Information and work

Paul Cockshott

September 28, 2004

Contents

0 Problematizing labour **5**
0.1 Watt on work 5
0.2 Marx: The Architect and the Bee 8
Emergent buildings 10
Apian efficiency 10
0.3 The Demonic challenge 13
0.4 Entropy 14
Men and horses 18

1 Problematizing information **21**
1.1 The Shannon–Weaver concept of information 21
Information engines as heat engines 24
1.2 Entropy reductions in action programs 26
1.3 Alternative views of information 26
The Chaitin–Kolmogorov concept of information 26
1.4 Randomness and compressibility 29
Random numbers contain non-random ones 29
Randomness of a number as opposed to of a generator. 30
1.5 Information and randomness 30

Chapter 0

Problematizing labour

0.1 Watt on work

Prior to the eighteenth century, muscles—whether of humans, horses or oxen—remained the fundamental energy source for production. Not coincidentally, the concepts of work, power, energy and labour did not exist in anything like their modern form. People were, of course, familiar with machinery prior to the modern age. The Archimedean machines and their derivatives—levers, inclined planes, screws, wheels, pulleys—had been around for millennia to amplify or concentrate muscular effort. Water-power had been in use since at least the first century A.D.,[1] initially as a means of grinding grain; during the middle ages it was applied to a wide variety of industrial processes. But water-power, and its sister wind-power, were still special-purpose technologies, not universal energy sources. Limited by location and specialized use they did not problematize effort as such.

A note on terminology is in order here. The (admittedly not very elegant) verb 'to problematize' derives from the work of the philosopher Louis Althusser. Althusser coined the term *problématique* (problematic) to refer to the field of problems or questions that define an area of scientific enquiry. The term is fairly closely related to Thomas Kuhn's idea of a scientific 'paradigm'. So, to problematize a domain is to transform it into a scientific problem-area, to construct new concepts which permit the posing of precise scientific questions. In the pre-modern era engineers and sea captains would know from experience how many men or horses must be employed, using pulleys and windlasses, to raise a mast or obelisk. Millers knew that the grinding capacity of water mills varied with the available flow in the mill lade. But there was no systematic equation or measure to relate muscular work to water's work, no scientific problematic of effort. That had to wait for James Watt, after whom we name our modern measure of the ability to work.

Watt, the best-known pioneer of steam, did not actually invent the steam engine, but he improved its efficiency. As Mathematical Instrument Maker to the University of Glasgow he was called in to repair a model steam engine used by the department of Natural Philosophy (we would now call it Physics). The machine was a small scale version of the Newcomen engine that was already in widespread use for pumping in mines.

The Newcomen engine was an 'atmospheric engine'. It had a single cylinder, the top half of which was open to the atmosphere (Figure 1). The lower half of the cylinder was connected via two valves to a boiler and a water reservoir. The piston was connected to a rocking beam the other end of which supported the heavy plunger of a mine pump. The resting condition of the engine was with the piston pulled up by the counter-weight of the pump plunger.

To operate the machine, the boiler valve was opened first, filling the cylinder with steam. This valve was then closed and the water-reservoir valve opened, spraying cold water into the piston. This condensed the steam, resulting in a partial vacuum. Atmospheric pressure on the upper surface of the piston then drove it down, providing the power-stroke. The two phase cycle could then be repeated to obtain regular pumping.

Watt observed that the model engine could only carry out a few strokes before the boiler ran out of steam and it had to rest to 'catch its breath'. He ascertained that this was caused by the incoming steam immediately condensing on the walls of the cylinder, still cool from the previous water spray. His solution was to provide a separate condenser, permanently water cooled, and intermittently connected to the cylinder by a valve mechanism. The cylinder, meanwhile, was provided with a steam-filled outer jacket to keep its inner

[1] See Strandh (1979), Ste. Croix (1981, p. 38).

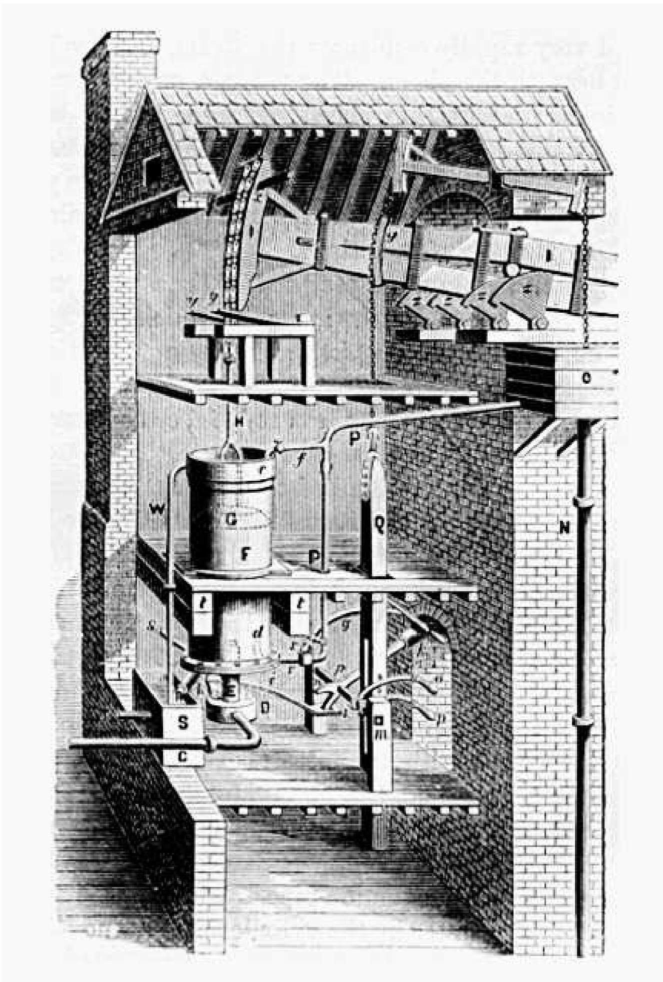

Figure 1: The Newcomen engine built by Smeaton (reproduced from Thurston)

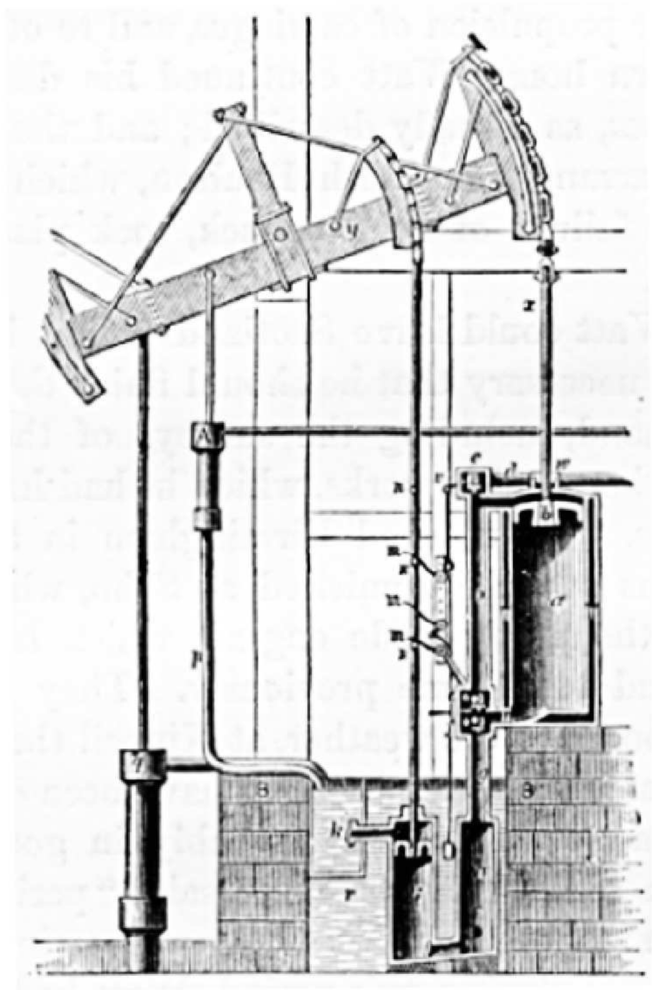

Figure 2: Watt's steam engine with separate condenser (reproduced from Thurston)

lining above condensation temperature (Figure 0.1). His 1769 patent was for "A New Method of Lessening the Consumption of Steam and Fuel in Fire Engines".

Watt's later business success was based directly on this gain in thermal efficiency. His engines were not sold outright to users, but were leased. The rental paid was equal to one-third the cost of coal saved through using a Watt engine rather than a Newcomen engine (Tann, 1981). This pricing system worked so long as the Newcomen engine provided a basis for comparison, but as Watt's engines became the predominant type, and as they came to be used to power an ever-widening range of machines, some system of rating the working capacity of the engines was needed. Watt needed a standardized scale by which he could rate the power, and thus the rental cost, of different engines. His standardized measure was of course the horsepower: users were charged £5 per horsepower year.

Watt's horse was not a real horse of course, but the abstraction of a horse, a standardized horse. The abstraction is multiple: at once an abstraction from particular horses, an abstraction from the difference between flesh and blood horses and iron ones, and an abstraction from the particular work done. The work done had to be defined in the most abstract terms, as the overcoming of resistance in its canonical form, namely raising weights. One horsepower is 550 ft lb/sec, the ability to raise a load of 1 ton by 15 feet in a minute.

While few real horses could sustain this kind of work, its connection to the task performed by Watt's original engines is clear. The steam engine was a direct replacement for horse-operated pumps in the raising of water from mines. But with the development of mechanisms like Watt's sun and planet gear, which converted linear to rotary motion, steam engines became a general purpose power source. They could replace water wheels in mills, drive factory machines by systems of axles and pulleys, pull loads on tracks. Engine capacity measured in horsepower abstracted from the concrete work that was being performed, transforming it all to **work** in general. Horsepower was the capacity to perform a given amount of work each second. By defining power as work done per second, work in general was itself implicitly defined. All work was equated to lifting. Work in general was defined as the product of resistance overcome, measured in pounds of force, by the distance through which it was overcome.

Mechanical power seemed to hold the prospect of abolishing human drudgery and labour. As Matthew Boulton proudly announced to George II: "Your Majesty, I have at my disposal what the whole world demands; something which will uplift civilization more than ever by relieving man of undignified drudgery. I have steam power."[2] To a world in which human muscle was a prime mover, this equation of work in the engineering sense with human labour was exact. Work on ships, in mines, at the harvest, was work in the most basic physical sense. Men toiled at windlasses to raise anchors, teams pulled on ropes to set sails and hauled loads on their backs to unload cargo. Children dragged coal in carts from drift mines, women carried it up shafts in baskets on their backs. The 'navigators' who built canals did it with no mechanical aid more sophisticated than the wheelbarrow (a combination of lever and wheel, two Archimedean devices).

As horsepower per head of population multiplied, so too did industrial productivity. The power of steam was harnessed, first to raise weights, then to rotate machinery, then to power water-craft, next to trains—and eventually, through the mediation of the electricity grid, to tasks in every shop and home—while human work shrank as a proportion of the total work performed. More and more work was done by artificial energy, yet the need for people to work remained. A steam locomotive might draw a hundred-ton train, but it needed a driver to control it. Human work became increasingly a matter of supervision, control and feeding of machines. Thus the identification of work with the overcoming of physical resistance in the abstract, and of human labour-power with power in Watt's sense, contained both truth and falsehood. Its truth is shown by the manifest gains flowing from the augmentation of human energy. Its falsity is exposed by the residuum of human activity that expresses itself in the control, minding and direction of machinery.

