

The Use of Hierarchies for Action Selection

Toby Tyrrell

Plymouth Marine Laboratory

Plymouth PL1 3DH, UK.

email: tt@pml.ac.uk

Abstract

Several researchers of animal behaviour, such as Tinbergen and Baerends, have proposed hierarchical mechanisms for action selection. Maes, amongst others, has argued against mechanisms of this type because of the top-down control and lack of robustness inherent in these *hierarchical decision structures*, or rigid switching mechanisms, in which decisions are made at every level. Two alternatives exist to these hierarchical decision structures: **(1)** Maes and others have designed non-hierarchical action selection mechanisms, and **(2)** Rosenblatt & Payton have outlined a hierarchical mechanism which does *not* make decisions at every level. In this paper the use of the two types of hierarchies for action selection is discussed and the Rosenblatt & Payton approach with free flow of information, combination of evidence and the ability to select compromise candidates is supported. It is also argued that the problem of action selection is by nature intrinsically hierarchical, and so Rosenblatt & Payton-like hierarchies (*free-flow hierarchies*) are more suitable for action selection than non-hierarchical mechanisms. Evidence from observation of animal behaviour and from computer simulation testing is used to give some support to these claims.

1. Introduction.

A long-standing interest among ethologists, psychologists, roboticists and others has been the question of 'how can behaviour be generated and controlled?'. When considering the computational structures or mechanisms that may underly behaviour, a common suggestion has been that hierarchies may be involved. This paper examines the use of hierarchies for generating behaviour (i.e. for selecting actions), and in particular contests a commonly-held assumption about how hierarchies should be used to select actions.

Before commencing with the body of the paper it is necessary to define some terms. What is meant by the term *hierarchy* will be discussed first, followed by definition of the term *action* and some other related terms.

The definition of a *hierarchy* used here is taken from [Dawkins, 1976]. For two nodes A and B then A 'is boss of' B if A has a direct causal influence on B . *Direct* is meant to imply that A is immediately above B in the hierarchy (i.e. its influence on B is not mediated through another node), and *causal influence* is meant to imply that the state of node B is to some extent dependent on the state of node A . For a set of nodes A, B, C , etc, A is said to be *superior* to B if (1) A is boss of B , or if (2) A is boss of a node which is superior to B . Using this recursive definition of the term superior it is now possible to define a hierarchy as a set of nodes A, B, C , etc. which satisfy:

- (i) there is no node in the set which is superior to itself, and
- (ii) there is one element in the set which is superior to all others in the set.

Figure 1 shows some example hierarchies. The hierarchical mechanisms discussed here are all composed of sets of overlapping hierarchies (one for each system) as in figure 1(e). For simplicity, the different mechanisms will be referred to as hierarchies, although strictly speaking they are each sets of overlapping hierarchies.

Several other terms are illustrated in figure 2 and in figures 5 and 6.

The term *action* here refers to an entity at the 'lowest' level, that of the behavioural final common path [e.g. McFarland, 1975]. That is to say, the demands of each action on the effectors of the animal are mutually exclusive so that only one action may be carried out at any one time. Actions are shown in figure 2 and also to the right-hand side of figures 5 and 6. Some example actions in figure 5 are retrieving eggs, trampling, lining of the nest and combing of the feathers. *Action selection* is the process of choosing, by whatever means, one out of a set of candidate actions.

It should be noted that the *action selection mechanisms* (ASMs) considered here are responsible only for the calculation of what the next unit of behaviour should be. The mechanisms are not responsible for calculating how the chosen action is implemented via movements of the limbs of the animal. These mechanisms compute action selection, not motor control.

Animals are vehicles for the transmission and propagation of their genes [Dawkins, 1989]. The global *action selection problem* for an animal is therefore how to choose successive actions so as to maximise the number of copies of its genes in future generations. Action selection mechanisms have evolved to provide a solution to this problem, i.e. to make the animal reproduce, and to make it help conspecifics with similar sets of genes (e.g. siblings, parents, offspring) to do the same. In order to achieve the above the animal needs to stay alive and to help conspecifics to do so as well. The main action selection problem thus engenders more specific *sub-problems* such as the need to obtain food, the need to avoid predators, the need to care for young and the need to find mates. Each of these sub-problems

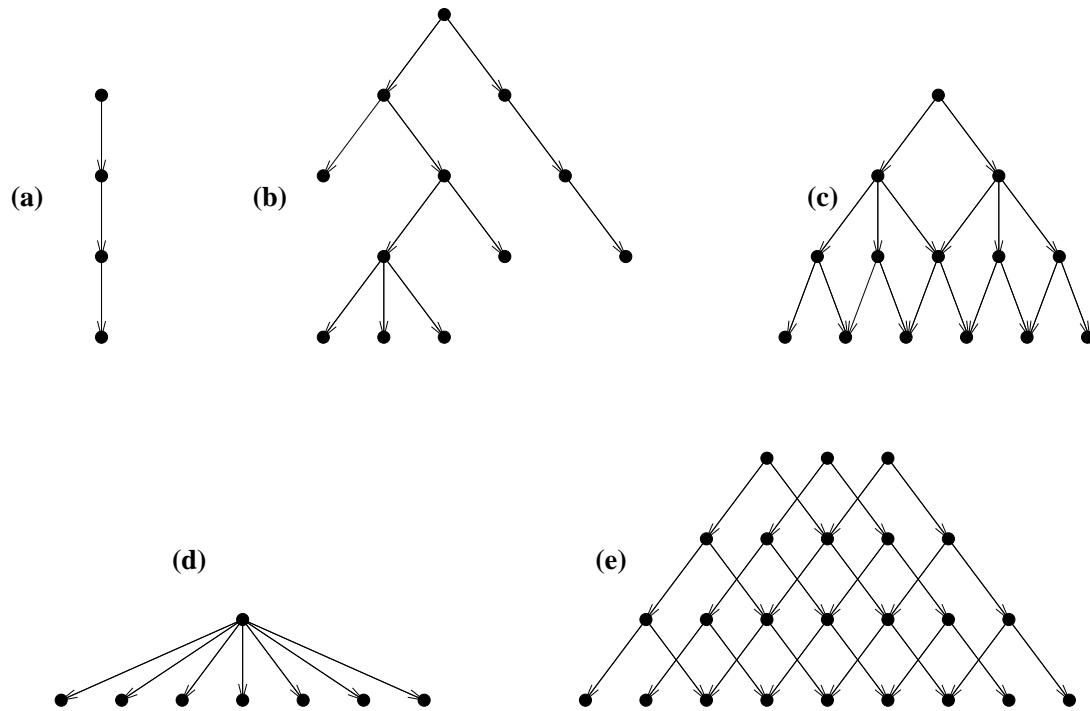


Figure 1: Example hierarchies. **(a)**, **(b)** and **(d)** are non-overlapping hierarchies, **(c)** is an overlapping hierarchy because some nodes have two bosses, and **(e)** is a set of overlapping hierarchies (after [Dawkins, 76]).

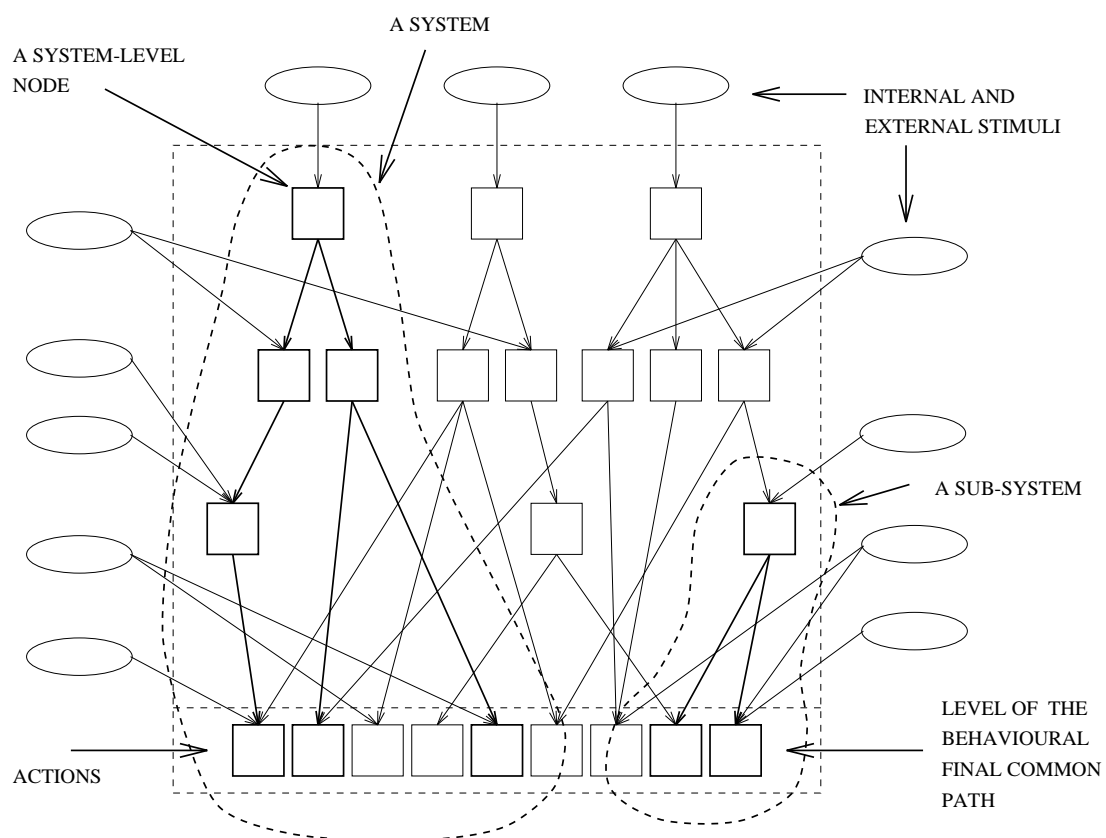


Figure 2: An example of the structure of a hierarchical action selection mechanism. The terms system, sub-system and action are illustrated.

will have its own requirements on the animal’s time which will need to be interleaved by the action selection mechanism.

Most mechanisms have separate parts devoted to each of the separate sub-problems, which will be referred to here as *systems* (see figure 2). In particular, with the hierarchical mechanisms considered here, each system corresponds to one of the set of overlapping hierarchies. A system-level node is the highest-level node of a system, i.e. the top node of one of the overlapping hierarchies. Some systems of the mechanism shown in figure 5 are nesting, escaping and preening. Parts of the mechanism at any level between that of the system level and the action level will be known as *sub-systems* (see figure 2). Each system takes input from many different internal and external stimuli.

The terms just defined, while being very similar to those used by Baerends, differ from those defined by McFarland [1975].

The rest of this introductory section is given over to a discussion of other work relevant to the theme of this paper. Firstly two hierarchical mechanisms for action selection, due to the ethologists Tinbergen and Baerends, are reviewed. They are both the product of extensive observation of animal behaviour. Following this is a description of a different type of hierarchical mechanism, proposed by Rosenblatt & Payton, two roboticists.

1.1. Tinbergen

This mechanism [Tinbergen, 1950 & 1951] was put forward as a “working hypothesis” and as such was only described fairly vaguely, rather than being specified exactly. There is a hierarchy of nodes (see figure 3), each of which (except for the very top-most one) is gated by an *innate releasing mechanism*. The highest level nodes correspond to systems such as ‘reproduction’ and ‘cleaning/preening’. The lowest level nodes correspond to actions such as ‘lay eggs’ and ‘comb feathers’.

The top-most nodes are activated by motivations such as ‘internal food deficit’. Their activation energy will then pass down to lower nodes in turn as long as the innate releasing mechanism (IRM) for each node is activated (an inactivated IRM ‘blocks’ its node). Each IRM consists of a device for summing several incoming stimuli and a threshold. If the sum of the stimuli exceeds the threshold then the IRM ‘frees’ the node, allowing it to pass activation down to lower layers. The IRMs act to block any nodes not relevant to the current situation.

In addition to the IRM threshold, each node has its own separate threshold which needs to be surpassed by the activation it receives from several different sources (see figure 4). The most important of these inputs to a node, for the purposes of this paper, are the inhibitory inputs from other nodes at the same level. It is unclear exactly how Tinbergen expected the intra-layer inhibition to work, except that “as a rule, the strong activation of instinctive behaviour of one kind prevents the functioning of another pattern” [Tinbergen, 1951] and “centres of the same level mutually suppress each other’s activities” [Tinbergen, 1950]. The desired effect of the IRMs and the inhibition amongst nodes at the same level is that in most cases only one node at any level is allowed to be active. That is to say, there is a *winner-take-all* operation at each level. It is this aspect of the mechanism that is of interest here.

