Chapter 4

Information, meaning and context

David Chapman

Note that this is a un-edited prepublication draft of Chapter 4 of Ramage, M. & Chapman, D., ed. (2011), *Perspectives on Information*, Routledge, New York and Abingdon.

Abstract

Shannon's 1948 paper on the mathematical theory of information provides the foundation for understanding and engineering communication links, and for that reason alone should be seen as a classic work of the information age. Arguments have raged, however, as to whether and in what way it has significance beyond engineering. This chapter supports a wider relevance of Shannon's work by bringing together the insights of layered architectures as used by communications engineers and the concept of 'meaning' from semiotics to suggest that the 'selective information' of Shannon and 'semantic information' are of the same nature.

Introduction: Foundations of information theory

Authors addressing the question of 'what is information' invariably at some point make reference to the work of Claude Shannon. For some it is merely to distance their own concept of information from that of Shannon, but most see the work of Shannon as the beginning of the exploration of 'information' as a concept that can be addressed scientifically. Borgmann (1999), for example, says: "The birth certificate of information as a prominent word and notion is an article published in 1948 by Claude Shannon" (p.9).

Shannon's paper, "A Mathematical Theory of Communication" (Shannon, 1948), in the words of David Mackay "both created the field of information theory and solved most of its fundamental problems" (Mackay, 2003, p.14). First published in the Bell Systems Technical Journal, it was reproduced, together with a paper by Warren Weaver entitled "Recent contributions to the Mathematical Theory of Communication" (about which I shall have more to say later), in a 1949 book published as "The Mathematical Theory of Communication" (Shannon and Weaver, 1949).

That "The Mathematical Theory of Communication" is still in print today is one piece of evidence for the far-reaching consequences of Shannon's work, but more dramatic evidence can be seen from how the work has been cited. A search in 2010 using the citation database *Web of Science* revealed more than 7700 documents that cited the 1948 paper, and more were appearing at the rate of around 500 a year. Shannon was working in the telecommunications industry, so it is not surprising that the bulk of the citations come from the disciplines of engineering or computer science, but more intriguing are the publications citing Shannon that come from disciplines as diverse as ethics, archaeology, sport sciences and art. In all, *Web of Science* finds citations in publications from more than 200 different subject areas, according to their classification scheme.

In this chapter I summarise some of the key results of Shannon's paper, and use ways of thinking derived from more recent work on telecommunications, in order to explore how Shannon's insights might be applied to "the general problem of communication" (to use the terminology of Weaver, 1949, p.97).

Outline of Shannon's theory

For Shannon, communication was about conveying a sequence of messages from an information source to a destination (Figure 4.1).

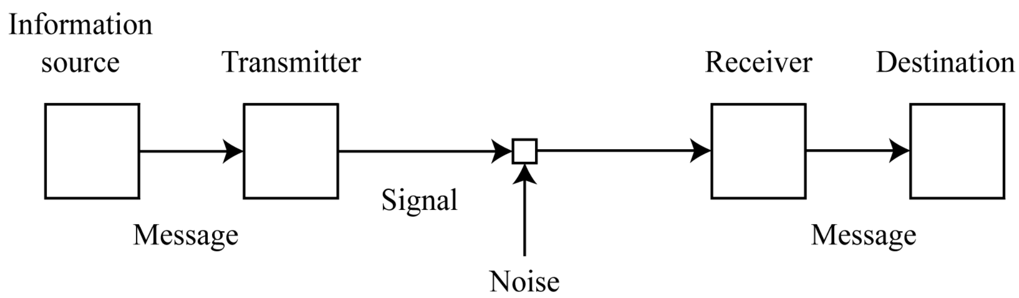

Figure 4.1 Model of communication (after Shannon, 1948)

The information source has a set of possible messages to send to the destination. On any given occasion, it selects *one* of these possible messages, and the message is encoded by a transmitter to make the signal, which is sent over the channel and decoded by a receiver, thus recovering the message. On the way through the channel the signal might become degraded leading to the possibility that the receiver can't decode it correctly. Shannon modelled the degradation by the addition of 'noise' to the signal, and I'll be discussing that in more detail later.

The Shannon-Hartley definition of information

A key feature of Shannon's model is the way in which he separates the information from the coding. To understand Shannon's concept of information, we first need to consider what constitutes a message.

The simplest message is the answer to a question which has just two possible answers: yes or no, up or down, win or lose, boy or girl. As used by Shannon (building on work by Hartley, 1928 and Nyquist, 1924; 1928), information is measured with units of bits ('information bits', which are related to, but not that same thing as, binary digits). These information bits are calculated such that an answer to a binary question contains one bit of information. So, each time a baby is born, one bit of information is needed to convey the gender. This could be represented by a binary digit as used by computers, with, say, a '0' meaning a boy and '1' meaning a girl.