Indeed, the introduction of powered machinery had the effect of lengthening the working day while making work more intense and remorseless. The cost of powered machinery was such that only men with substantial wealth could afford it. Cheap hand-powered spindles and looms could not compete with steam-powered ones. Domestic spinners and hand-loom weavers had to give up their independence and work for the owners of the new steam powered 'mules' and looms. Steam power brought no increase in leisure for weavers or spinners. The drive to recoup the capital cost of the new machinery brought instead longer working hours and shift-work, to a rhythm dictated by the tireless engine. The fact that the machinery was not owned by those who worked it, meant that it enslaved rather than liberated.

A particular pattern of ownership was the social cause of machine-enforced wage slavery, but that is only half the story. We may ask why the new machine economy needed human labour at all. Why did 'self acting'—or as we would put it now, 'automatic'—machines not displace human labour altogether? A century ago, millions of horses toiled in harness to draw our loads. Where are they now? A remnant of their former race survives as toys of the rich; the rest went early to the knackers. Why has a similar fate not befallen human workers? Why has the race of workers not been killed off, to leave a leisured rich attended by their machines?

Watt's horsepower killed the horse, but the worker survived. There must be some real difference between work as defined by Watt, and work in the sense of human labour.

0.2 Marx: The Architect and the Bee

Karl Marx proposed an argument which seems at first sight to get to the essence of what distinguishes human labour from the work of an animal or a machine, namely purpose.

---

[2] Compare Antipater of Thessalonika's eulogy on the introduction of the water mill:

> Stop grinding, ye women who toil at the mill
> Sleep on, though the crowing cocks announce the break of day
> Demeter has commanded the water nymphs
> to do the work of your hands
> Jumping one wheel they turn the axle
> Which drives the gears and the heavy millstones

> An immeasurable interval of time separates the state of things in which a man brings his labour-power to the market for sale as a commodity, from that state at which human labour was still in its first instinctive stage. We pre-suppose labour in a form which stamps it as exclusively human. A spider conducts operations that resemble those of a weaver, and a bee puts to shame many an architect in the construction of her cells. But what distinguishes the worst of architects from the best of bees is this, that the architect raises his structure in the imagination before he erects it in reality. At the end of every labour process we get a result that already existed in the imagination of the labourer at its commencement. He not only effects a change of form in the material on which he works, but he also realises a purpose of his own that gives the law to his modus operandi, and to which he must subordinate his will. (Marx, 1970, pp. 177–8)

This suggests that animals, lacking purpose, can be replaced by machines, but that humans are always required, in the end, to give purpose to the machine. We cite Marx's statement because it articulates what is probably a rather widely held view, yet it has several interesting problems. This is an issue where it is difficult to go straight for the 'right answer'. It may be profitable to beat the bushes first, to scare up (and shoot down) various prejudices that can block the road to a scientific understanding.

First, are animals really lacking in purpose? The spider may be so small, and her brain so tiny, that it seems plausible that blind instinct, rather than the conscious prospect of flies, drives her to spin. But it is doubtful that the same applies to mammals. The horse at the plough may not envisage in advance the corn he helps to produce, but then he is a slave, bent to the purpose of the ploughman. Reduced to a source of mechanical power, overcoming the dumb resistance of the soil, he is readily replaced by a John Deere. The same cannot be said of animals in the wild. Does the wolf stalking its prey not intend to eat it? It plans its approach with cunning. Who are we to say that the result—fresh caribou meat—did not "already exist in the imagination" of the wolf at its commencement? We have no basis other than anthropocentric prejudice on which to deny her imagination and foresight.

Turn to Marx's human example, an architect, and his argument looks even shakier. For do architects ever build things themselves? They may occasionally build their own homes, but in general what gives them the status of architects is that they don't get their hands dirty with anything worse than India Ink. Architects draw up plans. Builders build. (In eliding this distinction Marx showed an uncharacteristic blindness to class reality).

An office block, stadium or station has, it is true, some sort of prior existence, but as a plan on paper rather than in the mind of the builders. If by collective labour civilized humans can put up structures more complex than bees, it is because they can read, write and draw. A plan—whether on paper or, as in earlier epochs, scribed on stone—coordinates the individual efforts of many humans into a collective effort.

For building work then, Marx is partially right, the structure is raised *on paper* before it is raised in stone. But he is wrong in saying that it is built in the imagination first, and in implying that the structure is put up by the architect. What is really unique to humans here is, first, the social division of labour between the labour of conception by the architects and the work of execution by the builders, and second, the existence of *materialized plans*: configurations of matter that can control and direct the labour of groups of humans.

While insect societies may have a division of labour between 'castes', for example between worker and soldier termites, they do not have a comparable division between conception and execution, between issuers and followers of orders. Nor do insects have technologies of record and writing. They can communicate with each other. Dancing bees describe to others the whereabouts of flowers. Walking ants leave scent trails for their companions. These messages, like human speech, coordinate labour. Like our tales, they vanish in the telling. But, not restricted to telling tales, we can can make records that persist, communicated over space and time.

Our tales are richer too. The set of messages that can be expressed in our languages is exponentially greater than in the language of bees. Each works by the sequential combination of symbols—words for us, wiggles for bees—but we have many more symbols and can understand much longer sequences. The number of distinct messages that can be communicated by a language is proportional to v^m where v is the number of distinct symbols that can be recognized in the language and m is the maximum message length. If bees have a repertoire of six types of wiggles and can understand 'sentences' of three wiggles in succession then they can send $6^3 = 216$ different messages. A human language with a vocabulary of 3000 words and a maximum sentence length of 20 words could convey about $3.486784401 \times 10^{69} = 348,678,440,100,000,000,000,$ $000,000,000,000,\ 000,000,000,000,\ 000,000,000,\ 000,000,000,000,000$ distinct sentences. Of course, not all of these would be grammatically correct, and a rather small proportion of those would make any sense, but the number of messages is still astronomically greater than what insects can manage. And we can keep piling on the sentences until the listener loses track.

All this leaves open another interpretation of what Marx had to say. True enough, architects may not build theatres themselves, any more than Hadrian built his wall[3] or Diocletian his baths. But Hadrian caused the wall to be built and Diocletian's architect caused the baths to be built to a specific design. (This use of the word 'built' is of course common in class societies, where real builders get no credit for their creations. Their labour contributes instead to the fame of a ruler or architect.) If the architect creates only a paper version of a theatre, can we say, at any rate, that he creates this drawing in his mind before setting it down on paper? This interpretation of Marx's story of the architect and the bee seems to make sense, but it's not clear that it's a true description of what an architect actually does.

Emergent buildings

Some individuals, autistic infant prodigies or 'idiot savants', do seem to have the ability to hold in their minds almost photographically detailed images of buildings they have seen. Working from memory they are able to draw buildings in astonishing and accurate detail.[4] But it is questionable whether professional architects work this way. Some may, but for others the process of developing a design is intimately tied up with actually drawing it. They start with the broad outlines of a design in their minds. As this is transferred to paper, they get the contexts within ¡¡¡¡¡¡¡ bee.tex which the mind can work to elaborate and fill in details. The details were not in the mind prior to starting work, they emerge through the ======= which the mind can work to elaborate and fill in details. The details were not in the mind prior to starting work, they emerge through the ¿¿¿¿¿¿¿ 1.6 interaction of mind, pen and paper. Pencils and paper don't just record ideas that exist fully formed, they are part of a production process that generates ideas in the first place.

At any one time our consciousness can focus on only a limited number of items. On the basis of what it is currently conscious of, its context, it can produce responses related to this context. In reverie the context is internal to the brain and the responses are new ideas related to this context. In an activity like drawing a plan or engineering diagram, the context has two parts

(1) an internal state of mind; and
(2) that part of the diagram upon which visual attention is fixated,

and the response is both internal—a new state of mind—and external—a movement of the pencil on the paper.[5] Where in reverie the response, the new idea, slipped all too easily from grasp, paper remembers.[6] Architecture exchanges for the fallibility and limited compass of memory the durability of an effectively infinite supply of A0. One might say that complex architecture rests on paper foundations.

If the idea of the architect as creating buildings spontaneously out of the imagination is dismissed as an almost religious myth, redolent of the Masonic characterization of the deity as the *Great Architect*, what then remains of the antithesis between architect and bee? Well, how do the bees shape their hive? We can be sure there are no drawings of hexagons, made by the 'queen',[7] and executed by her worker daughters. We are talking here of *apis mellifera* not the solitary bumble bee. The labour of the honey bees is collective, like that of workers on a building site, yet although they have no written plans to work from they create a geometrically precise, optimal and elegant structure.

Apian efficiency

Consider the problem to which the honeycomb is the answer: to come up with a structure that is interchangeably capable of storing honey or sheltering bee larvae, is waterproof, is structurally stiff, provides a platform to walk on and which uses the minimum material. Given this design brief it is unlikely that a human engineer could come up with a better structure.

The structure has to be organized as a series of planes to provide access. Within the planes, the combs, the space has to be divided into approximately bee-sized cubicles. These could be triangular, square, or hexagonal (the only three regular tessellations of the plane). Our architects have a predilection for the rectilinear, but the hexagonal form is superior.

---

[3] It was of course the rank and file legionnaires who built the wall; see Davies (1989).

[4] It may be worth seeing if we could reproduce some images by such autistic artists

[5] The reader may notice that this argument is a thinly disguised version of Alan Turing's famous argument of 1937.

[6] Cite the passage in Tacitus, I think it is in the Annals, where he says that civilization depends upon Papyrus.

[7] The breeding female is no more an architect or Caesar than the Pope is the genetic father of his followers. Monarchy and patriarchy project dominance relations onto genetic relations and vice versa. Apian Mother becomes queen, the Vatican monarch, Holy Father.

Figure 3: Tessellation of the plane using hexagons

A tessellation of unit squares has a wall length of 2 per unit area, since a single unit square has four sides of unit length, each shared 50 percent with its neighbours. A tessellation of hexagons of unit area has a wall length of $\frac{2}{\sqrt{3}}$ per unit area, a reduction by a factor of $\sqrt{3}$ (see Figure 4). The honeycomb structure used by bees is thus more efficient in its use of wax than a rectilinear arrangement would be.

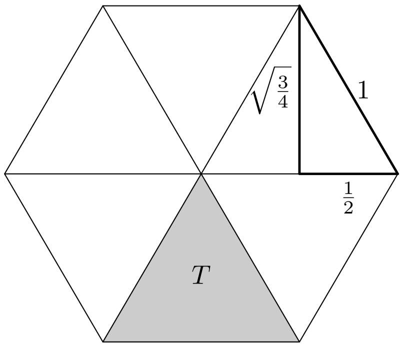

Figure 4: Apian efficiency

(1) A hexagon of unit side is made up of 6 identical equilateral triangles, thus its area is $6T$ where T is the area of an equilateral triangle of unit side.

(2) The area of an equilateral triangle of unit side is $\frac{1}{2}bh$ where b the base $= 1$ and h the height $= \sqrt{\frac{3}{4}}$. So $T = \frac{1}{2}\sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{4}$.

(3) The area of one hexagon is then

$$6\frac{\sqrt{3}}{4} = \frac{3\sqrt{3}}{2}$$

(4) The hexagon's six sides are each shared 50% with a neighbour.

(5) Wall per unit area for a hexagonal tessellation is then $3/\frac{3\sqrt{3}}{2} = 2/\sqrt{3}$ which is better than the wall to area ratio for squares.

The fact that hexagonal lattices minimize boundary lengths per unit area means that they can arise spontaneously, for example in columnar basalts.[8] Here the tension induced in rocks as they cool encourages cracking, preferentially giving rise to six sided columns. We might suspect that the beehive too, gained its structure from a process of spontaneous pattern formation analogous to columnar basalts or packed arrays of soap bubbles. But this doesn't tally with the way the cells are built up, or with the uniformity of their dimensions. In a partially constructed honeycomb the cells are of a constant diameter; those in the middle of the comb are all of uniform height while towards the edge the depth of the cells falls. The bees build the cells up from the base, laying wax down on the upper margins of the cell walls, just as bricks are added to the upper margin of a wall by a bricklayer. The construction process takes advantage of the inherent stability of a hexagonal lattice, allowing the growing cells to form their own scaffolding. But the process also demands

[8] Should we have a photo of the rocks around Fingal's Cave?

that the bees can deposit wax accurately on the growing cell walls, and that they stop building when the cells have reached the right height. That is, it depends on purposeful activity on the part of the bees.

