed to move large numbers of students between lecture locations on different floors of the building. The chain of cabins went upwards for students going up, kept upright at the top of the building, and then went downwards on a parallel path for students wanting to go down. Because almost all the cabins can be full at the same time, it can carry more people than a more conventional lift, but there is a problem of choice. Which cabin? Try to get on the one that's just leaving, or wait for the next one? When one lecture finishes and the next is about to begin, students spill out of the lecture theatres and queue at the lifts, so there is usually a queue of students pushing from behind to get on. Each person feels the pressure to decide whether to get on to the cabin just leaving or wait for the next one, but the decision has to be made. After two students were killed using the Paternoster, it was replaced by conventional lifts that came when you called and could wait while you decided whether or not to get on.

Chapter 2. Choice without preference.

The way people think

Making a decision when there is absolutely no difference between the alternatives is difficult. It is a problem of choice but without any preference to point the decision one way or the other. The difficulty is that as the reasons for choosing between the alternatives get more and more evenly balanced, the decision takes longer and longer to make, until the point where there is no difference, and no decision can be made. In the real world choices have to be made, you have to decide which lift cabin to take, and given a plane hijacked by terrorists, you have to decide whether to shoot it down, or take the chance that they won't fly it into a nuclear power station after all. The decision has to be made, even if there is no preference between the choices. Since we make choices all the time it seems obvious that there must be a solution. Why not plan beforehand to take a particular one of the choices if you see no difference? That way you would waste no time deciding.



In front or behind?

This doesn’t work, because you still have to decide whether there is a difference between the two paths or not, and this turns out to be just as hard as deciding between the original two.

Take a practical example. If you are joining a main road, with a car already on it, then you would normally move in front of the car if you were already well ahead of it, and fall in behind if you were behind. It can take a long time to decide, if you are just level, or only a little bit in front, and that might lead to a crash. Why not make up your mind in advance to have a special ‘can’t decide’ zone where you always to drop behind?

Now instead of making up you mind whether you are level or not, you are going to have to recognise whether you are in this zone or not. How far ahead are you? Are you far enough ahead to move in front or not? It’s just as difficult to decide, as it was to decide whether to go in front or drop back in the first place.

What you’ve done is move the place where a decision has to be made, but not remove the choice. At the boundary where you have to decide whether to go forward or back the decision can still take forever.

Dividing up the choices into ‘in front’, ‘behind’, and the third one ‘don’t know so go behind’ doesn’t work, and is part of the natural tendency of people to try and simplify something, which can’t easily be simplified. Philosophers have wrestled with this difficulty for over a thousand years, as we shall see, but the process of understanding what making a choice really means has added to the sum human knowledge.

The history of Buridan's Ass

Discussions of choice without preference have a long history. The problem is often known as the paradox of Buridan's Ass, after Jean Buridan, the 14th century Rector of Paris University who is believed to have used it to illustrate his discussions of the works of Aristotle. In his example a dog, or an ass, is hungry, and is placed midway between two identical bales of hay. He assumes that there is no reason for the animal to prefer one or the other bale, yet it must eat one of them, or starve. The philosophical argument is about whether any reasonable agent, animal or human, can choose a course of action without a preference. On first sight it appears not. If there is no reason to choose one or the other, how can a choice be made? The paradox is that the result of not making a decision is that the ass starves to death, and this is such an extreme outcome that we cannot believe it could happen, so somehow, even in the absence of any preference a choice must have been made.

Its fascination for philosophers lies in the difficulties of the human mind in resolving this apparent paradox, and it has been used by the Greek philosophers to argue for the position of the earth at the centre of the cosmos, by the Arabs to argue whether the world is eternal, or created by God, and even to discuss the nature of God's will. If a man has to make such an impossible choice, does he have the free will necessary to make it? In Buridan's time people drew a distinction between the ability of animals and the ability of men to exercise free will. Buridan found it difficult to show any difference between the freedom of our will, and the lack of freedom of a dog when placed between two bowls of food.

Philosophy and science are not a high priority when people are threatened by famine, disease and war. It is at times of relative stability and plenty that philosophers are able to find the time to meet together, debate, classify knowledge and to record it for the following generations. The Greeks knew the Paradox of Buridan's ass, and an illustration based on it appears in one of Aristotle's works, *De caelo*, - on the heavens.

Aristotle

Aristotle was the recorder of the knowledge of the golden age of classical Greece He was the son of the personal physician to Amyntas III, king of Macedonia and became friendly with Philip, king Amyntas's son, who was almost exactly the same age. When Aristotle was about ten years old his father died, so Aristotle could not follow his father as a doctor and, since his mother seems also to have died young, Aristotle was brought up by his uncle. In 367 BC, at the age of seventeen he went to Athens to finish his education, as he wanted to study with the scholars at Plato's Academy. It is said that when he finally found Plato, he was told that if he wanted to find the Academy he should walk Northwest through the Diplyon gate, follow the river for a mile or so, and there he would see a garden with a high wall. "If you hear students making great talk beneath the trees, apply your uncouth ear, and if you cannot understand a word of it you have found the right place." Aristotle became a student, and then a teacher at the Academy, and he was to remain there for twenty years.

While Athens was the centre of art and philosophy, Aristotle was a collector, teacher, and organiser of knowledge rather than a philosopher in his own right, and wrote on physics, metaphysics, and psychology, as well as describing the heavens, in a work known to us as *De Caelo*. In it he discusses the position of the earth in terms of its *natural place*. He is arguing against Socrates who believed the earth must be in the centre of the universe simply to maintain symmetry.

"The reason for the Earth's position is not its impartial relation to the extremes; that could be shared by any other element, but motion towards the centre is peculiar to the earth... If… the place where the earth rests is not its natural place, but the cause of its remaining there is the constraint of its 'indifference' (on the analogy of the hair which, stretched strongly but evenly at every point, will not break, or the man who is violently but equally hungry and thirsty, and stands at an equal distance from food and drink, and who therefore must remain where he is)"

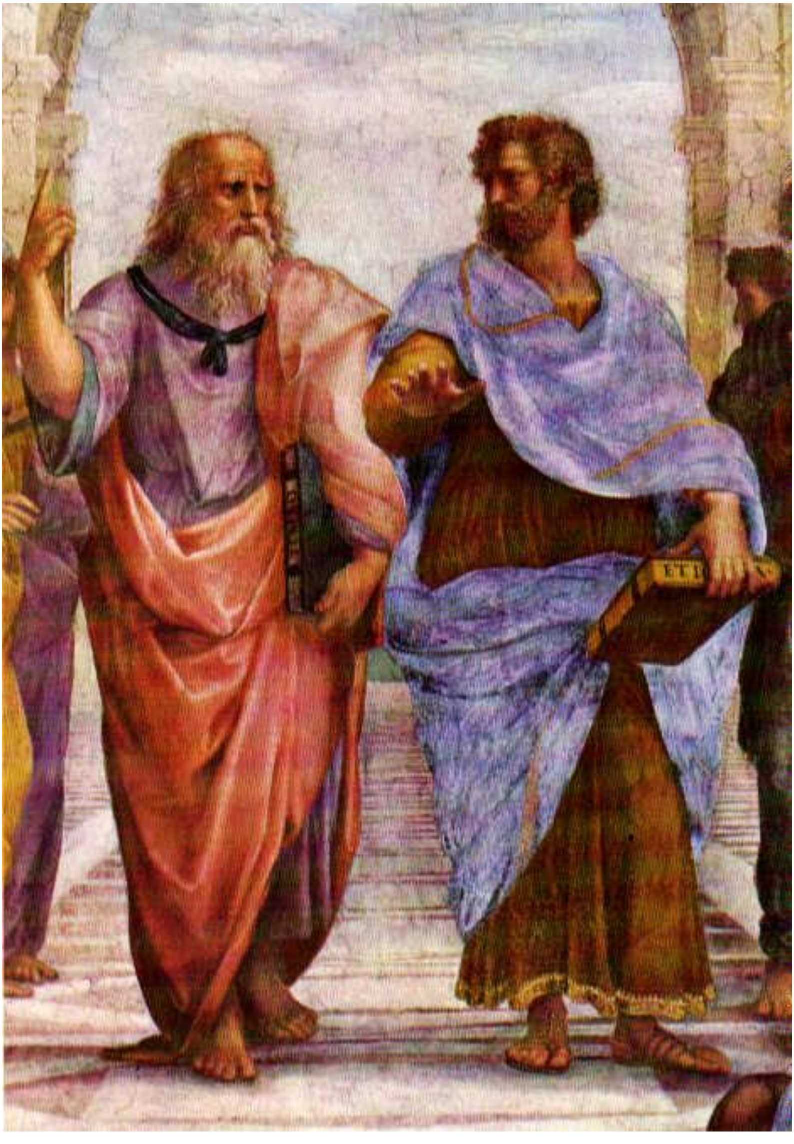
The belief that the earth is at the centre of the universe was held for nearly another 2000 years, but is supported here by only the idea that it must be its 'natural place', because we know that otherwise it would move one way or the other, in the same way that we might know a man will eventually satisfy his hunger or thirst. The example of the man torn between the equal attractions of food and drink and consequently unable to move must have been well known at the time.

Towards the end of Aristotle's twenty years at the Academy his position became difficult due to the political events of the time. Amyntas, the king of Macedonia, had died around 369 BC, a couple of years before Aristotle went to Athens to join the Academy. Amyntas's third son, Philip II came to the throne and showed himself to be a skilful politician as well as a soldier, expanding his power into northern Greece. This threatened Athens, and consequently there were suspicions about people with Macedonian connections, for example, Aristotle's friendship with Philip. When Plato died in 347 BC the leadership of the Academy was vacant, but Speusippus was elected, not Aristotle. Aristotle was opposed to the views of Speusippus and it seemed sensible to leave because of his unpopular Macedonian links.

Moving to Assos he began to develop a philosophy distinct from that of Plato who had said the kings should be philosophers and philosophers, kings. He wrote that

"It is not merely unnecessary for a king to be a philosopher, but even a disadvantage. Rather a king should take the advice of true philosophers. Then he would fill his reign with good deeds, not with good words".

Politics again intervened. The Persians attacked Assos, but Aristotle escaped to Macedonia. In 343 BC, remaining there for seven years. Philip was now at the height of his power but power was disputed by Alexander, Philip's son who with Aristotle's support now became king. Returning the favour, Alexander helped Aristotle to found a rival to the Academy in Athens, called the Lyceum.



Aristotle and Plato, Aristotle's hand is level to the Earth symbolizing his realist view of Nature; Plato's hand pointed towards the heaven symbolizing the mystical nature to his view of the Universe.

The collections of writings that have come down to us were not published by Aristotle himself in his lifetime. Aristotle probably never intended them to be published, as they are almost certainly the lecture notes from the courses given at the Lyceum

One of the most important of Aristotle's legacies was the idea of arguing in a series of logical steps from an initial assumption towards a conclusion that you would like to demonstrate to be true. Aristotle believed that logic was not a science but rather had to be treated before the study of every branch of knowledge. Aristotle's name for logic was "analytics". He believed that logic must be applied to the sciences, a view that helped bring rigour to scientific arguments, but also put abstract argument before the empirical facts. His work on logic also contains an appendix on 'sophisticated argument' or the pitfalls in arguments where one might successfully deceive one's opponent, or perhaps, oneself.

Aristotle was an advisor and friend to Alexander the Great, King of Macedon, for 12 years. "To my father," said Alexander, "I owe my life; to Aristotle how to live worthily." At the beginning of Alexander's reign the Persian empire covered all of the middle east, from Egypt in the south, to Turkey in the North, and Afghanistan in the east. In a series of battles, Alexander defeated King Darius of Persia and razed its capital, Persepolis, to the ground. His own capital, Alexandria, was to be built in Egypt, a seaport with a lighthouse that was one of the wonders of the world. After his death in 323 BC Ptolemy, the king of Egypt, created a library of over 400,000 manuscripts in the worlds first museum, to contain the works of Aristotle and Plato, but anti-Macedonian feeling in Athens was such that Aristotle retired to his former home in Northern Greece. He died the following year from a stomach complaint at the age of 62. Some say that as a man he was kindly, and affectionate with not much self-importance, though admirable rather than amiable. His will, makes references to his happy family life and takes care of his children, as well as his servants. During his life he wore rings on his fingers and cut his hair fashionably short. His detractors say he was spindle-shanked and overbearing.

The museum and library in Alexandria stood for nearly a thousand years but the Hellenistic monarchs of Macedon were defeated by the Romans in 197 BC. Eventually Alexandria was absorbed into the Roman Empire as the capital of the province of Aegyptus. The Roman priorities were triumphs, and colonization, plumbing and circuses, rather than knowledge and philosophy, so while the spirit of scientific enquiry did not die, science was not developed much further.

The end of civilization

Climate change, with consequent plague disrupted the classical word in the 5th Century. Rome was by that time in decline and in no position to resist the invasions of Germanic tribes who were themselves attacked by the Huns. China and Europe descended into the chaos, of war, and famine forced people to look to their own survival, and consider what sins they had committed to bring ruin on to themselves.
In Britain, according to Bede;

“The fires kindled by the pagans proved to be God’s just punishment on the sins of the nation, just as the fires kindled by the Chaldeans destroyed the walls and buildings of

Jerusalem. For, as the just Judge ordained, these heathen conquerors devastated the surrounding cities and countryside, extended the conflagration from the eastern to the western shores without opposition and established a stranglehold over nearly all the doomed island. Public and private buildings were razed; priests were slain at the altar; bishops and people alike, regardless of rank, were destroyed with fire and sword, and none remained to bury those who had suffered a cruel death. A few wretched survivors captured in the hills were butchered wholesale, and others, desperate with hunger, came out and surrendered to the enemy for food, although they were doomed to lifelong slavery."

People turned to religion for support. If these evils were the punishment of God for the sins of the people, then perhaps strict observation of the scripture would bring salvation. Christianity gained many converts in Europe, and Islam swept across Alexander's old empire.

The western empire based on Rome fell in 410 AD, but the effects of the barbarian invasions were not so severe in the east. By 814, Baghdad had grown to be the second city in the Eurasian world, exceeded only by Ch'ang-an in China. The library at Alexandria survived, and its knowledge was available to the Arabs.

The shift of concern from material to spiritual matters meant that the interests of the great Arabic philosophers were different. Instead of trying to provide a logical framework for the arrangement of the heavens, they were interested in providing support for their Muslim faith, and when the Paradox of the ass next appears it is used by the great Arab philosopher, Al Ghazali, to defend the idea that the world was created by the Will of God.

The Arabs

Abu Hamid Ibn Muhammad Ibn Muhammad al-Tusi al-Shafi'i al-Ghazali was born in 1058 AD, in what is now Khorasan, Iran (then called Tus). At the age of 34, however, he became rector and professor of the Nizamiya madrassa (muslim academy), located in Baghdad. It was during this period that he produced two of his most important texts, "The Intentions of the Philosophers," and its companion piece, "The Incoherence of the Philosophers." At this point, Al-Ghazali had come to the conclusion that Aristotle's logic did not work for him, and left his position, to make pilgrimages, spending the next ten years in or between the cities of Damascus, Mecca, Medina, and Tus (his birthplace). He died in 1111 AD.

Al-Ghazali understood the Aristotelian methodology, as did his contemporaries, known as the Arab Aristotelians, al-Farabi (Alfarabi) and Ibn Sina (Avicenna), but he did not believe, as they had argued, that Islam (and religious belief in general) could be proved or disproved by logic. He set out to show, in "The Incoherence of the Philosophers" that the effort to do so resulted in little more than incoherent pseudo-justification for belief.

The argument of the Aristotelians was that if the world had been created by God, there must have been a time when it did not exist, and a time when it did. God had a choice, to create, or not to create the world, but no reason to choose one or the other, and since there is no reason to make a choice, the choice cannot be made. The world is here, so it must have always been here. The religious would answer this by saying, "It is the Will of God, which is eternal and unknowable, that determines the moment of creation". The Aristotelians then reply that the Will of God is not required if the world is eternal.

This is blasphemy, and must be countered. At that time Islam was relatively tolerant, other religions might have burnt

Aristotle's books, if not the Aristotelians themselves. It didn't happen, though it must have been a close thing.
Ghazali's response contains the first clear expression of the difficulty of choice without preference, and, in fact the paradox should probably be known as the paradox of Al Ghazali's dates. He has to show that there is some way of distinguishing between two apparently equal courses of action, so that the Will of God can be admitted.

"Suppose two similar dates in front of a man who has a strong desire for them, but who is unable to take them both. Surely he will take one of them through a quality in him, the nature of which is to differentiate between two similar things. All the distinguishing qualities, like beauty or nearness or facility in taking, we can assume to be absent, but still the possibility of the taking remains. You can choose between two answers: either you say that an equivalence in respect of his desire cannot be imagined – but this is a silly answer, for we assume it is indeed possible – or you say that if an equivalence is assumed, the man will remain for ever hungry and perplexed, looking at the dates without taking one of them, and without the power to choose or to will, distinct from his desire. And this is one of those absurdities which are recognized by the necessity of thought. Everyone, therefore, who studies, in the human and the divine, the real working of the act of choice, must necessarily admit a quality, the nature of which is to differentiate between two similar things."

There we have the first use of choice without preference as a way of settling an argument, on the one side, the belief that choice is possible without any reason demonstrates the existence of free will, and on the other the denial of any concept of will without reason. The argument continued, with some agreeing that choice is possible, but only because there will always be a reason or a difference of some kind, and others believing that in the complete absence of any difference between the dates the man will have to starve.

Concerned to demonstrate that their point of view was superior to Aristotle's or that Aristotle could be used to show that the others were wrong several commentaries on his works were written. One of the best known is by Ibn Rushd, known as Averroes in the West. Averroes was born in Cordoba in approximately 1128 AD. At that time Cordoba was the largest city west of Constantinople, but was passing its peak as the centre of the Muslim Caliphate of Cordoba, and the Christians of the North had started the re-conquest of Spain. He was born into a well-educated family of judges and was educated in religious law (Islamic scripture and Hadith), medicine, science, philosophy, and mathematics, leaving his homeland for Morocco at the age of 27. It was there that between 1169 and 1195 Averroes completed his translation of Aristotle as well as writing commentaries on both Aristotle and Plato. He would later be known by the title "the Great Commentator".

Books by commentators like Averroes found their way north to a Europe emerging from chaos and war. Flickerings of light had from time to time illuminated the dark ages in Northern Europe. A library was established in Northumberland in 680, with 'A great mass of books' but places of learning were later looted and burned out by Viking and Magyar raids and invasions between the 8^{th} and 10^{th} centuries. Eventually the Vikings settled, new land was opened up and by the 11^{th} century there was, maybe not prosperity, but at least stability. Up until the 8^{th} century Islam had been on the offensive, conquering most of Spain and attacking France. Defence was the first priority in Europe. The subsequent economic recovery meant two things, the Christians had the resources to go on the offensive, pushing Islam back in Spain, and carrying the war to the Holy land by means of crusades. Then the increase in wealth and the absence of war in Northern Europe enabled the development of Cathedral schools into universities.

Chapter 3. The will of God.

The first steps towards a better understanding of the natural world, as opposed to the spiritual world were begun in Europe in the 13th century. To a limited extent, ideas like the meaning of choice could be debated in the logical framework inherited from Aristotle.

The medieval universities

Between 1150 and 1450 universities sprung up all over Western Europe, both Paris and Oxford appearing some time before 1200. Aristotle's works were translated into Latin, as well as the commentaries of the Arabs, and his ideas of logical argument and reason were used to explore the nature of the universe. This did not go without opposition from the Bishops who saw that the whole system of philosophy had been created without reference to Christian beliefs. In 1210 notes and lectures were burnt, and any public or private study of Aristotle was completely forbidden because not only did his works suggest that God was to be found throughout creation, and not outside it, but there was even a suggestion that the world could not have been created from nothing.

Later, in 1231, and 1263, the ban was reiterated but now only 'until the works on natural philosophy have been examined and found free of error'. This applied only to Paris University, but had the effect of both stimulating interest, and raising an argument that to oppose such error, it was necessary to have studied Aristotle thoroughly. In 1366 the Pope gave in, and Aristotle became obligatory for the study of Master of Philosophy.

The ideas of the Greeks were embraced enthusiastically, but not uncritically. One of the major issues was Plato's theory of forms, which holds that the world we see is not reality, but only the shadow of the pure forms that exist in the realm of ideas. For everything we see, there is a perfect form, every individual person is an example of the species of man, and everything must be classified in systems that show how these things are related.

Petri hyspani.

Sequitur arbor Porphiriana.

‘Porphyry’s Tree’

The kind of classification used is illustrated in ‘Porphyry’s Tree’, which shows how, Man, was seen. The roots of the tree consist of individuals, Socrates (Sortes), Johannes, Henricus, Nicolas, Petrus, and Plato, which are in reality outside the system, since individuals are only shadows and cannot be defined. Homo is a species, and together with another species, Beast, forms a Genus called Animal. What distinguishes these species are specific differences, in this case reason for Homo, and lack of reason for Beast. Such systems of classification can take over, becoming a substitute for rational thought, and the followers of Aristotle disagreed with those of Plato, dividing philosophers into two camps. The argument was about the reality of Universals such as Homo, Animal, organism, etc. Did they really exist?

The Realists, like Thomas Aquinas, thought they did, but the opposing camp, the Nominalists, which included William of Occam, thought them mainly concepts of the intellect, just names.

The disputation

One of the most important parts of an education at a medieval university was the obligation on a student to discuss his work with others in a formal disputation. A master would introduce the topic for discussion, and sum up the points, for and against, a particular conclusion, at the end. Buridans writings include discussions of such topics, which we can imagine as forming the basis for disputations.

It is the evening of a disputation day at Paris. The time has been posted, and the students assembled. A famous Master, John Buridan, known for his Nominalist views, and an advocate of the modern way of thinking enters.

Master: It is supposed that, in return for services you have rendered to me, that I promise to give you a horse, and I

go before a judge and solemnly swear that I will deliver a horse to you before Easter.
But I, being a clever sophist, argue as follows: If no horse can be found that is the one that I owe to you, then I do not owe you a horse. That there is no such horse, I can establish by questioning you on the witness stand. Is the horse Morellus such that I owe it to you?

Student: Yes.

Master: But since my debt can be paid by delivering Favellus, Brunellus, or some other horse, then it cannot be said that Morellus is owed to you.

Student: Perhaps

Master: Is the horse Favellus such that I owe it to you?

Student: By the same reasoning it cannot be said that Favellus is owed to you.

Master: Is it necessary that I run through all the horses that there are? For if I can establish that there is one horse which is not such that I owe it to you, I can establish it for every horse. Hence it follows that no horse is owed to you.

Student: I might agree you cannot deliver the concept of a horse, or some abstract universal horse, but there still must be a horse such that you owe it to me.

Master: Then if I owe you a horse, every horse is such that I owe it to you. The sentence "some horse is owe to you" is equivalent to "Morellus is owed to you, or Favellus is owed to you, or Brunellus is owed to you…". If at least one of these is true, then some horse is owed to you; but if all are false, then no horse is owed to you. There is no more reason that any one of them should be true rather than another, or that one should be true rather than another. Therefore if some horse is owed to you, every horse is owed to you; and if some horse is not owed to you, then no horse is owed to you.

Student: You did not promise me every horse, and therefore you do not owe me every horse.

Master: I willingly concede that this is so. Although every horse is such that I owe it to you, and giving it to you would absolve my debt, it does not follow that I owe you every horse. Nor does it follow from the fact that Morellus is owed to you that I owe you Morellus. But as the arguments showed, if I owed you a horse, then everything that is a horse is such that I owe it to you.

Here the impossibility of choosing one specific horse out of several equally desirable individuals is used to show that the Realist's abstract universal horse cannot be given to settle the debt, but for the nominalist, any and every horse will do.

John Buridan and the early renaissance

John Buridan was born sometime before 1300 in Picardy. He was bright enough awarded a benefice or stipend for needy students, and went to Paris to study. Receiving his Master of Arts degree and formal license to teach in the University of Paris by the mid-1320s, he rose to the position of rector first in 1328 and then again in 1340. The job of the rector was to the chair meetings between faculties, and to be the university's chief spokesman. To be elected to such a post means that he was seen as capable of reconciling the positions of apparently irreconcilable and opinionated colleagues to make an apparently coherent university policy. To be elected a second time means that those people had come to believe that the university policy was theirs rather than his. Not only clever, but sophisticated and urbane, such was his fame that there is a story that the King of France had him thrown into the Seine River in a sack because of a scandalous affair with the Queen. Another says that he hit the future Pope Clement VI over the head with a shoe while competing for the affections of the wife of a German shoemaker.

So far as Paris and Oxford Universities were concerned, the faculty of theology was the queen, since it required a special licence from the pope to be taught. But Buridan spent his entire career in the faculty of arts at Paris, without ever moving on to study for a doctoral degree in one of the higher faculties of law, medicine, or theology. Since university statutes forbade arts masters from teaching or writing about theology, Buridan produced no theological works, only those on logic, and natural science, but he avoided conflict with religion by distinguishing between conclusions, which are demonstrated by logic, and those, which are demonstrated by experiment. He believed that “The principles of natural science are not immediately evident, indeed we may be in doubt about them for a long time…but they are accepted because they have been observed to be true in many instances, and false in none.”
Dogmatists might argue that no one could know whether the laws derived from such evidence were the result of natural causes, or whether God was producing the effect. For them, only the absolute certainty of logic could produce scientific laws. Buridan disagreed “ Very evil things are being said by those who seek to undermine the natural or moral sciences because absolute evidence is not possessed by most of their

principles and conclusions, it being supernaturally possible for them to be rendered false. For in these sciences absolutely unconditional evidence is not required." His concept of science is one of seeking the truth rather than having certainty, but it also leaves room for God to overturn scientific laws, and avoids direct conflict with religion. All the same it is probably as much due to his personal and intellectual qualities that Buridan escaped persecution by the church, where others did not.

Freedom of the Will

The paradox of Buridan's ass does not explicitly appear in any of his writings, it was probably used in his lectures as an example to illustrate one of the issues of the time, such as the question of Freedom of the Will. The debate about free will was central to the issue of morality and religion. Given a moral choice between a good act and an evil one, do we slavishly follow our intellect, or do we have some other additional way of choosing called the will, which is the seat of morality? This question is linked to that of the Will of God. How does God choose whether a child lives or dies? Does God know with certainty and eternally what men will do? If everything is pre-ordained, there is no freedom of the will in men, and no point in behaving piously.

William of Occam's view was that God did not dictate every event, and that therefore the will of men was free to do good or evil.

Since God does not make choices for him, a man must use his intellect and his will, but how does this work when the choice appears to be equal? Should he give bread to the hungry beggar, or first tend to the sick? Buridan's answer is that the will does not decide from within its own resources, it is subject to reason. As reason suggests, the will must follow. Should two courses be judged equal, then the will cannot break the deadlock, all it can do is to suspend judgment until the circumstances change, and the right course of action is clear. He then says that if free will exists in men, it is hard to deny it to animals. 'For it would be difficult indeed to show that when our will is indifferent between to opposed acts, it could decide for one or the other without an external factor, where a dog could not.'

He left an idea of science in which the laws were simply the best description of observable facts, and the development of logic into a mathematical tool. These were not the only achievements of the 14th century. Buridan went on to provide the basis on which both Galileo and Newton built the laws of motion.

Aristotle had described the motion of a stone thrown upwards in the air as contrary to its natural movement towards the earth. In Aristotelian dynamic theory, this required a moving cause continuously in contact with it to keep it going after it had left the hand of the thrower, which, he supposed, came from the only thing in contact with it, the air. This is obviously rather feeble, and the Arab commentators were aware that something better was needed.

