Chunking in Soar. *Machine Learning*, **1**, 11-46.

Riolo, R. L. (1987a). Bucket brigade performance: I. Long sequences of classifiers. *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms* (pp. 184-195). Hillsdale, New Jersey: Lawrence Erlbaum Assoc.

Rivest, R. L. & Schapire, R. E. (1987). A new approach to unsupervised learning in deterministic environments. *Proceedings of the Fourth International Workshop on Machine Learning* (pp. 364-375). Los Altos, CA: Morgan Kaufmann.

Smith, S. J. & Wilson, S. W. (1989). Rosetta: toward a model of learning problems. *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 347-350). San Mateo, CA: Morgan Kaufmann.

Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. *Machine Learning: Proceedings of the Seventh International Conference* (pp. 216-224). Palo Alto, CA: Morgan Kaufmann.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, October, **59**, 433-460. Reprinted in *Computers and Thought*, E. A. Feigenbaum and J. Feldman, eds. (pp. 11-35), 1963.

van Heerden, P. J. (1968). *The Foundation of Empirical Knowledge*. Wassenaar, The Netherlands: Wistik.

Wilson, S. W. (1983). On the retino-cortical mapping. *Int. J. Man-Machine Studies*, **18**, 361-389.

Wilson, S. W. (1985a). Knowledge growth in an artificial animal. *Proceedings of the First International Conference on Genetic Algorithms and Their Applications* (pp. 16-23). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Wilson, S. W. (1985b). Adaptive "cortical" pattern recognition. *Proceedings of the First International Conference on Genetic Algorithms and Their Applications* (pp. 188-196). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

simulates natural light reflected from the objects. Stimuli could be as profuse as desired depending on the variety of the objects, the resolution and spectral sensitivity of the light detectors, and the range of vision. In such environments, the experimental procedure should be alert to the possibility that the system might form effective percepts quite different from, e.g., "tree 90 degrees to the right", (as stimuli are often directly given in Standard AI). Among other things, the percepts would reflect the bias of the animat's needs.

Besides top-down/bottom-up methods, there is another class of techniques for dealing with stimulus profusion. Since the animat has effectors that can change the sensory input (as when the animat moves), he can learn ways of doing so that select information contingently. This ability becomes useful in cooperation with detectors that condense stimuli to different degrees in different parts of the sensory field. Then the animat can move so as always to place the highest-resolution detectors at the point of greatest interest, leaving the rest of the field relatively diffuse.

The focal-peripheral vision of some animals is an example of stimulus condensation and coordinated action that should be investigated using animats. In primates, for instance, the environment is seen in detail in a central region but with falling definition toward the visual periphery [see, e.g., Wilson (1983)]. This permits vast reduction of stimulus detail over most of the visual field, a primitive form of generalization, but the reduction can always be reversed by moving the eyes. Such a move would be triggered by current needs and internal state in combination with a stimulus cue sufficient to get through the peripheral generalization (Wilson, 1985b).

Perception is one of the hardest human abilities to understand. Progress in machine perception has been slow. The animat approach offers a fresh perspective because well-defined experimental mechanisms can be investigated in contexts that retain essential characteristics of real organisms and environments.

5. Summary

This paper has outlined "the animat path to AI", a strategy for progressively understanding intelligence or the relation of mind to brain that differs significantly from Standard AI, and from the natural science approaches to the same problem. The approach is not new, in that examples of prior work exist and are somewhat known. This paper however attempts to bring out the value of the approach, calls for a more systematic effort, and offers some working themes.

Fundamentally, the animat approach advocates maintaining the holism of the situation of real animals in real environments, while progressively but efficiently increasing animat complexity only as necessary. The approach's hypothesis is that this program is feasible, and will ultimately lead to understanding of intelligence, adaptation, and perception at high levels. In support of the program, the paper proposes:

(1) Creation of a theory/taxonomy of environments, based on a "sensory-state machine" formalism;

(2) Establishment of criteria of animat efficiency in terms of need satisfaction and costs;

(3) The hypothesis that efficient animats will have architectures that deal with frequent, important situations by "virtual stimulus-response";

(4) The suggestion that the problem of "stimulus profusion" can be reduced through stimulus-condensing sensoria and contingent action.