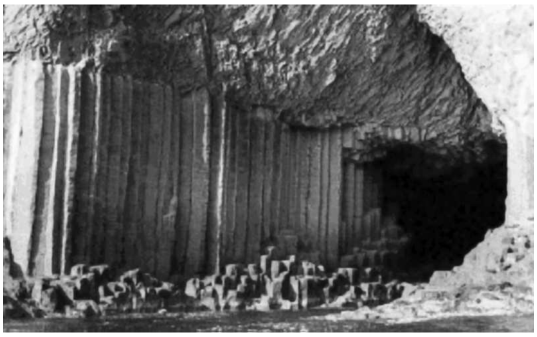

Figure 5: Nature is the architect of the hexagonal columns of Fingal's cave (Photo by Andrew Kerr)

A similar process takes place in the human construction of geodesic domes, hexagonal lattices curved through a third dimension. These have an inherent stability that becomes more and more evident as you add struts to them. You build them up in a ring starting at ground level. The structure initially has a fair bit of play in it, but the closer the structure comes to a sphere the more rigid it is. Human dome builders, like bees, exploit the inherent structural properties of hexagonal lattices, but they still need to cut struts to the right length and put them in the correct place. The bees likewise must select the right height for their cell walls and place wax appropriately.

Spontaneous self-assembly of hexagonal structures similar to geodesic domes does occur in nature. The Fullerenes are a family of carbon molecules named after Buckminster Fuller, the inventor of the geodesic dome. The first of these to be discovered, C_{60}, has the form of a perfect icosahedron (see Figure 6). Condensed out of the hellish heat of a carbon arc, it depends on thermal vibrations to curve the familiar planar hexagonal lattice of graphite onto itself to form a three dimensional structure. No architect or bee is required. Atomic properties of carbon select the strut length. Thermal motion searches the space of possible

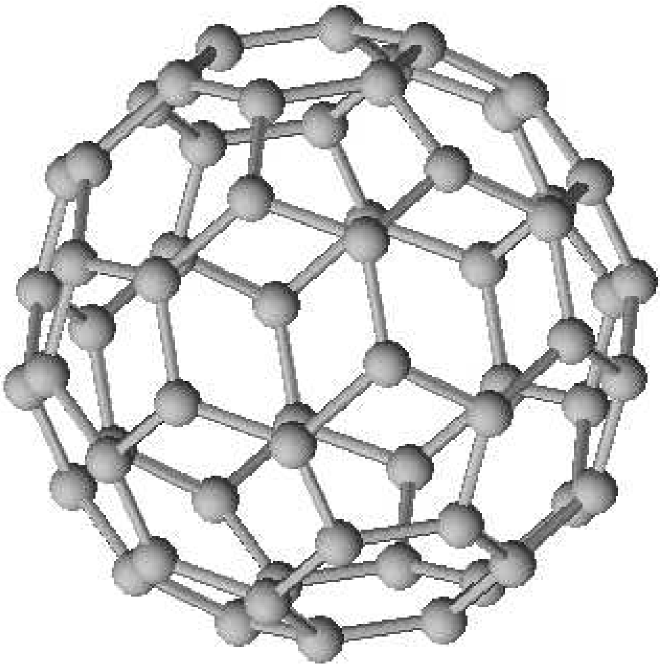

Figure 6: C_{60} a spontaneously formed dome structure

configurations; a small fraction of the molecules settle into the local energy minima represented by C_{60} and its sisters.

If the bees can't rely upon spontaneous self-assembly to build their hives, must they have a plan in mind before they start? Since they can't draw, the mind would have to be where they held any plans. While we can't rule this out, it seems unlikely. The requirement is that they can execute a program of work. A bee arriving on the construction site with a load of wax must, in the darkness, find an appropriate place to put it, for which they need a set of rules:

> If the cell is high enough to crawl into, put no more wax on it,
> otherwise if the cell has well formed walls add to their height,
> otherwise if it is a cell base smaller than your own body diameter, expand it,
> otherwise start building the wall up from the base. . .

No internal representation of a completed comb need be present in the bee's mind. The same rules, simultaneously present in each of a hive full of identical cloned sisters, along with the structural properties of beeswax, produce the comb as an emergent complex structure. The key here is the interaction between behavioral rules and an immediate environment that is changed as the result of the behaviour. The environment, the moulded wax, records the results of past behaviour and conditions future behaviour. But for rules to be converted into behaviours by the bees, the bees must have internal 'states of mind', and be able to change their state of mind in response to what their senses are telling them. A bee that is busy laying down wax is in a different state of mind from one foraging for pollen and their behavioral repertoire differs as a result.

As we have argued above, what an architect does is not so different. Architects produce drawings, not buildings or hives, but producing a drawing is an interactive process in which the architect's internal state of mind, his knowledge of the rules and stylistic conventions of the epoch, produces behaviour that modifies the immediate environment—the paper. The change to the paper creates a new environment, modifying his state of mind and calling into action other learned rules and skills. The drawing is an emergent property of the process, not something that pre-existed as a complete internal representation before the architect put pencil to paper.

0.3 The Demonic challenge

Purposeful labour depends upon the ability to form and follow goals. A goal is a representation of a state of affairs that does not exist plus a motivation to achieve it. Although bees do not have the goal processing capabilities of the human mind, they nonetheless follow simple goals. Goal processing, from simple, reactive programs hard-wired in the neural circuitry of insects, to the much more adaptive and sophisticated rational planning capabilities of humans, is the mechanism that distinguishes the constructive activity of humans and bees from the blind efforts of Watt's engines. An engine transforms energy in one form to another, but it does not act to achieve states of affairs, unlike bees that build or humans that labour.

There is a hidden connection between purposeful labour and work in the engineering sense. Any purposeful activity overcomes physical resistance and involves *work*, measured in watts, for which we must be fueled by calories in our food; the hidden connection comes from the realization that, at least in principle, purposeful labour could itself be a source of fuel.

Recall that Watt's key invention was the separate condenser for steam engines, which saved fuel by preventing wasteful condensation of steam within the cylinder of the engine. In the years after Watt's invention, it came to be realized that the thermal efficiency of steam engines could be improved by maximizing the pressure drop between the boiler and the condenser. A series of inventions followed to take advantage of this principle: Trevithick's high pressure engine, the double and then the triple expansion engine. These had the effect of increasing the amount of effective work that could be extracted from a given amount of heat. But successive gains in efficiency proved harder to come by. The amount of work obtained per calorie of heat could be increased, but not without limit.

It was understood that work could be converted into heat, for instance through friction, and heat could be converted back into work, for instance by a steam engine. But if you convert work into heat, and heat back into work, you always end up with less work than you put in. In converting work into heat, the number of calories of heat obtained per kilowatt hour of work is constant—conversion of work into heat can be done with 100 percent efficiency. The reverse is not true. Heat can never be fully converted into useful work.[9] The practical imperative of improving steam engines gave rise to the theoretical study of the laws governing heat, the laws of thermodynamics.

One of the first formulations of the second law of thermodynamics was that heat will never spontaneously flow from somewhere cold to somewhere hot.[10] This implied that, for instance, there was no chance of transferring the heat wasted in the condenser of a steam engine back to the boiler where it would boil more water. Thermodynamics ruled out perpetual motion machines.

But James Clerk Maxwell, one of the early researchers in thermodynamics, came up with an interesting paradox.

> One of the best established facts of thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which temperature and pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. But if we can conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being would be able to do that which is presently impossible to us. For we have seen that the molecules in a vessel full of air at a uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without the expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics. (James Clerk Maxwell, 1875, pp. 328–329)

The configuration of the thought experiment is shown in Figure 7. As the experiment runs the gas on one side heats up while that on the other side cools down. The end result is a preponderance of slow molecules in cavity A, fast ones in cavity B. Since heat is nothing more than molecular motion, this means that A has cooled down while B has warmed up. No net heat has been added, it has just re-distributed itself into a form that becomes useful to us. Since B is hotter than A, the temperature differential can be used to power a machine. We can connect B to a boiler and A to a condenser and obtain mechanical effort. An exercise of purposeful labour by the demon outwits the laws of thermodynamics. (Norbert Wiener coined the term 'Maxwell demon' for the tiny 'being' envisaged in the thought experiment.) It seems that the second law of thermodynamics expresses the coarseness of our senses rather than the intractability of nature.

0.4 Entropy

One perspective on the devilment worked by Maxwell's demon is that it has *reduced the entropy* of a closed system. The idea of entropy was introduced by Clausius in 1865 (see Harrison, 1975) with the equation

$$\Delta S = \Delta Q/T \tag{1}$$

where ΔS is the change in entropy of a system consequent upon the addition of a quantity of heat ΔQ at absolute temperature T.[11] According to Clausius's equation adding heat to a system always increases its entropy (and subtracting heat always lowers entropy) but the magnitude of the change in entropy is inversely related to the initial temperature of the system. Thus if a certain amount of heat is transferred from a hotter to a cooler region the increase in entropy in the cooler region will be greater than the reduction in entropy in the hotter, and overall entropy rises. Conversely, if heat is transferred from a colder to a hotter region entropy falls. Clausius's concept of entropy as an abstract quantity allowed him to give the second law of thermodynamics its canonical form: the entropy of any closed system tends to increase over time.

Using (1) we can readily see that Maxwell's demon violates the second law of thermodynamics. Suppose the demon has been hard at work for some time, so that B is hotter than A, specifically B is at 300° Kelvin and A is at 280° Kelvin. He then transfers ΔQ =1 joule of heat from A to B. In doing so he reduces the

[9] Carnot was able to show that the efficiency of heat engines depended on the temperature difference between heat source, for example the boiler, and the heat sink, for example a steam engine's condenser.

[10] This formulation was due to Clausius in 1850; see Porter (1946, pp. 8–9).

[11] At this stage the concept of entropy remains firmly linked to the sort of practical considerations, namely steam engine design, that gave rise to thermodynamics. Later, as we shall see, it becomes generalized.