If the choice is between more than two options, the information conveyed can be more than one bit. Suppose, for example, a hospital has a coding scheme for the health of the babies born and that they use four categories: green meaning 'healthy', yellow meaning 'some cause for concern', orange meaning 'needs medical attention' and red meaning 'emergency'. With binary digits this might be represented by a pair of digits, say 00 for green; 01 for yellow; 10 for orange and 11 for red.

It is not necessarily the case, however, that the health information is worth two information bits. In terms of Shannon information, the amount of information in a message is determined by the probability of the message (the probability of the answer to the question). If the four categories were all equally probable, they would indeed all be worth two bits of information, but that is not how it is here – babies needing emergency treatment are much less common than healthy babies. In general terms, Shannon's analysis determines that less-probable messages convey more information than more probably messages. This means that a message saying a baby needs emergency treatment conveys more information than a message saying a baby is healthy, which is in some ways reasonable because the hospital will carry on with its normal business when the healthy baby is born, but will have to spring into action in response to a baby needing emergency treatment. The rare event has a bigger impact. The parents of course might have a different perspective: the fact that their baby is healthy is big news to them, even if it is the same news that most other parents get.

Shannon (based on the work of Hartley and others) quantified the information content of a message using a logarithmic measure, and although the details are not important here, it is useful to illustrate his work with some numbers. The specific formula is:

$$I_k = -\log_2(p_k)$$

Here, a message that occurs with probability p_k delivers information I_k. With two equally-probable messages, like the example of 'boy' or 'girl', each message has probability of ½. Putting this into the formula, we see that $I = -\log_2(0.5) = 1$, giving the result we used earlier that a message about the gender of baby is worth 1 bit of information. With four equally-probable messages each message has a probability of ¼, and $I = -\log_2(0.25) = 2$, again agreeing with what we saw earlier. Suppose, to illustrate the calculation, that the probabilities of the health messages comes out as 80% for green (8 out of ten babies are healthy), 15% yellow, 4% orange and 1% red. These probabilities translate to 0.32 bits of information from a green message, 2.74 from a yellow, 4.64 from an orange and 6.64 from a red message. The point to notice here is that whereas with four equally-probable messages each message delivers 2 bits of information, in this case one message, the most common message, delivers rather less than 2 bits (less than 1 bit in fact) while the other messages deliver more than two bits.

For Shannon's purposes, in analysing a communications channel, the information value of a single instance of a message was not what was important. Shannon envisaged a channel sending a sequence of such messages: some one answer, some another, and what was important for the analysis was the average information per message. Over a long enough time, all possible messages would sent, but they would occur with different frequencies, depending upon their relative probabilities. Messages from a hospital about the gender of the babies born there are relatively straightforward, since the probability of a boy or a girl are roughly equal, each ½, so there will be approximately as many messages saying 'boy' as 'girl' (that is to say, in this case the relative frequencies of the different messages do *not* vary). Furthermore, each message, as we saw before, carries 1 bit of information, so the average information per message is also 1 bit per message.

With the health messages, most will be the green messages delivering 0.32 bits of information but some will be delivering more, and if you do the calculation it comes out with an average of 0.92 bits per message. If the four messages had occurred with equal probability they would all have delivered 2 bits and the average with have been 2 bits per message.

The average information per message is an important parameter in Shannon's analysis, and was called by Shannon the *entropy* of the source.

Shannon's noisy channel coding theorem

Having established a quantifiable measure of information, Shannon was then able to derive a formula for the rate at which this information can be sent over a communication channel. Specifically, he derived his famous formula for the capacity of a noisy communication channel:

$$C = B \log_2(1+S/N)$$

Understanding the maths is not important, but what is important is that a number, C, can be calculated for the channel capacity, measured in information bits per second. This is could be, for example 10 bits per second or 1 kbit/s (kilobit per second which is 1000 bits per second). The formula shows how you can calculate information capacity from the physical parameters of the channel, using knowledge of the bandwidth of the channel (B), the power being used to transmit the signal (S), and the power of the interfering noise (N).

To those without a technical background, these parameters, B, S and N, might seem mysterious, but the important thing to appreciate is that they are all physical properties of the channel: they can all be measured with test equipment. S and N are powers which could be measured in units of Watts, just like power needed to run a vacuum cleaner or a kettle. B, the channel bandwidth, is the range of frequencies that can be sent through the channel. The idea of frequency is less of an everyday concept than power, but is encountered in the context of analogue radios, where you tune to different stations by selecting different frequencies. The range of different frequencies that you can select on the radio – the range over which you can tune it – is the bandwidth of the radio.

Shannon's achievement was to show how knowledge of these physical parameters of a communications channel can be used to determine the maximum theoretical capacity of the channel, and shows the trade-off between different parameters. Calculation of the capacity of a communications channel is a fundamental underlying theory to the work of all communications engineering, and indirectly to the achievements of 'information age'. Whenever we connect to the internet or use our mobile 'phones, we are making use of communication channels which have been engineered to carry as much information as possible. Shannon's formula tells us what is possible, what can't be done, and what needs to be changed to get more capacity.