Raising the question, Buridan suggests that the stone has acquired a power of motion, which he calls impetus. "This impetus", he says, "would endure for ever if it were not diminished or corrupted by an opposed resistance or something tending to an opposed motion." He then goes on to consider impetus to be a function of mass and velocity in exactly the same way that Newton defined momentum, mv. Discussing how gravity affects a body in free fall, he says that the force of gravity causes it to gain successive increments of impetus during the fall. Since the mass does not change when impetus increases, the velocity does, and from this we can deduce that force equals mass times acceleration. This is a real break from the past. The idea that the natural state of a body is to be at rest, has been replaced by the understanding that the natural state is one of continuation. What follows from Buridan's suggestion is Galileo's observation that since the force due to gravity is proportional to the mass of an object, all objects falling form a fixed height hit the ground at the same time. The building blocks were in place for the laws of motion that we have today.

The Black Death

Though well known at the time as a very distinguished man and a celebrated philosopher, he was remembered later only for Buridan's ass. That we do not recognise him as the Renaissance man he was, is probably due to the events of 1348 and their aftermath. What follows is drawn from the chronicle of Jean de Venette, a Carmelite friar, who probably died in 1369.

"In 1348 the people of France, and of virtually the whole world, were assailed by something more than war. For famine had befallen them, and then war, so now pestilences broke out. As a result of that pestilence a great many men and women died that year and the next in Paris and throughout the kingdom of France, as they also did in other parts of the world. The young were more likely to die than the elderly, and did so in such numbers that burials could hardly keep pace. Those who fell ill lasted little more than two or three days, but died suddenly, as if in the midst of health - for someone who was healthy one day could be dead and buried the next. Lumps suddenly erupted in their armpits or groin, and their appearance was an infallible sign of death. Doctors called this sickness or pestilence an epidemic. Such an enormous number of people died in 1348 and 1349 that nothing like it has been heard or seen or read about. And death and sickness came by contact with others and consequent contagion, for a healthy person who visited the sick hardly ever escaped death. In many towns and villages the result was that the cowardly priests took themselves off, leaving the performance of spiritual offices to the regular clergy, who tended to be more courageous. In many places not two men remained alive out of twenty. The mortality was so great that, for a considerable period. More than 500 bodies a day were being taken in carts from the Hotel-Dieu in Paris for burial in the cemetery of the Holy Innocents. The saintly sisters of the Hotel-Dieu, not fearing death, worked sweetly and with great humility, setting aside considerations of earthly dignity. A great number of the sisters were called to a new life by death and now rest, it is piously believed, with Christ.

It is said that this mortality began among the infidel and then travelled to Italy. Afterwards it crossed the mountains and arrived in Avignon, where it attacked various cardinals and carried off their entire households. Then it gradually advanced through Gascony and Spain and into France, advancing town by town, street by street, and finally from house to house - or, rather, person to person.

Men ascribed the pestilence to infected air or water, because there was no famine or lack of food at that time but on the contrary, a great abundance. One result of this interpretation was that the infection, and the sudden death which it brought, were blamed on the Jews, who were said to have poisoned wells and rivers and corrupted the air. Accordingly the whole world brutally rose against them, and in Germany, and in other countries, which had Jewish communities, many thousands were indiscriminately butchered, slaughtered and burnt alive by the Christians. The insane constancy shown by them and their wives was amazing. When Jews were being burnt mothers would throw their own children into the flames rather than risk them being baptised, and would then hurl themselves into the fire after them, to burn with their husbands and children.

It was claimed that many wicked Christians were discovered poisoning wells in a similar fashion. But in truth, such poisonings, even if they really happened, could not have been solely responsible for so great a plague or killed so many people. There must have been some other cause such as, for instance, the will of God, or corrupt humours and the badness of air and earth. The mortality continued in France for most of 1348 and 1349 and then stopped, leaving many villages and many town houses virtually empty, stripped of their inhabitants. Then many houses fell quickly into ruin, including numerous houses in Paris, although the damage there was less than in many places."

Victims of the plague being buried in coffins.

Following such a catastrophe the world is different. Learning is less important than surviving, and people take their solace from simpler things.

“When the epidemic was over the men and women still alive married each other. Everywhere women conceived more readily than usual. None proved barren, on the contrary, there were pregnant women wherever you looked. Several gave birth to twins, and some to living triplets. The world, alas, has not been made any better by its renewal. After the plague men became more miserly and grasping, although many owned more than they had before. Also few men could be found in houses, towns or castles who were able or willing to instruct boys in the rudiments of Latin.”

The aftermath of the plague was a setback, but not a permanent one. William of Occam died in 1349, of the plague, and Jean Buridan around 1358 probably in a second outbreak. The resumption of the renaissance lay 100 years away, and Galileo’s rediscovery and demonstration of the laws of motion 250 years away. When it came it was again condemned by the church as heretical.

Chapter 4 As long as it takes

The mechanics of decisions

In the Middle Ages decisions were made by God and by men. Beasts were not capable of reason, and still less were mechanical devices. Only in the 19th and 20th centuries could the idea of a chess playing engine be at all credible, and only in the 20th could it actually be made. To build such systems requires not just a detailed understanding of the laws of physics, but also of the nature of a decision.

Exactly what is a decision? You must weigh the evidence, and then come to a conclusion. But the evidence and the conclusion may have an essentially different character. The evidence is often a measurement, and the conclusion one of a limited number of alternatives. They can both be expressed as numbers, for example a measurement might be 30.3mm or alternatives might be option 1, option 2 etc. as numbers they look much the same, but this is deceptive. A quantity like 30.3 mm is the nearest we can get to a real distance. It might not be exactly that, and if we had a more accurate measuring device we might have got 30.3138 mm. On the other hand, option 1 is always option 1 and not option 2. It doesn't depend on accuracy of measurement. This was only fully understood at the end of the 19th century.

Even if you know what you have to do, there's still the question of how to do it, and all engineers need a sound set of laws on which to base their dreams.

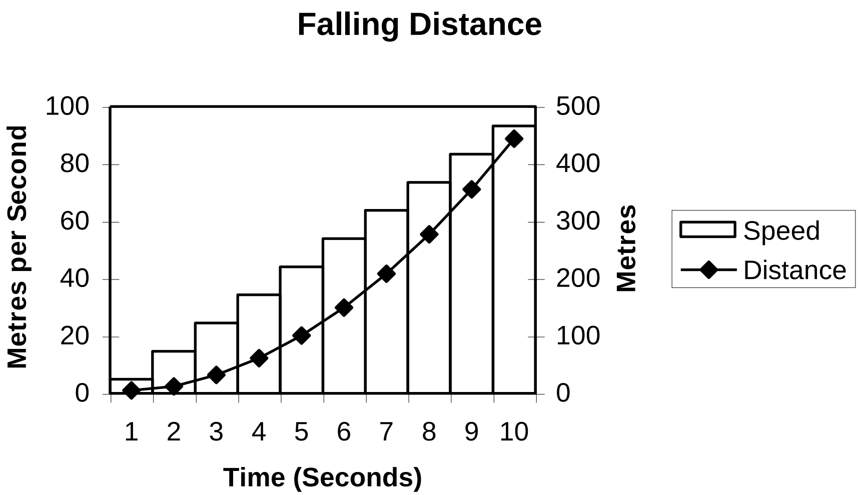
If you have to make a mechanical decision maker you need first to understand the physics of motion properly, and the first person to do that was Galileo.

The laws of motion

How much Galileo Galilei (1564-1642) was influenced by John Buridan's ideas of motion is not known, but a version of Buridan's ideas was commonly taught in the universities of the time. Buridan had understood that unless a force acts on a moving object, its inertia will cause it to continue in a straight line and that a body falling under gravity accelerates uniformly, but there was no convincing theory, still less any practical demonstration of its truth. Galileo is famous for his support of view that the earth is not the centre of the solar system, as Aristotle had said, but that the motion of the planets was much more simply explained by Copernicus. Copernicus had suggested that the sun was at the centre and the earth was spinning on its axis. To the average man, this is obvious nonsense. The earth could not possibly rotate once a day, otherwise the clouds would forever be disappearing over the western horizon, but Galileo countered with the example of a horseman who throws a ball straight upwards in the air. Even though he is galloping along the ball can still be caught as it comes down, because it continues to move along as well. Thus the clouds move with the earth. Using observations of Venus through a telescope he confirmed the truth of Copernicus' hypothesis. After publishing his work, Galileo was summoned by the Holy office to Rome The renaissance had seen the acceptance of classical thought, and the Church was now not only reconciled to Aristotle's views, but regarded them as essential doctrine, with the likes of Buridan just adulterers of the pure vision of Aristotle. Furthermore the Bible was quite clear that Joshua had asked god to stop the sun over Gibeon, if the sun did not move, that could not have been true. The tribunal passed a sentence condemning him as a heretic, and compelled him to solemnly abjure his theories.

Newton, who was not subject to the inquisition, could publish theories about the motion of the planets. The laws of motion themselves, as Newton described them, are very simple:

1. Every object in a state of uniform motion tends to remain in that state of motion unless an external force is applied to it.
2. The relationship between an object's mass *m*, its acceleration a, and the applied force *F* is $F = ma$.
3. For every action there is an equal and opposite reaction



The story about Galileo, that he used the tower of Pisa to drop two objects of different weight, is probably apocryphal; in fact, he used inclined planes to demonstrate that the time taken to fall from the same height is independent of the weight. It's easy to show that from the laws of motion. Because there is a constant force due to gravity, an object in free fall is accelerated at a constant rate. So if during the first second after it starts, it averages 5 metres per second, over the next second it averages 15 meters per second, then 25, then 35 and so on. The graph above shows how the speed (left hand scale) increases linearly with time, and the distance it travels in a given time (right hand scale) can be found from adding up all the distances travelled at each speed.

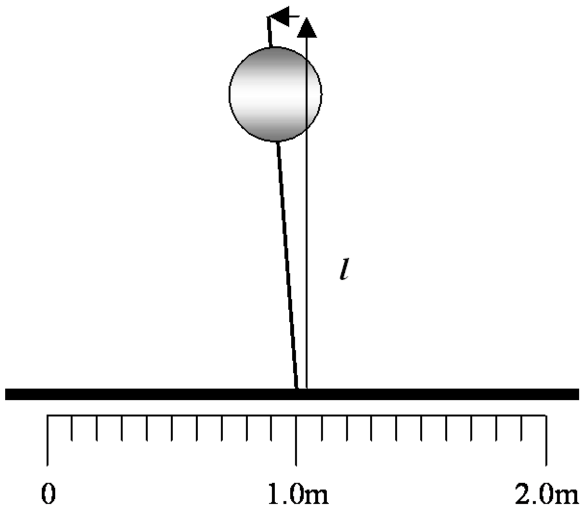
The total distance *l*, a heavy object has fallen after *t* seconds from a tower, is $4.9.t^2$. The time and the distance are linked together, so that for every leaning tower with a fixed height there can be only one time of fall, no matter how much the object you drop weighs. In one second objects fall 4.9 metres, in two, they fall 12.3 metres, and in 3 seconds 31.9 meters. If there is no external force, such as that provided by air resistance, a one metre drop always takes 0.45 s. Anyone could go to the tower and drop weights, which are heavy enough to make the air resistance negligible, and show they always hit the ground at the same time.

Time and the pendulum

Pendulums make use of much the same idea, if you pull the bob on a metre long pendulum to one side; it takes about 0.5 seconds to fall to the centre. In this case it doesn't matter how far to one side you pull it, the further you pull, the faster it goes when you let go, but it always takes the same time to get to the centre. Then when it has reached the centre, the bob keeps going, as Newton says it will, in his first law because there's nothing to stop it. It carries on until it has reached the same height on the other side as it started from, which, of course, takes the same time as it did to get to the centre in the first place. That's why pendulums used to be used in clocks, you can make them out of almost any material, but the time the bob takes to go from one side to the other is constant and depends only on the length.

To make a decision maker, that starts with a small input, and decides whether it is one thing or another, you could do worse than use an inverted pendulum. The small difference between the top dead centre and the actual starting position of the pendulum is the input, and the final outcome when you let go, either it hits the ground on the left, or the right, is the decision. This simple apparatus tells you whether it started on the right, or the left of top dead centre, no matter how close to the centre it starts.

The way it works is that a small movement to one side causes a force from gravity pulling it *away*, rather than *towards* the centre, and as it moves away it picks up more momentum, and goes faster towards the side it was pushed to in the first place, eventually, or in fact, rather quickly, falling to the ground on that side.



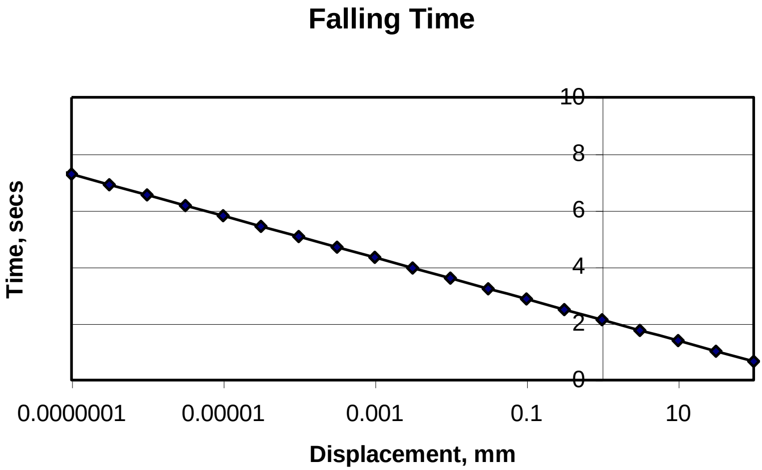
Inverted Pendulum

Like Buridan's ass, a very small bias towards one side or the other causes it to fall on that side, like a donkey moved slightly towards one bundle of hay will eat that one first. Old fashioned scales work on this principle, telling you if the goods in one scale pan are lighter or heavier than the weights in the other.

So it looks as if the problem is solved, but a closer look at the laws of motion suggest there might still be a difficulty.

If you displace the pendulum bob by a small distance, the force of gravity acts to pull it straight down. But the rod supporting it is at an angle to gravity, so there is a small net force pulling it further away. The further away the bob is from dead centre, the bigger the force. So the bob accelerates away from the centre at a faster and faster rate. This is not like gravity, where the rate of acceleration is constant; here the acceleration gets more and more the further away the bob gets from the centre.

Something that increases faster and faster as it gets bigger is called an exponential; it is the natural law of growth. The bigger it is, the bigger it grows, like a credit card debt. If you don’t pay it off, it increases, and the more the debt increases, the faster the rate of increase. If, instead of debt you have capital, you can invest it in government bonds at a fixed rate, and re-invest the interest. Now the more savings you have, the faster it grows. Of course, if you don’t have debts, and don’t have money to invest, your financial situation will stay the same forever, at zero. Savings of zero cannot grow or decrease



In the case of the inverted pendulum, with the bob straight up, there is nothing pulling it one way or the other, nothing happens either, it never gets to one side or the other because it never moves. With only a small displacement, there's very little force so it takes a very long time, but as the angle builds up it gets faster and faster. A 1 metre inverted pendulum takes about a second to hit the ground on one side or the other if it starts from a 3mm displacement, 2 seconds from 1mm off centre, 5 seconds from 0.0001mm, and 10 seconds if it's only 10^{-11}mm off. It's difficult to get it exactly upright, but if you could, it would take forever to fall. Arbiters follow an exponential law where every extra fraction of a second allowed to make a decision, enables the discrimination to be improved by a factor.

Sorts of numbers

The arbiter gives some insight into the philosphers' difficulties with choice, but no preference. They drew sharp distinctions between the case where there is a difference between the alternatives, and the one where there is no difference. No such distinction exists.

For a two-degree displacement the fall to one side takes only twice as long as the free fall of a weight under gravity, but only if the bob is exactly upright, does the decision take forever.

If it were possible to get the bob upright, it would take forever, but it's impossible to get the bob exactly upright. There's a continuous increase in the time taken to fall as you get closer to the centre, but you never quite get there. It's the same with the problem of the ass. He only starves to death if he's placed incredibly close to midway between the two bundles of hay. In fact in our inverted pendulum example you would need to get the displacement to within less than the width of an atom to get it to stay up for 7 seconds. So despite the fact that it could stay upright forever, in practice you can't get it to do that.

What our decision making device does, is to take a continuous measure, the starting position of the bob measured horizontally, along the ground, and turn it into one of two discrete outcomes, left, or right. If this distance is 1 metre when the bob is exactly upright, then for any starting point less than 1.0 m it falls to the left, and for any starting point greater than 1.0 m it falls to the right. Instead of left and right we could call left, position 1 and right position 2. These are whole numbers because it is impossible for the pendulum to fall in between 1 and 2 and it has to end up at one or the other.

There are really two different sorts of numbers. There's the sort you use for counting: one apple two apples, three apples etc; and the sort you use for measuring: 1.1 Kg. 366 grams 25.4 grams. People use both together without realising that it can be difficult to mix them. A particularly mean-spirited neighbour once rang me to ask if I could get her half a

Kilogram of apple for a pie. Her friends were coming to visit and she had agreed to bake an apple pie as her part of a joint lunch (you can see how mean she is). She knew that the supermarket had apples in at 99p for 1Kg, so she generously gave me a 50p coin and said that it would be more than enough. Since I was going there anyway, and anxious not to upset her, I checked out the apples at the store. Sure enough 99p for 1 Kg, but all the apples seemed to weigh about 200g. Should I get two apples and risk being accused of swindling her out of 10p, or should I get three and stand the difference myself? She would be sure to notice and disapprove whichever option I chose. Alternatives like asking the store to cut one in half do not seem to be an option here either, I still want to keep on good terms with the place I get all my groceries, and though I could sort through all the apples trying to find two at 250g, or three at 167g, I'm not sure that I want to spend the time.

One sort of number, the number of apples, is discrete, you can only have one two, three, etc., and the other sort, the measurement of the weight is continuous, any value between 0 and 1 Kg, for example 166.67 g, is possible. A discrete quantity is like two apples or two sides is absolutely precise, but a measurement is less precise, 199.9 grams could easily be confused with 200 grams on the weighing machine.

Just as there are two sorts of number, correspondingly, there are two kinds of infinity.

Chapter 5 Kinds of infinity

Language is one of the defining features of the human race. No other species has developed such a sophisticated method of communication, but what can be communicated by the words and structure of a language only reflects what people can think about. Concepts such as colours, music, classification of kinship, etc. appear in all the languages of humanity, and are universal concerns. In any list of ideas held in common by all languages, numbers make up a very small proportion. Most, but not every, language has a means of describing one object (singular) or many (plural). Most languages have words for the numerals one, two three and so on, but some do not have words for numbers greater than three. Beyond a few discrete numbers and even less fractions such as half, there is no inbuilt feeling for numbers. The proof of this is all around us in the shops, which advertise goods for sale at £1.99, £2.99, and £9.99. If people did not believe that these prices are very different from £2, £3 and £10, it would not be worth doing. Ideas like the decimal point have to be learnt, and the nature of infinity is certainly not easily understood. Nevertheless, it is necessary to understand how to reason about continuous measurements and infinitely small quantities to deal with apparent paradoxes like Zeno's Arrow and Buridan's ass.

Harmony and rationality

Few ideas are created out of nothing, and Platonic theories about ideal forms as well as Aristotle's attempt to rationalise all thought by using logic had their roots in earlier feelings that the natural world must be ordered and harmonious. For the ancient world the key to this was the concept of number.

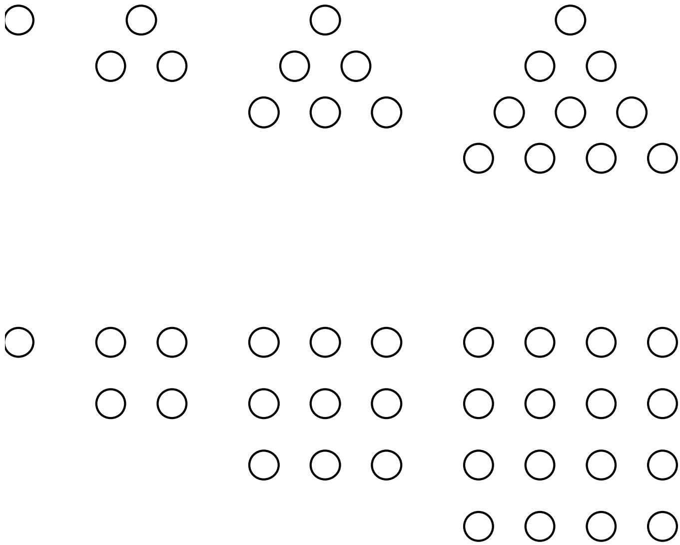
Around 500 BC Pythagoreans taught that the universe was ruled by whole numbers. Numbers had a satisfying perfection. Two plus two always made four, and not only did calculations provide certainty in an uncertain world, but the relations between numbers had hidden significance. Odd even, divisible, indivisible all ratios had meaning. This was their *arithmetike*, and what proved its importance was that it could be demonstrated in music and astronomy. In music, intervals between notes in a chord that are pleasant are related to whole numbers, but those that are not related to whole numbers sound discordant.

It was said that Pythagoras on his long voyages listened to the music of flapping sails, and the wind whistling and whining through the ship's rigging and playing a melody on the ropes. He decided then to investigate the connection between the sound of the wind and the vibrating strings. In another version he was strolling through the village of Croton, deep in thought, listening to the musical sounds of hammers striking anvils in a blacksmith's shop; when suddenly, tripping on a taut string that some children had stretched across the street, when he got the inspiration for an experiment.

Using stretched strings of different lengths placed under the same tension Pythagoras found the relation between the length of the vibrating string and the pitch of the note. He discovered that the octave, fifth, and fourth of a note could be produced by one string under tension, simply by "stopping" the string at different places: at one-half its length for the octave, two-thirds its length for the fifth and three-fourths its length for the fourth. For the Pythagorean philosophers his great discovery was the *tetrachord*, where the most important harmonic intervals were obtained by ratios of the whole numbers: 1, 2, 3, and 4.

The music of the spheres

In the relation of number and music, Pythagoreans believed they had found the pattern that controlled the movement of the planets through the heavens. The sun and the planets were spheres, moving through the sky on perfect circular orbits, one inside the other, separated by the harmonic ratios of musical intervals. Time and space were related by mathematical ratios, wheels within wheels, with the individual planets emitting the individual notes of the perfect chords of the heavens.

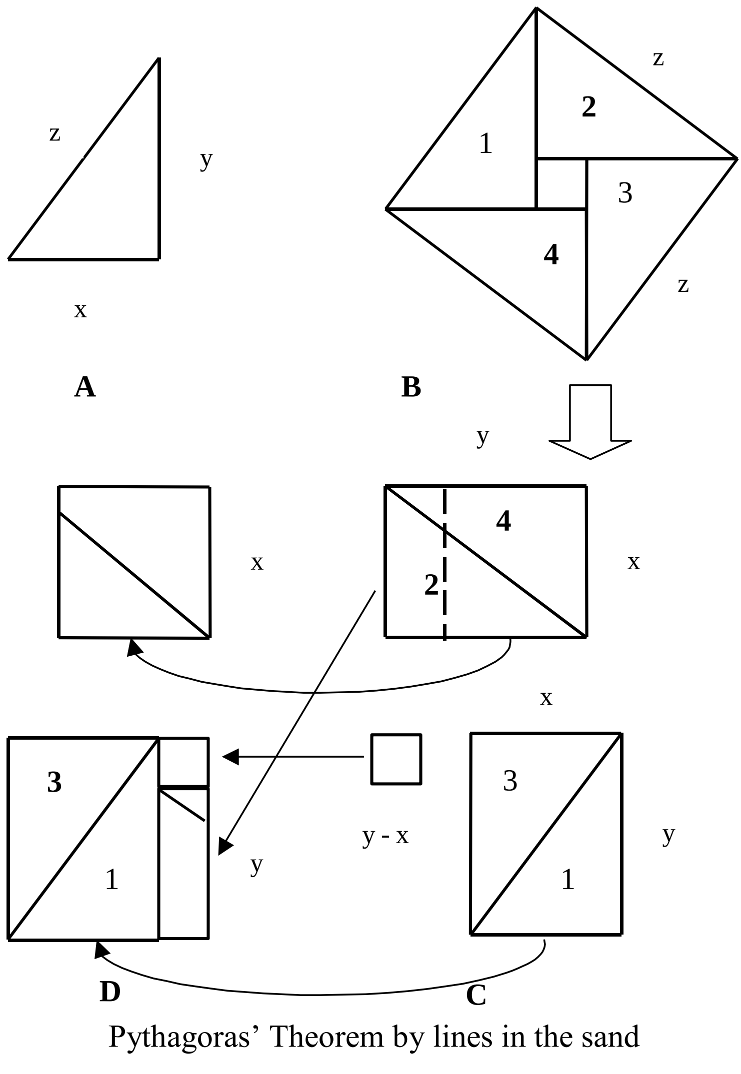


Pebbles in the sand

Numbers could also be built into shapes by placing pebbles in the sand. There were triangular numbers, 1, 3, 6, 10; square numbers, 1, 4, 9, 16; cubes 1, 8, 64, 125; etc.

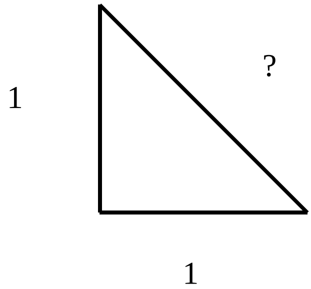
Perhaps all shapes, and even the laws of the universe could be reduced to whole numbers or the ratios of whole numbers. The pyramids had been built on the knowledge that a rope triangle gave a perfect right angle if its sides were in the ratio 3, 4, and 5. Not only that, but it was possible to show by geometry that for all right angled triangles $z^2 = x^2 + y^2$. What you do is to rotate the triangle in A four times, and the square formed, in B, has sides of z by z. It is made up of four of the original triangles, and a square of size (y-x) by (y-x). The four triangles and one small square can also be reassembled as two rectangles x by y, and the (y-x) by (y-x) square, in C. Then if you cut a bit off one of those rectangles and stick it on the side of the other together with the y-x square, you finally have the two squares x by x, and y by y, in D. The z square is made out of the same bits as the x and y squares. All of this could be done by reason and by lines drawn in the sand, and its most useful aspect was that you could mark out an exact right angle for the corner of a new building with three pegs, and a loop of rope 3 + 4 + 5 units long

The Pythagoreans thought that they had the key in their grasp, but there were one or two irritating difficulties. Right angled triangles were few and far between, 3, 4, 5 was well known. Lesser known was the 5, 12, 13 triangle but what about one with sides of 2, 2 2/3, 3 1/3? Obviously you can make new triangles with smaller sides simply by dividing all the sides on existing ones with a whole number. In this case it's obvious that if you multiply all the sides by three it's just a smaller version of the 3, 4, 5 triangle where the lengths are expressed as a fraction. So that was OK, all you needed to do was to multiply all the fractions in a small triangle by a big enough number, and you get whole number sides where $z^2 = x^2 + y^2$.



Pythagoras' Theorem by lines in the sand

But they didn't seem to be able to do that in every case, and what they really needed to build a pyramid was something that would help them get the corners of the foundation square.