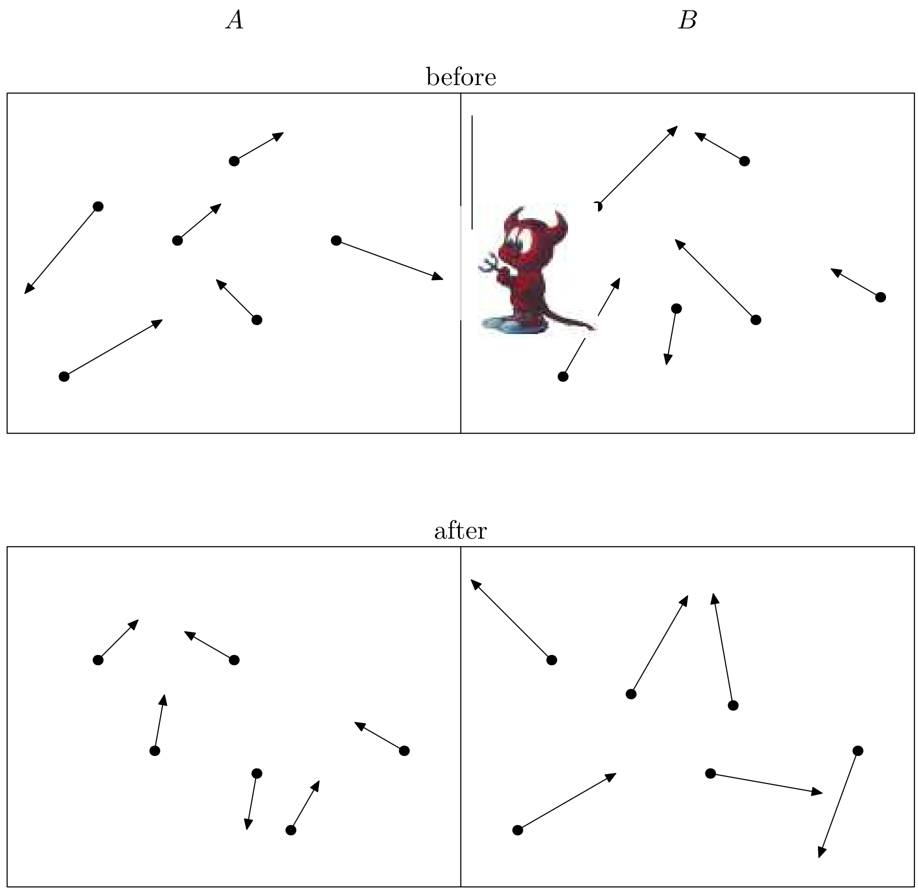

Figure 7: Gas initially in equilibrium. Demon opens door only for fast molecules to go from A to B, or slow ones from B to A. Result Slow molecules in A, fast in B. Thus B hotter than A, and can be used to power a machine.

entropy of A by $\frac{1}{280}$ joules per degree and increases the entropy of B by $\frac{1}{300}$ joules per degree giving rise to $\Delta S = \frac{1}{300} - \frac{1}{280} = -\frac{1}{4200}$, a net reduction in entropy, contrary to the second law.

Clausius's formulation of entropy did not depend in any way upon the atomic theory of matter. Maxwell's proposed counter-example to the second law was explicitly based on atomism. With Boltzmann, entropy is placed on an explicitly atomistic foundation, in terms of an integral over molecular *phase space*.

$$S = -k \int f(v) \log f(v) dv \tag{2}$$

where v denotes volume in six-dimensional phase space, $f(v)$ is the function that counts the number of molecules present in that volume, and k is Boltzmann's constant.

The concept of phase space is a generalization of our normal concept of three-dimensional space to incorporate the notion of motion as well as position. In a three-dimensional coordinate system the position of each molecule can be described by three numbers, measurements along three axes at right angles to one another. We usually label these numbers x, y, z to denote measurements in the horizontal, vertical and depth directions. However each molecule is simultaneously in motion. Its motion can likewise be broken into components of horizontal, vertical and depth-wise motion which we can write as m_x, m_y, m_z, representing motion to the left, up and back respectively. This means that a set of six coordinates can fully describe both the position and motion of a particle.

In Boltzmann's formula, the letter v denotes a range of possible values of these co-ordinates. For example, a volume 1 mm cubed on the spatial axes and 1 mm per second on the motion axes. The function $f(v)$ would then specify how many molecules there were in that cubic millimeter with a range of velocities within 1 mm per second in each direction. Boltzmann's formula relates the entropy of a gas, for instance steam in a piston, to the evenness of its distribution in this six dimensional space: the less even the distribution the lower the entropy. This point is illustrated in simplified manner in Table 1. Suppose we have just two cells in phase space, and eight atoms that can be in one cell or the other. The table shows how the entropy depends on the location of the atoms, lowest when all 8 are in one cell, and highest when they are evenly divided between the cells. (Note that the minus sign in Boltzmann's formula is needed to make entropy increase with the evenness of the distribution, consistent with Clausius's earlier formulation.)

Contents of cells 1, 2	$f(1)\log f(1) + f(2)\log f(2)$	Entropy, S
8, 0	$8(2.079) + 0 = 16.636$	$-16.636k$
7, 1	$7(1.946) + 1(0) = 13.621$	$-13.621k$
6, 2	$6(1.792) + 2(0.693) = 12.137$	$-12.137k$
5, 3	$5(1.609) + 3(1.099) = 11.343$	$-11.343k$
4, 4	$4(1.386) + 4(1.386) = 11.090$	$-11.090k$

Table 1: Boltzmann's entropy: Illustration

Boltzmann also showed that it is possible to reformulate the idea of entropy using the concept of the 'thermodynamic weight' of a state:

$$S = k \log W \tag{3}$$

The thermodynamic weight W is the number of physically distinct microscopic states of the system consistent with a given 'macro' state, described by temperature, pressure and volume. This concept is the key to understanding the second law. Recall that the entropy of closed systems tends to increase, that is they move into macro-states of progressively higher thermodynamic weight until they reach equilibrium. States with higher weight are *more probable*. So the second law of thermodynamics basically says that systems evolve into their most probable state.

A simple analogy may be helpful here. Suppose a 'fair' coin is flipped ten times. What is the most likely ratio of heads to tails in the sequence of flips? The obvious answer, 5/5, is correct. Now, what is the most likely specific sequence of heads and tails? Trick question! There are $2^{10} = 1024$ such sequences and they are all equally likely. The sequence featuring 10 heads has probability $\frac{1}{1024}$; so does the sequence with 5 heads followed by 5 tails; so does the sequence of strictly alternating heads and tails, and so on. The reason why a 5/5 ratio of heads to tails is most likely is that there are more specific sequences corresponding to this ratio that there are sequences corresponding to 10/0, or 7/3, or any other ratio. It's easy to see there is only one sequence corresponding to all heads, and one corresponding to all tails. To count the sequences that give a 5/5 ratio, imagine placing the 5 heads into 10 slots. Head number 1 can go into any of the ten slots; head number 2 can go into any of the remaining 9 slots, and so on, giving $10 \times 9 \times 8 \times 7 \times 6$ possibilities. But this is an over-statement, because we have treated each head as if it were distinct and identifiable. To get the right answer we have to divide by the number of ways 5 items can be assigned to 5 slots, namely $5 \times 4 \times 3 \times 2 \times 1$. This gives 252 possibilities. Thus the 'macro' result, equal numbers of heads and tails, corresponds to 252 out of the 1024 equally likely specific sequences, and has probability $\frac{252}{1024}$. By the same reasoning we can figure that a 6/4 ratio corresponds to 210 possible sequences, a lower 'weight' than the 5/5 ratio.

The number of possible states of a real gas in six-dimensional phase space is hard to visualize, so to explicate the matter further we'll examine a simpler system, namely a two-dimensional *lattice gas* (Frisch et al, 1986). The 'molecules' in such a stylized gas move with constant speed, one step along the lattice per unit time (see Figure 8). Where the lines of the lattice meet, molecules can collide according to the rules of Newtonian dynamics, so that matter, energy and momentum are conserved in each collision. The different ways in which collisions occur can be summarized by two simple rules:

(1) If a molecule arrives at an intersection and no molecule is arriving on the diagonally opposite path, then the molecule continues unimpeded.

(2) If two molecules collide head on they bounce off in opposite directions, as shown in Figure 9.

Lattice gases are a drastic simplification of real gases, but they are useful tools in analysing real situations. The simple rules governing the behaviour of lattice gases make them ideal models for simulation in computer software or special purpose hardware (Shaw et al, 1996).

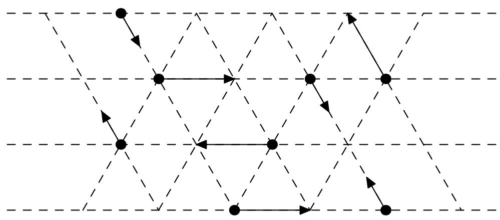

Figure 8: The molecules in a lattice gas move along the lines of a triangular grid with fixed velocities

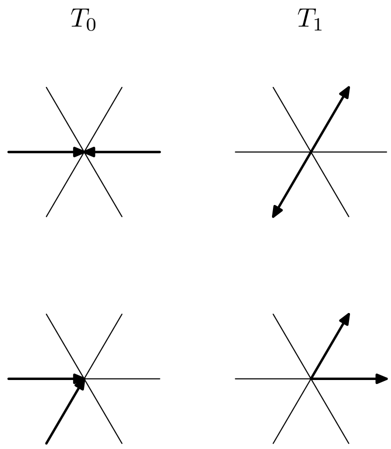

Figure 9: Collisions in a lattice gas: 'Molecules' colliding head on bounce off at 60° angles (above). In other cases the collision is indistinguishable from a miss (below). In all cases Newtonian momentum and energy are conserved.

Since the velocity of the molecules in a lattice gas is fixed, the temperature of the gas can't change (this would involve a rise or fall in the molecules' speed). So Maxwell's original example of a being with precise senses, able to sort molecules by speed, is inappropriate. But we can invent another demon to guard the trapdoor. Instead of letting only fast molecules through from A to B, this being will keep the door open unless a molecule approaches it from side B. Thus molecules approaching from side A are able to pass into B, but those in B are trapped. The net effect is to raise the pressure on side B relative to A while leaving temperature unchanged.

A lattice gas has only a finite number of lattice links on which molecules can be found, and since the molecules move with a constant velocity, Boltzmann's formula (3) simplifies to:

$$S = -kn \sum_i p_i \log p_i \tag{4}$$

where p_i is the probability of the node being in state i and n is the number of nodes. The weighted summation over the possible states has the effect of giving us the mean value of $\log p$. Suppose we have a very small pair of chambers, A and B, each of which initially has n nodes, and each containing $3n$ randomly distributed molecules. Then each of the six incoming paths to a node will have a 50 percent chance of having a molecule on it. We have $6n$ incoming paths to our nodes, and each of these has two equally likely states: a particle is or is not arriving at each instant. Each incoming path contributes $k \log 2 = 0.693k$. The total entropy of the chamber is then six time this or:

$$\text{Entropy of A in equilibrium} = 4.158kn.$$

Now suppose that our demon has been operating for some time, letting n particles pass from A to B, so that A now contains $2n$ particles and B contains $4n$ particles. In A, the probability of a molecule coming down any one of the paths is now only $\frac{1}{3}$. We can calculate the current entropy contribution of each incoming path as follows:

Number of particles	probability, p_i	$\log p_i$	entropy, $-kp_i \log p_i$
0	$\frac{2}{3}$	-0.405	$0.27k$
1	$\frac{1}{3}$	-1.098	$0.366k$
total			$0.636k$

The entropy of A after n particles have been transferred by the demon is $3.816kn$ which is less than before he got to work. By symmetry of complementary probabilities the entropy of chamber B will be the same,[12] thus the whole closed system has undergone a reduction in entropy.

This establishes that when an initially dispersed population of particles—the gas molecules in our case—is concentrated, entropy falls.[13] This is because there are a greater number of possible microstates compatible with dispersion than with concentration, and entropy is just the log of the number of microstates.

Consider in this light the work of the bees building their hive. There are two aspects to the work:

(1) The bees first have to gather wax and nectar from flowers dispersed over a wide area and bring it to the hive.

(2) They must then form the wax into cells and place the concentrated nectar in these as honey.

Both processes are entropy-reducing with respect to the wax and the sugar. The number of possible configurations that can be taken on by wax within the few litres volume of a hive is enormously less than the number of possible configurations of the same wax, dispersed among plants growing over tens of thousands of square meters of ground. Similarly the chance that the wax, if randomly thrown together within the hive, should assume the beautifully regular structure of a comb, is vanishingly small. That the wax should be in the hive in the first place, is, in the absence of bees, highly improbable; that it should be in the form of regular hexagons even more so.

The second law of thermodynamics specifies that the total entropy in a closed system tends to increase, but the bees and their wax are not a closed system. The bees consume chemical energy in food to move the wax. If we include the entropy increase due to food consumed, the second law is preserved.