There has been a lot of debate about the relationship between Shannon's 'information' and other concepts of information – indeed this book is part of the ongoing debate. Shannon's formula for the capacity of a noise channel seems to be making a link between two very different categories of things: material, physical things like signal powers, and something rather less physical, which Shannon called 'information'. Some people have objected even to the use of the word information in this context, or at least insist on a qualifier such as 'selective information'. Whatever word or phrase is used, however, there is something intriguing about what the equation brings together. Remember that these are selections or choices among alternatives. If the channel capacity is calculated to be, say, 100 bits per second, we can convey the equivalent of 100 binary decisions every second. This surely is taking us closer to the realm of thoughts and ideas, things that are legitimately classed as information, than the world of matter and energy.

Later in this chapter I shall explore ways in which we might move from Shannon's information to broader concepts of information through the use of the layering metaphor of engineering, but first I want to digress briefly with a story that emphasises the remarkable significance of Shannon's equation, even just in terms of engineering.

Turbo coding and the limits to channel capacity

Knowing the theoretical channel capacity is only the first step for the communications engineer. Shannon's work allowed engineers to determine the theoretical maximum capacity, but engineering a real system to achieve that maximum is a different matter. The key, as Shannon himself showed, is in the design of the encoding and decoding.

As I said earlier, the effect of noise on the channel is that the messages might not arrive correctly. Consider communicating the message 'boy' or 'girl' from the maternity unit. The simplest coding over a digital communication channel for this would be, for example, to use a binary '0' to mean a boy and a '1' to mean a girl. On one occasion a message saying that the child was a girl, sent using the sign '1', might get corrupted and received as a '0' with the result that the destination thinks that a boy was born. Better coding can reduce the chance of this happening. So, for example, the sign used for a girl could be three 1s (111) and for a boy three 0s (000). Now if one of the 1s in the girl's sign is received as a 0, the decoder gets, say, 101 instead of 111. Although this is wrong it is more like the sign for a girl than the sign for a boy, so it is correctly decoded to mean girl. This is an example of a repetition code – you just repeat '1' three times for a girl and repeat '0' three times for a boy.

With very simple codes, we find that in practice we cannot get very close to the channel capacity. If Shannon's formula suggests that the capacity is, say, 1000 bits per second, we might find that in practice we only send as few as 10 bits per second. If we try to send more, too many of them get corrupted.

There are much more sophisticated codes that can be used, however, and by using better codes we can get closer to the limit given by Shannon's formula. Nevertheless, for almost 50 years it proved impossible to approach the Shannon limit very closely, and there seemed a practical limit that was rather less than Shannon's.

All that changed in 1993 when three researchers from a France presented a conference paper (Berrou et al, 1993) which described a coding technique which they called 'turbo coding', that they claimed almost achieved the Shannon limit. So remarkable was this claim, that the first time a paper was submitted it was rejected as being too good to be true (Burr, 2001)! Since then turbo codes and other related techniques have become widely used, and Shannon-limit performance is no longer pure theory.

Layered architectures

A number of authors have addressed the bridge between the technology of a communications link and human communication by invoking layering metaphors. Warren Weaver did this in his article that appeared alongside Shannon's paper in the 1949 book, talking of three levels corresponding to what he referred to as the technical problem (level A), the semantic problem (level B) and the effectiveness problem (level C). These three levels roughly corresponded to the levels that Cherry discussed in his 1957 book "On human communication". Cherry drew from the field of semiotics, and identified layers that address syntactics, semantics and pragmatics. Although Cherry drew a distinction between his layers and those of Weaver, they were broadly speaking taking a similar approach, with the bottom layer (technical/syntactics) being where the work of Shannon applies, the semantic level being where we start to talk about meaning and the top level (effectiveness/pragmatics) where we consider the impact on people.

In this chapter, however, I come to layering from a slightly different direction, drawing on the way in which layered metaphors have been used in engineering (see, for example, standard texts on data communications such as Tannenbaum, 2002, or Stallings, 2010).

In Figure 4.2 I have redrawn Figure 4.1 to show layering.