References

Barto, A. G. & Sutton, R. S. (1981). Landmark learning: an illustration of associative search. *Biological Cybernetics*, **42**, 1-8.

Geman, S. & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6*, 721-741.

Gould, S. J. (1989). *Wonderful Life*. New York: W. W. Norton.

Grefenstette, J. J. (1988). Credit assignment in rule discovery systems based on genetic algorithms. *Machine Learning* **3**, 225-245.

Grossberg, S. (1987). Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science*, **11**, 23-63.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, **42**(1-3), 335-346.

Holland, J. H. (1986). Escaping brittleness: the possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In R. S. Michalski, J. G. Carbonell & T. M. Mitchell (Eds.), *Machine learning, an artificial intelligence approach. Volume II.* Los Altos, California: Morgan Kaufmann.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., and Thagard, P. R. (1986). *Induction: Processes of Inference, Learning, and Discovery.* Cambridge, MA: MIT Press.

Meyer, J.-A. & Guillot, A. (1990). Simulation of adaptive behavior in animats: review and prospect. *These Proceedings*.

Minsky, M. (1967). *Computation: Finite and Infinite Machines*. Englewood Cliffs, NJ: Prentice-Hall.

Mitchell, T. M. (1980). *The Need for Biases in Learning Generalizations* (Technical Report CBM-TR-117). New Brunswick, NJ: Rutgers University, Dept. of Computer Science.

Laird, J. E., Rosenbloom, P. S., and Newell, A. (1986).

3.2 *Efficiency*

The second part of the problem side has to do with animat needs and the efficiency with which they are satisfied. An animat can have (1) somatic needs (food, shelter), (2) reproductive needs, (3) additional needs like play, exploration, and prediction. A particular problem may address just one or a few of these, depending on how reinforcement is defined. In addition, hierarchies of secondary needs can in principle result from the primary ones, though this should be explicated experimentally. In the end, satisfaction of all needs can be viewed as in the service of reproduction/survival.

Efficiency of needs satisfaction is the grounds for choosing one solution over another. The best way to do this is probably using a competitive, evolutionary approach in which solutions have costs, niches exist or can form, etc. Though data exist from natural science, determining costs will be difficult. Simulations may tend to use computational in contrast to true somatic costs. This could ultimately turn the animat approach away from nature and toward artificial worlds, where the implications for natural intelligence may not be clear. In any case, one would like to have a reasonable theory of animat efficiency in terms of need satisfaction that will take into account costs and provide criteria for preferring solutions.

4. The Solution Side

4.1 *Architecture*

Here there appears to be a great deal of choice. However, the animat approach (going slowly "upward") should permit a strong criterion of "necessary and sufficient". The progression should at some level of abstraction parallel what actually exists in nature, but that is conjecture; the parallels are unlikely to be obvious, given the apparent role of accident in evolution (Gould, 1989). We can expect, however, that if the SSM description of environments is valid and useful then the best architectures at each stage should be those that most efficiently cope with increasing environmental non-determinacy.

For example, stimulus-response (Classes 0 and 1) environments should imply any of a set of associative memories, which could be implemented with networks, etc. However, as soon as the sensed environment does not uniquely characterize its state (Class 2), the animat can only reach optimal performance using some form of short-term memory, which suggests recurrent networks, classifier systems, etc. Further Class 2 complication will occur when reinforcements are highly delayed and the system must form and retain an intention, and its subordinate intentions, etc., until reinforcement is obtained. Efficiency may then require the introduction of higher-order internal states, modularity, etc. The animat strategy offers a way to bring these in naturally.

One interesting hypothesis is that the most efficient systems will be those that convert every frequently encountered important situation to one of "virtual stimulus-response" in which internal state (intention, memory) and sensory stimulus together form a compound stimulus that immediately implies the correct next intention or external action. This would be in contrast to a system that often tends to "figure out" or undertake a chain of step by step reasoning to decide the next action. The latter more contemplative system would presumably possess increased flexibility in the face of an uncertain environment. However, the present hypothesis is that greater overall efficiency will be found in systems that set up generalized S-R methods in the above sense. The motivation for the hypothesis is that in animals and people, even complex behavior, if frequent and important enough, tends to become reflexive. Standard AI has addressed the question of whether knowledge should be "interpreted" or "compiled" (Laird, Rosenbloom & Newell, 1986). The animat approach offers a new and perhaps more natural context in which to address it.