According to Pythagoras, the formula was $1^2 + 1^2 = ?^2$. Clearly $?^2 = 2$, but what was ?.
One squared is smaller than two, and two squared is too big at four, but there's no whole number in between. If we want to use whole numbers, we can try multiplying all the sides by 10. Now we get $14^2 = 196$, or $15^2 = 225$. Closer, but still not right. $141^2 = 19881$, or $142^2 = 20164$. Still no exact fit. The followers of Pythagoras struggled with the problem, it's easy enough to draw the diagonal, of a square, but no matter what they multiplied by, they could not get whole numbers that represented the ratio of the diagonal to the side.

Irrational numbers

One of the group, Hippasus of Metapontum, finally demonstrated that it was impossible to represent this ratio in terms of whole numbers.[1]

$? = m/n$
$2 = m^2/n^2$
$m^2 = 2n^2$

This is really bad news for the rational universe, because it means that at least one simple ratio cannot be expressed in terms of whole numbers. Others were to follow, triangles like 1, 2, $\sqrt{5}$; the ratio between the circumference and the diameter of a circle, π, and many more.
Because they could not fit them into their philosophy, the Pythagoreans called such numbers *irrational* and swore to keep them secret, for the discovery of these irrationals was surely something that would be explained in time, and with the right theorem.
Hippasus himself did not believe the problem would go away, and. broke the oath of secrecy. When they expelled him from the circle of Pythagoreans, he promptly set himself up as a public teacher of geometry, and revealed the irrationality of numbers to the world.
Actions like this are deeply threatening to those who believe they are the sole guardians of truth. Shortly afterwards he was drowned in a mysterious "accident" at sea. Some said that a storm had struck his ship as a direct vengeance from the gods; others, that he had been pushed overboard by agents of the Order of Pythagoreans. The next step in explaining irrationals took 2000 years

---

The only way we can get m^2 to be a multiple of 2, is if m is a multiple of 2, call it 2q

$$4q^2 = 2n^2$$
$$2q^2 = n^2$$

So, n is a multiple of 2 as well
That means, both m and n are multiples of two, which is impossible, because m/n was reduced to its lowest terms. So, we have proved that the square root of two cannot be expressed as a fraction

Newton and Leibniz

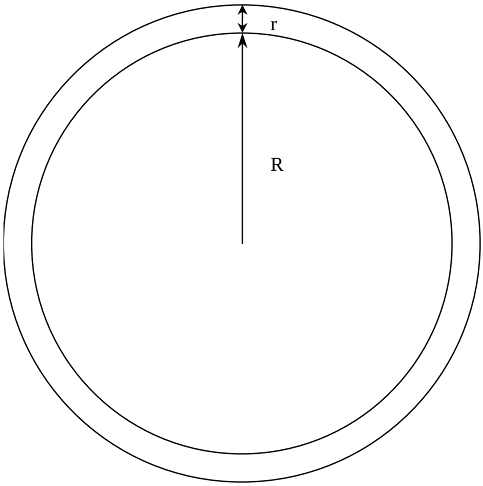
One of the things that worried the Greeks was the idea of the infinite. Zeno's paradox is the best example of this concern, because it seems to suggest that motion is impossible: -
If an arrow moves from an archer at A to its target at B then before it reaches B it has to pass through the mid-point, say B1 of AB. Now to move to B1 it must first reach the mid-point B2 of AB1. This argument can be continued to show that there are an infinite number of points that must be passed, and therefore an infinite number of distances that must be traversed before the arrow can get to its target.
Arrows do not fly at an infinite speed, so each distance must take a finite time, and since there are an infinite number of distances to travel, the arrow can never get to the target.
The argument seems watertight, but is obviously wrong, so what's wrong with it?
The thing that's wrong, is that while there are an infinite number of distances, each of them is infinitely small. Reasoning about infinity times zero, is not just difficult, it's impossible. Better ways of reasoning about a large number of small quantities were the response of mathematicians in the 17^{th} century, who needed to explain how the earth travels round the sun, and how to calculate the trajectory of a cannon ball. The Greeks were able to make sense of straight lines, but the ellipse of the earth's orbit and the parabola of a projectile are curves which could only be explained by thinking about them as an large number of very short straight segments. By the 17^{th} century it was necessary to solve this problem in order to explain how gravity determined the earth's orbit, and Newton used a method which he called fluxions, in which he calculated the instantaneous speed of travel of an object in two dimensions, say height and distance.

The ratio of the upwards motion to the motion along gives the rate of change of height, or how many metres upward a body moves for every metre it moves along the ground.
Calculating the rate of change from a formula giving the position is called differentiation, and in a tract written in 1666 Newton discussed the converse problem, given the relationship between x and the rate of change of y, how do you calculate y?
This is the kind of problem you need to solve if you are calculating the area of a circle, something the Greeks could only do approximately. Using Newton's method of fluxions it could be done by seeing what happens when circle is increased in size by a very small amount. If you know the rate of increase in the area for a particular value of the radius, you know the area added on by increasing the radius, and by adding up all the increases from zero to the value you want you get the total area.

Zeno would say that this is impossible, the answer must be infinite, but Newton could do it by observing that a circle of radius R, which increases by a very small amount r, will increase in area by $2\pi Rr$ because the ring that's the area of the ring. So the rate of change of area with increase in radius is $2\pi R$. Newton also knew, that it was possible to go the other way, a rate of change of $2\pi R$ was given by an area of πR^2, so πR^2 must be the area of the circle[2]

---

[2] To see this, just subtract the area of the smaller circle from the larger one: $\pi(R+r)^2 - \pi R^2 = 2\pi r(R+\frac{r}{2})$. If r is small enough compared with R it can be neglected, and the difference is $2\pi rR$



Area of a circle

The argument can only be justified by making the small increase in radius r so small that the area of the extra ring is very close to $2\pi Rr$. Of course, it never gets exactly there, but as r/R approaches zero, the approximation can be made as near as you like, so in practice, adding up an infinite number of infinitely small rings can be made to give a sensible answer.

Newton did not fully publish his findings until 1687, so unaware of his work, Gottfried Leibniz published the first account of methods he called the differential calculus in 1684 and then published the explanation of integral calculus in 1686. Since Leibniz was the first to publish, he was given the credit for the discovery for a number of years, which led to furious accusations of plagiarism from Newton. Leibniz had had contacts with those who had seen Newton's work, but he had been thinking of the foundations of the calculus very differently from Newton.

The idea of infinitely small steps also applies to Buridan's Ass, and Leibniz makes it quite clear that in his view, it is never possible to have absolute indifference of choice. He writes:

"All our unpremeditated actions are the result of a concurrence of petites perceptions, and even our habits and our passions which so much influence our deliberations, come there from. ….And yet, I see that among those who discuss freedom of the will, there are some who, taking no notice of these unperceived impressions, which are capable of inclining the balance, and imagine an entire indifference in moral actions, like that of the ass of Buridan, equally torn between two meadows."

So it is always possible to find smaller and smaller motivations, and the choice can never quite become exactly balanced.

The diagonal proof

Georg Cantor

Georg Cantor was the man who showed conclusively that continuous and discrete numbers were different. Born in St Petersburg on 3 March 1845, to well to do and talented parents. His father was a broker in the St Petersburg Stock Exchange and both parents loved music and the arts. Georg himself was an outstanding violinist. When he was eleven years old the family moved to Germany, but Cantor remembered his early years in Russia with great nostalgia and never felt at ease in Germany, although he lived there for the rest of his life.

At first they lived in Wiesbaden, then they moved to Frankfurt. Though Cantor's father wanted him to be an engineer, Georg eventually persuaded him to let him study mathematics at university in Zurich, moving to the University of Berlin in 1863. While at Berlin Cantor became involved with the Mathematical Society, being president of the Society during 1864-65. After a short spell as a school teacher, he joined Halle University being, promoted to Extraordinary Professor at

Halle in 1872 and in that year he published a paper in which he showed that irrational numbers could be defined in terms of infinite convergent sequences of rational numbers.
It’s not obvious that the number of fractions is the same as the number of whole numbers, there seem to be many more of them, because they fill in the gaps between the whole numbers, but in 1873 Cantor proved that this is indeed the case.
His idea was to count all of them. If you can do that, then each fraction, or rational number, labelled with different count. The first is 1; the second rational is 2, and so on. It doesn’t matter that you have to go on counting forever, if you can show it’s possible to count all the rationals, then there is a whole number for each, and therefore there is the same number of rationals and whole numbers. What he did was to lay them out in a grid with all the possible numerators along one side, and all the possible denominators along the other.

Denominator

Numerator

	1	2	3	4	5	6	7	8	9
1	1	1/2	1/3	1/4	1/4	1/6	1/7	1/8	1/9
2	2	2/2	2/3	2/4	2/5	2/6	2/7	2/8	2/9
3	3	2/2	3/3	3/4	3/5	3/6	3/7	3/8	3/9
4	4	4/2	4/3	4/4	4/5	4/6	4/7	4/8	4/9
5	5	5/2	5/3	5/4	5/5	5/6	5/7	5/8	5/9
6	6	6/2	6/3	6/4	6/5	6/6	6/7	6/8	6/9
7	7	7/2	7/3	7/4	7/5	7/6	7/7	7/8	7/9

It's pretty clear that all the possible rational numbers lie on the grid, and if they do, you can number each square on the grid. That means you can count them. Just start in the top left corner, the first four can be counted in the square with 1 or 2 as numerator or denominator, the next 5 by considering 1, 2, and 3 as numerator or denominator, and so on. Each square on the grid can be labelled with a unique whole number, so if you do that you've counted them, and there must be a new whole number for every rational.

Can you do this with real numbers, that is, the numbers we use when measuring distance and time? Usually those are expressed as decimals, 1.35 metres, 23.11 seconds, 1.4142 etc., but the important difference between a fraction and a real, is that the real can go on for as long as you like, each extra decimal place giving more accuracy, and incidentally increasing the possible number of numbers that can be made. If, in fact, the number of decimal places is infinite, are the number of possible reals still countable?

Georg Cantor proved that the real numbers were not countable in 1874, simply by showing that it is not possible to list them all, and since you can't list them, you can't count them.

Suppose we try to make a list of all the possible reals between 0.00 and 1.00. It seems straightforward. Start with 0.0000… yes, it goes on for an infinite number of places but you can assign the first row in the list to that number, and that's all that matters. The next one is 0.1000…, and the next 0.2000… When you get to 0.9000… the next is 0.0100… then 0.1100…, 0.2100…, and so on until all the decimal places have been through 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. That's all possible reals in a list. Or is it? Well no, actually. If you tell me you've made a list, by any method you care to name, then all I have to do is to change the first digit in the first row, the second digit in the second row, and so on.

Decimal Places

Reals

0.1	5	6	3	2	8	7	3
0.9	0	2	7	2	8	2	0
0.2	0	4	7	1	9	6	4
0.6	1	2	9	8	5	6	3
0.8	7	5	1	5	9	8	7
0.4	9	6	4	8	1	9	6
0.1	0	4	7	9	2	9	6
0.0	1	2	9	5	5	5	6

Decimal Places

Reals

0.2	5	6	3	2	8	7	3
0.9	**1**	2	7	2	8	2	0
0.2	0	**5**	7	1	9	6	4
0.6	1	2	**0**	8	5	6	3
0.8	7	5	1	**6**	9	8	7
0.4	9	6	4	8	**2**	9	6
0.1	0	4	7	9	2	**0**	6
0.0	1	2	9	5	5	5	**7**

All the reals?

Now by making a real out of this diagonal, I have a new number that is different from all of yours by at least one digit. The first digit differs from the first digit in the first number, the second digit with the second digit in the second number etc. So there are more reals than in any list you can devise, or to put it another way, the reals cannot be counted.

Having shown thus that the number of points on a continuous line must be infinitely greater than the number of discrete grains of sand in the universe Cantor wondered whether there was yet another sort of number which was bigger again. He asked

“Can a surface (say a square that includes the boundary) be uniquely referred to a line (say a straight line segment that includes the end points) so that for every point on the surface there is a corresponding point of the line and, conversely, for every point of the line there is a corresponding point of the surface? I think that answering this question would be no easy job, despite the fact that the answer seems so clearly to be "no" that proof appears almost unnecessary.”

He became engaged to Vally Guttmann, a friend of his sister, in the spring of 1874 and they married in August. It was the happiest and the most productive time in his life, and in 1877 he was able to show that the number of points on a line was the same as the number of points on a surface. So all measurements were the same, and there was no new type of number. He said

“I see it, but I don't believe it!”

One man who did not see it and did not believe it was Kronecker, who had lectured to Cantor when he was a student at Berlin. Kronecker vociferously opposed Cantor’s work. He insisted that arithmetic should be based on whole numbers saying

“God made the natural numbers; all else is the work of man”.

As a result of opposition to his work, and lack of support from his colleagues, Cantor began to succumb to depression. Towards the end of May 1884 he was clinically depressed, recovering after a few weeks but unable to continue his research.

“I don't know when I shall return to the continuation of my scientific work. At the moment I can do absolutely nothing with it, and limit myself to the most necessary duty of my lectures; how much happier I would be to be scientifically active, if only I had the necessary mental freshness."

He tried to heal the rift between himself and Kronecker in 1891, but his attempts to reconcile their opinions by getting Kronecker to speak at a meeting in Halle failed. Kronecker could not come because his wife died, and Cantor became depressed again, never again producing mathematical work of much significance after 1893. He was in and out of hospital up until his death in a sanatorium in 1918.

Chapter 6 Decisions at the speed of light

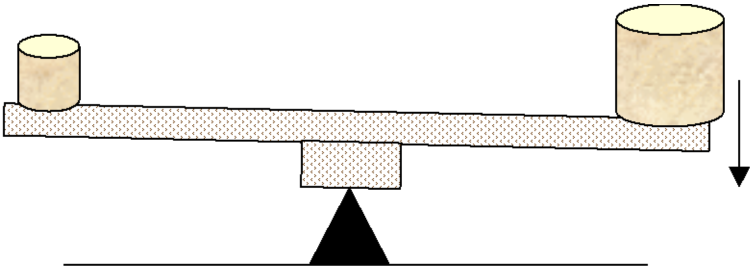
Can it really take forever?

Deciding between two alternatives means making comparisons. Is this bale of hay nearer than that one? Am I ahead slightly ahead of that car or a bit behind it? We have to look at the evidence and decide, but sometimes it's a very fine judgement. You need to look very closely at the evidence to make up your mind, and very small things can matter. This is what makes decision-making difficult, because sooner or later you'll get a comparison that is very close to equality. Not exactly equal, but very close. How close can the comparison get? Infinitely close, and as it gets closer, the time needed to decide which is bigger or smaller gets longer.

Even for a simple measurement, then there is an infinity of possible values that it can have. The choice that has to be made might be left or right, up or down, red, yellow or blue, but all of these are countable, and the number of possible comparisons between measurements, is not countable, so there are an infinite number of inputs for every distinct output, all of them different. To get the right answer takes time, the finer the comparison the longer the time, until we the comparison is so close to zero, that the time taken increases without limit. The chance of getting the distances between two bales of hay close enough that the donkey could take days, or weeks to decide is quite small, but if it does happen it really could starve to death.

If we are going to rely on machines to make our decisions then, it would be a good idea to know what the chances are of a long decision time.

Mechanics and electronics

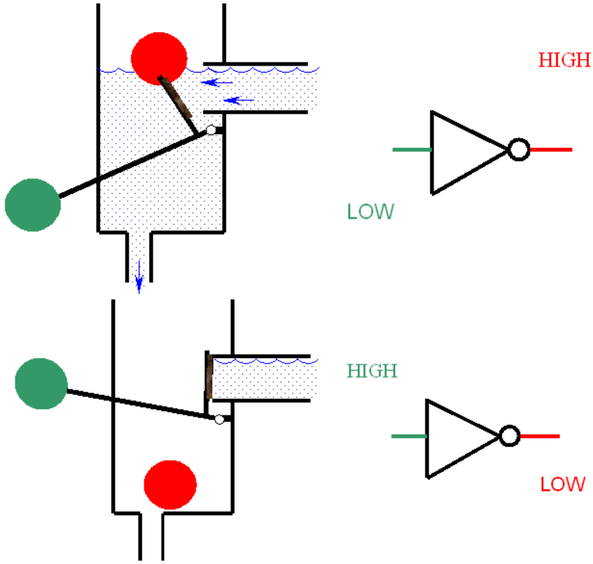


A weighing machine

The oldest mechanical decision maker is a set of scales, used in the market place for millennia to weigh objects for sale, grain, meat, metals, or anything else that we want to assign a fair value. If one is made from a plank of wood placed on top of a wedge, so that its centre of gravity is higher than the apex of the wedge, and the two bales are put one on each end of the plank, the heavier end will fall. If both bales are very nearly the same weight, it could take a long time for the heavier one to move towards the ground, but once it is there, it stays there. The plank has two stable states, left side down, or right side down, but the balanced state is unstable, because its centre of gravity is above the pivot point.
Mechanical devices like this are slow. Comparing two weights can take seconds, or even many seconds, because the speed of the plank is limited by its inertia and the inertia of the bales of hay, but if there's plenty of time to do the weighing there's usually no problem. Until the 20th century mechanics were used for instruments like clocks, calculators, and even the earliest computer, the difference engine, proposed by Charles Babbage in the 19th century because that was the only way to do it. In the 20th century, two things happened. The first was

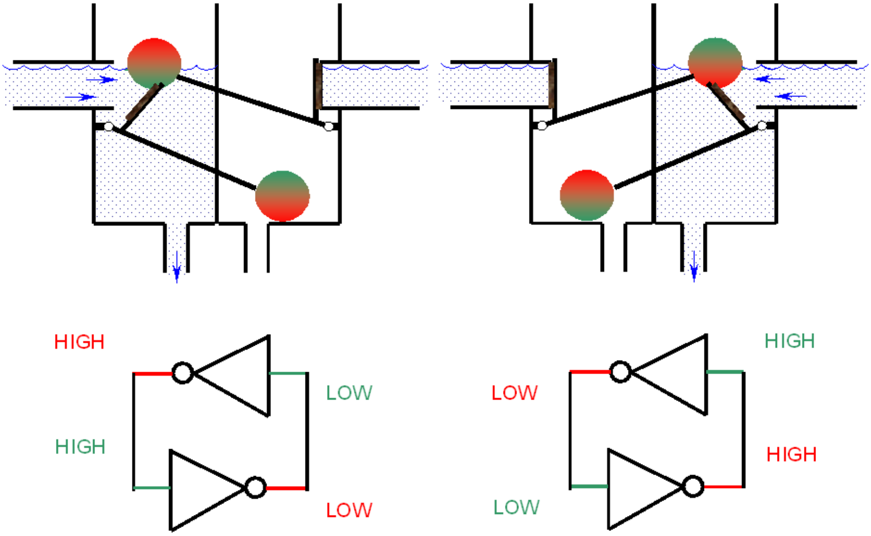
the invention of electronic devices that worked much faster than mechanical ones, and the second was an urgent need for more and faster decisions.

At the start of the second world war code breaking equipment was made from of hundreds of mechanical devices, making thousands of decisions in an hour, but at the end of the war, a few electronic systems making millions of decisions per minute had been developed. Electronic decision makers have very similar characteristics to simple mechanical ones like the weighing machine, but they're much faster. The speed of light, which is an electromagnetic wave, is about 1 million times faster than the speed of sound, and it is the speed of sound in materials that in the end limits the speed of mechanical things. If you push one end of a lever, that force needs to be transmitted to the other end to do its work, and it can't go faster than the speed of sound. Transistors are also much smaller than levers, microns rather than metres or millimetres, so the electrical signals don't have to travel as far on a silicon chip as the levers move in a mechanical clock. The smaller something is, the faster it can go, so we can expect that an individual electronic decision maker will work in nanoseconds – thousand millionths of a second – rather than seconds.
All decision makers in computers or indeed any electronic systems are ultimately based on a circuit called a flip-flop, which is made of two very simple inverter circuits. An inverter gives an output that is either high, or low, and is the opposite (or inverse) of its input. It uses transistor switches, which control a flow of electrical current to do that, but it can be described in terms of water flow rather than the flows of electrons that it actually uses.



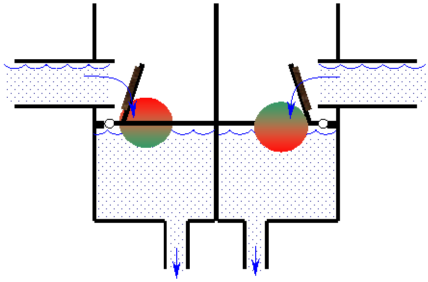
The inverter

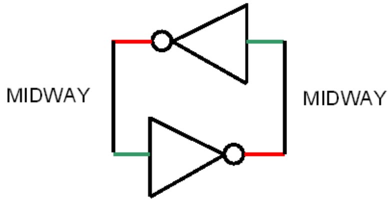
Using a tank and valve we can make an inverter, that is, something that has an output that is high when the input is low, or a low output when the input is high. In the diagram, a lever attached to a green ball controls the input valve, and any water in the tank drains off slowly though an outlet at the bottom. If the ball controlling the input is low, the valve is open, more water comes in than can get out, and the tank fills up. The output is a red ball floating on the water in the tank, which floats up to the top of the tank when water comes in, and falls to the bottom when it flows out. A low input ball fills the tank and gives a high output ball, and a high input shuts off the inlet, the tank drains and gives a low output.



A flip-Flop with two stable states

A flip-flop, the circuit that can store one binary bit, is made from two inverters connected back to back. Connecting two inverters in a loop means that if the input of the first is low, its output is high. Now because its output is the same as the input of the second, then the second inverter output must be low. Connecting back that output to the first input keeps that input low so the circuit is in a stable state. It could also be in a stable state with the first inverter's output high, and the second input low, so the flip-flop has two states that can store the information represented by one bit, 1 (high) or 0 (low). To change states you can force one input to high if it was previously low, and one inverter tank empties causing the second inverter tank to fill, and then you can let go, it'll maintain itself in the new state.





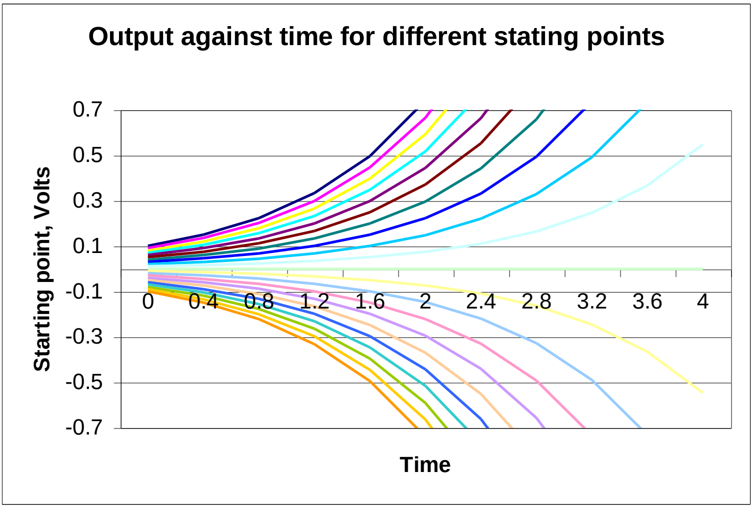
The metastable state

Unfortunately, if you don't hold the input up long enough, it's possible to get into a state where both tanks are part full, and the same amount of water comes in at the top as goes out at the bottom. This is not stable, but if the water levels are exactly the same there's nothing to tip it one way or the other. Eventually it'll end up in one of the stable states, but the nearer it is to balance the longer it takes to get there.

Mostly computer designers try to ensure than this can't happen by ensuring that the flip pulse that set it to a 1 or a 0 is long enough for it to get to a stable state, but there are circumstances where that can't be done

Early computer designers assumed that you could always get over this problem by making sure that there was always enough of an input pulse to push the flip-flop to the other stable state, and if there wasn't that it would only stay there for a limited time anyway. Neither of these is true. It turned out that flip-flops can behave in the same way as any other real physical decision maker, getting stuck for an arbitrarily long time in the metastable state. It has to take time - and the time taken depends on how close it is to balance.

Near balance, where there is little or no preference for one state or the other the time taken to get a result depends partly on the innate response time, of the device, which we call τ, and partly on where it starts. For our water tank, the time taken depends on how long it takes the water to fill or empty the tank, usually several minutes, though if you make the tank small enough it could be seconds. In mechanical systems, τ might be about 0.5 seconds, but for an electronic flip-flop, τ can be as short as 20 pico seconds.



Response times of a flip-flop

Starting from a point distant only 0.1 times the full output from the centre, which might correspond to quite a large input pulse, it takes about 2τ to move to a 1. If we make the pulse smaller, say 0.01, it might take 4τ. If the input pulse sets the flip-flop very near to the centre, it can take 7, 8 or even 30 times the normal response time to get to a stable output, either a 1 or a 0. Naturally, these long response times don't happen very often, but if it can happen, it will happen, and if you haven't considered what the system should do there is going to be trouble.

The birth of the baby

One of the first computers in the world was built at Manchester University in 1948 by F. C. Williams and Tom Kilburn. It was built from ideas, which were first developed during the second world war. The Bletchley Park code breaker, Colossus, had shown that millions of calculations could be done reliably by electronics, but what was missing was a cost effective means to store results, and hold the program itself. At the Telecommunications Research Establishment (TRE) at Malvern, F. C. Williams led a small group that worked on electronic circuits, mainly for radar. Immediately after graduating in 1942 with a first in Mathematics from Cambridge, Tom Kilburn was called-up to serve in the war. He was ordered to take a City & Guilds crash course in electricity, magnetism and electronics, and then to report to Williams at TRE where they formed the working relationship that led to the Manchester computers.

When the war ended the immediate demand for circuit solutions to radar problems disappeared and F. C. Williams took up a chair in the Department of Electrical Engineering at the University of Manchester. Tom Kilburn went with him to work on computers. No practical stored program computers had yet been built, because the key to building a computer was the memory. Williams and Kilburn were interested in storing information on a Cathode Ray Tube as a way to make the memory, and there were hundreds of war surplus CRT's available which could provide the technology. By the end of 1947 they had succeeded in storing 2048 bits on a CRT. There was no better way of demonstrating that the CRT store could provide a reliable and effective main memory for a computer, than building a computer, and getting it to store and run a program.

Tom Kilburn with FC Williams at the console of the Mark1

Now it was Kilburn rather than Williams, who led the work on designing and building a Small Scale Experimental Machine, "The Baby". On June 21st 1948, the first computer in the world that could hold any user program in main memory ran a program that Tom had written. In 1992 he recalled:

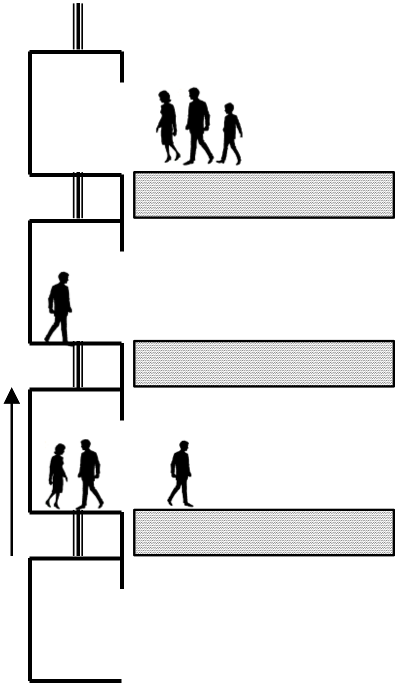
"... the most exciting time was June 1948 when the first machine worked. Without question. Nothing could ever compare with that."