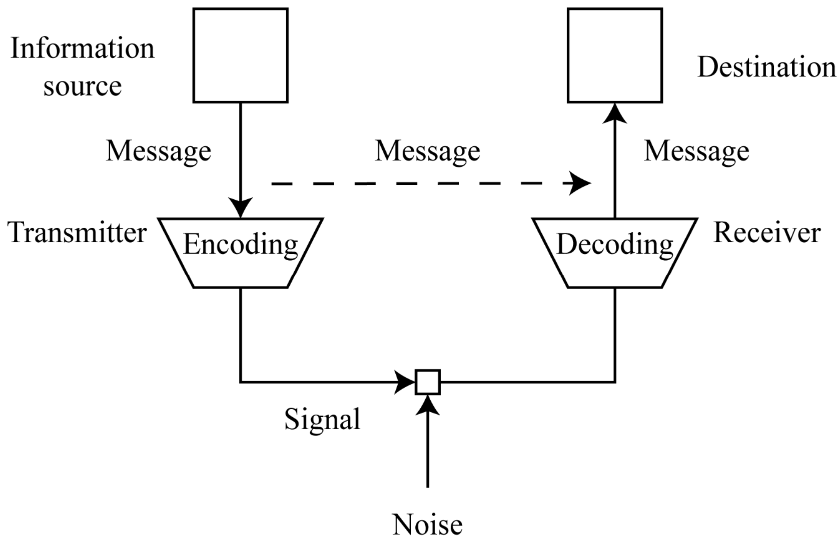

Figure 4.2 Model of communication, redrawn to emphasise layers

Within the metaphor, the encoding and decoding take place ‘below’ the information source and destination. This vertical relationship is indicated by the positioning of the transmitter and receiver on the diagram, and in Figure 4.2 I am also using the convention of drawing the boxes for the encoding and decoding as trapeziums. A feature of this layered metaphor is that we talk about ‘virtual communication’ taking place at higher levels, which is shown on the diagram with a dashed line (a ‘virtual channel’). So there is a virtual communication channel for the message, while the physical communication takes places at the lower level.

For example, suppose the message is a colour, and that we are communicating a selection from the four possible colours of green, yellow orange and red. Suppose further that on one particular occasion the message is yellow. This, ‘yellow’, is what is conveyed over the virtual channel, while the transmitter might code it as 01, which is what is carried over the physical channel.

The idea of layers introduces a new dimension to the communication, and we can extend up and down to bring in more aspects of the communication, which is what has been done in Figure 4.3. ‘Downwards’, the 01 that I previously called the signal would have some physical representation. It might be pulses of light in an optical fibre, or it might be a radio signal. The process of generating a physical signal we call modulation, and detection of the physical signal is demodulation, so we have the modulation/demodulation layer below the binary data.

Extending upwards takes in what the colours ‘mean’. Using the example from earlier – the health of newborn babies – yellow means ‘some cause for concern’. The virtual channel at the top is now the state of health of the baby, this is colour coded in the layer below and then digitally coded in the next layer before the modulation.

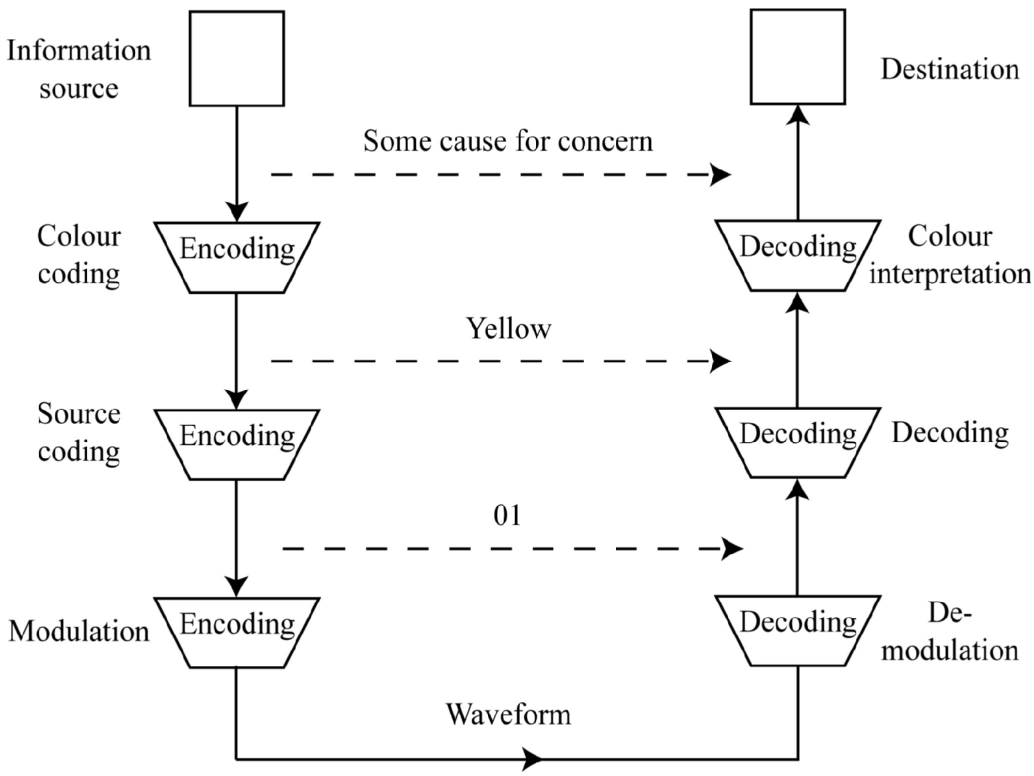

Figure 4.3. Further layers in a communication channel

Within the context of communication systems engineering, standardised layered architectures have been defined, such as the seven layer model for Open Systems Interconnection (the 'OSI model') and the 'TCP/IP protocols' of the internet. (The Transmission Control Protocol, TCP, is at a level above the Internet Protocol, IP. Likewise, the Hypertext Transfer Protocol, HTTP, familiar from web page addresses, is at a level above TCP. The details of these protocols is not relevant for the present discussion). For some purposes, working within a standardised framework is important – it allows equipment constructed by different manufacturers to work together, for example (that was the meaning for the word 'open' in 'Open Systems Interconnection'). However, the way of thinking in terms of layers permeates the design of many engineered systems, and not just communication systems. It is that way of thinking and some of the conventions used in engineering that I want to draw upon while working towards some insights into the nature of information.