4.2 *Perception*

The preceding discussion of architecture bypassed the issue of sensory profusion, tacitly assuming inputs are few and well-defined, as in Standard AI. Perception—which might be defined as knowing what in the environment is relevantly the case—has proved very difficult to imitate computationally. It has a chicken and egg quality: How do you know what aspects of a complex profuse input to select or combine into patterns until you know how to view the input so as to find them, which in turn means knowing where or what they are in the first place.

One approach with some success is a combination of top-down and bottom-up processing in which, iteratively, fragmentary data from below suggest candidate remembered percepts above which in turn guide the lower search for confirming or disconfirming additional data [see, e.g., Grossberg (1987)]. Most uses of these and the related relaxation techniques [e.g., Geman & Geman (1984)] have occurred in the analysis of scenes or images from specialized domains, and so are subject to Standard AI's brittleness. In contrast, the animat approach, retaining stimulus-profuse environments but aimed at simpler percepts, should permit the development of more general and adaptive top-down/bottom-up strategies.

Woods-like environments containing simple objects ("tree", "food", etc.) offer an interesting test-bed for such strategies if the animat receives a stimulus that

$$\{E(t{+}1)\} = f\,(E(t),\, A(t)).$$

Here the next stimulus is indeed a function of the current stimulus and action, but, as indicated by curly brackets, $E(t{+}1)$ will be one of a *finite set* of possible stimuli, not unique. Although the set is determined by the function f, the particular member of the set that occurs is not. Thus the above relation expresses a *non-determinacy* of the environment with respect to the variables E and A.

We might term this kind of environmental description a *sensory-state machine* (SSM). In fact, for every FSM there is an SSM that can be derived by straightforward (though perhaps tedious) examination of the FSM state diagram. The SSM, trades the determinacy of the FSM for a formalism that expresses the environment—or the animat's problem—solely in terms of variables that the animat knows about. Furthermore, the SSM's non-determinacy is familiar: the reaction of an environment to an action is very often not fully predictable from knowledge of that action and one's immediate sensory situation.

Let us note that the SSM as defined above is an *incomplete* description of the environment. It can be derived from the environment's FSM, but the FSM cannot be derived from it (in general). Nevertheless, the SSM appears to be a more useful construct for understanding levels of environmental difficulty, as we now attempt to show.

Consider an environment which an animat detects through extremely limited sensory apparatus. For example, the animat might have only a single small touch detector, pointed straight ahead. The SSM for this environment-cum-sensory-apparatus would be extremely non-determinate, since a large number of object shapes would be consistent with stimulation of the single small touch detector. Should, for example, the animat turn 30 degrees to the right, subsequent stimulation of the detector would be nearly unpredictable.

On the other hand, consider an animat in the same environment but having elaborate stereoscopic vision. In this case the SSM would contain little non-determinacy, since for example the visual stimulation subsequent to the same 30 degree turn or most other actions would be a unique function of the current image and therefore predictable.

Predictability of the results of actions in the context of sensory stimulation is the foundation of an animat's survival and, indeed, prosperity. Attainment of reinforcement depends on the ability to choose actions that lead to reinforcement, whatever the sensory circumstances. The examples above suggest that the degree of non-determinacy of an environment's SSM is an important measure of the environment's relative difficulty.

A simple and tentative taxonomy of environments can be constructed based on SSM non-determinacy.

Class 0. Environments with completely determinate SSMs and in which for every sensory stimulus there exists at least one action which if taken will result in positive reinforcement. This might be called a pure stimulus-response environment, meaning that the optimal action in each situation is a function only of the current stimulus. The "landmark" environment of Barto & Sutton (1981) is an example of a Class 0 environment.

Class 1. Environments with completely determinate SSMs in which for only some sensory stimuli does there exist at least one action which will result in positive reinforcement. This could be called a stimulus-response environment with sparse or deferred reinforcement. The 288-state environment of Grefenstette (1988) and the maze environment of Sutton (1990) are examples of Class 1 environments.

