The human brain as a slimming club

Much of the workings of the human brain may be seen as a process of slimming down information by reducing unnecessary 'fat' in what we see and hear and think about. There is potential here for new thinking in the development of artificial intelligence.

Gerry Wolff

CognitionResearch.org

When I was a student, studying psychology and other 'natural' sciences at Cambridge University, I was fascinated by some lectures given by Dr Horace Barlow—a great grandson of Charles Darwin who was later Professor Horace Barlow FRS and has now sadly died.

What I found so interesting in those lectures was the idea that much of the workings of the human brain may be understood as compression of information, meaning that it has the effect of reducing the quantity of information coming in from one's eyes, ears, and other senses, by reducing unnecessary repetition or *redundancy* in that information.

Close one's eyes and open them again

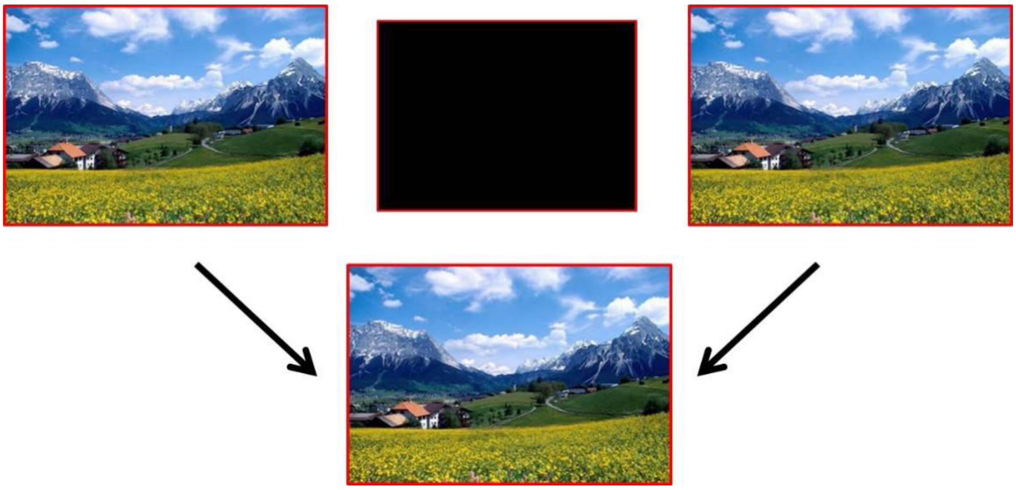
Here is an example of that kind of compression of information by our brains which seems so natural and obvious that its 'slimming' function may be overlooked. If we are looking at a scene, then close our eyes for a moment and open them again, we normally see the 'same' scene as we were looking at before. This means that, for a given person, the 'before' and 'after' views have been merged in his or her brain, thus avoiding the need to store the same information twice.

If we simply retained the 'after' view and discarded the 'before' view, we would not have the strong sense in many cases that they are the same. And, when they are different, we would not be able to detect those differences.

Seeing both views as the same is shown schematically in the figure where the 'before' and 'after' views are shown on the left and right, the black rectangle between those two views shows when our eyes are shut, and the image at the bottom shows the result of merging the 'before' and 'after' views.[1]

[1] The landscape reproduced with permission from Wallpapers Buzz (www.wallpapersbuzz.com).

Contact: Dr Gerry Wolff PhD CEng MIEEE, CognitionResearch.org, Menai Bridge, UK. *Phone:* +44 (0) 1248 712962, +44 (0) 7746 290775, *Skype:* gerry.wolff, *Email:* jgw@cognitionresearch.org, *Web:* www.cognitionresearch.org.



This merging of views is quite different from what happens in an old-style cine camera, where every view of a scene that lasts more than a few seconds is stored, very inefficiently, in two or more separate frames that are the same as each other or nearly so.

How we use language

Another thing that seems so simple and obvious that we do not normally think of it as a way of economising in our use of information can be seen in the way we use language. For example, the city of New York is a big complicated thing that would need a large book to describe everything about it: where it is, all its streets, buildings, and open spaces including Central Park, the people that live and work there and all their activities, the history of New York, and so on. And yet we don't need to create a copy of that big book on each occasion we want to talk about New York. Instead, we simply use the short name "New York"!

This idea is well known in the world of information processing and is called *chunking-with-codes*. In our example, the big book which describes New York with all its geography, history, and so on, is called a *chunk* of information, and "New York"—the nice short name for the chunk—is called a *code*.

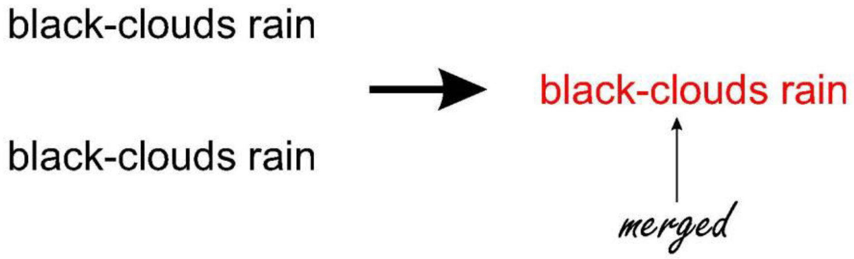
This idea is used everywhere in every natural language like English, Swahili, Mandarin, and so on. In general, every noun, verb, adjective, or adverb, may be seen as a short code for a relatively large and complicated chunk of information. This means that we can write or talk about things, actions, etc., in a way which is very much more economical than if we had to describe each thing fully every time we wanted to refer to it, somewhat like the language of the Ents in Tolkien's *The Lord of the Rings*.

