

# Today the earwig, tomorrow man?

David Kirsh

*Department of Cognitive Science C-015, University of California, San Diego, La Jolla, CA 92093, USA*

Received November 1987

Revised January 1988

## *Abstract*

Kirsh, D., Today the earwig, tomorrow man?, *Artificial Intelligence* 47 (1991) 161–184.

A startling amount of intelligent activity can be controlled without reasoning or thought. By tuning the perceptual system to task relevant properties a creature can cope with relatively sophisticated environments without concepts. There is a limit, however, to how far a creature without concepts can go. Rod Brooks, like many ecologically oriented scientists, argues that the vast majority of intelligent behaviour is concept-free. To evaluate this position I consider what special benefits accrue to concept-using creatures. Concepts are either necessary for certain types of perception, learning, and control, or they make those processes computationally simpler. Once a creature has concepts its capacities are vastly multiplied.

## **Introduction**

Is 97% of human activity concept-free, driven by control mechanisms we share not only with our simian forbears but with insects? This is the challenge proposed by Rod Brooks and fellow roboticists to mainstream AI. It is not superficial. Human activities fall along a continuum. At one extreme are highly reactive, *situationally determined* activities: walking, running, avoiding collisions, juggling, tying shoelaces. At the other extreme are highly *cerebral* activities: chess, bridge playing, mathematical problem solving, replying to non-obvious questions, and most discursive activities found in university research laboratories. It is an open question just where to draw the line between situationally determined activity—activity that can be initiated and regulated by smart perception-action systems—and activity that requires thought, language-like conceptualization, and internal search.

Brooks' position is that if we consider precisely what sensing is required to intelligently control behaviour in specific tasks, we make the startling discovery that in most cases there is no need, or next to no need, for symbolic

representation. Reasoning in the familiar sense of retrieving cases, drawing inferences, and running through possibilities ahead of time is costly and unnecessary. In fact representations often *get in the way* of behaviour control. Accordingly, efficiency and parsimony dictate using action control systems that are representation free.

Moreover, unless we *first* understand the 97% of behaviour that is non-representational, Brooks argues, we will never correctly understand the remainder. The trouble with AI so far is that it makes false abstractions. Theorists don't study the genuine requirements of intelligent behaviour. Instead of finding out exactly what vision and the rest of our sensors should deliver to permit the intelligent control of behaviour, AI researchers have cavalierly defined nicely formal models of the world—the alleged true output of the senses—and have simply assumed that somehow sensory systems can build these up. Within these false castles AI theorists have tried to solve their own versions of the planning problem, the learning problem and so on. But, of course, the assumptions of these models are false—so false, in fact, that no step by step relaxation of assumptions can bring them closer to reality. The models are false and so are the problems: cognitive phlogiston.

In what follows I will question these claims. I am not yet convinced that success in duplicating insect behaviours such as wandering, avoiding obstacles, and following corridors proves that the robotics approach is the royal path to higher-level behaviours. Insect ethologists are not cognitive scientists. There is a need for the study of representations. Nor do I think that existing research in reasoning is foundationless. Whatever the shape of robotics in the future it will have to accommodate theories of reasoning roughly as we know them. Abstractions are necessary.

My primary focus will be the claim that the majority of intelligent activity is *concept-free*. I use the term concept-free rather than representation-free, as Brooks prefers, because it seems to me that the deepest issues posed by the robotics approach really concern the place of conceptualization in intelligent activity, rather than representation per se.

The concept of representation remains a sore spot in foundational studies of mind. No one is quite sure exactly what the analysis of “state  $X$  represents the information that  $p$  is  $H$ ” should be. A glance at Brooks' robots shows that they are riddled with wires that carry messages which covary with equivalence classes of earlier signals (e.g. an edge covaries with an equivalence class of pixel configurations) and which often covary with properties in the environment (e.g. real edges, hand manipulations). If covariation is sufficient for representation then Brooks too accepts the need for representations.

It is clear that by representation, however, he means symbolic, probably conceptual representation. Let us define a symbolic representation as one which can be combined and manipulated. This condition adds the notion of *syntax* to representation. To get systematic generation of representations it is

necessary to have a notation that is sufficiently modular that individual elements of the notation can be combined to make molecular expressions. In this way, ever more complex structures can be constructed and used by a finite system. Semantic discipline is maintained on these symbol structures by enforcing Frege's requirement that however complex the symbol, its meaning is a function of the meaning of its parts and their syntactic arrangement.

If an agent has symbolic representations in the sense just defined, we may assume it has concepts.<sup>1</sup> But too little is understood about the nature of computation to require that all concept-imbued creatures operate with language-like internal notational elements. In principle, there could be computational architectures which implement the cognitive capacities we suppose concept-using creatures to have, but which do not pass notational elements around. These systems have the capacity for systematic representation in that they can systematically predicate property referring states—that is *predicates*—with states that refer to individual subjects—that is, *names*. But they do not have local notational structures which we can readily identify with symbols.

This capacity to predicate is absolutely central to concept-using creatures. It means that the creature is able to identify the common property which two or more objects share and to entertain the possibility that other objects also possess that property. That is, to have a concept is, among other things, to have a capacity to find an invariance across a range of contexts, and to reify that invariance so that it can be combined with other appropriate invariances. Moreover, combinations can be considered counterfactually. Thus if an agent has the concept red then, at a minimum, the agent is able to grasp that apples can be red, paint can be red, and so on.<sup>2</sup> The agent knows the satisfaction conditions of the predicate. Similarly, if an agent has the capacity to make a judgement about an individual—a person, number, or an object in the visual field, for example—then the agent must be able to make other judgements about that individual too. For instance, that 5 is prime, that it comes after 4, that it is a natural number.

In the same spirit, it is because we have concepts that we can make judgements of identity, as when we decide that the person we see in the mirror is the same person we see over there. Or again, because of concepts we can reidentify an individual, recognizing that the object or person in front of us now is the same one we met on other occasions.

Animals which have such capacities clearly have extra talents, though just what these extra talents are, is not entirely understood. Human newborns are largely devoid of them, but soon acquire them; dogs may have elements of

<sup>1</sup> I am fully aware that identifying syntactic symbol manipulation with possession of concepts begs the question of symbol grounding and the philosophical problem of reference. For the purposes of assessing Brooks' position, however, the identification is fair since one of his targets is clearly declarative representations.

<sup>2</sup> Providing of course that it has the other relevant concepts, apples, paint . . . .

them; chimps certainly do, and praying mantises certainly do not. Possession of concepts in a full-blooded form appears only some way up the evolutionary ladder.

