21

A Look at Biological and Machine Perception

R. L. Gregory
Brain and Perception Laboratory
University of Bristol

The study of perception is divided among many established sciences: physiology, experimental psychology and machine intelligence; with several others making contributions. But each of the contributing sciences tends to have its own concepts, and ways of considering problems. Each – to use T. S. Kuhn's term (1962) – has its own 'paradigm', within which its science is respectable. This can make co-operation difficult, as misunderstandings (and even distrust) can be generated by paradigm differences. This paper is a plea to consider perceptual phenomena from many points of view, and to consider whether a general paradigm for perception might be found. We may say at once that the status of perceptual phenomena is likely to be odd, as science is in general concerned with the object world; but perceptions are not objects, though they are in some sense related to objects. It is this relation between perceptions and objects which is the classical philosophical problem, and it cannot be ignored when we consider perception as a territory for scientific investigation. This territory is essentially odd: its phenomena tend to be illusions – departures from the world – rather than facts of the world. It requires a conceptual somersault to accept illusions as the major facts of a science! But this we must do; and once this decision is taken, we can hardly expect physics (or physiology) to provide the paradigm for perception.

Machine intelligence and cognitive psychology (though certainly not all perceptual theories) may agree in regarding perceptions as inferences – inferences based on strictly inadequate data. Our reason for seeing perceptions as inferences is that perception is predictive. Perceptual prediction is of two kinds. First: to properties of objects which cannot at that time be sensed directly, or 'monitored'. This applies to hidden parts of objects, to the three-dimensional form of an object given as a plane optical projection, and to non-optical properties such as hardness and mass, which in biological perception are vitally important. The second kind of prediction exhibited by biological perception is prediction to the immediate future. This allows neural conduction and processing time to be cut to zero, so that in skilled performances there is typically zero 'reaction-time'. (This implies that the classical stimulus-response notion, though applicable to reflexes, is not appropriate for perception.) Prediction to the future can also allow behaviour to continue appropriately through gaps in the available sensory data. Prediction to the future is vitally important because dangers, rewards, problems and solutions, all lie in the future.

Although it is convenient for experimental purposes to think of perception in stimulus-response terms, the immense contribution of stored data, required for prediction, makes us see perception as largely cognitive. Current sensory data cannot be sufficient for perception or control of behaviour: it must select relevant facts and generalisations from the past, rather than control behaviour directly from present stimuli.

The importance of prediction – which requires stored knowledge – makes us see perception in cognitive as well as in physiological terms. Although there must be physiological mechanisms to carry out the cognitive logical processes, of generalising and selecting stored data, the concepts we need for understanding what the physiology is carrying out are not part of physiology. We do not derive a cognitive paradigm from physiology, though every move is made by physiological components. (We find this situation in games, such as chess. The moves are physical moves, of pieces on a board; but we do not understand the game from the moves, without knowing the rules and where success and failure lie.) We may suggest that it is just this essential cognitive component of biological perception – which unfortunately is very difficult to investigate – which makes machine intelligence potentially important for understanding biological perception. But, again unfortunately, programs adequate for comparable machine perceptual performance do not as yet exist. Judging from biological perception, perceiving machines will not be of intellectual interest, or effective, until they are capable of using sensed data for making these two kinds of predictions, based on stored knowledge of objects. To increase the sensory capacity of machines, to try to avoid this cognitive component, is futile as a solution because there will always be hidden aspects and properties of objects, and the future cannot in principle be monitored. So to produce machines with accurate ranging devices, or other features for reducing the ambiguity of images or other sensory data, is merely to postpone the problem. However ingenious the engineering may be they are conceptually dull and may hide the essential problem by their (limited) success.

