

An Evaluation of Maes' "Bottom-Up Mechanism for Behavior Selection"

Toby Tyrrell

Plymouth Marine Laboratory

Plymouth PL1 3DH, UK.

email: tt@pml.ac.uk

Abstract

Maes has proposed a mechanism for action selection (behavioural choice) [Maes, 89; 90; 91a]. This mechanism is reviewed here and is evaluated using a simulated environment. This simulated environment is a detailed and complex generalised model of the action selection problem faced by an animal in the wild, and presents a rather severe test for an action selection mechanism. The results of testing Maes' mechanism in the simulated environment are discussed, some observed deficiencies in the mechanism's operation are described, and the computational reasons underlying the deficiencies are explained. It is argued that some central aspects of the design of Maes' mechanism mean that it is not well able to deal with animal-like action selection problems.

1. Introduction.

The question of how behaviour is produced and controlled in animals is of interest to ethologists, psychologists and neuroscientists. Similarly, researchers in artificial intelligence (and more particularly robotics) are interested in how behaviour can be artificially generated in machines. The study of action selection (AS), i.e. behavioural choice, is therefore of interest to many researchers in many different fields.

Ethologists and roboticists in particular have put forward proposals about the structure of the action selection mechanisms (ASMs) that might underly observed animal behaviour [e.g. Lorenz, 50; Tinbergen, 50; Baerends, 76], or be able to produce adaptive behaviour in robots [e.g. Kaelbling, 87; Rosenblatt & Payton, 89]. There has been some difficulty in evaluating the suitability of the different alternative theories of AS though. Ethologists can only resort to ‘thought experiments’, and this verbal speculation about mechanisms when there is no means of testing the speculations has been criticised as “armchair theorising” [McFarland & Houston, 81]. Roboticists can test their theories in behaving robots, but these robots usually have only a rather limited perception of the environment and a rather limited action repertoire (often consisting of locomotion only). Verbal speculation about mechanisms when there is no means of testing the speculations has been criticised as “armchair theorising” [McFarland & Houston, 81]. One way of overcoming some of these limitations is through the use of simulated environments (SEs), one of which is described and used in this paper. Although there are also limitations associated with the use of SEs (which are discussed later in the paper), they can be used to provide a complex and challenging test of AS; one which may be capable of exposing hitherto unsuspected deficiencies in the performance of ASMs.

This paper looks at one particular mechanism for calculating AS, that proposed by Maes [89; 90; 91a]. Maes’ action selection mechanism (MASM) was originally proposed as an improvement over previous approaches to AS in the field of artificial intelligence. In particular, it was proposed as an improvement upon traditional planning systems (see [Agre, 90; Agre & Chapman, 87] for a review and critique of such systems) and reactive systems [e.g. Kaelbling, 87]. The former are unsuitable for complex and multi-faceted animal-like AS problems for a number of reasons, one of which is that they proceed on the basis of rigid, pre-formed plans and pay little or no attention to changes of circumstance or new opportunities in the environment. The latter are unsuitable because they err in the opposite direction. They are stimulus-driven and take little or no account of internal motivations when formulating their choices of actions. MASM is an improvement over these approaches because it bases its calculations on both motivations and external stimuli.

Although MASM was originally proposed in the AI literature, and as an alternative to existing AI approaches to AS, Maes considered the model (along with other AI models) to be of more general relevance:

“Computational models of action selection are important both for producing actions in an artificial agent (e.g. an autonomous agent mobile robot) as well as for understanding the action selection behavior of biological agents (human and animal). In previous papers we argued that none of the models of action selection that have been presented in the lit-

erature on AI present a satisfactory solution to the problem (Maes, 1991a, 1991b) . . . In (Maes, 1989a, 1989b, 1991a, 1991b) we presented an action selection algorithm which tries to combine the best of previous models.” [Maes, 91c: page 108]

In particular, one of the papers on MASM [Maes, 91a] contains a discussion of how MASM deals with an animal-like AS problem, one in which there are motivations of CURIOSITY, LAZINESS, HUNGER, THIRST, AGGRESSION, FEAR and SAFETY; and appetitive and consummatory actions such as GO_TO_FOOD, EAT and SLEEP. A similar, but more complex, animal-like AS problem will be used to test MASM in this article.

Some previous tests of MASM have looked at its performance on a simple problem of choosing actions in the correct sequence so as to sand a board and spray-paint itself [Maes, 89]; at its performance on a ‘blocks world’ problem [Maes, 89]; on the animal-like AS problem just mentioned [Maes, 91a]; and on an AS problem for a Lego vehicle [Pebody, 91]. In addition: “We have evaluated the behavior selection mechanism presented above empirically by performing a wide series of experiments in simulations and robots” [Maes, 91c: page 110]. No important deficiencies were reported from any of these tests, beyond an occasional tendency for the mechanism to get stuck in loops [Maes, 90].

The fairly challenging test of AS employed in this paper did point to some inadequacies in the way that MASM selects actions. It is argued that MASM, while being an improvement over previous approaches, can itself be improved upon in some respects. At the same time, it is noted that the MASM ideas are presented in a commendably formal and explicit manner, thus making it easier to understand and critically analyse the mechanism.

The rest of this paper is arranged in the following way: section 2 considers the nature of the AS problem faced by animals; section 3 describes an embodiment of a generalised ‘animal-like’ AS problem in a computer simulation, and looks at how this SE can be used to test the validity and suitability of different ASMs; section 4 describes how MASM calculates selections of actions; section 5 describes how MASM was implemented to deal with the specific AS problem posed by the SE; section 6 presents the results of testing MASM in the SE, and traces the deficiencies that became apparent to their fundamental causes (which are shown to be due to aspects of the mechanism rather than due to the particular test employed); section 7 compares some aspects of MASM with the corresponding aspects of other ASMs; and section 8 gives some brief conclusions. For those unconcerned with the details of the testing and with some prior knowledge of MASM, it is possible to omit sections 3, 4 and 5 of this paper.

2. The Action Selection Problem for Animals.

Before examining the suitability of MASM for ‘animal-like’ AS problems, an ‘animal-like’ AS problem must be defined:

Animals are vehicles for propagating their genes into future generations [Dawkins, 89]. The ‘purpose’ of an animal’s existence is to maximise the expected number of its genes in future generations. This can be done in either or both of two ways: *(i)* reproducing as frequently and as effectively as possible themselves, and making sure that they stay alive in order to be able to continue to do so; and/or *(ii)* helping close genetic relatives (individuals with very similar sets of genes, such as offspring, siblings and parents) to do the same. For some animals, such as worker ants, their genes are propagated entirely by proxy, but for most animals it is the former way that predominates. Their main purpose is to maximise the number of offspring they produce during their lifetime, which engenders a need to live for a long time.

The paragraph above gives a high-level description of the sort of behaviour an animal has to produce in order to maximise its future expected genetic fitness. This high-level AS problem in practise decomposes to many *sub-problems*. In order to stay alive the animal has to concern itself with these different sub-problems, such as how to minimise the risk of detection by predators, how to get enough food, how to get enough water, how to keep its body temperature within acceptable limits and how to avoid hazardous areas of the environment.

What sub-problems, or groupings of functionally-related behaviours, are commonly seen in animals? Several references [Dewsbury, 78; Kilmer *et al*, 69; Morris, 90; and McFarland, 85] were scrutinised and the following list of commonly-observed sub-problems was extracted: care of offspring (including feeding them, protecting them, cleaning them, etc), body maintenance (cleaning of self by licking, preening, self-grooming, etc), communication with conspecifics (for example, alarm calls, threatening, appeasing), escaping from predators (by freezing, fleeing or fighting), getting food (including exploring for food, foraging, preparing food, predation and eating), getting water, grouping together with conspecifics, reproduction (including courting displays and copulation), sleeping, thermoregulation, territorial defence and vigilance (for predators).

For most of these different sub-problems there are *consummatory* (achieving) and *appetitive* (preparatory) alternatives. For instance, for the getting food sub-problem the consummatory alternative is eating, since this directly increases the amount of food in the animal’s stomach, from where it will proceed automatically into the bloodstream. Some appetitive alternatives, depending on the particular animal, might be exploring for food, stalking prey, approaching food or preparing food by removing the shell or rind. These appetitive behaviours do not benefit the animal directly, but pave the way for it to be able to perform the consummatory alternative.

An *action selection mechanism* (ASM), as its name implies, has the purpose of selecting the most appropriate action at each moment in time. An *action* is the smallest, most primitive, unit of behaviour that is considered. Using the assumption of the *behavioural final common path* [e.g. McFarland & Sibly, 75], it is assumed that all actions are mutually exclusive, and therefore that only one action can be performed at a time. A *system* of an ASM is the part of the ASM that is devoted to the control

of behaviour related to a particular sub-problem. An ASM will therefore consist of a number of different systems (one to generate food-getting behaviour, one to generate vigilance for predators, etc). Some of these terms are illustrated in figure 1, which shows Baerends' tentative ASM to account for the observed behaviour of the herring gull [Baerends, 76].

To summarise, at any moment in time an animal (a mammal such as a dormouse, a reptile such as a monitor lizard, or a bird such as a herring gull) is faced with a decision as to which out of a set of mutually exclusive actions (appetitive and consummatory) it should choose, and this choice will be determined by the requirements of a set of sub-problems.

An example instance of an animal-like AS problem is shown in figure 2. The animal needs to take account of many factors (including the urgency of different sub-problems and the opportunities available in the environment) in order to make a sensible choice of action. The most suitable action in this case might be to run away to the north-east because of the urgency of the "escape from predators" sub-problem.

The AS problem can also be thought of as one of time-allocation. An animal has a limited amount of time available in each day, and has to apportion that time so as to satisfy the needs of all of the many sub-problems. The demands on the animal's time from different sub-problems will differ in their urgency, their frequency and their degree of periodicity [Tyrrell, 92].

2.1. Desiderata for an Action Selection Mechanism.

From this brief discussion of the AS problem for an animal, some desiderata for an ASM (requirements of an ASM) can be inferred:

1. The probability of consummatory or appetitive actions in a system being chosen should increase as the urgency or degree of need (the motivation) for the effects of that system increases. For example, the animal should be more likely to choose to eat food, or to approach food, or to explore for food, as it starts to run out of food (as it becomes more hungry).
2. The probability of an action being chosen should increase as the relevance of the action to the current situation increases. For example, the animal should only choose to eat food if there is food there to eat, and should be more likely to choose to approach food if it can perceive a lot of food close at hand, and the food is appetising.
3. Consummatory actions are more profitable than appetitive ones. A consummatory action provides an immediate direct benefit to the animal, whereas an appetitive action will always require further expenditure of time before any benefit is realised. Also, it is possible for an appetitive-consummatory sequence to be interrupted and left uncompleted, making the initial expenditure of time on the appetitive action(s) worthless. Therefore, if both a consummatory and an appetitive action are relevant to the current situation then, all other things being equal, the consummatory action should be chosen. This applies whether the consummatory and appetitive actions being arbitrated between are in the same system or in different systems.
4. Because the benefits of appetitive actions can be lost if the animal does not