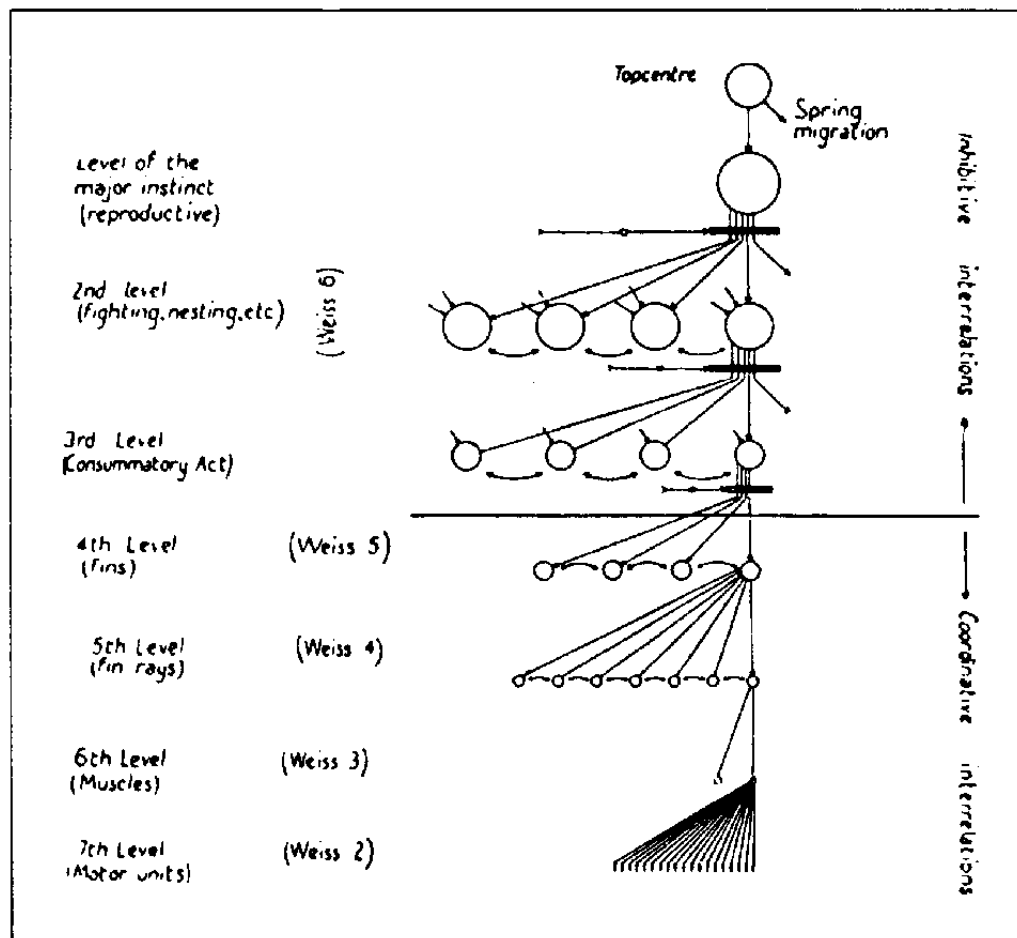


Figure 3: Tinbergen's 'Hierarchy of Instincts' (the bottom half shows motor control and can be ignored). Circles = centres, downwards arrows = causal factors, horizontal arrows = innate releasing mechanisms, and two-way arrows between centres on the same level = mutual suppression. Taken from [Tinbergen, 51] (© acknowledged to Clarendon Press).

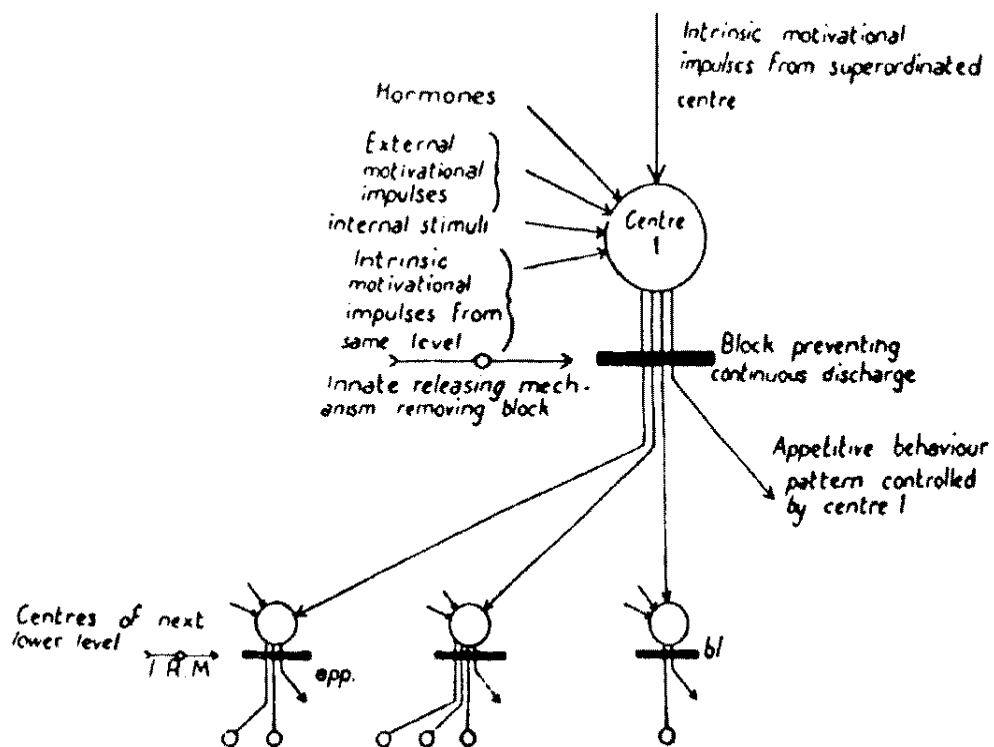


Figure 4: Tinbergen's 'tentative representation' of a node or centre from an intermediate level in the hierarchy. Taken from [Tinbergen, 51] (© acknowledged to Clarendon Press).

1.2. Baerends

Figure 5 shows a model that Baerends postulated to account for the incubation behaviour of the herring gull [Baerends, 1976], and figure 6 shows a similar diagram for the reproductive behaviour of the digger wasp [Baerends, 1976]. The details in the two diagrams are not important, only the type of architecture employed (which is the same in both cases). As with Tinbergen's 'hierarchy of instincts', there is a feeding of activation down through the hierarchy (or rather from left to right in Baerends' diagrams). There are no IRMs, but system-level nodes inhibit one another so that only one is active, and below that level each higher-level node controls (in some unspecified way) which one of its subordinates receives activation. Baerends specifies that internal and external stimuli affect which node (at any level) becomes active, and also that once a node becomes active, it tends to remain active for some time.

Again, the details of the mechanism are not important here, only the fact that (as with Tinbergen's mechanism) only one node at any one level is allowed to be active. That is to say, a *decision* is made at each level as to which system or sub-system is the most appropriate and all others are subsequently disregarded. This type of mechanism is called here a *hierarchical decision structure* (HDS).

1.3. Maes' Criticisms

Maes [1991] has argued against the use of such hierarchical decision structures, in particular because of the top-down, preprogrammed sort of control they exert. Maes designed a mechanism of her own [Maes, 1991] which is distributed and non-hierarchical with no central control. The behaviour of the network is an emergent property of interactions amongst the nodes, and of interactions between the nodes and the outside world (see figure 7). Space does not allow for a detailed description of this mechanism but it should be noted that there are no system-level nodes. Maes' mechanism selects between entities which are all assumed to be at the same level, somewhere below the system level.

While Maes' mechanism was reported to cope satisfactorily in a fairly simple simulated environment [Maes, 1991], it is not obvious how well it could cope with the more complicated scenarios of figures 5 and 6 in which it would seem natural to group together actions into different systems and sub-systems. Maes' mechanism has also been tested by the author in a more complex simulated environment (described in section 3.1), and shortcomings with it were uncovered. Some of these shortcomings may be derived from difficulties in trying to solve what is intrinsically a hierarchical problem with a non-hierarchical architecture. For a detailed account see [Tyrrell, 1993].

1.4. Rosenblatt & Payton

An alternative to a hierarchical decision structure (Tinbergen, Baerends) is what shall be called here a *free-flow hierarchy* (FFH), one in which decisions are not taken at every level, but only at the level of the behavioural final common path. The spread of activation between nodes is unrestricted and is able to 'flow down' to all lower nodes via weighted connections.

The idea of a free-flow hierarchy was proposed by Rosenblatt & Payton [1989]. In their mechanism nodes *express multiple preferences* for each of a set of lower level candidates rather than making a *decision* as to which one is most suitable. This is advantageous because it allows for *combining of evidence* from different nodes and for selection of *compromise candidates* (see figure 8). Each of the nodes in this case (see figure 9) are fairly

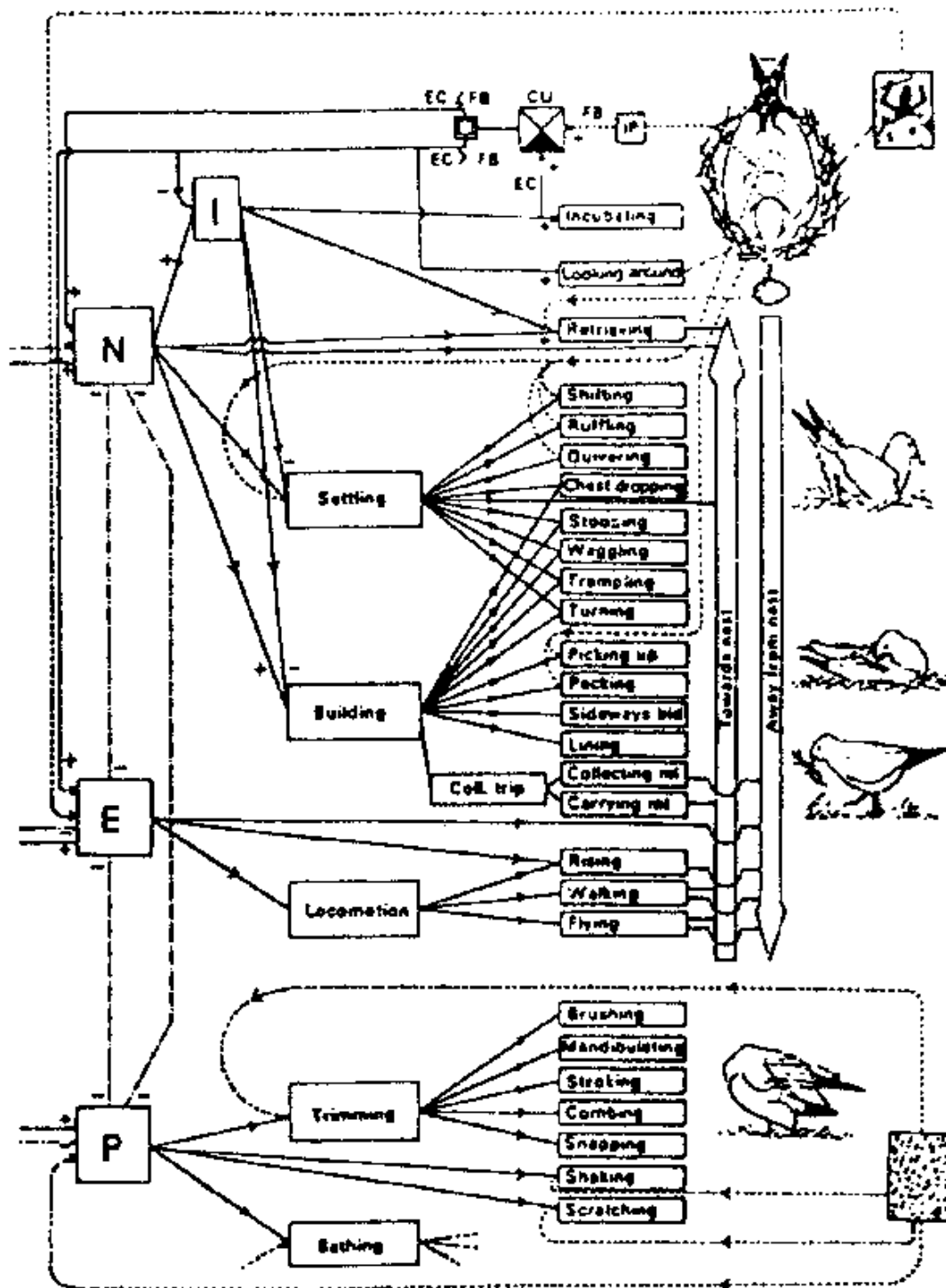


Figure 5: Baerends' model to account for incubation behaviour of the herring gull. Actions or fixed action patterns are in the far right column. "Superimposed control systems" of higher order are to the middle and left. N = incubation system, E = escape system and P = preening system. The three systems mutually suppress each other. Taken from [Baerends, 76] (© acknowledged to Bailliere Tindall).

