**NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:**
The copyright law of the United States (title 17, U.S. Code) governs the making of photocopies or other reproductions of copyrighted material. Any copying of this document without permission of its author may be prohibited by law.

COMPUTER SCIENCE DEPT.
TECHNICAL REPORT FILE

SSX
005

CABINET

# EXPLAINING INFANT PERCEPTION:

# INSIGHTS FROM ARTIFICIAL INTELLIGENCE?

Julie C. Rutkowska

Summer 1984

CABINET

Cognitive Science Research Paper

Serial no: CSRP 005

University of Sussex
Cognitive Studies Programme
School of Social Sciences
Falmer
BRIGHTON BN1 9QN

CABINET

UNIVERSITY LIBRARIES
CARNEGIE-MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213

ROOM USE ONLY
UNTIL ____

## REFERENCES


Ball, W. and Tronick, E. (1971) Infant responses to impending collision: optical and real. Science, 171, 818-820.

Ball, W and Vurpillot, E. (1976) La perception du mouvement en profoundeur chez de nourrisson. Annee Psychologique, 76 383-400.

Ball, W. & Vurpillot, E. (1981) Action and perception of displacements in infancy. In: G.E. Butterworth (ed.) Infancy and Epistemology. Sussex: Harvester Press.

Bobrow, D. and Raphael, B. (1974) New programming languages for artificial intelligence research. ACM Computing Surveys, 6, 155-174.

Boden, M.A. (1972) Purposive Explanation in Psychology. Cambridge, Mass.: Harvard University Press.

Boden, M.A. (1977) Artificial Intelligence and Natural Man. Sussex: Harvester Press.

Boden, M.A. (1981) Minds and Mechanisms: Philosophical Psychology and Computational Models. Sussex: Harvester Press.

Boden, M.A. (1983) Artificial Intelligence and animal psychology. New Ideas in Psychology, 1, 1-33.

Bower, T.G.R. (1966) The visual world of infants. Sci. Amer., 251, 80-92

Bower, T.G.R. (1972) Object perception in infants. Perception, 1, 15-30.

Bower, T.G.R. (1979) Human Development: San Francisco: Freeman.

Bower, T.G.R. (1982) Development in Infancy. 2nd edition. San Francisco: Freeman.

Bower, T.G.R., Broughton, J.M. and Moore, M.K. (1970a) The coordination of visual and tactual input in infants. Perception and Psychophysics, 8, 51-53.

Bower, T.G.R. Broughton, J.M. and Moore, M.K. (1970b) Demonstration of intention in the reaching behaviour of neonate humans. Nature, 228, 679-681.

Brainerd, C.J. (1978) Piaget's Theory of Intelligence. Englewood Cliffs, N.J.: Prentice-Hall.

Brazelton, T.B., Koslowski, B. and Main, M. (1974) The origin of reciprocity: The early mother-infant interaction. In M. Lewis and L.A. Rosenblum (eds.) The Effect of the Infant on its Caregiver. Wiley.

Bruner, J.S. (1968) Processes in Cognitive Growth: Infancy. Clarke University Press.

Bruner, J.S. and Koslowski, B. (1972) Visually preadapted constituents of manipulatory action. Perception, 1, 3-14.

Butterworth, G.E. (1981a) Infancy and Epistemology: An Evaluation of Piaget's Theory. Sussex: Harvester Press.

Butterworth, G.E. (1981b) The origins of auditory-visual perception and visual proprioception in human development. In: H.I. Pick and R. Walk (eds.) Perception and Experience, Volume II. New York: Plenum.

Butterworth, G.E. and Cicchetti, D. (1978) Visual calibration of posture in normal and motor retarded Down's syndrome infants. Perception, 7, 513-525.

Butterworth, G.E. and Hicks, K. (1977) Visual proprioception and postural stability in infancy. Perception, 6, 255-262.

Campos, J.J. Hiatt, S., Ramsay, D., Henderson, C. and Svejda, M. (1978) The emergence of fear on the visual cliff. In: M. Lewis and L.A. Rosemblum (eds.) The Development of Affect. vol. 1. New York: Plenum Press.

Campos. J.J., Langer, A. and Krowitz, A. (1970) Cardiac responses on the visual cliff in prelocomotor human infants. Science, 170, 196-197.

Clowes, M.B. (1971) On seeing things. Artificial Intelligence, 6, 79-116.

Cohen, L.B. (1979) Our developing knowledge of infant cognition. American Psychologist, 34, 894-899.

Cohen, L.B. and Strauss, M.S. (1979) Concept acquisition in the human infant. Child Development, 50, 419-424.

Costall, A. (1981) On how so much information controls so much behaviour: James Gibson's theory of direct perception. In G.E. Butterworth (ed.) Infancy and Epistemology. Sussex: Harvester Press.

Dasen, P.R. (1977) Are cognitive processes universal? A contribution to cross-cultural Piagetian psychology. In N. Warren (ed.) Studies in Cross-Cultural Psychology, Vol. 1. London and New York: Academic Press.

Day, R.H. and McKenzie, B.E. (1977) Constancies in the perceptual world of the infant. In: W. Epstein (ed.) Stability and Constancy in Visual Perception: Mechanisms and Processes. New York: Wiley.

Deregowski, J.B. (1978) In search of a wider perspective: Cross-cultural studies. In: A. Burton and J. Radford (eds.) Thinking in Perspective. London: Methuen.

Draper, S.W. (1980) Using models to augment rule-based programs. Proceedings A.I.S.B., Amsterdam, Holland.

Fodor, J.A. (1968) Psychological Explanation, New York: Random House.

Fodor, J.A. (1980) Methodological solipsism considered as a research strategy in cognitive psychology. Behavioral and Brain Sciences, ]>_, no. 1.

Fodor, J.A. and Pylyshyn, Z.W. (1981) How direct is visual perception?: Some reflections of Gibson's "ecological approach". Cognition, 9, 139-196.

Furth, H.6. (1969) Piaget and Knowledge: Theoretical Foundations. Prentice-Hall.

Gibson, E.J. and Rader, N. (1979) Attention: The perceiver as performer. In: G.A. Hale and M. Lewis (eds.) Attention and Cognitive Development. New York: Plenum Press.

Gibson, J.J. (1950) The Perception of the Visual World. Boston: Houghton-Mifflin.

Gibson, J.J. (1968) The Senses Considered as Perceptual Systems. London: George Allen and Unwin.

Gibson, J.J. (1979) The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.

Gratch, G. (1975) Recent studies based on Piaget's view of object concept development. In L.B. Cohen and P. Salapatek (eds.) Infant Perception: From Sensation to Cognition. Volume II: Perception of Space, Speech, and Sound. Academic Press.

Haugeland, J. (1978) The nature and plausibility of cognitivism. Behavioral and Brain Sciences, J_, 215-260.

Hinton, G. (1979) Some demonstrations of the effects of structural descriptions in mental imagery. Cognitive Science, ^5 231-250.

Huffman, D.A. (1971) Impossible objects as nonsense sentences. Machine Intelligence, 6>, 295-323.

Humphrey, N.K. (1974) Vision in a monkey without striate cortex. Perception, 3, 241-255.

Jacobson, S. (1979) Matching behaviour in the young infant. Child Dev., j>0, 425-430.

Kaye, K. (1982) The Mental and Social Life of Babies: How Parents Create Persons. Sussex: Harvester Press.

Lee. D-N. and Aronson, E- (1974) Visual proprioceptive control of standing in human infants. Perception and Psychophysics, JJL' 529-532.

Lettvin, J., Maturana, H., McCuUoch, W. and Pitt, W.S. (1958) What the frog's eye tells the frog's brain. Proceedings I.R.E., ^7, 1940-1951.

Mackworth, A. (1983) On seeing things, again. Proceedings 8th IJCAI, vol. 2, Karlsruhe, W. Germany.

Mandler, J.M. (1981) Representation. U. Cal San Diego, C.H.I.P. Report no. 103.

Maratos, O. (1982) Trends in the development of imitation in early infancy. In: T. Bever (ed.) Regressions in development. Hillsdale, N.J.: Lawrence Erlbaum.

Marr, D. (1977) Artificial intelligence - a personal view. Artificial Intelligence, 9, 37-48.

Marr, D. (1980) Visual information processing: the structure and creation of visual representations. Phil. Trans. Royal Soc. London, Series B, 290, 199-218.

Marr, D. (1982) Vision. San Francisco: Freeman.

Meltzoff, A.N. (1981) Imitation, inter-model coordination and representation in early infancy. In: G.E. Butterworth (ed.) Infancy and Epistemology. Sussex: Harvester Press.

Meltzoff, A.N. and Borton R.W. (1979) Inter-modal matching by human neonates. Nature, 282, 403-404.

Meltzoff, A.N. and Moore, M.K. (1977) Imitation of facial and manual gestures by human neonates. Science, 198, 75-78.

Mendelson, M.J. and Haith, M.M. (1976) The relation between audition and vision in the human newborn. Monographs S.R.C.D., 41, no. 4.

Michaels, C.F. and Carello, C. (1981) Direct Perception. New York: Prentice-Hall.

Minsky, M. (1980) Decentralized minds. Behav. Brain Sci., 3, no. 3.

Miller, L. (1978) Has artificial intelligence contributed to our understanding of the human mind? A critique of arguments for and against. Cog. Sci., 2, 111-127.

McKenzie, B.E. and Day, R.H. (1973) Perceptual shape constancy in early infancy. Perception, 3, 315-326.

Mounoud, P. and Hauert, C.A. (1982) Sensorimotor and postural behaviours: Their relation to cognitive development. In: W.W. Hartup (ed.) Review of Child Development Research. Volume 6. Chicago and London: U Chicago Press.

Mounoud, P. and Vinter, A. (1981) Representation and sensorimotor development. In: G.E. Butterworth (ed.) Infancy and Epistemology. Sussex: Harvester Press.

Neisser, U. (1967) Cognitive Psychology. New York: Appleton-Century-Crofts.

Neisser, U. (1976a) Cognition and Reality. San Francisco: Freeman.

Newell, A. (1980) Physical symbol systems. Cognitive Science, 4, 135-183.

Newell, A. and Simon, H.A. (1976) Computer science as empirical inquiry: Symbols and search. Communications ACM, 19, 113-126.

Nickerson, R.S. and Adams, M.J. (1979) Long term memory for a common object. Cognitive Psychology, 11, 287-307.

Oatley, K. (1978) Perceptions and Representations: The Theoretical Bases of Brain Research and Psychology. London: Methuen.

Olson, R., Yonas, A. and Cooper, R.G. (1980) Development of Pictorial Perception. In: M.A. Hagen (ed.) The Perception of Pictures, vol. 2. New York: Academic Press.

Piaget, J. (1953) The Origin of Intelligence in the Child. London: Routledge and Kegan Paul.

Piaget, J. (1955) The Child's Construction of Reality. London: Routledge and Kegan Paul.

Pope, M. (1982) Visual proprioception in infant development. Annual Conference B.P.S. Developmental Section, Durham.

Prazdny, S. (1980) A computational study of aperiod of infant object-concept development. Perception, 9, 125-150.

Pylyshyn, Z.W. (1978) Computational models and empirical constraints. Behavioral and Brain Sciences, 1, 93-127.

Pylshyn, Z.W. (1980) Computation and cognition: Issues in the foundations of Cognitive Science. Behavioral and Brain Sciences, 3, 111-169.

Regan, D., Beverley, D.I. and Cynader, M. (1979) Stereoscopic subsystems for position in depth and for motion in depth. Proc. Royal Soc. London B. 204, 485-501.

Reichardt, W. and Poggio, T. (1979) Visual control of flight in flies. Cited in D. Marr, Vision, 1982.

Roberts, L. (1965) machine perception of three-dimensional solids. In: J.T. Tippett, D.A. Berkowitz, L.C. Clapp, C.J. Koester and A Vanderburgh (eds.) Optical and Electro-Optical Information Processing. Cambridge, Mass.: M.I.T. Press.

Rock, I. (1977) In defense of unconscious inference. In: W. Epstein (ed.) Stability and Constancy in Visual Perception. New York: Wiley.

Runeson, S. (1977) On the possibility of "smart" perceptual mechanisms. Scandinavian Journal of Psychology, 18, 172-179.

Russell, J. (1981) Piaget's theory of sensorimotor development: Outline, assumptions and problems. In: G. Butterworth (ed.) Infancy and Epistemology: An Evaluation of Piaget's Theory. Sussex: Harvester Press.

Searle, J.R. (1980a) The intentionality of intention and action. Cognitive Science, 4, 147-170.

Shapiro, B.A. (1974) A Survey of Problem Solving Languages and Systems. National Science Foundation, Technical Report No. TR 235.

Shaw, R. and Bransford, J. (1977) Introduction: Psychological approaches to the problem of knowledge. In: R. Shaw and J. Bransford (eds.) Perceiving, Acting and Knowing: Toward an Ecological Psychology. Hillsdale, N. J.: Lawrence Erlbaum.

Shaw, R.E. and McIntyre, M. (1974) Algoristic foundations of cognitive psychology. In: W. Weimer and D.S. Palermo (eds.) Cognition and the Symbolic Processes. Hillsdale, N.J.: Lawerence Erlbaum.

Simon, H.A. (1969) The Sciences of the Artificial Cambridge, Mass.: M.I.T. Press.

Sinha, C. (in press) A socionaturalistic approach to language acquisition. In: G.E. Butterworth, J.C. Rutkowska and M. Scaife (eds.) Evolution and Developmental Psychology Sussex: Harvester Press.

Sloman, A. (1978) The Computer Revolution in Philosophy. Sussex: Harvester Press.

Sloman, A. (1980) What kind of indirect process is perception? Behavioral and Brain Sciences, 3, 401-404.

Sloman, A. (1982a) Computational epistemology. Cahiers de la Fondation Archives Jean Piaget, Geneva, June, No. 2-3.

Sloman, A. (1982b) Image interpretation: the way ahead? Paper presented at Rank Prize Fund Conference on 'Physical and Biological Image Processing'. London: The Royal Society, 29th September.

Spelke, E.S. (1976) Infants inter-model perception of events. Cog. Psychol., 8, 553-560.

Spelke, E.S. (1978) The infants acquisition of knowledge of bimodally specified events. Unpublished manuscript. U. Pennsylvania.

Spelke, E.S. (1979) Perceiving bimodally specified events in infancy. Dev. Psychol., 15, 626-636.

Spelke, E.S. (1982) Perceptual knowledge of objects in infancy. In: J. Mehler, E.C.T. Walker and M. Garrett (eds.) Perspectives on Mental Representation: Experimental and Theoretical Studies of Cognitive Processes and Capacities. Hillsdale, N.J.: Lawrence Erlbaum.

Steels, L. (1982) Constraints as consultants. Proceedings ECAI, Orsay, France.

Trevarthen, C. (1974) Conversations with a two-month old. New Scientist, 62, 230-235.

Trevarthen, C. (1980) Neurological development and the growth of psychological functions. In: H.J. Sants (ed.) Developmental Psychology and Society. London: Macmillan.

Trevarthen, C., Hubley, P. and Sheeran, L. (1975) Les activities innees du nourrisson. La Recherche, 6, 447-458.

Turvey, M.T., Shaw, R.E., Reed, E.S. and Mace, W.M. (1981) Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn. Cognition, 9, 237-304.

Ullman, S. (1979) The Interpretation of Visual Motion. Cambridge and London: M.I.T. Press.

Ullman, S. (1980) Against direct perception. Behavioural and Brain Sciences, 3, 373-381.

Uzgiris, I.C. and Hunt, J. McV. (1975) Assessment in Infancy: Ordinal Scales of Psychological Development. Urbana: U. Illinois Press.

Wagner, S.J. (1982) A philosopher looks at Artificial Intelligence. In: D.L. Farwell, S.C. Helmreich and W.D. Wallace (eds.) Issues in Cognitive Science. Linguistics Students Org., U. Illinois

Waltz, A. (1975) Generating semantic descriptions from drawings of scenes with shadows. In: P. Winston (ed.) The Psychology of Computer Vision, New York: McGraw Hill.

Weimer, W.B. (1977) A conceptual framework for cognitive psychology: Motor theories of the mind. In: R. Shaw and J. Bransford (eds.) Perceiving, Acting and Knowing: Toward an Ecological Psychology. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

White, B.L. (1969) The initial coordination of sensorimotor schemas in human infants. In: D. Elkind and J.H. Flavell (eds.) Studies in Cognitive Development. New York: O.U.P.

Wilcox, S. and Katz, S. (1981) The ecological approach to development: an alternative to cognitivism. Journal Experimental Child Psychology, 32, 247-263.

Wilks, Y. (1982) Machines and consciousness. University of Essex, Cognitive Studies Centre Memorandum - 8, February.

Winograd, T. (1972) Understanding Natural Language. New York: Academic Press.

Winograd, T. (1977) On some contested suppositions of generative linguistics about the scientific study of language. Cognition, 5, 151-179.

Winograd, T. (1980) What does it mean to understand language? Cog. Sci., 4, 209-242.

Winston, P.H. (1977) Artificial Intelligence Reading, Mass.: Addison-Wesley.

Yonas, A. (1979) Studies of spatial perception in infancy. In: A.D. Pick (ed.) Perception and Its Development: A Tribute to Eleanor J. Gibson. Hillsdale, N. J.: Lawrence Erlbaum.