The search for meaning

Looking back to Figure 4.3 and specifically at moving up the stack on the right hand side, notice how we get successive levels of meaning. The physical signal is taken to mean the binary bits 01. The binary 01 is taken to mean yellow, and yellow is taken to mean 'some cause for concern'.

Shannon, however, in his 1948 paper, was careful to distance himself from any talk of 'meaning':

> Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. (Shannon, 1948, p.379)

Warren Weaver, in his paper, “Some recent contributions to the mathematical theory of communication” which was published alongside Shannon’s 1948 paper in the 1949 book (Shannon and Weaver, 1949) commented on that statement of Shannon’s:

> But this does not mean that the engineering aspects are necessarily irrelevant to the semantic aspects. (Weaver, 1949, pp.99-100)

The problem seems to be the distance between a purely physical (material/mechanical) mechanism – sound waves in the air; voltage on a wire; light level in an optical fibre; or the amount of ink the page, for that matter – and a meaningful concept in the brain of a human being. Donald Mackay, speaking on BBC radio in 1960, put it like this:

> The original speaker, we may suppose, means something by what he says … yet in the next stages … (the generation of sound waves and all the rest of it) all signs of his meaning seem to have disappeared. Discussion at this level proceeds in exactly the same terms whether the air is handling the outpourings of a genius of the jabber of a monkey. Yet finally, when the message reaches the ear of the human listener, its ‘meaning’ seems to pop up again from nowhere … There are in fact two awkward discontinuities in this way of telling the story: a jump from meaningful utterance to meaningless air vibrations; and then back again. (Mackay, 1969, p.20)

Mackay went on to address the problem in terms of the selective function of messages on “the state of conditional readiness for behaviour” of the brain, and argued that “this opposition of ‘meaningful’ and ‘mechanical’ is false” (Mackay, 1969, p.21).

Expressed in the way that I did at the beginning of this section (the physical signal means the binary bits 01, 01 means yellow, and yellow means ‘some cause for concern’), however, it seemed quite painless to move from the physical signal all the way up to the very human concept of ‘some cause for concern’, with no apparent awkward discontinuities. Maybe if there is a discontinuity, it is right at the bottom where material is used to carry ‘Shannon information’, and the discontinuity is bridged by Shannon’s equation for the capacity of a noisy channel, as I discussed earlier. Above that we have entered the realm of information and left the material world behind.

We need to be careful, though, because the concept of ‘meaning’ is not straightforward. Saying that something ‘means’ something else is open to various different interpretations. The usage here is broadly speaking a semiotic usage, which I shall be looking at a little more later. First, I want to explore some more trapeziums.

More layers

The system for reporting the health of babies was artificial, presented to illustrate some aspects of Shannon's communication model. Figure 4.4 looks at a more realistic example of communication, in which one person (the originator) communicates with another person (the recipient) using text carried over a telecommunications system.