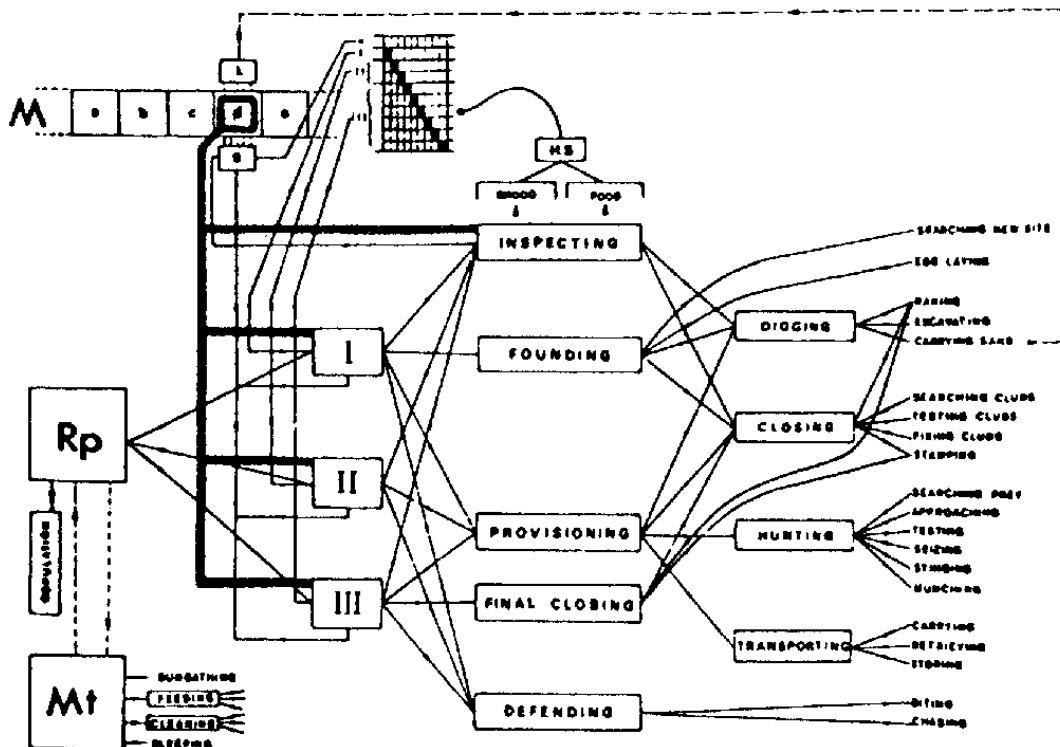


Figure 6: Baerend's model to account for nest provisioning behaviour of the digger wasp. As in the last diagram, actions or fixed action patterns are in the far right column and systems and sub-systems are to the left and middle. **Rp** = reproductive system, and **Mt** = 'maintenance' system. **I**, **II** and **III** are sub-systems controlling the completion of three separate phases of nest provisioning, each of which require the execution of several different lower-level sub-systems. Taken from [Baerends, 76] (© acknowledged to Bailliere Tindall).

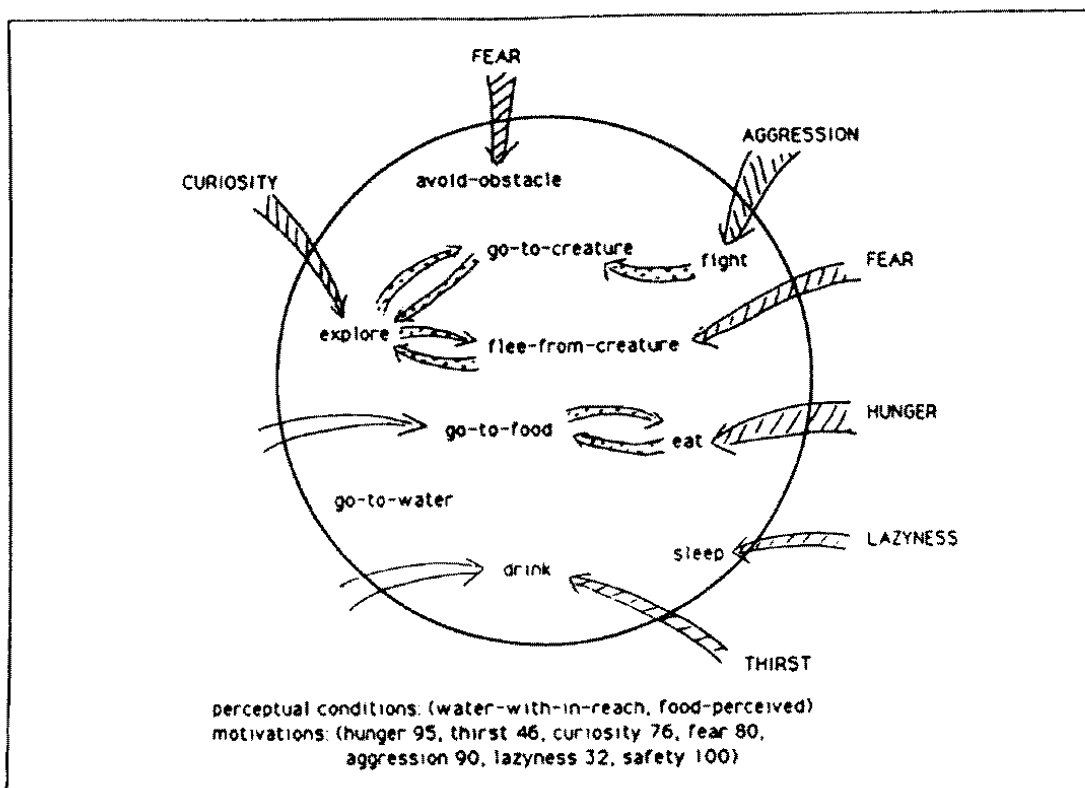


Figure 7: Maes' non-hierarchical, "Bottom-up Mechanism for Behaviour Selection". Excitation flows into the network from perceptual conditions (if true) and motivations. Excitation is also passed between nodes in the network. Taken from [Maes, 91] (© acknowledged to MIT Press).

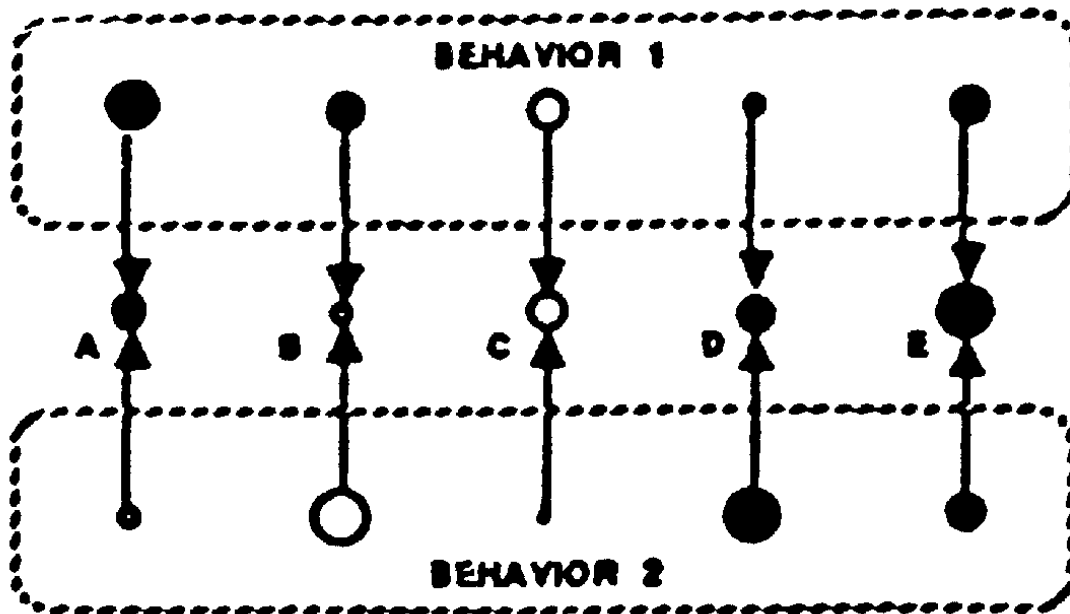
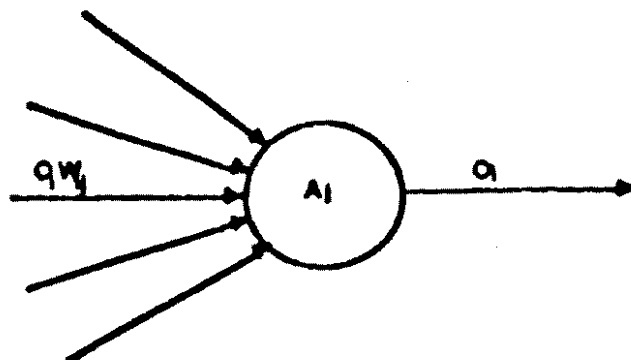


Figure 8: Combining evidence in a Rosenblatt & Payton hierarchy. The two behaviours (1 and 2) are expressing preferences, or ‘voting’, for five lower-level alternatives. Taken from [Rosenblatt & Payton, 89] (© acknowledged to IEEE).

standard artificial neurons (as used in the field of neural networks), except that their rule for combination of weighted inputs is not necessarily a strict summation.

A hierarchy designed according to these principles will include combination of evidence (rather than winner-take-all) at every level, except at the level of actions (because only one can be executed). The action-level node receiving most activation gets chosen. Any number of nodes can be simultaneously active to differing degrees at every other level.

Other work relevant to this paper has now been described. Alternative models of action selection, such as Lorenz’s ‘psycho-hydraulic model’ are not relevant to the current topic and are not considered here. The rest of this paper will consist of a discussion of the relative merits of Tinbergen-like hierarchical decision structures and Rosenblatt & Payton-like free-flow hierarchies. Arguments will then be presented in support of the claim that the latter type of hierarchies are preferable. Some evidence from animal behaviour observations and computer simulation testing will be given in support of these arguments.



O_i is the output of unit i .

W_{ij} is the weight on the link from unit i to unit j .

$A_j = f_A(O_1 W_{1j}, \dots, O_n W_{nj})$, is the activation level of unit j with n weighted inputs.

$O_j = f_O(A_j)$, is the output of unit j .

Figure 9: A node from Rosenblatt & Payton's hierarchy. Taken from [Rosenblatt & Payton, 89] (© acknowledged to IEEE).

2. Computational Reasons for Preferring Free-Flow Hierarchies.

Several computational arguments are presented in this section. These arguments explain why a free-flow hierarchy is better than a hierarchical decision structure at being able to calculate the most appropriate actions.

2.1. Robustness.

As mentioned by Maes [1991], top-down control is undesirable. An HDS is brittle and non-robust in the sense that the initial decisions are vitally important and if they are incorrect then the whole performance collapses. There is no winner-take-all process between the systems in an FFH, which makes it more robust. With an FFH there is less top-down control, more *influence* from above.

2.2. Sensory Bottleneck.

In order to make correct initial decisions in an HDS, the top level has to receive and process all of the sensory information relevant to the lower layers, leading to what has been called a 'sensory bottleneck' (see figure 10). For any system of an action selection mechanism there are several relevant internal / external stimuli. Some of these stimuli may be relevant to a whole system or sub-system (e.g. 'perception of partially built nest' to nest building), whereas others will be relevant to only one specific action (e.g. 'perception of egg outside nest' to egg retrieval) [e.g. Hinde, 1953].

So the desirability of carrying out any specific action will usually depend on several stimuli, some of which are relevant only to that action, and some of which have more widespread relevance. In a hierarchical decision structure the desirability of each system needs to be

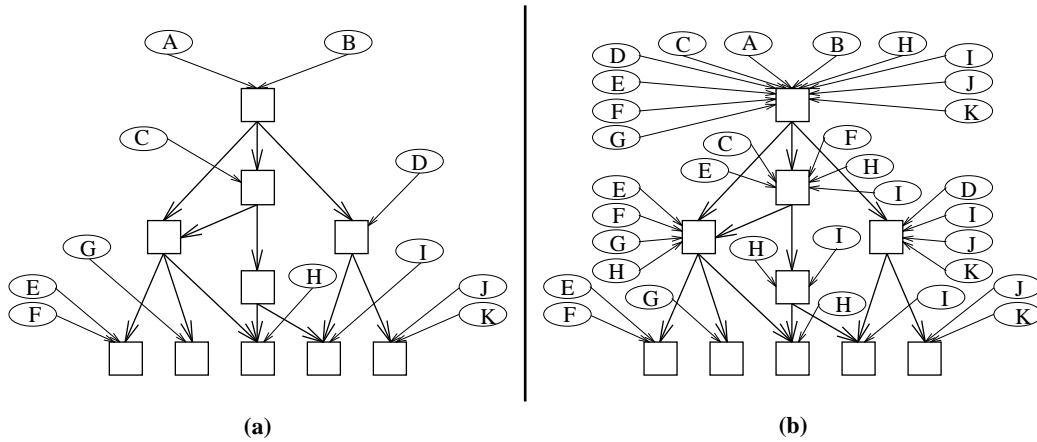


Figure 10: ‘Sensory bottleneck’ — different stimuli are relevant to different subsets of the whole system. In a free-flow hierarchy, these stimuli can impinge on the relevant nodes (as shown in (a)). In a hierarchical decision structure, higher-level nodes need information about the stimuli to nodes that they are superior to, in order for the higher-level nodes to be selected between appropriately (as shown in (b)).

calculated, so that the systems can be properly selected between. The desirability for each system will be related to the maximum of the desirabilities of each of the actions of the system. Therefore, the system-level nodes must each receive all of the relevant stimuli for all of the actions in their system, which produces the ‘sensory bottleneck’ mentioned above.

It should be noted that Tinbergen’s mechanism does include stimuli which impinge only on the relevant lower-level nodes of the hierarchy and does also select between nodes at each level, but this is sub-optimal. For instance, consider a system which contains a consummatory act for which the appropriate stimulus situation exists (e.g. a ‘getting food’ system might have a consummatory act ‘eat’ for which the appropriate stimulus situation is the presence of food). With Tinbergen’s scheme, the existence of the stimulus situation appropriate to the consummatory act cannot be taken into account when calculating the likely benefit to the animal of choosing that system as opposed to others.