## INTRODUCTION AND OVERVIEW

This paper is concerned with the relationship between developmental psychology and artificial intelligence (henceforth AI). The general question characterizing the work which it represents is: How should we conceptualize the abilities of young infants? The focus for this question is the problem of infant perception and the current attempts to explain it in terms of Gibson's (1950, 1968, 1979) theory of "direct perception[11] and the "ecological psychology" which has developed from it (e.g. Michaels & Carello, 1981; Turvey, Shaw, Reed & Mace, 1981; Wilcox & Katz, 1981). The present paper presents a case for a computational cognitivist account of the infant, and aims to show that developmentalists smitten by the appeal of Gibson's theory should not accept the jaundiced accounts of processing or computational approaches to which it has given rise. Specifically, it argues that the concept of action provides an appropriate organizing structure for infant ability, and that action is best conceptualized computationally.

Section A outlines a subset of the empirical findings and theoretical concerns of contemporary infancy research. These focus on the conflict between Piaget's cognitivist and constructivist account of the infant and Gibson*s theory of direct perception. It is argued that empirical observations which appear to demonstrate the superiority of one theory as against the other remain poorly understood in terms of the mechanisms which could underlie infant performance. Pursuit of either/or conflicts engendered by the debate cannot prove useful since they are based on an inappropriate analysis of a complex system.

Levels of explanation in AI and ecological psychology are contrasted in Section B, with particular emphasis on the virtual/physical machine distinction and the frequently misunderstood computational concepts of symbolism, representation and generativity. Contrary to ecological psychology's view of cognitivism, it is argued that AI models of vision - characterized by scene-analysis and the work of Marr - make possible a psychologically meaningful decomposition of vision, without implying the degeneracy of environmental information assumed by constructivist theories such as Piaget's. However, such models cannot explain the process of perceiving unless they take into account the selective use of information which Is characteristic of action systems, and it is suggested that a computational analysis can be extended to clarify ecological psychology's notion of "affordances".

Section C outlines the relevance of this approach to understanding and explaining infant action systems. It is argued that the notion of computation as rule-governed structure manipulation must be taken to include environmental as well as intra-subject structures. Viewing the causal embedding of the infant in the environment in terms of the program/machine distinction provides a unique vantage point on behaviour, which is viewed as overt program-governed process, and on the subject-environment relationship, which is viewed as a transaction. This computational framework can handle both causal and purposive aspects of action more coherently than alternative approaches. Following on from this, the framework is used to analyse two issues which are central to the infancy literature discussed in Section A: the "abstractness" of perception, and functional perceptual-behavioural relationships.

## A. UNDERSTANDING INFANTS: THE APPEAL OF GIBSON AND SOME PROBLEMS

Most mainstream developmental research has been concerned with testing theories, emphasising the construction of tasks on which the psychological mechanisms the subject is assumed to need for successful performance are predicted by one theory as opposed to another. In Miller's (1978) terms, it is preoccupied with "theory demonstration" - the belief that scientific progress is made by collecting empirical evidence for the truth or falsity of theories - in contrast with "theory development". The latter focusses on the construction of models of mechanisms which are sufficient to produce the performance of interest, and it is largely unconcerned with empirical verification in the traditional sense.

In the absence of competing, large-scale theories explicitly angled at cognitive development, much effort has been expended on empirical explorations of aspects of Piaget's (1953, 1955) theory of the nature and development of knowledge. Recently, however, research which fails to confirm what appear to be Piagetian predictions has come together with studies influenced by Gibson's (1950, 1968, 1979) theory of direct perception to suggest a new, Gibsonian framework for understanding the infant (see, e.g. Butterworth 1981a). The apparently striking contrasts between Piagetian and Gibsonian theories readily lend them to exploration within a theory demonstration framework with the ᶠCan infants X or not?[1] type of questions which it poses. In this section, it will be argued that empirical research which has been taken to provide a convincing demonstration of the validity of Gibson's theory as contrasted with Piaget's is, in fact, fairly vague with respect to the actual mechanisms involved. This does not mean that Piaget's theory Is right after all. It does mean that it is a mistake to try to decide empirically between two equally insufficient theories and that theory development is needed instead.

### A.1 The main theoretical disagreements between Piaget and Gibson

The basic opposition between these two theories revolves around the concepts of perception, action and representation. Piaget's (1953, 1955) theory stresses the misleading nature of perception and the independence of the senses at birth. For him, the term perception closely corresponds to static sensory inputs which provide no information on spatial or temporal relationships between objects and events in the world- Thus, the young subject's world lacks organization: the infant does not relate the object s/he sees to the object s/he touches or hears, objects are not recognized when they move or are encountered at different times, and "out of sight is out of mind". Starting with such momentary sense images, the infant must depend on action for knowledge of the world. The exercise and coordination of initially unrelated voluntary activities, such as looking at, listening to and manipulating objects, brings meaning and order to the infant's initially chaotic experience.

Piaget poses a deficit model of this "sensory-motor" period and of action, emphasising that the infant must be freed from the limitations imposed by perception and action through development to a conceptual-representational level *of* functioning. He equates this with the formation of new mechanisms which model absent objects and events and are assumed to develop around the second year. Evidence for such a qualitative transition is provided by the infant's successful performance on tasks such as delayed imitation of facial gestures and search for an object whose route to its hiding place cannot be seen because, for example, it is placed in a container which is then moved to a series of locations at one of which the object is tipped out.

Such claims may seem bizarre for those coming from outside developmental psychology. As Neisser (1976 ) observes, Piaget's views appear to owe more to philosophers than to observation of infants. But it is precisely because many of Piaget's empirical observations show a high degree of replicability (Dasen, 1977; Gratch, 1975; Uzgiris and Hunt, 1975; White, 1974) that his theory has proved so influential. For example, infants of about 9 months can be relied upon to search where they initially found a hidden object when they subsequently see it hidden in a new, equally accessible location: the ubiquitous AB error. However, Piaget's claim that this 'error' marks the infant's transition from erroneous object knowledge based on previously successful action to objective understanding through representation of the object remains controversial (see e.g. Gratch (1975) and Harris (1975) for reviews of competing interpretations).

In contrast with Piaget, Gibson (1950, 1968, 1979) argues, from an evolutionary framework, that the infant's knowledge of the world lies in preadapted perception. Perceptual systems evolved to exploit ecologically valid information to guide activity, so how can perception be mistaken, misleading or meaningless? Infant perceptual systems are assumed to be "tuned", from the outset, to "pick up" what Gibson terms "higher-order invariants": for example, relationships within the optic array and its transformations which provide unambiguous information about the structural and transformational properties of objects and events and the activities they "afford" the infant. Gibson stresses that perception does not start from meaningless sense-data which must be supplemented by non-perceptual processes such as motor acts or mentalistic concepts acquired through experience. For example, one can just see that an object is solid or graspable without having to associate visual sensations with tactual sensations through experience.

Gibson rejects any version of constructivism, including Piaget's theory, and this emphasis is reflected in the "ecological psychology" which has developed from his theorizing. Michaels and Carello's (1981) overview places this new approach in opposition to both traditional and modern "information processing" theories which contend that (a) the input does not provide a sufficient basis for knowledge of the environment, and (b) the organism provides the "embellishment" necessary for perception, usually in the form of some kind of memory. Following from this, a central contention of the ecological approach is that the appropriate level of analysis for a psychological account of perception should concern laws relating the organism and its environment, not processes within the organism as cognitivist positions, including Piaget's, have claimed. As Turvey, Shaw, Reed and Mace (1981) put it: "an adequate theory of perception requires not more psychology but more physics of the kind appropriate to living things and their environments. Perception is not in the province of mental states or formal languages of representation and computation but in the province of physical principles at the scale of ecology (p.254)".

Gibson's approach has variously attracted adjectives such as "revolutionary" (Turvey, Shaw, Reed & Mace, 1981) and "suicidal" (Fodor & Pylyshyn, 1981). Empirical research with infants which has led researchers to apply the first classification may be divided into two main areas which will be considered in turn- Section A. 2 discusses evidence that infants use information which is not limited to literal, static sense arrays and that there are innate links between what have traditionally been considered as separate sensory modalities in developmental accounts such as Piaget's. Section A.3 goes on to evidence that infants show organized and object-appropriate functional activity prior to experience *or* practice in coordinating perception with their behaviour.

## A.2 Sense/Modality Links

Several experimental paradigms reveal relations between seeing and hearing. For example, sound systematically affects the visual scanning of newborn infants, in light or dark, in ways which cannot be explained in terms of heightened activation or arousal due to the extra stimulation it provides (Mendelson & Haith, 1976). Newborns may be sensitive to the match between visual and auditory information for an objects spatial location: they sometimes turn to the source of a sound, and are more likely to do so if a visual target is congruent with it (Butterworth, 198V't; Castillo & Butterworth, 1979). But do such findings show that the infant has an "expectancy" that something heard implies something to see, as the question has frequently been phrased in challenging Piaget's theory? We lack an adequate language for discussing such notions effectively. Alternatively, are these findings compatible with the view that infants possess some "sensitivity" to correspondence between information from these modalities which serves as a basis for exploring and learning about the objects and events they will encounter?

Spelke (1976, 1978, 1979) has favoured the latter interpretation in her work on infant reactions to bimodally specified, auditory-visual events. She found that 4-month old infants who are shown two adjacent motion films will spend a significantly longer time looking at the one depicting events which match a single, centrally located soundtrack. This ability to discriminate correspondence appears to be based on amodal, temporal parameters of events, such as tempo and/or simultaneity of sounds and visual impacts. When sound tempo is varied, infants look longer at the film of a toy bouncing at the same rate, even if film and soundtrack are not synchronized. Infants look longer at the film of a toy whose bounces coincide with impact sounds when the alternative film shows a toy bouncing at the same rate but out of phase with the soundtrack. This effect is replicable with novel objects; during very brief experimental sessions, infants appear to be able to relate objects and sounds so that later, when a sound is briefly presented, they first look quickly to the object it previously accompanied.

Quite dramatic evidence of infants relating what they see and feel comes from Meltzoff and Moore's (1977) demonstration that neonates will "imitate" adult facial gestures such as tongue protrusion which they can feel but not see themselves perform. Also, very young infants who suck one of two differently shaped objects (either a smooth sphere or a sphere with protrusions) will subsequently spend longer looking at the one they sucked when the two objects are presented side-by-side for visual inspection (Meltzoff & Borton, 1979).

While this research is convincingly 'non-Piagetian', both these and Spelke's findings have prompted further questions, such as: Is the mechanism underlying these modality relations some cross- or inter-modal matching process or a single, "supramodal" system? Bower (1974, 1979, 1982) argues for the latter in suggesting a "primitive unity of the senses" in which the infant cannot discriminate between information from the different modalities, each of which provides interchangeable information on the object. Meltzoff (1981) acknowledges the difficulty by stressing that he uses terms such as "inter-modal matching" as a convenient convention for labelling the problem involved and not as an intended description of the underlying processes infants might use to solve it. The terminology used in other discussions of imitation-related mechanisms ranges from the "innate releasing mechanism" (Jacobson, 1979) to a "concept" or "notion" (Mandler, 1981).

Overall, what do such findings mean? If one uses this research to ask if Piaget is right or wrong, the answer is clear: such results are incompatible with his theory, illustrating abilities which he would not expect the young infant to possess. But disconfirming Piaget does not, by default, support the validity of Gibson's position. The question of what mechanisms actually <u>are</u> involved is far more difficult to address. Infants appear to be capable of using abstract information as opposed to sensory-specific images, and of relating modalities. But, as discussed above, we are not clear on how they do this. A further complication comes from the Spelke and Meltzoff et al. finding that information presented in one modality can affect behaviour related to a <u>different</u> modality and at a <u>later</u> time. Can this be explained by notions of direct perception? It seems more compatible with the Piagetian emphasis on representational processes. Meltzoff (1981), for example, suggests that his group's research points to a mechanism involving stored, abstract representations, although he emphasises that their work does not address its precise nature. Thus, the discontinuity between Gibsonian and Piagetian interpretations is not as definitive as one would like. One thing is clear: phrasing one's questions in the terms represented by a Piaget versus Gibson framework drastically limits the mechanism possibilities which can be considered, and it does not appear to do justice to the full complexity of the phenomena involved.

### A.3 <u>Functional Perceptual-Motor Activity</u>

The infancy literature now contains many demonstrations of highly specific perceptual information being coordinated with pre-adapted behavioural components to achieve specialized ends, suggesting that the infant's world is - at some level - one of objects and events and not a flux of unrelated sensory inputs. One example is provided by the prehension and manipulation of objects, where newborns can show a form of visually elicited reaching and grasping. This is more frequent when objects are within reach and are real objects, as opposed to photographs presenting the same size retinal image, and there is evidence that the degree of finger separation is appropriate to the object's shape before it is contacted (Bower, 1972; Bower, Broughton & Moore, 1970a & b). This "neonatal reaching" soon disappears to be succeeded, after some months, by more sophisticated visually guided reaching; characteristic hand/arm movement patterns which are observed before infants become capable of this phase of reaching are significantly more common when an object of graspable size, as opposed to a larger object, is in front of the infant (Bruner, 1968; Bruner & Koslowski, 1972).

A further instance of functional perceptual-behavioural coordination is found in what appears to be defensive behaviour towards "looming" objects. In newborns of many species, including humans, symmetrical expansion of an optical contour centred in the midpoint of the visual field - in Gibsonian terms, the "transformational invariant" which specifies an object approaching the face on a hit course - results in backward movements of the head which may be accompanied by the hands being raised in front of the face (Ball & Tronick, 1971; Bower, Broughton & Moore, 1970a; Bait & Vurpi Hot, 1976).

The control of posture and locomotion has also been shown to be intimately connected with visual information. Infants who are just able to either sit or stand lose their stability in that posture when the walls and ceiling of the small room surrounding them are moved: if the room appears to be moving towards the infant, the infant falls backwards; if the room appears to be moving away from the infant, the infant falls forwards (Lee & Aronson, 197^; Butterworth & Hicks, 1977; Butterworth & Cicchetti, 1978). The appropriateness of the direction of fall suggests that this phenomenon is related to what Gibson termed "visual proprioception" - visual information specifying the body's position and movement - and the infant is compensating for a perceived (albeit 'false[1]) loss of balance, not suffering from shock evoked by the walls and ceiling suddenly moving! As with adults, the visual input responsible for these effects appears to be the pattern of optic flow in the periphery of the visual field; if just the central portion of the wall an infant is facing moves backwards or forwards, s/he does not overbalance (Pope, 1982).

In contrast with manipulatory behaviour towards inanimate objects, which was noted above, very young infants have been shown to engage in what appear to be attempts at communication when they are faced with people (Brazelton, Koslowski & Main, 1974; Trevarthen, 197*4, 1975, 1980; Trevarthen, Hubley & Sheeran, 1975). Both repeated patterns of face and body gestures and distinctive mouth and tongue movements, which Trevarthen refers to as "prespeech", have been recorded. Detailed observations do not support the view that the infant is imitating the adult in these exchanges. Indeed, although claims of learning through a mechanism such as imitation may appear preferable to innatist assumptions it should be clear from section A.2, above, that this mechanism would be neither simpler nor easier to explain.

Here, as with the examples of Section A. 2, we come across the problem of getting from demonstrations of infants[1] successes and failures on well defined tasks to agreement on the mechanisms involved. Again, the problem is aggravated by the lack of an adequately precise vocabulary in which to phrase discussion. (Useful reviews and discussion of pertinent disputes are provided by Ball and Vurplllot (1981), Bower (1982), Butterworth (1981b) and Yonas (1979)). These examples of infant functional activity certainly seem to fit Gibson's (1979) and ecological psychology's (Michaels & Carello, **1981**) predictions that infants should be "preattuned" to information which is universal for the human species and perceive objects not in terms of meaningless physical properties but in terms of the adaptive activities they "afford" the infant. Relationships in the optic array appear to specify that something is graspable or can be communicated with, that one is about to be hit in the face or is falling over. But innateness is of limited relevance to the mechanism issue.

In developing a form of empiricism called "evolutionism", ecological psychology claims that only compatible things can coexist; thus, appropriate activity implies veridical perception. But it does not follow from this that the psychological mechanisms which evolve must be those proposed by the theory of direct perception. There is an implicit and erroneous assumption that a Piagetian-type theory of the infant must be the prototype for all cognitivist positions and that they must entail a "deconstructed" view of the infant. Ecological psychology may be correct in claiming that it is impossible, in principle, for perceptual systems to evolve sensitivity to meaningless inputs, only later evolving processes which make those inputs meaningful and useful. However, in Section B it will be argued that this is a misleading characterization of the essentials of a computational cognitivist approach to perception.

### A.4 What kind of explanation do we need?

One suggestion which has been made is that theoretical progress might be achieved by assuming that infants begin with abilities based in "direct perception" and subsequently develop the type of "conceptual-representational" abilities highlighted in Piaget's theory. Butterworth (1981a), for example, adopts this approach in conceptualizing infant development as a transition from direct perception to objective knowledge. From this perspective, the type of phenomena discussed in A.2 and A.3 are considered to be "precursors of objective knowledge in perception and action". Thus, both Gibson and Piaget are right, but in connection with different abilities at different points of development.

But neither of these theories can be shown to provide a sufficient explanation of perceptual or representational phenomena. In any case, different levels or types of ability cannot be easily or clearly distinguished. For example, in A.2 it was noted that Meltzoff felt the necessity for a representational interpretation of sense-modality links in neonatal subjects, even though he was unable to specify the necessary mechanisms. Also, even if one agrees that the direct perception approach is best seen as an explanation of everyday behaviours such as walking and grasping (Turvey, Shaw, Reed & Mace, 1981), it remains difficult to draw a rigid demarcation between these and supposedly more cognitive abilities. Such 'basic' abilities undergo important changes during human development, and these may best be viewed as changes in control and organization (e.g. Mounoud & Hauert, 1982) rather than an addition of 'representational' processes. This directs our attention to the need for a framework which can deal convincingly with transformation between ontogenetically earlier and later mechanisms, and this necessitates a degree of compatibility between theoretical assumptions at the different developmental levels involved. But cognitivist and ecological styles of explanation are fundamentally and significantly different in ways (to be discussed more fully in Section B) which augur poorly for either merger or juxtaposition.