Having shown that it could be done, Kilburn and Williams wanted to build a more powerful and useable computer, and the government had become interested. A local Manchester company, Ferranti, had a factory near the University, at West Gorton, Manchester. In the autumn of 1948, Ferranti Ltd. were commissioned to build a commercial machine "to the instructions of Professor F. C. Williams", known as the Manchester Mark 1. The Mark 1 was fully operational by around October 1949 at the University, and the commercial version, the Ferranti Mark 1 first ran a short time later. In early 1951 the Manchester Mark 1 was replaced by a Ferranti Mark 1, and used by the university in its own Computing Machine Laboratory.

At this point Williams turned back to Electrical Engineering and was no longer actively involved in computer design, but Kilburn drove forward with his plans to build better and faster computers. A self-contained, slightly shy man, he was also very determined, and did not allow problems, technical or financial to get in the way of his ambitions. He wanted to build bigger, faster, and better computers. In 1956 the MUSE project started, with the idea of making a computer 1,000 times faster than Mark 1. Transistors had replaced valves as the switching devices, and core memory was much more reliable that the CRT. Ferranti joined in 1959 and the project was renamed Atlas. By this time Tom Kilburn had become a Reader, and his team included around 20 university people as well as 8 from Ferranti. A computer with an instruction rate of less than 1MHz does not seem much now, but in 1957 there were those who said that it was an unnecessary project, as it would provide far more computing capability than the country needed, then, and for the foreseeable future. The Atlas computer was finally inaugurated at the Electrical Engineering department in 1962, and was acknowledged to be one of most powerful and sophisticated computers in the world at that time.

Synchronous and asynchronous

In the beginning, the design of computers did not follow any particular principles, the engineers just made it up as they went along. This applied to the organisation and timing of the computers as much as to anything else. To obey an instruction, the machine had to fetch it, decode it to find out what to do, fetch the numbers needed, do the operations on them, and then put the result back before going on to the next instruction. Each of these things happened pretty much when they needed to happen. When one thing finished the next thing was started. Arranging the timing like this is now called *asynchronous*, the times are not fixed, only the sequence of events. If each individual time period is different, complex machines can be very difficult to design and understand, so as they got bigger, a simpler design style, called *synchronous* came to be adopted by some groups, in which all the timing is fixed by a clock signal. Each operation, down to the simplest must always start on one clock tick, and finish on another. Because all the operations complete on a clock tick, the designer knows that the result can be used immediately, by another part of the machine that starts on a tick, and all the parts fit together in a simple way.



A paternoster lift

A paternoster lift in which all the cabins are linked together is an example of a synchronous system. All the cabins move at the same speed from floor to floor driven by a common rope, and all the cabins reach a floor at the same time. People get on, and people get off on all the floors, and the lift moves at a constant speed.

A highway full of cars is an asynchronous system. Cars can join the highway at any time, and the distance between junctions is not fixed. Each of the cars on the highway can move at it's own speed, and one can overtake another, but after moving out to the fast lane to overtake, the driver may want to get back to the original lane. This involves a decision about which gap in the traffic is sufficiently big, or far enough ahead to get into, so decisions are not avoided in asynchronous systems, and there may even be more of them required.

In an asynchronous computer, two requests for the same processor can coincide, and you then need to decide which one is taken first. Similarly on a highway, two cars can be intending to go into the same space, one continuing on the slow lane, and one trying to cut in from the fast lane. Who gets the priority, and how often do they collide? There is no chance of the cabins in a paternoster colliding, accidents can only happen when the people, who can arrive at the lift in their own time need to decide which cabin to get in. Synchronous systems don't have internal timing problems, only synchronization difficulties when an external event happens along. Asynchronous systems can have internal timing collisions as well, so are more vulnerable to the problems of choice.

In the 50's and 60's there was little contact between computer development groups in the US and the UK. Similar advances took place, but the ideas that drove these advances came from different individuals and arose mainly out of local influences. One key difference between the continents was that the US computer systems became largely synchronous, that is, there was a central clock which controlled the timing of all information flow in the processor, while in the UK, and particularly at Manchester, processors were mainly asynchronous. Events were individually timed, and could happen at any time. The result of this was that UK computers were more vulnerable to coinciding events, simply because such events were more frequent.

Ivor Catt

Ivor Catt joined Ferranti, West Gorton, the factory that collaborated with Tom Kilburn's group on computer design. When he graduated. Between 1959 and 1962, he worked with the logic design of the first Ferranti transistorised computer, Sirius. It had 2,000 logic gates and 40,000 bits of memory and sold for £25,000. Most of the early Ferranti computers, such as the Sirius, were asynchronous systems in which the timing of input/output peripherals, such as the card reader, was independent of the CPU. This means that when the peripheral has to interrupt the sequence of events in the CPU in order to be serviced, there has to be a circuit that decides whether to do the service immediately, or wait until the next time the CPU is ready. A special 'push button' circuit had been devised for this, which was said to be more reliable than the usual circuit to decide whether the card reader or the CPU got dealt with first. Ivor realised that this special circuit hadn't really solved the problem figuring out that the period of indecision could, in theory, last for a very long time, making digital computers fundamentally unreliable.

He knew that this would not go down well with the manufacturers of computers, because one of the major reasons why digital computers are better than their forerunners, analogy computers, is that analogy computers deal in continuous values, like voltage, which is used as the analogy of the numbers that are being computed. If the computed voltage varies a bit, maybe because the power supply for the circuits is a bit down, it's not possible to tell the difference between the value it actually represents, and what it should be. There's no way of correcting the error, so errors inevitably build up. A digital computer only deals in whole numbers, and you can't have a whole number that is a bit out. If a voltage is supposed to represent a whole number but it is actually 1.99V or 2.01V instead of 2, we know the number should be 2, and can correct the voltage. Digital computers were supposed to be reliable, that was a key selling point, but now there was a potential source of unreliability in the new digital computers that could not be removed. The customers would have to be told, and it could affect sales

In 1962, Ivor took his family to the U.S.A., departing for Los Angeles and his new job, in Ampex, with high expectations, but found that there, and subsequently at Motorola Phoenix Arizona in 1964, no one seemed to be interested in the problem interrupt circuits. They did not believe a problem existed, and if it did, they didn't want to know. Nevertheless, he wrote a short note about it and got it published in an IEEE journal in 1966. The note is written rather obscurely, and has errors in it, but the drift is clear, there's a problem.

A year later, he was at a conference and exhibition in Chicago demonstrating a product about which he knew little, and cared less, when Tom Chaney and Warren Littlefield walked by. It soon became apparent that they were working on synchronization, and had also noticed that the circuit used to synchronize outside events with the central clock sometimes had a problem. Clearly it was the same problem as Catt's and they were annoyed that Catt had succeeded in publishing in a recognised journal first – even before they had produced their own internal document. Their work 'beware the synchronizer' did not in fact appear until 1973, because Chaney and Littlefield had had considerable difficulty persuading the reviewers to accept their paper, referees would say things like "if this was a problem, I would have heard about it. I haven't so I don't think it exists". It was only after a special Workshop on Synchronizer failures was held by Charles Molnar, Director of the Computer Systems Laboratory of Washington University, St Louis to publicise the work that there was any acceptance of the existence of the problem in the US.

Chapter 7 I don't believe it

Does it matter?

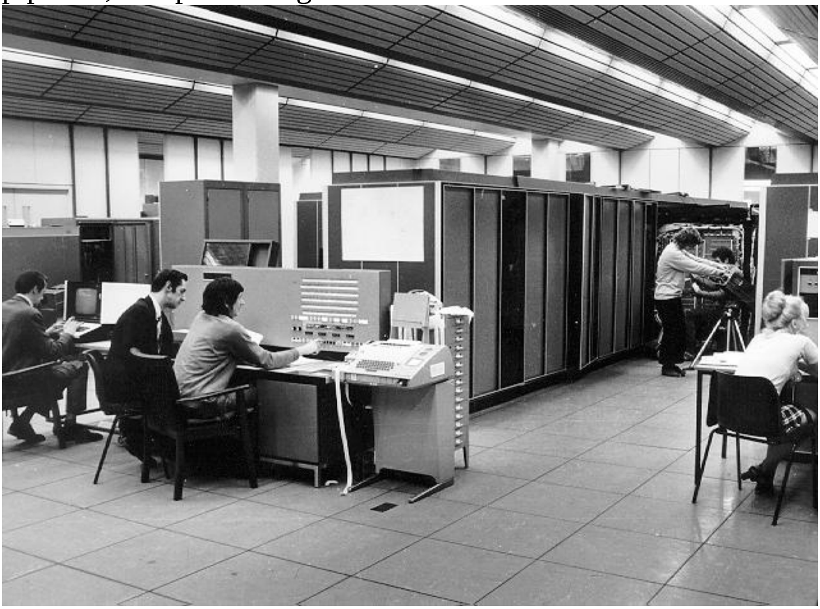
Until the second half of the twentieth century, the problem of choice with little or no preference was only a debating point used by philosophers to score against each other. It had no practical significance, used only to argue for or against the idea of free will, or the existence of God. With the invention of the computer, and the design of ever more complex systems, which could have to make tens millions of decisions every second, it would have practical importance as well. In computers, decisions where there were very small differences between the two alternatives could actually happen, and that when they did, the consequences could be catastrophic. The issue was simple but for many, accepting that such a thing could happen was impossible. Some philosophers refused to accept that decisions could take long enough for a person to starve to death, and some of the best computer designers dismissed even the possibility of their machines failing for want of a decision.

The MU5

By 1966 Tom Kilburn was one of the most respected figures in the world of computing. Made professor of Computer Engineering in 1960 and a Fellow of the Royal Society in 1965 he had also started a new Computer Science department at Manchester, which ran the first undergraduate Computer Science degree in the country. His new computer project was planned have speeds 20 times that of Atlas, and outperform all its rivals. Part of the gain would come from the factor of 8 improvement in circuit technology that was available from new integrated circuits, but the main objective was to produce a new fast architecture capable of running high-level language programs efficiently. Special local memories were to hold the numbers needed for calculations near where they were needed, and aggressive pipelining of the instructions was to give high throughput. Everyone in the new department was expected to contribute to the new ideas that would keep the group ahead of the world, not least the software group who had to produce completely new compilers for the computer and a new operating system making use of new ideas in managing the memories. No university, at that time, or since has taken on such a mammoth task, which often takes hundreds of man-years of effort in a commercial environment.

The computer was to be called MU5, as Tom argued it was the fifth Manchester University designed computer. The climate for research funding seemed favourable so an application for Government funding was made, and in 1968 the Science Research Council awarded a five-year grant of just over £630,000 (probably equivalent to about £10,000,000 in 2004) to the department. At the meeting where the grant proposal was discussed there were considerable misgivings about the amount of money involved, one of the biggest grants ever awarded to one group. The chairman tipped the balance "I think we ought to do it" he said. A team of 16 staff and 25 research students from the new department had started work in 1966 and continued on the project up until 1971. At that time ICL, who had taken over Ferranti's computer division, agreed to give assistance in terms of production facilities and manpower, and the group was increased with 19 engineers from ICL. The project was not finished until mid 1974 when it became possible to run Algol and Fortran programs under operating system control, but earlier, in 1970, success, or even completion of a working computer was not at all clear.

Progress on the hardware and software had slowed to a snail's pace, with each of the many innovations in hardware and software throwing up problems. Most of the new circuits worked well, but some of the high-speed connections involving groups of line matching resistors were not reliable, causing frequent faults. As a result software development had also slowed, and some of the more ambitious features of the operating system had to be cut. It became increasingly clear that the very long pipeline might not have been a good idea. The idea of a pipeline is to increase the rate of processing instructions by doing several things at once. Each instruction is broken down into a number of steps, for example, fetch the next instruction, decide how to do it, fetch the numbers needed, do the calculation, and then store the result. Rather than waiting for all the steps to complete before starting the next instruction, a new one is fetched while the first is being decoded to decide what to do, then another while the numbers are being fetched for the first, and while deciding what to do with the second, and so on. This works well as long as you know what the next instruction should be, but sometimes choosing which comes next depends on some result that has not yet been calculated. When everything is going well a pipeline is very efficient, 25 million instructions could be completed every second, a factor of 70 better than Atlas, but that happened only rarely. Whenever results from finished instructions were needed in order to decide what the next instruction should be, the pipeline could do nothing until the right result came out of the end. In MU5 this meant that all the partially completed instructions were useless, and had to be flushed out of the pipeline. The line could not even be restarted until the result had been brought back to the front end, a physical distance of some 25 meters. Electrical signals travel at close to the speed of light, but even so, that added to the wait, so every time there was a disruption to the pipeline, the pace of execution of instructions slowed to barely 3 times the Atlas speed. It would not matter if disruptions were not common, but in the four years since the start of the project in 1966, programmers had been encouraged to change to a more structured style, in which the code was divided up into short, manageable chunks. Every time the processor had to move from one chunk to the next, there was a disruption to the pipeline, and processing slowed to a crawl.



Commissioning the MU5, 1974

In mid-1970 the strain of such an ambitious project was beginning to show. The move to a new computer building, specially designed with an electricity substation in the basement to supply the 100KW of power required by MU5, and 10 tons of air conditioning on the roof to remove the power when it was finished with, was imminent. There would be a gap in the commissioning of the hardware while it was moved, so it was important that the best use was made of the time available in the old building to get things into shape. People worked nights to try and get back on schedule. Lack of sleep, artificial lights, and constant noise from the forced-air cooling did not help concentration, and then strange things began to happen in the hardware.

The buffer store

At the end of the pipeline, when many of the arithmetic calculations were performed, a small memory – the operand buffer - had been located to ensure that it was not necessary to go to the main store (some 10 meters away) for every number. A number could be stored in the buffer in case it was needed again for some new calculation. Sometimes new results were written into the buffer from the local processing units at the end of the pipeline, sometimes they had to come from the main memory, sometimes they were sent for local processing, and sometimes they had to be sent back up to the start of the pipeline, or to the main store. The timing of reading and writing to the buffer was entirely asynchronous, so any of these requests to read or write could happen at any time, and the control had to cope with that. A simple enough design task, it was done by looking at what requests were outstanding when the buffer had finished each read or a write, and deciding on which of the next had the highest priority. The design worked fine on the simulator, but kept failing when the hardware was delivered and tested. Worse than that, it did not fail in a simple and repeatable way, and because of that it was hard to find out why. Typically it would run for a random number of seconds and then stop with the control apparently either trying to do two requests at once or none at all. It should only do one, and if it was working properly, the control logic could never get in that state, so what had gone wrong?

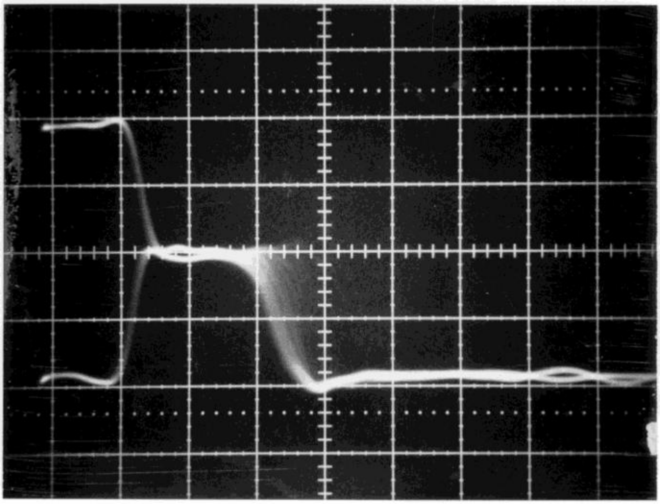
In one second, the buffer could process as many as 10 million requests, so finding the one that had caused things to go wrong seconds before with a conventional oscilloscope that only showed the last few nanoseconds seemed impossible. Most likely it would turn out to be another bad connection, but checking them all did not uncover anything obvious. Could it be that the flip-flops used to hold the request status were taking longer than they should? There were memories of flip-flops 'boggling' in earlier Ferranti computers, and Ken Johnson who was Ivor Catt's mentor at Ferranti, was visiting the University from time to time to keep in touch with the work. What if the flip-flops were the source of the problem? Ivor's paper describing how decision flip-flops could take a long time was a mess. He had approached the problem in a strange way, which did not seem appropriate to the MU5 circuits, but could he have something? Back of the envelope analysis of a better circuit model of the flip-flops showed that indeed it could take a very long time to decide whether there was a request present or not. All you needed was a very close timing between the end of one buffer store cycle, and the arrival of the next request.

In late 1971, Tom Kilburn asked two of his staff to attend a seminar on the teaching of computer science at Newcastle University. The new undergraduate course had been running for 2 years, and someone should go to present the department's new course. Tom rarely went anywhere himself, and he wanted to make sure that his rivals did not have the field to themselves. The seminar was a bit of a distraction from the commissioning effort, but one of the speakers, David Wheeler, from Cambridge, seemingly at a loss for a topic on the teaching of Computer Science, spoke about a problem that he characterised as "rarely well taught", and "apparently difficult to appreciate". He had previously read Ivor Catt's paper, recognised Ivor's problem as one that they had had at

Cambridge during the development of their computers. At the end of the talk, a member of the audience remarked that the effect of teaching that details of timing could affect the behaviour of computers would be to destroy their belief in abstraction. Could they be reassured? The succinct answer was "No". In this case reliability and time are intimately connected; and certainty is only available if you are prepared to wait forever.

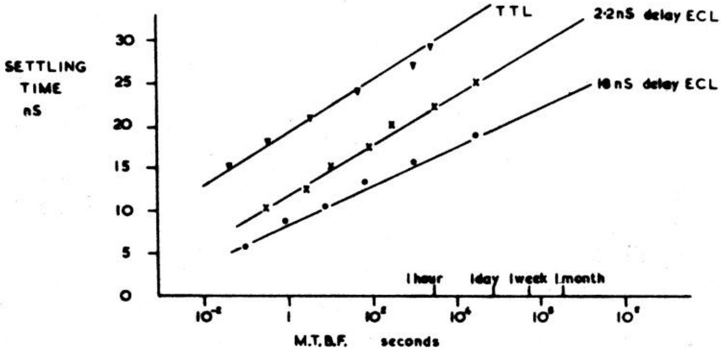
By now it was clear that what was happening in MU5 was connected with long flip-flop settling times. In order to understand what to do it was necessary to get a better handle on theory, and then test it on real hardware. More work on the circuit model showed that the closer in time the next request and the end of the previous buffer store cycle got, the longer the flip-flop would take to make up its mind. If you looked too early at its output you might see a half request, but later the half request might have disappeared. That could be why the control ended up in a strange state, part of it seeing something and part not. But why was the time between failures apparently random? An explanation for that might be the random nature of the timing between store cycle and request, producing different flip-flop timings, sometimes fast and sometimes slow.

A week spent building some test hardware confirmed the theory. Using two asynchronous inputs from independent signal generators for the end of buffer store cycle and the requests then trapping out all responses that gave different results at different times showed that the flip-flop could indeed produce two different results. The photo below shows what happened. As time progresses the trace of the flip-flop's output moves from left to right. The input request is sometimes high representing a request and sometimes low, or no request. At the end of a buffer store cycle, no more requests can be allowed, and the logic has to decide which of the outstanding requests to service, so any new request is locked out by the end of cycle (second vertical line from the left). Now if there has only just been a new request the flip-flop output might have almost got half way between a 0 and a 1 before the request is frozen out, so the flip-flop is nearly balanced, and doesn't decide to go back to 0 until after the 4th vertical line.



Metastability.
The output of a flip-flop sits halfway between 1 and 0

Normally flip-flops only took about 3 nanoseconds (3 thousand millionths of a second, or about 1/3 of a square in the photo) but if you look at the output at 3 nanoseconds, and then again after 10 nanoseconds, (two squares) the results can be different. The output first goes to the metastable state, half way between a 1 and a 0, and in this case the half level is taken by the subsequent circuitry to be a 1. But it eventually resolves to a 0.



Waiting time against mean time between failures

So how long do we have to wait to avoid a crash? The more often the requests come, the more likely you are to have at least one very small time between request and end of cycle, and the more frequently you need to get to the buffer the more likely you are to see the problem. Of course faster circuits help, but the fastest flip-flops that could be bought still needed about 1/3 of the whole buffer store cycle time to give a reasonable reliability, dropping its performance by 30%.

It wasn't just one buffer store that had a problem, all the fast stores had asynchronous inputs and deciding on the presence or absence of a request was fundamental to their operation. It looked like there had to be a trade off performance with reliability, and with any reasonable reliability, the whole computer performance would suffer.

Tom Kilburn had recently become Dean of the Faculty, so was not as closely associated with the project as he might have liked. He heard that there was yet another problem and called a meeting of those most closely concerned to try and sort it out. It seemed to him to be a simple logical problem, and if no one else could solve it he would have to. The meeting stretched out all day, becoming acrimonious at times. A solution that involved loss of performance was not acceptable to Tom, but he was not able to immediately offer a better solution. It was put to him that it could be demonstrated that there could be no solution that did not involve the comparison between two identical times, and if this happened it would inevitably result in an infinitely long flip-flop response, but he did not believe it. The next day he came back with a solution in which the request time had to be recognised as close to the end of cycle time, and then the request could be put off for a short time. This still required a decision between ‘ close to the end’ and ‘not close to the end’ that was equally difficult. The following day he had a new solution, but was less confident, and when that was demolished it was two days before his third proposal was put forward. Only after two weeks did he accept that the problem might be a fundamental difficulty with no simple logic based solution.

An occasional visitor from ICL was Ken Johnson, who had originally pointed out to Ivor Catt the special flip-flop circuits that had been used by early Ferranti computers. Always full of good ideas, he had not done as well as he might have at Ferranti because he understood technical issues better than many of his superiors in management and sometimes came into conflict with them as a result. In one meeting where his ideas were again causing them to rethink it was said of him “The problem is, the ******* is always right!”

He suggested that if you could detect when the flip-flop was moving away from the central, metastable state, it would not be long before it actually got there. In hindsight this was obvious. An unstable system in balance only stays there a long time if it is very near balance. Once it is any distance away, it moves very quickly, so you can detect when it has moved off the centre even if you can’t control how long it takes to move off centre. Once you know it has moved off centre you can guarantee it will get to a stable state in a fixed time, and cannot

take forever. The second piece of the puzzle is that 99% of the flip-flop response times are normal, and it's only the very rare, 1 in a million, responses that are as long as 10 times normal, so on average, starting the buffer store when the request flip-flop had resolved to a stable state would be almost as fast as planned. Waiting for the relatively few cases that needed the extra time would ensure reliability.
This was a solution that cost hardly anything in performance, and, contributed to the final completion of the MU5 project was in 1974. It was the only Manchester computer, which did not have a direct commercial counterpart, although many of the architectural concepts of the ICL 2900 series were derived from those of MU5

The third world war

Later, in 1973, after Charles Molnar's workshop in St Louis, more people began to accept that computers had been, and were being designed that were unreliable because the designers did not fully understand the problem of metastability. In an article in Scientific American that year a Digital Equipment Corporation Engineer is quoted as saying "Ninety-nine out of a hundred engineers would probably deny the existence of the problem. They believe that conservative design is all you need; you simply allow enough time for a flip-flop to reach a decision. But as computers are designed to run faster and faster the safety factor gets squeezed out… the problem looks unresolvable.."

Meanwhile Ivor Catt had returned to the UK, and while working or defence related computers, had noticed that the design of those systems did not allow enough time for synchronising the data exchanged between them. Instead of asynchronous requests from one part of a computer to another, there were asynchronous requests between computers. His book, published in 1974 entitled 'Computer Worship' says, in a typically extravagant passage that "There is theoretical evidence…that a computer will periodically go wild. This evidence is ignored by the computer industry because those few who have come across the theory find it wiser in the short term to ignore it". An excitable and loquacious individual his message that at some "random point in the future" a computer may malfunction and precipitate a nuclear war went unheeded, mainly because he blamed a priestly hierarchy of "Computercrats who are more concerned with their own holy rites than their flock". "Decadence, financial disaster and confusion" was his analysis.

As usual, his language was wild, but the basic message was right, and in 1980 two full-scale nuclear alerts brought the world to the edge of holocaust. Computer controlled displays at the Norad centre in Colorado indicated that a multi-missile attack had been launched against the US from submarines in the South Atlantic, and that the nuclear warheads would fall on US cities in less that 5 minutes. The US General Accounting Office had previously suggested "A basic circuitry problem" in the computer systems could cause problems, but they were unable to investigate in detail because they were denied access by the Joint Chiefs of Staff.

'Well, Dammit, That's Not What This Printout Says!'

It's not over yet.

Much of the early work to characterise metastability was done at Washington University where Thomas Chaney did extensive measurements on metastability characteristics of latching devices using a wide variety of discrete devices available in the 70's.

Despite that, and the work done by a few other pioneers, not many people even knew there was a problem, and companies were designing inherently unreliable hardware. As late as 1987, many of the major microprocessors manufacturers, AMD, Zilog, Intel, and Signetics all experienced failures, and a "Computer Design" magazine article in August of 1985 blamed Motorola and Intel for not drawing attention to the problem. This was in spite of over 80 papers, which had been published by that time including at least one by an Intel employee describing a system that failed every 4 to 10 minutes while arbitrating between two users. One of the more spectacular failures was the Texas Instruments Meta-Flops™ Bus Interface Logic family announced in July of 1988, which had been specially designed for synchronization, but without a proper understanding of the maths involved. Soon after release of the product line most of the parts were withdrawn.

Even now papers are still being published with "solutions" to the synchronizer problem, which claim to have perfect reliability, and designs are produced with unacceptable failure rates.

Now most synchronization circuits are buried deep inside the silicon, hidden amongst the other circuitry, and not much information has been made available to the users, who are unaware of the failure rates. With each generation of new products the synchronization problem gets worse. Some of the more sophisticated microprocessor chips now incorporated in hand-helds and cell phones have over 2000 synchronizers on board. Maybe the manufacturers know how often they are going to fail, and maybe they don't.

Chapter 8 Beware the synchronizer

System timing

Both synchronous and asynchronous computers can get into difficulty ordering events, but the way that it comes about is different. In a synchronous computer you don't have any problems once the input is synchronised to the clock, but you only have a limited time to do the synchronization, which means that the synchronizer is bound to fail sometimes. In an asynchronous computer everything can be stopped while a decision is made, there's no limit to the time available to decide, but because of that the decision could take forever. Bad news if you are in a plane about to fly into a mountain, and the navigation computer is unable to decide what to do. If the computer is synchronous, it sometimes crashes, and if it's asynchronous it lets you crash.

Which way to do the design? MU5 was an asynchronous computer, and in asynchronous computers is that the timing is hard to understand. Trying to organize what happens to an instruction in a complex computer is bad enough, especially when several things are happening at once. Then add the problem of never knowing the time that each stage in each instruction is going to take because it is all self timed, and you have a design and commissioning nightmare. On the other hand synchronous systems are simpler. The efficient working of the industrialised world relies on the rule of the clock. Trains and planes run to a fixed schedule, originally published nationally, and now globally. If you turn up on time, you catch your plane, and if you don't you don't get to where you want to be. Meetings are scheduled to the minute, and deliveries are arranged to be 'just in time'. Without a clock, it couldn't happen the way it does, but that doesn't necessarily mean that you have to use clocks in computers.