The problem which I see Brooks posing is this: At what point in a theory of action must we advert to concepts? Which activities presuppose intelligent manipulation of concepts, and which do not? Accordingly, this is not simply a question of the role of model-based planning in intelligent activity. It is a question of the role of *thought* in action.

There are many ways of thinking that do not presuppose use of an articulated world model, in any interesting sense, but which clearly rely on concepts. Recall of cases, analogical reasoning, taking advice, posting reminders, thoughtful preparation, mental simulation, imagination, and second guessing are a few. I do not think that those mental activities are scarce, or confined to a fraction of our lives.

Nor do I think they are slow. When a person composes a sentence, he is making a subliminal choice among dozens of words in hundreds of milliseconds. There can be no doubt that conceptual representations of some sort are involved, although how this is done remains a total mystery. As an existence proof, however, it establishes that conceptual reasoning can be deployed quickly. Yet if in language, why not elsewhere?

Brooks' own position is extreme: at what point must we advert to concepts?—almost never. Most activity is thought-free, concept-less. It is this view I shall be questioning.

My paper has two parts. In the first I spell out what I take to be the strongest reasons for extending the domain of concept-free action beyond its usual boundaries. There is in Brooks' work, the outline of an alternative theory of action well worth understanding. It has clear kinship lines with associationism, ethology, the theory of J.J. Gibson, and the Society of Mind theory of Minsky. But it departs from these in interesting ways.

In the second part I consider what conceptualization buys us. More particularly, I explore the motives for postulating conceptual representations in (1) a theory of action; (2) a theory of perception; (3) a theory of learning; and (4) a theory of control.

## 1. Action and conceptualization

From a philosophical point of view the idea that concepts might not play an essential role in a theory of human action is unthinkable. According to received wisdom, what differentiates an *action* from a mere movement such as twitching or wincing is that the agent knows what he or she is doing at the time of action. The action falls under a description, understood by the agent, and partly constituting its identity. Thus the qualitative movement of raising an arm

might at one moment be a communicative act such as gesturing goodbye, while at another moment be an act of stretching. Just which act is being performed is a function of at least two factors: the agent's intention, and the social context.

For an agent to have an intention, and hence to know the action performed, it is not necessary that he or she be aware of the action's description or that he or she consciously think before acting. Few agents are aware of putting their words together in sentences before they speak, or even of mapping between words in different languages when they fluently translate. This absence of conscious thought does not prevent them from saying what they mean and from translating aptly. Yet, any reasonable account of their practice must refer to their concepts, ideas, presuppositions, beliefs, etc. Introspection is misleading, then, as an indicator of when concepts and beliefs are causally involved in action.

Philosophy has bequeathed to AI this legacy of unconscious beliefs, desires and rational explanation. AI's signal contribution to action theory, so far, has been its computational revamping. In practical terms, this has meant that an agent acts only after planning, and that in order to plan, the agent must call on vast fields of largely unconscious beliefs about its current situation, the effects of actions, their desirability, and so forth.

Brooks' rebellion, not surprisingly, stems from a dissatisfaction with this approach in dealing with real world complexities and uncertainties. Surely children do not have to develop well-formed beliefs about liquids, however naively theoretical, in order to drink or go swimming. Even if we do require such implicit theories of children we cannot require them of gerbels or sea lions. The two forms of knowledge—theoretical and practical—can be divorced. But if we do not need an account of theoretical knowledge to explain the majority of animal skills and abilities, why invoke concepts, models, propositional reasoning—declarative representations more generally—to explain the majority of human action?

There are really three issues here which it is wise to distinguish. First, there is the question of what someone who wishes to explain a system—say, the designer of an intelligent system—must know in order to have proper understanding of its behaviour. Must he have an explicit theory of liquid behaviour in order to understand and design competent systems? If I am right in my interpretation of the doctrine of mobotics, pursuit of such theories is fine as an intellectual pastime but unnecessary for the business of making mobots. It is not evident what *practical* value formal theories of naive physical, social, geometrical, mechanical knowledge can possibly have for experienced mobot makers.

Second, there is the question of whether declarative representations, even if these are not truly concept-based declaratives, are required for intelligent control of activity.<sup>3</sup> Not all declarative representations that appear in the

<sup>3</sup> A declarative is not an easy entity to define. For my purposes I will assume that if information is encoded in a state, structure or process in a form that can be interpreted in a model-theoretic

course of a computation are conceptual. When a vision system creates intermediate representations, such as edges, texture fields, depth gradients, we need not suppose that it has concepts of these entities in the full-blooded manner in which I defined conceptual representations earlier, that is, as being subjects or objects of predication. Information is certainly being represented explicitly, but it is not the sort of information that can be used in thought; its significance is internal to the specific phase of visual processing taking place at that moment. Thus it cannot be shunted off to a long-term memory system because the representation is in the language of early vision. It fails to qualify as a predicate, since it is not predicable of anything outside its current context. The agent does not know its satisfaction conditions.

Brooks' stand on the need for these intermediate representations in a theory of intelligent action is less clear. One difficulty is that he does not explicitly distinguish representations that are *non-conceptual* declaratives from those that are *conceptual* declaratives. Consequently, much of the rhetoric that, in my opinion, is properly directed against conceptual declaratives is phrased in a manner that makes it apply to declarative representation more universally. Thus he deems it good design philosophy to avoid at all costs extracting higher visual properties such as depth maps, 3D sketches, and most particularly, scene parsings. Mobots are constructed by linking small state FSM's that sample busses with tiny probes, e.g. 10 or 20 bits. The assumption is that this approach will scale up—that a mobot can gain robustness in performance by overlaying more and more specialized mechanisms, without ever having to design fairly general vision systems that might extract edges or higher visual properties. Accordingly, although some intermediate representations are inevitable—the readings of tiny probes—more general intermediate representations are outlawed even if some of these are non-conceptual.

Finally, there is the question of names and predicates. On these representations Brooks' position is unambiguous: declarative representations of individuals and properties is positively pernicious for efficient robotics. Flexible activity is possible without much (any) processing that involves drawing inferences, retrieving similar cases from memory, matching and comparing representations and so on. In virtually all cases these computations are complex, frail, prone to bottlenecks and they make false assumptions about the sparseness of real world attributes.