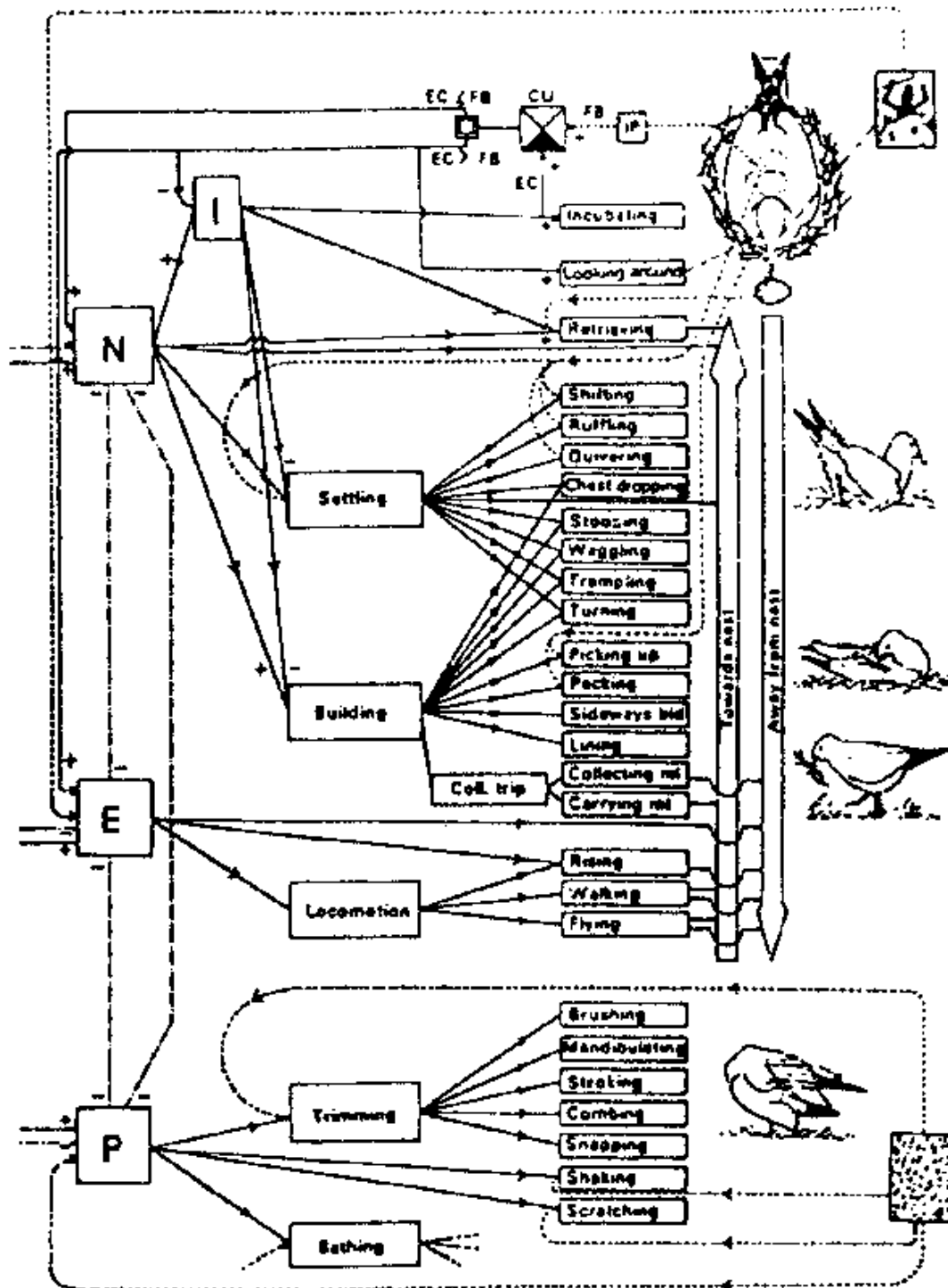


Figure 1: Baerend's model to account for incubation behaviour of the herring gull. Actions or fixed action patterns are in the far right column. "Superimposed control systems" of higher order are to the middle and left. **N** = incubation system, **E** = escape system and **P** = preening system. The three systems mutually suppress each other. The entities towards the top of the diagram (IP, EC, FB, CU, and I) form a control loop in which the feedback stimulation from the clutch (FB) is compared (in CU) to the desired stimulus value (EC). Different comparison results lead to the induction of different sorts of behaviour. Taken from [Baerends, 76] (© acknowledged to Bailliere Tindall).

INTERNAL STIMULI

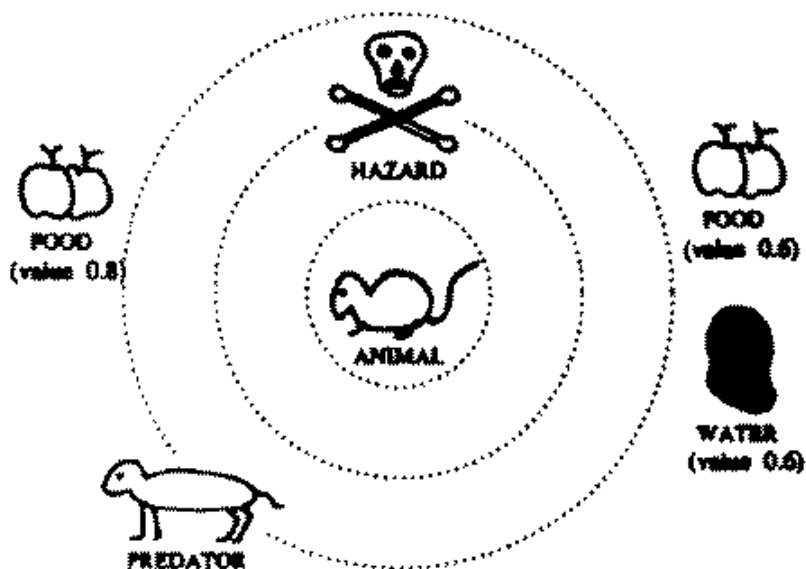
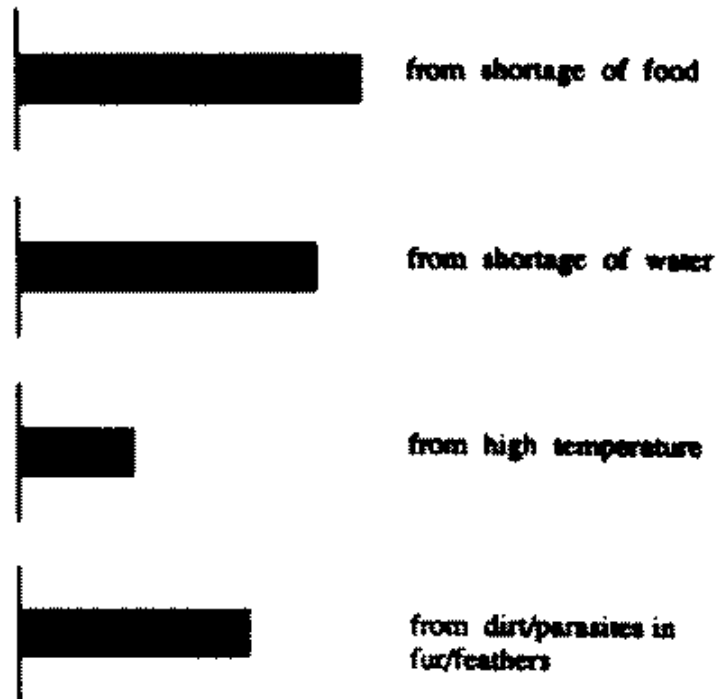


Figure 2: An example situation in which many different stimuli need to be taken into account in order to decide what the best action is. In this case the animal has several feasible actions to choose from: should it move south directly away from the hazard, move north-east directly away from the predator, move west towards the high-valued food source, move east towards the average-valued food and the average-valued water source, or stay where it is and clean itself? The stimuli shown here are offsets from optimal values, and the length of each bar is proportional to the seriousness of the offset.

follow through with the sequence up to the point of the consummatory action, there should be a tendency to continue with a sequence once started.

5. On the other hand, the mechanism should be flexible enough to interrupt an ongoing sequence if an urgent contingency arises (e.g. if the animal needs to engage in escape behaviour), or if a valuable opportunity becomes available.
6. If an action is common to two or more systems, and the motivations for the effects of those systems are high at the same time, then the action should acquire a higher level of activation (probability of being chosen) than it would have if only one of the motivations was high. That is to say, an if an action can contribute to two or more active goals simultaneously then it should be more likely to be chosen.

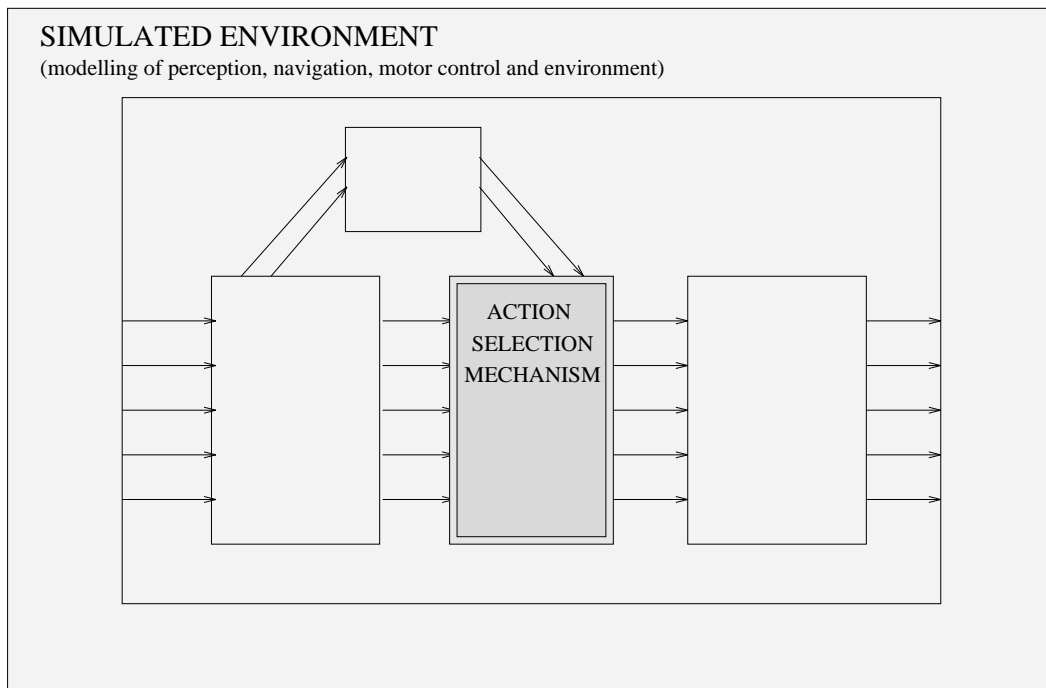


Figure 3: The relationship between the external environment, perception, navigation, action selection and motor control. The simulated environment models all the processes (and the interactions between them) except action selection, which is provided by an action selection mechanism.

3. A Simulated Environment for Testing Action Selection.

This section contains a brief description of a *simulated environment* (SE), a computer simulation¹ that models the AS problem faced by an animal in the wild and that can be used to examine the efficacy of the selections of actions of an ASM. This is followed by a discussion of how the SE should be used, and how results derived with it should be interpreted.

3.1. Description of the Simulated Environment.

The SE models the external environment of the animal (the outside world), internal variables (such as body temperature and food level), and the processes of perception, navigation and motor control (see figure 3). The dynamics of the interactions between all of these, and between them and the action selection mechanism (ASM), are all programmed into the SE. Only the process of selecting actions is not modelled in the SE. Instead, different ASMs, embodying different theories of AS, are ‘plugged into’ the SE and allowed to make the decisions as to which actions are chosen at each moment in time.

¹The software for the computer simulation, together with an implementation of MASM, can be obtained by request from the author

The AS problem provided 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 (larger animals which would ignore the modelled animal but may trample it inadvertently), sleeping at night in a den, not getting lost (not straying too far from familiar areas) and reproducing. This particular list of sub-problems was taken from the list of common types of behaviour in the previous section, but with all social types of behaviour removed, and with some other sub-problems added. Social sub-problems, apart from reproduction, were avoided (the animal was assumed to be solitary) because otherwise the SE (designed to test ASMs) would require built-in assumptions about the AS of conspecifics, which is rather close to including assumptions about the way that actions are selected in the ASM being tested. It was thought sensible to add a few extra sub-problems not mentioned in the previous section (such as the “not getting lost” sub-problem). It would seem that these sub-problems must be necessary for an animal in a natural habitat, but they are generally overlooked by ethologists, perhaps because they do not have associated unique and identifiable actions or fixed action patterns; they only cause the animal to move in a certain direction. One of the sub-problems, avoiding the edge of the SE, is artificial but is necessary unless a much larger environment is to be modelled.