Men and horses

Let us return to the question we asked in section 0.1: Why did the introduction of the steam engine, which made redundant the equine workers of the pre-industrial age, not also replace the human workers? We can make a rough analogy between the work done by horses in past human economies and the work done by the bees in transporting wax and nectar from flower to hive. This is in the main sheer effort, work in Watt's sense. Horses bringing bricks to a building site or bees transporting wax are doing similar tasks. What remains, the construction of the hive after the work of transportation is done or the building of the house once the bricks are delivered, is something no horse can do. Construction involves a complex program of actions deploying grasping organs, hands, mandibles, beaks etc., in which the sequence of operations is conditioned by the development of the product being made. Human construction differs from that of a bee or a bird in:

(1) the way in which the program of action comes into being;

(2) the way in which it is transmitted between individuals of the species; and

(3) the form in which it is materialized.

In the social insects the programs of action largely come into being through the evolutionary process of natural selection. They are transmitted between parents and their offspring genetically encoded in DNA, and they are materialized in the form of relatively fixed interactions between components of the nervous system and general physiology. In humans the programs of action are themselves products that can have a representation external to the organism, in speech or some form of notation. Speech and notation act both as a means of transmission between individuals, and as a possible form of materialization of work programs while the work is being carried out—as for example, when one cooks from a recipe or follows a knitting

[12] This will not generally be the case; we have chosen the particle densities so as to ensure this.

[13] This is true on the assumption that the potential, gravitational or electrostatic, of the particles is unchanged by the process of concentration as in our example.

pattern. The ability to make and distribute new work programs distinguishes human labour from that of bees and is the key to cultural evolution.

But even the work of transport requires a program of action, requires guidance if it is to reduce entropy. Transport is not diffusion. It moves concentrated masses of material between particular locations, it does not spread them about willy nilly. Without guidance there is no entropy reduction. A horse, blessed with eyes and a brain as well as big muscles, will partially steer itself, or at least will do better than a bicycle or car in this respect. But teams still needed teamsters, if only to read signposts.

The steam railway locomotive revolutionized land transport in the nineteenth century, quickly replacing horse traction for long overland journeys. Guidance by steel track made steam power the great concentrator, bringing grain across prairies to the metropolis. Railway networks are action programs frozen in steel, their degrees of freedom discrete and finite, encoded in points. Point settings, signaled by telegraph, coordinate the orderly movement of millions of tons according to precise published timetables. Human work did not all lend itself so readily to mechanization.

Chapter 1

Problematizing information

We have suggested that doing purposeful productive labour typically reduces entropy. Such entropy-reducing work requires information in two forms, an action plan or capacity for behaviour, and information coming in from the senses to monitor the implementation of the action plan. Productive labour also involves work in Watt's sense of overcoming physical resistance. As such it consumes energy and produces an entropy increase in the environment that more than compensates for the entropy reduction effected in the object of labour. We have also seen how Maxwell postulated that it should be possible to reduce the entropy of a gas if there existed a being small enough to sort molecules. In this case the being would be using information from its senses, and in its action plan, to produce an entropy reduction in the gas with no corresponding increase elsewhere. Up to now we have not rigorously defined what we mean by information. Once this is done, we shall see the deeply hidden flaw in Maxwell's argument.

1.1 The Shannon–Weaver concept of information

The philosopher Gaston Bachelard argues that the formation of a science is characterized by what he calls an 'epistemological break', which demarcates the language and ideas of the science from the pre-scientific discourses that appeared to deal with the same subject matter. Appeared to deal with the same subject, but did not really do so. For one of the characteristics of an epistemological break is a change in the *problematic*, which means roughly, the set of questions to which the science provides answers. With the establishment of a science the conceptual terrain shifts both in terms of the answers given and, more importantly, in terms of the questions that researchers regard as relevant.

The epistemological break that established information theory as a science occurred in the middle of the last century and is closely associated with the name of Claude Shannon. We saw how Watt, seeking to improve the efficiency of steam pumps, contributed not only to an industrial revolution, but to a scientific revolution when he asked questions about the relationship between work and heat. From this problematic were born both a convenient source of power, and our understanding of the laws of thermodynamics. Shannon's revolution also came from asking new questions, and asking them in a very practical engineering context. Shannon was a telephone engineer working for Bell Laboratories and he was concerned with determining the capacity of a telephone or telegraph line to transmit information. Watt formalized the concepts of power and work in an attempt to measure the efficiency of engines. Shannon formalized the concept of information through trying to measure the efficiency of communications equipment. Practice and its problems lead to some of the most interesting truths.

To measure the transmission of information over a telephone line, some definite unit of measurement is needed, otherwise the capacity of lines of different quality cannot be meaningfully compared. According to Shannon the information content of a message is a function of how surprised we are by it. The less probable a message the more information it contains. Suppose that each morning the radio news told us "We are glad to announce that the Prime Minister is fit and well." We would soon get fed up. Who would call this news? It conveys almost no information. "Reports are just reaching us of the assassination of the Prime Minister." That is news. That is information. That is surprising.

A daily bulletin telling us whether or not the Prime Minister was alive would usually tell us nothing, then on one day only would give us some useful information. Leaving aside the circumstances of his death, if an announcement were to be made each morning, there would two possible messages

0 'The P.M. lives'

Binary Code	Length	Meaning	Probability
0	1	False, False	$\frac{4}{9}$
10	2	False, True	$\frac{2}{9}$
110	3	True, False	$\frac{2}{9}$
111	3	True, True	$\frac{1}{9}$

Table 1.1: A possible code for transmitting messages that are true $\frac{1}{3}$ of the time

1 'The P.M. is dead'

If such messages were being sent over the sort of telegraph system that Shannon was concerned with, one could encode them as the presence or absence of a short electrical pulse, as a binary digit or 'bit' in the widely understood sense of the word. Shannon defines a bit more formally as the amount of information required for the receiver of the message to decide between two equally probable outcomes. For example, a sequence of tosses of a fair coin can be encoded in 1 bit per toss, such that heads are 1 and tails 0.

What Shannon says is that if we are sending a stream of 0 or 1 messages affirming or denying some proposition, then unless the truth and falsity of the proposition are equally likely these 0s and 1s contain less than one bit of information each. In that case there will be a more economical way of sending the messages. The trick is not to send a message of equal length regardless of its content, but to devise a system where the more probable message-content gets a shorter code.

For example, suppose the messages are the answer to a question which we know a priori will be true one time in every three messages. Since the two possibilities are not equally likely Shannon says there will be a more efficient way of encoding the stream of messages than simply sending a 0 if the answer is false and a 1 if the answer is true. Consider the code shown in Table 1.1. Instead of sending each message individually we package the messages into pairs, and use between one and three binary digits to encode the 4 possible pairs of messages. Note that the shortest code goes to the most probable message, namely the sequence of two 'False' answers with probability $\frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$. The codes are set up in such a way that they can be uniquely decoded at the receiving end. For instance, suppose the sequence '110100' is received: checking the Table, we can see that this can only be parsed as 110, 10, 0, or True, False, False, True, False, False.

To find the mean number of digits required to encode two messages we multiply the length of the codes for the message-pairs by their respective probabilities:

$$\frac{4}{9} + 2 \times \frac{2}{9} + 3 \times \frac{2}{9} + 3 \times \frac{1}{9} = 1\frac{8}{9} \approx 1.889 \tag{1.1}$$

which is less than two digits.

Shannon came up with a formula which gives the shortest possible encoding for a stream of distinct messages, given the probabilities of their individual occurrences.

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{1.2}$$

The mean information content of an ensemble of messages is obtained by weighting the log of the probability of each message by the probability of that message. He showed that no encoding of messages in 1s and 0s could be shorter than this. The formula gave him an irreducible minimum of the number of bits needed to transmit a message stream: this minimum was, he said, the real information content of the stream. Using Shannon's formula we can calculate the information content of the data stream encoded in the example above.

$$-\frac{4}{9} \times \log_2 \frac{4}{9} - \frac{2}{9} \times \log_2 \frac{2}{9} - \frac{2}{9} \times \log_2 \frac{2}{9} - \frac{1}{9} \times \log_2 \frac{1}{9} \approx 1.837 \tag{1.3}$$

Since our code used $1\frac{8}{9} \approx 1.889$ bits for each pair of messages, we see that in principle a better code may exist.

In his 1948 article Shannon notes:

> Quantities of the form $H = -\sum_{i=1}^{n} p_i \log p_i$ play a central role in information theory as measures of information, choice and uncertainty. The form of H will be recognized as that of entropy as defined

in certain formulations of statistical mechanics where p_i is the probability of a system being in cell i of its phase space. H is then, for example the H in Boltzmann's famous H theorem. We shall call $H = -\sum p_i \log p_i$ the entropy of the set of probabilities $p_1, ..., p_n$.

Shannon thus discovers that his measure of information is the same as Boltzmann's measure of entropy and decides that entropy and information are the same thing. Armed with this realization we can go back to the problem left to us by Maxwell. Could a sufficiently tiny entity violate the laws of thermodynamics by systematically sorting molecules?

Physicists have concluded that it is not possible. Leo Szilard, for example, pointed out that to decide which molecules to let through, the demon must measure their speed. He showed that these measurements (which would entail bouncing photons off the molecules) would use up more energy than was gained. Maxwell's demon, to vary the theological metaphor, was a *deus ex machina* (like Newton's God), able to know by immaterial means; Szilard's advance was to emphasize that knowledge or information is physical and can only come about by physical means. Leon Brillouin (1951) extended Szilard's analysis by pointing out that at a uniform temperature, black body radiation in the cavity would be uniform in all directions, preventing the demon from seeing molecules unless he had an additional source of light (and hence energy input).

It is possible, however, to build an automaton that acts as a Maxwell demon for a lattice gas. As we said before such gases can be simulated in software, or in hardware (see Figure 1.1), with each gas cell represented by a rectangular area of silicon and the paths taken by the molecules represented by wires. In such a system the demon himself is an automaton, a logic circuit, as in Figure 1.2. A circuit like this really does work: it transfers virtual gas molecules from chamber A to chamber B. Why does this work in apparent conflict with the laws of thermodynamics?

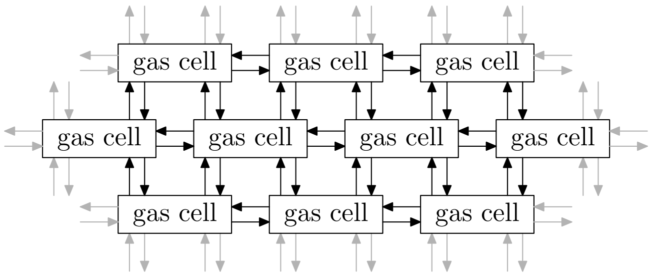

Figure 1.1: A lattice gas can be built in electronic hardware: each gas cell is represented by a rectangular area of silicon and the paths taken by the molecules are represented by wires.

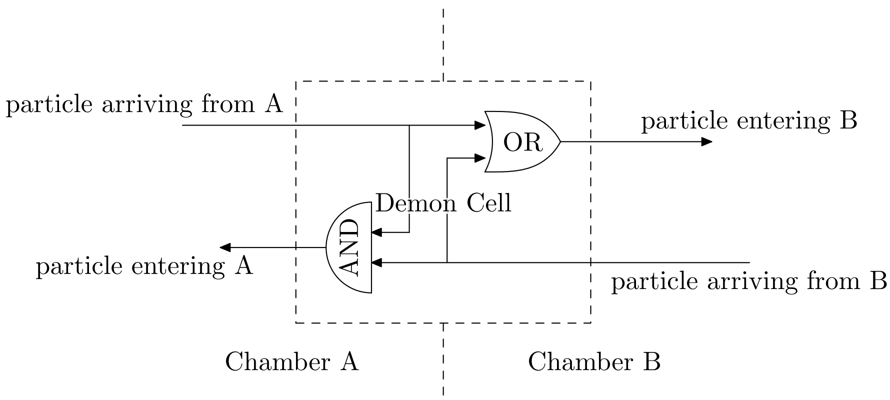

Figure 1.2: In a lattice gas, Maxwell's demon can be implemented with this logic circuit.