Probabilities via compression of information

There are many other examples of the way that our brains compress information but we'll take a short cut here to ask "Why is this so important in the way brains work?"

Part of the answer must be that, for any kind of animal including people, compression of information can help in reducing the space needed for the storage of information, and it can help in speeding up reaction times. But another answer which is at least as important is that there is an intimate connection between compression of information and concepts of probability, and a knowledge of probabilities can be very useful in finding food, finding a mate, avoiding danger, and so on.

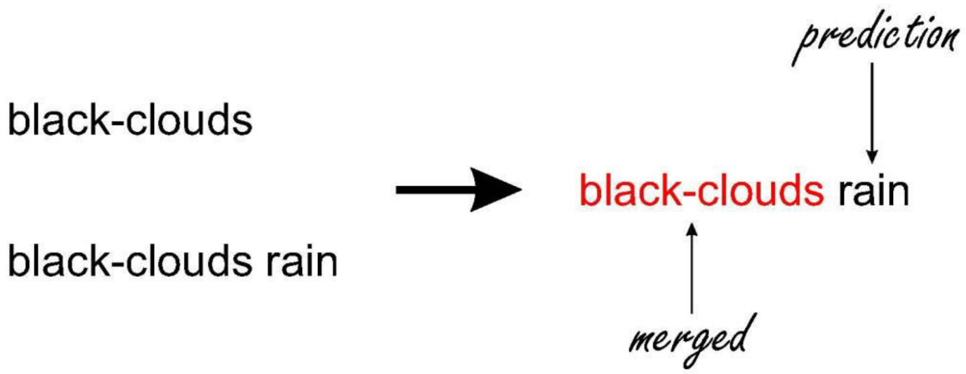
Here is a simple example. A familiar observation is that black clouds are often associated with rain. That means that we can compress that information by merging the repeated instances of the pattern ‘`black-clouds rain`’ to a single instance, as shown in the figure.



It is true that we could have a separate memory for every occasion that we experienced rain falling from black clouds. But that seems unlikely since if we used that kind of memorising for everything, there would be an enormous clutter in our brains, with a new memory for each time we saw a house, a tree, a car, a person, and so on. In general, we economise by only storing one copy of any repeated pattern, and it seems likely that there will be some kind of measure of how often the pattern occurs—and that measure can be useful in calculating probabilities.[2]

The merging of repeating patterns has the effect of compressing information but it brings with it another bonus: the possibility of making predictions and gaining advantage from those predictions. As with learning the association between black clouds and rain, the making of predictions works by merging patterns.

If we have learned that black clouds are associated with rain, and if we then see black clouds, a pattern like ‘`black-clouds`’ (corresponding to what we have seen in the sky) may be merged with the larger pattern ‘`black-clouds rain`’ (which we have learned from previous experience), so there would be an overall reduction in volumes of information, as shown in the figure.



But, in addition, this merging of patterns leads to the prediction that rain is coming, as can be seen in the figure. That prediction, which is quite probable but not certain, may lead us to postpone going out, or hurry home, or put up an umbrella—or if there has been a drought, we might put out some means of collecting the rain which we are expecting.

[2] Just to confuse matters, while we may have only one copy of each repeated pattern in our experience, it seems likely that our brains store two or three copies of everything we know as a backup or safeguard against the loss of our knowledge—for exactly the same reason that IT companies keep multiple copies of all their stored information.

Information compression and intelligence

The examples that have been described are just a few of many ways in which the slimming down of information by merging patterns that match each other may be seen to be integral to how the human brain works, not simply in how we avoid clutter in our memories, use language, or learn regular patterns and make predictions, but also in how we recognise things, retrieve information from our memories, do several kinds of reasoning, make plans, solve problems, and more. There is now much evidence that compression of information is the basis for many aspects of intelligence.

In this connection, it is interesting that, at an early stage, Horace Barlow pointed out that

> "... the operations required to find a less redundant code have a rather fascinating similarity to the task of answering an intelligence test, finding an appropriate scientific concept, or other exercises in the use of inductive reasoning. Thus, redundancy reduction may lead one towards understanding something about the organization of memory and intelligence, as well as pattern recognition and discrimination."

New directions for the development of AI

Those ideas from Horace Barlow's lectures stayed with me for several years and then showed their value in research that I was doing building computer models of how children learn their first language or languages.

Later, they have proved even more useful in a lengthy programme of research developing the *SP Theory of Intelligence* and its realisation in the *SP Computer Model*, with information compression at centre stage.

And they relate to arguments that much of mathematics, perhaps all of it, may be understood as a set of techniques for the compression of information, and their application.

But these are topics for another day.