The present paper concurs with Neisser (1976 ) that we will need to postulate cognitive structures because notions of "information pickup" and "information processing" both capture too much of the truth to be ignored, and because they offer a connecting link between perception and the (so-called) "higher" mental processes. Sections B and C of this paper outline the merits of a cognitivist account of the infant based on computational concepts. As a preliminary to this, it is necessary to articulate what kind of explanation is and is not being sought and why.

The close functional relations between perceptual and behavioural systems which were discussed earlier in A.3 lead to the assumption that the 'object' for explanation is a complex perceptual-behavioural system. What we need is an integrative framework within which the very intricate and subtle relations which exist between the availability of information, its pick up and use in purposive activity can be captured and explored.

As an example of the complexity of the system, we can consider infant performance on the visual cliff - a glass sheet, divided by a central board, with a patterned surface directly beneath half the glass (the "shallow" side) and some distance below the other half (the "deep" side). It is well known that infants who are able to crawl will only go to their mothers across the shallow side. Yet recordings of infant heart-rate reveal that they can discriminate between the two halves of the cliff long before they can crawl, but their response to the deep side changes across the first year of life, and it is somewhat counterintuitive. In the youngest infants, the change in heart rate on the deep side shows the pattern characteristic of interest or heightened attention and they actually cry less than they do on the shallow side; older infants show the pattern which is generally associated with fear (Campos, Langer & Krowitz, 1970; Campos, Hiatt, Ramsay, Henderson & Svejda, 1978). This suggests that infants can, initially, pick up the available information for depth without 'recognising' its conventional meaning or significance for adaptive activity (E.J. Gibson & Rader, 1979), such recognition requires the use of that information in crawling.

This example makes two things clear. Firstly, commonly asked questions of the form "Does/doesn't the infant perceive 'X'?" are too simplistic: information may be available to a system at different levels. There is obviously some sense in which both younger and older infants could be said to perceive the depth information available at the visual cliff, but their response to that information indicates that its apparent "meaning" for them differs. Secondly, the infant cannot be said to "just see" that the deep side of the cliff "affords" the dangerous activity of falling off (c.f. Costall, 1981). Perception cannot fully be understood as an independent or modular process, and any explanatory account must handle both the availability of information for the subject and its selective use in guiding activity. Cognitive and cross-cultural studies of adults also point to the dynamic, selective aspect of perception and its intricate connection with the goal-oriented aspect of activity. (Furthermore, they illustrate relations between this and "cognitive" processes such as evocation memory and classification (e.g. Nickerson & Adams, 1979; Deregowski, 1978).)

In order to address the issue of underlying mechanisms, what is needed is an explanation which is "systematic" (Haugeland, 1978) or "systemic" (Winograd, 1977): one in which the phenomena to be explained are produced by organized interactions between components which could not achieve their functional role independently of each other. In this case, the focus of such an explanation should be on what will here be called action systems which subsume perceptual and behavioural components. The cognitivist style of explanation and, in particular, the computational metaphor provides a suitable framework for reasons which will be clarified in Section B.

Gibson made a significant transition in his theorizing when he moved from preoccupation with invariants to the later emphasis on affordances - in which he stressed that meaning was equally determined by both information and the activities of the particular animal involved. This also points to the need for a systematic explanation, but his own work does not prove very successful in this direction. As Olson puts it, he "simply treats perceptual activity and performatory activity as two discrete forms of activity in his model joined by a line of unspecified function (1970:85)", (c.f. Neisser, 1976 ; Weimer, 1977). The line can be replaced by the concept of attention, where "attending" is defined as "perceiving in relation to a task or goal, internally or externally motivated (E.J. Gibson & Rader, 1979)." But this is only a rephrasing of the problem unless the mechanisms of knowledge are specified in such a way that attention can be defined with reference to those mechanisms and not merely with reference to the range and appropriateness of environmental information dealt with. The more recent ecological psychology has made a concerted effort to expand Gibson's concepts. However, it will be suggested in Section B that it emphasises metaphors which were well suited to Gibson's initial preoccupations but which fare poorly with notions such as affordances and the systematic explanation for which they call.

In keeping with the discussion so far, it will not be claimed that Piagetian theory can provide an acceptable framework. Piaget's notion that the infant is perceptually unorganized is clearly inappropriate in view of the research which has been outlined, and his conceptualization of mechanisms is inadequate. His claim that infant sensory-motor abilities must be understood in terms of action "schemes" - internal, organizing structures which explain the functional equivalence of behaviour patterns despite inevitable differences in the movements used in their execution and the contexts in which they are applied - has proved problematic for an empirically oriented psychology. In any case, the concept was not developed in sufficient detail since it is unclear how the sensory and motor aspects of action are related. The consequence of this has been that many simplistically equate his view of action with overt behaviour or motor response (e.g. Brainerd, 1978; Kagan, 1972), or reject it as circular and mentalistic (Wilcox & Katz, 1981). These difficulties have made Piaget's attempts at a transformational account of conceptual development - which he claimed involved the interiorization of action schemes - even more impenetrable.

However, this is not to say that Piaget was on the wrong track altogether. Lack of appropriate concepts prevented his arguments from being fully effective. As Boden (1979) suggests, Piaget's theorizing might have gone much further had he had access to recent computational concepts from AI instead of relying on biological metaphors. Wilcox and Katz's (1981) claim that the deficiencies of Piagetian concepts show the notion of mechanism to be unnecessary for prediction in developmental psychology is misleading. The present paper assumes that the goal of explanation is understanding through construction of a sufficient model of the phenomena of interest, not mere prediction. This view is compatible with more mainstream ecological psychology, and Wilcox and Katz's position will not be debated further here.

A contemporary, authentically cognitivist explanation of infant abilities which conceptualizes action systems in terms of the highly suitable computational notions available from AI has yet to be developed. (But see Bower (e.g. 1979) and Mounoud & Vinter (1981) for viewpoints which support a cognitivist approach through their emphasis on the need for a (non-Piagetian) representational account of even the earliest infant abilities.) Piaget is very much the "straw man" of the developmental aspect of the direct versus indirect perception debate. It is true that ecological psychologists have produced many apposite criticisms of the recent "information processing" school of cognitive psychology, but this uses the computational metaphor in a fairly loose sense and should not be equated with AI'. Work in the latter discipline is not alluded to at all in the discussions of Shaw and Bransford (1977), and barely so by Michaels and Carello (1981) and Turvey, Shaw, Reed and Mace (1981).

Section B of this paper suggests that AI concepts can provide general principles for understanding knowledge whilst permitting the individual theorist considerable freedom in the range of particular mechanisms which can be explored. Although it could be used to explore constructivist theories, AI is not committed to the constructivist pole of the direct versus indirect debate. AI concepts may be used to model action systems without subscribing to constructivist, "information processing" or Piagetian assumptions of the degeneracy of input.

## B. AI AND A COMPUTATIONAL FRAMEWORK FOR EXPLAINING ABILITIES

### B. 1 Levels of explanation: Processes and programs

FIGURE A: LEVELS OF EXPLANATION IN ARTIFICIAL INTELLIGENCE AND TRADITIONAL PSYCHOLOGY

```
        ARTIFICIAL INTELLIGENCE:                           PSYCHOLOGY:


             FUNCTION             < what? >          DATA FROM
                 *                             EXPERIMENTS/OBSERVATIONS
                 *                                      *      \
                 *                                      *  operationaliz
                 *                                      *    /
       MECHANISM (ALGORITHM)       < how? >           THEORY
             /*>   *                          MECHANISM <--?--> FORMALISM
   operationalizes *                                    .
             T\    *                                    .
             PROGRAM                                    .
       (HIGH LEVEL LANGUAGE)                            .
                 *                                      .
         ...............                                .
                 *                                      .
         ...............                                .
                 *                                      .
            MACHINE CODE                                .
                                                        .
             COMPUTER                         BRAIN/NERVOUS SYSTEM
       (ELECTRONIC CIRCUITS)                    (NEUROPHYSIOLOGY)
```

A.I. and ecological psychology share the view that it is necessary to distinguish between different levels of explanation of a system. However, they dispute the constructs which should be employed at the psychological level and the nature of its relationship with the biological level. While sharing the view that psychological explanations of perception must postulate something internal to the perceiver, they disagree over the relevance of computational processes and the appropriateness of the computational metaphor in general. In this section, it will be suggested that the computational approach offers the greater potential for effectively conceptualizing an active perceiving and behaving system such as the infant discussed in Section A.

The main levels involved in an AI explanation are schematized in Figure A, with the analysis of comparable levels in traditional psychology shown in the adjacent section. The central goal of AI is to build models of the mechanisms which underlie the kinds of activities we call intelligent - <u>functions</u> including perceiving, behaving appropriately, learning, and so on. This aim is not unique, but what is unique to AI is its contribution to our understanding of the <u>processes</u> which such mechanisms might involve. Computation is generally defined as "rule-governed symbol manipulation" (see, e.g. Sloman, 1982a), and a major assumption of AI is that both minds and computers are physical systems capable of having and manipulating symbols (Newell, 1980; Newell & Simon, 1976). For a particular function, the data or information involved and the operations performed on it are specified and the resulting system viewed as a model or theory of the mechanisms or principles underlying ability in the area concerned: the <u>mechanism/algorithm</u> level.

The computer provides a valuable tool for exploring AI theories. It is a type of physical symbol system since locations or addresses within it can be thought of as abstract symbols which can be assigned 'meanings'; these might be expressions about objects or states-of-affairs, or operations to be carried out. Thus, computational symbols act as the building blocks for theories of intelligent systems. The notion of symbol which is central to a computational approach should not be confused with the more restrictive use of the term in, for example, Piagetian theory. This distinction is closely related to the question of 'representation' and is discussed further in Section B.2. (See Boden (1977, chapter 1) for an introduction to the computational notion of symbol, and Newell (1980) for a more technical account.)

The structure which controls what a computer does - the <u>program</u> - can be thought of as series of instructions which, when they are activated, govern which processes the computer will perform and in what order. These processes manipulate the symbolic assertions which correspond to the system's information about its (internal and external) world. The same mechanism/algorithm can often be expressed in several programming languages - although some languages may be better than others for capturing the properties of the particular mechanism which is being modelled - and the theory-program distinction is eroded to the extent that programs can themselves embody a significant number of theoretically important principles (as opposed to ad hoc implementation details). Thus, there is a symmetry between function:theory and program:theory relations. As schematized in Figure A, psychologists' observations of functions, i.e. what subjects can do, lead to the formulation of theories of how they might do it, and these can be operationalized through experiment and behavioural observation; in AI, description of a particular function leads to construction of a computational model which can be operationalized by writing and running a program. This provides us with an important additional tool for operationalizing theories, and, in an interdisciplinary approach, should be viewed as complementary to behavioural techniques of operationalization. More uniquely, the program level provides a direct test of the sufficiency of one's theory.

In contrast with the computational, process emphasis of AI, ecological psychology's approach to perception stresses its immediacy by concentrating on Gibson's (1968) radio-receiver metaphor. Environmental objects and events are said to "broadcast" information through the way they structure light, and this information can be "picked up" by a suitably "tuned" perceiver who does not need to process the information but simply "resonates" to it. In a related argument, Runeson (1977) suggests that perception may involve "smart" mechanisms which are specialized for a single function and whose operation cannot be understood in terms of interconnections between more basic components. He develops an analogy with the polar planimeter, a "smart" machine which consists of an arrangement of rods and wheels for measuring the area of planar surfaces. The planimeter attains its function with great accuracy, but it does not do so by measuring lower-order variables such as length of side and multiplying them together. The ecological approach makes much of the immobility of computers and the physical/hardware limitations of present-day machines, arguing that the problems they pose are simply inapplicable to brains.

Such scepticism about the properties of computers fails to take into account the most significant distinction between levels of explanation in AI. AI is not about the physical or hardware characteristics of computers and stresses the distinction between the program and the machine on which it runs. A program may be considered as a "virtual machine" which is implemented on a particular physical mechanism - one's computer; it could be implemented on a number of computers with different hardware properties (Sloman, 1978). As Sloman puts it: "approaching a computer with a view to finding out what it can do is as silly as it would be for a physicist to study pencil and paper with a view to finding out what they can do. One approaches a computer in order to try to make it do something. . . . . by designing a program which will make it behave intelligently one constructs a theory, expressed in that program, about the possibility *of* intelligence. The failure of the theory is your own failure, not the computer's (1978:106)". Thus, it is being claimed that a necessary condition for a system to demonstrate intelligent functions is that it be a computational or symbol manipulation system, but nothing critical hinges on the particular mechanics of today's machines.

This distinction is crucial both philosophically and psychologically because concepts from the two levels cannot be inter-translated without losing explanatory power (e.g. Boden, 197Z, 1977 Ch.14, 1981 Ch.2). The high-level language in which a program is written usually goes through several levels or layers of translation which convert it into a machine code which actually 'works' the computer. The principles of functioning at the lowest level - what goes on inside the computer - are in terms of quite distinct concepts (e.g. from electronics) from those involved in discussing how the system achieves its function in program terms. The notion that program and machine are distinct but functionally related provides a computational framework with a novel approach to questions of the mind-body relation (see Boden, especially 1972, 1981). Furthermore, it provides a unique perspective on the nature and role of behaviour which will be discussed in Section C.1.

Of great psychological significance to the conceptualization and sufficient explanation of an active - and acting - system such as the infant is the fact that these two levels involve distinct forms of explanation: the functioning of the program level is susceptible to a purposive analysis, yet the program is implemented in the computer which is a causal machine. The commonly asserted view that a computer can "only do what it's been told to do" implies a linear, step-by-step execution of externally predetermined processes which fails to capture the characteristics of even simple program structures. Programs have important active properties. Not only can they run in order to activate processes, they can also be treated as 'objects' which are called, modified or inspected by other programs. (In the high-level programming language LISP, for example, an expression can be treated as a piece of program - to be run - or as a piece of text - to be operated on by other programs.) This enables programs to control each others' operation in complex, interacting (and sometimes unpredictable) ways. "Recursive" program structures are able to call themselves in order to execute some function.

Programs which could illustrate these active, coordinative aspects with particular relevance to infant action systems will be discussed in Section C.3. At present, it is important to note that the program:machine distinction, together with these program properties enables computational explanations to handle issues of agency, control and purpose in a coherent manner unavailable to ecological psychology. They can be addressed in terms of the structure and operation of programs. The purpose of a system, for example, can be analysed in terms of programs currently in operation and the wider context of computational structures and processes within which these may be embedded. Such issues prove far less straightforward for the ecological approach.

Ecological psychology proposes that constructs concerning the organism and its environment can be phrased at three "grains of analysis" provided by psychology, biology (including the brain) and physics, but it views these as alternative levels of explanation which cannot be causally related. Thus, asking how psychological constructs affect brain/body constructs and vice-versa is illegitimate. However, whether they are to be conceptualized as alternative or interacting levels of the system, ecological psychology's concepts for the psychological level and its machinery encounter unacknowledged problems because they are too passive and too molar. Both the radio and the computer satisfy Haugeland's (1978, p.216) criteria for being systems. Both are objects whose abilities must be explained <u>systematically</u> in terms of the organized cooperative interaction of their functional components. However, although the radio receiver appears to provide an elegant metaphor for picking up information, it loses its parsimony and effectiveness as soon as a systematic explanation of the relationship between the organism's perception and "effectivities" or potential purposive behaviours is sought. Although it is a system (as, indeed, is a mousetrap) it is inadequate for explanation of the abilities with which we are concerned. As Michaels and Carello (1981) note, this metaphor fails to capture the vital link with behaviour, and a radio receiver is passive insofar as it requires a "tuner" if it is to do more than resonate to a single frequency. Runeson's (1977) polar planimeter metaphor is similarly problematic, and he emphasises that the non-perceptual aspects of the system are, in some sense, the "user" or "person with planimeter". The potential advantages of the computational approach in dealing with purposive and behavioural aspects of the active subject may be illustrated through ecological psychologists' treatment of attention. Two things are found repeatedly in their attempts to formulate a sufficient explanation, and will be discussed in turn: additional, ill-defined concepts are introduced; explanation moves from the psychological to the biological grain of analysis.

(a) Introducing additional concepts

Gibson (1979) developed the concept of attention to account for the fact that different subjects can detect different information in the same object, or one subject can detect different information in the same object on different occasions, but is this more than a redescription of the fact that selective "tuning" occurs? One can still ask what controls attention and/or how are attention and behaviour coordinated? One way in which ecological psychology deals with this is to claim that part of the psychological level of explanation must address the issue of who perceives, the question being: "How might needs, intentions, desires and feelings be manifested in perception of and action upon the world? (Michaels & Carello, 1981, p.61)." It is suggested that such psychological states are a consequence of evolution, thus their inclusion in a scientific theory is "no less scientific than talking about liver cells (Michaels & Carello, 1981, p.75)." In one sense, this general conclusion is correct. We do want to clarify psychological constructs such as purpose. However, we should not need reminding of the problems which arise when we attempt to treat them as conceptual primitives or operationalize them. Similarly, Shaw and McIntyre's (1974) wish to conceptualize the "who" of perceiving as a "knowing agent" - "the algorist" - introduces novel terminology without clarifying the nature of purpose and the control of activitiy. It highlights the theory's inability to provide a sufficient explanation of the active subject. Thus, although rejection of the notion of computation initially appears to be a simplifying and parsimonious assumption, it proves to be quite the opposite.