The behaviour of the demon is summarized in Table 1.2. Notice that while there are 4 possible combinations of input conditions, there are only 3 combinations of output conditions. This implies that we are moving

input from A	input from B	output to A	output to B	Comment
No	No	No	No	No molecules involved
No	Yes	No	Yes	Door shut, molecule bounces back to B
Yes	No	No	Yes	Molecule goes from A to B
Yes	Yes	Yes	Yes	Molecules bounce off one another

Table 1.2: The action plan of the demon

x	y	x AND y	x OR y
false	false	false	false
false	true	false	true
true	false	false	true
true	true	true	true

Table 1.3: Tabulation of the functions x AND y, x OR y

from a system with a higher thermodynamic weight to one with a lower weight, which is what we would expect for an entropy-reducing machine. Just how much it reduces entropy depends on the probabilities of occurrence of incoming particles from each side.

Suppose that the system is in equilibrium and that the probability of occurrence of a particle on the incoming paths on each side is 50 percent in each time interval. In that case each of the 4 possible input configurations in Table 1.2 is equiprobable and has an entropy of 2 bits$= \log_2 4$. Applying Shannon's formula (1.2) to the output configurations we get

$$\frac{1}{4}\log_2 4 + \frac{1}{2}\log_2 2 \frac{1}{4}\log_2 4 = \frac{1}{4}\times 2 + \frac{1}{2}\times 1 + \frac{1}{4}\times 2 = 1\frac{1}{2} \quad (1.4)$$

an entropy reduction of half a bit per time step. The key to how this can happen lies in the nature of the components used, logic gates for the functions AND and OR.

Rolf Landauer (1961) pointed out that any irreversible logic gate must destroy encoded information and in the process must dissipate heat. An irreversible logic gate is one whose inputs can't be determined from an examination of their outputs. Consider gates with two inputs and one output, such as the AND and OR gates whose truth functions are tabulated in Table 1.3. Roughly speaking they take two bits in and generate one bit out, thus destroying information within the system defined by the lines connecting the gates. Landauer argues that the lost information, i.e., the entropy reduction within the logic circuit, results in an increase in the entropy of the environment. Each time a logic circuit of this type operates, the lost internal entropy shows up as waste heat. By applying Shannon's formula (1.2) to the output of the AND gate we get the following:

Output	p_i	$-p_i \log_2 p_i$
false	$\frac{3}{4}$	≈ 0.311
true	$\frac{1}{4}$	0.5
	1	0.811

The output has an entropy of *less* than one bit. Given that 2 bits of information went into the gate, a total of 1.189 bits are lost in processing the inputs. Since the probability structure of OR gates is the same, a similar information loss occurs going through these.

Information engines as heat engines

Boltzmann's constant (see equation 2) has the dimension joules per log-state degree Kelvin. Landauer saw that one can use this constant to convert entropy in Shannon's form, measured in log-states, to energy. The equation he established is

$$e = \ln(2)ktb \quad (1.5)$$

e represents the energy-equivalent, t is temperature in degrees Kelvin, b is the number of bits, and k is Boltzmann's constant, which has a value of about 1.38×10^{-23} joules per degree Kelvin. The remaining term in the conversion is the natural log (ln) of 2, to get us from the natural logarithms used by Boltzmann to the base-2 logarithms used in Shannon's information theory.

Using Landauer's equation we can calculate the heat energy, e_{AND}, generated by a single operation of an AND gate, in which 1.189 bits are lost:

$$e_{\text{AND}} = 1.189 \ln(2) kt$$

At room temperature, or roughly 300° Kelvin, this is 3.4×10^{-21} joules each time the gate switches. This is a very, very small quantity of energy which is at present mainly of theoretical interest. What it represents is the theoretical minimal energy cost of operating a two-input irreversible logic gate.

Now look again at the demon cell in Figure 1.2, which has a pair of input logic gates. The process of deciding whether to open or close the trapdoor must consume certain minimum Landauer-energy. The energy consumed by the logical decision to open or close the barrier makes the demon ineffective as a power source.

Watt started out investigating how to convert heat into work efficiently; he was concerned with minimizing the heat wasted from his engines. Since Landauer we have known that information processing, too, must dissipate heat, and that information processing engines are ultimately constrained by the same laws of thermodynamics as steam engines. We can calculate the thermodynamic efficiency of an information processing machine just as we calculate the efficiency of a steam engine. If a processor chip of the year 2000 had roughly 6 million gates and was clocked at 600Mhz, its dissipation of Landauer energy would then be $(600 \times 10^6) \times (6 \times 10^6) \times (3.4 \times 10^{-21}) = 16.3\mu$w, or 16 millionths of a watt. This is insignificant relative to the electrical power consumption of the chip, which would be of the order of 20 watts. It implies a thermodynamic efficiency of only around 0.0001%. As a point of comparison, steam engines prior to Watt had an efficiency of about 0.5%. The steam turbines in modern power stations convert around 40% of the heat used into useful work. Two centuries of development raised the efficiency of steam power by a factor of about 100.

In thermodynamic terms a Pentium processor looks pretty poor compared to an 18th century steam engine: the steam engine was 500 times more efficient! But if compare a Pentium with the Manchester Mk1, the first electronic stored program computer (Lavington, 1980), we get a different perspective. The Pentium has at least a thousand times as many logic gates, has a switching speed a thousand times greater and uses about one hundredth as much electrical power as the venerable valve-based Mk1. In terms of thermal efficiency, this represents an improvement factor of 100,000,000 in fifty years. If improvements in heat engine design from Watt to Parsons powered the first two industrial revolutions, the third has benefited from an exponential growth in efficiency that was sixteen times as rapid.[1]

We know from Carnot's theory that there is little further room for improvement in heat engines. Most of the feasible gains in their efficiency came easily to pioneers like Watt and Trevithick. We're now left with marginal improvements, such as the ceramic rotor blades that allow turbine operating temperatures to creep up. In the case of computers too, efficiency gains will eventually become harder to attain. There is still, to quote Feynman, "plenty of room at the bottom". That is, there is mileage yet in miniaturization. We have room for about a million-fold improvement before computers get to where turbines now are. However, as we take into account the growing speed and complexity of computers, the thermodynamic constraint on data processing will come to be of significance. On the one hand, if the efficiency of switching devices continues to grow at its current rate, they will be at close to 100% in about 30 years. On the other hand, as computers get smaller and faster the job of getting rid of the Landauer-energy, thrown out as waste heat, will get harder. In the 27 years following the invention of the microprocessor the number of gates per chip rose by a factor of some 3000. Processor speeds increased about 600-fold over the same period. Table 1.4 projects this rate of growth into the next century.

From being insignificant now, Landauer heat dissipation becomes prohibitive in about 30 years. A microprocessor putting out several kilowatts, as much as several electric heaters, is not a practical proposition. There is a time limit on the current exponential growth in computing power.

That is not to say that computer technology will stagnate in 40 years. Landauer's equation (1.5) has a free variable in *temperature*. If the computer is super-cooled, its heat dissipation falls. But once we're

[1] Heat engine efficiency improved about ten-fold per century. Information engines have been improving at a factor of about 10^{16} per century.

year	*gates*	*clockspeed*	*landauer watts*
2000	8×10^6	600Mhz	16.3μw
2005	3.4×10^7	1.9Ghz	230μw
2010	1.5×10^8	6.4Ghz	3.24mw
2015	6.4×10^8	21Ghz	45.7mw
2020	2.8×10^9	68Ghz	643mw
2025	1.2×10^{10}	224Ghz	9.06w
2030	5.1×10^{10}	733Ghz	128w
2035	2.2×10^{11}	2.4Thz	1.80Kw
2040	9.5×10^{11}	7.8Thz	25.4Kw

Table 1.4: Projected Landauer heat dissipation in 21st century computers operating at 300° Kelvin.

in that game the rate of improvement in computer performance comes to be limited by improvements in refrigeration technology, and these are unlikely to be so dramatic.

1.2 Entropy reductions in action programs

Maxwell's demon cannot exist for real gases, but it can for lattice gases. If the demon really existed, he would reduce the laws of thermodynamics to the status of an anthropocentric projection onto reality. Lattice-gas devils, on the other hand, are not a threat to physics. They reduce the entropy of the gas, but only because they use logic gates with an external source of power. Nonetheless, their structure suggests something important. The demon reduces the entropy of the gas thanks to an action program which has four possible input states and only three possible output states.

We would suggest that this is not accidental: it would seem that *all production processes that produce local reductions in entropy are guided by an entropy-reducing action program.* Consider the bee once again, this time in its capacity as forager. In Maxwell's original proposal, the demon used its refined perception to extract energy from chaos. In reality a bee uses its eyes to enable it to extract energy from flowers. Were bees unable to see or smell flowers, their energy would be expended in aimless wandering followed by starvation. The bee uses information from its senses to achieve what, from its local viewpoint, is a reduction in entropy—the maintenance of homeostasis—albeit at a cost to the rest of the universe. To achieve this it requires a nervous system that performs entropy reduction on the input data coming into its visual receptors. At any given instant the bee's compound eyes are receiving stimuli from the environment. The number of possible different combinations of such stimuli is vastly greater than the number of instantaneous behavioural responses that it has while in flight—the modulation of the beat strength of a small number of thoracic muscles. In selecting one appropriate behavioural response out of a small repertoire, in response to a relatively large quantity of information arriving at its eyes, the bee's nervous system functions in the same sort of way as the AND gate in the demon-automaton of Figure 1.2. Having fewer possible outputs than inputs, it discards information and reduces entropy.

1.3 Alternative views of information

We have come across two approaches to the idea of entropy so far, deriving from classical thermodynamics and Shannon's communication theory respectively. From the 1960s onwards a third version has developed: that of computational complexity. Where classical concepts of entropy derived from mechanical engineering, and Shannon's concept from telecommunications engineering, the latest comes from computer science. The key concepts appear to have been independently developed by Chaitin in the US and Kolmogorov in Russia. Their presentation, while not contradicting what Shannon taught, gives new insights that are particularly helpful when we come to consider the role that information flows play in mass production industries.

The Chaitin–Kolmogorov concept of information

Chaitin's algorithmic information theory defines the information content of a number to be the length of the shortest computer program capable of generating it. This introduction of numbers is a slight shift of terrain. Shannon talked about the information content of *messages*. Whereas numbers as such are not messages, all

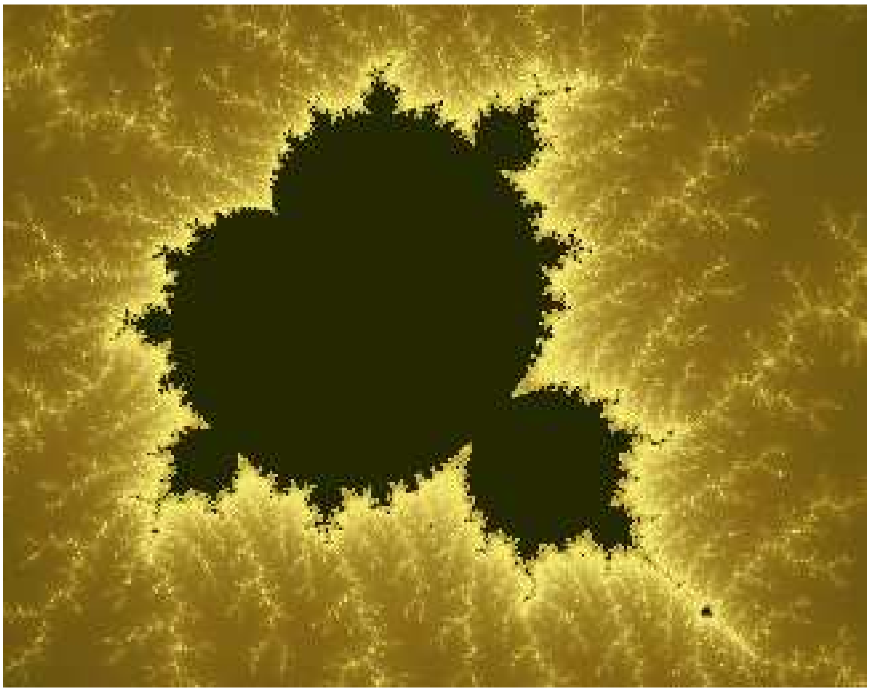

Figure 1.3: The Mandelbrot set, a complex image generated from a tiny amount of information.

coded messages are numbers. Consider an electronically transmitted message. It will typically be sent as a series of bits, ones and zeros, which can be considered as a binary number. An information theory defined in terms of numbers no longer needs the support of a priori probabilities. Whereas Shannon's theory depended upon the a priori probability of messages, Chaitin dispenses with this support.