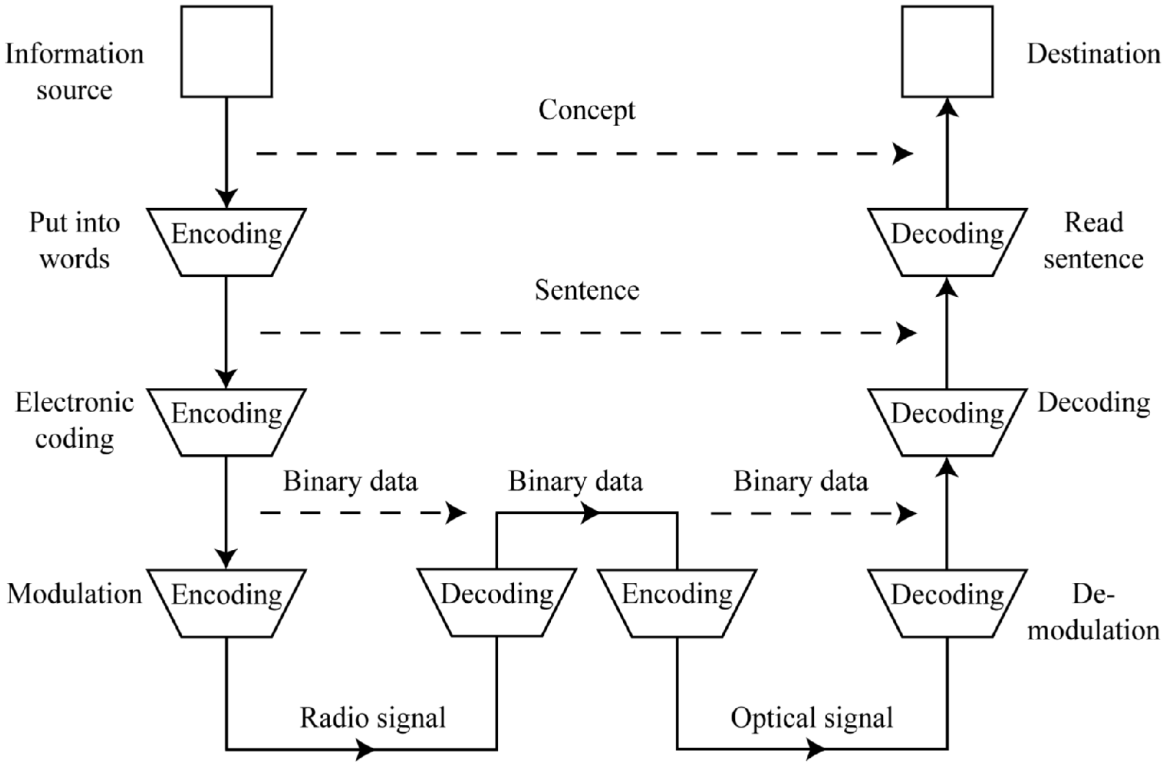

Figure 4.4 More advanced use of layers in communication

The originator has some concept to convey to the recipient, which they do by putting the concept into words which are then digitally encoded and modulated for transmission. At the destination, the physical signal is demodulated to recover the digital data, which is then decoded to get the words of the sentence. The sentence is read and interpreted by the recipient to extract the concept. 'Concepts' are, of course, much more complex things than words or binary digits, but that does not invalidate the layered model. The complexity of the concept is dealt with in the appropriate layer.

Figure 4.4 assumes that on the way from the source to the destination two different transmission media are used. First the signal is carried over a wireless (radio) link, then further on it uses an optical fibre. Because different modulation systems are needed for the different media, the signal is demodulated back to the binary data and modulated again for the new medium, which can be seen in the middle of the diagram. This intermediate stage which shows decoding and recoding without going all the way up the ‘stack’ again illustrates a way of thinking drawn from engineering, which can provide insights for the general problem of communication.

Semiotics

It is not appropriate, nor is there space, to explore semiotics in depth here, but there are a few basic ideas that we can exploit to extract further insights into the communication process as modelled by the layered diagrams.

Signs are modelled in different ways in different semiotic traditions, but the most basic is the dyadic, two-part, model of Saussure (1974). In this model (discussed by Chandler, 2002), a sign consists of the signified (the concept) and the signifier (the representation). This distinction is illustrated in Figure 4.5(a).

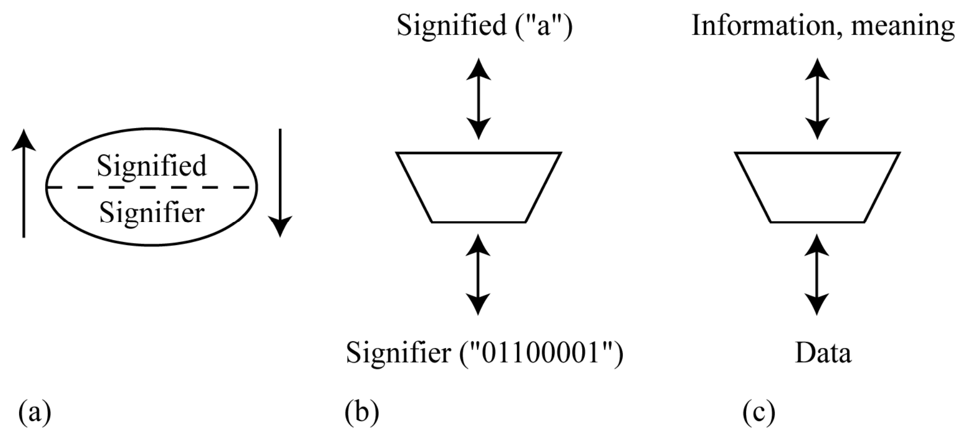

Figure 4.5. (a) Saussure’s model of a sign (b) As a trapezium (c) The relationship between data and meaning

Note that the sign is the combination of these two: the signifier alone does not constitute the sign. For Saussure both the signifier and the signified were psychological, although later usage has sometimes interpreted the signifier as the physical form of the sign. Roland Barthes (1957) gives the example of a bunch of roses (the signifier) to signify a suitor’s passion (the signified). The roses signifying passion constitutes the sign.