The computational approach is especially well-suited to operationalizing what would otherwise be considered "mentalistic" constructs (c.f. Newell, 1970), and need not postulate anything beyond the machine and the programs which control its operation. A key assumption of the approach is that of systemic interaction: our abilities are the outcome of interaction of processes. The knowledge, beliefs, intentions, and so on of any intelligent system are not primitive constructs; they are better thought of as intuitive, global labels applied to the properties of computational systems and their operation. As Minsky (1980) puts it, such notions are too coarse to support powerful theories, "they are not (like milk and sugar) objective things our theories must accept and explain; they are only first steps toward better concepts (p.439)." From the computational viewpoint, many conventional conceptual distinctions - perception and attention, perception and cognition, learning and memory, etc. - present arbitrary, reifying boundaries. At present we need to use concepts like perception or attention, but we should bear in mind that they are particular vantage points onto the operation of a complex system; they are not circumscribed 'things' and there may be considerable overlap in the processes which support them.

(b) Regression to the biological grain of analysis

The molar nature of ecological psychology's constructs is clearly illustrated by Michaels and Carello's (1981, p.71) discussion of the concept of attention. They note that since attending consists of producing behaviour which will facilitate the detection of appropriate information - looking, rubbing, listening, etc. - a person looking for a set of car keys might behave just like someone looking for a book of matches. How can they be said to be attending differently? In view of ecological psychology's claim that psychological and neurophysiological levels of description should be distinct, the answer is far from compelling in its absence of any psychological constructs: "In such a case, it can be supposed that it is a brain configuration that distinguishes the would-be driver from the would-be smoker. It would be desirable to be specific about how such brain configurations might be described; unfortunately, at this time, it can only be suggested that the term "brain configurations" means whatever it is about the brain that serves the same function as the interleaved metal plates (variable condensers) in old fashioned radios (1981, p.71)." It is, surely, a psychological configuration which should be sought here, particularly in view of the ecological psychologist's view that there is no sense in looking for causal connections between grains of analysis.

A computational approach would not deny a role to neurophysiology, and can legitimately embrace both levels of discussion. In psychological study, the nervous system may be seen as parallelling the computer in AI, and it follows from this that a neurophysiological model will be necessary to a complete account of human abilities. However, it cannot be either sufficient or simply an alternative to the program or psychological level. The key to a complete explanation of the relation between program and machine levels - and, by analogy, between psychological and neurophysiological and other bodily processes - would be discovery of the language(s) which intervene or translate between them. The complexity of this task should not be underestimated. For example, studies suggest that it is not possible to infer the function of a computer program - whether it is doing the accounts or planning the launch of a nuclear missile - from knowledge of the computer's physical states, unless both the machine code and high-level target language are known beforehand (Wilks, 1982).

The relation between the brain and human abilities is likely to prove even more complex than that between the computer and its programs, and challenges the explanatory burden which ecological psychology often appears to place on biological constructs alone. For example, one repeatedly encounters the notion that perception is unproblematic for human adults and infants because of the "biological design" and "optimal tuning" of the brain which enables information to simply "resonate" with "highly evolved neural structures" (e.g. Shaw & Bransford, 1977; Olson, Yonas, & Cooper, 1980). This confounds levels of explanation and has no genuinely explanatory status, being, at best, a plea to the inadequate criterion of "material instantiation" (Fodor, 1968). It is similar to a programmer who can get a computer to perform some function but can only explain how it does so by pointing to the existence of the working program; the machine exists, and it works, but how?

It is true, however, that AI has focussed largely on the topmost/program level of the metaphor and shown relatively little interest in how the physical as opposed to the virtual human machine is to be conceptualized. <1> The present paper disagrees with Fodor's (1980) influential "contingency" position which advocates that the mind can and should be studied independently of the the physical machine which supports it. The fact that a single virtual machine can be implemented on different physical machines extends the power of programming as an AI tool, but it does not mean that the details of the physical machine can be ignored when one's aim is to understand the human system. In any particular case, physical structure - including not only neurophysiology but also body form - will limit the processes which the program level can control. <2> This emphasis is of particular significance when the focal concern is a perceptual-behavioural system such as the infant, and is continued in subsequent sections of this paper. Sections B.3 and C.2 pay particular attention to Marr's computational approach to vision, which aims to be compatible with known neurophysiological substrates. Additionally, Section C. 1 will take up the issue of the relation between the body and internal processes, with particular reference to the theoretical status of overt behaviour.

## B.2 The centrality of representation

The structures and processes which underlie the function of a computational system can be equated with mechanisms of knowledge. Thus, a computational system is *representational* in the general sense of enabling the subject to interact with or 'know' certain objects and events in the world and not others. Newell (1980) characterizes the general relation between the symbols of a computational system and the entities to which they refer in terms of "designation": "An entity X designates an entity Y relative to a process P, if, when P takes X as input, its behaviour depends on Y (p.156)". Symbols can be input to many kinds of process, and the implications of a particular symbolization cannot be understood without knowing within what process complex it is embedded. The representational states of such a system are *intentional* in the sense of being "*directed* at or *about* objects and states of affairs in the world (Searle, 19~8(T748)", (c.f. Boden, 1972, 1977, 1981 ch.2). These senses of representation and intention are not as restricted as the usage common in developmental psychology. Nor do they mesh easily with ecological psychology's perspective on representation which was discussed in Section A. The computational senses of the terms are best clarified by addressing three related questions which are frequently posed by those concerned with infant abilities:

(a) Do we need notions of symbols, representation and intentionality to explain our infant subjects' abilities?

This question, which seems particularly pertinent when the subject concerned is the very young infant, is best answered by considering the assumptions underlying it. Both ecological psychologists and developmentalists tend to assume the focal terms must mean something akin to their usage in Piagetian theory. For Piaget (1953, 1955), all three properties are emphasised as goals of infant development, and it is one of the attractions of the Gibsonian alternative that it attempts a parsimonious explanation of the newly discovered infant abilities which does not identify them with these adult-like functions. But adopting a computational approach to the mechanisms of infant knowledge does not commit one to the view that they possess, from the outset, the kinds of abilities which Piaget argues they lack until the second year. More fundamentally, from a computational viewpoint it is not at all clear what such a claim would amount to.

For Piaget, the hallmarks of representation are evocation and deduction. Both properties are demonstrated through success on the invisible displacement task which is assumed to entail "an image of absent objects and their displacements (Piaget, 1955:4)" since "no perceived sign commands belief in permanency,........the vanished object is displaced according to an itinerary which the subject may deduce but not perceive (Piaget, 1955:85)." A symbol, for Piaget, is a material event - an image, imitative gesture, word - which is differentiated from and takes the place of some other event which is known but not present. The signal-based behaviour of the sensory-motor infant is similar to the reaction towards the thing signified, but symbols can be associated with other behaviours not necessarily invoked by their referents. Thus, even the 9-month-old infant who can retrieve a hidden toy still lacks representation and symbolism for Piaget because the cover acts as a signal or index of the presence of the hidden toy and is an undifferentiated aspect of the total context, Piagetian notions of intentionality, representation and symbolism are intimately connected since intentional behaviour involves having a differentiated goal which is approached via interchangeable means in a foresightful way.

But although Piaget[1]s observations provide an indication of interesting performance phenomena (functions), his theory fails to precisely define the mechanism or class of mechanisms which make them possible. Piagetian-style notions of representation, symbol and intention are best seen in terms of the notion of systemic interaction which was introduced in B.1, as emergent properties of interacting processes. They are not appropriate conceptual primitives. In the case of symbols, for example, Furth (1969) stresses that the relation between a differentiated symbol and its referent - the known thing, as opposed to the thing itself - presupposes knowing and does not explain it. The more fundamental computational concepts of symbol and representation should be seen as essential to an explanation of the process of knowing.

(b) Could a computational system have symbols, representation or intentionality in the same way that infants do?

The computational definition of these notions is internal to the concept of a physical symbol system, they are not intended to, nor do they map directly onto the developmentalists usage. But this does not mean that they cannot be used to model systems which are equivalent to those of particular developmental levels. The properties of the computational concepts enable AI to model a wide class of intelligent behaving[1] systems; these could include not only functions characteristic of the subject who is representational-symbolic in the "strict" Piagetian sense, but also those functions characterizing the younger perceptual-behavioural infant. From a computational viewpoint, both types of subject experience an organized, meaningful world. <3> Thus, AI may aim at an account of the structures and processes involved in perceiving and interacting with an object, say, or those underlying the understanding and use of language. Both can be modelled in terms of the manipulation of symbolic assertions, and both involve representation in the computational sense. Examples will be provided in the following part of this section of some of the types of symbol structure manipulations central to AI explorations of perception. An example of a computational model of the structures and processes underlying language use is provided by Winograd's (1972) early system SHRDLU. This "understands" natural language, as demonstrated by its ability to communicate about a miniature but changeable "blocks world". One of the most important aspects of the system's design is that it is not based solely on the manipulation of symbol structures representing knowledge about words. These interact with other symbol structures which represent the system's perception (essentially a symbolization of the current state of the blocks world), knowledge about the blocks world (e.g. what kinds of objects exist in it and what their properties are) and knowledge of its own actions (structures which can be accessed and manipulated under certain circumstances to check, for example, that a certain action is possible). With these components, Winograd is able to model the comprehension and answering of questions such as "how many blocks are not in the box?" and execution of commands such as "stack up both of the red blocks and either a green cube or a pyramid." <4>

Correspondingly, the computational metaphor can provide the conceptual tools for an analysis of purpose and intention. Again, there may be different types of system which share a fundamental property, in this case that of directedness at objects. Thus, Bower, Broughton and Moore (1970b) and Trevarthen's (197^0 insistance on ascribing intention to the neonate - on the basis of distress being exhibited when activities such as prehension and communicative interaction are disrupted - may not be incompatible with Piaget's argument that it is a much later, developmental addition to the infant's functions. They are possibly talking about different types of intentional system, both of which could usefully be conceptualized as controlled by programs. The computational problem is to specify what type of symbol structures and permissible operations or processes are involved in each case.

The range of behaviour which is purposive (and intentional in the computational sense) cannot be restricted to the Piagetian usage. Boden (1983) discusses the way in which animals may often produce complex chains of behaviour without the program which controls this behaviour including a representation of the outcome of the behaviour chain. A bird may show a sequence of activities such as nest building, food collection, cleaning debris from the nest and feeding the young without any part of its internal representation of the world being a "goal" of raising healthy young to maturity, i.e. a symbol structure or "explicit representation" corresponding to healthy, grown-up chicks which controls the execution, sequencing and termination of activities. Instead, individual behavioural components appear to be controlled by programs linking perceptual with motor activities. As the bird's behaviour alters the environment, new information will result, activating further programs of this type, etc. The control mechanisms of the earliest infant behaviour may well fit this type of model and only later incorporate explicit representations of potential behavioural outcomes.

For the computational approach, therefore, the appropriate question is not whether or not computers or infants do or don't have symbols, representations, or intentions in the senses in which psychologists have often chosen to define these terms. It is to clarify what type of system could exhibit the functions with whose presence these properties have been associated. This is just one aspect of the broader problem of detailing the properties of different types of knowledge system and the relations between them.

(c) What is to be gained by adopting these computational definitions in preference to the more familiar ones which we understand?

As suggested in the 'answers' to (a) and (b), we do not really understand these familiar uses that well: we do not always know what mechanisms they involve; we do not always use them consistently. A framework based on computational concepts provides the opportunity for a more exploratory approach to infant development, one which stops asking "do infants have X?", and asks instead "what is X?". A computational approach does not see simple answers to global questions about the nature of and relation between perception, concepts, representation, and so on. It encourages us, rather, to ask finer grained questions about what we really mean when we use such terms, and it provides an internally consistent language within which discussion can be phrased.

If our observations of infant performance convince us that there is a developmental progression from system "A" to system "B", we are unlikely to understand how infants make that transition unless we have a good grasp on what "A" is. Once we have achieved that, we can proceed to ask how "A" would have to be restructured to produce the performance phenomena which are associated with "B", and what developmental processes could produce this?

B.3 Constraints: generativity without degeneracy

So far, the general nature of a computational theory has been outlined. The notion of representation as the sum of the subject's computational structures for interacting with the world was emphasised in order to show that this framework is potentially as suited to explaining the infant abilities described in Section A as it is to the language using, thinking adult. Of central importance to a more detailed analysis are the properties of AI vision programs. Two aspects of the ecological critique must be addressed, and they need to be clearly distinguished. Firstly, is the pick up of information usefully conceptualized as involving computation on or inference from lower order variables, i.e. is it subject to "psychologically meaningful decomposition" (Ullman, 1980)? Secondly, is the information necessary for veridical perception present in the input, or must it be supplemented? It will be suggested here that a computational basis for perception provides us with useful concepts which form a sound bridge to an exploration of its role in action. However, the frequent assumption that an AI framework must presuppose degeneracy of perceptual information (e.g. Russell, 1981; Shaw & Bransford, 1977), thus being totally incompatible with Gibson's theory but compatible with constructivist positions is unfounded.

Overviews of the cognitivist approach have, it is true, tended to justify its relevance from the ambiguity of information. Pylyshyn (1980), for example, presents several versions of the view that "all physical events are intrinsically ambiguous, in the sense that they can be seen in very many different ways (p. 121)." But it is misleading to take Pylyshyn's "all" literally. It depends on what is meant by a "physical event", and ambiguity - or, more accurately, the possibility of multiple interpretations - is a question of level of analysis. Essentially, perception presents us with a version of the "equivalence class problem" (e.g. Mackworth, 1983): how can a set of things which are different under one description be identical under another? The unifying assumption of cognitivist frameworks is that the subject's symbolization or representation of the environment determines both what can be known and how it can be known, and AI accounts of perception emphasise that different representations of the same input are possible. Research in AI, like ecological psychology, has exploited the fact that objects which may be described in several ways while considered in isolation are unambiguous when the relationships between them are taken into account. The term "ambiguous" carries unfortunate negative connotations, but need not be associated with either vagueness or degeneracy of input to the system. A more appropriate and positive emphasis is on the fact that the possibility of multiple interpretation of elements at one level of representation, though not others, underlies the ability of computational systems to be generative, the general issue being how a finite system can cope with indefinitely many inputs. For example, the programs of a sufficient system for recognising triangles could only consist of a restricted, finite number of procedures, but we would want it to be capable of recognising an indefinitely large number of triangles of differing sizes, colours, and so on.

The basic notion of computational generativity can be simply outlined by considering early work in "scene-analysis". This introduces two further key notions of continued significance to vision programs: domains of structure and constraint analysis. The aim of scene-analysis is to produce computational models which can take an input in one structural domain - that of two-dimensional line drawings - and produce a description or interpretation of it in another domain - that of three-dimensional object structures or scenes. Any domain can be characterized in terms of its elementary units or "primitives" and the types of relationships which exist between them. Thus, in the picture domain we have simple lines combining into junctions, and line junctions depict vertices in the scene domain which are combined into three-dimensional polyhedra. Clowes (1971) and Huffman (1971) independently had the insight that it was useful to conceptualize the task of scene-analysis as similar to language understanding, following Chomsky's emphasis on the rule-governed yet infinitely creative or generative nature of language. Both developed line-labelling approaches in which individual lines could be labelled as concave ("-"), convex ("+"), or occluding ( ">" or "<" depending upon which object surface is visible from the camera angle) edges. The aim was to define a system which could label the lines of any picture and generate a unique and consistent description in the object domain for all possible or physically realizable objects: one in which every line retained the same label along its length and all lines were labelled. Here, there is an explicit parallel with the linguist's notion that the rules of a grammar should generate all and only those sentences which are deemed acceptable by native speakers of the language concerned. Correspondingly, a successful line-labelling system should fail to produce a consistent reading for anomalous or paradoxical objects, such as the devil's pitchfork or Penrose triangle, in a manner similar to a grammar failing to generate nonsense sentences. (This aim was achieved in some cases such as the devil's pitchfork.)