In an FFH, system-level nodes receive only motivational stimuli relevant to *all* of the actions in the system. Stimuli impinge on different nodes of the hierarchy. Some stimuli impinge on system nodes, some on sub-system nodes, some on action nodes (dependent on how much of the system they are relevant to). There is no ‘sensory bottleneck’.

2.3. Treatment of Information.

Another consideration is that of the proper treatment of information. By ‘shutting off’ all parts of the hierarchy in a hierarchical decision structure, except those beneath nodes winning the competition at their level, there is a *loss of information*, information which might be useful in selecting the most appropriate action. While redundant and irrelevant information can be discarded, relevant information must not be.

Information is treated correctly in a free-flow hierarchy. The preferences of systems and sub-systems for all lower-level candidates are always utilised until the level is reached at

which a decision has to be made (i.e. the action level). All relevant information is *preserved*.

2.4. Compromise Candidates.

This loss of information in an HDS results in an inability to choose *compromise candidates*. This deficit is particularly apparent in the case of *proscriptive demands* on the animal's actions. A proscriptive demand is 'ruling-out' one, which specifies which action or actions should *not* be performed. If an animal happened to be situated next to a hazard such as a precipitous cliff then the overriding need for it is to avoid that hazard. In a Tinbergen-like or Baerends-like system this will cause the 'avoid hazard' system to be chosen and to bring about a movement away from the hazard. In fact there are many actions that would not cause a movement towards the hazard, some of which would be beneficial because of other systems, but these cannot be considered simultaneously with the need to avoid the hazard. Another example occurs when the animal's internal body temperature is dangerously high. In this case it needs to avoid any strenuous actions which will increase its body temperature even further, but that still leaves several different actions that it can undertake. In order to choose usefully between these actions, the preferences of other systems need to be considered simultaneously.

Because of the preservation of relevant information in an FFH, it is able to choose compromise candidates (e.g. see figure 8). In particular, proscriptive demands can be meshed appropriately with preferences from other systems, as shown in figure 11.

While it is fairly obvious that the problem of action selection is by nature hierarchical, it has not been generally accepted that the best solutions to the problem (i.e. the best action selection mechanisms) should also be hierarchical. From a computational viewpoint, it would seem that the need to express and combine preferences between layers means that there *has* to be some kind of hierarchy in any valid action selection mechanism.

This need for combination of preferences has already been appreciated to some extent in robotics, where most of the repertoire of possible actions of a robot usually consists of movements in different directions. Traditionally, action selection mechanisms have been centralised planners with bureaucratic, top-down control. Researchers such as Brooks [1987] have argued against such centralised control and *potential fields* [e.g. Arkin, 1987] have recently become popular. A potential field is a scheme for combining preferences for movements in different directions.

An analogy can be made between the two hierarchical schemas for action selection considered here and the decision-making of a group of ministers or government officials. Imagine that each official is responsible for a particular aspect of policy, such as education, foreign affairs, unemployment or the economy. Imagine further that the group are meeting to decide upon the future political strategy of their nation, and that each official is constrained to evaluate the various strategy options from the standpoint of their individual department only. The prime minister or president resides over the council and makes the ultimate decision as to which of the alternative options gets chosen.

Decision-making with a hierarchical decision structure is equivalent to the prime minister or president listening to each of the various officials in turn and then selecting the one department for which the decision is most crucial. The future policy is then chosen to be that which is most beneficial to the particular department.

By contrast, decision-making with a free-flow hierarchy is analogous to the prime minister or president listening to each official outline the pros and cons of each action for their particular aspect of government, and then coming to a decision which takes into account the views of

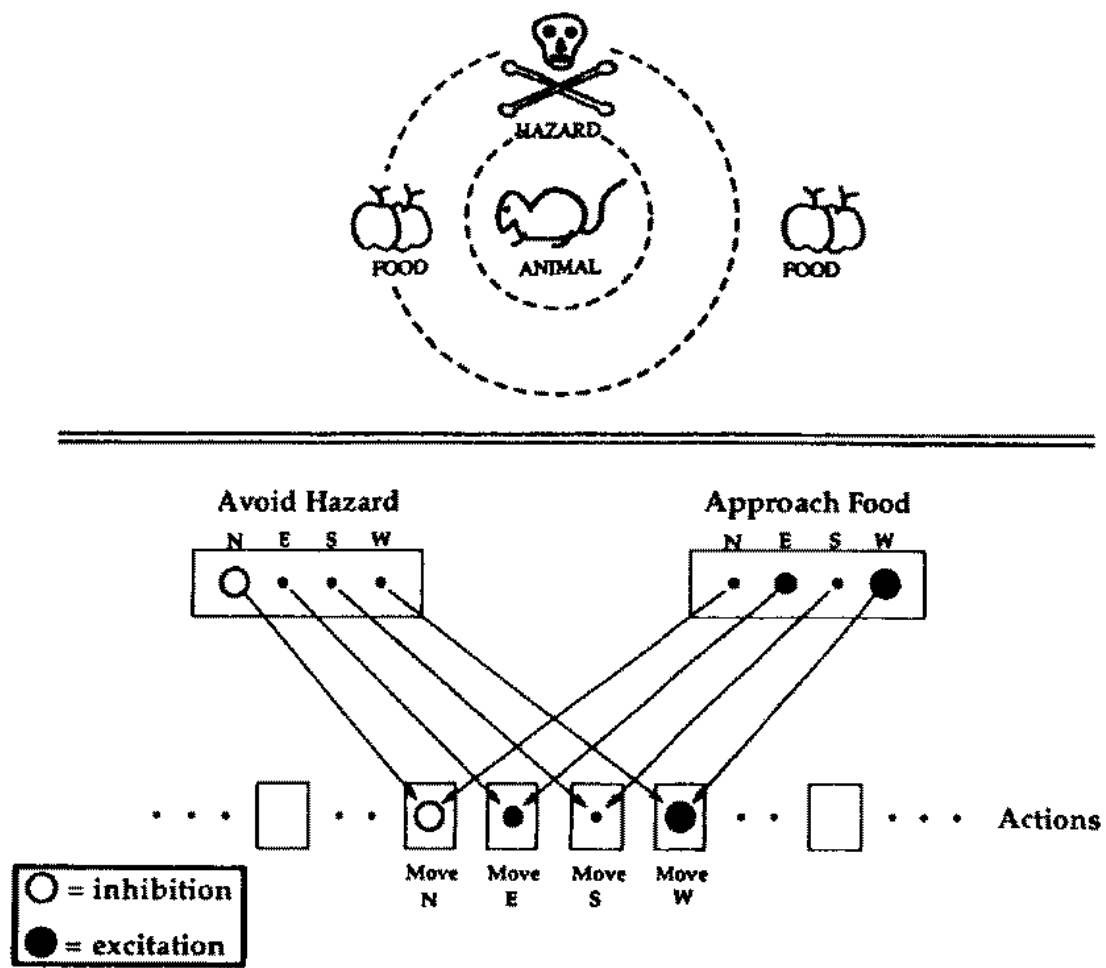


Figure 11: Combining preferences from proscriptive and prescriptive systems (the chosen action would be 'move south' with a hierarchical decision structure or would be 'move west' with a free-flow hierarchy).

all officials simultaneously. So, for instance, the official in charge of unemployment might express a strong disapproval of option A because of the number of people who would be put out of work, but be ambivalent as to which of the other options is chosen. The official for foreign affairs could mildly favour option B because it would ease diplomatic relations with another country, but express no preference or dislike for other options. And the various options might each be favoured to differing degrees by the official concerned with the economy, in accordance with the likely effects on the national finances. The preferences of each official for each policy option can be combined to calculate which option is most beneficial when all aspects are taken into consideration.

3. Supporting Evidence from a Computer Simulation Study.

Some computational arguments have now been outlined which, if valid, indicate that free-flow hierarchies (FFHs) are more suitable for calculating action selection than hierarchical decision structures (HDSs). This assertion was tested in a computer simulation of the action selection problem faced by an animal. The computer simulation was used to test the two types of action selection mechanism, and to generate performance measurements which could then be compared.

This section describes the computer simulation study, the results of which give some support to the claim that FFHs are better than HDSs at action selection. However, it will be argued that the computer simulation results can only provide a rather fallible test of this claim. Most of the material in this section is not directly relevant to the central claim of this paper, but is required in order to allow the reader to be able to gauge the strength of the evidence provided by the simulation testing.

This section is comprised of: *(i)* a description of the simulated environment (SE), a computer simulation of the action selection problem faced by animals, *(ii)* a description of the testing procedure used, i.e. the way in which the simulated environment was used to evaluate the performance (the ability to select actions appropriately) of the two different ASMs, *(iii)* an explanation of some extensions to the basic idea of expression of preferences, that were shown to be necessary in order for free-flow hierarchies to be able to select actions sensibly, and *(iv)* presentation of results from the simulated environment, which allow a tentative comparison to be made between the performances of the two different types of ASM.

3.1. The Simulated Environment.

The simulated environment models the external environment of the animal (the outside world), as well as the internal variables (such as body temperature and cleanliness of fur or feathers), and the processes of perception, navigation and motor control (figure 12). The dynamics of the interactions between all of these, and between them and the action selection mechanism, are all programmed into the simulated environment. Only the process of selecting actions (on the basis of perceptual and navigational information) is not modelled in the SE. Instead, different ASMs, embodying different theories of action selection, are ‘plugged into’ the SE and allowed to make the decisions as to which actions are chosen at each moment in time.

The action selection problem posed by the SE is composed of 13 different *sub-problems*: obtaining food, obtaining water, keeping clean, regulating body temperature, avoiding predators, being vigilant for predators, staying close to cover (protective vegetation), avoiding the edge of the SE, avoiding hazards (places where it is dangerous for the animal to go), avoiding irrelevant animals, sleeping at night in a den, not getting lost and reproducing. These sub-problems all derive from the more general problem for an animal of how to survive and also reproduce as frequently as possible. This list was compiled after looking at all the behaviours or behaviour groupings mentioned in [McFarland, 85], [Morris, 90], [Dewsbury, 78] and [Kilmer *et al*, 69]. The most generally applicable sub-problems were chosen, as long as they were not social ones (such as raising offspring, social grooming, or defending territory). Social sub-problems were avoided (the animal was assumed to be solitary and male) because otherwise the SE (designed to test action selection) would require built-in assumptions about the action selection of con-specifics, which is rather close to including assumptions about the action selection of the ASM being tested. One of the sub-problems, avoiding the edge of the SE, is artificial but is necessary unless a much larger environment

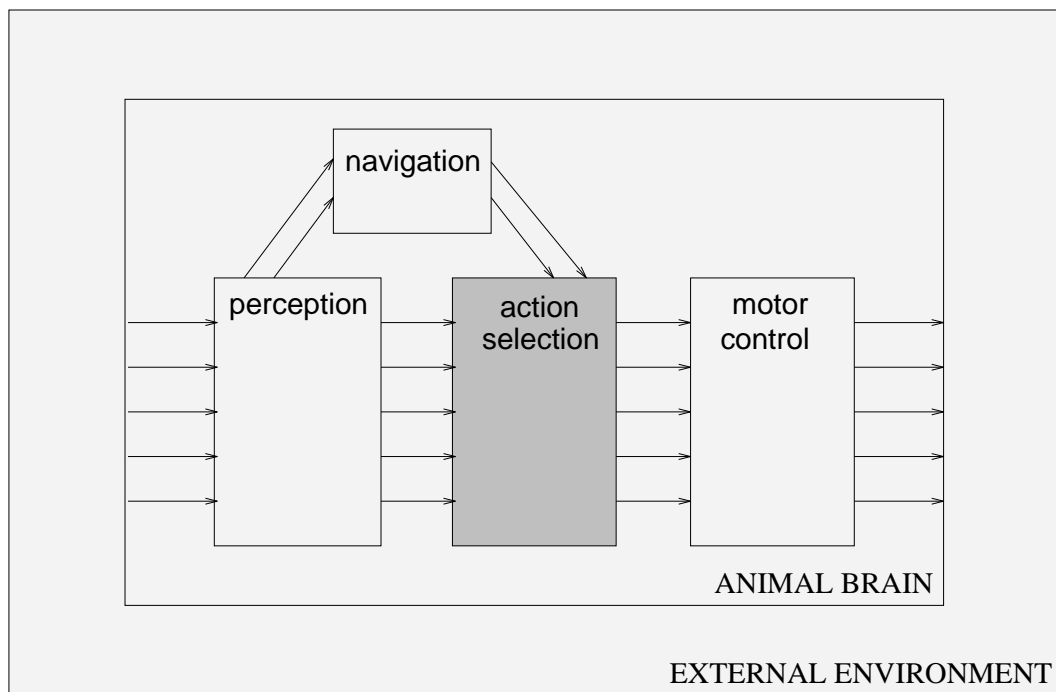


Figure 12: The relationship between the external environment, perception, navigation, action selection and motor control. The simulated environment models everything except action selection, which is provided by an action selection mechanism.

is to be modelled.

In order to include different sub-problems in the action selection problem posed by the SE, different features, internal variables and actions must be added as necessary. For example, in order to include the sub-problem ‘obtaining water’, the following must be added to the SE: a feature ‘water source’, an internal variable ‘water level’, an action ‘drink’, and appropriate dynamics and interactions between the three of them (for instance, when an animal drinks at a water source then the amount of water there should decrease, and at the same time the animal’s water level should increase).

