II.C.

Imagination as Description.

§II.C.1. Images, Percepts, Descriptions, and Computers.

The notion that experiencing mental imagery is like considering a description of an object or scene seems to have a very much briefer history than the idea that it is like looking at a picture. The earliest serious comparison of visualisation and description that I know of is as recent as 1952, in a paper by J.M. Shorter {1}. Shorter is not at all concerned with psychological theories here but rather with the 'logical grammar' of ordinary language expressions; the comparison between visualising and describing is supposed to help clarify for us the sense in which imagery can be vague {2*}, it is certainly not supposed to tell us much about what is going on in our brains. However, when Dennett {3*} picks up on Shorter's remarks he clearly does intend to be suggesting a possible psychological theory. The claim is not the highly implausible one that when we imagine something we are really describing it to ourselves 'under our breath', in English, but rather that the representations in the brain underlying visualisation are best conceived of not as (functional) pictures but rather as descriptions in a language-like representational medium of neural states (Dennett calls it "brain writing"). Even so, Dennett is still speaking here as a philosopher and is only giving a tentative and unelaborated suggestion as to how a seeming obstacle to his real, metaphysical concerns might be dealt with. If awkward mental entities like images could be eliminated by some other theory he would, I am sure, be happy enough {4*}. The effective originators of the descriptionist theory of imagery were not really these philosophers, nor even psychologists, but computer scientists.

As with pictorialism, descriptionism identifies mental images with some sort of internal representational product of the processes of perception. They are produced during perception itself and are in some way reinstated in the appropriate 'place' when we remember or imagine. The idea that the end-product of perception is somehow linguistic, that it eventuates in internal assertions or descriptions of what we have perceived is at least as old as Plato. In the **Philebus** Socrates suggests that when we perceive:

> the conjunction of memory with sensations, together with the feelings consequent upon memory and sensation, may be said as it were to write words in our souls. And when this experience writes what is true the result is that true opinions and true assertions spring up in us, while when the internal scribe that I have suggested writes what is false we get the opposite sort of opinions and assertions. {5*}.

However, the contemporary currency of this idea seems to stem from that most modern of disciplines, Artificial Intelligence (AI). Indeed, 'computer vision' as one of the aims of AI research, has recently been **defined** in these terms:

> **Computer vision** is the construction of explicit, meaningful descriptions of physical objects from images [i.e. from the machine equivalents of retinal images, such as those formed in a TV camera]. (...) Descriptions are a prerequisite for recognising, manipulating and thinking about objects. {6}.

This conception of vision is thoroughly uncontroversial in the AI community. An even more blunt statement to the same effect is given in a very recent book on 'robot vision' by Horn {7}, and new collection of quite various papers on 'computer vision' {8} can be quite unselfconsciously entitled **From Pixels to Predicates**.

It is no accident that AI researchers should adopt this line {9*}. A good, quick, rough and ready characterisation of AI would be that it is the project to get digital computers to think {10*}, and the notion that this might be possible depends on the conceptualization of computers not as machines for doing arithmetic but as general 'symbol manipulation' devices {11*}. (The other necessary premise is, of course, that thought is nothing more than a particular type of such symbol manipulation {12}.) It is important to note that "symbol" is being used here in a particular, quite precise sense. According to Boden, within the computational context:

> A symbol is an inherently meaningless cipher that becomes meaningful by having meaning assigned to it by a user, who thereafter interprets it in a particular way. (...) There need be no intrinsic similarity between a symbol and what it symbolises. {13}.