The fact that pictures are ambiguous, in the sense that any single line can be interpreted in several ways, may be seen in a positive light. It does not mean that the picture input is degenerate, but provides an important contribution to the flexibility of our perceptual systems, just as the ambiguity of individual words contributes to the flexibility of language. However, it is clear that there is a real problem if, working at the line level of analysis and, starting with a single line, one systematically works through all possible labellings. Since a line can have one of four possible labels, each junction type can have many possible labellings (four to the power of the number of lines making up the junction), and the search space of possibilities becomes vast as the number of junctions in a picture increases. The difficulty has successfully been overcome through a style of parallel computation called <u>constraint analysis</u>. This exploits the fact that there are relations, patterns, or "constraints" on the way entities in the physical world interact with each other. In the broadest sense, a constraint can be defined as "a relation between a set of qualitative or quantitative descriptions (Steels, 1982:75)", For example, a = b & c, is a constraint between a, b and c. Computationally, a constraint can be viewed as a set of program procedures that can complete a relation from partial descriptions and maintain it through changes to any of its parts. The processes which resymbolize input into a particular type of structural description consistent with constraints can be equated with the system's implicit knowledge of that constraint. In devising a system for recognising trihedral scenes, for example, there are two major constraints which can be exploited. Firstly, although there are potentially an infinite number of pictures depicting trihedral objects, these can only ever contain four types of junction (provided there are no shadows or cracks in the scene). Secondly, there are only 16 possible labellings of these junction types which can correspond to physical vertices in the trihedral world, a very significant reduction on the 208 types derived from considering all four possible edge labels at the four junction types. <5> Waltz (1975) discovered that an algorithm which implicitly embodies such constraints - it eliminates all impossible labellings for a line through a pairwise comparison procedure and extension or "propogation" of the results to neighbouring vertices - makes it unnecessary to use a search procedure because the number of possible labellings for each vertex is reduced to one. Thus, at a junction the interpretation of a particular (otherwise ambiguous) line is constrained, in much the same way as the value of an algebraic variable is constrained within an equation. <6>

Despite the influence of modern linguistics on computational approaches to perception, two significant differences between the frameworks must be noted. The first involves levels of explanation and linguistics' concern to specify the subject's "competence". The programs of AI vision systems attempt to model a mechanism which can actually perform the function of interest. A linguistic grammar, on the other hand, provides an abstract characterization of the sentences which constitute its focus of analysis. Under the psychological theory level, Figure A distinguishes between mechanism - structures and processes which model <u>how</u> an observed performance is produced - and formalism - abstract description of <u>what</u> is produced; these are joined by the symbol "<--?-->" to indicate that their relationship is unclear (Marr, whose work is discused below, develops a particular viewpoint on this relationship). A formal model such as a linguistic grammar does not "generate" sentences in the sense of producing them, it merely assigns descriptions to them. The descriptions appropriate to sentences are theirs quite independently of the mechanisms by which the sentences are produced (Winograd, 1977). The distinction between a scene analysis mechanism based on an exhaustive tree-search of all possibilities and one employing constraint analysis illustrates this AI concern, a concern which is relevant to psychology's focus on performance mechanisms. The second point is that Chomsky's original approach stressed that the input available to the language learner is fragmentary and degenerate, replete with incomplete and ungrammatical utterances, thus there must be a strong biological basis for grammar acquisition since it could not be inferred from the input. A computational approach to perception need not deny a biological basis claim, but it would challenge the prior degeneracy assumption.

In constraint analysis, the manner in which the structure of the system governs processing cannot be equated simply with either transduction or construction (based on either innate or acquired knowledge). It does not just involve changing the input into another physical form, nor does it rely on predominantly top-down processing in which the actual input to the system plays the role of providing cues which confirm hypotheses derived from knowledge outside the perceptual process itself (e.g. Neisser's (1967) "analysis by synthesis"). <7> The several domains of structure are best viewed as levels-of-processing. The structures in each domain involve different ways of parsing the input, each coding some relationships implicitly, some explicitly and some not at all. In scene analysis, an input in the form of an intensity array may be successively symbolized in domains whose primitives are lines, junctions, vertices and objects, but the symbolization at each level continues to designate the entity responsible for the input. This remains true no matter how many processes operate on the initial symbolization, since the symbolization or designation relationship is transitive (Newell, 1980). This lacks the "embellishment" or "adumbration" of data so criticised by ecological psychology. As Ullman (1980) puts it, processing does not create information but extracts it, integrates it, and makes it explicit and usable. (The "usable" aspect is particularly important and will be discussed further in Section B.4.) Not everyone agrees that this is different from constructing information (e.g. see Sloman, 1980), but it is not construction in the traditional sense originally challenged by Gibson since it does not imply awareness of meaningless sense-data which are subsequently supplemented with non-perceptual information in the form of ideas, memories or concepts to yield a meaningful perceptual experience.

Programs of the sort exemplified by scene-analysis might seem poor contenders for modelling the infant, relying, as they do, on processes which are restricted to implicitly embodying knowledge of possible vertices and their appearances. But more recent work in AI vision gets considerably closer to a more general model of human vision while expanding on rather than replacing the general concepts introduced above.

Marr's influential work on vision consolidates and extends previous AI work (see Marr (1980, 1982) for overviews), and will prove useful to the conceptualization of infant abilities in Section C. It is thoroughly opposed to traditional constructivist approaches while insisting on the necessity for a computational framework. Marr explicitly agrees with Gibson's view that the information available to the subject can specify real-world properties: perceiving involves recovering properties of the world from sensory input, not constructing them. However, he believes that Gibson seriously underestimated the difficulty of recovering invariant properties such as image surfaces. Marr emphasises that this really is a problem in information-processing, by which he means that it involves moving between two or more domains of structure which represent the input in different ways. Vision begins with an "image" of the world - a type of representation whose description of the world provides explicit information about intensity values - and its task is to produce from this an efficient symbolic description which is suitable for the subject's uses and purposes.

The significance of appropriate descriptions is captured by Marr's "principle of explicit naming". This is comparable to Newell's notion of "designation", which was discussed above, and has the same representational implications: once a collection of data has been given a "name" it can be manipulated as an entity in its own right, have properties assigned to it and be referred to by other external structures and processes. Thus, certain processes may be readily applicable at one representational level, but applicable only with difficulty (or not at all) at another which contains the same information in a more implicit form.

Marr argues that an effective information processing system should have a modular organization in which a large computation is split up into relatively independent parts. Modular design of a complex process means that a small change in one place does not have wide-ranging consequences elsewhere; this makes the process much easier to improve (by a human designer or natural evolution). This helps the system to follow "the principle of least commitment", extracting all the useful information in the image without prejudice to higher-level domains of structure, thus avoiding doing things which might later need to be undone. These assumptions lead to the proposal that early levels of the perceptual process need to be "data driven", with analysis moving from the bottom up, uninfluenced by "downward flowing" information which might be equated with hypotheses or expectations concerning what is present in the image.

Support for these ideas, including operationalization in the form of working programs, is developed through Marr's progressive elaboration of a vision system to provide a description of the world which is suitable for the efficient recognition of three-dimensional objects. The task of extracting all the information in an image is unlikely to be achievable in a single step, and Marr proposes several domains of structure or "intermediate representations" which describe the input in distinct yet related ways. At each of these representational levels, Marr's analysis distinguishes between three related levels of explanation of an information-processing system. The topmost level of "computational theory" aims at a specification of the information-processing task facing the system in terms of <u>what</u> it has to compute and <u>why</u> this is appropriate in terms of the principles of interaction or physical constraints in the physical domain concerned. This provides an account of the properties of the problem and principles of solution which is universal and independent of how a particular system might actually solve it, as do the competence models of linguistics which were noted above. Here, it is interesting to note that the aims at the topmost level - which Marr sees as the most important and the most ignored by AI, psychology and neurophysiology - show a commonality with ecological psychology's commitment to analysing what invariant structures of stimulation exist and describing them in terms of an appropriate geometry (specifying the class of transformations over which particular structures remain constant). However, any comparison breaks down at the second level. That concerns implementation or <u>how</u> the task can be achieved in terms of the construction of a description in a specific representational format by processes ("algorithms") making use of the relevant constraints. In contrast with the ecological approach, this involves more than mapping between each invariant and the neural structures underlying its detection. At the third and final level, the "physical realization" of these computations is sought. The main aspects of Marr's model are summarized below.

The aim of the first level of processing is considered to be making explicit important types of intensity change in the image because these provide information about discontinuities in the image (such as shadows, object boundaries, etc.). It results in a primitive but rich description of the image called "the primal sketch". The first stage in recovering this information interprets the continuous or analogue representation which is the "image" into a discrete symbolic representation in terms of intensity changes. The processes involved here are conceptualized in terms of the detection of "zero-crossings", locations where the value of a function passes from positive to negative; these are detected by filtering the image through a set of differential operators which are, essentially, local calculations applied at each location in the image and making use of the intensity there and in the immediate vicinity. The mathematical properties of the operators chosen need not be discussed in detail; they may be simply understood by noting that in differentiation an intensity change in the image will give rise to a peak or trough in the first spatial derivative and a zero-crossing in the second. It should be noted that that Marr was able to map these processes onto known neurophysiological mechanisms (involving cells with circular receptive fields), thus furthering his aim of explaining neurophysiological function in information processing terms and providing independent justification for his approach.

The outputs of zero-crossing detectors are combined to provide a new description in a vocabulary of types of intensity change, e.g. "edge segment", "bar" and "blob". The processes which achieve this grouping make use of a constraint on the way the world works: most intensity changes in an image are produced because of the spatial localization of physical phenomena. Thus, grouping processes are designed to implicitly embody a "spatial coincidence" assumption in that they take relations between the outputs of zero-crossing detectors to indicate the presence of an intensity change due to a single physical phenomenon such as a change in reflectance, illumination, depth or surface orientation. This emphasis is crucial; the operation of the grouping process is the system's implicit knowledge of spatial coincidence. The intensity changes which are recovered from the image are hierarchically grouped into units called "place tokens" which associate their properties such as length, width and brightness with positions in the image. The full primal sketch uses virtual lines to make explicit the local geometry of these place tokens.

This primal sketch, and not the image, is the input to further processing. The type of information it contains can be used by a wide variety of processes whose outputs result in the production of a description known as "the 2.5-D sketch". This explicitly represents information about surface discontinuity and the shape and orientation of visible surfaces with respect to the viewer. It can be produced without the need for any a priori knowledge about the shapes of viewed objects because it is restricted to the orientation and discontinuities in depth of patches of visible surface, arbitrary points on the image as opposed to the depth *or* orientation associated with particular objects which would be made explicit by trying to segment the image into "objects", "regions" and so on. (Systems which attempt to achieve the latter form of description at an early stage of processing have tended to be very unsuccessful, even with "added knowledge".) Marr attributed the success of this level of processing in his system to the fact that image intensities are principally determined by illumination, object reflectance properties, visible surface shapes and viewpoint. Thus, they implicitly embody information about local surface dispositions (e.g. orientation, depth) and local surface materials (e.g. texture, colour) which can be recovered via processes operating on the primal sketch.

The processes which relate these two descriptions may include those concerned with stereopsis (the objective mapping of corresponding positions in two images becomes possible at the primal sketch level, but cannot be achieved at the image/intensity array level), recovery of surface shape from motion (involving consecutive matching), texture and shading. Each process embodies its own constraints. For example, one of the "assumptions" embodied by stereopsis is "uniqueness": physical items can only be in one place at a time, thus each item from each image can only be assigned a single disparity value at most. The complete set of constraints need not be detailed here, but it is interesting to note that the use of texture proves far harder to analyse than Gibson believed.

The viewer-centred, local information of the 2.5-D sketch makes it unsuited for recognising a particular three-dimensional object as "the same as one seen earlier". The final, "3-D model description" has to make explicit readily identifiable geometric features of the object's overall shape, and it must be "object-centred", i.e. it must specify these features and their dispositions relative to the shape itself. Marr argues for the utility of a representation based on volumetric as opposed to surface properties. The geometric properties of the volume occupied by a shape can be used to construct a description of the relative spatial arrangement of the components of the object involved. The notion of a "generalized cone" - a surface which is created when a cross-section is moved along a given straight axis, the cross-section varying smoothly in size but not shape - provides a suitable element or primitive for such a description; having an axis as an integral part of their structure, generalized cones can be used to define a local coordinate system.

Complex shapes can be described in terms of what Marr calls "3-D models". For example, the complete 3-D model description for the human body consists of a hierarchy of 3-D models, each comprising a model axis for the shape involved together with its component axes. This permits the body to be described at differing levels of detail. The most general 3-D model in the hierarchy is a model axis corresponding to the gross size and orientation of the body, together with six component axes (torso, head, arms and legs). A more detailed description is achieved by developing 3-D models for each of these components, and so on. The axes of each 3-D model provide a local coordinate system; this preserves modularity, enabling the shape of parts (e.g. the hand) to be described irrespective of their relation to the body.

It is important to note that Marr's model of early vision achieves a generality beyond the recognition of solid three dimensional objects. For example, it provides interesting links with why and how we find it easy to perceive line-drawings, which are thought to contain information similar to that usually made explicit in the primal sketch, or recognize stick figures such as pipe-cleaner animals, whose component sticks may correspond to the axes of generalized cones which approximate the shapes of parts of the object.

The 3-D model description may seem much farther removed from the image than either the primal or 2.5-D sketches. However, Marr was able to establish that if a physical surface is a generalized cone, there are specific relationships between the points on its surface which give rise to occluding or bounding contours in the image and inflections (convexities and concavities) of that contour. Thus, data-driven processes which make use of these relationships constitute one method of recovering the natural axes of a shape from its image, providing the basis for the 3-D model description. It is only in the process of using this description for explicit recognition that information or knowledge concerning particular objects is brought to bear (in the form of "stored" shape descriptions). This will not be discussed further here because it is the general issue of how the information which has been "made usable" might be used which is of primary concern in developing a potential computational basis for a systematic explanation of action systems.

### B.4 Using information

Section A of this paper emphasized the complexity of the relationship between perception and purposive activity and the need for a systematic explanation of action systems. Marr's stress on the "modularity" of early vision does not contradict this when one recalls the distinction which was made between the availability of information and its use by the subject. Marr concentrates on the first of these, but he explicitly equates his model's modules or domains of structure with pre-attentive vision. Thus, it is consonant with his position to see early vision as automatically providing certain types of information which may or may not be referred to or used by extra-perceptual processes. The latter would be implicated in the control of vision, i.e. when the information it provides is sought and used rather than what information it can deliver, and this opens the door to a computational account of attention and of perception's role in action systems. Unfortunately, it is true that AI vision work has shown relatively little interest in these issues. However, Marr (1977) himself stressed the significance of what he termed "the reference window problem": how might the description of an object provide the potential for its use in quite different contexts, e.g. how can a newspaper be perceived as suitable for reading, fire-lighting or squashing irritating insects against a brittle surface. In suggesting a solution to this problem, it will be shown that ecological psychology's notion of "affordances" can be assimilated to a computational framework.

Neuropsychological studies of brain damage patients support the view that there are different and dissociable types of "use-knowledge". For example, left parietal lobe damage is associated with selective loss of the ability to provide or recognize the name, use, purpose and properties of an object, even though its shape can be correctly perceived. This suggests the existence of a particular type of representation which includes an explicit description of the activities an object might be useful for: a hierarchical (possibly linguistic) description which might be called an "integrated object description" and which is additional to the perception of the object. Evidence that a different type of use-knowledge is also represented comes from subjects who have lost the ability to name an object or its properties (e.g. a comb) but are quite able to "recognize" it in the sense of being able to engage in appropriate activities with it (i.e. combing their hair). Those who believe that computational approaches are inherently constructivist would tend to assume that a computational account of the use of information provided by early visual analysis must involve the first of these types of representation. But this is not the case, both types could be analysed computationally.

One need not view different domains of structure as a hierarchy of increasingly meaningful representations which progressively approximate to the 'real object[1] in the world, preceding a decision on what behaviour can and should be initiated. It is important to distinguish two computational positions, the second of which is compatible with ecological psychology's emphasis on affordances. On the one hand, we could argue that the perceptual system involves several domains of processing but the early levels involve structural descriptions which only make explicit information about, say, geometric properties, and this is of no use to human purposes. This view would go on to stress the necessity of additional structural descriptions which are appropriate to the use to which an object can be put, e.g. a description of an object which makes explicit that it is a chair and contains elements, perhaps in a "frame", of essential/universal and optional characteristics of chairs. The role of perception in an activity such as sitting could then be viewed as activating processes which attempt to match environmental objects with this abstract structural description of 'chairness'. (See Oatley (197<§) for this type of viewpoint).

The present paper doubts the necessity of such high level "integrated object descriptions" for guiding activities such as those described in Section A of this paper. An alternative argument is that all levels of perception are providing potentially useful information. Sloman (1982b) introduces an idea along these lines; Marr (1982) suggests each species will be capable of a set of description suited to its particular purposes, although some descriptions may be dedicated to specialist purposes while others are implicated in more general abilities. If we consider "simple" creatures such as the frog (Lettvin, Maturana, McCulloeh & Pitts, 1953) or the housefly (Reichardt & Poggio, 1979) it is clear that their representations of "a fly" differ markedly from ours. The frog's "net convexity detectors" explicitly provide only the information that there is a blob of a certain size moving across the visual field. This type of description is akin to Marr[f]s low level "primal sketch", but, in the context of the frog's feeding program, it suffices for guiding the tongue to the relevant location and type of object, even though it would be useless in a world of very slowly creeping insects. Similarly, the success of the housefly's flight to intercept others of its kind uses a description which makes explicit the direction and angular velocity of a patch in the visual field. If the fly's "assumption" that it is approaching a fellow-fly is ill founded (the description just might refer to a rapidly moving elephant a long way off) it will make the discrimination simply by failing to reach the target, never "knowing" why it failed. Humans may build and use such low-level descriptions, although they tend to play a less exclusive role in activity. The important general point is that the "salience" of a particular type of information is not a function of its level of processing but of what the subject is able and trying to do.

In this alternative formulation, the control of any particular activity could be viewed in terms of a program, part of which controls processes which are activated by or call upon a specific description or set *of* descriptions. For example, a program controlling an activity such as sitting might activate processes looking for a constellation of descriptions relating to size, surface orientation and material. The "meaning" of an object or event is not explicitly represented but is a product of the system's causal embedding in the world and of the structures and processes which mediate its appropriate behaviour. In Section C, the implementation of such a system will be discussed in further detail. We can envisage an action system with a generative control structure in (at least) two senses: governing the translation between levels of perceptual processing and governing the relation between information (and combinations of information) at these levels and the subject's behaviour. This provides for a parsimonious model with substantial parallelism, in contrast with the more redundant and linear model implied by the ecological approach. Within the perceptual component, the same representation may be used by different processes to arrive at new descriptions, as exemplified by the multiple processes which can refer to the information made explicit in Marr's primal sketch. The modularity of descriptions permits processes external to the perceptual component to be activated by or call upon different permutations of the set of available descriptions, without each type of information needing to be re-computed each time it enters into a new combination. The ecological model implies a cumbersome system in which every invariant and combination of invariants would have to be coded independently, and the relation with extra-perceptual activity would remain obscure.