I will have something to say about all these forms of representation. It seems to me that there is no escaping the fact that intelligent systems often frame or pose problems to themselves in a certain way, that they search through some explicit hypothesis space at times, and that they have a memory that contains

---

semantics, then it is declarative. Declaratives release their information upon being *read*, whereas procedures release their information upon being *run*. But clearly, there is no straightforward way of defining the difference between declaratives and procedures since in certain programming languages both can serve as first class objects. For related discussions see [7].

encoded propositions or frames or some other structured symbol, and that part of intelligence consists in knowing how to find the structures in memory that might be helpful in a task and putting those structures to use. Usually these processes make sense only if we assume that the creature has conceptual representations; but occasionally we can view them as involving intermediate representations alone. I believe, moreover, that there are clearly times when as designers we need an adequate domain theory to construct robots in a principled fashion. Accordingly, I will argue that all three forms of representation are necessary for an adequate science of robotics. But equally I think we should appreciate how far we can get *without* such representations. This is the virtue of Brooks' alternative theory of action.

## 2. An alternative theory of action

We may usefully itemize the core ideas underlying this alternative theory of action as follows:

- (1) Behaviour can be partitioned into task-oriented activities or skills, such as walking, running, navigating, collecting cans, vacuuming, chopping vegetables, each of which has its own sensing and control requirements which can be run in parallel with others.<sup>4</sup>
- (2) There is a partial ordering of the complexity of activities such that an entire creature, even one of substantial complexity, can be built incrementally by first building reliable lower-level behavioural skills and then adding more complex skills on top in a gradual manner.<sup>5</sup>
- (3) There is more information available in the world for regulating task-oriented activities than previously appreciated; hence virtually no behavioural skill requires maintaining a world model.<sup>6</sup> If you treat the world as external memory you can retrieve the information you require through perception.
- (4) Only a fraction of the world must be sampled to detect this task-relevant

<sup>4</sup> Many ethologists regard compiling an *ethogram* to be the first step in the description of a species. An ethogram is a behavioural vocabulary of a species which lists all the basic types of behaviours an organism can perform. These behaviours are unit-like in that they can be performed in sequences. Brooks' notion, it seems to me, departs from this more classical notion in being more task oriented. Thus, an activity may be a controlled collection of simpler activities, grouped together by their common purpose of, say, grasping coke cans. Ethologists too look for the function of activities and cluster more basic behavioural units together, but their definition of function is strongly tied to the concept of evolutionary adaption. See [13].

<sup>5</sup> Cf. [3]. Brooks diverges from Gallistel in treating the interaction of activities to be often more complex than that found in simple hierarchies.

<sup>6</sup> This idea I take to be Gibson's principle contribution to the study of sensory systems. Cf. [4, 5]. Where Brooks departs from Gibson on this point is in viewing the process as akin to information retrieval. Gibson supposes that the information is directly picked up.

information. Smart perception can index into the world cleverly, extracting exactly what is needed for task control without solving the general vision problem.<sup>7</sup>

- (5) the hardest problems of intelligent action are related to the control issues involved in coordinating the various behavioural abilities so that the world itself and a predetermined dominance or preference ordering will be sufficient to decide which activity layer has its moment in the sun.

In short, the theme of this alternative theory is that representation can be exchanged for control. If a creature knows where to look and when to look, and knows what activities to activate and deactivate, then it can approximate arbitrarily rational agents.

To take a rather simple example consider an insect which feeds off of sugar, and lives in an environment of wily but slow predators. Such a creature must be able to sense sugars or the probability of sugars at a small distance, “Feed” on those sugars when possible, “Move”, in a specified direction, “Run Away” when it gets too close to certain objects—particularly predators, “Stop Short” if it is about to hit an object directly in front of it, and be able to perform compounds of these low level abilities such as “Wander” so that it might improve its probability of finding food, “Avoid Obstacles” and “Follow Freeways” so that it may move through irregular terrain or flee predators without stumbling. Each of these activities is tuned to certain environmental conditions, such that the activity is turned on or off, amplified or diminished according to locally detectable conditions in conjunction with the internal switching circuitry. If all works well, the net effect is that as the world changes, either because the robot itself is moving through it, or because of external events, the robot will behave as if it is choosing between many goals. Sometimes it runs, sometimes it wanders, sometimes it feeds.

Obviously, the trick in making a mobot behave in a way that looks like it is choosing between many goals without it explicitly predicting the effects which the various behaviours would have on the world, is to design the right pattern of control into the circuitry. Certain pathways will carry messages which dominate the normal input to a module or which suppress the normal output. Accordingly, one goal of research is to find a way of minimizing the amount of this control. Each FSM should be tuned to the *right* stimuli so as to let the world force choice whenever possible.

Thus, for example, when the senses register a looming stimulus, the Stop Short module, takes command. Stop short was primed; it was in a state which acts on a looming stimulus and is hooked up to output so that its signal overrides any others that may also be transmitted. Similarly, if a system were

<sup>7</sup> Again compare [4, 5]. In Gibson’s view the senses are not passive receptors of information; they are active seeking mechanisms, searching out the information—often minimal information—required for effective action and avoidance of physical harm.



on a coke can collecting mission, the Move Hand module might take over as soon as the system sensed a halt in optical flow and a streak of red. A complex cooperative behaviour might emerge, therefore, simply because each component activity becomes primed for particular changes in the state of the world that matter to it. Hence, coordination is achieved automatically without posting requests on some central blackboard or relying on some active arbitrator to pass control to slave activities because the preference relations among activities have been built into the switching network of the system.

Let us call behaviour that is controlled by the situation in this way, *situation-determined behaviour*. Situation-determined behaviour can be considerably more complex than the stimulus driven behaviour found in behaviourist theory. For instance, humans, when putting together jig saw puzzles, may be said to be situationally determined if there is enough joint constraint in the tiles and assembled layout to ensure that they can complete the puzzle without wasted placements. No behaviourist theory can explain jig saw performance, however, because there is no readily definable set of structural properties—i.e. stimulus conditions—that are the causes of jig saw placements—i.e. responses. The agent is too active in perceptually questioning the world. On two confrontations with the same world the same agent might perceive different situations as present because it asked a different set of perceptual questions. These questions are a function of the state of the agent and its most recent interactions with the world.

We can say that jig saw puzzles are perceptually hard but intellectually simple. The actions are intentional but under perceptual rather than conceptual guidance. Thus it is the eye, not the thinking center, which must be trained to look for the salient corners that differentiate tiles and signal proper fit. It is a problem of perceptual search.

Viewing situation-determined behaviour to be a solution to a perceptual problem points out several worthwhile aspects of situationally determined tasks.