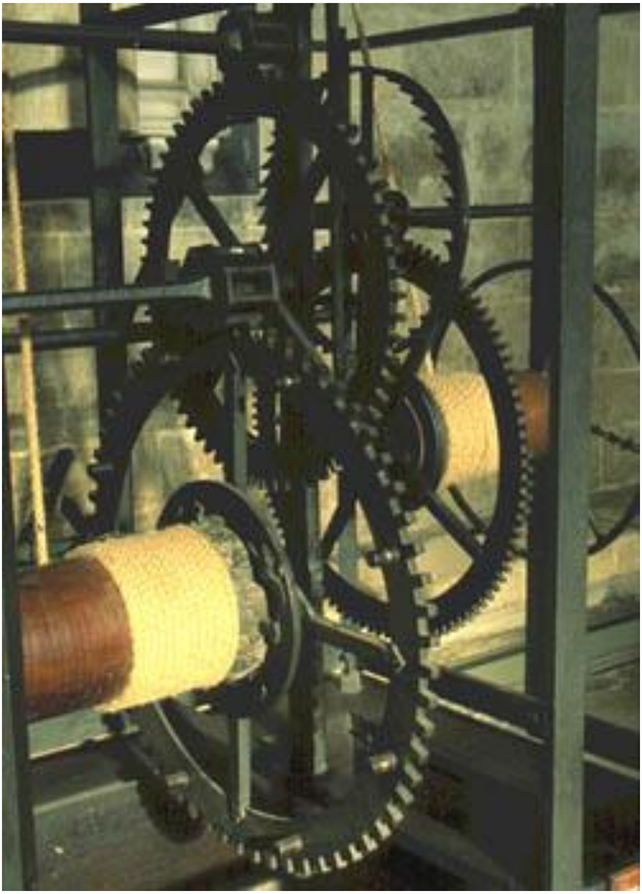
The tick of the clock

The industrial revolution seems to have been closely linked to the use of clocks as the means by which labour is organized. Before industrialisation, the church determined the pace of life in Western Europe. Monks rang the bells to call worshippers to church, telling them when it was time for the service. Great importance was attached by the church to regular hours of worship. This is shown by one of the stories in the early Medieval 'Life of Abbot Ceolfrid', which tells how the plague of 686 swept away all the choir monks who maintained the regular services, 'with the exception of the Abbott himself and one boy'. Because of this catastrophe, the Abbott decided to discontinue the normal observance of the hours. But after a week, and with the help of the boy, who became Abbott himself, and was later famous as the Venerable Bede, the two of them rang the bells, and restored the services, singing the psalms in full by themselves

The invention of mechanical clocks driven by weights in the later middle Ages made the monks life easier by enabling the bells to be struck on time and automatically, so that hours could be properly observed.

Each church clock in each village was set by the parish church to the diurnal rhythm, the rising of the sun, and its setting, and the sound of the church bells could be heard over the village fields, telling farm workers the time directly rather than them having to use the sun's position in the sky. Clocks were better than the sun, more precise, and not affected by the weather, but still weren't very accurate. Each village had its own time, with those in the west running a little later than those in the east because of the passage of the sun through the sky from east to west, but no one would notice, because you could not travel fast enough to observe the difference. In 1582 Galileo proposed using a pendulum to regulate the time shown by a clock more accurately, and later still much more accurate timepieces were needed to determine longitude for ships on the open ocean. In the 19th century train timetables compiled to the minute, and the linking of towns by telegraph meant that all village clocks had to tell the same national time.

The challenge of how to synchronize all the clocks to the national, or the international standard, was one that occupied the minds of inventors at the end of the 19th century. Patents poured in to the patent offices, particularly those in Switzerland, a country much concerned with clocks. A favourite method was to use the fastest method of signalling between the two clocks, a beam of light, in order to bring them into line. In 1902, Albert Einstein, a junior employee of the Swiss patent office in Zurich, worried about the differences brought about by any relative motion of the two clocks, showed how time and space were connected. He thought that each observer could have its own measure of the position and time of the other. His paper on relativity published in 1905 does not refer to any other scientist's work, and this almost obsessive internal focus has been said to have been the effect of Asperger's syndrome. More likely it all came from thinking about clocks.



In 1386, the first mechanical clock on record in Britain was installed at Salisbury Cathedral. It has no face and only strikes the hours.

In the 20th century with planes flying between continents, global stock exchanges and international television transmissions the whole world has become linked, and universal time is measured to million millions of a second. Clocks seem to be essential to our way of life.

The reason for this is because they make the organisation of interactions between people easier. Communication is simpler because events can be synchronised to the clock. If a meeting is arranged for 09:00 – 10:00 hrs, each person makes sure to get there at that time, and apologies have to be made if they don't. Usually that means that little time is wasted and the meeting can take place on time. If you know in advance when it is going to finish you can arrange other appointments later. Planning is easy. The alternative, to arrange a place for the meeting, and say it will start when everyone gets there, then finish when the business is done, is the asynchronous approach. Sometimes this makes things quicker, you might be able to start before 09:00, and finish early if the business is done, then go on to do other things sooner. Usually, the reverse is true. Life gets more difficult, and things take place at the pace of the slowest person. Imagine a train service in which the trains leave when they are three quarters full. In rush hours the service would be frequent because they would fill up quickly, but at other times they might stay in the station for a long time. The traffic on the tracks would be difficult to predict, and when you would actually arrive at your destination would be even harder to say.

Because of the ease of planning, where everything is timed to fit within the period between clock ticks, schedules based on clocks became dominant in computer design. One size, the time between clock ticks, was made to fit all operations, and the clock determined when everything happened. Now, to get a measure of how fast the latest PC will go, all you need to know is the speed of the clock, and this is why the advertisements for computers often announce the speed of their clock before they tell you anything else.

Synchronization

Though it's good to have a clock, it works best if everyone uses the same time. When they do, all the clocks are synchronized, but the world is not naturally synchronous. A committee appointed by the leading nations might decide on a global time frame, but it's all an illusion. The clock does not really decide the movements of the planets, and does not even determine every instant in our lives, even though we might think it does. People don't always arrive on time for their planes, and the clerk at the check-in has to deal with that difference between the body clock which didn't want you to get up early enough, and the airline's schedules. He has to synchronise you to the airline's time by making a decision to let you through if you are not too late for the plane, or stop you from boarding if you are. Even the airlines can't always keep to their own schedules. Planes get held up by adverse winds, and don't arrive on time, so travellers may fall out of the world of nice, simple synchronized schedules, into less than ideal hotel rooms and have to be rescheduled into new connecting flights. All the same the schedule works most of the time, and the rearranging your plans doesn't have to be done very often. The job of synchronization is mostly done smoothly, without too many fights as the check-in. Of course, the rule of the clock works best when all the events are well controlled and as many as possible are subject to the same clock. This is the way that most synchronous computers work. The clock says when every operation should finish, and new one should start, but the dictates of the clock can only extend as far as one computer. Connect it to another which has a different clock and there's a mismatch between the time the data becomes available from one computer and the time it can be used in the second computer. Ideally you could link the two clocks together, in the way that time is now linked in different countries, so that the well-organised traveller can travel reliably between them, but it’s not easy. What gets in the way is the time taken to communicate between the two clocks in order to adjust them. To be any use for communication they have to be accurate to a small percentage of a clock tick but this is comparable to the time taken by light to travel 30mm. If the computers are more than a few tens of millimetres apart, it becomes impractical. Even if it were realistic, there comes a point where for complete reliability of exchange of data, the user of the computer also has to be synchronized to the computer clock. Whenever any external input comes in, for example from the keyboard, there is no way that this input can be controlled by the internal computer clock. It’s typed in when the person doing the typing wants to do it. The computer clock cannot realistically control the world of the typist, and this causes problems.

If there are two worlds each with their own separate timing, there is no reliable way of synchronising them. The earth goes around the sun once a year, and it also turns on its own axis once a day. We can pretend that there are 365 days in the year, but it’s not true.

The number of days in one revolution of the Earth around the Sun is approximately 365.2422 days, but unfortunately it’s not a rational number. It’s not just that there are there infinitely more of these numbers than whole numbers, but it would take an infinite number of digits to specify the number of days in the year exactly. Any attempt to have rules about leap years so that the same date always comes at exactly the same time in every year is doomed to failure. That hasn’t stopped people trying, and the first major attempt to rationalise the number of days in the year was done in the name of Julius Caesar. In the Julian calendar a day was taken from the last month of the year (then February) to make the month with his name, July, have 31 days. February was further attacked by Augustus Caesar who also took a day from it to make his month, August, have 31 days as well. As some sort of compensation the leap-day every 4 years was given to February. They thought they had a system, which worked pretty well, and giving credit where credit is due, it lasted over 1000 years.

The Julian calendar had a year of average length 365.25 days. It also moved the start of the year from March 25 to January 1, supposedly the shortest day. A calendar year of 365.25 days is only slightly different from the actual length, which is closer to 365. 2422 days, but over the centuries the difference mounts up and by the 16th century it had become noticeable. Because the Julian year is longer than the actual year, the dates of spring, summer and autumn get earlier and earlier, at the rate of 1 day every 128 or so years. After 1300 years the calendar was 10 days out, and the spring equinox came on March 11 rather than on March 21.

If that was allowed to continue, the date of Easter would end up in mid winter, so the authority of the church was called upon to determine God's Will. Pope Gregory's 1582 reform to the calendar altered the rule for determining if a year should be a leap year by removing leap years for all centenary years unless they were divisible by 400, and dropping 10 days. Not many people can remember that, but it makes the mean length of the calendar year 365.2425 days, and is much closer than the Julian calendar to the actual number of days in the year. Now the error is 0.0003 days per year instead of 0.0078.

A catholic pope's reform to the calendar could not possibly be accepted in Protestant England, so the reforms were not adopted until 1752 in England, by which time the error had crept up to 11 days. When the reform could not be put off any

longer there were riots. People were concerned that 11 days were going to be eliminated from the calendar, and thought that their lives were being shortened to satisfy the papists. Mobs protested in the streets “Give us back our 11 days”.
Even after the reform, the error builds up at the rate of 3 days in 10,000 years, it hasn’t gone away. Remember that to express the number of days in the year as a ratio between whole numbers (which is what the calendar is attempting to do) is impossible unless you have an infinite number of digits to do it. So reform will have to come again and again if we want to keep pretending that a fixed set of rules exists for keeping the earth and the sun in step.

At the frontiers of timing

In a synchronous system it’s only the actual synchronization of data to the clock that can fail. Input to the first computer systems came from peripheral devices like card or paper tape readers, which were (relatively speaking) slow. Within the computer all the data going from one place to another would be controlled by the clock, which in the 60s and 70s might operate at 1 MHz or 1 million ticks per second. Inputs at tens or even hundreds of events per second still only need to be read every 10,000 clock ticks, and it doesn’t really matter whether the data is processed on one particular tick, or the next, but because the timing of the input data is not linked to the clock, it could occur at any time between one clock and the next. The two timing worlds interact only at their boundaries where data was passed from one to the other.
How does this cause a problem? Imagine catching train scheduled by the railway company. Their schedule is like the system clock. You haven’t checked the timetable because the trains run pretty often, and you have plenty of time, but when

you get to the station, there's one there, and it's just about to leave. Do you run to catch it or wait for the next? If you run to catch it you might save a bit of time, but it's a lot of effort. The alternative is to let it go and wait on the platform for the next train. Depending on the distance between you and the train, and when it starts moving, you have to make a choice, and you don't have much time to make it. You could end up running down the platform holding on to a carriage door still unsure whether to jump on or not.

Every so often, this scenario ends in a nasty accident, and that's also what can happen to the data being passed from the card reader to the computer. It has to catch the next clock tick, but if there's one already there when it arrives, it can end up half in and half out. A mangled piece of data is no good to anyone, and though it doesn't matter whether it catches one tick or the next, it has to be all on one or all on the other. To sort this out we have to have a good synchronizer, to ensure that the incoming data is synchronized to the clock. Unfortunately this bit of circuitry involves deciding whether the time of arrival of the data (an irrational number) is before or after the current clock tick, and we have to do it before the following clock tick comes along

Normally, the circuits that do the synchronization can come to a decision in much less than a clock period. They only get stuck in indecision, or metastability, for longer than the clock period if the time between data and clock is very small. This doesn't happen very often, so there's might only be a 1 in 100 million chance of any particular data item getting into trouble, and if there are only 10 data items per second being read, problems will only occur about once every 3 months, much longer than early computers stayed up and running normally, so in the early days no one noticed if the computer crashed as a result of data corruption in the synchronizer. But as computers got faster things got worse. The more data items in a second, the more likely becomes the crash, and the faster the clock rate the more likely the synchronizer will fail.

By the close of the 20th century, computers were talking to each other at rates of 10 million bytes per second, and clock rates had reached over 100 million ticks per second. If you try to do synchronizations a million times more often, the synchronizer fails a million times more often. If the synchronization circuits hadn't also got faster things could have been much worse, but even so, the sheer number of synchronizations per second makes reliability poor and a computer which fails once every 15 minutes is no longer acceptable.

Something has to be done. One improvement is to detect errors in the data as it arrives. Communication networks are designed to check the data when it gets to its destination, and ask for it to be retransmitted if it's wrong, but the ways in which data can be corrupted are infinite, and no error detection scheme can cope with that. A better way of getting round the problem is to allow the synchronizer more time. You don't necessarily need decide within one clock tick whether to accept a piece of data or not. If the data rate is quite slow, you may be able to take two or even more clock ticks to decide. That makes an enormous difference to the probability of failure. A one in a thousand chance of problems becomes one in a million. You need time in the station to decide whether you can get on the train, and the longer the time the less likely you are to fall off.

The downside is that both re-transmission and longer synchronization times involve waiting longer to get your data from A to B. Either you don't wait, and have to retransmit, or you spend two clock ticks synchronizing instead of one, the penalty for reliability is time.

Smaller and smaller, faster and faster

Every year the size of the transistors that can be made on a silicon integrated circuit gets smaller, and more can be put on the chip. Starting from simple circuits containing about 100 components in 1966 chips with over 50 million transistors were common in 2000. Relentlessly, every 18 months the capacity of the circuits put on the chips has doubled for more than 30 years. Not only has the capacity doubled every 18 months, but, at the same time, because they became smaller, the transistors got faster. In 2000 the smallest dimension of a transistor, its gate length, was around 0.2 millionths of a metre, small enough to fit 200 in the breadth of a hair. That sort of progress can't be sustained for ever, but if it did, in another 25 years we could be seeing transistors of 5 thousand millionths of a metre. (Or rather, not seeing, because this is much smaller than the wavelength of visible light). It's unlikely to happen because these dimensions would be comparable to the spacing between atoms in a crystal, so the idea of a transistor that is built on a conventional crystal surface breaks down, and maybe progress beyond 2015 would be much more difficult.
Even so, organising a million million transistors to work as a single coherent whole is a problem that will have to be solved well before 2015, and one the problems associated with organisation is scheduling the timing of what goes on and where.

When the clock ticks as fast as 10 thousand million times a second there's only 100 pico seconds (100 million millionths of a second) between the ticks. Synchronous computers rely on the tick being heard everywhere on the chip at the same time, or at least approximately the same time. To transfer a piece of data between two adjacent processors the accuracy of the timing needs to be much better than a single tick. But we are now up against the speed of light. Going at the speed of light a tick can only travel 3mm in 10 picoseconds. That's less than the size of a chip, and the clock ticks actually travel slower than that across the surface of the silicon. It's no longer possible to be sure that the common clock tells the same time everywhere, sometimes there's apparently more than enough time to move data from one place to another, and sometimes there seems to be no time at all.

The poor accuracy of clock distribution means that we can't have a global clock any more; the chip has to be split up into smaller zones where it is possible for an accurate clock to be distributed. Then, every time data goes from one zone to the next it might have to be re-synchronized to each local zone clock. With millions of things happening every second, in millions of places, every synchronizer has to be super reliable for the system to stay error-free. To be sure each of them is reliable enough, it could take several clock ticks to do the synchronization every time data passes into a new time zone, and the passage of data across the chip will get very slow if it has to pass through many clock zones, and even then, there's still some fundamental unreliability. We are used to thinking of computers as deterministic, with a given set of inputs, the output is predictable, but this is not and never will be the case. There's always the chance, however small, that the data could become corrupted in a synchronizer, and there are several other ways, like stray radiation, that could do much the same thing. These mechanisms get more important as the size of the switching devices goes down the energy involved in taking a decision gets less, and there are more decisions being taken. If synchronous systems don't work in the future is there another way? The alternative is to go back to the asynchronous design style. No clocks, no synchronization, no unreliability. We just don't know when the calculation will be finished

A world without clocks

To understand the difficulties, imagine a world without clocks, there would be no alarm clock to wake you up in the morning. You might get up as soon as you wake up, or if there is not much to do at work, lie in before setting off. Commuter trains would not follow a timetable, but would have to run according to the number of passengers. Only when one is full enough can it set off, and if there is not much demand, you will get fewer trains starting. When the train gets to a station, it stops, and when people have finished getting on and off, it starts again. Meetings could only be scheduled to take place at a particular location and not any fixed time so the start has to be determined by the organiser by phone, or email, with no watches, there could be no precise timing for the start or the finish, when the key people get there, the meeting starts, then it carries on until the business is finished. Shop would never completely shut, without a clock how could they show when they were open next? When a customer knocks on the door they would have to open, or tell the customer to go away. Large shops would man the checkouts according to the demand. This does not need clocks, neither do cinemas that run the same film again and again, when the film finishes, it is started again, so people can turn up when they like.
Radio and TV programmes could not run to fixed times, each programme would have to follow the previous one, as soon as one finished the next would come on. How would you know when to switch on? Well, many magazine type programmes already are like that, a collection of small items put together, and a resume of what will be in the programme is given at natural breaks in the schedule. You know the order, but not the exact time of one item. If you really want to see a particular item on the TV and only that one, a signal can broadcast that it is about to start, so that your video recorder starts up at the

right point, then stops at the end of the item. You look at the tape when you have time. No clocks needed, just the order in which things will happen. With no advance warning, difficult choices may have to be made. Football matches start when enough people have turned up, and finish when one team has got two (or some other number) goals.

Mostly it can be made to work, but sometimes you need to decide between two more or less simultaneous events. The problem comes when two people want to use the same resource at the same time. Two people arrive and want to use a public telephone. Who goes first? The train is scheduled to set off when there are 10 people on board. There are already 9 on board, one more appears, with another just behind. Does the train go with 10 people, or 11? The situation is not quite the same as the synchronization problem where you may have to run down the platform trying to get on the train, instead in an asynchronous system you may have to fight with the other people for priority in getting on, but the train will wait.

The difficulty occurs because of the two or more individually timed worlds, like the earth and the sun interacting to define the number of days in a year, or in this case two different people who both want to get on the train. In an asynchronous system you hold every thing up while the individuals concerned make up their minds who goes first. The train can't leave the station until everyone agrees they are either going to get on or not. It could take forever for then to decide, but it usually doesn't, and they don't always end up fighting. It's not a solution that can be applied to a train schedule fixed according to the clock, the train has to leave on time, whether or not you have got on it. In the asynchronous system time is lost while priorities are decided, while in the synchronous system you must allow more time when you get to the station in case you have to decide whether to get on or not, but the bottom line is that neither of them is a perfect solution.

Fundamentally all computers are just a little bit unreliable, or unpredictable in their timing, and often both of these things.

Chapter 9. The casting of lots

On a knife's edge.

If you build a house of cards, the higher you get, the more unstable it becomes, until the slightest tremble of your hand in placing the cards will bring the whole thing down. Similarly, the unstable scales, balanced on a knife-edge, are affected by even the smallest thing. What effect does the vibration from a distant storm, or the soft tread of a spiders foot on the beam of the scales have on the time taken for the scales to tip, as they surely will? Is the tipping quicker because of the distant thunder, or not? Decisions are affected by chance, even if the way they are taken seems straightforward and incorruptible. For example, asking the opinion of three experts on a simple proposition – is Paris the capital of France? – this seems to be bound to lead to a clear-cut answer. Two experts can disagree, but having a third will ensure a majority, one way or the other. Even if the answer is wrong the process seems robust. One obvious source of weakness is that the third might take forever to decide on an answer, and while he or she is deciding, the outcome is on a knife's edge where seemingly trivial things can have an influence, and then determine the outcome. Even the final answer, yes, or no, may be given by the spokesman of the expert panel at a moment when you are distracted, and misheard. This seems a remote possibility, but the more people you ask, the more likely it is that a few of the answers will be misheard, or biased by outside influences. Can this really affect the result? - Certainly.

The America presidential election

In the US the result of the election for President depends on the number of votes in the Electoral College. Each state ballots its citizens to decide how its votes in the electoral college will be cast, and the number of votes each state has is determined by its population, Thus, Vermont has 3, and New York has 33. Usually all the votes from a particular state go to the same candidate, an almost foolproof system, you might think. True, it's possible for one or more states to have a tied vote, but it's not very likely. Similarly it's just about possible for the electoral college vote to be split down the middle 269 votes out of 538 for one candidate and 269 for another, but before that happens the effects of random disturbances, or to use a more technical term, noise, start to affect the result.

November 2000

In the millennium year, 2000, the Electoral College vote was so close that a change in the results of any state would have swung the election (271 Electoral College votes for George W Bush and 266 for Al Gore). Although Gore got more popular votes than Bush, in America only the Electoral College chooses the President, so Bush was declared the winner, but before that happened there was a long drawn out and acrimonious dispute about the legality, or otherwise of the voters, the counting machinery, the size of the ballot papers, and, for some, the parentage of the lawyers involved in the proceedings.

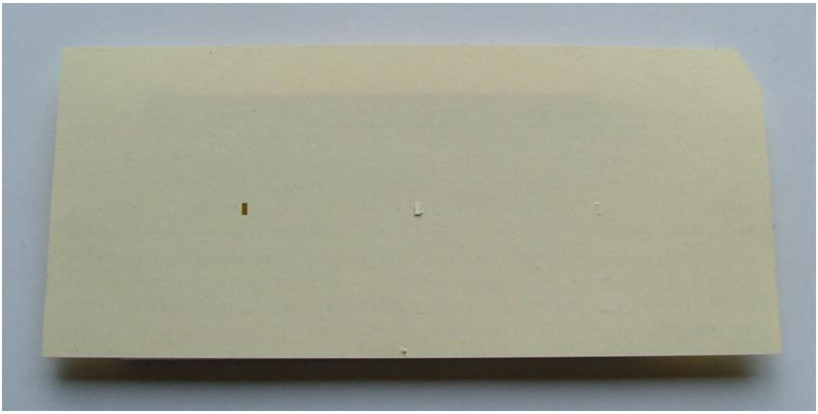
November 7, 2000 was Election Day, and shortly after midnight EST, it seemed that the election was going to Bush so at about 3 a.m. on November 8,the Democrat Al Gore telephoned Republican Party candidate George W. Bush to concede. At that stage Gore was ahead in the popular vote, by not quite one quarter of one percent of the 100 million ballots cast, it was in the balance. Then Florida was called for Bush. With 25 Electoral College votes Florida was the state that would decide, but Bush had only won it by a very small margin. So about an hour later, Gore retracted the concession because Bush's majority in Florida was slim enough to trigger an automatic recount. The Republican Party presidential ticket had received 2,909,135 (48.8%) votes and the Democratic Party presidential ticket 2,907,351 (48.8%) votes; a difference of less than 2000 votes. The automatic recount then produced the even smaller margin of less than a thousand votes, and the Gore team insisted on hand recounts in three large Democratic counties. All day reports had been coming in of voter confusion in heavily Democratic Palm Beach County, where a well-meaning election official was trying out a new ballot. In order to make the ballot easier for seniors to read, the font size had been increased, so that all candidates in the presidential race could appear on one page. This was later known as the 'butterfly ballot' where the layout was in two columns, with the punch holes down the centre, staggered in such a way that the second hole on the ballot, meant that a vote for Buchanan, whose name topped the second column, could have been confused with that for Gore, whose name was in the second place in the first column, but whose punch hole was the third down the centre. When the votes in Palm Beach County were counted, Buchanan, a right wing candidate, had more than 3000 votes, far and away his best showing in even the most conservative counties in Florida, and over 19,000 ballots were disqualified because of double punching. (The way the holes lined up, some voters thought a hole had to be punched for both Gore and his running mate, Lieberman, whose name, appearing directly under Gore's, was approximately even with Gore's punch hole. Some voters who may have thought they were punching holes for both Gore and Lieberman actually double punched for Buchanan and Gore, and thereby invalidated their ballots).

Now the lawyers were let loose. Manual recounts were requested by or on behalf of the Gore campaign on November 9 in Palm Beach, Broward, Miami-Dade, and Volusia counties. On November 11: Bush commenced a federal lawsuit to stop the recounts because of alleged equal protection and other constitutional violations which then triggered a host of action by other lawyers. By late November there were no fewer than 24 separate court cases pending related to the election in Florida, with appeals to the Florida Supreme Court, the Federal Circuit Court, and the U.S. Supreme Court.

Among the arguments on the Republican side were the number of overseas ballots missing postmarks or filled out in such a way that their validity under Florida law was disputed. The largest group of disputed overseas ballots were military ballots, which the Republicans argued to have accepted. Another issue was the television news late on the night of November 7 which called the state for Al Gore, while voters in the western panhandle (which is in the Central Time Zone) of the state were still voting, potentially depressing the voter turnout. This region of the state was mostly Republican.

The Democrat complaints included the number of voters missing from ballot lists who were unable to vote. These persons were disproportionately African-American and Democratic voters, as were the number of voters incorrectly listed as felons and whose votes were not counted.

First one camp and then the other to gained the upper hand, only to lose it in a subsequent ruling. The drama was heightened by Bush’s running mate Dick Cheney's mild heart attack and angioplasty procedure; a police-escorted convoy of ballots to the state capital in Tallahassee, and a near-riot by Republican recount observers in Miami-Dade County. Amidst the claims, counter claims and recount deadlines missed, the recounts staggered on through November, but they were not proceeding smoothly.



A punch card

Chads in stages of gestation cleanly punched, hanging and pregnant

Double punched ballots can be disqualified; receipt of postal votes may be subject to a deadline, but what to do about the infamous 'chads'? These are the tiny bits of paper which drop off, or are supposed to drop off, when a voter punches his or her ballot, but whether they actually drop out, or not, depends on how hard the voter pulls the lever, and the vagaries of the punch. There are hanging chads, where all but one of the four corners has been detached, swinging door chads, with two corners still intact, tri chads, with only one corner detached, and even pregnant chads, those pierced by the stylus but with all corners intact, and dimpled chads, ones where an indentation has been made but no hole punched. What was the voter's intent? How far does a chad have to be detached or dimpled to be sure that it counts? And what if an object falls on a hanging/pregnant/dimpled chad during the count causing its validity to change? Back to the lawyers. Judge Sanders Sauls ordered about 14,000 disputed ballots from Palm Beach and Miami-Dade counties to be brought to him in Tallahassee. What he did with them is not clear.

In the end, a 5-4 decision on December 12 by the United States Supreme Court ended the legal wrangling and certified Florida's election results, resulting in Bush's victory. Seeing no legal recourse from the U.S. Supreme Court ruling, Gore conceded on December 13.

Because of the closeness of the vote, the decision in Florida had actually taken over a month to make, where it would normally be done in a day or so, but just as interesting were the 14,000 disputed votes, about ten times the finally agreed majority. The intent of the voters was buried deep in the snow of random noise.

The throw of the dice.