In order to include different sub-problems in the AS problem posed by the SE, different features, internal variables and actions were added as necessary. For example, in order to include the sub-problem “obtaining water”, the following were 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 internal 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 4. A key to the different features is given in figure 5. 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. There are *costs* associated with each action, and so each selection of an action will influence several 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 6). 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. The animal can see nothing at night. Navigation is also error-prone, with the animal able to get lost (unable to find its way back to its den) 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 the process of motor

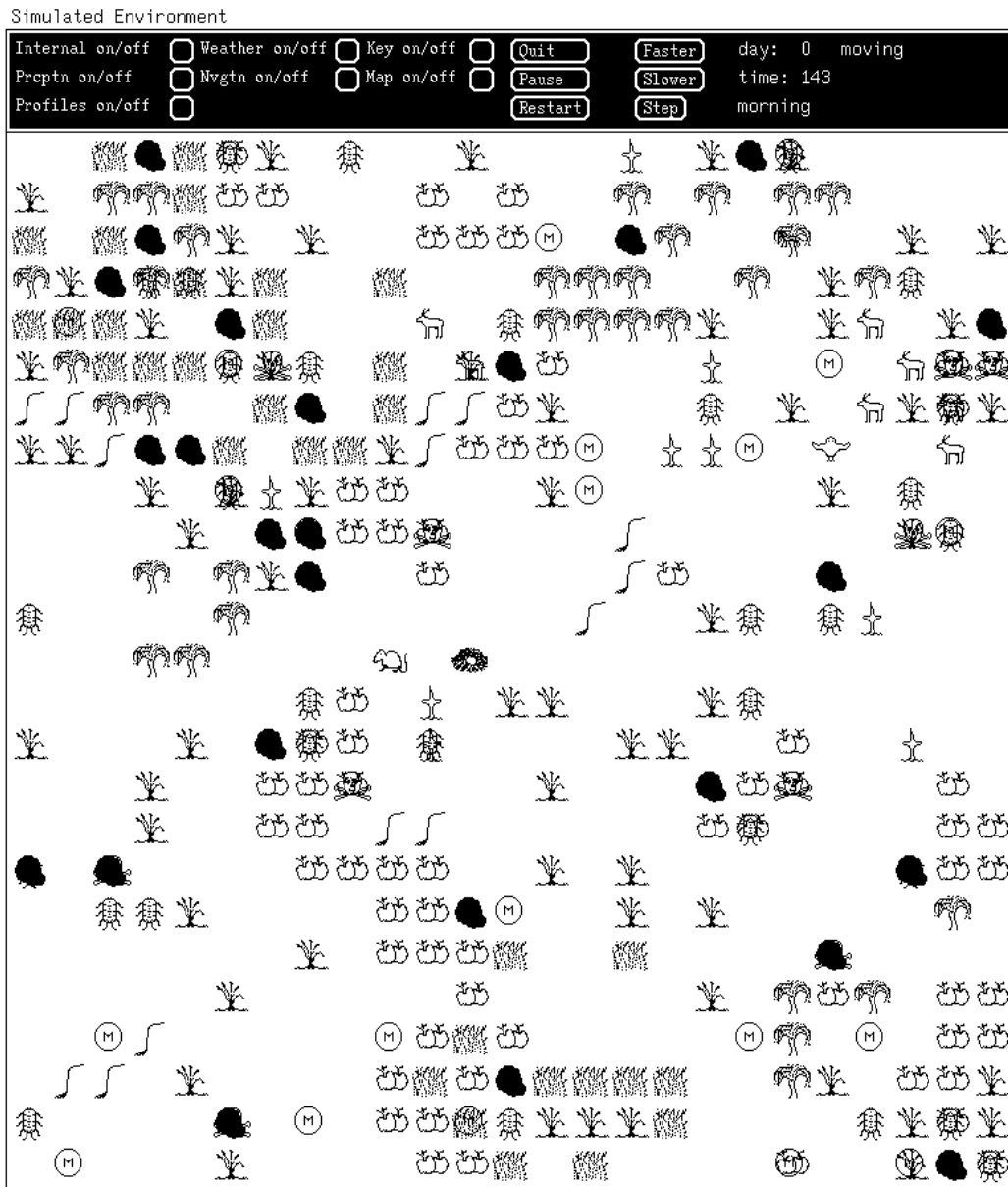


Figure 4: A randomly generated instance of the simulated environment showing the positions of features, the position of the animal, the current action of the animal, the time, the day and the current part of the day.

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 5: Key to features in the simulated environment. A toxic food or water source is one which will cause a temporary decrement in the animal's health when it consumes from that resource.

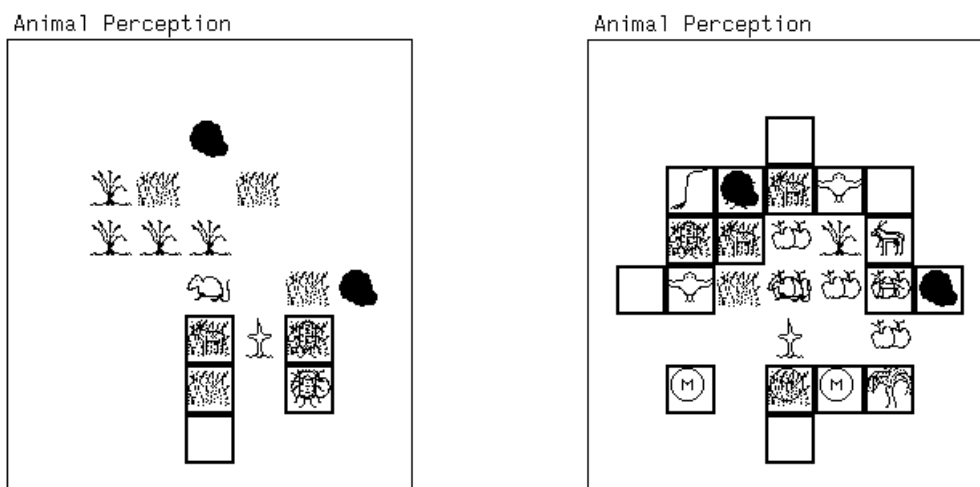


Figure 6: Two instances of the animal's perception of its local environment. Boxes around squares indicate that the contents of the square have been incorrectly perceived. The perception on the left occurred when the animal was not in vegetation and chose an action to look around. The instance on the 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).

control includes the possibility of the incorrect execution of an action; when the animal attempts an action, but is unsuccessful at it.

Performance in the SE is defined as equal 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.

The description of the SE that has just been given here is brief, but space does not allow for more to be included. More explanation, justification and validation of the SE is given in [Tyrrell, 93a]. In particular, a description is given there of many tests that were carried out on the SE to check the veracity of low-level details, the suitability of individual sub-problems and the ability of the SE as a whole to be able to discriminate between better and worse ASMs.

The 13 different sub-problems included in the SE lead to 14 different feature types, 6 different internal variables, 6 other non-internal variables (such as the proximity of nightfall), and 35 different actions (16 of which are movement actions). The SE that was used to test MASM in [Maes, 91a] included seven sub-problems which involved 4 types of feature, two internal variables and 4 actions plus movement actions.

3.2. Use of the Simulated Environment.

The SE can be used in one of two ways: either to compare the performances of

different ASMs, or to observe the selections of actions of particular ASMs and to look for gross deficits in the choices made by them (such as never choosing consummatory actions, or never attending to a particular sub-problem). This paper is concerned with one ASM only and so it is the latter method of using the SE that is of interest here.

The SE can be used to look for gross inadequacies in AS in the following manner:

1. an ASM is implemented so as to be able to select actions in the SE in the manner specified by the particular theory of AS.
2. the performance of the ASM in the SE is observed directly and any major deficits are noted.
3. the underlying reasons behind any major deficits are investigated and analysed to an arbitrary level of detail using various tools, including longitudinal profiles of node activations through time and *dbx* (a unix tool to help analyse and debug programs).
4. the deficits are corrected by making adjustments to the ASM, where this is possible without changing the whole nature of the ASM.
5. the last three steps are repeated until either (*i*) there are no major deficits in performance, or (*ii*) the remaining deficits cannot be removed without changing fundamental features of the ASM (the remaining deficits are inevitable given the major design decisions that have been made for the ASM).

3.3. Interpreting Results from the Simulated Environment.

Three potential concerns with results obtained in this way are: (*i*) the SE is only a *model* of the AS problem typically faced by animals in the wild, and its veracity cannot be guaranteed, (*ii*) the deficits observed in the SE might not be applicable in general, but may be specific only to this test of AS (for instance they may be due to peculiar interactions between the programming of the SE and the programming of the ASM), and (*iii*) the deficits might not have occurred if the ASMs had been implemented more competently (the deficits could be a property of the particular implementation of the ASM rather than of the ASM itself).

To guard against these concerns, care was taken during the testing to ensure that a deficit was only reported as such after: (*i*) the *reasons* behind the deficit had been understood, rather than just accepting the deficit at face value, (*ii*) the reasons could be shown to be generally applicable to animal-like problems of AS, and (*iii*) the deficit could be shown to arise as an inevitable consequence of essential features of the ASM. A deficit was not reported if it was possible to remove it by making fairly minor changes to the ASM.

These safeguards, taken together, ensure that the reported deficits are applicable to AS in general, and are a consequence of fundamental aspects of the ASM in question.

An analogy can be made between the use of the SE to test ASMs and the use in the past of canaries down coal mines because of their particular susceptibility to toxic gases. When a canary keeled over then this was used as a signal that poisonous gases were starting to build up to a dangerous level and the miners were evacuated from the

pit. However, a dead or unconscious canary was not an infallible indication of a build-up of gas. For instance, the canary's diet or age could have been the cause. Similarly, when a deficit in performance is observed for an ASM in the SE then this cannot be used as an infallible indication of an inadequacy in the ASM. Only when analysis is carried out, and the underlying causes of the observed deficit are determined, can any firm conclusions be drawn.

4. Maes’ “Bottom-Up Mechanism for Behavior Selection”.

For the sake of completeness, a description is given here of MASM. Maes’ own, more extensive descriptions of her mechanism are given in [Maes, 89; 90 and 91a]. This mechanism (MASM) combines ideas from artificial intelligence (e.g. distributed processing), from artificial life (e.g. emergent properties) and from ethology (e.g. motivations, consummatory and appetitive alternatives). MASM is fairly complex, but is specified very clearly and precisely; more precisely than any other mechanism or theory of AS. In contrast, the ‘thought experiments’ of Tinbergen and Lorenz [Tinbergen, 50 and 51; Lorenz, 50 and 81], for instance, are much vaguer descriptions, and are much less amenable to critical analysis.

MASM is a distributed, non-hierarchical network. There is a set of nodes which represents consummatory and appetitive alternatives. The nodes are generally at a level ‘lower’ than the system level and ‘higher’ than the action level (where system and action level are as shown in figure 1). The nodes are fairly simple, but are more complex and less realistic of individual biological neurons than the simple models of neurons usually used in artificial neural networks. There are two waves of input to the network — firstly from the sensors of the external environment (mostly *external stimuli*), and secondly from the motivations / goals (mostly derived from *internal stimuli* such as body temperature), as shown in figure 7. The sensors of the environment are binary-valued. That is to say, they measure the truth or falsity of logical statements about the environment. In the initial accounts of the mechanism [Maes, 89 and 90] the goals or motivations (the two terms will be used interchangeably throughout) were also either active or inactive (0 or 1), but in [Maes, 91a] the goals or motivations are allowed to take continuous values, and that amendment was included in this implementation of the mechanism (to be described in the next section). The goals only excite consummatory nodes directly. They also inhibit nodes which threaten to undo or counteract them. As well as the connections from goals and sensors of the environment, there are internal connections between the nodes which are of three types: predecessor, successor and conflictor links. After activation is passed in from the perception of the environment and from the motivations, then it is also spread around inside the network along the internal connections. The input and spreading of activation can be repeated many times before a decision is made.

The central idea of Maes’ scheme for AS is that the different types of links encode various relationships (e.g. consummatory/appetitive relationships between nodes, conflict relationships between nodes, goal-achieving relationships between nodes and goals, goal-counteracting relationships between nodes and goals and situation-dependency relationships between nodes and environmental sensors). It is then assumed that with this information hard-wired into the mechanism, excitation can be fed in from the current situation and the current goals and activity will then, after several iterations, come to settle in the node representing the most appropriate behaviour.