First, there is enough local constraint in the world to “determine” successful placement despite there being several tiles that can be successfully played at any moment. In a sense each move is underdetermined, hence no deterministic behaviourist theory can explain placement behaviour. Nonetheless, given a tile and an existing layout, the situation wholly determines whether or not the tile can be correctly placed at that time and where. There is no need to check downstream effects. In the jig saw game, successful placements are additive. Good moves do not interact hostilely with other good moves. There are no traps, dead ends, or loops that may stymie a player. The situation contains enough information to pre-empt the need for lookahead. This is the main point of assumption three.

Second, the perceptual problem is tractible in the sense that only a fraction of the visible world state must be canvassed to determine where to move. The point of sensing is to provide enough information to permit a creature to

choose between the actions it can perform next. In the case of jig saw it is conceivable that to solve the puzzle one must identify the overall shape of all the pieces first. If this were true, a jig saw puzzle would be a tedious game indeed, for either it would require colossal visual processing each move, or it would require tremendous visual memory of shapes. How much easier if complete shape identification is unnecessary.

Is this possible? Is it possible to decide which tile to place next by using a strategy of visually questioning the board that does not require computing the overall shape of each tile? The question is important because if perceptual questioning can be confined to simple features there will be no need for higher level intermediate representations.

Imagine a case where a player cannot decide which of five tiles to play in a particular opening. Each tile seems like it might be a proper fit, but it is hard to tell. An obvious aid to the problem is to have the player try to fit one of the tiles in the opening to let the world highlight the crucial feature that differentiates the proper tile from the near misses. The function of this test move is to focus the player's attention on the situationally *salient* features of the tiles. It is to identify the crucial differentiating features. Now a true expert of the game might not need this help; his perceptual system may be so tuned to the task that he can home right in on the relevant differentiating features. If so, this possibility affirms the point of assumption four: that if one knows what to look for, there is a fairly local feature which correlates with correct moves. Not only does the situation contain enough local constraint to determine good moves, these constraints are highly specific to the task and learnable.

It is worth dwelling on this issue for it emphasises the truth of assumption five: that control is the hard problem, and the methodological importance of assumption one: that behaviour can be partitioned into task oriented activities. These, I take it, are the backbone of this alternative approach to action.

It is standard in decision theory to treat perception as a bounded resource that must be guided in order to be used to its fullest. The problem which decision-theoretic accounts encounter, however, is that to know what question it is best to ask next, or which test it is best to perform next, the agent must know all the sources of information available now and in the future, all the decisions that might be taken now and in the future, their consequences, utilities, etc. To achieve optimality is clearly impossible in practice, for it requires knowing where you are most likely to get the information you want before you know exactly which decisions you must make. If one restricts the horizon of one's decisions to specific task-oriented activities the problem is simpler. Must I halt now? Can I proceed in that direction? Is there a predator nearby? For each of these questions there may be a straightforward test which is decisive, or nearly decisive, or indicative of what to test next. Once again the question is whether the test (or perceptual query) is computationally cheap.

In a situationally determined context such questions are necessarily cheap.

The environment can be factored into a set of partial states or *indicators* which correlate well with the presence or absence of the larger environmental factors which affect task performance. Thus for a robot whose environment contains doors with right angles it may be possible to discover an invariant *microfeature* of doors which under normal conditions can be seen from all angles. Relative to door entering activity, this invariant may be all that need be sought. Moreover, it may be simple—a top right and bottom left corner in suitable opposition for example. This fraction of doorness is sufficient for door recognition in this environment, as long as the robot remains upright, as long as no new doors are introduced, and so forth. It correlates with all and only doors. Consequently, one of the hardest problems for mobot designers is to discover these indicators, and the perceptual queries that best identify them. For each activity the designer must determine which possible indicators correlate well with the likelihood of success or failure of the activity given the current state of the world. This is a hard problem for most activities. But the key point is that without the assumption that behaviour can be partitioned into task-oriented activities, it would be impossible to discover these indicators at all.

This introduces the third and final respect situationally determined tasks are illustrative of the alternative theory of action: what is most salient in the environment is usually discernable and economically detectable from the *agent's perspective*. Most task indicators are egocentrically definable. This is a crucial factor in deciding how much of activity can be intelligently controlled without concepts because concepts are often held to be non-egocentric, public or quasi-public entities.

Developmental psychologists draw a distinction between the *egocentric space* of an agent and the *public space*, which as observers we see the agent performing in. The distinction is intuitive. In egocentric space, the agent is always at the spatio-temporal origin of its world. It sees the environment from its own perspective. Indexical terms such as beside-me, to my right, in front, on top, nearby, occluded-right-now, are all well defined, and depend essentially on the agent's location. They shift as it moves about.

In public space, by contrast, the world is understood almost as if viewed from nowhere. If the agent is included in the world at all it is included objectively as another entity in relation with objects in the world. This is done to facilitate useful generalization. Two people can see the same ball; a ball remains the same ball despite its currently being outside the agent's visual field; and it remains beside a companion ball whether partly occluded or not. Because we can count on the permanence of objects and on a consensual understanding of space-time we can usefully organize our experience of the world by appeal to public objects, public space, and public time. We can describe actions and strategies in a manner which allows people in different circumstances to use them; and we can talk about consequences of actions as if we were not there to see them. Thus, in describing the action of lifting a box

five feet in the air it is usually irrelevant whether the agent approaches the box from the right or left. Where the agent was positioned in the situation is less important than what it did to make the box go up. This can be stated in terms of the lawful changes which the objects in the environment undergo.

In the classical theory of action, the beliefs that were thought causally important in determining action were stated in the language of public objects and properties. Actions were defined as situation-action rules—transformations between pre and postconditions, and were understood as transformations over public states.

The practice of enumerating the troubles of situation action rules based on public concepts is by now a familiar pastime in discussions of AI planning. It is therefore regarded a virtue of the situationally determined account that the indicators which matter to situationally determined task performance are definable from an egocentric perspective.

J.J. Gibson, for example, argued at length that the genuine environment of action is not a world of objects and objective relations but a world of surfaces and textural flows as seen by the agent. Gibson, in his ecological approach to perception, emphasised that action and perception are not distinct processes. Animals and people do not passively perceive the world. They move about in it actively, picking up the information needed to guide their movement. This information is always available in an egocentric form, because as a result of the interlocking between perception and action, certain egocentric invariants emerge. Flies can find landing sites by detecting wiping of texture in the optic flow [17, pp. 215–218], chicks and babies can avoid precipices by detecting motion parallax and texture gradients [17, pp. 234–235]. These invariants can be picked up early. They do not require the level of visual processing involved in creating a full 3-D representation. The same it seems holds for most situationally determined tasks: the indicators which matter can be gleaned by relatively early attention to egocentric invariants, or properties.