Usually we would prefer a decision not to be a hostage to fortune, but sometimes the aid of fortune is deliberately sought. A common way out of a dispute is for the final arbiter to be that of chance. If there is a decision to be made, and to be made quickly, where all the circumstances seem to be equal between the courses of action, we revert to the toss of a coin. Who starts in a game of two sides, which team plays at which end, and so on. This does not mean that the decision can *always* be done quickly, the coin could land on its edge, and drawing straws from a closed hand has the possibility of the short straw being one of the last two. The person choosing one out of the last two straws is still faced with an indifferent choice. All the same it's usually quicker to do things this way than have an argument. In fact, sometimes choice by random selection can be built into the law of some countries. The New York Times of January 12 1959 reports that chance is the arbiter prescribed by Swedish law for breaking the votes in Parliament, though it has apparently never yet been necessary on any major issue. In the US itself, on the admission of Hawaii as the 50th state two new senators were elected, and the senate had to decide which of them had seniority, and which would serve the longer term. The first decision was made by the toss of a coin and the second by drawing of cards. (It's not reported how they decided who drew first)

The person credited with being the first to suggest in print that random selection devices, in his case, Lots, be used as a means of resolving the problem of indifferent choices was Thomas Gataker (1574-1654). He was an English scholar and cleric who published his study "Of the Nature and use of Lots" in 1616. In it he criticizes the view that "a Lot discovereth to men God's hidden will", saying that "Lots are not to be used in a question of fact, past and gone but for that is no ordinarie Lot able to decide, but where some question is who hath the right to a thing; in which case, notwithstanding the Lot is not used to determine who in truth hath the right to it, but who for peace and quietness sake shall enjoy it". He is firmly of the opinion that "concerning the matter or business wherein Lots may lawfully be used, the rule of Caution in general is this, that Lots are to be used in things indifferent only". His career in the church suffered as he appeared to the puritans to be favouring games of chance.

Thermal noise

Only at absolute zero, -273 °C, are the atoms in a crystal motionless. They are locked into the crystal lattice, each bonded to its neighbours. As the temperature rises, they begin to vibrate randomly, moving back and forth, up and down, and side to side. The hotter it gets, the more they move, until eventually they break free of their bonds and the solid becomes a liquid. Still loosely attracted to each other the atoms can flow more or less independently until they get hot enough to escape completely from the liquid and the liquid has become a gas.

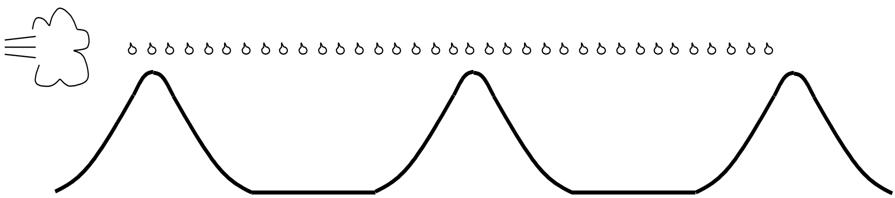
All matter is subject to this kind of thermal motion, and nothing is completely still except at absolute zero. In any conductor or semiconductor the thermal motion of the atoms and their electrons shows itself as a small random current or voltage variation, known as thermal noise. Noise is the stuff that looks like snow on the screen of a TV set when the signal is low. The TV tries to make the signal bigger by amplifying it, but if it's only just bigger than the thermal noise at the antenna, the noise gets bigger as well, and there's nothing more that can be done (apart from getting a bigger antenna)

Noise is not at present large enough to cause any problems in a computer, since the noise voltages are about a thousand times smaller than the voltages used to carry the 1s and 0s around the system, but it matters when a flip flop is in metastability, because then the very small random voltages can knock the circuit off balance, and determine the final outcome.

Noise in a synchronizer

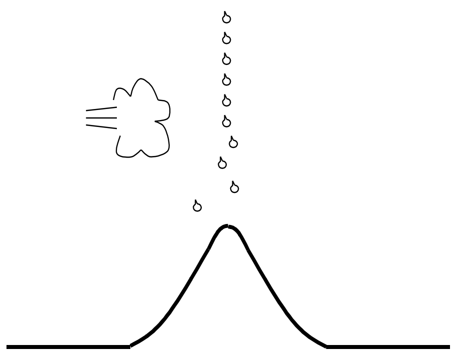
A person typing a large document into a computer is unaware of the clock ticking at many millions of beats per second, but

every time the ENTER key is pressed the synchronizer must choose which interval between clocks in which to accept the input. Each piece of data has to fall between one clock tick and the next to be safely accepted. This is like a single drop of rain falling onto ranges of mountains. If it falls on the peak of one mountain range it will flow down either one side or the other into a valley. Which valley it ends up in depends on whether it falls just to one side or the other of the mountain range. In a synchronizer, each clock beat is similar to a mountain range, so because the clock ticks continuously, there are mountain ranges as far as the eye can see, and each input is a drop of rain, which eventually ends up in one of the valleys between the ranges.
Most of the drops will fall straight into the valleys, or on to the side of a mountain range but a very few will hit the very top of a mountain, and it will be some time before they move one way or the other, to the valley on the left or on the right.
Because we don't know when the clock ticks will happen, we can't time our inputs to miss the exact time of the tick, and even if we did know, they are so fast we would not be able to respond to them, so the thousands of inputs made when you are typing are like a shower of rain drops thrown evenly on the peaks and the valleys alike. If a gust of wind blows, knocking one drop off its course away from a peak, it is just as likely to blow another one closer. The mountain peak gets just as wet no matter how hard, or how randomly the wind blows.



Rain on mountain ranges. The peaks are the clock beats, and each drop represents and input, each of which must end up in one of the valleys.

The wind is equivalent to the thermal noise in an electronic device, ubiquitous but unpredictable. So in a synchronizer the internal thermal noise makes no difference to how often an input lands exactly on a clock beat, thereby causing long decision times, some are pushed away from the tick, and others towards it, but the number that come very close to the tick is just the same as it is without the noise. But it does make a difference to each individual input, because it's no longer possible to predict which clock tick that is going to be used to process it, this one or the next. The noise can push one earlier and another later, at random. Some raindrops might end up in one valley or another, but we can't say which if it's windy.
It's different if all the drops are aimed straight at the mountaintop. Then they always land in the same place and the peak of the mountain gets very wet. Then they all take a long time to get to the valley floor. Unless, of course, there is a gusty wind, in which case the peak gets less wet, the sides get wetter, and most of the water doesn't take as long to get down the mountain.



Trying to make all the drops land in one place is difficult.

This second case is like building a very unstable house of cards. If there's no wind you have a better chance of putting the cards in exactly the right place, so that it stays up for a while, and if you do it often enough you might get it to stay up a long time. If it's windy, though, it's more likely that it will fall down quickly than it will stay up for any length of time. These are two situations are quite different, one where the chance of an indifferent choice is low, (the rain drop can be anywhere between the mountain ranges), and one where it is high (it hits the peak every time). If you always have an indifferent choice, for example which football team should kick off first, it makes sense to add some random input, like tossing a coin, otherwise you might never be able to start the game.

The ghost in the machine.

The amount of noise in a synchronizer is only enough to make the processing path for about one in 10,000 inputs unpredictable, but as the size of the transistors gets smaller, the effective size of the noise gets bigger, and we may not be able to say what happens in as many as 1 out of 1000 synchronization events. Now add that to the fact that there could be over 1000 synchronizers in a system as small as a mobile telephone, and you can appreciate that while we may know in general what the system is doing, we might not know at any particular time exactly what is going on where. The system is non deterministic. Noise can also turn apparently discrete inputs into continuous ones. If the inputs are numbers of votes for one candidate or another, then they are discrete quantities, and it should be possible to make a quick decision about who has won. But votes that are sometimes there, sometimes not, and sometimes halfway there like pregnant chads, turn the discrete measure into a continuous one. It's a matter of judgment whether the vote is valid. In systems which use electrons to carry signals from one place to another, thermal noise also smears over the discrete nature of flows made of one, two or three electrons, and makes them appear to be a continuous measure of current. Only at absolute zero can a perfectly discrete arbiter be built, and nothing much happens at absolute zero.

It is the aim of most engineers to build systems that are always predictable, and understandable, though there are occasions where random number generators are useful. Two examples where they can be useful are lotteries, and encryption, where the aim is to produce results that are either unpredictable, or at least completely incomprehensible to the outside observer. It's certainly possible to do that with a simple flip-flop, which normally responds to its inputs in a deterministic way, but if the inputs are deliberately tied together so that the flip-flop is in balance, and then let go, the output could go either way. It is non-deterministic, responding only to its own internal noise, and not to its history.

It is a circuit with a will of its own, listening to no one, and remembering nothing, it can make decisions that no one can hope to understand.

Chapter 10 The Moral maze

Making things simpler

Life is complicated. We all have to do the best we can to make our way in a world where choices have to be made every day between a bewildering array of alternatives. Even buying a tube of toothpaste is difficult when the supermarket has rack full of hundreds of different kinds, different brands, with different ingredients, in different sizes, at different prices. If you know exactly what you want, and where it is, there's no problem, but suppose they've run out, and you have to find an alternative it may not be easy. First you might have to look at all the possible substitutes. The more there are, the longer it takes to look at them all. The more there are the smaller the differences between one brand and another, and the longer it takes to make up your mind even when you've narrowed down the choice. In the end you may be reduced to choosing the red package rather than the green because you like the colour. Choosing the right groceries may not matter very much, but with lots of possibilities to consider it takes time to do the shopping.

In the old days, (the ones where the sun shone every day, and people were always polite), there wasn't so much choice, so life was simpler. There was only one sort of toothpaste, you either bought it or you didn't. The natural solution to complication is to simplify. How to cut out the stuff that doesn't matter so that decisions can become easier. But by simplifying, you may lose detail that can be important, for example you may need to read the list of ingredients on the pack because if you are allergic to one ingredient in the only toothpaste available, you can't buy it. The other problem is that most things in life do not involve a simple choice between easily distinguishable things like apples and pears, it’s apples at one price, and pears at another, so fine gradations in price can push the decision one way or another.

One the one side there is the urge to simplify, to make life easier, to make decisions quicker, and on the other, in the real world you may have ignored the crucial factor that makes all the difference and thrown out the baby with the bathwater. The inviolable principles that are used to discriminate between one course of action and another are rarely inviolable, and even if they were, the noise of misunderstanding often intervenes to blur the boundaries between one thing and another. It seems dramatic to say that fundamental moral principles can be brought into question by having to make a choice, but when it is a matter of life or death, it can happen.

A matter of life or death.

On October 1 2003 a High Court Judge ruled that Natalie Evans could not complete the *in vitro* fertilisation treatment she had started with her partner some months earlier. For Miss Evans the six embryos held in cryogenic storage represented her last hope of having more children of her own. After trying for 18 months with her partner, she had gone for medical checks, and was offered IVF. When her eggs were harvested Natalie's ovaries were found to have potentially fatal precancerous cells and were removed, destroying any chance of future conception by natural means. Her only chance of pregnancy was by using the embryos resulting from the IVF treatment. But in June 2002 her partner split with her, and refused to agree to the use of the embryos, so her only chance of pregnancy had disappeared. She offered to sign any document he wanted in which she would agree never to claim maintenance or involve him in their child’s life. He asked for time to think, but eventually said that the embryos should be destroyed, and the court ruled in his favour, not hers.

This true story raises a whole series of questions about moral rights and wrongs, which are not easy to resolve.

- Can it be right to destroy the embryos? Surely this is the taking of human life, and must be forbidden.

- In a marriage the husband and wife have may have a joint moral responsibility to look after any children. Even after a divorce, the husband is expected to continue to contribute towards their maintenance. Can the female partner insist on his duty to contribute towards children unborn at the time of the split?

- Should the rights of men and women be considered equal in this case? How would the two parties feel if the situation was reversed, for example, suppose the two partners had succeed in conceiving a baby before splitting up, and then the man had testicular cancer. Should it be possible for him to insist on her carrying the child to term, even if she did not want to, because it would be his only chance of a child?

These are just some of the questions, which are at least partly about when a human being comes into existence. Is there a moment before which there is nothing, and immediately afterwards there is a person? Catholics, as well as many other religious and moral authorities, believe that conception is the moment of ensoulment, when the new genome is created that will become a new human being. The US Government has taken this view and will not fund research on stem cells if new embryos have to be destroyed in order to extract them. Others argue that there is in fact a process that lasts between 24 and 48 hours between the time that one or more sperm penetrate the egg, and the complete control of that egg by one new genome. Sometimes two eggs are fertilized, and instead of developing into twins, fuse together in what is called a Chimera, where one person has two genetically different sets of cells. Does this person have one soul or two? More importantly, this new entity is in the first few days wholly dependent on it's mother to develop into a person. More than half of all eggs do not implant in the uterus and will never become a human being. So it's not clear when this blinding flash of ensoulment actually occurs, and the law has to deal with the uncertainty. At one end of the moral spectrum are those who hold that it is a woman's right to choose to terminate a pregnancy, up to the point when the foetus is viable outside the womb. At the other end, even contraception is a sin, as it seems to be an impediment to the natural process of conception, which starts with a couple's agreement to have sex.

Taking the view that all embryos are human, even those of just a few cells, they cannot be left to die because that is murder. It follows that men and women must do everything in their power to preserve life. In which case Miss Evans would be obliged to enable all six embryos to become children.

Taking the other extreme, it could be argued that the embryos are not yet human, they are her property, and she can do what she likes with them. But why are they all her property, and not at least partly her partner's? Does he not have some rights in the matter?

Neither of these two simplistic arguments produces a completely satisfactory solution, and the central question, of when a human being gets its soul seems to depend on time. It can be anywhere between the parents agreement to have sex, or create an embryo by IVF, and the day of the child's birth, so there is always uncertainty. Even if you could get agreement on an exact number of hours that the soul arrives after a sperm has been injected into an egg by IVF, the decision is not much simplified. Suppose a technician drops the test tube on the floor at just that moment. Is it manslaughter, or just a failed experiment?

Making the difference.

Suppose you have won a modest sum of money on the lottery, more than enough for your immediate needs, and want to give some away to your family and friends. How do you decide how much to give them? Leaving aside the question of who's your friend, and who isn't, you might decide to divide the sum up and give everyone an equal amount. But you know that Sarah Jane will probably invest wisely while Billy Jo will blow the lot on drink, drugs and fast cars. What to do? Perhaps keep back Billy Jo's share until he demonstrates he is a reformed character. And you have to tell him just what he has to do to demonstrate his worthiness. Just such a problem faced the British Government when it decided on a policy of spending money to improve the Health Service.

One of their Big Ideas was that they would increase funding to National Health Service hospitals on condition that they would reform to meet patients' needs. In order to ensure that the best hospitals were indeed worthy of the cash, they set up a committee of honourable people whose task was to devise a suitable performance rating system. Not too complicated, so that it would be obvious who should benefit, and who should not (at least, not yet). Not too simple, otherwise they would be accused of insensitivity to the needs of the nation.

Hospitals were assessed on their performance in nine key areas:

1. Accident and Emergency admission waits (12hrs)
2. Cancelled operations.
3. Financial management.
4. Hospital cleanliness.
5. Improving working lives.
6. Number of inpatients waiting longer than standard. (12-15months)
7. Number of outpatients waiting longer than standard (21 weeks)
8. Total time in A&E.
9. Two-week cancer waits.

Some of these targets are a bit vague, but most of them (1, 6, 7, 8, and 9) indirectly involve measuring time, by counting the number of people waiting longer than a certain period. The final performance rating was based on a star system, three stars for the best hospitals, which would get independence and extra funding, down to no stars for the worst, cast into the limbo of not yet, or worse, the hell of public humiliation.

Three stars were only awarded to hospitals meeting all nine targets, and so inevitably, there were hospitals that missed only one target out of nine, with 0.01% of their patients being outside the prescribed waiting time. Doctor’s prestige and money are at stake here, so anger at the rejection, and guilt at their own part in their humiliation stalked the corridors of the rejected hospitals. It must be the system that was at fault.

“The ratings are ludicrous, and should be scrapped”

“They bear no relation to the quality of care that the patient is receiving”

“Waiting times, and trolley waits are important to patients, but what they really need to know is how successful a heart surgeon has been at a particular hospital”

"We have improved in all other areas, and only missed this one by a narrow margin"

The essence of the arguments against the system were that the performance rating system does not work because individual cases are judged unfairly (lack of detail), that the wrong targets are used (too crude a judgement), and that a good performance in one area cannot be set off against a poor one somewhere else (too black and white). All of these criticisms have some truth, but it is not true that a pass or fail judgement cannot be effective in making improvements. Evolution shows that individual animals in a species may die in circumstances that seem unfair, or through chance, but on average, the fitter individuals survive more often than the unfit, and the species improves. It does not matter what targets are set, some injustices will be done, but most hospitals will adjust their mode of operation to meet new criteria, either by reducing waiting times, or by other means, such as delaying the letters to patients informing them that they are on the waiting list. Of course, those that resort to sharp practice can be reduced by more careful definition of 'waiting time', but only at the expense of small print in the targets, which makes them harder to understand. Whatever the set of rules there must still be successes and failures, those that meet individual targets and those that do not, or there is no incentive to improve. The tension is between complex criteria, which deal with the fine detail, but need teams of lawyers to draft them, armies of administrators to collect the statistics, and well paid wise men to adjudicate the results, and the obvious injustices of simple tests that everyone can understand but don't deal with some cases (such as most targets met by a mile, and only one missed by an inch).

There is no set of rules that can ever discriminate perfectly between one waiting list and another, and the comparisons that are made are always overlaid by the noise in the system, exactly *when*, did each person join the list, was the letter drafted and on the clerk's computer, or was it delayed in the post? Success or failure will sometimes depend on chance and the postal system.

Grading systems like this seem to be part of life. Exams results are graded A, B, C etc. when the actual results are a series of marks, so a board is set up to make the decisions. (It's never entrusted to a formula – too crude and unfeeling). Anyone who has ever sat on an exam board knows that the members spend most of their time arguing over the borderlines between grades, trying to find reasons for separating candidates who have very small mark differences. Some must have an A, and some a B. There is usually a prescribed borderline between the two grades, say 70% and above for A, 60% - 70% for a B. There will be a discussion over whether someone with 69.9% should actually get an A rather than a B. Hard luck stories are listened to, the weather on the day of the exam is discussed, and the peculiar difficulties of the paper compared with previous years are aired. In the end the borderline is always pushed downwards rather than upwards, because the members of the board like to think of themselves as benevolent souls, which they may well be, and not hard uncaring people willing to do the students down. The argument usually hinges on the principle of "finding a gap", thus if there are students with marks of 69.5%, 69.8%, 69.9%, and 70%, there is a bigger gap between 69.5% and 69.8% than there is between 69.8% and 70%, so the top three get A, but the one with 69.5% gets a B. The overt justification of this is that we know the accuracy of marks in the examinations is unlikely to be better than 1%, so the top three 69.8%, 69.9%, and 70%, cannot be separated, and since 70% gets an A "by right" so do the other two. In fact this is little more than a scramble to avoid responsibility for making mistakes, since the board members, feel the fourth candidate can be seen as much less deserving that the other three. A mistake in awarding an unmerited higher grade will not be questioned, where a mistake in not awarding the grade might be. The illusory nature of the "gap" can be seen when you consider how many candidates took the exam. If there were only 100, there may well be a gap of 0.3%, between 69.5% and 69.8% with no candidate in between, but with 1000 people taking the same exam, the gap of 0.3% would probably disappear, and there might be 40 marks between 69.5 and 70%. Now a gap, perhaps of only 0.03% might appear in quite a different place, probably nearer 70%, and the student who got 69.8% might not get an A in the same exam, marked by the same examiners with the same weather on the day, just because more people took the exam.

The vagaries of any testing method are such that a considerable amount of noise is present, what the examiner had for lunch, the positioning of a full stop in a question, the presence of pigeons in the roof of the exam hall, all affect the outcome. No matter how hard we try to take all these factors into account, it is impossible to reduce the noise to zero. Why then, have committees to discuss at length all these issues? It may make them feel that they have made every effort to be fair, but it is unlikely that it actually has much effect of the fairness of the result, so why not just publish the marks as they stand?

We could also ask why there is a star system for grading hospitals. Times are continuous quantities, and at the level of millions of pounds, money is (fairly) continuous, so it would be possible to produce a weighted average of how far each hospital has got inside the targets, and then hand out an appropriate amount of money.

The answer lies in the nature of peoples understanding. The league table of none, one, two, or three stars seems easier to understand and administer than a continuous variable, like, say, percentage rating. Of course some people might say that in a privatised system the customers take their money to the hospital that serves them best anyway, so there is a more direct relationship between the service and the reward, and no need for star ratings. Others might say that if all hospitals get the same funds and provide the same services, there is equally no need.

Find the scapegoat

Any system of simple rules can be found wanting, even one as tried tested, and complex as the English legal system. An as yet unresolved case involves two youths in car driven dangerously by one of them. It mounted the pavement and killed two innocent bystanders. Both young men ran away, but were caught later, one within hours, and the next the following day. Someone has committed an offence, causing death by dangerous driving at least, and maybe murder. But who? Each of the occupants of the car says that the other one was driving, and the only witnesses are both dead. One of them is lying, but there is obviously no incentive for the liar to tell the truth, he would immediately be found guilty. The other one is probably telling the truth, so why should he change his story? The problem is that we don’t know which one, and so the prosecution can’t proceed. There’s no crime of joint responsibility for dangerous driving, and even if there were it would involve unfairness for the one who wasn’t driving. Picking one at random to try for the crime won’t help; he would almost certainly get off for lack of evidence. This is the classic impasse, one of them must have done it, but there isn’t enough evidence to decide which one. What happens in this case is that we must either wait forever, for something to turn up, or accept the possibility of injustice, by assuming equal responsibility.

Most of the time a simple code against which we can measure things is a useful guide, but in the hard cases some fundamental issues may have to be faced, and this, in itself can be a hard decision. Hard decisions can give random answers and could take forever to resolve.

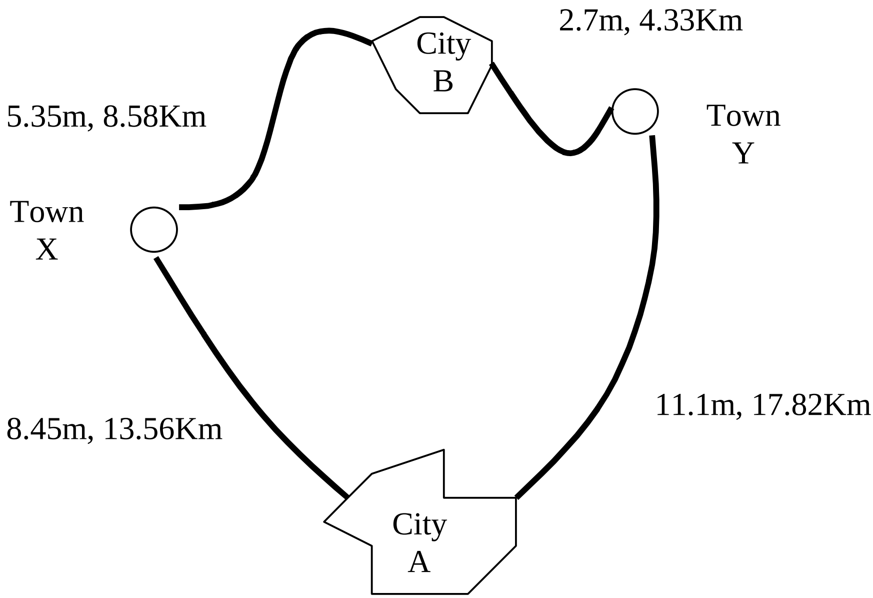
Chapter 11 The Search for Certainty

Understanding and simplicity

We don't always have time to look at the detail. Headlines are short and simple to catch our attention. Politicians talk in sound bites that we can absorb without thinking: "the average wage has gone up 5% whilst we have been in power", or "the majority of people have seen a reduction in take home pay since the present government took over". Every year a mass of new legislation is enacted aimed at closing loopholes in the existing law, making distinctions between different groups, and dealing with changing lifestyles. We can't understand it, we can't remember it, and we can't cope with the flood, so we just ignore it. People don't have time to get on top of the detail. They want simplicity. It's good, or it's bad. Usually good and bad statements can both made about the facts, and both can be true, but the politicians simply select the things they want in order to support their own views.

If 40% of the population earn £50,000, and 60% earn £20,000, the average is £32,000. An 8% increase in the £50,000 bracket to £54,000, and none for those earning £20,000 gives an average of £33,600, or an increase of 5%, so that one's true. Any increase in tax for the lower income earners will produce a reduction in their take home pay, and they are the majority, so the other one's true as well. It's all a bit complicated, and there isn't time to work everything out, so we just accept what the experts say.

Even if the facts are not deliberately selected to produce the answer you want, there can be conflict. Suppose you set out to go from your home in city A to the nearest city at B. You can go by way of town X or town Y, but you have two maps, one where the distances are measured in miles, and the other where they are measured in kilometres. On the first map the road from A to X is marked 8 miles, and X to B 5 miles. Going by way of Y, the map says A to Y, 11 miles, and Y to B, 3 miles. It seems that the quickest route is by way of X (13 miles) rather than Y (14 Miles). The other map tells a different story. A to X is marked 14Km, and X to B 9Km. Going by way of Y, this map says A to Y, 18Km, and Y to B, 4Km. If you believe that, it would be best to go by way of Y (22 Km) rather than X (23Km).



Who is to blame for this error? Surely one the mapmakers can be sued for misleading the public, or at the very least, confusing them. But the distance is actually much the same, and both maps are correct, it's just the simplification to an easily digestible figure that causes the conflict. 5.35m and 8.45m round down to 5m and 8m, but 8.58Km and 13.56Km round up to 9Km and 14 Km. Going by way of Y it's the other way round. Nobody really wants distances on a map printed to two decimal places, so there is a trade-off between accuracy and simplicity.

Moral codes and what they do

The confusion that results from simplification can go to the root of people's beliefs. Belief systems usually include a moral code in which a law is either obeyed, in which case the person obeying is correct, clean, pure and good, or disobeyed in which case sin is committed. Each act is an all or nothing affair and a clean line is drawn between heaven and hell.

There is a story about a priest teaching in a small catholic primary school in the suburbs of a city. He warned the boys that they would be taught many things when they went to the big school, some of which they should not believe. They might be taught that men had evolved from animals, and that the universe could be measured. "These things," he said, "are known to God, and it is a sin for men to think they can count the stars". A small boy stood up at the back and asked, "Father, how many stars can I count before it becomes a sin?" There was a pause. "One hundred" said the holy father.

Most people know that the merit of any particular act may not be measurable in black and white, or even in whole number terms. It may be partly good and partly not, but fundamentalists of all kinds insist on a strict division. Moral codes like those defining cruelty to animals are often defended (or attacked) in black or white terms. Killing people is wrong, because they have an immortal soul or are sentient beings. Is killing animals wrong? If they are sentient beings it must also be wrong. Are cats allowed to kill other sentient animals? Now according to the fundamentalists we must either condemn all cats to hell, or invent another rule. We might try one that says a cat is not a moral agent but a human being is. By this line of argument we would have cats, which are sentient but do not have the same concept of morals as humans in a sharply different category to people. Chimpanzees, who share 98% of their DNA with humans, frequently hunt and kill monkeys. Do they also share our ideas of morality? Do they have a soul? It seems unlikely that the answers could be packaged in a few words.

It is not that moral codes are not a good thing. To live one's life in a way that benefits others as well as oneself is a worthy aim, and some of the measures of a good life are universal to all mankind, and are unchanging because they are related to the human condition. But many of them are not. Not to kill another human being is a very common, but by no means universal rule. Things were different in the past, and even today, of 31 hunter-gatherer societies 64% fight their neighbours every two years. The morals of these societies are quite different from ours, with typically 40% of the men having killed or participated in killing. The Wari people of the Amazon regard anyone who is not a member of the tribe as edible. Cannibalism, which we see as anathema, was frequent in prehistory, and part of the religion of the tribes concerned, which fitted the circumstances in which they found themselves. Very recently, by contrast, the taboo on killing has been extended by some animal rights activists to a condemnation of any animal death resulting from human activity. Meat is murder.