Class 2. Environments with partially non-determinate SSMs. In contrast to Classes 0 and 1, reliable prediction can no longer be based on the current sensory stimulus and action. The environment "WOODS7" of Wilson (1985a) is an example of a Class 2 environment, as is the "Little Prince" environment of Rivest & Schapire (1987).

In many cases it will be possible to reduce or eliminate the non-determinacy of a Class 2 environment by taking into account some degree of recent history. For example, suppose that for a particular Class 2 environment we construct the *second-order* SSM:

$$\{E(t{+}1)\} = f_2(E(t),\, A(t),\, E(t\text{-}1),\, A(t\text{-}1)).$$

It may well be the case that this SSM is less non-determinate that the first-order one, the additional context of the prior time-step's stimuli and actions serving to reduce the uncertainty. We can further imagine that for *some* order of SSM, the non-determinacy is eliminated. Let that order be k. Then we could describe the environment in question as being of Class 2.k, with higher values of k standing for greater difficulty.

This concludes our discussion of environments, in which we developed the idea that environments could be ordered in difficulty according to the non-determinacy of their SSMs.

One complication that we have not mentioned, but will take up in Section 4.2, is the fact of "stimulus profusion" in real environments. The environment may not have a tricky SSM, but at the sensory interface it always has a *very* large one. Thus realistic environments pose a problem of selection of relevant data. At higher levels this can be a problem of "pattern recognition" and quite complex.

ment immediately. In others, the reinforcement is deferred, though the optimal action is still knowable from the stimulus. An example would be an odor gradient that reliably pointed toward the location of food.

In a somewhat more complicated environment, information-bearing stimuli are not as simply related to reinforcement as odor is related to food. Instead they may consist of more or less arbitrary cues like stimuli from a certain kind of bush that prey like to hide in or, to mention a human context, a certain kind of golden arch! In still more complicated environments, the optimal action is no longer knowable from the immediate sensory stimulus. Consider leaving your office and turning in the correct direction in accordance with a phone call received five minutes earlier. Or the environment of a stalking animal in which the current objective is temporarily out of sight. At such times the immediate sensory stimulus may contain no information at all relevant to attaining the objective. Further complexity is of course introduced by the presence of competing creatures with similar or different needs. All cases become more difficult if environment characteristics are statistical, or stimuli or reinforcements contain noise. These are just examples, but they suggest a bit of the range and subtlety of real environments.

Given this variety, a more formal characterization is desirable. A start can be made by noting that from an animat's (or animal's) point of view, the environment is a kind of machine that (in general) responds with a new sensory stimulus (which may include reinforcement) whenever the animat executes a motor action. One formal way to describe such an environment is as a *finite-state machine* (FSM) for which the motor actions are inputs and the sensory stimuli are outputs (Riolo, 1987; Rivest & Schapire, 1987). The behavior of a finite-state machine is defined by two equations (Minsky, 1967):

$$Q(t+1) = F(Q(t), A(t))$$

$$E(t+1) = G(Q(t), A(t)),$$

where A is the machine's input (in this case the animat's motor action), E is the machine's output (in this case the sensory stimulus), and Q represents the machine's (the environment's) current "state". Time t is assumed to be discrete. The variables A and E are in general vectors.

The first equation says that the environment's next state is a function F of its current state and the motor action. The second equation says that the next sensory stimulus to the animat is a function G of the current state of the environment and the motor action. The FSM formalism captures the idea that actions in a given environment result in new sensory stimuli; the state variable Q makes it possible for the machine to respond differently to the same action in different circumstances, a common property of real environments. That the FSM is "finite-state" means essentially that the number of different possible outputs for a given input is finite (though possibly very large), or equivalently, that the number of possible values of the state variable Q is finite. A strict finiteness property for real environments is perhaps debatable, but since large FSMs provide a good approximation in many problems of interest, the debate can be left for another occasion.

Besides reacting to animat actions with new sensory stimuli, real environments also sometimes present new stimuli in the absence of action (e.g., the clouds move while you gaze at them, other animats move in your field of view, etc.). This important property is not captured by the FSM formalism, and needs to be included in a fuller environment theory.