I want to suggest a parallel between a semiotic sign and the trapeziums of Figures 4.2, 4.3 and 4.4. At the bottom of the trapezium is the signifier, at the top the signified. For example, in ASCII code the bits ‘01100001’ (the signifier) signifies the letter ‘a’, as shown in Figure 4.5(b). In Figure 4.5(c), I have introduced another interpretation, building in the idea that 0110001 ‘means’ the letter ‘a’, so I am suggesting that the trapezium links the ‘data’ to the ‘meaning’. Or, maybe the ‘meaning’ of the data is the ‘information’ – which fits with the definition quoted by Bissell in Chapter 3 that information is comprised of data in context.

Alternative semiotic models of a sign are triadic and illustrated by triangles similar to that of Figure 4.6(a). There are many variations, but broadly speaking they extend the dyadic model by introducing explicit consideration of the process of interpretation of the sign. Thus, the bits '01100001' only signify the letter 'a' in a context where they are being used as an ASCII code, in the same way that a bunch of roses only signify passion when used for that purpose.

In Figure 4.6(a) I have chosen a variation which suits my purposes, although not matching any of the standard semiotic triangles such as Peirce's 'representamen, interpretant and object' (Hartshorne and Weiss, 1965) or Ogden and Richard's 'symbol, thought or reference and referent' (Ogden and Richards, 1949).

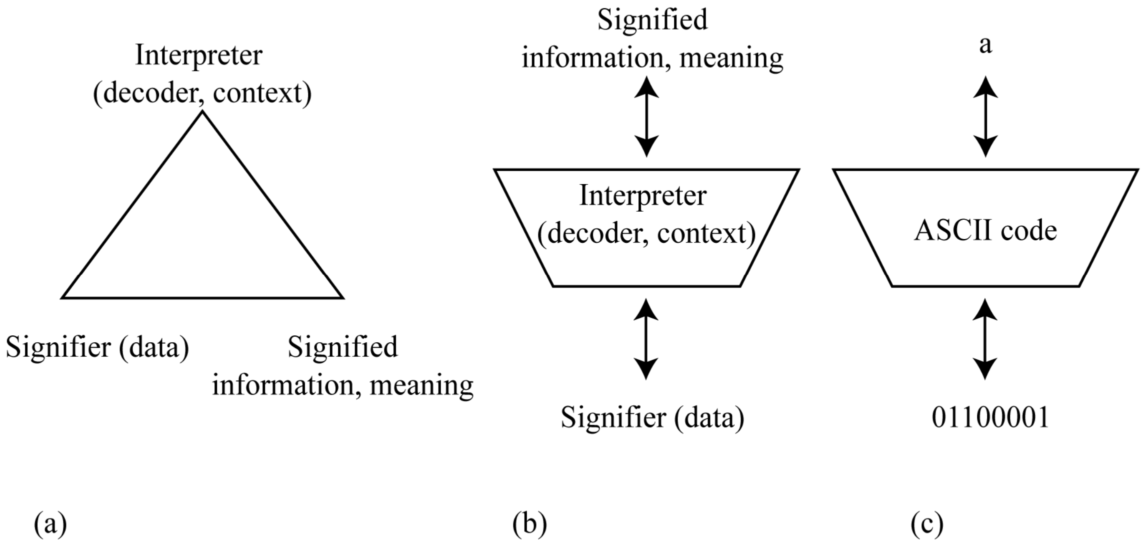

Figure 4.6. (a) A semiotic triangle, (b) Encoding/decoding as triadic sign (c) ASCII code as a sign

With trapeziums the interpreter is readily identified with activity of the trapezium itself, the encoder/decoder or the context, as shown in Figures 4.6(b) and 4.6(c).

Whether or not we can legitimately suggest that the encoding and/or decoding of a message 'is' a sign, the parallel helps the reading of the trapezium diagrams and provides some useful insights. In particular, an important insight of semiotics is that signs are meaningless in isolation: they only have meaning in relation to other signs. This is true of the trapeziums too. Binary data – a string of 1s and 0s – are completely meaningless without the context. In Figure 4.4, the trapezium can only decode the binary data into a sentence using a programme that reads binary data as text.

Each layer in the diagram corresponds to a sign-system. There is meaning *within* each layer by virtue of the relations between the signs within that layer, but trying to find the meaning from one layer within another layer is fruitless. It is akin to trying to find the concept of passion within the biology of roses – a kind of reductionism. The meaning of the letter ‘a’ cannot be found in the digits 01100001 and the meaning of the word ‘war’ cannot be found in the letters ‘w’, ‘a’ and ‘r’. Notice that this definition of ‘meaning’ (information) is relative rather than absolute. Meaning emerges at the layer boundaries, but the meaning is only relative to the layer below.

It is important to appreciate, however, what is and what is not implied by the ‘independence’ of layers. Meaning does not transfer between layers as we have seen, but the capabilities of a layer has an impact on the layer above in terms of the service that it can deliver. Also, it was Weaver’s contention that “the mathematical theory of communication… although ostensibly applicable only to level A [technical] problems, actually is helpful and suggestive for level B [semantic] and level C [effectiveness] problems” (Weaver, 1949, p.114).