The composition of a node is now described more fully. Each node (see figure 8) has a set of preconditions. These preconditions are logical conditions about the environment which are required to be true in order for the node to be executable; i.e. in order for it to be a valid, selectable alternative. For example, the precondition “water in square” must be true in order for the node “drink” to be executable. Requiring

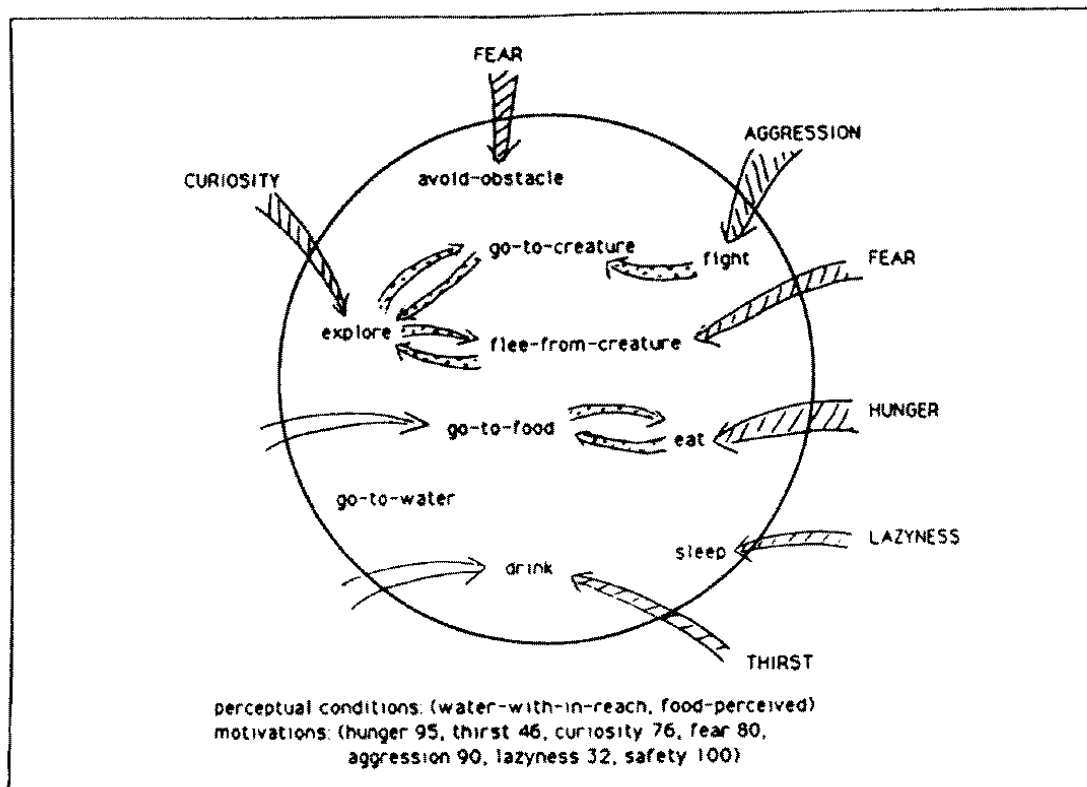


Figure 7: Maes' non-hierarchical, "Bottom-up Mechanism for Behaviour Selection". Excitation flows into the network from perceptual conditions (if true) and motivations. Excitation is also passed between nodes in the network along internal connections. Taken from [Maes, 91] (© acknowledged to MIT Press).

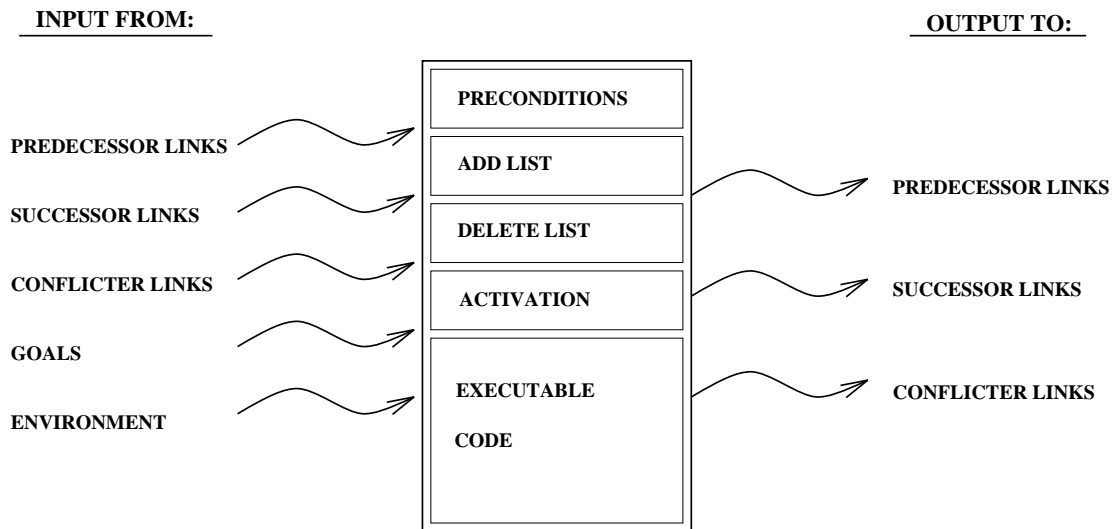


Figure 8: The make-up of a node in Maes' network.

all preconditions of a node to be true is a *conjunction* of all of the preconditions. Disjunctions have to be implemented through the use of disjunctive preconditions [Maes, 90: page 9]. As well as a list of preconditions, each node also has an add list and a delete list. The add list consists of conditions about the environment that the node is likely to make true (e.g. “water in square” for the node “approach water”). The delete list consists of conditions that are likely to be made false by the execution of the node (e.g. “no rivals nearby” for the node “approach rival”). The final two components of the node are the activation level and the code which gets run if the node is executed (this code will translate from a consummatory or appetitive alternative, such as “approach water”, to a specific action, such as “move south-east”).

The internal links are specified as follows:

1. **predecessor links** – if (i) proposition X is false, (ii) proposition X is a precondition of node A , and (iii) proposition X is in the add list of node B (i.e. if B can help A to become executable), then there is an *active* predecessor link (excitatory) from A to B .
2. **successor links** – if (i) proposition X is false, (ii) proposition X is in the add list of node A , (iii) proposition X is a precondition of node B , and (iv) the node A is executable (i.e. if A can help B to become executable), then there is an *active* successor link (excitatory) from A to B .
3. **conflictor links** – if (i) proposition X is true, (ii) proposition X is a precondition of node A , and (iii) proposition X is in the delete list of node B (i.e. if B stops A from becoming executable), then there is an *active* conflictor link (inhibitory) from A to B .

The external links providing input to the network are specified as follows:

1. **from sensors of the environment** – if (i) proposition X about the environ-

ment is true, and *(ii)* proposition X is in the precondition list of node A (i.e. if A is at least partially appropriate to the current situation), then there is an active link (excitatory) from the sensor of the proposition X to node A .

2. **from goals** – if *(i)* goal Y has an activation greater than zero, and *(ii)* goal Y is in the add list of node A (i.e. if A is likely to achieve goal Y), then there is an active link (excitatory) from the goal Y to node A .
3. **from protected goals** – if *(i)* goal Y has an activation greater than zero, and *(ii)* goal Y is in the delete list of node A (i.e. if A is likely to undo goal Y , or stop it from being achieved), then there is an active link (inhibitory) from the goal Y to node A .

The procedure used to select a node to execute at each timestep, as shown in figure 9, is:

1. calculate the excitation coming in from the environment and the motivations.
2. spread excitation along the predecessor, successor and confictor links.
3. normalise the node activations so that the average activation becomes equal to the constant π (see below).
4. check to see if any nodes are executable, and if so then choose the one with the highest activation, execute it and finish.
5. otherwise, if no node is executable, reduce the global threshold and repeat the cycle.

A node is executable if all of its preconditions are true and if its activation is greater than the global threshold. If more than one node is executable after a cycle then the one with the highest activation is chosen. When a node has been chosen and executed then its activation is reset to zero before the next timestep. The activation of the sender node before excitation comes in from outside in step 1 is used when calculating the excitation sent along the internal links in step 2.

Several global parameters can be used to ‘tune’ the performance of the ASM to a particular environment:

1. the mean activation value after each cycle (used in normalisation) $[\pi]$.
2. the initial value of the global threshold $[\theta]$, which is reduced by an amount (e.g. 10%) after each cycle if no executable node has an activation greater than it.
3. a constant determining the weighting of environmental sensor inputs (as opposed to other sorts of inputs), as well as the weighting of successor links (as opposed to other sorts of links) $[\phi]$.
4. a constant determining the weighting of goal inputs and predecessor links $[\gamma]$.
5. a constant determining the weighting of protected goal inputs and confictor links $[\delta]$.

The different inputs to a node are multiplied by the following: *(i)* environmental sensors by ϕ , *(ii)* goals by γ , *(iii)* protected goals by δ , *(iv)* successor links by $(\frac{\phi}{\gamma})$, *(v)* predecessor links by $(\frac{\gamma}{\gamma} = 1)$, and *(vi)* confictor links by $(\frac{\delta}{\gamma})$.

There are also division rules affecting how much activation gets added or subtracted from nodes as a result of the different sets of links in the network. In general, each

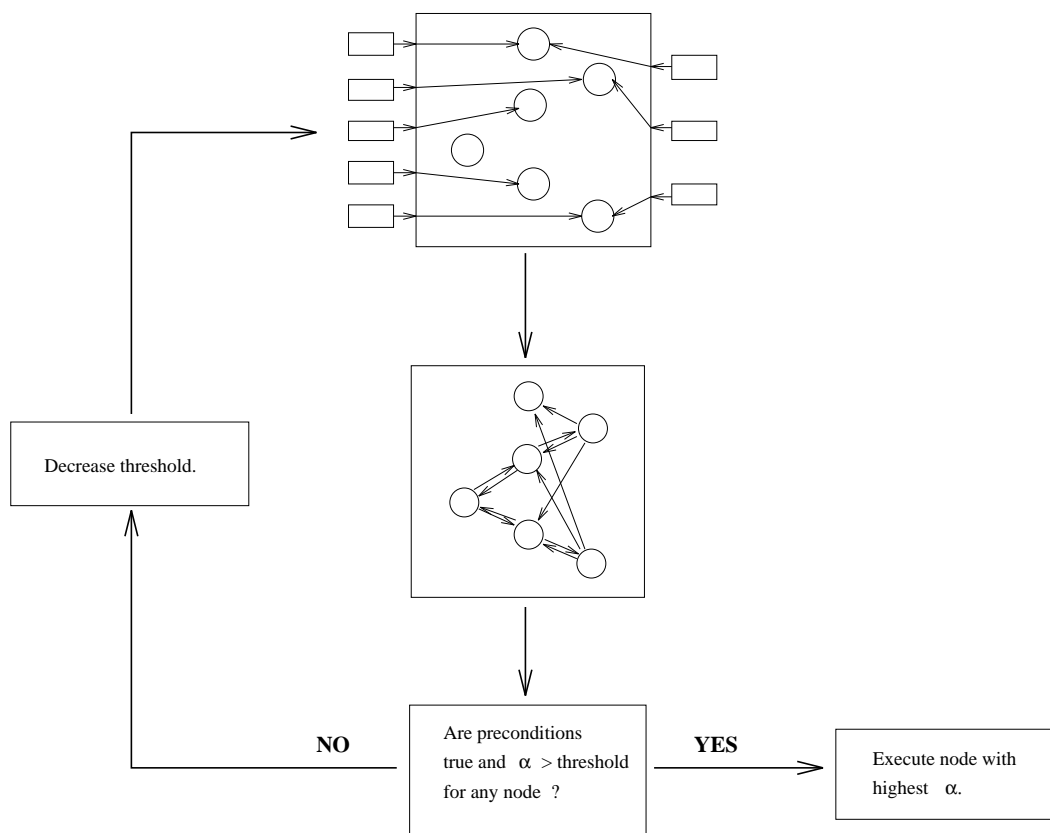


Figure 9: The procedure carried out at each timestep in Maes' mechanism.