Before continuing, it is necessary to be quite careful about the meanings of E and A. Knowledge about the environment comes only through the use of the sensory and motor apparatuses, each acting as a kind of communication channel. Because these tend to be fixed in phylogeny ("hardwired") it is often useful to define the sensory and motor channels as part of the environment. Then, a particular environment of interest might consist, for example, of a (physical) maze as detectable by two eyes of a certain retinal description and manoeuverable by four legs of a certain musculoskeletal description. From this, one would proceed to establish the appropriate functions F and G, treating the retinal outputs as E and the motor command signals as A.

Alternatively, one could treat the sensory and motor channels as part of the animat and not as part of the environment. Such a division might be desirable in problems in which the sensory and motor equipment was subject to an evolutionary process. However, for our current purposes we shall use the former approach in which the two channels are regarded as fixed and part of the animat's environment.

The FSM formalism has advantages and disadvantages. An advantage is that the environmental description can be as precise as desired, and it is necessary to be precise in order to program a simulation. A disadvantage is that the FSM description has a certain opaqueness from the point of view of understanding levels of environmental difficulty. A further disadvantage is that animats deal in stimuli and actions while the FSM also contains the state variable Q, which the animat can't detect. Though in some sense the animat should learn "the reality behind appearances", there is merit in examining an environmental formalism from which Q is absent.

Suppose we try to express the next sensory stimulus directly in terms of the current stimulus and current action. Examination of some FSMs will show that the result is in general not determinate, as indicated by the following relation:

close to what we have called the "animat" (or artificial animal) approach (Wilson, 1985a), and the child machine is an advanced form of animat. Rather than isolated competences, the animat approach is holistic, focusing on complete systems (simulated or, when possible, realized) that, like animals, exist in realistic environments and must cope with the varied problems that they present.

Obviously, we can't yet simulate human intelligence holistically. But the basic *hypothesis* of the animat approach is that by simulating and understanding complete animal-like systems at a simple level, we can build up gradually to the human. At each point we will be careful to include full connection with a sensory environment, together with maximum use of perception, categorization, and adaptation. Thus when we reach the human level these crucial abilities will not be missing. We hope to reach human intelligence "from below", instead of piecemeal through high-level competences as in Standard AI.

The animat approach also brings with it a needed element of pragmatism (Holland, Holyoak, Nisbett & Thagard, 1986). Survival needs and their derivatives are evidently the principal drivers of animal behavior, and so, at bottom, they must be for human beings. The effect is that needs have a powerful influence on the formation of percepts and concepts—in machine learning terms, they set the *inductive bias* (Mitchell, 1980)—yet this has been little acknowledged in AI work. The animat approach explicitly brings in needs by making them the drivers of system behavior.

Introduction of needs opens the way to operational definitions of intelligence since the efficiency of need satisfaction is in principle quantifiable. For example, some years ago van Heerden (1968) summarized his observations on human intelligence as follows:

> Intelligent behavior is to be repeatedly successful in satisfying one's psychological needs in diverse, observably different, situations on the basis of past experience.

With suitable changes, this definition can be applied from human to very simple animal levels. It brings in perception, categorization, and adaptation, and it bases degree of intelligence on rate of need satisfaction.

Our aim in this paper is to outline themes in the animat approach to AI—that is, to suggest "how to go about it", at least in first approximation. A number of efforts already exist [for a review, see Meyer & Guillot (1990); also see Smith & Wilson (1989)], and the interest is accelerating. We shall explain our view of what is needed, and suggest potential directions of formalization.

2. The Basic Animat Strategy

The basic strategy of the animat approach is to work toward higher levels of intelligence "from below"—using minimal ad hoc machinery. The essential process is incremental and holistic: given an environment and an animat with needs and a sensory/motor system that satisfies these needs to some criterion, increase the difficulty of the environment or the complexity of the needs—and find the minimum increase in animat complexity necessary to satisfy the needs to the same criterion. Alternatively, the environment could stay the same but the needs satisfaction criterion might be increased; again find the minimum animat complexity increase. In either case it is vital (1) to maintain the realism and whole-ness of the environment, however simple it is, so as to avoid special-purpose solutions; (2) to maximize physicality in the sensory signals, so as to avoid predefined symbolic inputs; and (3) to employ adaptive mechanisms maximally, to minimize the rate of introduction of new machinery and maximize understanding of adaptation.

