vivid imagination

Alan Dix

Lancaster University
alan@hcibook.com
http://www.hcibook.com/alan/

whence imagination

So how is that we come to have this strange and wonderful ability to project our images out to our senses? Any answer must be speculative as we cannot turn the clock back and observe our primitive ancestors (although possibly palaeontological studies like Mithen's may be able to cast light and also studies of other animals).

I believe the asnwer lies in semantic feedback loops for sensory disambiguation.

Early models of perception and attempts at computer vision, speach and text analysis used unidirectiobla 'pipeline' architectures. For example, visual stimuli go through stages of edge detection, object–ground separation, object identification etc. Similarly in language understanding there are lexical, syntactic and semantic levels of processing. The end point is some deep semantic representation that can be used for reasoning and planning.

As computational models these largely failed and, not surprisingly, our own perceptions are less pipelined than first thought. Our own perceptual systems and successful automatic system employ semantic feedback.

If you are talking to someone in a noisy environment, or over a phone with a poor connection, it is possible to make sense of one another even though many of the words are not completely audible. This is partly because of the redundancy of language so that many of the connectives, pronouns etc. can be inferred. However, this would not help us to hear the more significant nouns and verbs. In fact we can guess such words easily (and usually unconsiously) because we know what we are expecting to hear.

We are most aware of this when it leads to errors, we think we have heard something we expected to ghear, but what was said was different. Indeed many jokes use this effect to fool us and surprise us.

Without this feedback out sensory systems would be caught in a series of meaningless moments. As I write I am in a room with book coated walls and wedged between the books christmas cards, a tree, fairy lights ... so many patches of different colours that need to be interpreted as bookcases, greetings cards and books? If I look at the black corner of the computer screen

behind it leaning against the wall is a black guitar case. But I can 'see' the corner of the computer, even though the sensory data to disambiguate the sceen edge from the background is tenuous. This is because I 'know' it is there. Deeper levels of my mind 'know' there is a monitor so signal that the object being seen is a monitor and hence square edged, this feeds back to edge detection which now expects edges at particular places and so interprets slight variatiosn in shade, that otherwise may be interpreted as texture or a fold in the soft guitar case, instead as a sharp edge of the screen.

Of course, the process is not quite as algorithmic as this sounds, but the basics of feedback from semantic knowledge to lower levels of perception is there.

The need for this is perhaps even more clear in the 'wild' rather than the rectilinear world of fabricated items. Some while ago I visited Kruger wildlife park. We drove along the dirt roads and then everyh so often we would see a group of cars parked, their occupants strainming to see, eyes clamped to binoculars or long-lensed cameras. Sometimes the object of the observation was clear – a herd of elephants or a distant rhino. But sometimes, especially with lions, it was far more difficult. We would stop and at first look out and see nothing but straw-brown grass. Suddenly one of the group would say "there to the left of the tree". I'd look and still see nothing and then, like one of those strange patterned posters, the grass would refocus and there clearly dozing in the arid African sun would be a lioness lying in the long grass. But once I noticed the lion it was clearly theree, it doid not dissolove back into the jumble of grass again unless my eyes wandered, but even then, knowing where it was, it was easy to 'see' it again. The same visual stimuli interpreted difdferently by my eye once I knew what was there.

This abaility to 'see' something once I had worked out what it was is useful for the game tourist, but essential for the impala. Once the patch of grass has moved and the lion spotted it is a matter of life and death that the image of the lion is preserved even if well camoflaged.

In general our sensory data is partial and fragmentary and is interpreted over time using different clues, perhaps movements, perhaps knowledge, but then that interpretatiion is used to 'lock' the image, sound or other sensory object. Without it our sensory interpretations of thw world would be continually shifting, the world would be reduced to turmoil.

The locking occurs because the semantic feedback is a positive feedback loop, which naturally exhibit hysteresis. Imagine a confused image. The raw sensory data could be interpreted in several ways, but at some point one interpretation becomes more likely than the other. This interpretation is then fed back into the recognition process as the most likely candidate, so subsequent ambiguous data is interpreted in the same way.

This disambiguation and locking in on one interpretation means that higher levels of cognition have a stable 'image' to work with. Unfortunately it could also mean that a wrong

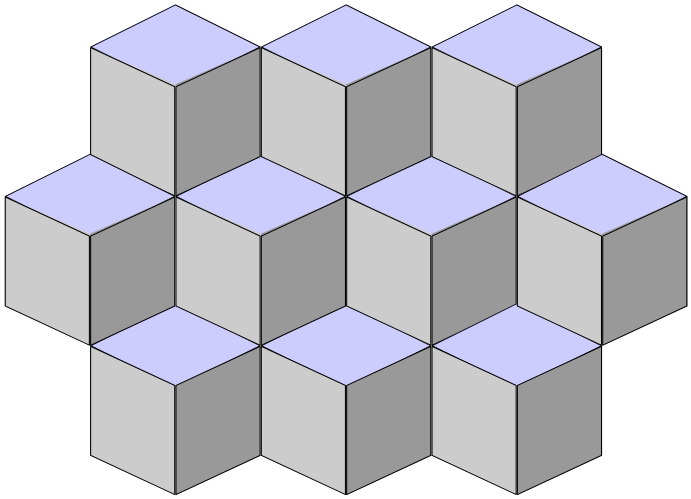

Figure 4. blocks optical illusion

interpretation could be prefered even when new sensory data would start to point to other interpretations. Happily we also have a secondary inhibitaory process that effectively 'resets' the interpretation periodically allowing other interprettations to 'get in'.

We can see this at work with optical illusions. If you stare at one of the central corners in figure 4 you will first see the corner 'pop out' and then 'sink in' to the pattern. There are two possible interpretations of the pattern (one with the light to th top left, the other lit from the bottom right) and your 'eye' (read your visual perception system) is flicking between the two interpretattions. For a while it 'locks in' to one interpretation (the effect of the positive feedback loop), but then after a few seconds the inhibitory mechanism kicks in and allows your vision to find an alternative interprettation. You can see the same effects with other ambiguous figures such as the vase/two faces (figure 5).

These sophisticated feedback machanisms are useful for all but the simplest animals, so it is reasobable to assume that it is qquite an ancient system. (I do not know the literature on non-human visual perception, so would be particularly grateful for

Figure 5. vase or faces

any feedback here.) This means that by the time the first animals began to develop imagination there were pre-existing pathways from semantic encodings through to vivid perception.

Notes

1. xxxx

Bibliography

[[E02]] xxx