When we start to compare the visual perception of animals with present machine perception – by regarding both as giving inferences about external objects from ambiguous sensed data – we strike a paradox. The perceptual performance of quite primitive organisms is greater than the most sophisticated machines designed for scene-analysis or self-guided response to objects. On the other hand even the most advanced brains are weaker at performing logical operations than are simple devices, of cogs or switches. Why is it that machines, such as computers, can perform logical tasks so well where brains fail; and yet cannot begin to compete with biological object-recognition? This may seem strange enough; but if Helmholtz (1963) was right (and there is every reason to believe him right) to regard perceptions as conclusions of 'unconscious inference', then we must face a truly odd situation – for if perception involves sophisticated inference why are brains weaker at performing logically than are machines? We might suppose that ability to infer was developed in brains, at the start of biological perception; but that this ability was locked away, and is not available for other kinds of problem solving. Or we might suppose that the processes of perceptual inference are different from what we call logic.

It is clear that the physical characteristics of components available for seeing machines are very different from those present in brains. The first are relatively free from drift and noise; but they are relatively large, and cannot carry so many inputs or outputs. This makes parallel processing convenient for biological computing, and serial computing more convenient for man-made computers. Can we suppose that these physical differences lead to the supposed different kinds of inference? If so, biological perception seems to demonstrate powers of *parallel* processing, while computers demonstrate very different powers of *serial* processing. In addition, we might argue that the biologically unique power of human logical problem solving is due (in whole or part) to language, and special symbolic aids, including: mathematical and logical notations, 'digit' fingers for counting, and the abacus – all helping us to infer serially.

Can this suggestion be supported? We might start by noting that in general the more an aid differs from what it aids, the greater the improvement it can confer. (To take an example: knives, forks and spoons are so useful as aids to eating, essentially because our hands are not like knives, forks or spoons. So, if we know that Martians eat with such utensils, we could make a shrewd guess that their 'hands' are not like our knives, or forks, or spoons.) If we accept this as a principle, that: *the greater the difference between an aid and the aided, the greater the possible improvement*, then we have some support for the notion that biological non-perceptual problem solving is by *parallel* rather than by serial processing – because it is *serial* aids which are so effective. Although this cannot be claimed as a strong argument, it seems worth some consideration. Finally, if brains are good at perceptual inferences through adopting *parallel* processing, perhaps it will be necessary to adopt parallel processing for machine perception.

However this may be, we have now learned, from painful experience, that machine perception is extremely difficult to achieve. This implies that we do not understand biological perception adequately to design corresponding machines; but perhaps we can learn some useful things from living systems, including ourselves, clearly capable of perception. Continuing our emphasis on the cognitive component of biological perception, we will assume that the details of the physiology are less important in this context than the kinds of inference (from stored and sensory data) and the cognitive strategies by which objects and the future are inferred. But of course perceptual inference is not infallible. There are errors, and these may occur systematically in certain situations. Much as the logician may use logical paradoxes for revealing the nature of logic, so cognitive psychology can use perceptual phenomena for revealing perceptual assumptions and inference procedures. This is however to assume that at least some perceptual phenomena are due to *misplaced strategies of inference* rather than to *physiological malfunction*. Such visual phenomena as after-images, we may safely attribute to the physiology of the system, because the pattern of intense stimulation of the retina transfers to (is superimposed on) any other pattern. Other visual phenomena are specific to the pattern or to its probable significance in terms of the object world. These phenomena we may attribute to inference strategies rather than to the mechanism carrying out the strategy.

It seems useful, if we regard perceptions as the results of inference, to call perceptions *hypotheses*. This draws attention to their similarity, on this view, to hypotheses in science. In both cases slender data serve to make decisions appropriate through inference to situations which cannot be monitored directly, and which may lie in the future. A detailed comparison of perceptions as hypotheses in this sense could be rewarding. Meanwhile, we are in a position to describe *perceptual phenomena* as 'inappropriate hypotheses'. By asking why inappropriateness is generated, we may learn something of the inference procedures (and sometimes the physiology) of perception. As an example, we shall consider some new visual phenomena in these terms. These phenomena are not distortions, but are perceptually created visual features. If they are generated by misplaced inference, we may call them: 'cognitive creations' (Gregory 1972). This loaded term will be used; but they will be considered in terms of alternative paradigms.