The upshot is that for situationally determined activity, perception, particularly egocentric perception, rather than conceptual reasoning is the determining factor of success. This holds because there is a reliable correlation between egocentrically noticeable properties of the environment and actions that are effective.<sup>8</sup>

Now, from both a scientific and engineering standpoint nothing but good can come from exploring in silicon and metal how much of intelligent activity can be duplicated following the principles of this alternative theory of action. Until

<sup>8</sup> The conception of situation determinedness I offer is stipulative. Others can be proposed. For instance, one could propose that a context situationally determines an action for an agent if the situation in conjunction with the inner state of the agent determines what he will do next. But to stretch the definition in that direction is to give up the distinction between situation determinedness and determinedness simpliciter. Agents are usually determined by the union of mental state and local environment. The noteworthy condition of true situation determinedness is that reasoning is not required for action.

we construct creatures which can have hundreds of procedures turned on and waiting, we cannot know how effective the world might possibly be in deciding the sequence of the procedures to use. There may be far more indicators in the world that are able to bias performance than we would have dreamed possible prior to designing creatures to run in the real world.

Nevertheless, as with most nascent areas of AI, it is easy to see early results as compelling evidence for strong conclusions. In Brooks' case, the success of this design strategy for simple insect-like creatures is meant to justify a host of methodological directives and criticisms for design strategies of far more complex creatures and behaviours.

Accordingly, let us consider some of the limits of situationally determined actions, and the attendant reasons higher-level creatures are likely to use concepts and representations in action, perception, and control.

### 3. The limits of situationally determined action

Situationally determined activity has a real chance of success only if there are enough egocentrically perceptible cues available. There must be sufficient local constraint in the environment to determine actions that have no irreversibly bad downstream effects. Only then will it be *unnecessary* for the creature to represent alternative courses of actions to determine which ones lead to dead ends, traps, loops, or idle wandering.

From this it follows that if a task requires knowledge about the world that must be obtained by reasoning or by recall, rather than by perception, it cannot be classified as situation determined. Principle candidates for such tasks are:

- Activities which involve other agents, since these often will require making *predictions* of their behaviour.
- Activities which require response to events and actions beyond the creature's current sensory limits, such as taking precautions now for the future, avoiding future dangers, contingencies, idle wandering—the standard motive for internal lookahead.
- Activities which require understanding a situation from an objective perspective such as when a new recipe is followed, advice is assimilated, a strategy from one context is generalized or adapted to the current situation. All these require some measure of conceptualization.
- Activities which require some amount of problem solving, such as when we wrap a package and have to determine how many sheets of paper to use, or when we rearrange items in a refrigerator to make room for a new pot.
- Activities which are creative and hence stimulus free, such as much of language use, musical performance, mime, self-amusement.

These activities are not isolated episodes in a normal human life. Admittedly, they are all based on an underlying set of reliable control systems; but these control systems are not sufficient themselves to organize the *global* structure of the activity.

Thus, to prepare tea requires coordinating both global and local constraints. At the global level, teamakers must be sensitive to the number of people they are serving, ensuring there is enough water, tea, cups, saucers and biscuits. Once these items are laid out more mobot-like control systems may take over, pouring the water, stirring etc. But the initial resource allocation problems are hard to solve. Animals are notoriously ineffective at them. Moreover, can we expect mobots to intelligently arrange plates on the tray? Arrangement or bin packing requires attention to a number of non-local factors, such as how many items remain to be placed, how well they can be expected to stack, and how stable the overall configuration must be, given the path to the parlor. Anticipation of the future is required. Hence, whenever global considerations enter the control of action, the creature must either be pre-tuned to the future, or it must be able to call on memories, reason about contingencies, ask for advice, and so forth.

In short, the world of human action regularly falls short of total situation determinedness. Most of our life is spent managing locally constrained choice.<sup>9</sup> It is at this management level that we can best appreciate the virtue of concepts and representations.

#### 4. The virtues of concepts and representations

Concepts are involved in the management of action because they serve at least three organizing functions in cognitive economies. At the *perceptual* level, concepts unify perceptions into equivalence classes. An agent possessing the concept of a dog, for instance, should be able to recognize dogs from different points of view. A dog is an invariant across images. It is also an object for the visual system in the sense that the visual field will be segmented into dog images and non-dog images, offering whatever attentional mechanisms reside in the perceptual system to be directed at specifics of dog images. Accordingly, one aspect of saying that a creature has a concept of dog is to say that he or she can identify dogs perceptually. This means that a vast array of perceptual circumstances can be simplified and reasoned about economically, and that a host of perceptual mechanisms are coordinated around the perceptual object dog.

At a more *conceptual* level, concepts license inferences. A dog is *not* identical with the set of its possible appearances. It is a spatially extended temporally enduring entity that can enter into causal relations with other

<sup>9</sup> For an outline of the virtues and problems with local choice, see [6].

objects. It is a possible subject of predication. Hence much of what is true of other objects—other possible subjects of predication—will be true of dogs. Many of these *inheritable* truths constitute the *presuppositions* which a creature able to have beliefs and thoughts about dogs will hold. In thinking about dogs, then, the creature will have in mind an entity that is alive, breathes, normally has four legs, and so on. This information is readily accessible, but of course need not be conscious. It enables the creature, however, to intelligently respond to invisible properties of dogs [18]. Thus, a child may resist striking a dog because it knows it would hurt the dog, despite the fact that the property of being open to hurt is not a perceptually present property of dogs.<sup>10</sup>

At a *linguistic* level, a concept is the meaning of a term. To know the meaning of ‘dog’ in English is to have the concept of dog, and to know that the English word signifies that concept. The concept dog is a semantic value; in the Fregean system, when coupled with another appropriate semantic value it constitutes a proposition, or truth bearer.

Now, when an agent has a concept it can do things and think thoughts it could not otherwise. As developmentalists have pointed out, once a child has the concept of an object, it can know that the same object can present different appearances. It can decide that what looks like a dog is not really a dog, but a misleading image of a bear. It can infer that your image of this dog is different than mine, but that we both know it is the same dog [14]. And it can infer that dogs feel pain because they are alive. Concept users understand a great deal about their environment when they conceptualize it.

There can be no doubt that the skills we identify with possession of concepts are of great value for certain forms of intelligent behaviour. But how widespread is this behaviour? Can we approximate most intelligent behaviour without concepts? This is Brooks’ challenge.