As an example of the algorithmic approach consider the Mandelbrot set picture in Figure 1.3. This image is created by a very simple computer program.[2] Although the image file for the picture is large, about 6 million bits, a program to generate it can be written in a few thousand bits. If one wanted to send the picture to someone who had a computer, it would take fewer bits to send the program than to send the picture itself. This only works if both sender and receiver have computers capable of understanding the same program. Chaitin's definition of information has the disadvantage of seeming to make it dependent upon particular brand of computer used. One could not assume that the length of a program to generate the picture would be the same on an Apple as on an IBM.

In principle one could chose any particular computer and fix on it as the standard of measure. Alternatively one could use an abstract computer, much as Watt used an abstract horse. Chaitin follows Watt, using a *gedankenapparat,* the Universal Turing Machine, as his canonical computer. Thus he defines the information content of a sequence S as the shortest Turing machine tape that would cause the machine to halt with the sequence S on its output tape.[3]

Randomness and pi

An unsettling result from information theory is that random sequences of digits contain more information than anything else. According to common sense, information is the very opposite of randomness. We feel that information should be associated with order, but Shannon's identification of information and entropy amounts to equating information with *dis*order. To illustrate this let's compare a long random number with

[2] In fact it uses the formula $z = z^2 + c$ where z is a complex number.

[3] There is, in principle, no algorithm for determining the shortest Turing Machine tape for a sequence. $3 \div 7$ is a rule of arithmetic, an algorithm that generates the sequence 0.428571428571. So this sequence is presumably less random than 0.328571428771 (we changed two digits). But we can never be sure. This is a consequence of Gödel's theorem, which showed we cannot prove completeness of a consistent set of arithmetic axioms. There will be true statements that cannot be proven. If there existed a general procedure to derive the minimal Turing machine program for any sequence, then we would have a procedure to derive any true proposition from a smaller set of axioms, contra Gödel.

π. We know from Shannon that 1 million tosses of a fair coin generates 1 million bits of information. On the other hand, from Chaitin we know that π to a precision of a million bits contains much less than 1 million bits, since the program to compute π can be encoded using much fewer bits. Thus π must contain less information than a random sequence of the same length.

But what do we mean by random? And how can we tell if a number is random? The answer now generally accepted was provided by Andrei Kolmogorov, who defined a random number as *a number for which there exists no formula shorter than itself.* By Chaitin's definition of information a random number is thus incompressible: a random number of n bits must contain n bits of real information.

A fully compressed data sequence is indistinguishable from a random sequence of 0s and 1s. This not only follows directly from Kolmogorov and Chaitin's results but also from Shannon, from whom we have the result that for each bit of the stream to have maximal information it must mimic the tossing of a fair coin: be unpredictable, random.

We have a paradox: one million digits of π are more valuable and more useful than one million random bits. But they contain less information. They are more valuable because they are harder to come by. They are more useful because a host of other formulae use π. They contain less information because each and every digit of π was determined, before we started calculating it, by π's formula. Thus in a sense the entire expansion of π is redundant if we have its formula. Valuable objects are generally redundant. We thus have three concepts that we must distinguish with respect to sequences: their information content, their value, and their utility.

Concept	*Meaning*
Information	Length of program to compute the sequence.
Value	Cycles it takes to compute the sequence.
Utility	The uses to which the sequence can be put.

The *value* of a sequence is measured by how hard we must work to get it. π is valuable because it is so costly to calculate. We can measure the cost by the number of machine cycles a computer would have to go through to generate it.[4] As with information content, this definition is dependent upon what we take as our standard computer. A more advanced computer can perform a given calculation in fewer clock cycles than a more primitive one. For theoretical purposes any Universal computer will do. Information theorists typically use machine cycles of the Universal Turing Machine (UTM) for their standard of work. We will follow them in defining the information content of a sequence in terms of the length of the UTM program that generates it, and the value of a sequence in terms of the UTM cycles to compute it.

Now the UTM is an imaginary machine, a thought experiment, living in the platonist ideal world of the mathematician. Its toils are imaginary, consuming neither seconds nor ergs; its effort is measured in abstract cycles. But any physical computer existing in our material world runs in real time, and needs a power supply. Valuable numbers—tomorrow's temperature for example—whose computation requires large number of cycles on the Met Office super computers, take real time and energy to produce. The time depends on clock speed, and the energy depends on the computer's thermodynamic efficiency.[5] If we abstract from changes in computer technology, information value in UTM cycles is an indication of the thermodynamic cost of producing information. It measures how much the entropy of the rest of the universe must rise to produce the information.[6]

Having traced the conceptual thread of entropy from Boltzmann through Shannon to Chaitin, it is worth taking stock and asking ourselves if Chaitin's definition of entropy still makes sense in terms of Boltzmann's definition. To do this we need to move from numbers to their physical representation. A material system can represent a range of numbers if it has sufficient well-defined states to encode the range. Will a physical system in a state whose number has, according to Chaitin, a low entropy, have a low entropy according to classical statistical mechanics?[7]

What we will give is not a proof, but at least a plausible argument that this will be true. As a *gedanken* experiment we will consider a picture of the Mandelbrot set rendered on digital paper. Digital paper is a

[4]We are identifying the value of a sequence with what Bennett calls its logical depth. The homology with Adam Smith's definition of value should be evident.

[5]The UTM plays, for computational complexity theory, the role of Marx's "labour of average skill and intensity" in the economic theory of value. Improvements in computer technology are analogous to changes in the skill of the worker.

[6]This is what Norretranders calls *exformation*.

[7]We need this step if we are to apply Chaitin's theory to labour processes that produce real physical commodities. We need an epicurean not a platonist theory.

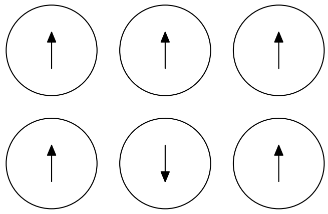

Figure 1.4: Configurations of parallel poles are unstable and tend to evolve towards the anti-parallel configuration.

proposed display medium made of thin films of white plastic. In the upper layer of the plastic there is a mass of small bubbles of oil, in the middle of each of which floats a tiny ball. One side of the ball is white and the other black. Embedded within the ball is a magnetized ferrite crystal with its North pole pointing towards the black end.[8] If the paper is embedded in an appropriate magnetic field all of the balls can be forced to rotate to have their white half uppermost, making the paper appear white. Applying a South magnetic pole to a spot on the paper will leave a black mark where the balls have rotated to expose their dark half. When it is passed through an appropriate magnetic printer, patterns can be drawn. A sheet of digital paper with a Mandelbrot set image on it nicely straddles the boundary between an industrial product and a number or information structure.

According to algorithmic information theory, the Mandelbrot set image represents a relatively low entropy state, since the length of the program to compute it contains fewer bits than the image. Does it also represent a low entropy state in statistical mechanics?

The second law of thermodynamics states that the entropy of a closed system is non-decreasing. So we would expect that a picture of the Mandelbrot state drawn on digital paper would tend to change into some other picture whose state would represent a higher entropy level. In fact there are good physical reasons why this will take place. If a local area is all white or all black, the magnetic poles are aligned as shown in the top of Figure 1.4. In this configuration the like poles tend to repel one another, and over time some of the poles will tend to flip to the configuration shown in the bottom half of the diagram.

The rate at which this occurs depends upon the temperature, the viscosity of the fluid in which the balls are suspended, and so on, but in the long run entropy will take hold. The image will gradually degrade to a higher entropy state, both in thermodynamic terms and in algorithmic terms. The program necessary to produce the degraded picture is bound to be longer than the program that produced the pristine one. Hence thermodynamic and algorithmic entropy measure the same scale.

The example we have given is stylized but the thermodynamic degradation of digital information is not hypothetical. Magnetic tape libraries have a finite life because of just this sort of flipping of the magnetized domains on which the information is stored.

1.4 Randomness and compressibility

You may find at this point that reason in you rebels at the idea that information content and randomness are equivalent. But this is what information theory teaches us, so it is worth considering and trying to resolve several apparent paradoxes that arise from information theory.

Kolmogorov identifies the randomness of a number with its incompressibility (via his “no shorter formula” proposition). There seems to be a contradiction—or at least a strong tension—between this conception of randomness as a property *of a number* and the “ordinary” conception of randomness as a property of a *mechanism for generating numbers.* (As in the statisticians’ talk of a “random variable” as a variable whose values are determined by the outcome of a “random experiment”.)

Random numbers contain non-random ones

To expose the tension, consider a random number generator (RNG). Suppose it’s a true quantum RNG, set to produce a series of uniformly distributed ten-digit numbers. The standard definition of randomness would be that every ten-digit number is produced with equal probability (and the drawings are independent, so the

[8] We are giving a somewhat stylized account of digital paper for the purposes of this argument.

equal probability condition holds not only in terms of marginal probability but regardless of conditioning information). Thus if we leave our RNG running for a while, it's bound to produce numbers such as 1111111111 and 0123456789. But these are not "random numbers" on the Kolmogorov definition. The paradox is then that the output of a random number generator (i.e. a device that generates numbers at random) is bound to include nonrandom numbers.

In these examples we have non-random sub-sequences of the output of the RNG. This is not a valid objection, as we have to take the entire output of the RNG up to some large number of digits, in order to obtain these sub sequences that appear non-random. So these short subsequences are not produced by the random number generator, but, strictly speaking, by a Turing machine program that is a prefix to the random number generator, and which searches for patterns like 1111111111111 in the output of the RNG. The Algorithmic Information Theory approach to this would be to add the information content of the program which generated the sequence to the program which selected for the "non-random" sub sequences.

Randomness of a number as opposed to of a generator.

In standard statistical parlance it doesn't really make sense to talk of a random *number* as such, as opposed to a random *variable* or a random number *generator* (where the adjective "random" attaches to the generator, i.e. it's a random generator of numbers rather than a generator of random numbers). Kolmogorov defines "random number", in a way that seems to conflict with the standard view.

But this is just a divergence between what we commonly understand as a number in statistics and how a number is defined in computational complexity theory. By number the Algorithmic Information Theory just means a sequence of digits. Since any sub-sequence of digits is also a number, formalisations in terms of numbers also provide for formalisation in terms of finite sequences of numbers. Thus a sufficiently large number can be treated as a generator of smaller numbers.

1.5 Information and randomness

To get at the second paradox we will report a little experiment. We have an ASCII file of the first eleven chapters of Ricardo's *Principles*: it's 262899 bytes. We ran the bzip2 compressor on it and the resulting file was 61193 bytes, a bit less than quarter of the size. Suppose for the sake of argument that bzip2[9] is a perfect byte-stream compressor: in that case the 61193 bytes represent the incompressible content of the Ricardo chapters. They measure the true information content of the larger file, which contains a good deal of redundancy. That idea seems fair enough.