A PARADIGM FOR COGNITIVE CREATIONS?

It is possible to devise simple line figures which evoke illusory contours and create large areas of enhanced or diminished brightness. Unlike the well-known brightness contrast effects, these illusory contours can occur in regions far removed from regions of physical intensity difference; and they can be orientated at any angle to physical present contours. Figure 1a is the figure described by Kanizsa (1955). An illusory triangle is observed whose apices are determined by the blank sectors of the three disks. The 'created' object appears to have sharp contours, which bound a large region of enhanced brightness.

We may discover what features are necessary for producing these creations by removing parts of this figure. Figures 1b,c,d, show such a sequence. Three dots spaced as the apices of an equilateral triangle (figure 1b) give no created contours, although they are readily seen as indicating a triangle. The broken triangle (figure 1c) does not evoke the figure (except perhaps slightly after the effect has been observed in figure 1a); but combining the equilaterally spaced dots with the broken triangle (figure 1d) does evoke the illusory object, though less markedly than with the sectored disks of figure 1a.

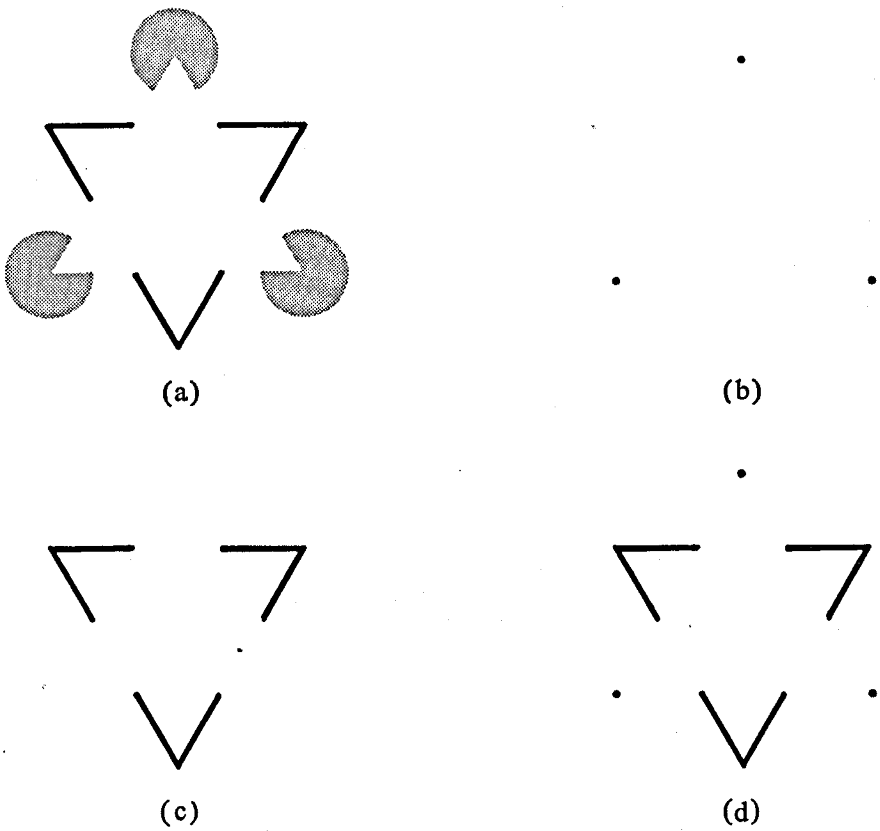

Figure 1

We can discount eye movements as important for these effects, for if the retinal image is stabilised (as by viewing the figures as after-images, produced with an intense electronic flash), then the effects are seen in the after-images, fixed to and moving precisely with the retina, with movement of the eyes. (When the initial positive after-image changes to a negative after-image, the whiter-than-white created area changes to a corresponding blacker-than-black area, just as when the figures are changed from negative to positive by optical means and viewed normally.)