An example of that is the discovery of Hartley and Shannon of the relationship between probability and information content: low probability of a message corresponds to high information content. This was important for the technicalities of data communication systems, but it is an idea that works at other levels too. The answer ‘yes’ to the question ‘have I won the lottery’ provides me with more information than ‘no’, but ‘winning the lottery’ is a high-level concept and a long way from the technical details of a communications link.

Conclusions

This chapter has brought together ideas from semiotics and the layered architectures of engineering in order to provide a framework that bridges the gap between communications technology and human communication. It has suggested that the concept of information as introduced by Hartley and Shannon is essentially of the same substance as the semantic information of semiologists, but that Shannon was addressing information at the lowest level in a layered hierarchy whereas semiotics is concerned with the higher levels where information interacts with the human mind.

The framework is essentially a re-working of the ideas proposed by Weaver and by Cherry, and provides a means of dividing up ‘the general problem of communication’ into coherent fields that can be analysed in isolation from each other. In this chapter we have seen how Shannon analysed one level in detail in his classic paper of 1948, and work on that level continues in the developments of computer and communications technologies to this day. We saw earlier in this chapter, for example, the extraordinary story of the development of turbo codes in the 1990s.

Other chapters in this book explore information at higher levels. Chapter 9 by Piwek, for example, explores information in the context of human conversations. Piwek is not interested in the digital coding or the physical representation of the words and sentences because his chapter is exploring information at a higher level than that addressed by Shannon. Chapter 11 by Corrigan with its focus on information laws and intellectual property is arguably at an even higher level.

At the end of his chapter in *The Mathematical Theory of Communication*, Weaver suggested of Shannon's work:

> [T]his analysis has so penetratingly cleared the air that one is now, perhaps for the first time, ready for a real theory of meaning. (Weaver, 1949, p.116)

Through the use of a definition of meaning based on a rather mechanistic interpretation of the meaning of semiotic signs, this chapter has described a way of talking about meaning that can be applied at all levels of the general problem of communication.

References

Barthes, R. (1957) *Mythologies*, translated by Annette Lavers 2009, London, Vintage Classics

Berrou, C., Glavieux, A. and Thitimajshima, P. (1993) 'Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes', in *Proceedings of the IEEE International Conference on Communications (ICC 93)*, Geneva, 23-26 May 1993, pp. 1064-1070.

Borgmann, A. (1999) *Holding on to Reality*, Chicago, University of Chicago Press.

Burr, A.G. (2001) 'Turbo-codes: the ultimate error control codes?', *IEE Electronics and Communication Engineering Journal,* vol. 13, no. 4, pp. 155-165.

Chandler, D. (2002) *Semiotics: The basics*, London, Routledge.

Cherry, C. (1957) *On Human Communication*, Cambridge, MA, MIT Press.

De Saussure, F. (1974) *Course in general linguistics*, London, Peter Owen.

Hartley, R. (1928) 'Transmission of Information', *Bell System Technical Journal*, vol. 7, no. 3, pp. 535–563.

Hartshorne, C. and Weiss, P. (eds) (1965) Collected Papers of Charles Sanders Peirce, Volume I: Principles of Philosophy and Volume II: Elements of Logic, Cambridge, MA, The Belknap Press of Harvard University Press.

Mackay, D. (1969) *Information, Mechanism and Meaning*, Cambridge, MA, MIT Press.

Mackay, D.J.C. (2003) *Information Theory, Inference, and Learning Algorithms*, Cambridge, UK, Cambridge University Press.

Nyquist, H. (1924) 'Certain Factors Affecting Telegraph Speed', *Transactions of the American Institute of Electrical Engineers*, vol. 43, pp. 412 -422.

Nyquist, H. (1928) ‘Certain Topics in Telegraph Transmission Theory’, *Transactions of the American Institute of Electrical Engineers*, vol. 47, no. 2, pp. 617–644.

Ogden, C.K. and Richards, I.A. (1949) *The Meaning of Meaning*, 10th edition, London, Routledge and Kegan Paul.

Shannon, C.E. (1948) ‘A Mathematical Theory of Communication’, *Bell System Technical Journal,* vol. 27, no. 3, pp. 379–423 and no. 4, pp. 623–656.

Shannon, C.E. and Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana, University of Illinois Press.

Stallings, W. (2010) *Data and Computer Communications*, 9th Edition, Upper Saddle River, NJ, Pearson Prentice Hall.

Tannenbaum, A.S. (2002) *Computer Networks*, 4th Edition, Upper Saddle River, NJ, Pearson Education.

Weaver, W. (1949) ‘Recent Contributions to the Mathematical Theory of Communication’, in Shannon, C.E. and Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana, University of Illinois Press, pp. 93-117.