These ideas are not incompatible with ecological psychology's stress on "affordances": Gibson's (1979) suggestion that infants should begin by perceiving objects in terms of "invariant combinations of variables" useful for activity rather than initially discriminating their qualities and then learning what combinations of qualities specify actual objects; Michaels and Carello's (1981) claim that "humans do not perceive chairs, pencils and doughnuts; they perceive places to sit, objects with which to write and things to eat (p.42)." However, this is not the only mode of perception available to human adults. We can also perceive individual "qualities" or a chair as a chair, and this flexibility is relevant to establishing the advantages of a computational explanatory framework.

Michaels and Carello (1981) stress that there can be no rapprochement between ecological and cognitivist psychologies - viewing the former as concerned with the "what" of information and the latter with the "how" - due to the ecological camp's total rejection of degeneracy. But this paper has been at pains to show that a computational approach need not be equated with traditional constructivism together with its assumptions of the degeneracy of data. Generativity, domains of structure, and the exploitation of physical constraints underlie a flexible and parsimonious system, not one struggling to make up for the deficiencies of the input.

Referring back to the radio-receiver and planimeter analogies, it can be anticipated that ecological psychology would still deny the utility of a computational approach to "how" information is acquired. Perception is supposed to be about registering variables of high information value to the subject, not discrete variables *or* properties basic to physics. Gibson (1979) stressed that it is easier to perceive an invariant combination of variables than to perceive the variables separately because the former is meaningful but the latter is not. But the computational account presented here is not claiming a progression from meaningless to meaningful experience. Higher-order invariants or higher-level representations may be derived from lower ones without our being aware of the processes or computations involved. In any case, as Rock (1977) emphasises, a "proximal" mode of perception is possible. For example, adults are able to perceive the size of an object either in a relational "constancy" mode or in terms of its extensity in the "proximal" mode; as one tracks a moving finger, one can become aware of the (stationary) background appearing to move in the opposite direction; thus, such information is available even if we rarely attend to it. The claim that higher-order invariants or relationships are detected by "smart" mechanisms would provide no place for such perception. Rock notes the similarity between these two modes and Gibson's (1950) distinction between the "visual field" and the "visual world", although Rock prefers to view proximal mode experiences as perceptions not sensations. However, he fails to find much of a role for the proximal mode, noting its subjective relation to observation conditions, lack of obvious relevance to behaviour, difficulty of description and so on. But the computational emphasis of this paper would question this general assessment by considering all representational levels as providing potentially usable information: that is what makes them meaningful and brings them into the domain of perception as opposed to sensation. It is important to recall the relation between available information, its salience and the subject's activities. The proximal mode may correspond to lower levels of processing, as in Marr's primal or 2.5-D sketches. Humphrey's (1974) studies of functional dissociation in primate vision are pertinent here because they focus on a subject which, he believed, could be characterized in terms of "visual field" perception with an absence of object constancy and recognition. His monkey subject still showed the ability to locomote, avoid obstacles and make a range of object discriminations which could serve as the basis for food choices. Besides providing a role for lower levels of processing, the computational account does not arbitrarily partition out the relevance of even higher levels. While adults can certainly represent an object such as a chair in what might be called the "affordance mode", as discussed above, they can also represent it in ways suited for recognition as "a chair" or "the same chair", taxonomic classification, etc. It is possible that the type of "integrated object description" noted earlier is implicated in such classificatory activities and develops as a consequence of cultural pressures for such activities.

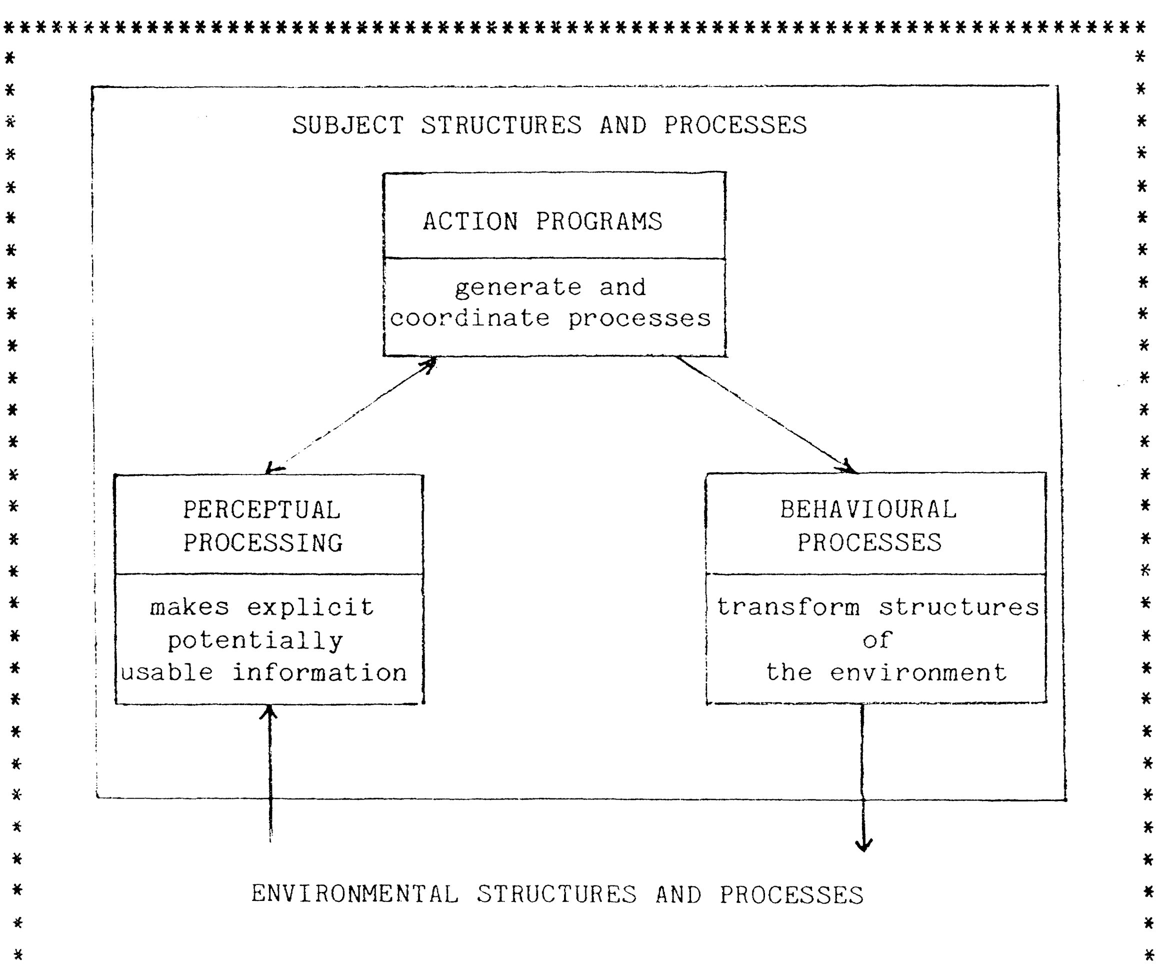
A final ecological criticism asks: how could such a "processing" as opposed to a "direct perception" system have evolved? This loses its force in the face of the computational assumption that there is no meaningless, non-informative basis for perceptual processing. (In fact, Simon's (1969) analysis of the mathematics of evolutionary processes supports the view that hierarchical systems which are organized in terms of components or parts at various levels will be the most likely to survive.) Thus, a computational approach can provide an alternative to ecological and traditional psychologies. Viewing perception as <u>either</u> direct <u>or</u> mediated involves a simplistic dichotomy. Perception <u>may be</u> direct in the sense of (potentially) providing veridical information about the world, but it may be indirect in the sense of involving processing and not being independent of other psychological systems.

## C. A SYSTEMATIC EXPLANATION OF INFANT ACTION

Section A of this paper stressed the lack of an adequately rich language for discussing the mechanisms underlying recently discovered infant abilities, and the aim of the present section is to explore the advantages of using the new concepts and ideas which were introduced in Section B. Section B contrasted styles of explanation in AI and ecological psychology, and established that a computational framework does not commit one to the type of cognitivism characterized by Piaget's theory. Computational cognitivism need not entail a deconstructed view of the infant, nor does it necessarily imply that young infants possess representational processes such as those implicated in evocation memory. In this section, the alternative, AI-influenced approach which was outlined is applied to the particular problem of explaining infant action systems in the context of three issues which featured in Section A. Section C.1 addresses the general nature of a systematic explanation of action systems within this computational framework. Sections C.2 and C.3 go on to consider how it can improve our understanding of the abstractness of infant perception and functional perceptual-behavioural relations respectively.

### C.1 The infant as a computational system

FIGURE B: KEY COMPONENTS OF THE INFANT ACTION SYSTEM



In relating the general computational framework of Section B to a systematic account of infant action, the main theme to be developed here is that action cannot be treated as a conceptual primitive. Action is best treated as an organizing concept which subsumes not only perceptual and behavioural components of the infant system, as was suggested in Section A, but also program and machine levels of explanation, which were introduced in Section B. This produces a unique perspective on the nature of behaviour, which will no longer be equated simplistically with action, and on the subject-environment relationship. Figure B schematizes the main aspects of a computational account of infant action based on the ideas which were introduced and developed through the course of Section B.

At the most general level, infant action systems can be thought of in terms of types of action. (Various classifications of infant activities have been suggested, though their relative merits will not be considered here. Trevarthen , for example, proposes a four-fold classification: exploratory/performatory, communicative, self-absorbed and cooperative.) Each action type can be envisaged as consisting of a set of action programs which provide a control level generating and coordinating processes. For example, manipulatory/exploratory programs would include those controlling localization, following, prehension and so forth. Such programs would underlie generativity in one sense discussed in Section B.4 by enabling a system with restricted perceptual and behavioural capacities to engage in a wide range of activities. Action programs will require selective access to the perceptual component, which is envisaged, following the analysis in Section B, in terms of autonomous processing providing modular descriptions or domains of structure each of which makes explicit some specific information and is thus potentially suited for use by particular action programs. For example, trajectory information may be useful for both following and prehension, but not for object recognition purposes. In Section C.2, the types of description apparently available to infants will be considered within the general framework provided by the notions of modularity and selective use. The relation between action programs and perceptual processing is viewed as bidirectional since such programs could either generate processes which search for, or be activated by the presence of a certain description or set of descriptions within the perceptual component. This point will be developed further in Section C.3, where the kinds of programs which might coordinate perceptual and behavioural processes will be discussed in greater detail.

The behavioural component is assumed to encompass eye-movements, reaching, grasping, and so forth. It's status within this computational framework is highlighted by emphasising the term "behavioural processes", paralleling the notion of perceptual processing. However, while Section B's discussion of AI vision systems concentrated on domains of internal structure in vision and the constraint-based processes relating them, computation need not be restricted to the manipulation of internal structures. As shown in Figure B, the symbol structures and program-governed processes of the infant system are embedded in an environment which can also be conceptualized in terms of structures and processes (c.f. Sloman's (1978) view that a mechanism or mind and its environment are simulataneously part of each other). <8> In general terms, both the subject and the environment may be conceptualized in terms of a "problem-space" or "state-space" analogy as consisting of a set of states together with the processes or "operators" which link one state to another. The subject's behaviours may be viewed as operators which have a privileged status with respect to other subject processes because they can transform structures external to the subject, i.e. states of the environment. Behavioural processes transform states of the world in two related ways. Firstly, they may function quite directly as operators, as in moving an object from one location to another. Secondly, we may view the location of the subject's body as part of the environmental state-space, and behavioural processes may transform it by altering the location of the body or parts of the body, as in locomotion or eye movements. Both types of transformation change the subject-environment relationship and provide new potential inputs to the action system. In the first case, we have the production of a new state-of-the-world; in the second case, access to a different aspect of the world. Thus, Figure B shows the direct link between action programs and the behavioural component as unidirectional, because behavioural processes exert their reciprocal effect on action programs indirectly through the perceptual component. It follows from the approach to perception developed in Section B that the new information which is made potentially available can only be considered as information from the subject's viewpoint if it can be symbolized, described or represented by the perceptual component, and if the subject possesses processes which can address, refer to, or, in general, use these descriptions (e.g. to generate further behaviour).

The conceptualization of computation and behaviour is of importance equally to the computational analysis of action, to ecological psychology's claim that cognitivism lacks a genuinely active subject, and to the more general issue of defining the nature of computational psychology. Both ecological and traditional cognitive psychologies have too commonly viewed questions of cause and effect in terms which reinforce an internal-external dichotomy. For example, concern about a purported causal relation between input and output of the system is illustrated by Michaels and Carello's question: "if X is information put into the machine and Y is the output act, what mathematical function relates Y to X (1981, p.168)." Alternatively, the cognitive psychologist has tended to ask what internal process(es) cause the overt behaviour we observe. Either of these approaches equates behaviour with the "what" of ability, seeing it in an undifferentiated and reified way as some type of external product, and this tendency underlies a common and inappropriate equation of "behaviour" with "action".

At one level, behaviour can be treated as a 'thing' and analysed morphologically in terms of the spatio-temporal characteristics of movement, but this cannot do justice to its role in action. In considering the intelligent functions with which developmentalists are concerned, it becomes clear that the "what" and the "how" of even apparently simple abilities are intricately connected, and that explanations must encompass states, processes and control. For example, in describing the infant's 'success' on a standard manual search task, we intuitively apply superordinate structural descriptions which consider behaviour as a transformational process mediating a change in the joint state-space constituted by subject and environment. We say "the infant found object x", which links the infant's possession of the object at time-t2 with an earlier event - our hiding of the object in some location at time-t1 - and the intervening processes generated by the infant. We do not describe the infant's 'success' merely in terms of states such as the outcome: "object x is in the infant's hand". Nor do we painstakingly delineate the arm and hand movements involved. Both of these are necessary to a complete account, but neither is sufficient if considered in isolation.

The transformational aspect of behaviour underlies the fact that it cannot be equated with movement or motor activity. Nor can action be equated with overt behaviour. The above description of infant performance on the search task may appear trivial, but it is not inevitable and philosophers could no doubt generate alternative accounts of mysterious forces taking over the infant and producing identical observed body movements. What such accounts would lack is the element of intentionality or purposiveness implied by the notion of the subject "finding" the object. This is readily captured within a computational framework through the virtual/physical machine distinction. In Section B, I talked in conventional terms about the computer:brain analogy, but it should be emphasised that "brain" is merely shorthand for "central nervous system". Thus, in the case of the human subject, the physical machine with which we are dealing goes beyond the brain to encompass the movable body. This is important, since it reinforces the notion that arm, leg or eye movements have the same theoretical status as the unobservable brain processes which we tend to equate with "information processing".

Analysis of the mechanisms underlying performance must involve locating the control of action at the level of programs which govern the activation and sequencing of processes, overt and covert. The control of action is then viewed in primarily "vertical" terms, as lying in programs which govern structure manipulating processes. This link between action programs and movements - the virtual and physical machine levels respectively - underlies the fact that behaviour can be conceptualized both purposively and causally. When we talk of behaviour in terms of movement, we are viewing it in the causal/process terms which are appropriate because of its implementation in body processes that are spatially and temporally located in the environment. When we talk of behaviour in terms of "acts" or "effectivities" - or call it "action" - we are stressing its intentional, purposive nature, which it attains by virtue of being controlled by programs.

The view of action as computation which is suggested here lacks the rigid demarcation between subject and environment which ecological psychologists consider characteristic of cognitivist accounts. The "what" of ability involves transformation, including manipulation of external environmental structures as well as internal symbol structures. The "how" of ability is, similarly, both external and internal, involving programs governing covert processes and external behavioural processes. From this perspective, it becomes difficult to apportion responsibility for aspects of task execution between infant and environment, since performance phenomena involve a transaction. The causal embedding of the infant system in the environment means that the adaptive properties of infant action systems are partly due to regularities in environmental structures and processes. Even a simple event may involve environmental processes some (but not all) of which are caused by the infant and some (but not all) of which may be symbolized by the infant as relevant to future activities. Thus, if we view the infant's representation or knowledge of a task in terms of the data structures and processes which s/he contributes to its execution, it is far from being an explicit copy of the environment: it is implicit and incomplete, but nonetheless functional given the scaffolding provided by a structured environment.

An implication of this transactional conceptualization of the subject-environment relationship is that the "contingency assumption" noted in Section B.1 is unlikely to simplify the task of understanding actual mechanisms as opposed to possible ones. Connections to the world and symbol manipulation should not be separated for the purposes of providing a theoretical domain for a science as Fodor and others have suggested. The transformational operations which the subject can execute are constrained as much by the range of motor activities available as by the programs that control those activities. A psychological methodology which exploits the externalization of processing characteristic of infant action may provide a valuable route to understanding psychological mechanisms.

This analysis of action links with Boden's (1981) computational approach to the structure of intentions, which also develops a systematic explanation in the sense defined earlier in this paper. She is able to show that a range of philosophical problems can be addressed by viewing any intention as having three aspects: motivation, which need not be explicitly represented but arises from the broad structure of plans and goals of the system; procedural schemata, which can be equated with plan structures at the program level discussed in the present paper; and bodily-operations, which can be purely "intellectual" or involve overt motor activity. In not restricting intention to the domain of formal manipulation of interpreted symbols, this analysis concurs with the present paper's opposition to Fodor's approach to computation and the contingency position to which it has given rise.

Thus, a computational approach can locate the subject in the environment, even though its treatment of their relationship will focus on psychologically relevant processes rather than the causal, physical principles demanded by Turvey, Shaw, Reed and Mace (1981). Instead, central issues will concern analysis of the mechanisms subjects possesses which enable them to exploit environmental stability and cope with environmental changes, including those engendered by their own behaviour. Further analysis of these mechanisms is the aim of the remaining parts of this section.