These effects have particular theoretical interest, for they might be explained in terms of at least three very different perceptual paradigms. They might, with at least initial plausibility, be described in terms of: (1) gestalt, (2) physiological, or (3) cognitive-hypotheses paradigms.

(1) The gestalt paradigm is satisfied by supposing that the 'created' shapes are 'good figures', having high 'closure' and so on (as accepted by Kanizsa (1955) for the first figure).

(2) The physiological paradigm would be satisfied with the supposition that feature detector cells of the striate cortex are activated by the edges of the disk sectors (or less effectively by the dots) to give the appearance of continuous lines, though only their ends are given by stimulation.

(3) The cognitive-hypothesis paradigm, in which perceptions are regarded as going beyond available data to give 'object hypotheses' (Gregory 1970), would be satisfied by supposing that the created features are 'postulated', as supposed masking objects, to 'account' for the blank sectors of what are likely to be complete disks and the breaks in the lines of what is likely to be a complete triangle. The sectors and gaps are supposed, on this cognitive paradigm, to elicit the hypothesis of a masking object, lying in front of the given figure, which is likely to be complete but if complete must be partially masked. Like all other perceptions, this is a cognitive creation; but in this instance it is not appropriate to the facts and so is an illusion.

These paradigm views of the phenomena each give a different logical account and a different logical status to the phenomena. They also each have different experimental predictions, and so can be regarded as scientific rather than metaphysical statements. All three rival paradigms allow that there is a physiological basis; so each can ask: 'where is the fiction generated?' Simple experiments provide clear cut answers which at least rule out several possibilities.

By adopting the technique devised by Witasek in 1899, of sharing parts of the figures between the two eyes with a stereoscope, it is easy to show that the effect is not retinal in origin, but must be after the level of binocular fusion. This follows because the effect holds when the sectored disks are viewed with one eye and the interrupted line triangle with the other eye, when they are stereoscopically fused. By changing the angles of the disk sectors, so that they no longer meet by straight line interpolation, we find that the effect still occurs. The created form is now changed, to give interpolation with a *curved fictional contour*. This may be seen in figure 2. This new effect seems to increase the plausibility of the cognitive fiction notion – for it seems unlikely that 'curved-edge' detectors would be selected by the mis-aligned sectors; and it seems that these concave-curved (and other) figures which are created are not especially 'good' figures in the gestalt sense.

The black line on white background figures give a homogeneous whiter-than-white fictional region. The corresponding negative white line on black background gives a blacker-than-black region. The illusory intensity difference can be measured with a reference light spot, as in a matching photometer. The measured brightness difference is about 10 per cent. Both the black and the white illusory triangles are reported by our subjects as appearing somewhat in front of the rest of the figure. We have measured this objectively, by matching a stereoscopically viewed marker light spot to the apparent distance of the physical and created parts of the figures. This is compatible only with the cognitive paradigm.

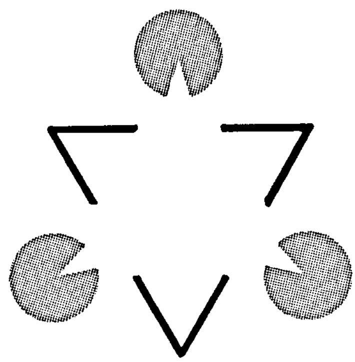

Figure 2

Not only contours, but large homogeneous areas of different brightness are created – but could such areas of different brightness be created by the line detectors of the physiological paradigm? Consider figure 1c. We have lines with gaps; but there is no observed difference in brightness between the inside and the outside of this triangle – and no contours between the gaps. So why should there be contours, and a brightness difference, with the created triangle? The lack of contours in the gaps of figure 1c and the absence of enhanced brightness show that aligned features are not sufficient for producing these effects. What seems to be needed is a high probability of an over-lying object which if it existed would give gaps by masking. This, however, would require processes of a logical sophistication beyond those believed to occur at the striate region; and concepts beyond those of classical physiology – the cognitive concepts of our third paradigm.