Attitudes change with the circumstances. A man whose subsistence crop is threatened by rabbits has no problem in killing (and eating) as many as he can. Similarly a tribe in competition with another for the only available resources may not want to recognise their competitors as human.

A clear, consistent and simple moral code is difficult to find, especially if you are looking for a small set of black or white rules, because you will have to make decisions to map the real, shifting, continuous world into the discrete unchanging world of the rules. To some extent you can avoid oversimplifying by not having that kind of simplification. We don't really have to accept that we are fully responsible for out pet cat's actions in killing other animals. The cat may have its own morals, but they may not be the same as ours. It may be more strongly driven by its instincts, and less by it's intellect, and yet still is a sentient being with some degree of freedom of will. In short its soul could be in some ways lesser than ours, but still exist. It could be more wrong for people to kill people than to kill mice, and it could be fairly normal for cats to kill mice.

A measured response to particular circumstances avoids the sharp conflicts of all or nothing, but a code of morals saying "Thou shalt not kill (unless in self defence, and provided that the methods used are appropriate to the situation)" does not have the same feeling of moral certainty, or simplicity, as "Thou shalt not kill", and will therefore not be easily accepted.

This is one reason why moral codes are simple. Complex rules are difficult to understand and remember. Do we really have to carry around a list of species in order to remember which can be killed and which not? Another reason is that rules that measure the quality of virtue, responsibility or guilt, are subject to drift. Changed circumstances may slightly alter how blame or good behaviour is perceived, and then alter the balance of the choice that has to be made, so we have to ask whether a moral code is laid down by God and fixed for all time, or whether it is subject to reinterpretation.

The fundamentalists say that once you start to unpick one corner of a moral code you cast yourself adrift on a sea of moral relativism. But others, such as the Islamic Philosopher Abdolkarim Soroush argue that interpretations of religious knowledge can change over time, or can be understood in their historical context. He feels that science cannot progress under totalitarian rule. The ayatollahs are uncomfortable with such ideas, which they feel take away the sacredness of religion and make it subject to human understanding.

Observing that the many religious codes are not always compatible, some Western religious philosophers would go further than Soroush, saying that no religious beliefs are literally true, and that religion simply supplies us with poetry and myths to live by.

The tension between the fundamentalists and the progressives can be seen in most religions, for example recently the Church of England has been divided on the question of the appointment of openly homosexual bishops. Should it be allowed or not?

The apostle Paul writes (1 Timothy 3) "Our leader, or Bishop, must be above reproach, married to one wife, sober, temperate, courteous…". St Paul was probably unaware that a future candidate for election to the office of Bishop could have left his one wife for a homosexual partner, and that the progressives would argue:

"The Bible should be reinterpreted in line with modern thinking" and "The Church is hopelessly irrelevant fighting about homosexuals"

On the other hand, the conservative wing of the church, view all sex outside marriage, particularly homosexuality, as outlawed by scripture.

This kind of division often leads to a schism, the Orthodox want to keep to the old ways and the code laid down by God, but the Reformed faction break away because they want to redefine the code. Jesus was one such revolutionary who said that "You have learned that they were told, 'eye for eye, tooth for tooth'. But what I tell you is this: 'If someone slaps you on the right cheek, turn and offer him your left'" (5 Matthew 38-39). He was also of the view that "Plain Yes, or No, is all you need to say, anything else is the work of the Devil." (5 Matthew 37)

The Last Judgement by Michealangelo.
Christ, in the centre, decides which of the souls rising from their graves on the bottom left shall go upwards or down to the underworld at the bottom right

The extremes that fundamentalists will go to in the pursuit of certainty is illustrated by the story of the Rabbi who kept to a strict kosher diet. He could only be sure that the food was Kosher if he had prepared it himself, so he made a rule to only eat Kosher food that he had made himself. His goodness was such that God Himself invited him to come and eat a meal in heaven, but he would only drink water saying, “how can I be *sure* that the food is Kosher?”

Why do people adhere to such, simplistic beliefs when they so often lead to contradictions? One answer is that the rules are simple because they involve only this, or that, and nothing in between, so they are easy to understand, and to follow. Another answer is that applying the rules admits you to the flock, or alternatively marks you as an infidel, to be condemned by those of the true belief. One of us, or one of them, we all have a need to belong. There is nothing in between. Even if you are slightly outside the rules, provided you know what to do to be saved, you can confess, try harder next time and be accepted by the faithful. The sons of Satan who live in ignorance, or in defiance of the word of God can easily be distinguished from the worthy, and attacked as unworthy. The danger of the shades of grey approach is that no one is very sure of what is moral, and who belong to the dammed. What proportion of divinity does the prophet have, and what proportion of error? To admit to anything other than complete infallibility in a prophet leads to drift. One generation might believe the prophet capable of one error, the next many and so on until he is not holy at all. That way the laws of the group become indistinct from those of other groups, and little by little the faith itself changes. Religions with black and white, true or false laws are incorruptible, as is the digital code on a CD-ROM disk. The problem they have is that the world is corruptible and changes, so interpreting between the faith and the reality is hard.

Trial by Jury

On August 20 2003, a 42-year-old police officer Detective Constable Brian Stevens walked free from court without a stain on his character, after having been accused of possessing images of child pornography. His name was one of those passed to investigators after his credit card details were found to have been sent to a Texas internet child pornography website for the purchase of paedophile images. The man who ran the site selling access to graphic details of child sexual abuse was eventually sentenced to 1100 years in prison, and the discovery of his activities had set off a world wide investigation. Several files of the repulsive images from the site were found on the officer's computer, and in interviews with investigators he admitted using internet chatrooms "posing as a 13 year old girl out of curiosity". In an interview published in the *Sun* newspaper, his first wife (he had been divorced and remarried) claimed he was a drunken, violent sex maniac with perverted tastes. Given that it is illegal to possess images of children being sexually abused, why did the case against him fail?

Earlier, the police had identified another credit card subscriber to the Texas site as a teacher in Exeter. They arrived one morning at 7 am, broke down his door and searched his belongings. He was then arrested and proceedings were started to take his two children into care. When he protested his innocence the police sighed, "They all say that", but in this case, no pornographic images were found on his computer. The police said that he must have wiped the hard disk. He had also kept a complete inventory of his credit card transactions, and had queried one, which he did not believe he had authorised. Eventually it was discovered that a school technician had copied his credit card details and used them to pay for images from the Texas site, and later 140,000 horrendous images of sexual abuse were found on the same technician's computer.

The evidence in the first case all points in one direction, but it is still within the bounds of credibility that someone else could have stolen the policeman's credit card details, and downloaded the images on to his computer. In this case the police officer had clearly benefited from a principle of English justice, that the accused is innocent until proven guilty. He had insisted that someone else had down loaded the images, and when the prosecution was unable to prove he was lying, (phone records had been destroyed, and the computer expert had made some mistakes in his analysis of the computer) the case against him failed. The question here is not whether he was guilty or innocent, but what level of proof is required to separate the guilty from the innocent. How many innocent people might be wrongly convicted for every guilty one that escapes justice? Or from the opposite point of view, how many criminals can be allowed to get away with it before one innocent person is wrongly convicted.

To argue for an incontestable proven verdict, guilty, or not guilty, in every case, is at best naïve, and at worst will lead to much longer court cases in which ever more detailed evidence is examined with ever lengthening delays in the justice system. While this may be good news for lawyers, whose fees would increase, it increases the administrative costs of the legal system as well with little additional benefit to society. In an increasingly litigious society, where accidents are not accepted, and blame must be apportioned in every minor issue, there has to be frequent recourse to the courts. Then when the lawyers charge higher fees, and take longer to settle each case there will be less and less time and money available for anything else.

Rigorous legal processes reduce errors, but never to zero. Once the time spent on examining evidence has been fixed, there will still be a number of contestable verdicts, but the more careful the analysis, the fewer these might be. Eventually, for some of the cases the quality of evidence is little more than random noise so that minor glitches in the legal process itself can obscure the result, and beyond that, the rate of dubious verdicts is irreducible. There are the errors you do see, like the computer expert's lack of care, and the ones that don't come to light, perhaps insufficient diligence in following up leads by the police. It is always the lawyer's first instinct to protect their reputations by not being seen to convict an innocent person, so the standard of proof for a guilty verdict is set to be very high. That does not eliminate the errors – in either direction – it simply biases the probabilities such that it is much more likely that a criminal will escape justice than an innocent man will go to jail.

A month after his acquittal, Brian Stevens was re-arrested with his former girlfriend, who had provided him with an alibi on the night he was alleged to have downloaded pornography. They were taken into custody on suspicion of perverting the course of justice.

The search for absolute certainty can be very costly. It is possible to reduce the influence of noise by doing the same comparison several times, either one after another, or all at the same time. If the noise is truly random, its effect is unbiased, it can push the result one way in one test or in the other direction in another, but the small piece of evidence that you want to uncover always influences the result in the same direction. Adding the results up causes the effects of real evidence to add but the effects of the noise do not add so fast, so with enough trials, you can reduce the noise by as much as you want. But this is very costly. Running all murder trials 100 times rather than once and adding up the results, plus one for acquittal, minus one for conviction to determine the result will reduce the number of errors by a factor of $\sqrt{100} = 10$, but at what cost? 100 sets of lawyers (they have to be different to ensure that the noise doesn't add up in the same way as the real evidence) 100 sets of expert witnesses, 100 different police investigations. Is the cost justified? The lawyers might say that saving one innocent man from conviction is worthwhile, but they're biased.

To some extent having 12 people in a jury rather than a single judge reduces the variability in the individuals that are involved in the decision-making by a factor of about 3 ($\sqrt{12}$) but the rest of the process is still very subject to error. Juries are confused by complexity, e.g. large amounts of data, and experts presenting balanced views. A balanced picture might say on the one hand it might be like this, or on the other like that. This is confusing, they need one figure, yes or no, certainty. Experts are forced by the system to take one side or the other. They do not present a balanced picture, the two sides each presenting an extreme position. The jury is reduced to deciding which expert they trust, and tend work on the body language of the witnesses rather than the facts presented.

Can we be certain?

Life is short so it is necessary to try to make things simple, and easy to understand, but it can't be done without making mistakes (or being deliberately misled). Along with the desire for simplification in the hearts of people is the need for a faith that they can trust, and against which against which all things can be measured and found good or bad. Such absolute certainty, as with absolute proof, is not to be found in nature. The facts are obscure, and the standards that we measure them against are, in reality shifting sands.

Chapter 12 Four ways to beat Buridan.

Don't make choices if you don't have to.

If the act of choosing may take forever then it pays to recognise when the choice is absolutely necessary, and when it might seem we have to choose, but the question does not actually involve choice. If there's no choice needed, you don't have to wait, so it's worth considering if there is a different way of doing things so that you can never be left in a quandary. On the other hand, if it is certain that you will always have a difficult choice, how can you try and keep the dithering time down?

It is very stressful to be unsure of what to do when time is limited, and you have to do something. Is there any way of avoiding a disastrous decision? There are some answers to these questions, which don't deal with all the problems of choice, but they do help with some. Even in the most famous example, the problem of Buridan's Ass, there is another way.

Marge and the Elephant.

All that is needed is to change the problem, so that instead of a donkey with two equidistant bales of hay, or a starving man with two dates in front of him, there is a thirsty man with two servants each of whom has a jug containing one cupful of water. He has a cup, which can be filled by either of the servants, and then he could drink, but he can only drink exactly one cupful at a time. Which servant does he call upon to fill his cup? This looks very similar to Al Ghazali's difficulty with the dates, but it isn't. He can avoid the choice by placing in front of him a bowl, which can hold at least two cupfuls of water, and then commanding both of his servants to put their water in the bowl. Then he drinks the water by scooping it out of the bowl one cupful at a time. By the time he comes to scoop it out, it is impossible to tell where the water in the cup has come from, so no choice is required. If no choice is needed he has no need to hesitate, and can drink straight away. The key to this is the merging of the two alternatives into a single one, thus eliminating the element of choice. Mark Greenstreet suggested this wait-free option in 1999, when he used the story of Marge and the elephant, which turned water into wine. Customers who want their bottles of water turned into wine give them to Marge, who pours them into the magic elephant. Bottles of wine then appear, the number of bottles of wine coming back from the elephant being equal to the number of bottles of water going in. The only problem with this is that the customers can't get their wine back so easily. If one customer turns up, gives his water to Marge, and the magic elephant does his thing, the wine can be collected, and everyone is happy. But if two people turn up, more or less at the same time their water gets mixed up, and first one bottle of wine is produced and then the other. Who gets the first one? Only in the limited circumstances that it does not matter whose water contributes to which bottle of wine, does it work. So if Al Ghazali is the sole consumer of the wine, he can get drunk as fast as he likes with the assistance of Marge and the suppliers of water, and provided Buridan's Ass can get some help to merge the two bundles of hay into one, he might never need to starve either.

The single track railway

How do you control trains in both directions on a single-track line? When two trains approach the same single-track section from opposite directions, only one can be allowed on. If they both get on it at the same time, they are heading straight towards each other, and there is likely to be a train wreck, so the choice that has to be made is which one is allowed on? The normal solution to this is to have a central control which makes the decision, but these controllers are not infallible, and are subject to all the usual indecision and dithering that all arbiters show.

In the late afternoon, of 30 August 1991 a train departed from Shelby, heading south to Laurel, Montana in the US. Earlier, at 1:50pm a westbound train had left Great Falls and after calling at an earlier station, approached a single line section just north of Laurel

The conductor of the southbound train was the first to see another approaching at a closing speed of 87 mph. He called a warning to the driver and jumped. Immediately putting his train into emergency, the driver paused to call a warning to the brakemen in the rear before jumping himself. The driver of the other train, saw the oncoming headlights, shouted a warning to his conductor and after putting the train into emergency he too jumped.

At about 5:50pm the two trains collided head on. Fuel spilt from the locomotive fuel tanks caught fire, and all the locomotives were destroyed along with 22 freight cars. Three train crew lost their lives in the accident.

Single track sections in Montana are controlled by Track Warrants which give the driver permission to occupy that section up to the next point, so provided that only one track warrant is issued at a time for a particular section, it should not be possible for two trains to be on the same section.

The Track Warrants were actually issued by radio from the dispatching centre 500 miles away in Oregon direct to the drivers. The dispatcher acted as the arbiter, and sat alone in an air conditioned room operating a computerised track warrant system and an indifferent radio, which was noisy and unreliable, issuing a Track Warrant roughly every 10 minutes, and logging them on the computer.

As the westbound train reached Dutton, 22 miles south of Ledger, the conductor called the dispatcher for permission to continue. Unfamiliar with the computer, and talking to the conductor at the same time the dispatcher said "Proceed from Dutton toooo ---------" and searched for the end of the controlled section on the computer muttering "whaaa" as he did it, and accidentally releasing the radio button giving a hissing note. The voice tape recording of the destination sounds like "toooowhaesh". Finally he found what he was looking for. Nine seconds later the dispatcher said "Ledger".

The driver of the train waiting at Dutton heard the warrant destination of Ledger, but the dispatcher did not hear any confirmation of it as at that moment the train moved out and could not be heard on the radio. As it left Dutton the conductor said to the driver "Was that West Yard Shelby" to which the driver replied "sounded like West Yard Shelby to me". When the dispatcher asked for confirmation by repeating the track warrant the conductor reported the "to" location as West Yard Shelby.

Why the dispatcher accepted that is not clear. Maybe he thought it was their final destination rather than the limit of the Track Warrant. The word Ledger was underlined on the computer as confirmation despite the response from the conductor.

Now the southbound train called for permission to proceed. With the other train authorised only to Ledger on the computer there was nothing to stop it also being authorised. The two trains hit 6 miles beyond Ledger.

Working procedures had got sloppy. A combination of unclear speech, transmission gaps while the dispatcher tried to enter of data into the computer system while holding a conversation with the train crew, and straightforward noise in the radio system had caused the Warrant to become corrupted, with fatal results. A later inquiry identified 15 bad practices in the use of the radios at Seattle, but also failures in checking by the crew and the computer system that permitted confirmation by matching two different words all contributed.

The idea of the computer and dispatcher is that together they are the authority which ensures that only one warrant is given, but in this case noise in the transmission, and the uncertainty of the dispatcher led to the crash. All that can be done with a system like that is to reduce the noise, and allow enough time to reduce to accidents to an acceptable level

But you do not always need an arbiter. 100 years ago single-track rail sections were controlled by a much more robust system consisting of single large key. If one train wanted to enter the section to go in one direction, the driver would have to stop the train, dismount, and get the key to unlock that section so that he could get on, no other train was allowed on without the key. While he was on the critical section he carried the key with him, and when he left it he left the key for a train going in the other direction. Provided there's only one physical key, this method ensures that there can only ever be one train on the critical section at a time. If the key is not in its appointed place when a train arrives, it can't get on the single track until the train coming in the other direction arrives and passes it on. A big iron key may not be as clever as a computer, but there's less to go wrong, so it's more reliable.

This method is still used on the roads when repair works restrict the width of a dual track highway to one lane, and a common solution is to use a convoy system. Nobody is allowed to go past the road works unless they are part of a convoy led by a special convoy truck. A convoy is formed out of a specific number of vehicles at one end of the road works, if there aren't enough the convoy waits until there are. Then the convoy sets off to the other end, and when it gets there, the vehicles making up the convoy continue their journey. Then a new convoy is formed to go in the other direction. No decisions are necessary about who can get on to the single lane, it's done strictly in order, one convoy in one direction followed by one in the other, and the word of the convoy leader is law. You might point out that the convoy leader still has to know how many vehicles make up a convoy, and that's a decision, but if he always limits the size to the same number of cars, that one is also avoided.

The obvious problem is not the safety aspect, but the insistence that up traffic must be exactly interleaved with down traffic. Every train going in one direction, carrying the key up the track must be followed by one going down, bringing it back. If the number of passengers going in one direction needs to be different from the number in the other you may have to run half empty trains, but you have to run the same number. To do better, you either have to make a decision when more than one train arrives at the section which one is allowed on, or avoid the decision by ensuring that the traffic is planned so that more than one train can never be there at the same time. Nowadays the choices implied by the first solution are made by computers rather than people, so while there is still a minute possibility of failure because of indecision, it's not the main source of errors. In the second solution, there are no choices, but it involves knowing exactly where every train is at every instant.

It's a toss up

If you know that you are going to be faced by a very difficult choice, and want (on average) to cut the decision time down, you can easily do it by increasing the role of the random element in the choice. Simply toss a coin. Board games like chess or Scrabble involve players taking turns until the game is finished, and often, the player staring first has an advantage, so how to decide who starts first?
You could try to even things up by planning number of games in which players take turns to start first, and this strategy is essentially the same as synchronization. There are no choices in a synchronized system, so once you are into the series, whether you start first in the next game, or your opponent starts first is pre-planned, but you still have to be synchronized into the series. Who starts first in the first game?
The time-honoured solution is the coin toss, or drawing a scrabble tile from the bag. Philosophers object to this as a way of avoiding choice without preference because that there could still be a draw, the coin could land on its edge, or both players could draw the same letter from the bag. If you wait long enough for the coin to fall, as it surely will, or each draw another tile form the bag the decision will be made. But can't it still take forever? Yes, it can, the tiles in the second round could still be the same, but the probability of that is very low, and most of the time it's quicker than arguing.
If you've got to choose between the car and the train, for a journey you can toss up. You might have really preferred to go by car but the coin says get the train, then you'll probably hesitate for a while, and maybe go by car anyway, maybe not. When your preferences are taken into account the number of hesitations of a given length is likely to be about the same as it was with or without the coin toss. However, if you know that the choice will always be difficult before you start and you don't want to allow anyone to influence it, it speeds things up to toss a coin, because without it there's always going to be a

long wait. The coin toss makes a very large random contribution, usually deciding things immediately, while letting two team captains fight it out for the kick off would in the end be decided by small random differences – like who could land the first knock out punch – and would take much longer.

Can I have my money back?

What do I look like in this? Shopping for clothes can be a nightmare, not enough time, not enough money and the disaster of looking a fright in front of all your friends if you choose wrongly. How do you decide what to buy? Some shops come to your rescue by telling you that if you don't like it, you can take it back, or if you find it cheaper elsewhere, you can have your money back.

The rationale behind this is that you'll feel more comfortable in the shop, and spend more time buying than dithering. It's important to the shop that what actually happens is that most purchases don't get taken back, and everyone wins because they save time overall. You may have to spend extra time taking things back occasionally, but spending less time in the shop, and feeling happier at the outcome more than compensates for the effort.
You are still making the choices, but you don't need to spend so much time at the shops. From the shop's point of view, even better than giving the customers his or her money back, is to give them tokens. This is a bit like the wine for water scheme, instead of a choice between a jacket and a pair of trousers, you've traded them in for tokens, which are neutral, and could stand for either when you've made up your mind. But you are committed to a purchase.

This one works because there is a recovery strategy. Take it back and get the money back. It wouldn't work where there is no recovery strategy. A murder trial in a country with capital punishment can't be finished early, and then rerun if more evidence comes in, once you've killed the accused, he stays dead.

So there are several ways to avoid making difficult choices, some avoid the choice altogether, and some reduce the consequences of long decision times, but you can have a win sometimes.

Chapter 13 What philosophers think

One of the concerns of philosophers is with the proof, or otherwise, of the notion of free will in men. Are all our actions determined by the world we see around us, and our experiences? Do our opinions depend only upon what we have learnt, and what we inherit with our genes? If determinism rules, then when we come to a situation in which we have to choose a course of action, for good or for evil, we are in effect, the prisoners of circumstances and our history. To some, this is a deeply depressing view because it seems to suggest that there is nothing we can do about the future. It's all predetermined, so there's no point in trying. The opposite view is that there is some spark in every individual that is free from such earthly matters as circumstances and history, call it a soul if you wish, but it enables that person to break free from such causal determinism, and to be in a small way a creator rather than a creation.

Choice and Free will

The issue is free will, which is hard to demonstrate in the abstract, but they thought that choice without preference was concrete enough to show the logic of their arguments. The question it poses is this: given nothing in the circumstances or in history to incline a donkey towards one pile of hay or the other, can it make a choice? If it can, then clearly it has some inbuilt mechanism that allows it to transcend the bounds of the real world, and if it could starve to death then it has not, and might even be considered to have no soul.

If I concede that he will [starve to death],' said Spinoza (1632-77), `I would seem to conceive an ass, or a statue of a man, not a man. But if I deny that he will, then he will determine

himself, and consequently have the faculty of going where he wills and doing what he wills.'
So the fact that no one has ever seen an ass stuck between two bales of hay, or a man starving to death with two dates in front of him can be taken as a proof of the existence of free will. The man or the ass must have some internal something that enables them to choose. While this could be taken as evidence that actions are not always deterministic, it is by no means a proof, because it is very difficult to judge a man’s state of mind from the outside. He may have been influenced towards one or the other of the alternatives by something in his childhood, that was itself determined by the genes of his parents, this is not always obvious because we do not know his detailed history, so how can we tell if the choice is completely free of preference?
Al Ghazali avoided the question of ‘how can we tell’, by assuming that that a state of no preference can be achieved, even if we can’t observe it, such a thing may happen, and if it happens, either the man starves to death, or he is guided by his will. He said:
‘Everyone, therefore, who studies, in the human and the divine, the real working of the act of choice, must necessarily admit a quality, the nature of which is to differentiate between two similar things.”
The implication is that this quality is restricted to God and men, and is not given to animals or machines. John Buridan was more cautious, he believed that in the case of free will, it would be hard to tell, and he was noticeably less dogmatic than Al-Ghazali.
He says: ‘For it would be difficult indeed to show that when our will is indifferent between to opposed acts, it could decide for one or the other without an external factor, where a dog could not.’
Spinoza is even more forthright in his conclusion:

`I grant entirely that a man placed in such an equilibrium ... will perish of hunger and thirst. If they ask me whether such a man should not be thought an ass, I say that I do not know - just as I do not know how highly we should esteem one who hangs himself, or ... fools and madmen...'

Descartes doubted everything, even the fact that he was sitting by the fire - he might be dreaming. But he did not doubt that he was thinking. His dictum, “Cogito, ergo sum” separated the brain, the physical mechanism of thinking from the mind - the separate immaterial thing that in his view could exist after death. This is what we might call the ghost in the machine, the separate soul that is usually identified with consciousness and free will.

Leibniz knew that absolute indifference could never be achieved, but still believed that “man is free, and the ass is not.” He conceded, however, that both ass and man were subject to the same absence of perfect equipoise, both would eventually make a choice, and so avoided the question of ‘what is free will?’

Ted Honderich expresses a more recent opinion in an interview for the Philosophers’ magazine.

He defines determinism as the doctrine that everything we do comes from our history. He calls it the causal sequence, and is very definite in ruling out any kind of chance. ‘The sequence is one of standard causation. Each event in it is a real effect…certainly not an event merely made probable by antecedents’. Is this compatible with free will? It seems not, but those who think it can be compatible, the compatibleists, say that a free action is one that is determined by internal state rather than external inputs, even though we know that any internal state is simply a result of history. A free decision in these terms is one, which comes from the desires of the individual however un-free these might be, rather than from the gun placed to a mans head by a blackmailer.

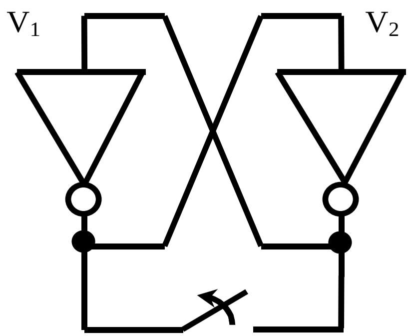
Those that hold that free will is incompatible with determinism, define free will differently. They look for origination, the idea that a decision could have been different at the time it was made. If I decide to have an apple or a pear for breakfast, the past could have been exactly the same as it was up to that moment, but I may decide differently. Here we have choice without preference. There is no explanation for the choice that I make, it must be purely internal, and therefore is my responsibility.

Origination

Philosophers seem to have difficulty arriving at a clear concept of origination, but Honderich says that at the point that a decision is made, to do this, or that, it would not be possible to predict what the decision might be, in his words 'There is no standard explanation of the decision I in fact make'.
To an electronics engineer, this is very close to the definition of a truly random sequence of 0 and 1 bits. At any point in a truly random sequence, it is impossible to predict to better than 50% accuracy what the next bit will be, 0 or 1. It is completely independent of its history, and therefore any internal state. If there is any explanation why the bit is 0 or 1, it cannot be random. This is not just a definition of some theoretical idea; real electronic random bit sequence generators have been made and are used in several applications. The best-known example is ERNIE the Electronic Random Number Indicating Equipment that is used for picking the winning premium bond numbers. Anyone who thinks that the sequence of numbers produced by ERNIE is predictable should start predicting and buying the right bond numbers.
Another application is the creation of codes for encryption. Security services and villains alike invest a great deal of money

in methods for breaking codes, the security of which depend on the impracticability of deducing the key. A completely random sequence of 128 bits can have 2^{128} combinations of 0s and 1s, and all the combinations are equally probable, so a search for the correct key might need to try 2^{127} combinations before finding the right one. That's a big number, 1 with 38 zeros after it. Even if you can try a million different combinations every second, it'll take you about 10^{25} years to get the right one. Non-random sequences are much, much easier to predict, and most people's computer passwords are a similar number of bits, but in non-random combinations, so the average hacker can easily guess what they are.

