Published paper

Madden, A.D. (2011) *Why is so much information uninformative?* CILIP Update, 10 (7). pp. 32-33.

Why is so much information uninformative?

James Gleick's recent book: "The Information: A History, a Theory, a Flood" provides a fascinating exploration of the implications of Claude Shannon's ground-breaking work on communication theory. It is one of many works that treat information as something quantifiable.

This article is a response to such works. It attempts to provide a clear description of Shannon's concept of information, and to explain why, useful though this is, it fails to describe information in a way that captures the range of interests of information professionals.

Uninformative Information

Some years ago I worked as a librarian in a London school. One lunchtime, I was approached by a Sandra[1], an intelligent, well-educated 16 year-old. She came up to my desk and held out a small white flower with a velvet yellow centre. "*Is this a daisy, Sir?*" she asked.

Not long beforehand, I had been reading Lévi-Strauss's 'The Savage Mind'. In Chapter 1, he comments on the extensive botanical knowledge displayed by various tribes around the world. Where Sandra struggled to recognise one common flower, other girls her age who may never have seen a school, were able to identify dozens, perhaps even hundreds, of plants. As well as naming each plant, they could explain its culinary and medicinal value, and describe how it might be used for making clothes and tools, and for building shelters.

Many people nowadays believe that we receive far more information than our forebears did, and conclude that Sandra belongs to the best informed generation ever to have lived. But was she better informed? If so, in what way? Sixteen year old girls from remote tribes don't have access to the Internet, but what exactly does the Internet deliver? Can something be considered information if it is uninformative? Can information be quantified in any meaningful way?

Some informations are more measurable than others

Numbers are powerful. Where they have been reliably assigned to phenomena, they have made comparisons easier and have helped people to make accurate predictions concerning those phenomena. When something is expressed as a

[1] Not her real name.

number however, there are always factors to consider. What units are being used? Is the number a constant? Under what conditions is it constant? Is the number an approximation? What are the largest and smallest possible values between which it lies? If these factors are not taken into account, the number can become meaningless and misleading.

Unfortunately, once a number has been attached to something, people often stop considering the thing and just concentrate on the number. This can lead to factors being overlooked that are critical if the number is to be appropriately interpreted. When people talk glibly about quantities of information and claim that we receive more information than our forebears, they are guilty of not examining the nature of the "thing".

Comparing bits

Imagine you are in a shoe shop. Beside you is a pair of shoes that would be ideal if only they were bigger. You turn to the assistant and ask: "*Do you have a larger pair?*"

Now imagine that you are President of the USA. You are surrounded by devices that are blipping and flashing. On a screen, a sinister shape looms large. You turn to the nearest aide and ask: "*Is this a nuclear missile I see before me?*"

In both cases you have asked a yes/no question and are eliciting one binary digit (bit) of information; but where the first bit may affect your wardrobe, the second could affect the world. Except in a very limited and technical sense therefore, it is not sensible to compare the two bits of information.

Science however, deals with limited and technical senses. One of the problems that scientists have in communicating with the public is that terms which, to scientists, have a specific, tightly defined meaning, are often derived from words in common use. The meanings are related, but the general definition is broader and includes uses that are excluded from the scientific definition. An example of such a word is "information".

Shannon Information

Claude Shannon, who features so largely in Gleick's book, produced his most ground-breaking work whilst at Bell Labs in New York. Unlike his peers, he focussed on the nature of messages rather than the mechanics of transmitting them. He identified the key components of a communication system as being:

Information source; Transmitter; Channel; Receiver; Destination.

Suppose I was visiting you and left a message on your telephone answerphone to confirm details. In that case, I would be the Information source; you would be the Destination; my telephone handset would be the Transmitter; the Receiver would be your handset; and the Channel would be the telephone system that conveyed the signal.

Unfortunately, when you come to play the message, it is unclear. All you hear is:
“*I will be arriving at 9* [crackle] *m on Friday* [crackle] *of February.*”

Luckily however, as well as leaving you a phone message, I posted a note, which is handed to you by a rueful postman.
“*Sorry boss*,” he mutters. “*It fell in a puddle.*”

Engineers refer to the crackles and hisses that interfere with sound transmissions as “noise” [2]. When Shannon generalised the components of a communication system, he also generalised the factors that can affect a message during transmission. In his scheme therefore, “noise” includes not only crackles on a telephone, but also water damage to a letter.