One of the most important uses of concepts is to organize memory. Whether or not a system has limited memory, it has a need to index memories in a manner that facilitates recall. In action management, an effective creature will benefit from its performances in the past. It will remember dangers, failures, helpful tricks, useful sub-goals. It may recall unexpected consequences of its previous performances. These memory accesses need not be conscious. Nor need they be complete. Someone describing a particular pet dog may not have accessed all the related information he or she knows about the animal. Some information lies untouched. But this information is *primed* in the sense that retrieving that related information in the near future takes less time than had the topic never been discussed [10].

<sup>10</sup> Gibson argues that sentience and the like are perceivable properties of an animal. But in his system, there is almost no action-relevant property that is not perceivable. Thus, post boxes have the perceivable property of affording letter posting. Just how many non-obvious properties can be perceived or registered is a deep question which the alternative theory of action raises. But I think we may safely say that the line must be drawn short of *all* action-relevant properties. Default reasoning will be valuable for these.

In general, if memory is deemed useful for an action it is less plausible to call that action situation determined. The strong empirical claim Brooks makes, then, is that to access organized memories takes too long for most actions. Given the pressing exigencies of the real world, there is no time to retrieve and reason with conceptualized information.

Short of knowing the actual time a particular creature takes for accessing memory it is impossible to argue for or against Brooks' thesis. But we can have intuitions. For instance, in tasks where the time to react is very short, recall will be costly; some recall may be possible but it must be directly applicable to tasks without much reasoning.

Yet how much of life is reactive? In driving home, for instance, I am often on autopilot, but I do come to genuine choice points, where I must decide whether to take, for example, Torrey Pines Blvd. or the highway [9]. In assessing my options I have conceptualized the possibilities. My preferences are over world states, conceived sometimes as my possible future experiences, sometimes as objective states of the world. My response is not reactive, it is thoughtful. My decision depends on how I think of the future.

The point, here, is that if I wish to accommodate my present action to events, objects or actions that are distant in time and space, I shall have to anticipate them now. A perception-driven creature can only anticipate the future if there is evidence of the future in its present. With memory, however, it can remember that *Y* follows *X*, and so coordinate its actions to a broader environment than that perceptually given.

If the future is a simple function—possibly Markovian—of the perceptually present, a system of linked FMSs might cope with simple futures. FSMs have state and so can encode information about the future. But the future they encode cannot be complicated or complexly branching. When the future is complex simple FSMs will be unreliable. For it is inevitable that one set of future states which correlates with the present will recommend action in one direction, while others, also correlating with the present, will recommend action in other directions. How is choice to be made? Prudent decision-making in such situations requires an all things considered approach. It requires balancing the recommendations, and setting a course of action which may involve the future coordination of a complex network of acts. It is hard to see how this could be done without the simplifications of the world which conceptualization gives us.

This capacity to accommodate the future ties in with a second ability that comes naturally to systems with concepts: to *take advice*, and to learn by imitation [11]. It is characteristic of humans that if they are in the middle of a task that has several parts they can make use of hints or suggestions. These need not be linguistic clues because often it is enough if someone shows us manually what to do, or shows us a technique or move that is similar to the one we must perform. New ideas can bias performance. This implies that whatever



means we have for controlling our behaviour it must be *permeable* to new information.

What makes this permeability hard to capture in models built on the alternative theory of action is that hints and advice are often offered from a non-egocentric point of view. Hence there is no reason why a hint or a suggestion should be meaningful input to the home system. This is not to say that concept using agents have no trouble assimilating advice. Advice can be more or less ready for use. Hints phrased from an objective, perspectiveless orientation may be hard to put into practice by agents wholly immersed in their own perspective. But some form of this translation problem is solved every time we understand that other agents see the world differently.

Advice taking also has a sensory side. Suppose I am told that a friend has been in a car accident and broken his legs. I now expect to see a person on crutches. Hence I can recognize him at a distance, and not be deceived by appearances.

This adaptation of future expectations is impossible to explain without concepts. There must be some device in an agent which functions like an indexed long-term memory of objects which keeps track of changes and which allows it to update expectations about the behaviour and appearance of objects in a controlled manner. Somehow it must be able to systematically change the attributes that an object may be assumed to inherit or possess by default.

This same idea applies to behaviour in strategic environments where the effectiveness of an action often depends on the interpretation which other agents impose on it. To take advantage of these dependencies requires knowing the interpretations of others. It presupposes that the agent can understand its opponents or colleagues as systems whose behaviour is a partial function of its current and future behaviour. It is hard to see how the effects of this recursive interpretation can be achieved without conceptual representations. First, it will require understanding other agents as agents in a common world playing in a common field, hence operating in a public domain rather than an egocentric one. Second, it will require understanding them counterfactually, in terms of how they might interpret the agent if it were to do *X* instead of *Y*.

## **5. The need for representation in a theory of perception**

I have been describing the importance of representation, particularly conceptual representation, for a theory of action: there are limits to how subtly a system can act if it is entirely situation determined. The ability to frame and test hypotheses about the future and about other agents' behaviour is essential for survival in human style environments. But there are equally strong reasons to suppose that representation is important for a theory of perception.

The field of computational vision has done much to explode the myth that vision is strongly under the influence of expectations, memory and inference *at early stages of processing*. But there are few who believe that extraction and identification of shapes can occur without at least some models of shape in memory.

Shape models are not the same as concepts. They constitute equivalence classes of perceptions, but they carry no implications of objecthood. Accordingly, it is not until *scene parsing*, where items are identified in the visual scene and conceptualized as organized, that we are justified in claiming that a system imposes concepts.

Brooks' position on this is, I believe, much like Gibson's. Organisms detect without inference or reconstruction those properties of things which they need to achieve their goals. In general, this will not require visual processing to the point of 3-D shape recognition, and certainly never to the point of scene recognition. Brooks believes it will never require more processing than that required for a viewer-centered representation of objects; and most often the information needs of action can be fulfilled by special purpose detectors.<sup>11</sup>

The trouble with this view, however, is that it doesn't make clear how some of the interesting visual properties that need detecting can actually be accomplished.

For instance, how can an object seen from one orientation be recognized as the same object when viewed from another orientation. This is necessary for backtracking in the world.

A key assumption of the alternative theory of action is that the world is benign: ineffective moves can be tolerated because seldom is it the case that they lead to irrecoverable states. If an ineffective move is made, the creature can either just continue from its new position, or backtrack in the world—provided, of course, the creature could remember its path. But this is the problem: how can the creature recognize where it's been if it cannot recognize the same object from both front and back?