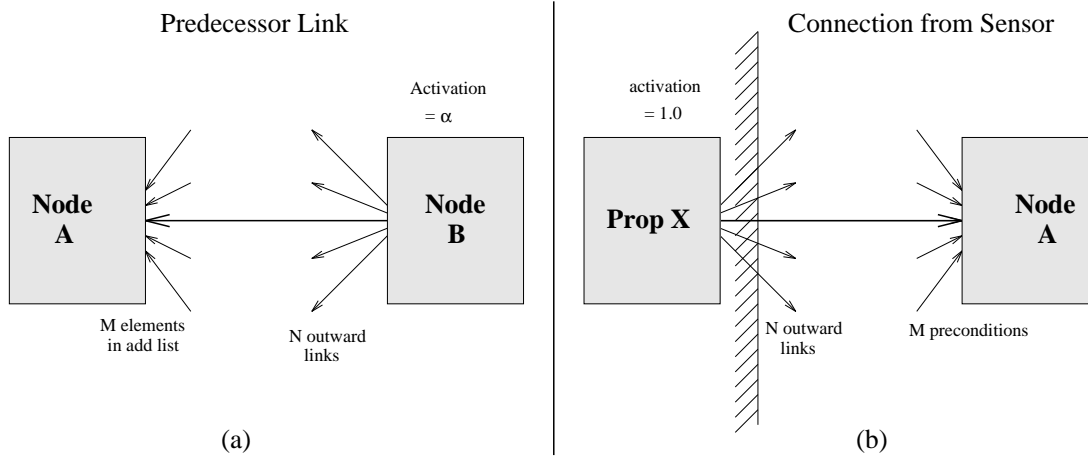


Figure 10: The division rules applied to excitation transferred along (a) predecessor links, and (b) inputs from the environment. In (a) the activation of A is incremented by $(\frac{\gamma}{\gamma} \frac{\alpha}{NM}) = (\frac{\alpha}{NM})$; in (b) the activation of A is incremented by $(\phi \frac{\alpha}{NM})$. The symbols ϕ , γ , α , N and M are explained in the text.

amount of activation that gets passed to a node is divided by the number of outputs of that type (N) from the source, and by the number of inputs of that type (M) to the receiving node. Consider the situation in figure 10a, in which there is a predecessor link from node B to node A , due to a proposition X being in both the precondition list of B and the add list of A . Node A has M elements in its add list (one of which is proposition X) and there are N different nodes that can make X true (i.e. that receive predecessor links from B because of proposition X). In this case the predecessor link increments the activation of A by an amount $(\frac{\gamma}{\gamma} \frac{\alpha}{NM} = \frac{\alpha}{NM})$, where α is the activation of node B .

Similar division rules are prescribed for the successor and conflictor links, and also for the inputs to the network from outside it. For instance, when activation is fed in from a sensor of the environment due to a proposition being true (see figure 10b), then node A 's activation is in this case incremented by $(\phi \frac{\alpha}{NM} = \frac{\phi}{NM})$, where $\alpha = 1$ (because the proposition is true), N is the number of nodes that have proposition X in their precondition list, and M is the number of preconditions of node A .

It should be noted that MASM has nodes at only one level, which is below the system level but in some cases above that of the action level. The only distinction between nodes is that between appetitive and consummatory. Consummatory nodes receive direct motivational input, whereas the appetitive nodes only receive that input indirectly through the consummatory nodes or through other appetitive nodes closer in the sequence to the consummatory node. All nodes can receive direct input from sensors of the state of the environment. The effects of internal and external stimuli (mediated through internal connections for some nodes) are always summed.

5. Implementation of Maes' Mechanism.

A brief description of the SE was given in section 3, followed by a description of Maes' ideas about how behaviour can be generated in section 4. This section now describes how Maes' very explicit specification of her scheme was used to construct an implementation (a particular ASM) that was tailored to the particular AS problem presented by this SE.

The precondition lists, add lists and delete lists for each node of the implementation are shown in tables 1, 2 and 3, and the resulting network is shown in figure 11. A full description of each of the sub-problems in the SE could not be given here, and so some of the details of the implementation of Maes' ideas may seem slightly confusing. However, all of the deficits in MASM will be shown to be general, rather than specific to this particular problem of AS, and so explaining this implementation in very fine detail is not necessary. A few of the more pertinent details are given though to help explain the implementation: *(i)* some actions (pounce, court, mate and clean) are not allowed when the animal is in its den, *(ii)* as the animal's health approaches zero (death), then the probability of small perturbations in health taking it below the threshold increases, and so it becomes more important for the animal to engage in actions which will result in an immediate increase in health (eat actions, drink or clean), and *(iii)* there are three different sources of food (cereal type food, fruit type food, and prey) and so there are three different consummatory actions for eating.

Some variations of the network in figure 11 were also tested, with similar results. For each test in the SE of each implementation of MASM, there was also experimentation with different values of the 5 global parameters described in the previous section. However, because of the large number of possible network configurations and the large number of different combinations of parameter values it was not at all possible to carry out an exhaustive or rigorous search of the space of different mechanisms. It is argued that this possible limitation of the testing is not critical though because the faults that are found in MASM are examined analytically and shown to be independent of details of the testing procedure. If analysis shows that the faults really exist then it does not matter how they were discovered.

The implementation and dynamics of a single example system ("obtaining water") will now be described. The system is shown in figure 12 (connections from other systems are not included). If there is water in the animal's square then the node "drink" will be executable and will receive activation from both the sensor and from the goals (assuming that at least one of the goals have activation greater than zero). Because "drink" is executable, no activation will be passed down the line to "approach perceived water" and so that node cannot receive more activation than "drink". "Drink" should get more activation than the other nodes in the system if it is executable, because it will be the only node to receive any activation from the goals. Similarly, if "water in square" is false, but "water perceived" (at a distance) is true then, since "drink" is not executable, the predecessor connection becomes active and the goal activation gets passed along to "approach perceived water", but not further. The general idea is that the first node in the sequence which is executable will receive the most activation (out of the executable nodes) and will win the competition, in its system at least.

Node	Precondition List
Eat Cereal Food	(Cereal Food in Square)
Eat Fruit Food	(Fruit Food in Square)
Pounce (Eat Prey)	(Prey in Square), (Den Not in Square)
Approach P. Food	(Perceived Food)
Approach R. Food	(Remembered Food)
Drink Water	(Water in Square)
Approach P. Water	(Perceived Water)
Approach R. Water	(Remembered Water)
Rest	(Shade or Shelter in Square)
Approach P. Shelter	(Perceived Shelter)
Approach P. Shade	(Perceived Shade)
Look towards P1	(P1 Perceived Weakly)
Freeze	(P1 or P2 Perceived), (Shelter in Square)
Run away from P1	(P1 Perceived)
Look towards P2	(P2 Perceived Weakly)
Run away from P2	(P2 Perceived)
Avoid Dangerous Place	(Dangerous Place Perceived)
Avoid Irrelevant Animal	(Irrelevant Animal Perceived)
Approach P. Mate	(Perceived Receptive Mate)
Court	(Receptive Mate in Square), (Den Not in Square)
Mate (Copulate)	(Courted Mate in Square), (Den Not in Square)
Sleep	(Den in Square), (Sunset or Night)
Approach P. Den	(Perceived Den)
Approach R. Den	(Remembered Den)
Clean	(Den Not in Square)
Avoid Edge	(Edge Perceived)
Explore	<i>None</i>
Look Around	<i>None</i>
Approach P. Cover	(Perceived Shelter)

Table 1: Preconditions for each node in Maes' ASM.

Node	Add List
Eat Cereal Food	(Food OK), (Health Not Too Low)
Eat Fruit Food	(Food OK), (Health Not Too Low)
Pounce (Eat Prey)	(Food OK), (Health Not Too Low)
Approach P. Food	(Cereal Food in Square), (Fruit Food in Square), (Prey in Square)
Approach R. Food	(Perceived Food)
Drink Water	(Water OK), (Health Not Too Low)
Approach P. Water	(Water in Square)
Approach R. Water	(Perceived Water)
Rest	(Not Too Hot), (Not Too Cold)
Approach P. Shelter	(Shelter in Square), (Shade or Shelter in Square)
Approach P. Shade	(Shade or Shelter in Square)
Look towards P1	(Recently Scanned), (No Predator Since Last Scan)
Freeze	(No Predator1s Perceived), (No Predator2s Perceived)
Run away from P1	(No Predator1s Perceived)
Look towards P2	(Recently Scanned), (No Predator Since Last Scan)
Run away from P2	(No Predator2s Perceived)
Avoid Dangerous Place	(No Dangerous Places Perceived)
Avoid Irrelevant Animal	(No Irrelevant Animals Perceived)
Approach P. Mate	(Receptive Mate in Square)
Court	(Courted Mate in Square)
Mate (Copulate)	(Just Mated)
Sleep	(Sleeping in Den)
Approach P. Den	(Den in Square), (Variance OK)
Approach R. Den	(Perceived Shelter), (Perceived Shade), (Den in Square), (Variance OK)
Clean	(Clean), (Health Not Too Low)
Avoid Edge	(No Edges Perceived)
Explore	(Perceived Food), (Perceived Water), (Remembered Food), (Remembered Water), (Perceived Shelter), (Perceived Shade), (Perceived Receptive Mate), (Den Not in Square)
Look Around	(Recently Scanned), (No Predator Since Last Scan)
Approach P. Cover	(Close to Cover)

Table 2: Add list elements for each node in Maes' ASM.

Node	Delete List
Eat Cereal Food	<i>None</i>
Eat Fruit Food	<i>None</i>
Pounce (Eat Prey)	<i>None</i>
Approach P. Food	<i>None</i>
Approach R. Food	(Den in Square)
Drink Water	<i>None</i>
Approach P. Water	<i>None</i>
Approach R. Water	(Den in Square)
Rest	<i>None</i>
Approach P. Shelter	<i>None</i>
Approach P. Shade	<i>None</i>
Look towards P1	<i>None</i>
Freeze	<i>None</i>
Run away from P1	<i>None</i>
Look towards P2	<i>None</i>
Run away from P2	<i>None</i>
Avoid Dangerous Place	<i>None</i>
Avoid Irrelevant Animal	<i>None</i>
Approach P. Mate	<i>None</i>
Court	<i>None</i>
Mate (Copulate)	<i>None</i>
Sleep	<i>None</i>
Approach P. Den	(Den Not in Square)
Approach R. Den	(Den Not in Square)
Clean	<i>None</i>
Avoid Edge	<i>None</i>
Explore	(Den in Square)
Look Around	<i>None</i>
Approach P. Cover	<i>None</i>

Table 3: Delete list elements for each node in Maes' ASM.

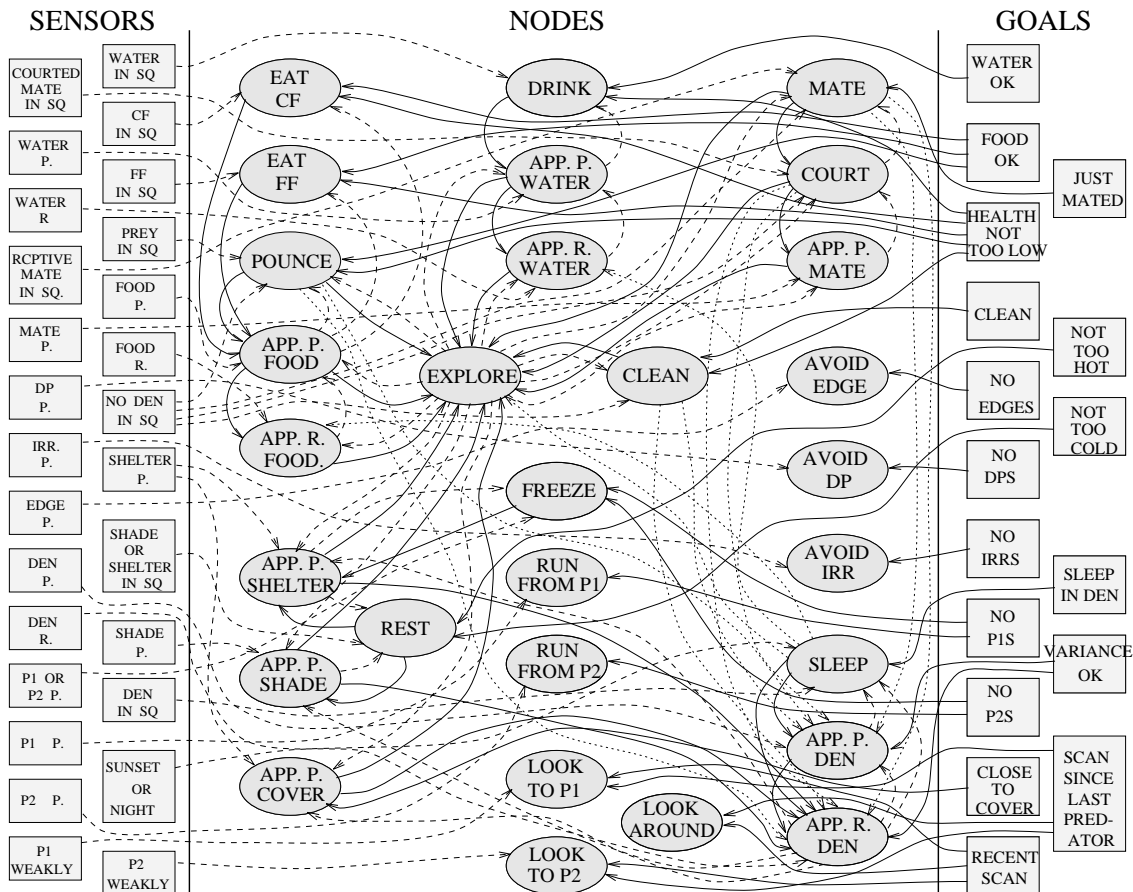


Figure 11: The implementation of Maes' ASM to solve the action selection problem posed by the SE. Solid lines denote goal or predecessor connections, dashed lines denote sensor or successor connections, and dotted lines denote protected goal or conflictor connections. 'P' stands for 'perceived', 'R' for 'remembered', 'APP' for 'approach', 'SQ' for 'square', 'P1' for 'predator type 1', 'P2' for 'predator type 2', 'CF' for 'cereal type food' and 'FF' for 'fruit type food'.