## C.2 "Abstractness" and perceptual structures and processes

In addition to the studies discussed in Section A, discrimination and habituation methods have provided evidence of a wide range of information which infants are capable of symbolizing or representing. An exhaustive listing would be neither possible nor helpful in the present context. This part of Section C has a more restricted aim: to consider whether contemporary findings can be mapped usefully onto the framework for perceptual processing, derived from AI and Marr's model in particular, which was developed in Section B. The computational emphasis on modularity and domains of description allows an organized exploration of the range of representational formats available to infants. In general, analysis of infancy studies will suggest that descriptions up to and including the type characterized by Marr's 2.5-D sketch level may be generated by newborn and very young infants, and aspects of 3-D model descriptions may be available to them. It should be emphasised that Marr's specific suggestions for the content of different domains of description are not considered to be the only or ultimate possibilities. Nevertheless, they do provide ideas which may serve a useful heuristic function in furthering our understanding of infant perception.

From Section A, it will be recalled that much contemporary infancy literature, in disputing the Piagetian view of infant object understanding, emphasised the "non-literalness" or "abstractness" of infant perception, both within and between modalities. In terms of the computational framework of Section B, this claim is ill-defined: non-literalness and abstractness cannot simply be equated. For example, each representational level of Marr's system has different properties, but none of them are literal copies of input; even the raw primal sketch embodies a type of abstractness. Those representations which are object-centred might be considered more abstract than those which are viewer-centred insofar as they are less specific to input modality and to specific relations between subject and object. On the other hand, the earlier levels of processing make explicit general qualities shared by all objects, and might be considered the more abstract in terms of this greater generality. Something like this latter notion forms the implicit foundation for Bower's (1979) claim that the infant's representation of objects develops from abstract to specific and not vice-versa.

The "abstractness" issue is closely related to the problem of what is meant by "object understanding". The computational framework throws light on some of the inconsistencies that have been generated through attempts to explore the basis of infant perception in terms of <u>bipolar</u> distinctions such as "proximal versus distal" or "features versus patterns" which tend implicitly to assume that there is <u>a</u> critical dimension of "objectness" whose presence or absence may be demonstrated. Especially persistent has been the notion that shape or form is the vital dimension of "real" object understanding. Conflicts regarding the presence and significance of shape perception are more apparent than real when they are contextualized within a computational framework, replacing a quest for <u>the</u> object with a mapping of the range of descriptions or representations available and the uses or purposes to which they are suited. This emphasis is complementary to that of ecological psychology which also stresses the wide range of information provided by objects and disputes the role which has been allocated to shape and pattern recognition as the focus of object understanding. It is distinct in providing an explanation of the equivalence classes which the subject establishes for input in semantic, representational terms, in addition to the mathematical, geometric focus of the ecological approach.

To begin with an example from Section A, a phenomenon such as "looming" may be conceptualized in terms of a relatively early level of processing. Relevant to this is Marr's extension of his model to show how pre-primal sketch level zero-crossing detectors could demonstrate "directional selectivity", providing a single source of information about movement, its direction. [The direction of movement of a zero-crossing can be ascertained from its contrast and the sign of its time derivative. Marr attempts to substantiate the psychological validity of this algorithm by arguing that psychophysical studies of the transient channels in vision and neurophysiological recordings of the Y retinal ganglion cells, to which they are thought to correspond, demonstrate that these channels are, indeed, measuring this time derivative.] The raw outputs of directionally selective units could underlie a rapidly functioning "looming detector" based on comparison of the signs of movements at corresponding points on the two retinas (these often correspond to nearby points on a single moving object), in accordance with the motion-object relations outlined by Regan, Beverley and Cynader (1979). When these signs indicate incompatible movement directions for the zero-crossing detector involved, it is moving in depth relative to the observer: approaching if it is moving away from the nose, retreating if it is moving towards the nose. Motion of the detector to the right on both retinas corresponds to a miss path to the observer's left, and motion to the left a miss path to the right. In order to see the shape of moving surfaces, a system which combines motion and contour analysis is preferable, according to Marr, but this more complex analysis is inessential to providing the approach information central to the looming phenomenon.

This account is compatible with demonstrations of infant avoidance behaviour to a film showing a continuously expanding object which changes shape during its apparent approach, from a solid square to a Y form, for example (Ball and Vurpillot, 1981). It lends support to Ball and Vurpillot's tentatively expressed conclusion that, for the infant, perception of an event such as movement in depth may not imply perception of a constant object. However, their Piagetian notion of "constant object" would require considerable differentiation. The properties of the type of description to which this term is to be restricted must be elaborated. This account contrasts with those computational approaches, more closely allied with traditional constructivism, which propose a series of hierarchically structured descriptions or representations becoming increasingly relevant to the control of action. In Oatley's framework, for example, which was noted previously in Section B.4, a high level of processing and of explicit object representation would be implicated in infant behaviour on the looming paradigm. The relevant domain of description would be assumed to make explicit information about four-dimensional entities in the world, moving objects, embodying relations such as "dangerous" and "approaching", and possessing attributes such as "to be avoided". The type of information provided by the low level of processing which has been discussed here might act as a "cue", activating a "hypothesis" of a dangerous, three-dimensional, approaching object, but would not itself be considered to be in an appropriate domain for the guidance of avoidance activity.

Spelke (1982) has reported an extensive series of investigations into the way in which infants of 3-4 months segment visual scenes to arrive at descriptions of the boundaries of objects, clarifying the role of various types of information in this process. Her findings might be viewed appropriately in terms of a 2.5-D sketch type of description since her questions ask which parts of a display are grouped together and are thus concerned with discontinuities in visible surfaces rather than more explicit descriptions of the structure of particular objects. Her detailed studies are relevant to the types of perceptual structure implicated in the prehension phenomena of Section A, which provide preliminary evidence for the infant's discrimination between solid objects and pictures. Spelke's work highlights the role of information provided by separation of surfaces in depth and motion for the infant. For example, when an object is suspended over a background, infants do not appear surprised when the whole object moves back and forth, but they are surprised if half of the object and its associated background moves in this manner. This effect is obtained even if both object and ground are the same colour and texture (e.g. both covered with orange carpet). However, if an area of one colour and texture is encompassed by an area of a different colour and texture, both in the same plane, no surprise is shown when half of the display moves back and forth 'violating' texture and colour boundaries. Such findings appear to fit well with Marr's analysis which considers colour to provide relatively peripheral information about objects. More interestingly, it was noted, in Section B, that the formal analysis of what subjects should be trying to do and why at this level of processing had revealed the use of texture to recover information about visible surfaces to be problematic. The infant data suggest that texture may not be a usable source of information in this context, even if it seemed plausible from the Gibsonian viewpoint. (It seems feasible that texture may prove more relevant to the detection of properties such as degree of hardness of a surface.)

Unlike texture and colour, motion does appear to provide a crucial source of information for infant grouping processes. For example, when infant looking time habituates to repeated presentations of back and forth movements of a rod whose centre is occluded, recovery of visual attention is subsequently shown to a split rod but not to a complete rod which exhibit the same motion. This implies that the two segments of the initial display were perceived as a unitary entity, and the broken rod stimulates recovery from habituation because it appears novel in terms of the original representation. This pattern cannot be established with static displays, thus emphasising the significance of "common movement". Indeed, Spelke has demonstrated, using the same procedure, that infants will group together areas which move back and forth in this manner even if they differ in colour, texture, and shape, and their contours and major axes are not aligned. She discusses this phenomenon in terms of "perceiving the unity of a partly hidden object", but the vocabulary of surface descriptions appears more appropriate, particularly in view of the fact that the infant's description does not discriminate between cases where there is and is not a unitary object to perceive at a level of processing suitable for 3-D object recognition.

From these and other examples, Spelke also argues that infants do not make use of Gestalt-like principles involving shape or form in order to segment an array into individual objects. Besides "common movement", processes which recover information concerning the "connectedness of surfaces" appear to be involved. Infants appear to treat two objects as such only if they are separated in depth, they treat them as a single object when they are adjacent and touching, even if each has a simple, continuous boundary. This again fits with Marr's claim that the types of processes associated with description at the relatively early 2.5-D sketch level should not involve a priori information about the possible shapes of objects. The type of description involved is abstract in the second of the senses discussed earlier because it provides a suitable format for any of the many particular objects which ah infant may encounter. However, caution is needed with respect to Spelke's conclusion that shape or form is not an "essential" property of objects. Her studies concern only the role of simplicity of shape or form, and with reference to a single use or purpose: segmentation of an array into unitary bounded entities. Marr's analysis suggested that shape, in the sense of the type of information made explicit by the 3-D model description (as opposed to simplicity of form), is an essential object property to recover if one's purpose Is to establish the complex correspondences associated with recognition or classification.

The question of whether *or* not infants are capable of describing input in terms similar to Marr's 3-D model description level can be asked by looking at investigations of their capacity for maintaining or generalizing a description in the face of variation in viewer-centred object properties such as orientation. For example, Cohen and Strauss (1979) used habituation of looking at photographs of faces to study what they termed "concept acquisition" in infants from 18- to 30-weeks of age. Repeatedly presented with a photograph of the same face in the same orientation, infants at all ages increased their looking time equally to either the original or a new face presented in a new orientation. However, if the original face had been presented in a different orientation on each trial, the 30-week olds showed this increase of looking time only to the novel face; the *younger* infants, again, increased their looking time to both. This leads Cohen and Strauss to conclude that only the oldest group was responding to the specific face shown, younger infants of an age comparable to those of Spelke's studies, were responding to the less abstract dimension of orientation. Also, only the oldest group indicated some type of description underlying a category of "faces in general": when a series of different faces were presented in a single orientation for the habituation trials, this group showed no increase of attention to either a novel or familiar face with a new orientation. The fact that these phenomena are obtained with photographs, eliminating motion information, suggests that processes which were not in operation in Spelke's studies are functioning here. However, it would be inappropriate to conclude that elements of an 'abstract', object-centred description are unavailable to infants younger than 30-weeks or so of age. Neither the issue of what such a description would entail, nor the importance of specifying its purpose have been paid the necessary attention.

It will be recalled, from Section B.3, that Marr argued for the merits of a description which made explicit features of an object's overall shape in terms of its volumetric as opposed to surface properties. If we pose the problem of how the orientation of a face might be recovered, instead of intuitively assuming it to be a low order, relatively literal feature of stimulation, one answer might be provided by Marr's volumetric scheme. Thus, the sensitivity of young infants to a face's orientation may be evidence that its shape is described in terms of its main model axis, corresponding to its gross size and orientation. Age differences might lie in the degree of differentiation of such a description, or in the relevance of and capacity for classificatory recognition.

What, then, should be made of studies suggestive of orientation-independent representation and "shape constancy" for 2- and 3-dimensional objects in newborn and very young infants? Cohen (1979) has proposed a distinction between "simple" patterns or forms, which even the youngest infants treat as compounds and as equivalent across orientational change, and "complex" ones, such as faces and coloured forms, which they do not. But this largely post hoc distinction seems incompatible with the neonatal imitation phenomena, which were introduced in Section A and are discussed further below; even neonates appear capable of what were agreed to be "abstract" descriptions of "complex" objects such as faces and hands. The fact that the younger infants description of a face does not support recognition of it as "the same face" does not entail that complex body descriptions suited to other purposes are not available to the infant. The answer may be that this type of data is, in fact, much more problematic than is generally thought once the properties of the stimuli used and the types of description onto which they might be mapped are considered in more detail.

For example, Schwartz (1975: cited by Day and MeKenzie, 1977; Cohen, 1979) found that very young infants whose looking time habituated to a square appeared to "classify[11] it as the same when it was presented rotated through 45 degrees. However, it is possible that such infants are not displaying a higher, more abstract level of description with this simple stimulus than they could achieve with the more complex face: the opposite may be the case- Interestingly, this type of shape constancy is not found in adults. For them, the square and diamond are not equivalent and this has been linked to representational formats including "significant directions" (e.g. Hinton, 1979), a notion which is compatible with Marr*s 3-D model scheme. Infants may not be able to generate such a description with these kinds of stimuli. Schwartz's and similar studies have tended to use 2-dimensional black and white stimuli which seem ideally suited to description in terms of the vocabulary of types of intensity change which characterizes Marr's primal sketch. It may not be possible to recover the type of volumetric information which could form the basis for an object-centred representation from such 2-dimensional stimuli due to the nature of their bounding contours and the consequent absence of an appropriate correspondence between contour inflections and potential points on the surface of a 3-diraensional object. It would be interesting to explore this issue using 2-dimensional contour figures similar to Picasso's [1]Rites of Spring[1] which Marr has used to illustrate properties of his 3-D model description and whose formal properties are better understood.

A shape constancy phenomenon involving "simple" 3-dimensional objects also proves problematic in an interesting and different way within the context of Marr[f]s model. Bower (1966) found that very young infants who were conditioned to turn their heads In the presence of a solid rectangle in the parallel plane would generalize their rate of responding to the same size rectangle presented now at an angle of *H5* degrees to their line of sight. This pattern of generalization was not found to a trapezoid in the parallel plane generating the same retinal shape as the rotated rectangle, or when the stimuli were presented as photographic slides rather than solid objects. Similarly, McKenzie and Day (1973) discovered that infants showed identical patterns of habituation with either repeated presentations of a single cube or a series of presentations of the cube in different orientations; this did not occur with cut out photographs of the cube in different orientations. Day and McKenzie[f]s (1977) review of these and other constancy studies concludes that it is slant information derived from motion parallax which underlies this early capacity for establishing equivalence.

MarrVs tripartite explanatory system proves highly relevant here, and suggests that the developmentalists* conclusions concerning either what infants are achieving or how they are achieving it must be in error. Subject and/or object movement may certainly provide a basis for the recovery of information about surfaces, but computational work should lead us to be cautious about equating this simply with object shape. Section B's summary of Marr's model noted that physical constraints on object movement can provide one basis for processes constructing the 2.5-D sketch. For example, Ullman (1979) conceptualizes this in terms of processes which establish correspondence between elements in a set of sequential views and recover three-dimensional structure from the motion which they have undergone. He has been able to provide mathematical proof that if at least 3 views of 4 points on a rigid body, lying on different planes, are available, their three dimensional arrangement in space is uniquely determined. Marr (1982) discusses also the role which optic flow, the velocities of elements on the retinal surface, might play. This provides a more formal footing for Spelke's notion of a "common movement principle" in infant perception, and it is compatible with the idea that an object which is changing its position or orientation would continue to be seen as a unitary entity in motion.

There appears, however, to be no way in which such processes could establish the equivalence of object shape across discrete presentations of an object in different orientations, unless one were dealing with apparent movement phenomena as opposed to the temporally distinct trials of the infancy studies. At best, a series of 2.5-D sketch type descriptions might be possible, but these would not be equivalent across different viewpoints in the absence of a description which could make explicit the object's internal structure. Thus, the information stressed by the developmentalists is simply inadequate for object shape constancy under the conditions in which they have tested it and in the sense in which they generally intend it. The fact that these equivalence phenomena cannot be achieved with static two-dimensional displays is particularly important, for it reinforces the view that mechanisms more akin to those studied by Spelke and compatible with 2.5-D sketch level analysis may be operating here.

In conclusion, it can be suggested that it is possible to map between our understanding of infant perception and a computational framework, and that this is a useful thing to do in several respects. To the extent that parallels are found to exist between infant perception and AI specifications of information processing systems, such as Marr's, the greater degree of detail of the computational framework enriches our understanding of what infants may be doing. It also provides an integrative framework for a potentially wide range of phenomena which may, initially, appear unrelated or incompatible. As exemplified by the "shape constancy" example, in spelling out the importance of distinguishing between available information and description generating process, it restrains us from concluding oversimply that if information is available it must be used or even be usable for constructing a representation which would be adequate for a particular purpose.

### C.3 Perceptual-behavioural relations: Action programs and the pick up and use of information.

The distinction between availability and use of information has been emphasised throughout this paper, and the aim of the present section is to explore the way in which the notion of action programs can increase our understanding of the selective pick up and use of information which is evidenced in functional perceptual-behavioural activities of the type discussed in Section A. It has been suggested, in Sections B.4 and C.1 respectively, that ecological psychology's notions of "affordances" and "effectivities" are compatible with the computational framework which is being developed. However, the ecological movement's approach to the control of behaviour, especially its key concept of attention, has been shown to have significant limitations. Here, the potential advantages of the active, coordinative aspects of programs will be discussed further. It is argued that these can make it possible to operationalize and develop sufficient accounts of attentional phenomena and the close structural correspondences which exist between environmental information and the infant's activities.

The so-called procedural languages, which were originally designed for systems engaged in problem-solving (Bobrow & Raphael, 1974; Shapiro, 1974), may serve to illustrate the kinds of programs which a computational analysis of action would require. Such languages consist of two interdependent aspects: a data-base or set of symbolic assertions which describe objects, events or states of affairs, and procedures which, when activated, generate processes which change, delete or add to the symbolic assertions in the system. In such languages the representation of knowledge is primarily implicit in procedures and not explicit in extensive declarative data-bases. Thus, considering action programs in such terms is highly compatible with the notion of action as a form of tacit or implicit knowledge.