To cope with the memory demands of search in the world humans trail-blaze—they leave markers of where they've been. They can then reuse pre-existing procedures, such as, go to the first visible landmark that you haven't already visited. But there are obviously environments where trailblazing is impractical. In such cases, a snapshot of the relevant portion of the world state is required. This is akin to episodic memory. But the episode is not recorded as a simple snapshot. For if the creature is to use the snapshot, it

<sup>11</sup> The justification for this claim, it seems to me, is the Gibsonian theory that perception involves active exploration by an organism. Instead of asking how an organism infers the structure of its environment from the pattern of activations on its receptive field, we ask how the organism picks up what it needs to know by moving through the "ambient optical array" containing the information. By dynamically sampling this optical array the creature is supposed to be able to get whatever information it needs about objects and layouts to fulfill its objectives.

must record the scene in a perspective neutral way. Otherwise the image will be the wrong orientation to resemble what the agent sees as it backtracks. Records more abstract than agent-oriented images are required.

As a rule animals do not rely on such sophisticated perception and perceptual recall. I do not know whether they do much controlled search in the world but they can easily determine whether they have visited a spot by scent. They are locally driven machines. Such is not always the case with humans. Early in our evolution we traded olfactory prowess for visual intelligence, with the attendant advantage that we now can determine whether we have visited a spot without sniffing it at close range.

In the same way our abilities to handle complex objects without practice also feed off of our advanced visual intelligence. Funny shapes require funny grasps. Unevenly distributed masses require prudent grips, and heavy objects require appropriate force. We don't approach a weighty textbook the way we do a paper container.

How do we determine our approach to these objects without performing enough computation to determine (1) the center of mass of the object, (2) a set of points or regions of opposition, and (3) the texture of the surface so that we can make a good guess about the object's material and hence its weight? One possibility is that we use a vast table look up which associates shapes with grasps. Yet grasps vary with hardness, smoothness and weight too. These too will have to be built into a table. The net effect would be a table of enormous complexity, Accordingly, the obvious alternative is to invoke intermediate representations and compute solutions on-line. These intermediate representations are not conceptual; they represent properties that are relevant for grasping. But they do emphasise that perception must solve big problems, and frequently in a way that is general. At the very least, the complexity of vision argues for the need to analyse the problem at a general level, if only to construct the look up table.

## 6. The need for representations in a theory of learning

Skill generalization is a further area that may pose problems for the robotics approach. One reason we currently believe that representations—of both the conceptual and non-conceptual variety—are vital to learning is that we know of no other way of *simplifying* situations so that what is similar between situations is easy to note. Obviously we want systems that can apply existing knowledge to new tasks, systems that can *transfer* expertise. Unless mobots can generalize stimuli they will have to be reprogrammed to perform what are essentially the same tasks on slightly different objects. If a mobot can pick up a coke can it should be able to pick up a coffee cup.

The trouble with coordinated FSMs is that they are each carefully tuned to

the particular properties of specific tasks. If a hand-control system that regulates coke can grasping focusses on specific coke can properties—a red streak, a shiny “circular” surface—then it is not easy to see how that control system can be used for grasping coffee cups. The issue is not whether some of the constituent modules of the coke grasping reflex can be used; it is, rather, that one or several FSMs depend on specific perceptual microfeatures of coke cans.

Now sometimes this task specificity is justified. Perhaps the ability to pick up cans is different than the ability to pick up cups with handles, or to pick up flyswatters. But how are we to know this? The mobot engineering philosophy is to test out designs to see what is common across tasks. If coke can grasping does not work on coffee cups then add extra control layers. This same process will continue until someone decides that the grasping system is too complex. At that moment, a redesigned system will be constructed that simplifies the system on the basis of what has been learned.

There is nothing objectionable in this familiar engineering approach. But it is based on two rather strong assumptions. First, that it is imprudent to pursue prior analysis because one cannot know what are the natural groupings of grasping until one knows how a grasper relying on microfeatures might work. Second, generalization of the grasping system can be achieved without extracting higher-order structural properties.

The virtue of representations, both intermediate and conceptual, is that they let us see similarity in superficial disparity. Two objects may differ in almost all their microfeatures, but be deemed relevantly similar at a more abstract level. Thus the generalization problem: is  $X$  relevantly similar to  $Y$ , is easy to solve if we have characterized  $X$  and  $Y$  in a relatively sparse feature space, but hard in a dense lower level space. The questions: “What are the task-relevant properties common across objects?”, “What properties of objects must be made explicit to simplify control?” are what the study of representation is all about. Only in a rhetorical sense, then, can moboticists contend that they abjure representations.

## 7. The need for representations in a theory of control

Any system that is to forever substitute control for representation must be able to:

- (1) cope with increasingly *complex desire systems*; and
- (2) resourcefully recover from failure.

If we are ever to build the much awaited household robot, it will have to be designed with both these abilities. I think designers, however, will have an

impossibly difficult time building in such abilities without using conceptual representations. Consider desire systems first.

Any household robot worth its salt must be able to make us a midnight snack. Before I rely on such a device, though, I will want it to be able to operate with complex goal systems. I want it to be able to balance competing desiderata when it reaches the fridge. The trouble is that mobots, as we envisage them today, operate with an impoverished goal system and so are limited in their performance.

Basically, a mobot-inspired creature would work on what might be called the *refrigerator model of desire*. Open up the refrigerator, look in, and let the contents and some simple capacitor notion of wants decide what to select. This has the nice property that the creature doesn't have to have a fixed idea of what to select in advance, it can let the possibilities decide for it. Thus the choice problem is solved in the simplest way possible: thirst is valued more than appearance and less than gut hunger. If hunger has been largely satisfied so that the *capacitor* measuring hunger is low, then thirst prevails, and so forth.

The problem with this approach is that if the creature is to cope with many desires it is not at all clear how a ranking can be provided in so simple a fashion. Given a choice between filboid sludge for breakfast and taking a chit for a five course lunch at Panache, I'll choose the chit. My top level goal may be the allayment of hunger but how I subgoal may be complex and sensitive to many desiderata, such as taste, appearance, comfort, diet, to name a few. Desires do not just compete in a simple winner take all fashion, because in complex desires system it is not possible to rank desires according to a small number of lexicographically ordered dimensions. There are real limits to the capacitor concept of desire.

What this means is that when desire systems get large there must be some type of desire management, such as deliberation, weighing competing benefits and costs, and so on. This applies whether the mobot is out there in the field doing my bidding or it is an autonomous creature with its own set of desires. Without representation, desires lack the *modularity* to be reasoned about, or even flexibly assembled. If the representations are not conceptual they will not be about enduring states of the world that can be entertained and reasoned over. Conceptual representation is necessary for desire management. Without desire management, mobots will be little more than insects or lower animals.

Now consider the value of belief systems for flexible control.