We find that at least some of the classical distortion illusions can be generated by these created contours and created regions of different brightness. Figure 3 shows a kind of Poggendorff figure, in which the usual parallel lines are physically absent but are generated by four sectored disks, placed well away from the interrupted oblique figure. Figures such as figure 4 also evoke apparently cognitive contours and they also produce distortion illusions. This seems significant, for how could these distortions be generated by physiological processes, signalling borders, if these borders are not signalled directly by sensory patterns but are, rather, called up from store as fictional hypotheses? It would seem that these effects are not due to physiological interaction effects, such as lateral inhibition; but they are compatible with the notion (Gregory 1965, 1970) that distortions can occur as a result

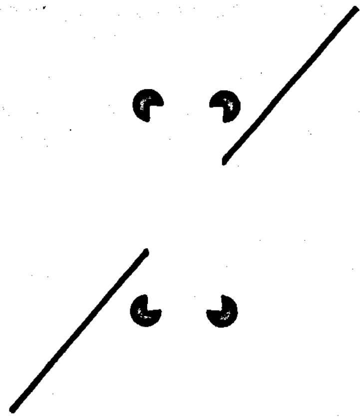

Figure 3

of scaling being set inappropriately to the surrounding objects, or line figures, by following usually predictive assumptions which do not apply to the given objects or line figures. (This happens especially when perspective features occur on flat picture planes.) On this view, it is not the physiology which generates the errors: it is the strategy which leads to error, though of course the strategy is carried out by physiological events. If we do not understand the strategy we will not understand the phenomena, even though we may

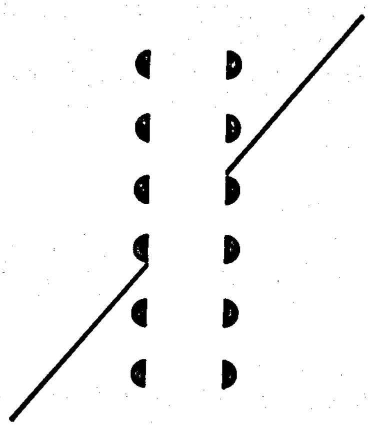

Figure 4

understand the physiology in every detail. If this view is approximately correct, we have an example where simply taking over paradigms of physics or physiology may be unhelpful, or seriously misleading. This remains true though we are not violating any physical or physiological principle with our cognitive paradigm.

The phenomena of perception deserve consideration though they may appear trivial. There is nothing new in this: some of physics' most dramatic successes came from questioning, and using as tools trivial-looking phenomena which may have been children's toy things. Studying how pith balls, suspended on silk threads, are affected by rubbed amber; and how lode stones, floating freely, point to the only fixed star in the sky, lead to new ways of seeing – to new paradigms of the physical object world. Perhaps only phenomena (and not philosophy) have the power to suggest new paradigms, – to break down old barriers which prevent us seeing how we see.

REFERENCES

Gregory, R.L. (1965) Inappropriate constancy explanation of spatial distortions, *Nature*, **207**, 4999, 891–3.

Gregory, R.L. (1970) *The Intelligent Eye*. London: Weidenfeld & Nicolson.

Gregory, R.L. (1972) Cognitive contours. *Nature*, **238**, 5358, 51–2.

Helmholtz, H. von (1963) *Handbook of Physiological Optics* (1867) (ed. Southall, J.P.C.S.). New York: Dover Publications, Inc.

Kanizsa, G. (1955) Margini quasi – percettivi in campi con stimulazioni omogonea. *Rivista di Psicologia*, **49**, 7, 30.

Kuhn, T.S. (1962) *The structure of scientific revolutions*. Chicago: University of Chicago Press.