For each individual test that is carried out, a randomly different instance of the SE is created by distributing sets of features semi-randomly about a 25×25 space, as shown in figure 13. A key to the different features is given in figure 14. The animal has a choice of 35 different actions that it can choose to undertake, such as ‘drink’, ‘clean self’, ‘sleep’, ‘move north’, ‘move south-west’ and ‘look around’. By choosing appropriate actions in appropriate situations, the animal can influence some control over the values of its internal variables, and its local environment. Each choice of an action will also influence other things, such as how well the animal can perceive its local environment, how likely it is to be perceived by predators, and how rapidly it depletes its levels of food and water.

The perception of the animal is local and error-prone (figure 15). The animal is less likely to perceive something correctly (or to perceive it at all) as the distance from the animal increases, or if vegetation interrupts the line of sight. Navigation is also error-prone, with the animal able to get lost if it explores too long without returning to familiar areas, and with the possibility of not recognising previously encountered features on subsequent occasions. In addition, the model of motor control includes the possibility of incorrect execution of an action, in which case the animal attempts an action, but is unsuccessful at it.

Performance in the simulated environment is equivalent to genetic fitness, i.e. the ability of the animal to propagate its genes into future generations. Because of the lack of social sub-problems and the presumed solitary lifestyle of the animal, a *non-inclusive* genetic fitness measure is used here. That is to say, the performance of the animal is measured solely by the number of times it manages to mate before it dies. This means that the animal needs to concentrate on mating, but it also needs to concentrate on surviving so that it can continue to mate in the future.

This rather brief description of the SE is all that can be given here. More explanation, justification and validation of the SE is given in [Tyrrrell, 93].

Many tests were carried out on the SE in order to check specific details of the simulation, as well as to investigate the overall capacity of the SE to distinguish between better and worse ASMs. One test in particular is of relevance here, and the results of it are shown in figure 16. In this test, a simple ASM was slotted into the SE and tested many times to give an average performance (the left-most point in the graph). Following this, several more instances of the same ASM were also tested, but with the difference that varying degrees of noise were added to the calculations in the ASM. It was expected that, as the amount of noise increased in magnitude, the performance of the mechanism would be increasingly impaired (because its calculations would be increasingly interfered with). The results of the tests with increasing noise (all but the left-most point in the graph) indicate that this was indeed the case, and that the SE does possess a capacity to discriminate between ASMs that are inherently better or worse at action selection.

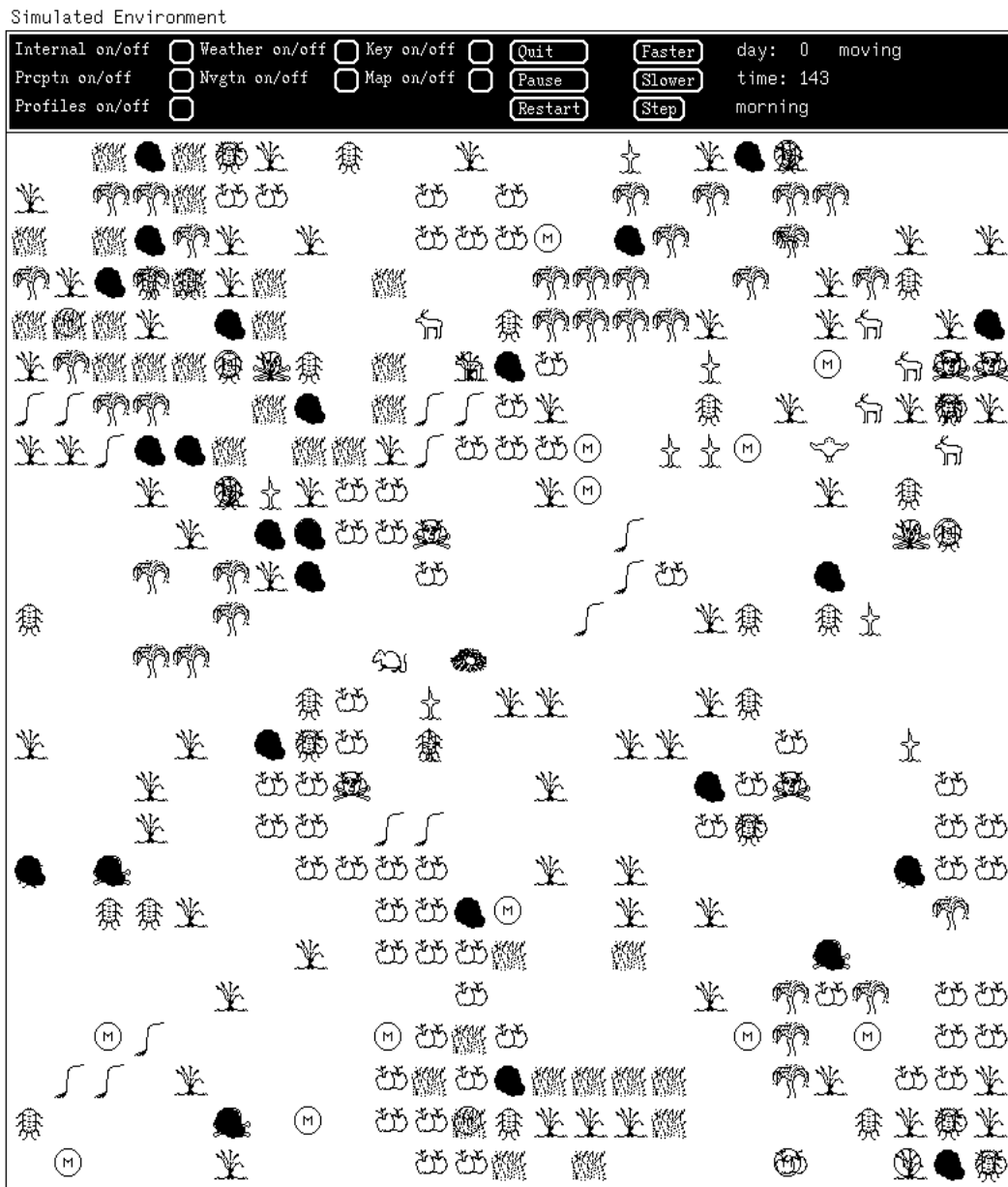


Figure 13: A randomly generated instance of the simulated environment showing the positions of features, the position of the animal, the current action of the animal, the time, the day and the current part of the day.

Key to Feature Types

Lighter shades imply greater value to the animal













-  - water source
-  - toxic food or water source
-  - cover
-  - shade
-  - dangerous place
-  - landmark
-  - cereal type food
-  - fruit type food
-  - den
-  - irrelevant animal (just needs avoiding)
-  - mate
-  - predator (type 1)
-  - predator (type 2)
-  - prey
-  - animal whose behaviour is being modelled

Figure 14: Key to features in the simulated environment.

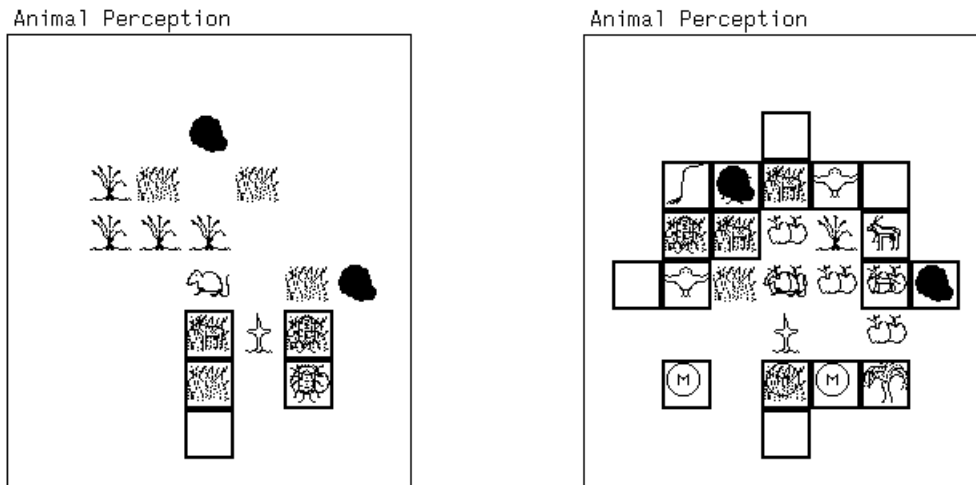


Figure 15: Two instances of the animal's perception of its local environment. Boxes around squares indicate that the contents of the square have been incorrectly perceived. The perception on the left occurred when the animal was not in vegetation and chose an action to look around. The instance on the right occurred when the animal chose an eating action while in a square containing fruit type food (which is a type of vegetation as well as food).

3.2. The Testing Procedure.

The procedure that was used to evaluate an ASM in the SE is now described:

- (1) a theory of action selection, or a more precise specification, was used to construct a mechanism which could deal with the 13 sub-problems in the SE.
- (2) the performance of the ASM was observed directly in the SE, and any obvious and gross inadequacies in the choices of actions (such as never choosing consummatory actions, or never attending to a particular sub-problem) were noted.
- (3) adjustments were made to the ASM so as to remove any gross deficiencies in performance, if this was possible without changing the whole nature of the ASM.
- (4) after removing any gross problems in the action selection, then the performance of the ASM was fine-tuned by adjusting parameters by hand. For instance, in the case of a free-flow hierarchy, connection strengths (weights) would be altered so as to try and maximise performance.
- (5) once the parameter values of the ASM had been finalised then its performance was measured in the SE by carrying out about 1650 tests, each with a randomly different initial set-up of the simulated environment. The average of the 1650 tests was taken as the measure of the performance of the ASM.

There are several factors which reduce the significance of the results obtained: *(i)* the choice of the sub-problems in the SE is biased away from social interactions, *(ii)* the authenticity of the models of perception, navigation and motor control cannot be guaranteed, even though they were carefully constructed, *(iii)* the parameter values in each ASM had to be set by hand and fine-tuned using trial and error, and so cannot be guaranteed to be optimal, and *(iv)* because of the randomness in the SE, the results obtained are only approximate measurements of the ability to select actions appropriately.

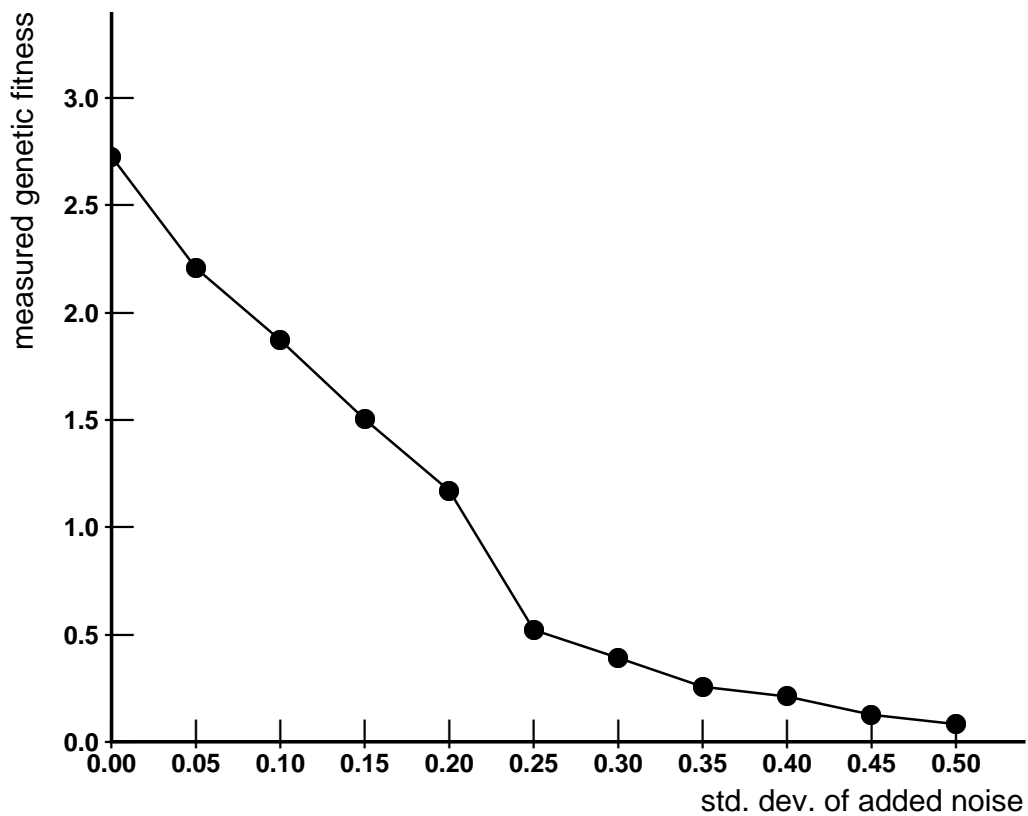


Figure 16: Performance (as measured by the simulated environment) of several mechanisms of the same type, but in which the calculations have been interfered with to varying degrees. Each point in the graph was obtained by averaging the results of 300 tests with that amount of added noise.