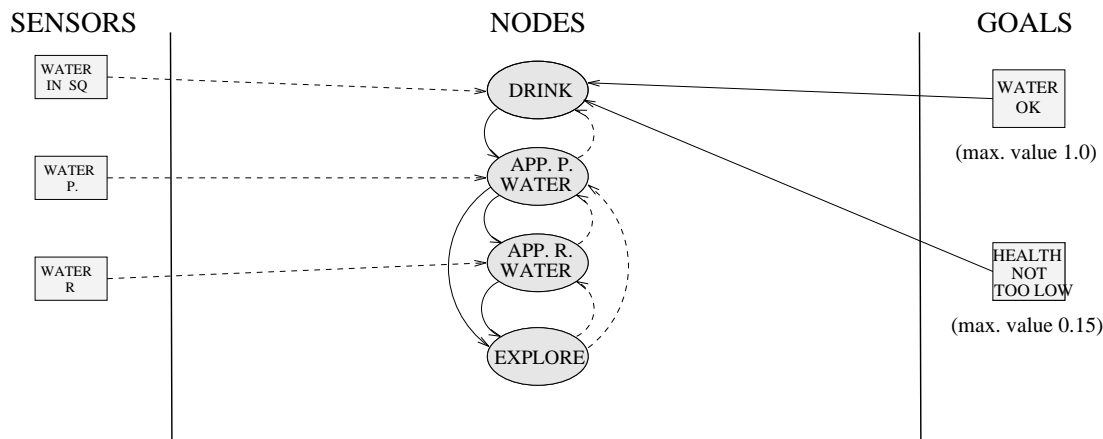


Figure 12: One system in Maes’ ASM — the one which is responsible for generating behaviour to rectify a shortage of water (i.e. which is responsible for dealing with the “get water” sub-problem). Solid lines denote goal or predecessor connections and dashed lines denote sensor or successor connections. Connections to and from other systems are not shown. ‘P’ stands for ‘perceived’, ‘R’ for ‘remembered’, ‘APP’ for ‘approach’ and ‘SQ’ for ‘square’.

6. Results of Testing Maes’ Mechanism.

The testing of MASM in the SE was carried out initially by examining the performance of an implementation that was in complete accordance with the specification in [Maes, 91a]. When deficits in performance were discovered for the first implementation then changes were made to it in order to try to remove the observed deficits, and this generated the second implementation. Similarly, subsequent versions of the ASM were produced in order to try and rectify new or remaining deficits in the performance of the previous versions.

6.1. First Implementation.

The first implementation of MASM was exactly as described in section 5, and the division rules and all other aspects were as specified in [Maes, 91a]. Binary sensors of the environment such as “cereal food in square” or “water perceived” were set to ON when the perceived values (amount of cereal type food or water) of the entities exceeded a certain threshold; otherwise they were set to OFF. Goals were real-valued.

There was an obvious problem (see figure 13) with the performance of this initial implementation, which was discovered by detailed examination of the activations passing between nodes in the ASM as it performed in the SE. Because of the division by the number of outputs (of a similar type) from the sender node (N in figure 10), there was a prejudice against nodes which received input from sensors which also excited other nodes. Nodes such as “approach perceived cover” and “clean” received very little activation and were never chosen because their sensors inputs were shared with other nodes and were therefore diluted in strength. Nodes such as “drink” and “avoid dangerous place” received a lot of activation and were chosen frequently

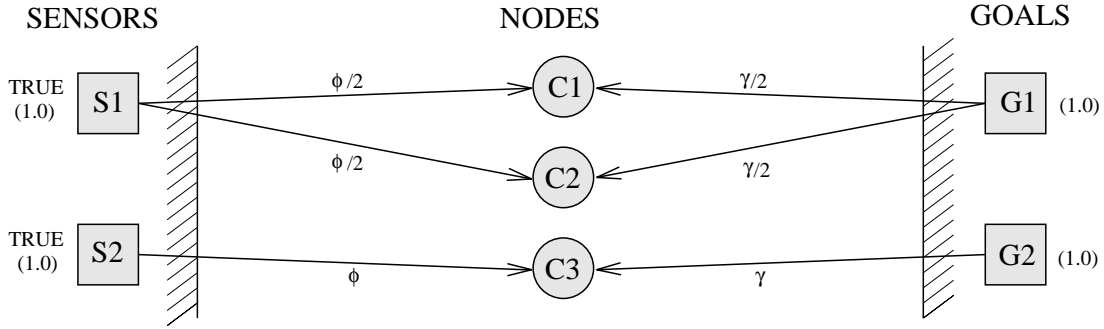


Figure 13: Unbalanced competition between consummatory nodes because of division by the number of outputs. $C1$ and $C2$ receive inputs of $(0.5\phi + 0.5\gamma)$ whereas $C3$ receives input of $(\phi + \gamma)$ in every cycle, though there is no reason to prefer $C3$. Goal and sensor activations are shown in parentheses beside them, and the weights on the connections are shown next to the connections. ϕ and γ are global parameters, whose meaning is explained in the text.

because their sensor inputs were undiluted and therefore stronger.

This principle of division by the number of outputs (or to be more exact, by the number of other nodes also receiving input from the sender node because of the proposition involved) was abandoned. Similarly, excitation from a goal or inhibition from a protected goal was not divided by the number of recipients. This was because (i) if a goal can be achieved by several different nodes then each node should receive a full complement of excitation if it is to be able to compete effectively against other nodes (the alternative is to prejudice against nodes for which the goal that they achieve can also be achieved by one or more different nodes), and (ii) if a sensor is relevant to many nodes then that should not make a difference to the amount of excitation each node receives from it (the alternative is that nodes which receive inputs from ‘widely-used’ sensors are penalised against).

Maes states that the division by the number of outputs is necessary because “we want modules that achieve the same goal or modules that use the same precondition to compete with one another to become active (we view them as representing a disjunction or choice point)” [Maes, 90: page 9]. This argument is not valid. The modules will still compete with each other whether or not the excitation they receive is divided by the number of outputs. If the excitation from a sensor or goal is divided up amongst the recipient nodes then these nodes will not be able to compete on equal terms against other nodes which are the only ones capable of fulfilling the goal they achieve, or on equal terms against nodes which do not share their preconditions with other nodes.

6.2. Second Implementation.

To create the second version of MASM the division by the number of outputs on the sender side was removed. And so, for example, the new connection weights for the two cases in figure 10 were changed to $\frac{\alpha}{M}$ and $(\phi \times \frac{\alpha}{M})$ respectively. With these new division

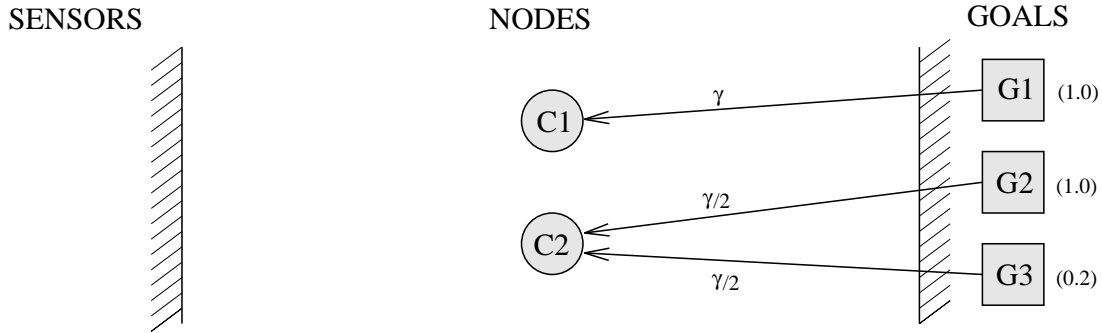


Figure 14: Unbalanced competition caused by division by the number of inputs. $C1$ receives input of $(1.0 \times \gamma)$ in every cycle, $C2$ receives input of $(0.6 \times \gamma)$ in every cycle, even though it can achieve goal $G3$ as well as goal $G2$. γ is a global parameter, whose meaning is explained in the text.

rules in place MASM was tested again but a second problem became apparent, again through observation of the detailed dynamics of MASM while it was selecting actions in the SE. This time the problem lay with the division by the number of inputs, or to be more exact, the division by the size of the receiving node's precondition list for environmental sensor or successor inputs, the size of the receiving node's add list for goal or predecessor inputs, or the size of the receiving node's delete list for protected goal or conflictor inputs. Because of this division there was a prejudice against nodes with many inputs of a particular type, since all of the inputs to those nodes had to be large in value in order for the nodes to accumulate as much activation as a node with only one input, which had a large value (see figure 14). For instance, the nodes "look around" and "drink" (which both receive two goal inputs, one of which can be fairly large, the other of which is always fairly small) are penalised because the total input they receive from goals is equal to the average rather than the sum of the inputs. This is not obviously a problem for inputs from sensors of the environment or successor links (since they are binary), but it is a problem for inputs from goals and inputs along predecessor links.

6.3. Third Implementation.

In order to remove the problem caused by division by the number of goal-type inputs, and in order to allow nodes which can help achieve more than one goal to receive more rather than less activation, a change was made for the third implementation. The division by the size of the receiving node's add list (for goals and predecessors), and by the size of the delete list (for protected goals and conflictors) was removed. This change was not made to inputs from the environment or along successor links though. And so, to give two examples, the excitation fed in along each link from a goal was now changed to (γG) , whereas the excitation added to a node due to a successor link was left unchanged at $(\frac{\alpha \phi}{M \gamma})$, where G is the strength of the goal, α is the activation of the sending node, M is the number of preconditions for the node, and ϕ and γ are global parameters (see section 4).

Division Rule	Deficit in Action Selection
Division by N	penalises nodes ‘sharing’ inputs
No Division by N	————
Division by M (successor and environment links)	————
Division by M (goal and protected goal links)	penalises nodes with high M
No Division by M (goal and protected goal links)	————
Division by M (predecessor and conflictor links)	penalises nodes with high M , where inputs are due to different goals
No Division by M (predecessor and conflictor links)	favours nodes with high M , where inputs are due to the same goal

Table 4: The deficits caused by different variations of the division rule (see section 6.3). M and N are defined in section 4. ‘Penalises’ means that a node receives too little activation, so that it may lose out in the competition against other nodes on occasions when it should win. ‘Favours’ means that a node receives too much activation.

This third implementation was still not completely successful because nodes such as “approach perceived food”, which received many predecessor inputs, always had an unjustified advantage in the competition against other nodes. This problem was due to the change in the division rules made after the last implementation. Since there is no division by the number of predecessor inputs, nodes with many of these inputs are no longer penalised against but instead, in some cases at least, come to dominate the competition (i.e. they nearly always get selected, even when they are not sensible choices).