Some symbols, notably pictures, are generally taken to symbolise their objects intrinsically or naturally, because they resemble them in some way (although whether this can ever really be the case is open to question {14}). Other symbols fairly obviously only mean what they do in virtue of some arbitrary convention, and the paradigmatic examples of this are the words of ordinary languages (with just a few onomatopoeic exceptions). Within an electronic digital computer the basic "ciphers", the physical tokens which are to be interpreted as symbols, are the on-off states of certain electronic switches, and all of the symbols which a computer is able to manipulate must be made up of sequences and concatenations of sequences of these. The things which such sequences most obviously **resemble** are the numerals of binary arithmetic, which is no doubt one reason why computers are so widely taken to be essentially arithmetic machines {15*}. But the sequences certainly can be, and are, assigned as symbols for anything that one might choose. It is this latter sort of meaning assignment which makes possible all forms of non-numerical computing, including (if it **is** possible) AI. But the obviously arbitrary assignment of meaning at the fundamental level means that the natural model for these symbolic systems is language, not pictures. In a seeing computer then, if such a thing can exist, the internal representation of the input visual scene will be linguistic in this sense. A linguistic representation of any complex object which goes beyond the trivial level of merely naming it, and in some way reflects its structure, is clearly a description.

This model of perceptual representation carried over very naturally into the earliest discussions of 'computer imagery' {16}. Furthermore, the fact that neurones resemble a computer's memory elements at least to the extent of being two-state electrical devices lends plausibility to the carrying over of the same sort of argument to the case of brains as well. Our seeing something, it is claimed, is not a matter of having some sort of picture produced inside us; rather it is a matter of our being caused to hold certain beliefs {17*}, certain propositions about the visual scene before us. These propositions are to be expressed in the brain in some "highly abstract language" {18*} analogous to a computer's internal symbol system. Imagining is not the production of internal pictures but the entertaining of such descriptive propositions {19}. This view is, indeed, widely known as the "propositional" theory of imagery.

Now computers clearly can produce pictures. Computer graphics programs, not to mention the Kosslyn & Shwartz imagery simulation, do just that. But these pictures must themselves be generated from descriptions couched in the symbol system of the currently running program, even if they are just descriptions specifying which points are to be illuminated or marked and which are not. There is a sense, then, in which Kosslyn's quasi-pictorial model is itself a form of descriptionism; if it were not it could not be embodied in a computer program. Such considerations (if not the mere fact of his producing a running computer simulation) have led certain figures

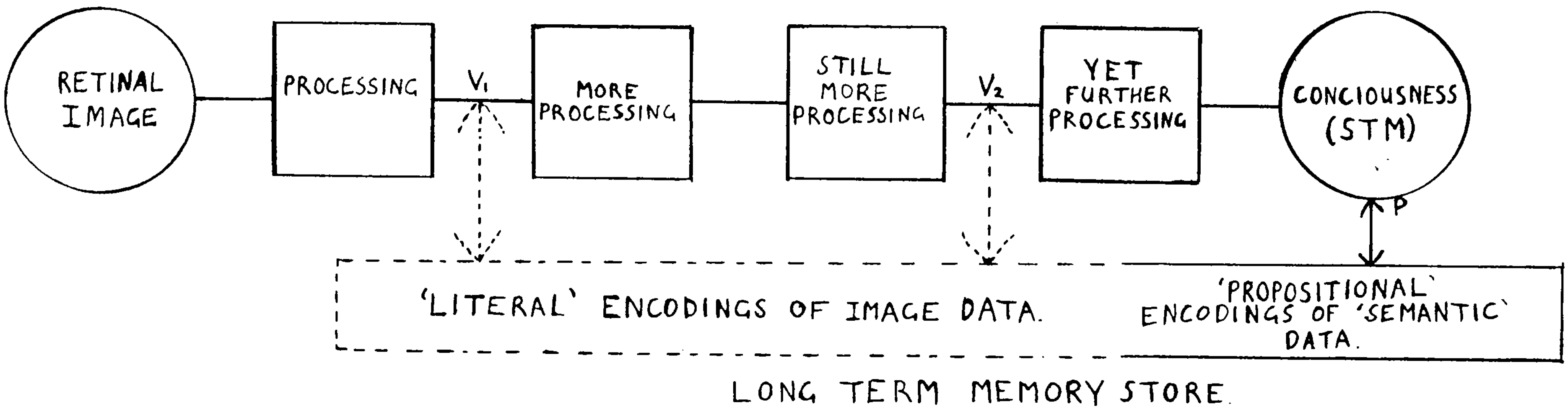

Figure II.C.1-1.
(loosely based on Neisser [1976 p.17].)