Two important points need to be mentioned. Firstly, a program is not merely a fixed series instructions. The symbol structures which represent objects (e.g. the descriptions central to the discussion of vision) or activities to be executed (either internal or behavioural processes) can contain variables which permit them to represent a general class of objects or a general strategy or pattern of action. Variables can take a range of values (which may qualitative in some cases and quantitative in others). For example, Sloman (1982a) presents the following (schematic) example of a piece of program which could be involved in forming a plan or executing an action:

```
PROCEDURE ([GRASP ?X])
   PRECONDITION: CLEARTOP X
   PRECONDITION: EMPTY HAND
   MOVETO X
   GRASP
END
```

This contains a single variable "X" designating an, as yet, unspecified object. Secondly, a procedure is activated when a value is given to the variable(s) it contains, providing an input to that procedure. But, this need not imply a unidirectional relation between data and procedures, a fact which is of particular relevance to the operationalization of attention. For example, in pattern directed invocation of procedures, part of the definition of a procedure is a general pattern (as in the example above) which could be matched by a range of symbolic assertions, and the process which the procedure governs is activated whenever an input in the form of a particular symbolic assertion which matches this pattern is added to the data-base. But the processes which lead to this happening can originate either outside the system or internally as the consequence of the operation of some other procedure, and procedures can be set to actively search for a particular input at a given time.

Prazdny (1980) outlines a working program which includes pattern directed invocation in the form of IF-ADDED and IF-REMOVED procedures to model a general strategy for infant anticipation of a moving object which disappears from view behind a screen. For example, the IF-ADDED procedure named OCC1 is triggered whenever a two element pattern (DISAPPEARED ?OBJECT-X) is added to the data base. The first element of the pattern is fixed (DISAPPEARED), the second element is a variable (?OBJECT-X) which, once replaced by a particular value or object description, remains constant for that run of the procedure. Once activated, it will search the data base for the presence of a three-element pattern (PAROCCLUDES ?OBJECT-Y OBJECT-X) whose first element is fixed, third element is the known value of X, and whose second element is another variable for which a value will be available if the missing object has recently moved behind another. Succesful matching of this pattern leads to further alterations in the data base and to invocation of another procedure whose role is to check for a pattern representing partial occlusion in the region of the opposite edge of the occluding object. In terms of the discussion of program-process relations in Section C. 1, it should be emphasised that the process which is set in motion when the latter procedure is invoked would need to be a behavioural process (involving head and eye movements), though this is not specified by Prazdny.

One may have many reservations about the specific content of this program. For example, does the choice of symbolic assertions indicating disappearance and partial occlusion really make it necessary to explicitly add a pattern (OCCLUDES OBJECT-Y OBJECT-X) to the data base? Nevertheless, it suffices to illustrate three things which are relevant to the operationalization of attention. Firstly, the active nature of the program handles the perplexing issue of "who" tunes the receiver, to use the direct perception metaphor. The procedural representation provides a sufficient account of responding to and actively seeking information in the context of the general program-governed (though only implicitly represented) goal of maintaining contact with the object. Secondly, it shows how a computational approach can cope equally with both the overt/behavioural and internal/central aspects of attention without regressing to a discussion of brain states. The "how" as well as the "what" of attention is ammenable to a psychological analysis in computational terms. Finally, it counters ecological psychology's criticism that for processing approaches information is always in the computer rather than the environment. In fact, the computational notions discussed here are well suited to modelling a system whose functioning relies upon picking up and looking for information in the environment.

The attention issue focusses on the control of the selective relation between perception and behaviour. But ecological psychology has also emphasised the intricate structural correspondences which exist between perception and behaviour. For example, much of the information which the subject uses appears to be "body-scaled". As Michaels and Carello put it, for the example of someone hitting a baseball, the relevant information is body-scaled in terms of time to contact, not absolute time. From the ecological perspective, the subject needs no monitoring device: "Regulation is a consequence of the fit between the optic array and the activity (Michaels and Carello, 1981, p.54)." However, the claim that there is a relationship of "mutual constraint" between effectivity and affordance structures is more descriptive than explanatory, and it is necessary to ask how this structural correspondence could be achieved.

The type of perceptual-behavioural correspondence revealed in the prehension and imitation phenomena which were discussed in Section A indicate how the computational analysis needs to be extended. Both suggest that some representation of the body, a "body schema", may need to be an integral part of the behavioural component. Section C.2 noted the possible link between prehension and Spelke's studies of the infant's perception of bounded surfaces. The 2.5-D sketch type of description with which her analysis was compatible would seem quite appropriate for this particular purpose. Not only does it make explicit the size and orientation in depth of surfaces with reference to the viewer, but also the way in which it represents objects could produce the same 'errors' as young infants if it were incorporated into a mechanism for prehension. For example, infants cease reaching for an object if it appears to lose a boundary by being placed on and in the same plane as a larger object (Bower, 1979). Here, it is notable that the infant appears not only to re-parse the visual array, but also ceases reaching in the presence of the new 'larger' object. Only descriptions of bounded surfaces of "graspable size", where this appears to be defined in terms of the infant's hand-size, are coordinated with the behavioural processes relevant to prehension (c.f. Michaels and Carello's example of the praying mantis).

One way in which such phenomena could be modelled would be to view the type of perceptual description being discussed as a precondition within a program such as the simple grasping program which was mentioned earlier. The action program involved might be structured so that the presence of the description in the data base invoked the procedures associated with reach and grasp behaviour. Further preconditions could be introduced to take size into account, such as:

PRECONDITION: SIZE X LESSTHAN SIZE HAND

When the description changes - as a consequence of the manipulations in Bower's study, for example - the preconditions are no longer met and the procedures governing the relevant behaviours are not activated. It should be noted that no explicit representation of the consequences of applying procedures are necessary to guide action, i.e. the infant system does not need to 'know' that it will fail to grasp an object if it is too large; the structure of the action program involved ensures that the attempt will not be made.

An important issue is raised by the nature of the description which makes explicit information about object size and its relation to the body. The simple precondition described above would be expected to subsume many procedures which would certainly not produce a simplistic description of object size as smaller than, equal to or greater than the hand. For example, the more detailed treatment of looming in Section C.2, showed how movement with respect to the nose could be made explicit without any explicit representation of the nose per se. Thus, the need for a "body schema" might be challenged in the present example. However, its usefulness is clearly shown in attempting to clarify the nature of imitation.

Imitation is particularly interesting since it reveals not only correspondence between perception and behaviour, but also isomorphism in terms of both the body parts involved and the parameters of their movement. From Section A, it will be recalled that Meltzoff argued for the necessity of a very abstract representation to permit the infant to match the movements of an adult hand, say, with those of its own. Even if both were visible to the infant, their size, texture and so forth vary considerably and deciding what kind of mechanism could "categorize" them as the same is non-trivial. Appropriate imitation involving unseen body movements suggests strongly that a "body schema" or representation of the infant's body must be included in a complete account of the infant (c.f. Mounoud and Vinter, 1981), but this leaves open the issue of what such a schema might consist of.

A possibillity is that the correspondence between body parts which characterizes imitation provides strong empirical evidence that the type of description which was exemplified by Marr's "3-D model description" (Sections B.3 and C.2) is implicated. This, it will be recalled, is based on volumetric primitives (generalized cones) whose geometric properties make it possible to build a description of the relative spatial arrangement of an object. Thus, this type of description is "abstract" in the sense of being object-centred or independent of the subject's viewpoint. A 3-D model description for the hand would, at one level, consider it as a single volume, thus making explicit it's main axis which, at a further level, would serve as the coordinate system for the axes of the fingers, each of which may be considered as an independent volume with its own axis.

Specific to imitation, we could conceptualize it in terms of the descriptions involved, together with procedures governing matching processes between them. Earlier, emphasis was placed on the fact that symbolic assertions or patterns which describe objects, events or activities can contain variables as well as fixed elements. Additionally, Marr viewed object description in terms of shape with other aspects of the description (such as size, colour, etc.) "hung-off". Thus, one aspect of imitation could be conceptualized in terms of procedures for matching fixed elements relevant to shape in descriptions within the perceptual component and behavioural components (the perception of the adult hand and the "body schema" description respectively). This would explain the choice of appropriate body part. In addition, the role of variables in these descriptions could underlie the isomorphism of movement which characterizes imitation.

A recurrent criticism of AI models has been that they cannot model analogue transformations such as "openning and closing". But work by Hinton (1979) shows that this is not the case. Hinton argues for the central role of structural descriptions - as opposed to analogue representations - in both perception and imagery, and provides experimental evidence that the relative difficulty of imagining certain transformations is systematically related to structural descriptions assigned during perception. He argues that there is no reason why such descriptions cannot include "real-valued" variables. For example, the orientation of an edge can be represented by the visual angle it makes with the "assigned directions" of the structure containing it. Transformations such as dilation or rotation can then be modelled by continuously changing the real-valued labels on the relevant structural description.

The 3-D model description of a hand, discussed above, could represent the position of the fingers in terms of real-valued variables, the angles their volume makes with the main axis of the hand. A movement such as openning would then involve the activation of procedures which increased the values of these variables. In addition to procedures for matching fixed elements in the description of adult and infant hands, procedures which compare real-valued variables can be envisaged. The movement aspect of imitation may be viewed in terms of this comparison, together with procedures for generating behavioural processes (body movement) which will bring the description of infant and adult hands into alignment.

Even though it is sketched very roughly, this computational treatment of imitation can already provide a clearer understanding of what Meltzoff, following Piaget, refers to as "active matching with intelligent confusion". Also, the nature of structural descriptions, together with their matching between perceptual and behavioural components, may account for the infant's 'near misses[1], such as when Jacobson finds tongue protrusion in reaction to the movements of a pencil. The reliance of imitation on structural correspondence between the mechanisms of perception and production is consistent with Maratos (1982) evidence that neonates only imitate gestures which they can perform spontaneously. While this discussion has been confined to a single example, the approach may have potential for conceptualizing and exploring intermodal relations such as auditory-visual coordination. Perceptual inputs in general may be able to specify the values for real-valued variables involved in motor control via appropriate coordinating programs.

In previous sections, the conceptual advantages of a computational framework for infant abilities have been elaborated. The examples developed in the present section begin to address the use of a computational approach to provide sufficient explanations of difficult notions such as "attention" and "body schema". While attention is central to the "active perceiver" of ecological psychology's metatheory, the notion of "body schema" is more usually linked with traditional cognitivist positions which emphasise the role of copy-like representations. Yet it has been shown that both might profitably be treated as outcomes of the interactions between computational processes. This reinforces the earlier claim that computational ideas can provide an effective medium for discussing mechanisms without endorsing a constructivist position. The structures and processes which were discussed are not copies, they do not bear a literal resemblance to the world outside the subject, nor do they replace it.

## D. CONCLUSIONS

The main concern of this paper has been to establish that we need more powerful and explanatory theories of the mechanisms underlying infant abilities, and that appropriate concepts and methods for the development of such theories can be found within a computational cognitivist framework based on recent research in Artificial Intelligence. This framework has been contrasted with the increasingly influential ecological psychology movement, which rejects the relevance of information processing approaches to perception and seeks to emphasize the preadapted functioning of the infant in a structured, information rich environment.

It was argued that a systematic explanation of infant perception is required in terms of the role of perceiving in action systems, and that treating action as an organizing structure from a computational viewpoint can subsume many of the concerns of ecological psychology. Ecological and computational psychologies cannot be treated as alternatives. Both share a concern with formalizing the structure of environmental information. And many of ecological psychology's emphases - such as the "active subject" - are metatheoretical and can be conceptualized as well or better in terms of computational cognitivism. Rapprochement may be possible and should be advantageous to both ecological psychologists and AI workers.

Ecological psychology can benefit from AI "processing" approaches to perception, and the present paper stressed that the generativity of current AI vision models does not rely on the embellishment of inadequate information which characterizes traditional constructivism. Additionally, there are advantages in handling the control of action - issues of intention and purpose, the nature and role of behaviour, attention and the selectivity of perception - in terms of computational notions such as the program/machine distinction and attempting to operationalize these ideas by implementing them in working programs.

Correspondingly, AI can gain from taking ecological psychology seriously. The present paper has argued that a computational psychology of infant ability must emphasise the notion of causal embedding and the nature of computation as a process of rule governed structure manipulation - including environmental structure - as opposed to a purely intra-subject process of symbol manipulation. This is consistent with ecological psychology's stress on the "active subject", but at variance with the concerns of most contemporary AI. Thus, the framework developed here has attempted to show that ecological psychology's action-based notions of "affordances" and "effectivities" can and should be understood computationally. This contrasts with AI critiques of the theory of direct perception such as Ullman's (1980) challenge which fails to consider the action-based nature of perceiving and discusses neither of these concepts.

Thus, the present paper favours computational concepts, but it is not uncritical of AI. Little of the discussion was devoted to 'clever' programs which can do what infants do - although programs exist which can, for example, recover shape information from optic flow or model the process of visual proprioception. It is the more fundamental, recurrent concepts and principles of AI which, at present, are of greatest utility for our attempts to conceptualize infant phenomena and clarify our ideas. These general concepts need applying to particular content areas, and such applications are notably uneven in current AI as is shown by the range of ideas available for considering vision in contrast with relatively unexplored notions such as "body schema". It is here that the methods of psychology and AI can come together to explore the possibility of an interdisciplinary methodology. Important questions which link ecological and cognitivist concerns can be established by considering infancy. Considering the infant as an example, par excellence, of a causally embedded system may enable us to generate a methodology which is relevant to actual as well as possible intelligent mechanisms.

Footnotes

1. Recently, Pylyshyn (1980) has advocated taking the computational metaphor for the mind literally. He explores the possibility of making a principled distinction between fixed functions or capacities which can be described as instantiating physical or biological laws - the "functional architecture" - and those which can only be explained in terms of computation.

2. It is interesting to note that developmental psychologists often implicitly invoke forms of the contingency assumption. One example is found in the belief that manual or visual search tasks provide equivalent access to the infant's knowledge of object existence constancy.

3. Despite the intimate connection between computational concepts of representation and symbol, the latter has been virtually ignored outside the field of cognitive science. Newell (1980) bemoans the way in which this key notion has been ignored in information-processing style cognitive psychology texts. This neglect is also a feature of the developmental cognitive science literature where no texts explicitly compare and contrast cognitive science and developmental psychology uses of symbol. For example, Mandler (1981) effectively contrasts representation as knowledge with representation as the use of words, drawings or other artifacts to stand for or refer to some aspect of the world or of one's knowledge of the world. However, she views the latter as symbol use, thus restricting her definition of that term to a sense in which the representational process implied by "x represents y" has a communicative function telling others or the self that when "x" is used it is meant to stand for a piece of shared knowledge. Thus, she goes on to argue that much of our thinking may not involve any symbol manipulation at all. From the cognitive science viewpoint this is, by definition, not true, but from Mandler's perspective it is intended as a criticism of the Piagetian theory of thought. Piaget himself distinguished between the two senses of representation noted - what he called representation in the "broad" and "strict" senses - but he provided no analogue to the computational sense of "symbol".

4. There are complex and unresolved questions about whether the words of a language using computer could be said to "mean" or "represent" in the same way as human words do. Since the present paper is concerned with the pre-verbal infant, this issue will not be considered further here. Relevant discussions can be found in Winograd (1980) and Wagner (1982).

5. There are 6 types of "ell", 3 types of "fork", 4 types of "tee" and 3 types of "arrow". Winston's (1977) introduction to constraint analysis discusses 18 legal junction arrangements, but 3 of the forks are identical under rotation.

Carnegie Mellon University Libraries
3 &H&E D1323 ISEfc,


6. Ames* much maligned psychology has been subjected to extensive criticisms of being "unecological". But his work does not establish "the unreliability of Npeep-hole' vision (Costall, 1981:43)", rather it provides excellent demonstrations of some of the constraints exploited by our perceptual processes: the types of information we use and how they interact. Thus, the well known "Ames* room" demonstration - in which faked rectangularity cues lead to one person being seen as half the size of another despite appearing at the same distance from the viewer - exposes the functioning of a constraint which computes the size of any object from information about its distance from the observer and its retinal image size. This is not as important in showing that perception can be "misled" as it is in confirming that perceptual phenomena such as size perception can be viewed computationally in terms of constraints without involving 'expectancies' or acquired knowledge about the size of particular objects. Indeed, the power of this demonstration lies in the fact that we really perceive one person as half the size of another, even though we know they are the same size. What we perceive is bizarre, but not unstructured or meaningless, much like the well known sentence "colourless green ideas sleep furiously".

7. There remain complex, unresolved issues concerning how such programs should be conceptualized. For example, Draper (1980) argues that many, if not all, vision programs have been "model-based" - interpreting input by attempting to match it to specific structures believed to characterize the particular object domain concerned - even when they give the appearance of being based on a general purpose method applying rules in a bottom-up manner. He notes that programs for interpreting line-drawings have included successively smaller models, from Roberts[1] (1965) use of complete, simple, convex polyhedra, e.g. bricks or wedges, through possible vertices and their appearances in line-labelling approaches, to the planes of his own ELLSID (Exhaustive Line-Labelling using SIDedness reasoning) program. Certainly, all such programs are knowledge based, but the use of "model" may be misleading. It is important to maintain the distinction between a program whose procedures have a range of application restricted to a specific range of inputs and one whose domains of structure include a representation of legitimate cases against which inputs are matched. Constraint analysis is more like the former, as, indeed, is the ecological psychology notion of an information specific resonator.

8. The kinds of structures and processes which constitute the environment merit additional analysis beyond the scope of the present paper. They cannot be limited to the straightforward spatio-temporal domain of physical objects and events, which may be conceptualized as structures but are not symbol structures; the domain of other program-governed symbol manipulation systems, such as persons, must also be included. The interaction between these two domains vis-a-vis the infant is likely to prove particularly interesting, as shown by recent attempts to conceptualize both what and how the infant learns about the physical world as mediated by social processes (e.g. Kaye, 1982; Sinha, in press).