This problem, that both division by the number of predecessor inputs as well as the lack of such division produces unsatisfactory results (see table 4), is a complex one and is due to the way in which motivational inputs are propagated through the network in MASM. Instead of a goal or motivational variable sending excitation to all of the appetitive and consummatory alternatives in the system (figure 15b), as is the case with Beer’s artificial cockroach [Beer, 90] (see figure 16) and other mechanisms [e.g. Baerends, 76; Halperin, 90; Tyrrell, 93a], in MASM the goal excites only the consummatory node, and then that consummatory node passes back activation to the appetitive node if it isn’t executable (figure 15a).

The dilemma that occurs in Maes’ scheme for propagating motivational inputs arises

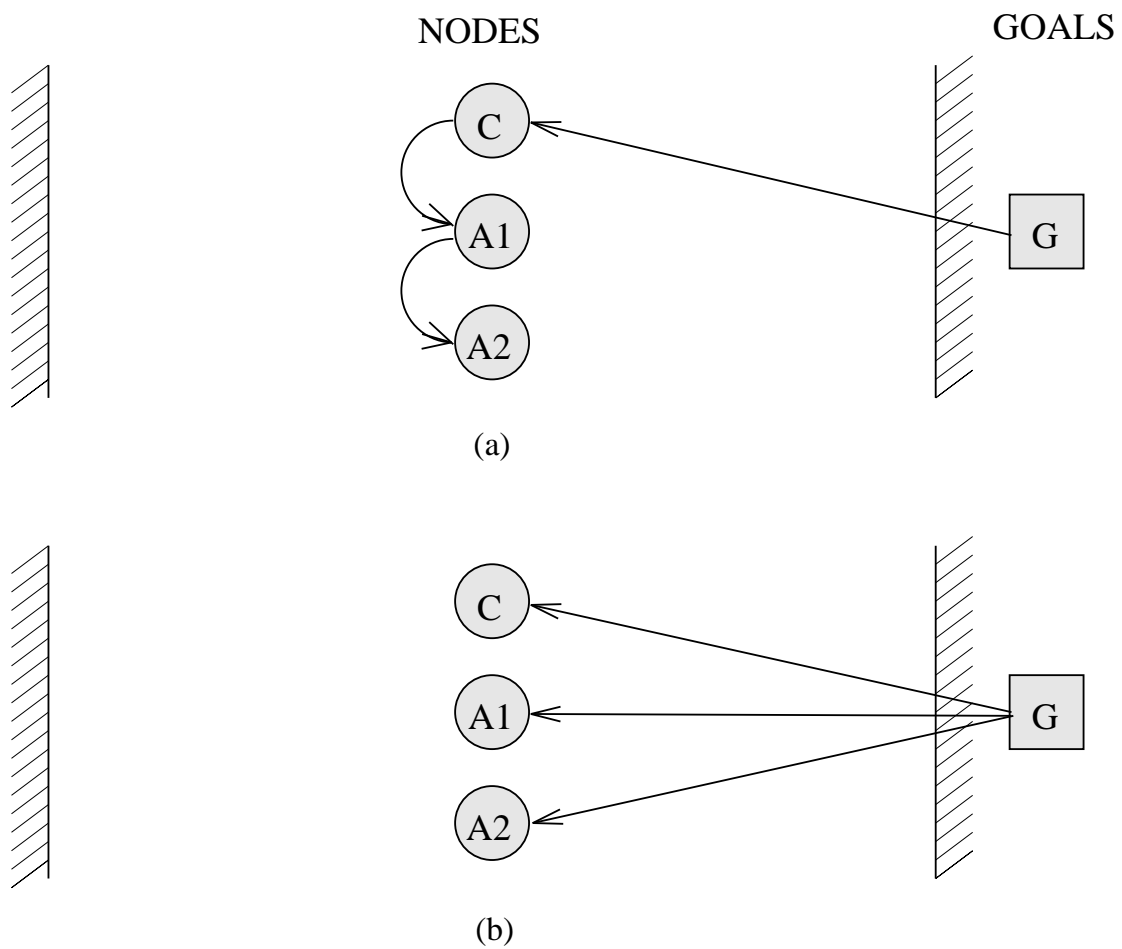


Figure 15: Two possible ways of propagating motivational inputs to appetitive nodes **(a)** via the consummatory node and then in a chain to successively less immediate appetitive actions, or **(b)** directly.

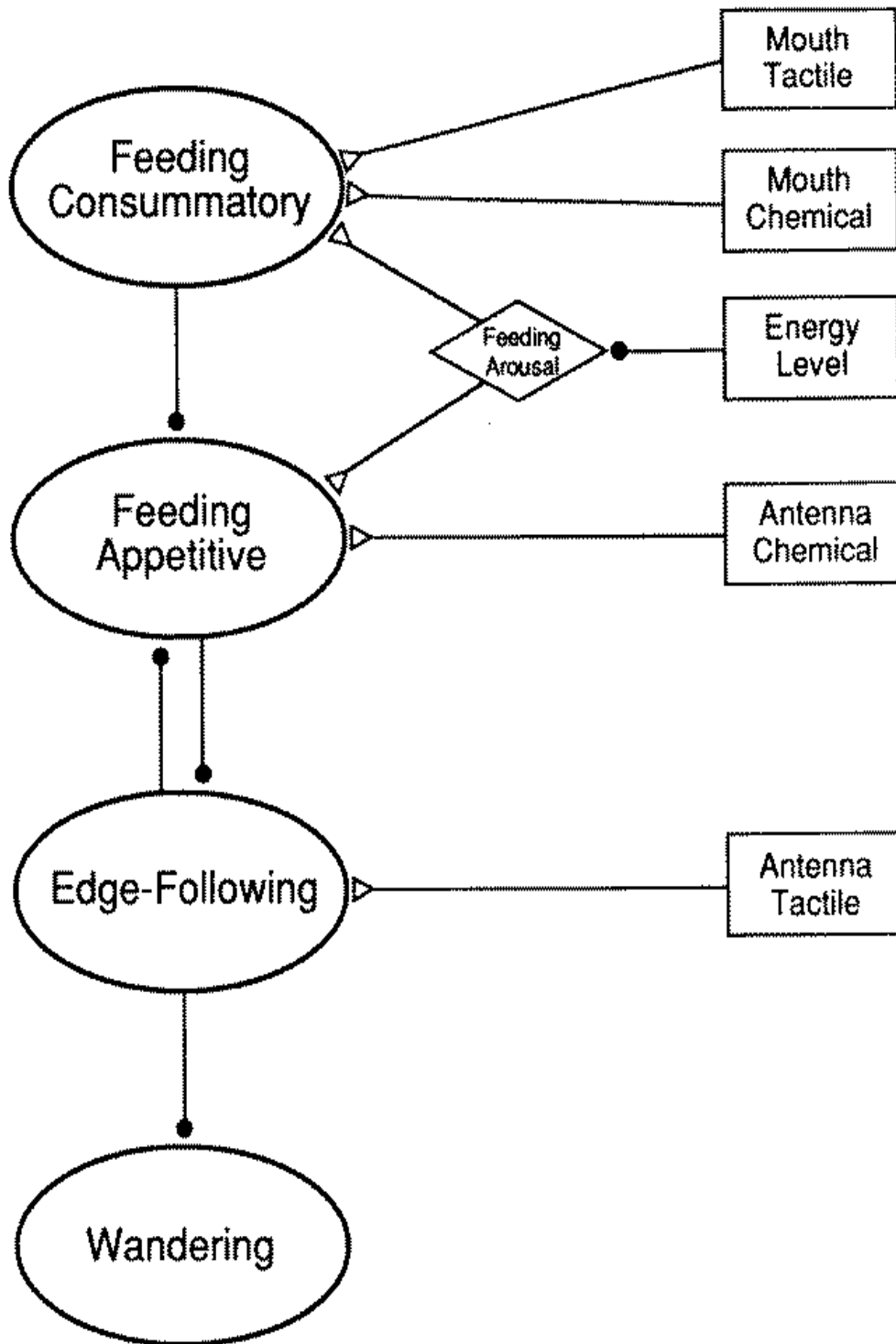


Figure 16: Action selection mechanism for Beer's artificial cockroach. Filled circles denote inhibition and empty triangles denote excitation. The goal or motivational stimulus is "feeding arousal". Taken from [Beer, 90] (© acknowledged to Academic Press).

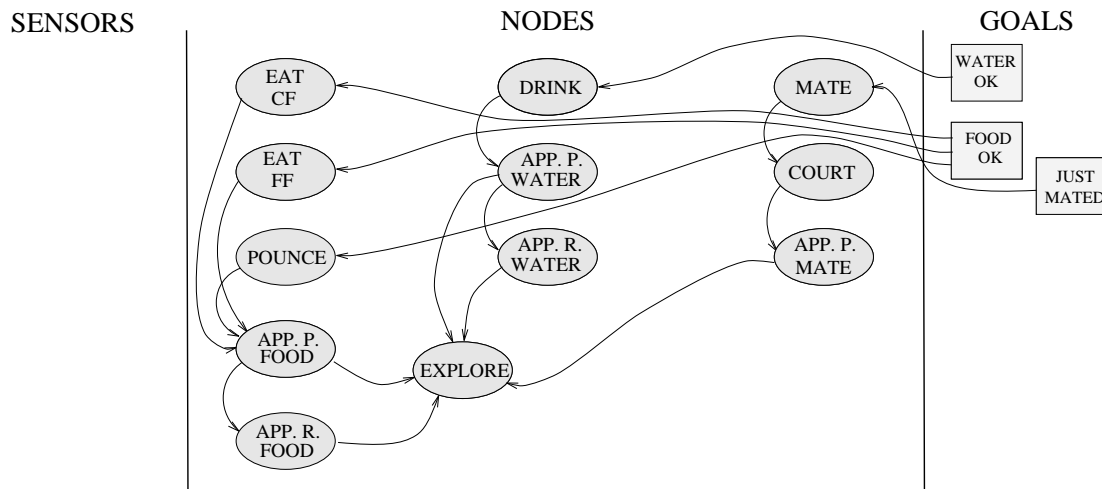


Figure 17: Primary goal and predecessor connections to the ‘get food’, ‘get water’ and ‘reproduce’ systems. ‘Approach perceived food’ receives three predecessor connections from nodes in the same system. ‘Explore’ receives five predecessor connections from nodes in three different systems. ‘P’ stands for ‘perceived’, ‘R’ for ‘remembered’, ‘APP’ for ‘approach’, ‘CF’ for ‘cereal type food’ and ‘FF’ for ‘fruit type food’.

when deciding whether or not to divide by the number of predecessor inputs to a node. Consider the part of the network shown in figure 17, in which the goal or motivational variable “food okay” (i.e. internal food deficit) is propagated to the three consummatory nodes “eat cereal food”, “eat fruit food” and “pounce” (eat prey). When these three are unexecutable then they each pass activation through predecessor links to the node “approach food”. In this case it is proper for activation from the three inputs to “approach food” to be divided by the number of inputs, three, since the three eat nodes have all been fed activation by the same goal (i.e. they all act in response to the same motivational variable). On the other hand, the goals “food okay”, “water okay” and “just mated” can all be propagated through consummatory and appetitive nodes to the ‘furthest removed’ appetitive node “explore” (when no food, mate or water are perceived). In this case it is not proper for activation from the different inputs to “explore” to be divided by the number of them (three), since the nodes providing input to explore all try to achieve different goals, and the execution of “explore” would contribute to three goals simultaneously.