The factors just outlined mean that it is not possible to use the SE to make a totally reliable comparison between the performance of two ASMs, but the detrimental effect of some of the factors above can be reduced by: (i) some animals (e.g. snakes, hedgehogs, tigers) do lead lives which are for the most part solitary, (ii) the ASMs were tested in several different versions of the SE, in which different models of perception, navigation and motor control are incorporated. If differences in performance are maintained across different versions of the SE, with different models of the three processes, then the differences are likely to be due to inherent differences in the ability to select the best actions at each moment in time, (iii) approximately equal time was given to the optimisation of parameter values for each ASM, and (iv) the statistical significance of a difference in average performance between two ASMs was gauged according to the value of the term

$$\left(\frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right) \quad (1)$$

where the two ASMs have been tested n_1 and n_2 times respectively, producing average results of \bar{x}_1 and \bar{x}_2 and variances of s_1^2 and s_2^2 . If the value of this term is greater than 1.96 then the observed difference between \bar{x}_1 and \bar{x}_2 is, with 95% probability, due to an underlying difference in performance, rather than a random fluctuation between the results from two ASMs with essentially similar performances. Likewise, there is a 99% probability of a real rather than an apparent difference if the value of this term is greater than 2.576.

As a rule, comparisons between average results in the SE only constitute fairly weak evidence of a real, inherent difference between the ability of two different ASMs to choose actions appropriately in the real world. Although the SE described in the previous section is probably fairly realistic, and although the testing procedure described in this section is fairly thorough, the SE is still only a *model* of what is going on in the real world and so performance results obtained with it cannot be guaranteed to be applicable to the real world, and should be treated with a degree of scepticism.

3.3. How to use Free-Flow Hierarchies.

When a simple first attempt was made to implement a Rosenblatt & Payton-type FFH, and this ASM was then tested in the SE, it became obvious that some changes needed to be made to it because its performance was very poor, and it was observed that the ASM hardly ever selected consummatory ('achieving') actions. The changes that had to be made, together with the reasons why, are described in this section.

The initial and simple formulation of the Rosenblatt & Payton-type FFH was such that all preferences were combined according to the rule

$$A_j = \sum_{i=1}^N (P_{ij})$$

where A_j is the activation of node j , N is the number of positive and negative preferences for node j , and P_{ij} is a preference for node j from node i . That is to say, the activation of node j was calculated as a straightforward sum of all the preferences for it. When the implementation of this FFH was tested in the SE the performance was extremely poor. It could be seen that the animal nearly always chose move-type actions (such as 'move west' or 'move south-east'), and only very rarely chose consummatory actions (such as 'drink', 'clean' or 'mate'). And so the animal moved around the environment a lot, and frequently approached useful water sources, useful food sources or mates, but hardly ever stopped to utilise them.

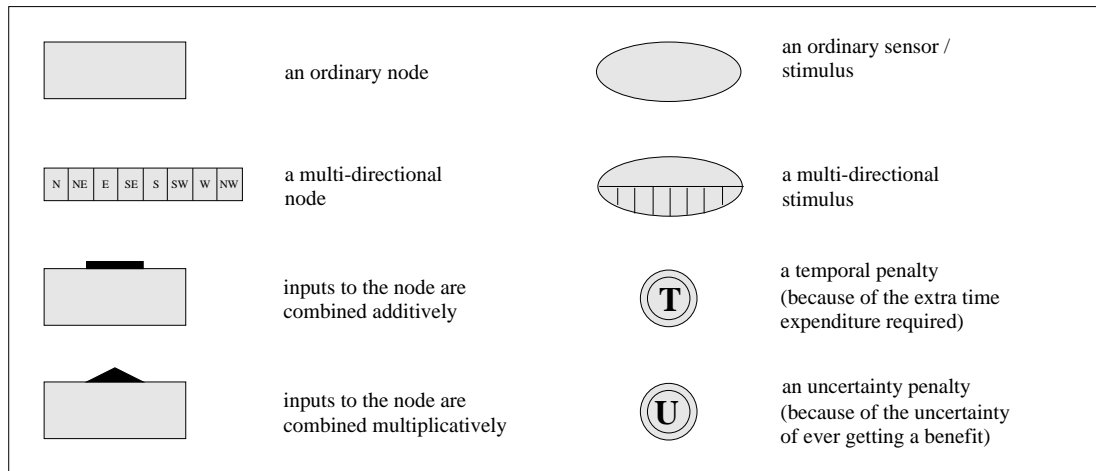


Figure 17: Key to free-flow hierarchy figures. When a node is shown with neither a solid triangle nor a solid rectangle above it then the node has a more complicated rule for combining inputs. Multi-directional nodes and stimuli are ones with eight components, one for each of eight compass directions.

There are three computational reasons behind this deficiency in action selection:

- (1) The first fault with this simple implementation of an FFH, is that there is no in-built tendency to prefer consummatory over appetitive options. All other things being equal, appetitive (preparatory) options are less desirable than consummatory (achieving) ones because (i) a consummatory option will produce a real and immediate benefit to the animal (e.g. from ingestion of food), whereas an appetitive option will only help to bring about a situation in which a consummatory option can be undertaken, and so will only lead to an actual benefit after a *further expenditure of time* (e.g. approaching food is only valuable after the animal has also expended some more time eating the food), and (ii) an appetitive option may not necessarily be followed by a consummatory one. The potential benefit that is produced by the appetitive option (e.g. the opportunity to eat food that is produced by approaching food) may never be realised if, for example, the animal is interrupted by the appearance of a predator and has to move away.
- (2) The second fault is that the preferences are summed, even though they may all come from the same system. Imagine the case in figure 18. Here the action node ‘move west’ is receiving preferences from ‘approach perceived food’, ‘approach remembered food’ and indirectly from ‘explore for food’, but these three inputs are all originally triggered by the same goal. But if the animal encounters the perceived food first then it will very possibly become satiated and not continue to approach the remembered food or explore for food afterwards. The utility of proceeding in a particular direction in order to approach perceived food, approach remembered food and explore simultaneously is not equal to the sum of the utilities of each individually.
- (3) The third fault with the naive formulation of the Rosenblatt & Payton FFH is that preferences are summed, even when they all come from appetitive nodes. Consider the situation in figure 19. Here the action ‘move north-east’ receives preferences from the nodes ‘approach receptive mate’, ‘approach perceived water’ and ‘approach

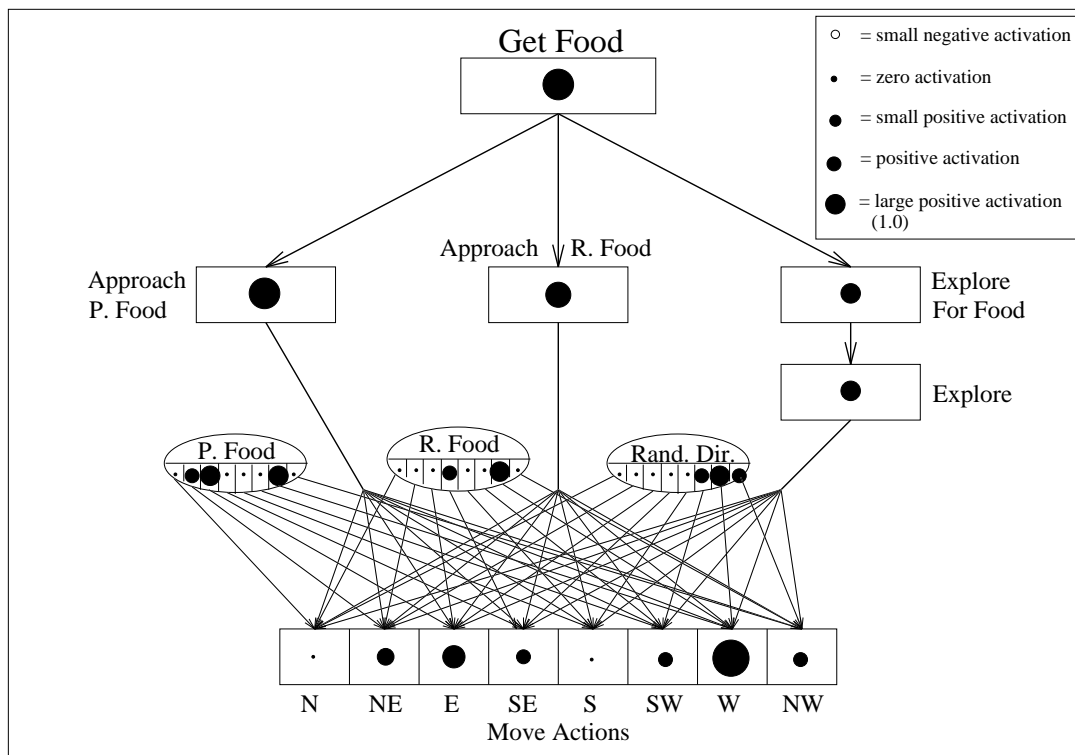


Figure 18: Preferences from the same system onto the same action node. 'move west' receives excitation from 'approach perceived food', 'approach remembered food' and 'explore for food'. An explanation of the meaning of some of the symbols in the diagram can be found in figure 17. Each preference for each of the move actions (e.g. for 'move north-west') is the product of an input from a higher-level node and an input from the relevant component of a multi-directional stimulus. Only the appetitive alternatives for the system are shown here, and inputs from other systems are not shown. 'P' and 'R' refer to perceived and remembered respectively.

shade'. The three preferences all originate from different goals in this case, but the utility should still not be equal to the *sum* of the preferences from the three nodes. An appetitive alternative has utility because it brings about a situation in which a consummatory action can be performed. But because of the constant possibility of interruption of any appetitive/consummatory sequence (by the need to avoid predators or some other urgent demand) the utility of performing an action which brings about three appetitive options simultaneously is much less than the sum of the utilities of performing them each individually. This is because the probability of performing the joint appetitive action and then eventually continuing to perform consummatory actions for all three systems is fairly remote. For instance, by the time the animal has approached, courted and copulated with the mate, and has then approached and drunk the water (assuming it gets to do so) then the probability of it still being uninterrupted and able to approach the shade is fairly low. In any case, its body temperature may have improved of its own accord by then, making it no longer necessary to approach shade. In general, the longer the likely gap between an appetitive and a consummatory action, the less likely that the one will lead to the other (because of the dynamic, rather unstable and unpredictable nature of the environment), and so the lower the expected value of performing the appetitive action. For this reason, an action which receives preferences from several appetitive nodes should have a utility which is significantly less than the sum of all the individual preferences for it.

These three faults have the effect of making nodes which receive many preferences (particularly the 'move' actions in the case of the SE) get more activation than they should.

Two changes were made to the implementation to remedy the faults just outlined:

- (1) A scheme for inhibiting appetitive options was introduced, as shown in figure 20. The scheme is based on Sutton & Barto's idea of temporal discount factors [e.g. Sutton, 88]. Each appetitive option is inhibited by a *temporal penalty* (T in the figure), which is proportional to the amount of extra time that the animal will probably have to expend in order to obtain a benefit, as well as an *uncertainty penalty* (U in the figure), which is proportional to the probability of the appetitive option *not* eventually leading to the execution of the consummatory option.
- (2) The rule for combining preferences was changed from

$$A_j = \sum_{i=1}^N (P_{ij})$$

to

$$A_j = \left(\frac{\max_i(P_{ij}^+) + \alpha \sum_{i=1}^{N^+} (P_{ij}^+)}{1 + \alpha} \right) + \left(\frac{\min_i(P_{ij}^-) + \beta \sum_{i=1}^{N^-} (P_{ij}^-)}{1 + \beta} \right)$$

where P_{ij}^+ and P_{ij}^- are the positive and negative preferences from node i for node j , N^+ and N^- are the numbers of such preferences for node j , α and β are constants, and $\max_i()$ and $\min_i()$ are functions which calculate the most positive and most negative of the preferences for node j . This new rule for combining preferences worked best in the SE with fairly low values of α and β (e.g. $\frac{1}{(N^+)^2}$ and $\frac{1}{(N^-)^2}$ respectively). This new rule reduces the activation received by a node with many strong preferences for it, although the activation received is always greater than or equal to the size of the largest positive preference (assuming that there are no negative preferences).

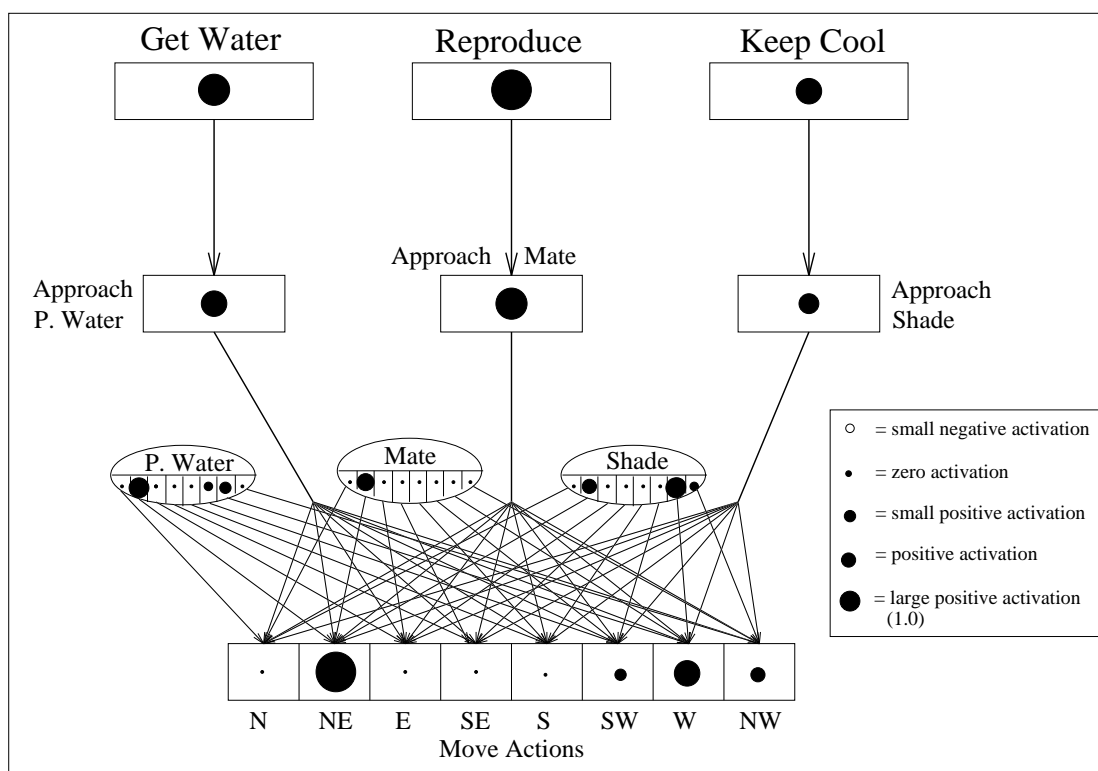


Figure 19: Preferences from appetitive nodes in different systems to the same action node. 'move north-east' receives preferences from 'approach perceived water', 'approach perceived mate' and 'approach perceived shade'. An explanation of the meaning of some of the symbols in the diagram can be found in figure 17. Inputs from other systems are not shown.