Note that the strategy has a "problem side" (harder environments, increased efficiency) and a "solution side" (new architecture: sensory/motor, internal, adaptive). Changes in the problem side can be due to the experimenter, but also to (co-)evolutionary effects if the environment is evolving. Similarly, changes in the solution side can be deliberate or evolved, based on a genotype and selection.

Research on animat-like systems has tended to emphasize the solution side. A certain experimental environment is selected as being in some sense interesting, but most of the work goes into testing and refining a particular architecture in that environment. The result is often a successful system, but accompanied by insufficient insight from a formal point of view into the properties or difficulty of the environmental problem that has been solved, and with what efficiency. A major aim of this paper is to suggest the need for a more systematic understanding of environments.

We now discuss the problem and solution sides in more detail.

3. The Problem Side

3.1 Environments

Environments differ enormously in their complexity, uncertainty, and degree of reinforcement. Needed is a formal theory and taxonomy that will order environments and reveal their differences in difficulty. For example, some environments (e.g., some food concentration gradients) can be thought of as pure stimulus-response: the local environmental signal directly indicates the optimal action and provides reinforce-

From J. A. Meyer and S. W. Wilson (eds.), *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behavior*, Cambridge, Massachusetts: The MIT Press/Bradford Books (1991)

The Animat Path to AI

Stewart W. Wilson
The Rowland Institute for Science
100 Cambridge Parkway
Cambridge, MA 02142 USA
(wilson@think.com)

Abstract

A research methodology is proposed for understanding intelligence through simulation of artificial animals ("animats") in progressively more challenging environments while retaining characteristics of holism, pragmatism, perception, categorization, and adaptation that are often underrepresented in standard AI approaches to intelligence. It is suggested that basic elements of the methodology should include a theory/taxonomy of environments by which they can be ordered in difficulty—one is offered—and a theory of animat efficiency. It is also suggested that the methodology offers a new approach to the problem of perception.

1. Introduction

There are two broad approaches to the scientific understanding of intelligence, or how mind arises from brain. One is the natural science approach, analyzing and experimenting with phenomena of life, mind, and intelligence as they exist in nature. In this there are two main branches: physiology and especially neurophysiology, in which living systems are subject to detailed internal investigation; and experimental psychology, including studies of animals, in which living systems are studied through their external behavior. Related to the latter, but more observational, are fields such as linguistics and anthropology.

In contrast, the second broad approach to intelligence may be termed synthetic and computational, in which the objects studied are constructed imitations of living systems or their behavior. In "Computing machinery and intelligence", Turing (1950) suggested two possible directions for the computational approach:

> We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child. Things would be pointed out and named, etc.

Turing's first proposed direction led to "standard AI" or computational cognitive science. Standard AI is basically competence-oriented, modelling specific human abilities, often quite advanced ones. However, while many AI programs exhibit impressive performance, their relevance for the understanding of natural intelligence is, in several respects, limited.

In addressing isolated competences, AI systems typically ignore the fact that real creatures are always situated in sensory environments and experience varying degrees of need satisfaction. Furthermore, the systems attach less importance to such basic natural abilities as perception, categorization, and adaptation than they do to algorithmic processes like search and exact reasoning. This leads eventually to problems connecting the arbitrary symbols used in internal reasoning with external physical stimuli ("symbol grounding" (Harnad, 1990)), and "brittleness" (Holland, 1986), the tendency for AI systems to fail utterly in domains that differ even slightly from the domain for which they were programmed.

AI systems also have an arbitrariness: it is often not clear why one program that exhibits a certain intellectual competence is to be preferred over some other one exhibiting the same competence, especially since the field has not agreed on—or too much sought—a clear definition of intelligence. In a sense, the programmer's facility for imitating a high-level fragment of human competence is a kind of trap, since from a natural science perspective there is usually no strong relation to nature.

Turing's second proposal, for a "child machine", received, over forty years, little attention or resources, perhaps because it seemed fantastic. Yet the child machine was to be situated from the start in a real sensory environment and was to learn through experience. It would have emphasized precisely the abilities that standard AI minimized. Turing's proposal is in fact