You take the damp card, lay it somewhere dry, and wonder what use it will be. After all, my first message was unhelpful, and was only missing two crackles-worth of information. Most of this message has been reduced to smudges and smears.

Between the blurs, you read:
“ _____ ___ *arriv*_____ ____ *9 pm* ___ *Fri*_____ _____ *14* ___ *Feb*______ ”

Despite its dousing, the card has proved more informative than the largely intact phone message.

One of Shannon’s key insights was that not all parts of a message are equally important. Often, what has gone before will determine what comes after. In English, a word beginning *arriv* is likely to continue with *a, e,* or *i*.

Sometimes, what has gone before wholly determines what comes after. If you were receiving a message character by character, and the transmission had got as far as: *“I will be arriv…”,* you could easily guess the fourth word, making the next three letters unnecessary.

[2] Another example of a word with a general meaning and a scientific meaning.

Shannon introduced the concept of redundancy[3] to describe such situations. Where characters from a finite set are being transmitted, a probability can be assigned to each character. Prior to Shannon's communication theory, a signalling system that transmitted the letters *arriv* acted as though any letter of the alphabet was equally likely to follow the *v*. Adoption of Shannon's ideas made it possible to assign a high probability value to the letters *a,e,i,* and a low value to all the others.

Where part of a signal is fully determined by what has gone before, (as in "*I will be arriv…*"), that part of the signal is wholly redundant. If however, the earlier parts of the signal give absolutely no indication of what comes next, then each character has an equal chance of following. According to Shannon, such a circumstance represents total entropy.[4]

In the context of Shannon's communication theory therefore, entropy is directly related to information. Where there is a high level of redundancy, we can guess the next part of a signal, so when it arrives, we learn little. Where there is high entropy (and consequently, low redundancy), every part of the signal tells us something new.

Shannon's best known concept was that of the binary digit. It enabled signalling to become a sort of Yes/No guessing game. Given the letters "*arriv*", Shannon's predecessors would have adopted an approach equivalent to asking:

1. "Is the next letter in the range 'a' - 'p'?" (letters 1 – 16) (Yes)
2. "Is the next letter in the range 'a' - 'h'?" (letters 1 – 8) (No)

And so on, until only one letter (*i*) remains. This approach needs five bits (Yes/Nos) for each letter in the signal. If, however, relationships between characters exist and are used to target guesses (i.e., redundancy is taken into account) it becomes possible to improve the guesses and so the same signal can be sent using fewer bits of data.

The specific scientific meaning of information therefore, is derived from signalling. According to this definition where a series of characters is being transmitted, if the next character is wholly unpredictable, it carries more information than if it falls within a predicted range. Consequently, the next letter in the sequence "*I will be arriv…*" carries no information, while the next letter in the sequence "*zlpwcfwkcyjf*" carries 5 bits.

[3] Yet another example of a word with a general meaning and a scientific meaning.
[4] Shannon was not content with co-opting two widely used words (information and redundancy) and giving them specific scientific meanings. He also took a word with one specific scientific meaning and gave it a second one.

Informative and uninformative information

In 1949, Shannon co-authored a book based on his ideas. In it, his co-author, Warren Weaver, stressed that "*In particular, information must not be confused with meaning… In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent… as regards information*."

Regrettably, it is a distinction that is lost on people who assume that we are well-informed because we receive an abundance of messages. Shannon information is related to predictability. According to this definition of information, the more predictable a signal is, the less information it contains.

In the water-splattered note, the first legible part was "*arriv*". This could be followed by one of three letters (*a,e,i*). The first omission from the phone message was either *a* or *p*. Here there is a choice of two letters, so it is easier to predict what goes in this gap than what fills the gap after *arriv*. Less Shannon information is needed therefore, to tell you whether I meant *am* or *pm* than to tell you whether I meant *arrival, arrive*, or *arriving.* However, you will be far better informed about my intentions if you know whether I plan to come in the morning or the afternoon.

Information and boundaries

When people talk about information, they generally assume that someone is being informed. Once, sending a message was expensive and people took trouble to ensure that their messages would inform. A metonymic association arose, which was reinforced by Shannon. "Message" became synonymous with "information".

Today, communication is cheap and the cost of sending messages is negligible. Consequently, the scientific meaning of information is often at odds with its more general meanings, so now it is clearly possible to receive an abundance of information, much of which is uninformative.