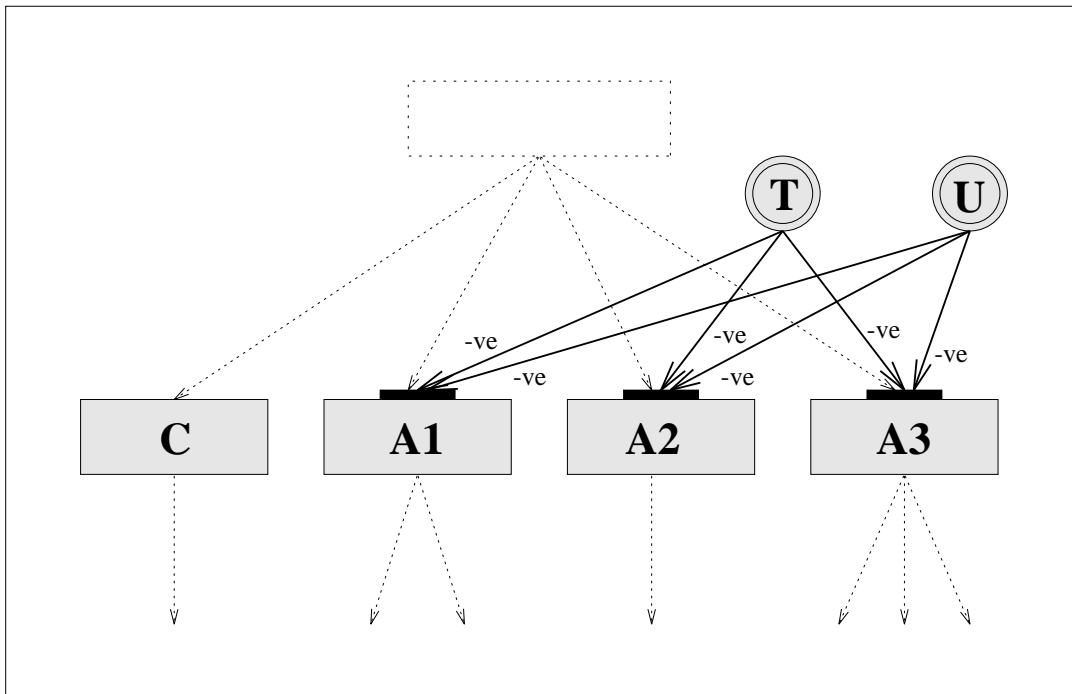


Figure 20: A scheme to allow consummatory alternatives to be favoured over appetitive ones, and more certain and more immediately rewarding appetitive alternatives to be favoured over others. The connections from the 'T' node convey temporal penalties, the connections from the 'U' node convey uncertainty penalties. The penalties are larger for *A3* than *A2*, and larger for *A2* than *A1*.

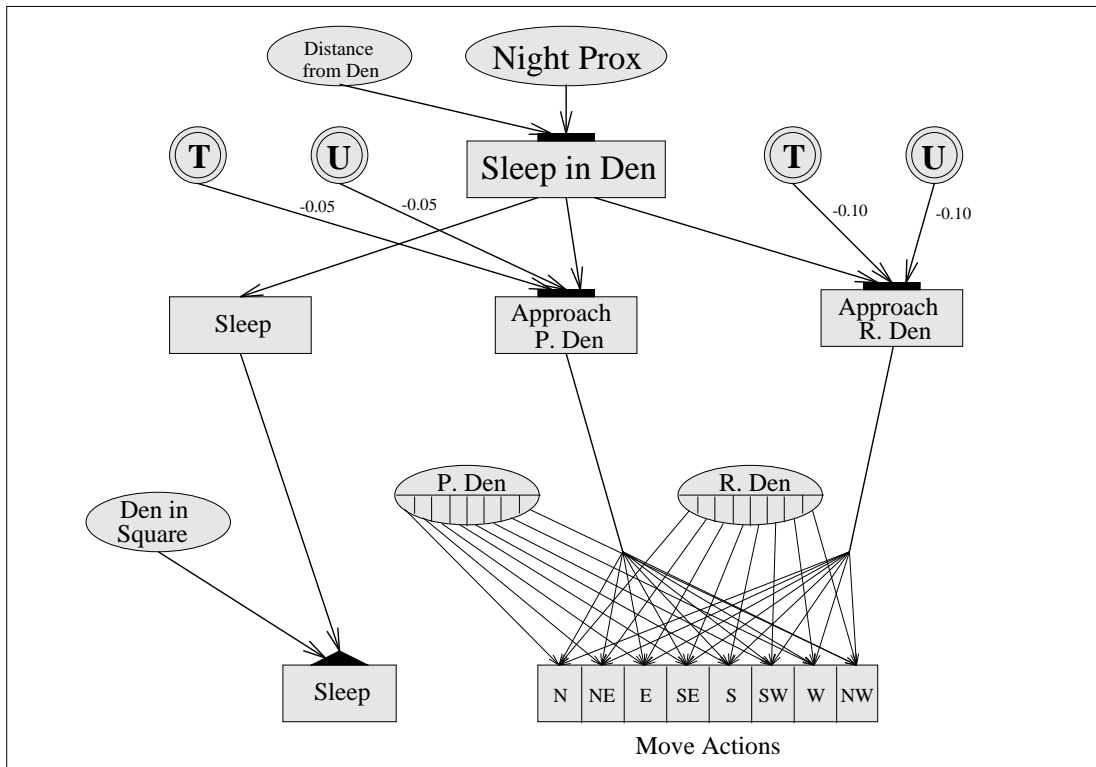


Figure 21: A possible system from an extended Rosenblatt & Payton-type free-flow hierarchy. This system makes the animal tend to move towards its den as nightfall approaches, and then sleep in the den for the rest of the night. The stimulus ‘night proximity’ is 0.0 at daybreak, then increases as nightfall approaches, and stays at its maximum value through the night. The stimulus ‘den in square’ is 1.0 if the animal is in the same square as its den, and is 0.0 otherwise. ‘P’ and ‘R’ refer to perceived and remembered respectively. The two penalties have activation 1.0, which is multiplied with a small negative connection weight to provide inhibition to the appetitive nodes.

The modifications that needed to be made to the first implementation of Rosenblatt & Payton’s ASM have now been justified and described. More extensive justifications of the use of temporal and uncertainty penalties are given in [Tyrrell, 93], together with further discussion of the particular form of the rule for combining preferences that was chosen.

Two possible systems from a viable FFH are shown in figures 21 and 22.

A hierarchical decision structure was also implemented. The HDS which was tested was not a direct implementation of either Tinbergen’s or Baerends’ ideas, both of which are somewhat imprecise and unclear, but instead was an implementation of the general concept. The HDS had an identical structure to that of the modified FFH (including the temporal and uncertainty penalties), but the weights on the connections (between nodes and from stimuli to nodes) were optimised separately. The rule for combination of preferences was changed to

$$A_j = \max_i (P_{ij}^+)$$

and negative preferences were not allowed. Negative preferences would serve no purpose

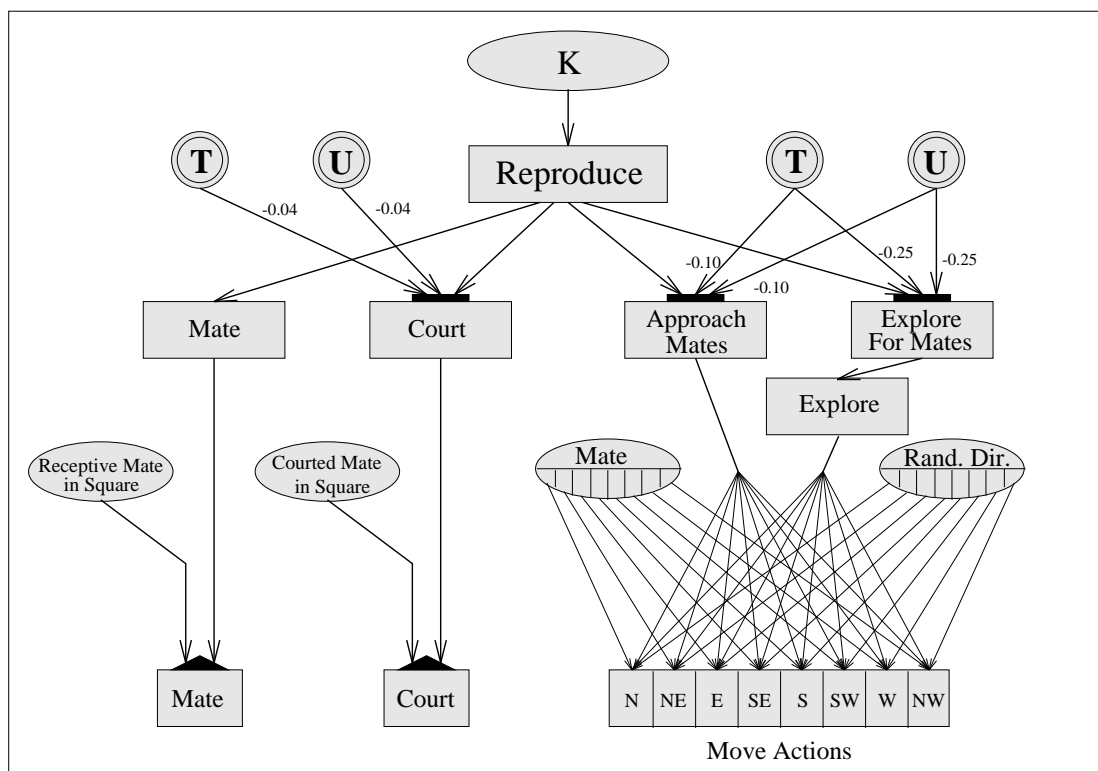


Figure 22: Another possible system from an extended Rosenblatt & Payton-type free-flow hierarchy. This system makes the animal reproduce, when it influences the animal's selection of actions. 'K' is a constant stimulus which provides an unvarying input to the system-level node.

		Standard SE	SE v1.	SE v2.	SE v3.	Average
HDS	av. fitness (\bar{x})	6.69	3.48	6.48	7.79	6.11
	std. dev. (s)	6.24	3.77	6.39	7.43	5.96
	num. tests (n)	1641	1679	1693	1587	total no. = 6600
	\bar{x} std. dev. ($s_{\bar{x}}$)	0.15	0.09	0.16	0.19	0.15
Extended R&P	av. fitness (\bar{x})	8.09	3.61	8.16	13.38	8.31
	std. dev. (s)	7.06	3.63	6.66	9.16	6.63
	num. tests (n)	1649	1693	1653	1605	total no. = 6600
	\bar{x} std. dev. ($s_{\bar{x}}$)	0.17	0.09	0.16	0.23	0.16

Table 1: The average genetic fitnesses for the extended Rosenblatt & Payton-type FFH and for the HDS, when tested in the four versions of the SE. Each test in a version of the SE consists of n runs, which have a mean performance of \bar{x} and a standard deviation of s . The expected standard deviation of \bar{x} (i.e. $s_{\bar{x}}$) is equal to (s/\sqrt{n}) .

in an HDS because there would be no positive preferences from other systems for them to counteract. The lack of a positive preference in an HDS has the same effect as a negative preference would.

3.4. Performance Results in the Simulated Environment.

The improved Rosenblatt& Payton-type FFH (with temporal and uncertainty penalties, and with a more complex rule for combining preferences) was tested in four versions of the simulated environment: a standard one, and three other versions with different models of the processes of perception, navigation and motor control. The HDS described in the previous section was also tested. In both cases there were no obvious or major deficits in performance. The performance measurements (averages from more than 1500 runs in each case) are given in table 1 and shown graphically in figure 23. Both types of hierarchy performed better than Maes' ASM.

It can be seen that the FFH selects actions more appropriately than the HDS in the SE, and that the superiority in performance is maintained throughout all four versions of the SE. Putting the average results for the FFH and the HDS into equation 1 (section 3.2) gives

$$\left(\frac{|(8.31 - 6.11)|}{\sqrt{\frac{69.6}{6600} + \frac{37.3}{6600}}} \right) = 13.4.$$

13.4 is greater than 2.576 and so there is a probability of greater than 99% of a real difference between the action selection of the two ASMs.

