# Information and mass production

## Babbage to Boltzman

Paul Cockshott[1]

[1]Department of Computer Science

Faraday meeting Nov, 2007

# Outline

1. Motivation
   - The Basic Problem of Information
2. Previous Work
   - Maxwell
   - Boltzmann
   - Shannon
   - Chaitin
3. Our Results/Contribution
   - Information can be applied to industrial processes
   - Copying technologies

# Outline

- What is information
- How does it relate to entropy
- How does it relate to work
- How does it help us understand production

# Outline

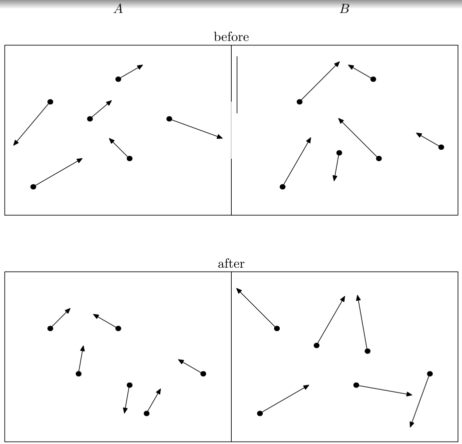


Figure: Gas initially in equilibrium. Demon opens door only for fast molecules to go from A to B, or slow ones from B to A. ⟶ Slow molecules in A, fast in B. ⟶ B hotter than A, and can be used for power.

# Outline

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*.

$$S = -k \int f(v) \log f(v) dv \tag{1}$$

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.

# Outline

# Shanons equation

The communications engineer Shannon introduces the concept of entropy as being relevant to sending messages by teletype.

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{2}$$

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

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.

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

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.

# Mean length

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:

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

which is less than two digits.

# Outline

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 coded messages are numbers. 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.

# Mandelbrot

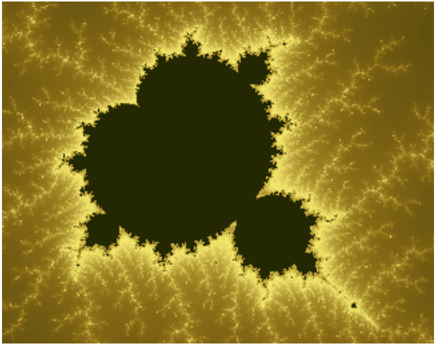

Figure: The Mandelbrot set, a complex image generated from a tiny amount of information. In fact it uses the formula $z = z^2 + c$ where $z$ is a complex number.

# Outline

There are three fundamental ways by which the flow through any production process can be increased.

1. Accelerating the production cycle.
2. Parallelizing production.
3. Eliminating wasted effort.

These are all well studied. What
is less understood is the role that entropy and information play in this.

# Paper making

If we transfer what we have learned to the example to ordinary paper and the process of producing a book we see that the production process encompasses two opposite phases.
First, we have the production of the paper. This is an entropy-*reducing* process. The blank sheets of paper obviously have low information content with respect to human language, but they also constitute a low entropy state with respect to the raw material. In a sheet of paper the cellulose fibres are constrained in both orientation and position. This implies a reduction in the volume of state space that the fibres occupy, and thus, from Boltzmann, a corresponding reduction in entropy.

# Printing

Next, there is the writing of the text—whether by hand, as in the distant past, or using a printing press. This is an entropy-*increasing* process. Imagine that the text to be printed exists as binary data in a file on disk, encoded using ASCII or UNICODE
Clearly the book contains this information, since by sending the book in the post to someone we enable them to recreate the relevant binary file. Thus, by the equivalence of information and entropy, we have increased the entropy of the book relative to the blank sheets of paper.

# 2 phases

In the first phase a low-entropy material is created; in the second phase the entropy of this material is increased in a controlled way. Initially *natural* information is removed; subsequently *anthropic* or human-created information is added. The natural information removed in the first stage is of no interest to us, while that added in the second stage is dictated by our concerns.

# Is this general?

It is easy to see the relevance of information theory to the printing industry. Its product, after all, contains information in the everyday as well as the technical sense of that word. Does this approach provide insights into how other production processes function?

# Weaving and Spinning

For a rather different example, consider the process of producing cloth. The starting material is wool or cotton fibres in a random tangled state. This is first carded to bring the fibres into rough alignment, and then simultaneously twisted and drawn to spin the fibres into yarn. In the yarn both the volume and orientation are sharply reduced. Energy is used to reduce the entropy of the cotton. The weaving of the cotton then increases the entropy by allowing two possible orientations of the fibres at right angles to one another (or more if we take into account the differences in possible weave).

# Car production

Other industries that use thin, initially flat materials clearly have a lot in common with printing. The manufacture of car body parts from sheet steel, or the garment industry, share the pattern of producing a low entropy raw material and adding information to it. In pressed steel construction, added information is encoded in the shape of the dies used to form the car doors, roof panels etc. We can quantify it using Chaitin's algorithmic information theory, as proportional to the length of the numerically controlled machine tool tape that is used to direct the carving of the die.

# Outline

# Replication

If steam powered the industrial revolution, the technologies of replication were the key to mass production. The classic example of the importance of accurate replication was in the production of the Colt revolver in the mid 19th century. Prior to Colt establishing his factory the gun trade was dominated by handicraft manufacturing techniques. The different parts of a gun's mechanism were individually made by a gunsmith so that they fitted accurately together.