One of the lessons learned from first generation expert systems is that unless an agent has some understanding of why certain if-then rules work it will be unable to respond flexibly when it finds that it has no rule that will apply in its current context or when it discovers that one of its rules fails to have the desired outcome. Models of underlying relations are important.

To take a simple example, if a radio repairman is unable to fix a broken set by standard tweaks, he will try to discover by reasoning the cause of the

system's observed behaviour. The customary imputation is that experts have levels of understanding: for standard cases they operate with an abstracted representation of a device or possibly a set of precompiled procedures. But when necessary they can reflect on the rationale of those procedures, on why they work in certain cases and why they may fail in others; they may even reason from first principles.

Now in a typical robotic system there can be no more than a small number of fixes one could try in problematic situations. In some cases this strategy will work. It achieves a type of robustness: a system that announces it doesn't know what to do is more resilient than a system which is determined to try something, no matter how ham fisted.

The problem is that if one wants to do better than giving up, the fix has to be appropriate to the case. The lesson of second generation expert systems is that such fixes require being selective about choosing what additional information to seek. This is a hard problem and requires a fairly deep understanding of the situation. But it is unclear that Rod's robots can have this kind of understanding without having the equivalent of models of the domain. How can a system whose response to failure is to try a simpler behaviour achieve this innovative resilience?

The reason this is a hard problem is that the response the system has to make varies across situations. The same behavioural failure can require different fixes. That means that at a superficial level there is not enough information to determine an action. The system must conjecture and test. Since the range of conjecture is vast, the state space of FSM's would have to be correspondingly vast. But once again this vast space would not be systematically generated, except, of course, by the designer who used concepts and compiled his answer to hide the systematicity.

## 8. Conclusion

I have been arguing that although AI can substantially benefit from greater attention to the richness of perceptual information, this richness will never replace the need for internal representations. Any plausible household robot, even one that does not have the full improvisational skills of a human, will have to rely on symbolic representations at least sometimes.

This is especially obvious if we consider how language use can accelerate evolution. No one understands how closely language is tied to vision, or how closely it is tied to reasoning. But it is widely recognized that once language is acquired certain forms of learning and reasoning become possible and certain other forms are accelerated.

For instance, with linguistic communication comes the possibility of identifying and storing very precise information. Without language it is hard to draw

someone's attention to a particular perceptual fact; for it is difficult to specify *which* condition of the situation is the salient condition. The problem becomes exponentially more difficult if the condition is abstract. Imagine trying to draw someone's attention to the *bluntness* of a particular pin.

Similarly, once arbitrary amounts of knowledge can be stored and passed on from generation to generation, we can accelerate the rate at which our abilities grow by learning from the lessons of others. Cultural transmission of information is much faster than genetic transmission of information. This might explain the shockingly brief time it took for man to develop his higher mental skills when compared with the great length of time evolution took to develop sophisticated motor skills.

Thus, is 97% of life concept-free? The answer depends on how you count abilities. If an ability is defined relative to an environment, then the richness of the human environment suggests that there are wildly more tasks that can be done in the human world, than in environments characteristic of less language, norm-ridden creatures. Once language-like communication emerged the rate at which we could acquire new abilities rose dramatically because we could identify, create, and teach new abilities.

The magic that made this take-off possible was the ability to remember facts, rules, norms, strategies and the like. With specific cases in mind we could avoid pitfalls, with norms and rules we could cleave to the conservative but safe path; with strategies and plans we could find our way where random search would be disastrous. And of course with the ability to communicate—which these higher order abilities presuppose—we could also take advice.

These goods seem to flow from the ability to internally represent facts and to reason explicitly. Any theory that asserts that we can get by without conceptual representation will have to *explain away* these goods by showing that they are not necessary for intelligent activity.

### Acknowledgement

In writing this paper I have benefited considerably from insightful discussions with Peter Cudhea, Pattie Maes, Jean Mandler, Eric Saund and Bob Stalnaker.

### References

- [1] R.A. Brooks, Intelligence without representation, *Artif. Intell.* **47** (1991) 139–159, this volume.
- [2] R.A. Brooks, A robust layered control system for a mobile robot, *IEEE J. Rob. Autom.* **2** (1986) 14–23.
- [3] C.R. Gallistel, *The Organization of Action* (Erlbaum, Hillsdale, NJ, 1980).

- [4] J.J. Gibson, *The Senses Considered as Perceptual Systems* (Houghton Mifflin, Boston, MA, 1966).
- [5] J.J. Gibson, *The Ecological Approach to Visual Perception*, (Houghton Mifflin, Boston, MA, 1979).
- [6] D. Kirsh, Managing local choice, in: *Proceedings AAAI Workshop on AI and Rational Choice* (1989).
- [7] D. Kirsh, When is information explicitly represented?, in: P. Hanson, ed., *Information, Content and Meaning* (UBC Press, Vancouver, BC, 1990).
- [8] N.J. Mackintosh, *Conditioning and Associative Learning* (Oxford University Press, Oxford, 1983).
- [9] G. Mandler, *Cognitive Psychology* (Erlbaum, Hillsdale, NJ, 1985).
- [10] G. Mandler, Memory: Conscious and unconscious, in: P.R. Solomon, G.R. Goethals, C.M. Kelley and B.R. Stephens, eds., *Memory: Interdisciplinary Approaches* (Springer, New York, 1989).
- [11] J. Mandler, How to build a baby: on the development of an accessible representational system, *Cogn. Dev.* **3** (1988) 113–136.
- [12] M. Minsky, *The Society of Mind* (Simon and Schuster, New York, 1986).
- [13] V. Reynolds, The origins of a behavioural vocabulary: The case of the rhesus monkey, *J. Theor. Social Behav.* **6** (1976) 105–142.
- [14] M. Siegel, *Knowing Children: Experiments in Conversation and Cognition* (Erlbaum, Hillsdale, NJ, 1990).
- [15] E.S. Spelke, Where perceiving ends and thinking begins: the apprehension of objects in infancy, in: A. Yonas, ed., *Perceptual Development in Infancy* (Erlbaum, Hillsdale, NJ, 1988).
- [16] R. Sternberg, ed., *Advances in the Psychology of Intelligence* **4** (Erlbaum, Hillsdale, NJ, 1988).
- [17] B. Vicki, and P. Green, *Visual Perception, Physiology, Psychology and Ecology* (Erlbaum, Hillsdale, NJ, 1985).
- [18] H.M. Wellman and S.A. Gelman, Children's understanding of the nonobvious, in: R. Sternberg, ed., *Advances in the Psychology of Intelligence* **4** (Erlbaum, Hillsdale, NJ, 1988).