Solid parts indicate aspects more or less accepted by both modern 'quasi-pictorial' and 'descriptional' ('propositional') theories of imagery. Dotted parts are peculiar to 'quasi-pictorialism'. Arrows indicate directions of information flow. V1 and V2 are meant as **alternative** possibilities for the point where, on a 'quasi-pictorial' account, image specifying data might be abstracted from or inserted into the visual information processing stream, alternative sites for Kosslyn's "visual buffer". I believe that Kosslyn's theory is importantly ambiguous as to whether this is at an early or a late 'stage of processing'. **P** is the point at which descriptive 'propositions' are taken to store or reinserted into consciousness. According to 'descriptionist' theories their presence here is sufficient for experiencing imagery.

from within the mainstream of AI to look quite favourably on the work of the psychologist Kosslyn {20}. Even George Baylor, who made the first real attempt at computer simulation of imagery {21}, and whose work clearly did much to inspire the better known advocates of descriptionist or 'propositional' image theories {22}, has recently recommended the Kosslyn & Shwartz model above his own {23}.

Nevertheless, Kosslyn and the more committed 'propositionalists', notably Pylyshyn, still refuse to be reconciled {24}, and I think rightly. Although both sides work within the 'information processing' 'paradigm' and both want their theories to be programable there are important substantive differences in their views. If we return to figure II.B.1_3 [which for convenience is reproduced here as figure II.C.1_1] the points of agreement and disagreement should become apparent. **Both** sides here regard the immediate end-product of both vision and imagination to be "abstract propositions" in our current awareness or Short Term Memory (STM) {25*}. To hold such propositions in STM just is to be aware, conscious, of something. The two approaches do not differ in any vital respect over what goes on in vision: it consists in a progressive reduction of stimulus information to propositional descriptions in STM. Subsequently these propositions, or a selection of them, will go to storage in the Long Term Memory (LTM).

According to the 'propositional' or 'descriptional' theory imaging simply consists in retrieving propositions from LTM and reinstating them in STM, which Pylyshyn {26} describes as a sort of "cognitive workspace" for problem solving. Provided these propositions are ones about, and originally produced by, perceptual appearances they will induce an appropriately quasi-perceptual, imagistic, experience when they return to the workspace of consciousness. However, not all the original perceptual propositions may get there, and other, non-perceptual propositions may be present too. Furthermore, new descriptions may be concocted from perceptual propositions coming from quite unrelated experiences. In such ways the 'propositionalist' can neatly explain the image indeterminateness, the "cognitive penetration" effects and the composite images which raised such difficulties for quasi-pictorialism in the previous chapter.

In Kosslyn's theory, of course, the image specifying data, which are stored in a rather less 'processed' form, are not inserted directly into STM from LTM. They must first go through some of the later stages of 'visual information processing'. What this means is that although the image is still based on a description it does not function as such since information **implicit** in it may still be extracted, and by non-inferential, specifically **visual** processes {27}. This is certainly a substantive difference.

We have already seen what problems this essential difference causes for quasi-pictorialism. It raises other problems for descriptionism, I believe, and we will soon turn to them. Before we do that we will try, in the next section, to get a more concrete sense of what is meant by "propositions" and "descriptions" in these contexts. Before we do even that, however, we should take note of at least one significant way in which **neither** of these theories scores over the other. The account of our conscious experience in both quasi-pictorialism and descriptionism is **the same**. In both cases consciousness is a matter of having propositions in STM. Whether they have got there from LTM directly or indirectly makes no difference; in either case they will be equally like the propositional end-products of perception itself. The visual processing activity postulated by Kosslyn's type of theory will be quite unconscious (just like the rest of visual processing) and so, perhaps surprisingly {28*}, neither theory will give a better account of the immediate experience of imagery (or perception) than the other. The corollary of this is that **both** theories stand or fall together on the viability of this 'propositional' view of the mind. That issue will be deffered until the final section of this chapter.

§II.C.2. Simulating Imagery in Terms of Descriptions.

In the early days of AI research in the U.S.A. there was a certain amount of amicable disagreement between