Most random number generators are based on thermal noise, and one can easily be made from a flip-flop. Just connect the two wires together for long enough for it to reach a stable equilibrium, then open the switch. If the voltage difference between the wires is less than the level of thermal noise (which it usually is), the final state of the device, 0 or 1, will be determined almost entirely by the noise.



Random bit generator

There is a small internal memory effect, but it is small, and it can easily be reduced to a negligible level, and the resulting output bit stream will be random as far as the most sophisticated security checks can determine. Non-determinism is anathema to a computer architect, computers should be dependable, which means that that the hardware must be deterministic, you know what you are going to get. Generally speaking, computers are reliable, and can do things that people cannot do. Their great strength is their accuracy and dependability, but they have limitations. All computers have to communicate with the outside world, and therein lies one of their fallibilities. Data has to move from one timing domain to another. It can only be done with a finite reliability in a finite time, and however hard we try to make them deterministic, non-determinism creeps in at the edges. In the few cases where non-determinism is actually required, such as generation of keys for cryptography, it's easily done, and if it is not perfect, it can be made as close to perfection as you might like.

Christopher Taylor and Daniel Dennett believe that free will and determinism are compatible. To them the idea that in a deterministic universe one can never truthfully say "I could have done otherwise" is false. A famous comment by John Austin illustrates the point:

"Consider the case where I miss a very short putt and kick myself because I could have holed it. It's not that I should have holed it if I had tried: I did try and missed. It's not that I should have holed it if conditions had been different, that might of course be so, but I am talking about conditions as they precisely were, and asserting that I could have holed it. There's the rub. Nor does 'I can hole it this time' mean that I shall hole it this time if I try anything else; for I may try and miss, and yet not be convinced that that I could not have done it; indeed further experiments may confirm my belief that I could have done it that time, although I did not"

Here is Buridan's Ass again. The ball goes in, or it does not but the factors which decide whether it does or does not are continuous and include genuinely random ones. Taylor and Dennett point out that the range of possibilities implicit in 'further experiments' can only exist because of the random factor, but find this "strangely dissatisfying as a new foundation for human freedom and dignity"

They use an absolutely deterministic world based on a computer that is sufficiently reliable to always reproduce its results when the same program is run twice. In the toy world two chess playing robot compete, but unless there is some additional random element, all games with program A taking the white pieces, and B taking the black will be identical. They add a random element by generating a pseudo random sequence of bits so that all games are subtly different.

Now, a pseudo random number generator is not the same as a random number generator based on thermal noise, with a given starting point (assigned when the computer boots up) it always produces the same sequence so it is completely deterministic. Similarly the computer system is assumed to be deterministic throughout. In this toy world, they claim, it may be possible to make the pseudo random number generator indistinguishable from a real one and the computer big enough to support complex organisms rather than just chess playing automatons. In their world these organisms can adapt to their environment, and learn from their mistakes. In that sense they are able to make a difference to the world that they inhabit, and are entitled to claim some originative value.

In fact their toy world, as they freely admit, is much too simple. It has two flaws, one is that the internal state of the computer is made up of a string of digits. A long string of course, but the number of states possible is countable, whereas we must use continuous numbers to represent the real world, which are not countable. Mapping between the two means that a computer system big enough to contain many complex organisms as well as their environment is likely to have a degree of indeterminism itself, so we do not need a pseudo random generator, both kinds of free will exist in the machine. And if a simple machine can show the spark of origination, how can we deny it to a man?

Is the world just a computer simulation?

A recent suggestion is that the world we know is just a computer simulation run by a higher intelligence. Such a thing is possible only if the exponential growth in computer power over the past 50 years is projected forward for many years ahead. Each new decade has seen computers with an order of magnitude increased power over the previous decade. If this carries on (and the implication is why not?), we could soon be simulating not just aircraft performance, virtual dinosaurs, and the weather, but whole people, populations and worlds. If we are all just part of a computer simulation, then time travel becomes possible, the simulation can be run backwards as well as forwards, but the bad news is that we have no free will. The future, as well as the past is completely determined.

To get the computer power required using the same technology of today's systems on silicon would need the suspension of a number of the laws of physics. Let's, for the moment, suspend reality, and assume that the power of a computer might be sufficient to simulate not just the billions of cells in a human being, but the billions of interacting human beings as well, and all the atoms and molecules that make up each of the cells. Another 20 orders of magnitude at least would be needed. No, problem, you might say. Only another 200 years before computers are good enough. But to do that, would require the size of the gates and memory bits to be reduced in size by another 10 orders of magnitude, making them smaller than the size of an atom. Again put that on one side, and assume it could be done. The energy involved with each decision would have to get much less, also around 10 orders of magnitude smaller than it is now, and well before you could get to simulate one person, it would become comparable to the thermal noise, not just for difficult decisions, but for all decisions. Given that, the simulation is anything but deterministic, it might be hard to tell if we are part of a computer simulation or part of a real non-deterministic world. In either case our free will still exists because the computer world is non-deterministic in the same way as the real world.

Fatalism and fortune telling

If origination exists in people, there is a strong argument that it is the result of purely random processes in the brain, and not a manifestation of the hand of God. Things happen in a living organism that is the consequence of pure randomness. The immediate cause of many cases of cancer is a random mutation in the DNA of a cell, and this itself may have been the result of the random breakdown of an unstable atom, giving out a radioactive particle, which damages the structure of the DNA. Measured over the lifetime of an individual, the ability to repair such damage reduces, and habits like smoking, can also increase the chance of cancer, but exactly when, or whether it strikes, is not predetermined. Some people live longer; others live shorter lives than average. The brutal Soviet dictator Stalin, who was responsible for the deaths of millions of Russians in the 1930's, died of a stroke in 1953. If not quite random, there is a strong element of randomness about the occurrence of strokes. It is quite easy to imagine that better control of blood pressure by Stalin's doctors, together with some good fortune might have enabled him to live until the mid sixties. How would he have dealt with the Cuban missile crisis of 1962? In 1944, at a critical moment in the second world war, a bomb plot just failed to remove Adolf Hitler from command. He was injured, and deafened, but not killed or completely incapacitated, so the war continued. Both of these near random events affected the lives of millions, for better, or for worse. In the long term, when we are all dead, there are likely to have been just as many bad men whose lives were cut short as those whose lives ere spared, so the overall effects of randomness are reduced, but never to zero.

It seems, then, that strict determinism, in which any kind of chance is ruled out, does not exist. We can never know exactly what the future will be, and the further ahead that we try to see, the more uncertain it becomes. Predicting the weather in the short term is hard, but possible, because if we take enough measurements of the current state and the history of atmospheric pressures we can rely, more or less, on the deterministic laws of physics to predict the next few days.

In the very long term, climate change is also predictable, because random events average out, and large scale influences dominate the noise, but it may never be possible to predict which days in the year will receive more rain than others, and exactly where a shower will fall, because the system is chaotic. The beat of a butterfly's wings, or the random movement of molecules in the air can be amplified to the level of a tropical storm in one place and not another. Because noise always exists, and the further ahead you want to prognosticate, the more its effects are felt, there is no future for fortune telling or indeed, for fatalism. We are not necessarily all doomed.

Origination: Noise in the brain?

The triumph of the will

We all hope for a better future. If determinism is true it seems we don't need to bother trying to make it happen, it will either happen or not happen anyway.
Daniel Dennet argues that even in a deterministic world, it is possible to evolve organisms that avoid harm to themselves and so live longer lives than those that do not. He cites the computer game of life in which a simple set of rules determines whether pixels on a screen are switched to black or white according to their neighbours. Complex patterns of pixels

evolve which are resistant to destruction. They have within their structure the capability to avoid death. So it is possible to acquire knowledge about the rules from the environment. Electronic systems like his are called finite state machines, and finite state machines include all computers. They are entirely deterministic in that the set of rules and the starting state (which pixels are black and which white) will completely determine what happens for all time. The fact that the number of different patterns you can get is finite doesn't mean that there aren't a large number of them. Even a modest screen with 1000 x 1000 pixels has $2^{1000000}$ possible states. If you regard human brains as a collection of neurons which make a finite sate machine, we can also learn about our environment from our own experience and that given to us by previous generations by our parents or in libraries, then use this in our brains, which have evolved in a way which should help us deal better with the world.

The fact is that we try to do just that. And it seems to make a difference. If that were not so, mankind would not have advanced in health wealth and control over his environment so much in the last 10,000 years.

The history of the world has shaped our species, and our life and experiences have shaped our outlook. Out of this internal state comes the will to succeed. We are built to survive, and we cannot survive without other people. Despite the inherent randomness and unfairness of the world around us, and even that in our own heads, we continue to make decisions that are aimed at our long-term health and success. The future is not all predetermined, and even the smallest action by an individual can make a difference. We can argue that even animals have freedom of will because they are able to make their decisions on where to go for food. Cats can decide whether to obey their genes, and hunt for mice, or whether to rely on food put out by humans. They will even attempt to influence their future prospects by approaching humans and purring deceitfully. Humans are only different in that they have a greater capacity for gathering information, and a greater capacity for influencing their environment. In that sense they have a greater freedom of will. The more we know about the world the easier is the decision. With hard decisions it is often possible to delay the decision until we have enough information to overcome the noise, so we don't always need to be the victims of random processes, but only if there is enough time. We don't need to be destroyed by an errant asteroid if we can measure its path accurately enough, and soon enough to deflect it with a small and well placed force. The more information we have about the present, and the more we have assimilated from the past the better will be our control of the future, which according to Shakespeare, lies not in our stars, but in ourselves.

Chapter 14 Of men and machines

In history the question of choice without preference has thrown some light on what people understand by choice, and has been used to inform the discussion of free will, in people. Humans are distinguished from other species by amongst other things, the size and organisation of their brains. The function of the brain may one day be imitated by the development of computers, systems made not by nature but by people. Whether this in fact happens, and whether or not it is possible to have a truly artificial being, will depend on the quality of the imitation, whether the arrangement of the neurons in the brain can be imitated by hardware and software, whether that arrangement can organise itself so as to learn from the environment, and more controversially, whether that organisation can take responsibility for its actions.

Machines already take decisions, and we can have answers to the question of how they do it and what happens when the choice is indifferent. These answers go part way to answering some bigger questions like: do machines have intelligence, could they be conscious, and might they have free will.

By putting some of the simple things we know about decisions in machines alongside those things we know about decisions that people make, we may also be able to in understand something about the way people think and whether a brain could be considered a machine.

Here we revisit some of the things discussed earlier to see if there is anything to be learnt from the comparison between men and machines. One important question which should be asked, is why people find the concepts of Buridan's Ass or Zeno's paradox difficult to understand, because not understanding can lead to illusions about the nature of choice and free will. According to Leibniz, writing in the 17th century, there can be no such thing as choice without preference, but even now, there are people who insist on asking what happens when they are presented with an indifferent choice, and some of those will go on to describe what they would do when it happens.

Why there aren't many starving donkeys

Because we can measure the responses of an electronic decision maker for all likely inputs, we can make computers with known reliability and known decision times. It is possible to propose a law that predicts, if only statistically, how long they will take to decide. Are the decisions of animals taken according to different laws? In the past people have thought so. Since no-one has ever seen a starving donkey between two bales of hay, it seems obvious that it can't happen, and we can go on to postulate that it is God's will that determines that it can't happen. The only issue open to question is whether only God and humans have the divine spark able to make the choice when all things are equal, or whether that divinity extends further down the pecking order to sentient animals, or perhaps as far as all living things.

In the case of an electronic arbiter, since it is clear that it has no hope of divine intervention in making up its mind we can use a different approach to understand how it is done. The time taken for decisions in and electronically controlled robot depends on how near the input is to the balance point, or how close the robot is to half way between two equally desirable objectives. In this case the time to decide increases by a fixed amount, 2.3τ every time the input difference falls by a factor of 10. The fact that it has some internal noise may alter an individual decision, but over many decisions it makes no difference to the number take longer than 2.3τ, (about 10%) or 4.6τ (about 1%) etc. This is because it is equally likely that a decision may take a little longer, or a little shorter time because of noise. It's harder to do the measurements because you can only do one every 10 seconds or so, and life is too short to do enough to establish an iron law. Suppose, though, that the same law still holds, and that τ is about 0.3 seconds, about the same as reaction time. What that would mean is that 90% of decisions can be made in 0.69 second, 99% in 0.92 seconds, and 99.9% in 1.15 seconds. Now if you are walking towards someone in a narrow corridor at a combined speed of 2.5 metres per second, you probably have 1.5 – 2 seconds to decide which way to go, and in that time 99.99% of the choices can be made, so you don't bump into each other very often. If you both got on motor cycles, and rode at each other at 10 metres per second closing speed you'd be lucky to get past 2 times out of 3. You can argue about the exact figures, but most people would agree that the less time you have the more likely you are to crash.

So it seems the same law of diminishing errors with time applies to the chances of not being able to select one out of two bars of chocolate in the time it tales to die of hunger.

How long does it take to die of hunger? Maybe 2 months, and the chances of not deciding in that time are down to about 1 in $10^{10,000,000}$. Since there are only 10^{10} people alive on earth today, and there are unlikely to have ever been more than 10^{12} people that ever lived, the chances of anyone ever having seen someone starving to death between two bars of chocolate are pretty remote.

As far as any reasonable person is concerned it can't happen, and never has happened, but there's still a finite chance that it could happen.

Divide and conquer

People are very good at dealing with complex tasks. They usually do it by recognizing that a complex task is often just a collection of simple ones. If you can deal with each of the simple ones in turn, then putting then together will solve the complicated one.
You have to be careful though, how you do the decomposition, for example if your problem is to decide which of two dates to eat first, you may want to divide the problem into three cases:
1. If the left hand date is nearer eat that first.
2. If the right hand date is nearer eat that first
3 If the distances to both are the same then eat neither (or god will choose for you).

The third case is a strange one, because it has a singularity in the time required to make the choice. As the distances to each of the dates become closer, the time grows without limit. The best-known example of a singularity is a black hole. A black hole is star, which has collapsed to the point where its mass is pulled inwards by its own gravity and has a gravitational field, which increases without limit as you approach it. Reasoning about singularities is more difficult than reasoning about situations where the normal assumptions apply. Assuming that a particular task will always take much the same amount of time is not something you can do in the presence of a singularity in time, so that decomposition is not helpful, and the best way to think about the choice is to divide it into only two cases, not three, either the left date is nearer, or the right date is nearer. The third case in fact does not exist since the two distances are never exactly equal. They can be close, but the probability of them being closer than a given distance is proportional to the distance. When the distance is exactly zero, the probability of it happening is exactly zero, so it never happens, as Leibniz knew.

But that doesn't stop people believing that it does, for example philosophers like Nicholas Rescher who distinguishes between what he sees as valid reasons for a choice and those which he sees a invalid reasons
"If a person were offered two dollar bills, the only perceptible difference between which is a difference in the serial numbers we would be greatly astonished of this selector could offer us a reason for choosing one rather than the other."
He sees this as a demonstration of the possibility of choice without preference, but there is still a difference between the dollar bills, people do have preferences for one number rather than another. It's just that this difference is small compared with noise, and in the electronic device, any difference less than noise results in a more or less random response.

Abstraction and divisibility

Sometimes changes in the way people do business can change the problems they face in life. Before the invention of money, barter was the way in which goods were traded. If you provide me with a service I will give you a horse. Coins are a halfway house between pure barter, and a purely abstract payment system like a credit card
The sophist, who argued that having promised a horse in return for services rendered, he could not deliver, could be forced to pay if he had signed the credit card bill. He argued that: "If no horse can be found that is the one that I owe to you, then I do not owe you a horse."
Because all horses are the same, it is impossible to pick one rather than another, and the concept of some abstract universal horse could not be delivered, so he did not have to pay.
Now if instead of agreeing to supply an actual horse the agreement is to supply the value of a horse, there is no need to

select a particular animal at the time of delivery. The cash is handed over, the recipient buys himself a horse, and the debtor is forced to sell a horse in order to eat.

Aha! The sophist might say, "which actual coins do I select to hand over out of the number that I have" to which the modern creditor might reply, just do it by electronic funds transfer, there is no need to hand over real objects, the transaction is a purely abstract debiting of your account and a crediting of mine.

The need of the sophist to select a horse to sell in order to replenish his bank account still exists, but it is of no concern to the transaction, which can be completed to the satisfaction of his creditor without any difficulty.

Being able to avoid choice is concerned with the divisibility of an object you wish to take. If you really, really must choose between two discrete objects and take one rather than the other, then the need to turn the continuous choice criteria into a discrete result ends up giving a variable choice time.

To avoid the necessity for such choices you can make the objects divisible (glasses of water), or not actually pass them (horses), or fix the ordering of the choice (synchronization, single track railway). Otherwise because of the choice you have to live with unreliability. Either it is possible that it takes and infinite time or there is a crash. Money back just gets you out of the shop, you still have to choose your clothes, and that could still take forever. Exactly the same techniques can be used in computer systems, but in the end the apparently perfect reliability of a digital system is always, to some extent, compromised by the synchronization of events, which entail necessary choices.

Them and us.

People yearn for the perfect reliability of a moral arbiter, but this seems not to be possible either for men or machines. The guidance given by religious authorities appears clear, but is not. Pope John Paul II, in 1996 spelled out the Catholic position. In his view “Man is the only creature that God has wanted for his own sake” “Man is called to enter into a relationship of knowledge and love with God himself” “it is by virtue of his spiritual soul that the whole person possesses such a dignity”, and “Theories of evolution which…consider the spirit as emerging from the forces of living matter… are incompatible with the truth about man. Nor are they able to ground the dignity of the person”.

This puts a clear line around the human species, dividing animals into two discrete classes, those for whom which communion with god is possible, and those for whom it is not. It is in fundamental conflict with the idea of evolution, which is based on a slow continuous development of man from other animals. If such a line exists it must have appeared at some time, a time before which the human race did not exist, and after which it did. Quite logically, the Pope has said that evolution is incompatible with the idea of the spiritual soul.
Most religions are concerned with differentiating one thing from another, and put a moral line between those things they like, and those that they don’t. Some define as moral what most satisfies the preferences of sentient creatures. If we believe that all animals, including man are sentient, then this definition leads naturally to vegetarianism and animal rights. Extreme adherents to this religion use as a slogan ‘meat is murder’, and they put the line in a different place, but still have a line between sentient animals and non-sentient things. Others believe fish have no feelings and are outside the pale, but it is the sharpness of the line that leads to inconsistencies. Everyone inside is ‘us’ and everything outside is ‘them’ but

being included in the moral and sentient community is unlikely to persuade all animals to accept vegetarianism. Cats in particular would probably want to continue hunting and killing other animals. Maybe we can accommodate this by saying that cats are not moral agents, and, therefore are not bound by the moral code. By this stratagem a cat can have a vote in creating the moral code, but is not bound by the code itself. If we truly believe that animals are sentient in the same way that we are, then must they not also be moral?

It is only possible to be consistent by accepting degrees of sentience and morality. The world is at least partly continuous and not all discrete. Humans are both more sentient, and more morally responsible than cats, so cats have less of an influence on the definition of morals than humans, and it is mostly the greater good of the human race that we should consider. At the same time cats are much less bound by a code that they did not create.

This will not suit animal rights activists because the moral code is not absolute. The status of a species depends on how you weigh each of the votes and that can easily shift over time. A code based on absolutes can never vary, if animals are equal to humans, they are always equal. The fundamentalist likes absolute certainty, so that drift can be corrected, as it always can with any system consisting a limited number of values or a few rules, and acceptance of the creed forces people to choose, you are either with us, or against us.

Repressive regimes rely on loyalty to keep people in line consequently they try to prevent the publication of new ideas which may discredit their morality as Galileo found to his cost. Restricting information by government censorship or religious dogma reduces the freedom of the individual to choose, and hence to influence their future. In a totalitarian system, where the laws and the moral code are absolutely fixed, new knowledge is impossible. Scientific development cannot proceed because no new ideas can be admitted. But on the other hand moral degeneration is equally impossible. Standards are upheld, and transgressions, however slight are severely punished. In George Orwell's Nineteen Eighty-four one of the purposes of Newspeak, the language of the Party, was remove all ambiguities and shades of meaning, so that it became impossible to think any thoughts that were disloyal to the state.

There are times when innovation is useful, and times when it is dangerous. In a world where famine or disaster are not close, experimentation is possible because there is a margin for error. Then, totalitarianism is a disadvantage because new knowledge cannot be used to increase wealth further. If both good and bad ideas are freely available people will usually try ideas likely to improve their situation, and thus freedom of thought is part of a virtuous circle. More ideas produce more possibilities, and increase wealth faster. The increase in wealth provides more security within which more freedom can flourish.

At times of stress, war, famine, plague, and poverty, the reverse is true, strict observance of the code prevents the new idea that might, just might, turn hunger into famine, and a strong line holds back decline, decay, and lawlessness. Otherwise society descends more rapidly into the abyss of crime followed by falling living standards, disease, and then abandonment of knowledge. This view was famously put by Thomas Hobbes following the chaos resulting from the civil war in the 17^{th} century.

"Hereby it is manifest that during the time men live without a common power to keep them all in awe, they are in a condition which is war; …In such a condition there is no place for industry, because the fruit thereof is uncertain: and consequently no culture of the earth; no navigation, nor use of commodities that may be imported by sea; no commodious building; no instruments of moving and removing such things as require much force; no knowledge of the face of the earth; no account of time; no arts ; no letters; no society; and which is worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short"

It cannot be an accident, then, that fundamentalist religion and despotic authority is strongest in the poorest areas of the world.

Why people keep making the same mistake

The idea of market pricing, rather than exchange of goods by barter is very new in human history. It relies on mathematics, and the idea of abstracting the value of an object away from the object itself. It is not natural to human societies, and is still deeply suspect to some. People would rather exchange a horse to pay off a debt than agree a sum of money. A bunch of flowers in exchange for looking after a cat is somehow more honourable than cash. There is more understanding of the tangible goods that are exchanged than the intangible figures. People pile up debt on a credit card because it seems less real to buy on a credit card than to hand over banknotes, and somehow the banknotes themselves are less real than an obligation to a friend.

The English language contains a great many words, but it is very limited in its ability to express number. There are words like sole, and single, which can be used for 1. Pair and double convey the idea of 2, but beyond that we have only a few concepts. Few: 'a small number'. Several: 'more than two, but not many'. Many: 'a large number of'. The distinction between many and infinite is available at a formal level, but is not usually used in ordinary speech, where informal terms like 'umpteen' and 'more 'n you can shake a stick at' serve for both. Even people who understand numerical abstractions at the intellectual level sometimes do not understand at the emotional level. Tom Kilburn and Kronecker, both eminent

mathematicians, saw the difference between whole numbers and continuous quantities, but they did not understand it.
If you don't really believe that time or distance are infinitely divisible, because you only believe in whole numbers, it's very difficult to understand problems like Zeno's arrow or Buridan's ass. They seem to be paradoxes.

Can machines think?

The psychologist B.F. Skinner in 1904 said, "The real problem is not whether machines think but whether men do."
Skinner had little interest in understanding the human psyche. He believed everything we do is shaped by our experience of punishment and reward. In his view the mind (as opposed to the brain) consciousness and the soul are unobservable and so useless for a scientific psychology, so they don't really exist. Instead, Skinner recommends that psychologists concentrate on observables, that is, the environment and our behaviour in it. Skinner was known for being provocative, controversial, and an excellent publicist of his ideas.
As believer in the environment as the shaper of behaviour, Skinner did not want to admit to any inbuilt aspects of personality, but even those who argue strongly for inborn personality traits, such as Steven Pinker, writing in "The blank slate" do not accept the concepts of either self, or soul. He describes a demonstration of the illusion of the single unified "self" in which surgeons cut the corpus callosum joining the two hemispheres of the brain, so that each could exercise independent free will. When the right half initiates some action as a response to a signal that the left half cannot see, the left half is asked why did you just do that? A plausible explanation is always forthcoming, even though there can be

no communication between the halves, and the left half has no responsibility for the actions decided by the right half.
He calls this the boloney generator. Even more bizarre is the existence of human chimeras, where two non-identical twins have fused in the womb to form one individual. Some of the cells contain DNA from one twin, and others contain the DNA from the other. An apparently normal woman in Massachusetts was recently discovered to be a chimera as a result of tissue typing her sons. At first it seemed that they could not possibly be the result of a natural conception, because they did not match her blood cells, but a closer look showed that most of her blood cells must have come from one twin, and some of the cells in her ovaries had come from the other. The report does not say what mixture of cells make up her brain, but it seems likely that her brain is a partnership rather than an individual, so in her case at least, the concept of self cannot be unique
These observations underpin the view of some neuroscientists that the concepts of consciousness is an illusion, there is no self or soul and they suggest that living beings are powered by molecular clockwork. The processes we call thought are, according to this way of thinking, governed only by the laws of physics and the structure of the brain, and there is no reason to think of the brain as other than a mechanism.

So can machines think? We are used to machines that are intended to be deterministic. Press the accelerator, and the car goes forward. A car that is non-deterministic is a joke. People laugh at Basil Fawlty, who started to beat his car with a branch when it broke down. Everyone knows that machines are not sentient, and have no free will, so hitting a broken down car makes no sense. Because this idea is ridiculous it does not follow that unpredictability could not be deliberately built into a machine, as it is already in the fruit machines of Las Vegas. Computer engineers go to some pains to reduce the indeterminacy in synchronizers affecting the reliability of their systems, but it is only possible to have a completely deterministic computer if it never interacts with other independently timed systems (for example humans at the keyboard). Random, unplanned events are a characteristic of both people and machines, and if origination is the mark of the soul, both have it.

Both men and machines can also respond to their environment, and try to prolong their existence. This is shown by the hackers who design computer viruses. These viruses are, as yet, primitive, but they can multiply, and infect new hosts over the internet, the more new hosts they infect the more they prolong the life of their species. Fortunately they do not yet evolve, and they rely on their host computer systems to survive. Even so, in order to continue working correctly, the host computers have had to be equipped with virus detectors which function in the same way as immune systems in an animal's body. Most people would regard a computer virus as a machine rather than a life form, computer viruses do not yet exchange code with each other, and they are not yet as complex as those built from the same molecular clockwork that we are, but this is simply a matter of scale. Even emergence of code with a scale of complexity similar to a human genome (3 billion base pairs) is not difficult to imagine, there are already plenty of computers with a main memory big enough to hold code that big, and researchers into artificial intelligence try to define the nature of consciousness in order to replicate those essential functions in a machine.

Many people find this idea repugnant, but if people are just machines, and machines can breed and originate their actions, what is there to prevent them having thoughts and consciousness?

Further reading

Spinoza, Appendix to *Ethics* 2, in *The Collected Works of Spinoza*, ed. and trans. Edwin Curley (Princeton, NJ: Princeton University Press, 1985).

The world of Medieval Learning, Anders Pilz, Basil Blackwell, 1981

What Philosophers think Ed Julian Baggini and Jeremy Stangroom, Continuum London 2003, ISBN 0-8264-6754-7

O'Connor, John J. The Rise of Calculus. St Andrews, Scotland, 1996.

“The free will handbook” Ed Kane, Oxford University press.

N Rescher, “Choice without preference: A study of the logic and the history of the problem of Buridan’s Ass” Kant-Studien, 1959/60, No. 51, pp142-175.

E A Moody ‘Jean Buridan’ in Studies in medieval Philosophy, science and logic: collected papers 1933-1969 (Berkeley, Calif., 1975

David J. Kinniment ‘Synchronization and Arbitration in Digital Systems’, Wiley and Sons, Chichester, 2007

Mark Greenstreet

Steve Furber