As discussed previously, the simulation testing can only give rather weak evidence of the superiority of free-flow hierarchies over hierarchical decision structures. However, it should be kept in mind that the simulation results do agree with the hypothesis, and the simulated environment is a fairly painstaking attempt to model the action selection problem for animals. The SE provides a complex and demanding test of action selection, with sub-problems which differ in many ways (some are homeostatic, some are not, some are urgent, some are not, some are prescriptive, some are proscriptive, etc [Tyrrell, 92]), and it is important that the improved FFH described was able to choose actions appropriately in this context, and was able to do so more effectively than the HDS.

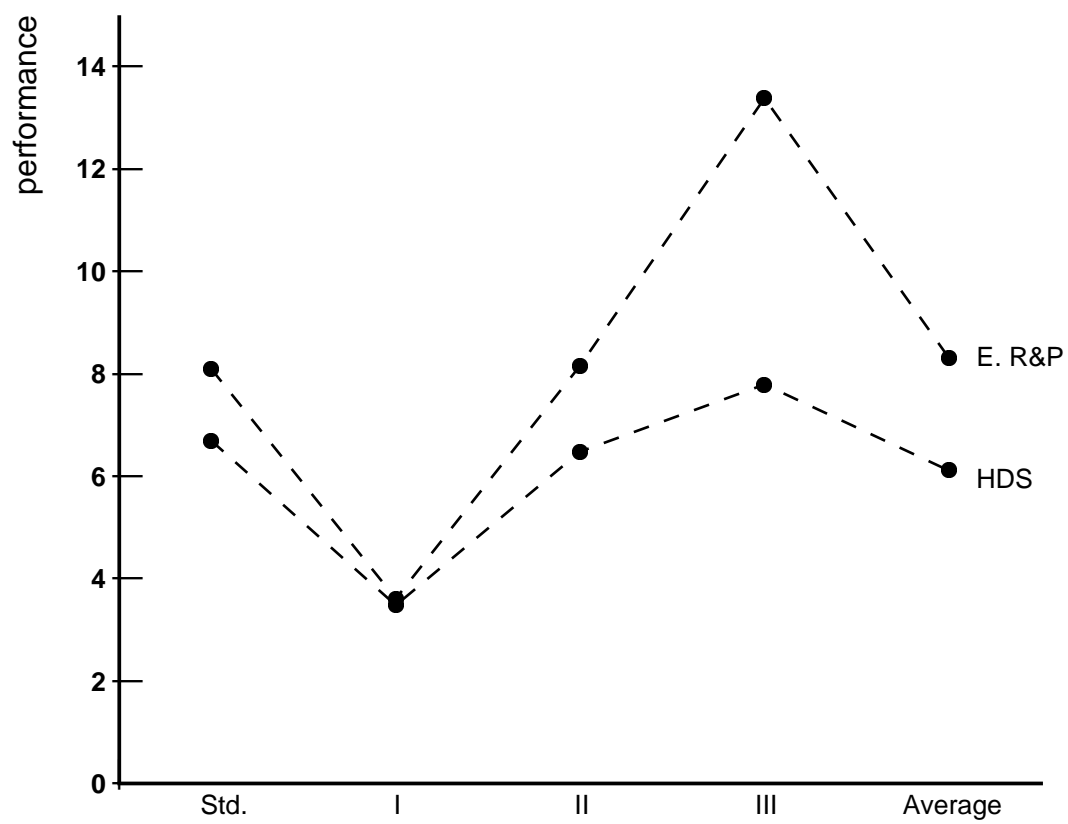


Figure 23: The results (genetic fitnesses) of testing the FFH and the HDS in a computer simulation of the action selection problem for an animal. **Std** stands for the normal SE, **I**, **II** and **III** stand for the three altered versions of the SE. Each point in the graph is the average of more than 1500 evaluations.

4. Supporting Evidence from Animal Behaviour.

This section presents some observations about animal behaviour, which give support to the central hypothesis of this paper, that the degrees of preference for different alternatives are propagated through all systems simultaneously, rather than the system with the highest activation shutting down all others. The three examples below come from pages 243–245 of [Lorenz, 1985].

1 : The simultaneous firing of antagonistic muscles due to simultaneous excitation of conflicting systems:

Additive superposition is found even in cases in which the two independent motivations activate antagonistic muscles. ... A conflict between motivations in Anatidae, one demanding a forward extension of the neck, the other a retraction, which can occur in a goose wanting to eat grain offered in a human hand and not quite daring to do so, produces a violent trembling of the neck. ... The “trembling neck” of the goose probably is effected by conflicting innervation of the antagonistic muscles, but we cannot be certain about this.

2 : The simultaneous operation of opposing sets of fins due to the strong activation of conflicting systems:

In territory disputed, certain cichlids (*Etroplus maculatus*) position themselves opposite another, threatening across the border separating their territories. As in every threat, aggressive motivation is contending with that for escape. Whenever one of the adversaries moves a short distance forward into enemy territory, it appears as if he were swimming into a current, the speed of which rapidly increases as one proceeds upstream. This effect is produced by the action of the pectoral fins, which are sculling in reverse, and doing so more and more intensely the farther the fish moves into the other’s territory. The tail fin is under the control of aggressivity and the pectorals under that of escape, and the observer cannot help feeling – ridiculously – that the pectorals are more afraid than the tail, because they are nearer to the enemy.

3 : The facial expressions of a dog can simultaneously display the intensity of two separate, conflicting motivations, fear and aggression (figure 24).

These three examples all serve to illustrate the point that the activation levels of different nodes in a hierarchy are likely to be maintained, as in a free-flow hierarchy, rather than only one at each level surviving, as in the winner-take-all operation of a Tinbergen-like hierarchical decision structure. As well as the three examples stated above, Lorenz went further to state that

In nature, behaviour activated by a single motivation is found at least as rarely as hybrids differing in only one gene. A higher animal in its natural habitat must always be ready to undertake a great number of different and – as often as not – mutually exclusive actions, and what it finally does is almost always a compromise made among several necessities.

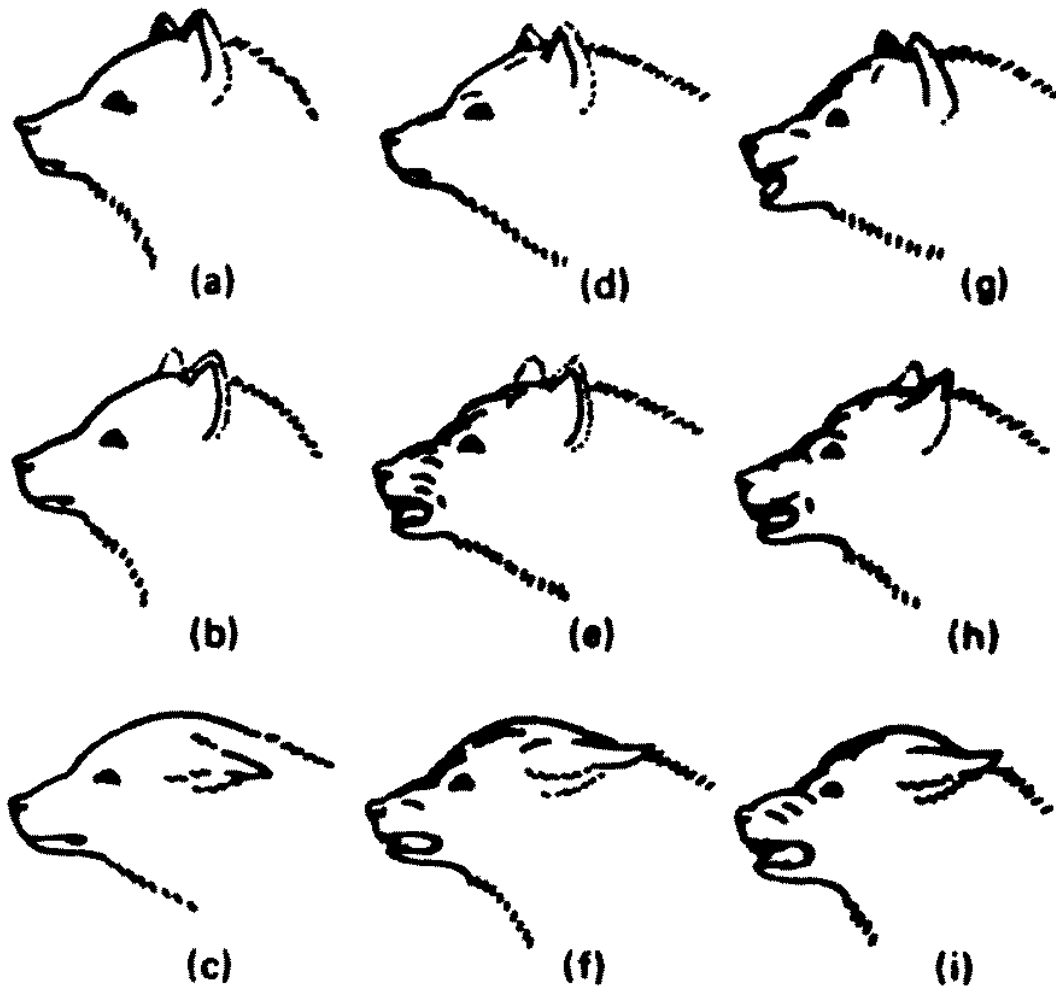


Figure 24: Facial expressions of a dog showing differing degrees of fear and aggression. Readiness to flee increases from top to bottom, aggression increases from left to right. Taken from [Lorenz, 81] (© acknowledged to Springer-Verlag).

5. Conclusions.

The question that has been addressed in this paper is ‘how to use hierarchies for action selection?’. Two alternatives have been described and compared: the traditional *hierarchical decision structure* (such as has been proposed by Tinbergen and Baerends), in which switching decisions are made at every level; and the more novel *free-flow hierarchy* (proposed by Rosenblatt & Payton), in which there is unrestricted flow of information, combination of preferences and the possibility of compromise candidates. Several computational arguments have been put forward in favour of free-flow hierarchies, and supporting evidence has been described, coming from a computer simulation study and from animal behaviour observations.

Some necessary extensions to the simplest form of free-flow hierarchy were described, but the most important conclusion is that these modified free-flow hierarchies are better at action selection than hierarchical decision structures. This conclusion challenges a long-held preconception about how hierarchies should be used for calculating action selection.

References

- Arkin, Ronald C. 1987. Motor Schema Based Navigation for a Mobile Robot: An Approach to Programming by Behavior. In *Proceedings of the 1987 IEEE International Conference on Robotics and Automation*. Computer Society Press of the IEEE.
- Baerends, Gerald P. 1976. The Functional Organization of Behaviour. *Animal Behaviour*. 24:726–738.
- Brooks, Rodney A. 1986. A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation* 2:14–23.
- Brooks, Rodney A. 1987. Intelligence Without Representation. In *Proceedings of the Workshop on Foundations of Intelligence*. MIT Press.
- Dawkins, Richard. 1976. Hierarchical Organisation: A Candidate Principle for Ethology. In *Growing Points in Ethology*. Bateson P. P. G. & Hinde R. A. eds. Cambridge University Press.
- Dawkins, Richard. 1989. *The Selfish Gene* (2nd edition). Oxford University Press.
- Dewsbury, D.A. 1978. *Comparative Animal Behaviour*. McGraw Hill.
- Hinde, R. A. 1953. Appetitive Behaviour, Consummatory Act, and the Hierarchical Organisation of Behaviour – with Special Reference to the Great Tit (*Parus Major*). *Behaviour* 5:189–224.
- Kilmer, W.L., McCulloch, W.S. & Blum, J. 1969. A Model of the Vertebrate Central Command System. *International Journal of Man-Machine Studies* 1:279–309.
- Lorenz, Konrad. 1985. *Foundations of Ethology*. Heidelberg: Springer-Verlag.
- Maes, Pattie. 1991. Bottom-Up Mechanism for Behaviour Selection In An Artificial Creature. In *Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, (JA Meyer & SW Wilson, eds). MIT Press/Bradford Books.
- McFarland, D.J. & Sibly, R. 1975. The Behavioural Final Common Path. *Phil. Trans. R. Soc. Ser. B*. 270:265–293.
- McFarland, D.J. 1985. *Animal Behaviour*. Longman.
- Morris, D. 1990. *Animal Watching*. Jonathan Cape.
- Rosenblatt, Kenneth J & Payton, David W. 1989. A Fine-Grained Alternative to the Subsumption Architecture for Mobile Robot Control. In *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks*.
- Sutton, R.S. 1988. Learning to Predict by the Methods of Temporal Difference. *Machine Learning* 3:9–44.
- Tinbergen, Nikko. 1950. The Hierarchical Organisation of Nervous Mechanisms Underlying Instinctive Behaviour. *Sympos. Soc. Exper. Biol.* 4:305–312.
- Tinbergen, Nikko. 1951. *The Study of Instinct*. Clarendon Press.
- Tyrrell, T. 1992. Defining the Action Selection Problem. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.
- Tyrrell, T. 1993a. *Computational Mechanisms for Action Selection*. PhD thesis. Available as In-House Publication (technical report) no. EUCCS/PHD 72, Centre for Cognitive Science, University of Edinburgh.