The second part of the experiment was to generate another file of 262899 bytes of printable ASCII characters (the same length as Ricardo), this time using a random number generator[10], and running bzip2 on the resulting file produced a compression to slightly over 80 percent of the original size.

The first question is why we get any compression at all on the "random" ASCII files?

Our bytes are printable characters. These are drawn from a subset of the possible byte values[11], and as such all, the possible byte values are not equiprobable. Thus the stream is compressible.

The next question concerns the information content of the various files. Suppose we have already accepted the idea that the 61193 bytes of bzipped Ricardo represent the irreducible information content of the original Ricardo file. Then by the same token it seems the 218200 (or so) bytes of bzipped rubbish from the random number generator represent the true information content of the (pseudo-)random byte stream. The rubbish contains almost four times as much information as the Ricardo. This is very hard to swallow.

The point here is that standard data compression programs use certain fixed algorithms to compress files. In this case an algorithm known as Lempel-Ziv[12] is used. Lempel-Ziv does not know how to obtain the maximum compression of the stream—which would be an encoding of the random number generating program. One can not make a general purpose compressor that will obtain the maximum possible compression of a stream. One can only produce programs that do a good job on a large variety of cases.

We make the distinction between information as such and utility, and in those terms it's clear that the Ricardo is of much greater utility than the rubbish. Even so, intuition rebels at the idea that the rubbish carries *any* information. We have a conception of "useless information" alright, but it seems doubtful that a random byte stream satisfies the ordinary definition of useless information. In ordinary language information

[9] A publicly available data compression program.
[10] the `rand()` function in the GNU C library
[11] There are 256= 2^8 possible values for 8 bit bytes.
[12] Ziv 78.

has to be *about* something; and it's useless if it's about something that is of no interest. For me, the weekly guide to Cable TV programming may contain useless information. It's of no more interest to me than a random byte stream. Nonetheless, I recognize that it does contain (quite a lot of) information; it is certainly about something.

In classical political economy use-value is neither the measure nor the determinant of value, but nonetheless it's a *necessary condition* of value. If a product has no use-value for anyone then it has no value either, regardless of how much labour time was required for its production. Can we say that the utility of a message is not the measure of its information content, but if a "message" is of no potential use to anyone (is not about anything) then it carries no information, regardless of its incompressible length?

No. Information exists even if it is not useful. Take the case of hieroglyphs prior to the discovery of the Rosetta stone.[13] They were meaningless until that was discovered, useless in other words. Once it was discovered they became useful historical documents. Their information content was not created *ex-nihilo* by Champollion, but must have been there all along. Similarly, the works of Ricardo in Chinese contain no information to me, are of no use to me, but they still contain information.

In the end, whether information is useful to us concerns our selfish thermodynamic concerns. Does it enable us to change the world in a way that saves us work or produces us energy. This is an anthropospective projection. It is not a property of the information it is a property of the user of the information, which is cast back onto the information itself. Information theory in its epistemological break, had to divest itself of anthropospective views, just as astronomy and biology had to.

The "digital paper" example suggests one further paradox on the issue here. Let's go back to the ASCII Ricardo. Its incompressible length was (according to bzip2) 61193 bytes. Now suppose the hard drive is exposed to radiation that results in random bit-flipping, which changes some of the bytes in the Ricardo file. At some later point we try compressing the file again. We find that it won't compress as well as before. Its information content has increased due to the random mutation of bytes! Meanwhile, of course, its value as representation of what Ricardo said is eroding. Is it possible to make any sense of this?

Yes. The degraded work contains more information since to reconstruct it one would need to know the trajectories of the cosmic rays which degraded the stored copy, plus the original copy. We may not be interested in the paths of these cosmic rays,[14] but it is additional information, provided to us courtesy of the Second Law.

Althusser, Louis and Étienne Balibar *Reading Capital*, London: New Left Books, 1970.

Babbage, C., *The Economy of Machinery and Manufactures*, London, C. Knight, 1832.

Bachelard, Gaston *La philosophie du non: essai d'une philosophie du nouvel esprit scientifique*, Paris: Presses Universitaires de France, 1970.

Bennett, Charles H. 'Notes on the History of Reversible Computation', *IBM Journal of Research and Development*, vol. 32, no. 1, pp. 16–23, 1988.

Bennett, Charles H. 'Logical depth and physical complexity', in *The Universal Turing Machine: a Half Century Survey*, Oxford, 1988.

Brauer F., Nohel J. A. *The Qualitative Theory of Ordinary Differential Equations*, Dover Publications, 1989.

Brillouin, Léon 'Maxwell's Demon cannot operate: Information and Entropy', *Journal of Applied Physics*, vol. 22, pp. 334–7 & 338–43, 1951.

Brillouin, Léon *Science and Information Theory*, New York: Academic Press, 1956.

Chaitin, Gregory *Information, Randomness and Incompleteness*, Singapore: World Scientific, 1987.

Davies, Roy *Service in the Roman Army*, Edinburgh University press, 1989.

DiVincenzo, David, Quantum Computation, *Science*, Vol 270, pp 255..261, Oct 1995.

Duménil G. 'Beyond the transformation riddle: a labor theory of value', *Science and Society*, 47, pp. 427–450, 1983.

Farjoun E., Machover M. *Laws of Chaos, a Probabilistic Approach to Political Economy*, Verson, London, 1989.

Feynman, Richard, There's plenty of Room at the Bottom, lecture to the American Physical Society, 1959, reprinted in in Hey, Antony, ed. *Feynmann and Computing* Westview Press, Oxford, (2002).

[13] The inscription on the Rosetta Stone, is a decree for King Ptolemy V Epiphanes dating from March 196 BC. It is repeated in hieroglyphs, demotic and Greek. By using the Greek section as a 'key' scholars realised that hieroglyphs were not ideograms, but that they represented a language. Jean-Franois Champollion (AD 1790-1832), realised in 1822 that they represented a language which was the ancestor of Coptic.

[14] In other circumstances, archeological dating for example, such radiation damage gives us useful information.

Feynman, Richard, Simulating Physics with Computers, *International Journal Theoretical Physics*, Vol 21, Nos. 6/7, (1982), reprinted in in Hey, Antony, ed. *Feynmann and Computing* Westview Press, Oxford, (2002).

Fisher, Douglas,*The Epic of Steel* , 1963, Harper and Row, New York.

Foley D. K. 'The value of money, the value of labor power, and the marxian transformation problem', *Review of Radical Political Economics*, 14, pp. 37–47, 1982.

Foley D. K. 'On Marx's Theory of Money', *Social Concept*, 1, pp. 5–19, 1983.

Frisch, U., B. Hasslacher and Y. Pomeau 'Lattice Gas Automata for the Navier Stokes Equation', *Physical Review Letters*, vol. 56, pp. 1505–08, 1986.

Gode D., Sunder S. 'Allocative efficiency of markets with zero intelligence traders: markets as a partial substitute for individual rationality', *Journal of Political Economy*, 101, pp. 119–137, 1993.

Harrison, M. J. 'Entropy Concepts in Physics', in Kubat and Zeman (1975) pp. 41–59.

Hutchison, Terence *Before Adam Smith, The Emergence of Political Economy, 1662–1776*, Oxford: Basil Blackwell, 1988.

Kubat, Libor and Jiri Zeman (eds) *Entropy and Information*, Amsterdam: Elsevier, 1975.

Kuhn, Thomas *The Structure of Scientific Revolutions*, Second edition, Chicago: University of Chicago Press, 1970.

Krause U. *Money and Abstract Labour*, Verso, London, 1982.

Landauer, Rolf 'Irreversibility and Heat Generation in the Computing Process', *IBM Journal of Research and Development* vol. 5, No. 3, pp. 183–91, 1961.

Landauer, Rolf 'Information is Physical', *Physics Today*, May 1991.

Landauer, Rolf, Irreversibility and Heat Generation in the Computing Process, *IBM J. Res. Dev.*,**5**,183-191, (1961).

Landauer, Rolf, Information is Inevitably Physical, in Hey, Antony, ed. *Feynmann and Computing* Westview Press, Oxford, (2002).

Lardner, D., 'Babbage's calculating engines', *Edinburgh Review* 59 (1834), 263-327.

Lavington, Simon *Early British Computers*, Manchester: Manchester University Press, 1980.

Marx K. *Capital*, Vol. 1, Progress Publishers, Moscow, 1954.

Marx K. *Capital*, Vol. 3, Progress Publishers, Moscow, 1972.

Marx, Karl *Capital, A Critique of Political Economy*, London: Lawrence and Wishart, 1970.

Maxwell, James Clerk *The Theory of Heat*, Longman: London, 1875.

Norretranders, Tor *The User Illusion*, Harmondsworth: Penguin, 1998.

de Prony,*Notice sur les Grandes Tables Logarithmique &c* printed in a quarto pamphlet 'Recueil des Discourses lus dans la seance publique de L'Academie Royale Des Sciences, 7 Juin 1824.' Translated from the French by C.J.D.lRoberts

Ricardo, David *Principles of Political Economy and Taxation*, vol. 1 of The Works and Correspondence of David Ricardo, ed. P. Sraffa, Cambridge: Cambridge University Press, 1951.

Roemer J. E. *A General Theory of Exploitation and Class*, Harvard University Press, Cambridge, Massachusetts, 1982.

Rubin I. I. *Essays on Marx's Theory of Value*, Black Rose Books, Montréal, 1973. First edition published in Russian, 1928.

Porter, A. W. *Thermodynamics*, London: Methuen, 1946.

Rae, John *Life of Adam Smith* [1895], with an Introduction by Jacob Viner, New York: Augustus Kelley, 1965.

Shannon, Claude 'A Mathematical Theory of Communication', *The Bell System Technical Journal*, vol. 27, pp. 379–423 and 623-56, July/October 1948.

Shaw, P., P. Cockshott and P. Barrie 'Implementation of Lattice Gases Using FPGAs', *Journal of VLSI Signal Processing*, pp. 1251–56, 1996.

Smith, Adam *The Wealth of Nations*, ed. Andrew Skinner, Harmondsworth: Penguin, 1974.

Sobel, Dava,*Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*, Penguin, 1996.

Ste. Croix, G. E. M. de *Class Struggle in the Ancient Greek World*, Duckworth, 1981.

Steedman I. *Marx after Sraffa*, Verso, London, 1981.

Strandh, Sigvard *Machines, an Illustrated History*, Göteborg: AB Nordbok, 1979.

Szilard, Leo 'Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen', *Zeitschrift für Physik* vol. 53, pp. 840–53, 1929. ('On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings'. Translation in *Behavioral Science* vol. 9, 1964, pp. 301–310.)

Tann, Jennifer (ed.) *Selected Papers of Boulton and Watt*, London: Diploma Press, 1981.

Thurston, R. H. *A History of the Growth of the Steam-Engine*, New York: D. Appleton and Company, 1878.

Turing, Alan 'On Computable Numbers, With an Application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society*, Series 2, vol. 42 pp. 230–65, 1937.

¡¡¡¡¡¡¡ biblio.tex Whiteseides, G., M., Kim, E., The Use of Minimal Free Energy and Self-Assembly to Form Shapes, *Chem. Mat.* 1995, 7, 1257-126.

Whitesides, G.M., Self-assembling Materials, *Scientific American*, September 1995, pp. 146-149.

======= Wannier G. H. *Statistical Physics*, Dover Publications, 1987. First edition published by Wiley, New York, 1966.

¿¿¿¿¿¿¿ 1.11 Ziv, J and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inform. Theory*, vol IT-24, no 5, 1978, pp. 530–536.