For parts to be interchangeable they must be made to very precise tolerances. This improvement in accuracy of production involves the parts having a lower entropy, occupying a smaller volume of phase space, than the old hand made parts. Again by the equivalence of information and entropy this means that the standardized parts embody less information than the hand made ones.

# Accuracy and information

In the 19th century, prior to the introduction of numerically controlled machine tools, replicated parts had to be composed of circular and planar elements which could be produced on lathes or milling machines. The limited information content of these can be seen when you consider that in turning a smooth bore gun barrel one only has to specify the inner and outer radii and its length. If an axle and a bearing are being produced separately to fit together, then one wants the uncertainty in the surface of the bearing, given the surface of the axle, to be reduced below a certain limit.

# Accuracy continued

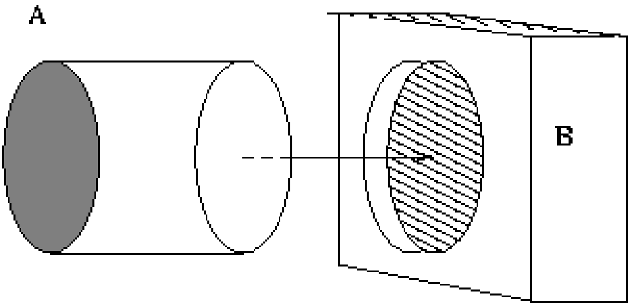


Figure: When inserting axle A into bearing B we want to minimize the conditional information $H(B|A)$, between B and A.

# Binary cross section

```
00000011111111111111111111111111111000000
00000011111111111111111111111111111000000
00000001111111111111111111111111110000000
00000000111111111111111111111111100000000
00000000011111111111111111111111000000000
00000000001111111111111111111110000000000
00000000000011111111111111111000000000000
00000000000000111111111111100000000000000
00000000000000000111111100000000000000000
00000000000000000000000000000000000000000
00000000000000000000000000000000000000000
00000000000000000000000000000000000000000
```

Figure: A slice through the axle.

# Generator

```
program circ ;
const
    b: array [boolean ] of char =( '1' , '0' );
    c =20;
    r =18;
var
    a: array [-c ..c ,-c ..c ] of boolean ;
begin
    a← √(ι0² + ι1²) < r;
    write(b_a);
end .
```

$$a \leftarrow \sqrt{\iota_0^2 + \iota_1^2} < r;$$

We cannot guarantee to have found the shortest such program.[1]

[1] The program is in Vector Pascal which is fairly concise

# Bearing

Clearly if the bearing exactly fitted the axle the expanded encoding for a slice through the bearing would be an array similar but with 1s and 0s swapped round. This can be produced by a trivial change to the program circ, the addition of a single statement.

**write**($b_{\textbf{not}a}$);

replaces the line:

**write**($b_a$);

This must come close to minimizing the conditional entropy of the two parts.

# Imperfect construction

00000011111111111111111111111111111000000
00000011111111111111111111111111111000000
00000001111111111111111111111111110000000
00000000111111111111111111111111100000000
00000000011111111111111111111111000000000
00000000000111111111111111111100000000000
00000000000011111111111111111000000000000
0000000000000001111**0**111111000000000000000
0000000000000000000**0111**000000000000000000
00000000000000000000**1**00000000000000000000
00000000000000000000000000000000000000000

Figure: Part of a pin with a fault on its circumference.

# means added code

we would need to add the following lines to the generator of A' to make the bitmap for B:

$a_{r,1} \leftarrow$ *false*;

$a_{r,2} \leftarrow$ *false*;

$a_{r+1,0} \leftarrow$ *false*;

$a_{r,0} \leftarrow$ *false*;

$a_{r-1,-1} \leftarrow$ *true*;

**write**($b_{\mathbf{not}a}$);

This contains extra information, to correct the bitmap of A' to generate that of B. In pre-industrial production, the extras steps in the generator program would imply additional steps of filing and grinding to make parts fit.

# Pottery

1. Hand formation
2. Turn on wheel – only have to supply diameter as information
3. Casting – parallel information transmission, the Roman samian ware industry based on such casting is the first parallelised mass production of consumer goods

# Law of reproduction

The algorithmic information in repeated production grows by a law of the form $H(P) \leq H(c) + \log n$ where $P$ is the total product made up of $n$ repetitions of $c$. If we look at the process of reproduction as a whole there are two terms the first given by the complexity of the original and a second logarithmic term given by the number of repetitions.

# Law of reproduction - pottery

In the case of the Samian ware pottery there is the original work of producing the master or pattern piece which corresponds to $H(c)$, but then the number of copies that could be made grows exponentially with the number of successive steps of copying: if the master is used directly to produce the pots then $L$ pots can be made, where $L$ is the lifetime of the master. If the master produces moulds which in turn produce the pots then $L^2$ pots can be made, etc. Invert the relationship and we find that the number of successive steps of copying will be related to the number of pots produced $n$ as $\log_L(n)$, a relationship suggested by the predictions of information theory.

- 1832 Babbage, C.: 1832, *The Economy of Machinery and Manufactures*, London.
- 1988 Barnsley, M.: 1988, *Fractals Everywhere*, Academic Press.
- 1999 Chaitin, G.: 1999, Information and randomness: A survey of algorithmic information theory, *The Unknowable*, Springer, Signapore.
- 1875 Maxwell, J. C.: 1875, *The Theory of Heat*, London.