In short, because the various predecessor inputs to a node can all be from nodes in the same system (trying to achieve the same goal), or can be from several systems (trying to achieve several different goals), there is no general division rule for predecessor links that works for all situations. This is illustrated in figure 18. Division by numbers of predecessor inputs causes problems for nodes like “explore”, and the lack of such division causes problems for nodes like “approach perceived food”. The decision as to whether or not to divide by M requires information which is not available to the nodes, i.e. knowledge about whether the different predecessor inputs are due to the same or to different goals. This difficulty with the division rules stems partly from the decision to feed motivational stimuli only to consummatory nodes and indirectly from

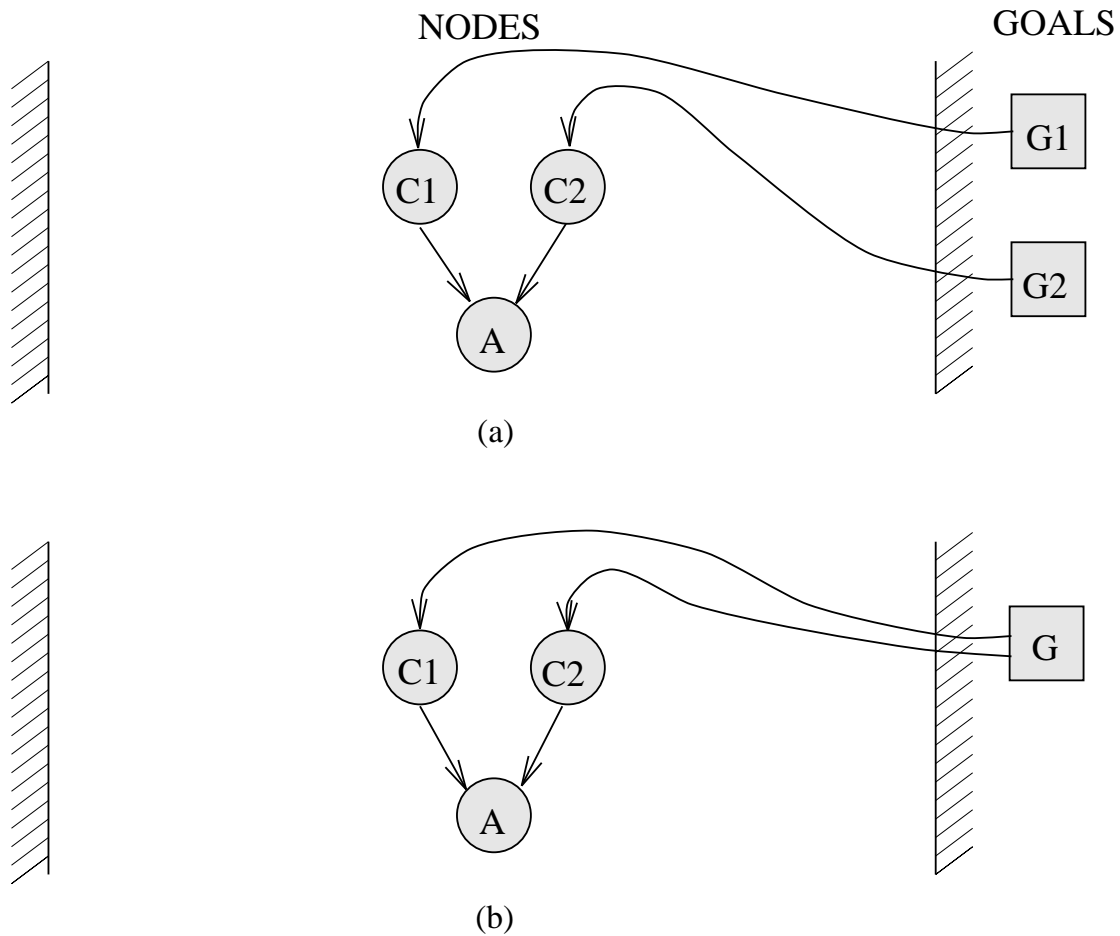


Figure 18: Two examples showing the dilemma when deciding whether or not to divide by the number of predecessor inputs to an appetitive node. Because there is no way of knowing whether or not the inputs derive from the same goal or not, there is no way of knowing whether there should be division by the number of inputs or not.

there to appetitive nodes, rather than directly to both consummatory and appetitive nodes, as is the case in other mechanisms. Section 7.2 contains a discussion of how excitation can be combined when goals are fed directly to appetitive nodes. The shortfalls caused by various configurations of the division rules are summarised in table 4.

6.4. Fourth Implementation.

There is no easy and principled way of resolving this dilemma without making fundamental changes to MASM, but a trade-off can be implemented which takes the average of the input value with division and the input value without division. So, each predecessor input in the fourth implementation was changed so that it was calculated

according to the formula

$$I = \frac{1}{2}(\alpha + \frac{\alpha}{M})$$

and each conflictor input was calculated according to the formula

$$I = \frac{1}{2}(\alpha \frac{\delta}{\gamma} + \frac{\alpha}{M} \frac{\delta}{\gamma})$$

The arguments in section 6.3 do not apply to goal or protected goal inputs, and so in these cases the lack of any division by the number of inputs or outputs was maintained.

Using the new division rules for predecessor links, the fourth implementation was tested with better results than with any of the previous implementations. The performance of the ASM was still poor in comparison to other mechanisms though, and there were problems over and above those caused by the unsatisfactory trade-off for the predecessor inputs. In particular, the main problem was that consummatory nodes seemed to get chosen only very infrequently. When the goal they could help to achieve had a high strength and when all their preconditions were true then they still lost out, on most occasions at least, to appetitive nodes in other systems. So, for instance, on occasions when “approach water” and “eat cereal food” were both executable (when all of their preconditions were true), and goal strengths for “food okay” and “water okay” were of similar value then “approach water” (the appetitive, and therefore less favourable choice) was selected. The optimal selection is always to choose a consummatory node over an appetitive one, all other things being equal. This is because the benefits from consummatory behaviour are more immediate and more certain. Appetitive behaviour by definition does not contribute to fitness unless and until consummatory behaviour follows it.

The reason for this undesirable predominance of appetitive over consummatory nodes is illustrated in figure 19. In this simple example case there are two goals $G1$ and $G2$, each of strength g , and each of which can be achieved only by one consummatory node ($C1$ and $C2$ respectively). For each consummatory node there is one appetitive node which can bring about the truth of its single precondition. If all nodes start off with zero activation then it can be shown, by calculating node activations for successive iterations of the spreading activation process, that $A1$ will come to obtain more activation than $C2$ after three iterations of the cycle, regardless of the values of G , γ and ϕ (see table 5). This is because of a positive feedback loop that exists between $C1$ and $A1$ ($C1$ is not executable), but not between $C2$ and $A2$ ($C2$ is executable and so the predecessor and successor links are not active). The situation is in fact a little more complicated than this because of the effects of normalisation after each cycle, which have not been shown in the table. Normalisation may increase the number of cycles until $A1$ overtakes $C2$, but will not remove the tendency for that to occur.

This phenomenon of exponentially-increasing positive feedback between an unexecutable consummatory and an executable appetitive node means that executable appetitive nodes can often be preferred over executable consummatory nodes in other systems, if more than a few cycles are required to obtain a choice of behaviour.

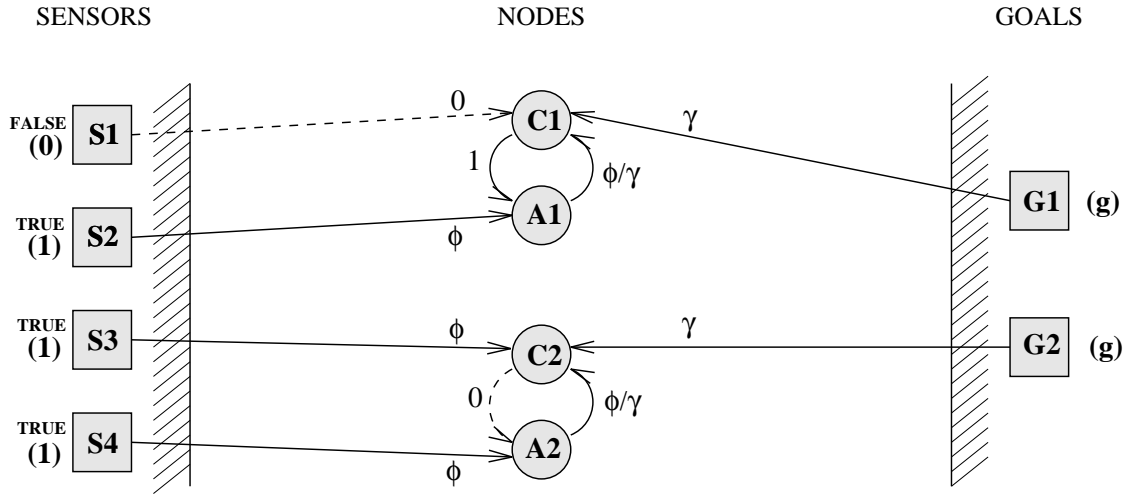


Figure 19: Two identical hypothetical systems, both with one consummatory and one appetitive node. Sensor and goal activations for the current state are shown in parentheses beside them, and connection weights are shown next to the connections. Dashed lines indicate inactive links because of the current state (false sensors input no activation, predecessor links pass activation in response to a false precondition in the sender node, successor links pass activation in response to a false precondition in the receiving node). ϕ and γ are global parameters whose meaning is explained in the text.

	Cycle 0	Cycle 1	Cycle 2	Cycle 3
$C1$	0	γg	$2\gamma g + \frac{\phi^2}{\gamma}$	$3\gamma g + \frac{3\phi^2}{\gamma} + \phi g$
$A1$	0	ϕ	$2\phi + \gamma g$	$3\phi + 3\gamma g + \frac{\phi^2}{\gamma}$
$C2$	0	$\phi + \gamma g$	$2\phi + 2\gamma g$	$3\phi + 3\gamma g$
$A2$	0	ϕ	2ϕ	3ϕ

Table 5: Node activations for successive iterations of the mechanism shown in figure 20. All nodes are started with zero activation. $C2$ should always be ahead of $A1$, but it can be seen that in fact $A1$ accumulates more activation after three cycles.

It was decided at this point that although there were still inadequacies in the performance of MASM, it was not sensible to try and remove these since the alterations required would have entailed radical changes to the architecture and central concepts of MASM, with the removal or alteration of the different types of connections and/or major changes in the way information is passed between nodes.

6.5. Summary of Testing.

The performance of the final implementation of MASM is shown in figure 20, alongside that of some other mechanisms that were tested in the SE. Each mechanism was tested in four slightly different versions of the SE in order to try and make some slight variations to the type of test presented by the SE. The average performance values for MASM in the four versions of the SE were 0.16, 0.22, 0.18 and 0.43 (giving an average over all of the four tests of **0.25**). The expected standard deviations of the four average results (each calculated over 1650 evaluations in the SE) were 0.01, 0.01, 0.01 and 0.02. The averaged results over all of the four versions of the SE (a total of 6600 runs in each case) for the other ASMs shown in figure 20 were: (1) Lorenz = **2.71**, (2) HDS = **6.13**, (3) Drives = **6.23**, and (4) ERP = **8.31**. The performance values for each mechanism should be considered somewhat sceptically though, due to some sources of arbitrariness in the testing and implementation of each mechanism. The graph is only included here in order to show that the deficiencies in MASM did lead to a lower level of performance in the particular test of AS presented by the SE. A full description of the implementation of these other ASMs, the testing procedure employed and the statistical significances of the differences in performance is given in [Tyrrell, 93a].

The two main findings that were uncovered by the testing in the SE were that: *(i)* the exact division rules specified by Maes are untenable, and further, no satisfactory set of division rules can be derived without making use of information about which goals the predecessor inputs are ultimately due to, and *(ii)* the dynamics of the interactions between consummatory and appetitive nodes means that MASM cannot be guaranteed to choose appropriately between consummatory and appetitive alternatives in different systems. These two deficiencies have been discussed in detail above and have been shown to be problems that are not just specific to this test in the SE; they will occur in general whenever there are multiple goals, consummatory and appetitive alternatives, and many nodes to choose between. It was also shown that the deficiencies are significant in their effect, and that they are firmly linked to a fundamental aspect of the mechanism; namely the decision to pass goal-type stimuli in a chain-like fashion, to the consummatory node first and then indirectly to the appetitive nodes in turn if the respective predecessor connections are active.

Noting these two drawbacks, it is now worth looking back at the list of desiderata in section 2.1. Unlike the planning and reactive systems that MASM was designed to supersede, MASM takes account of both motivations and external stimuli (points 1 and 2). With respect to point 3, MASM selects appropriately between consummatory and appetitive alternatives in the same system, but it has been shown that it is not guaranteed to do so when the alternatives are in different systems. MASM does have a tendency to complete sequences when started (point 4), and is also able to

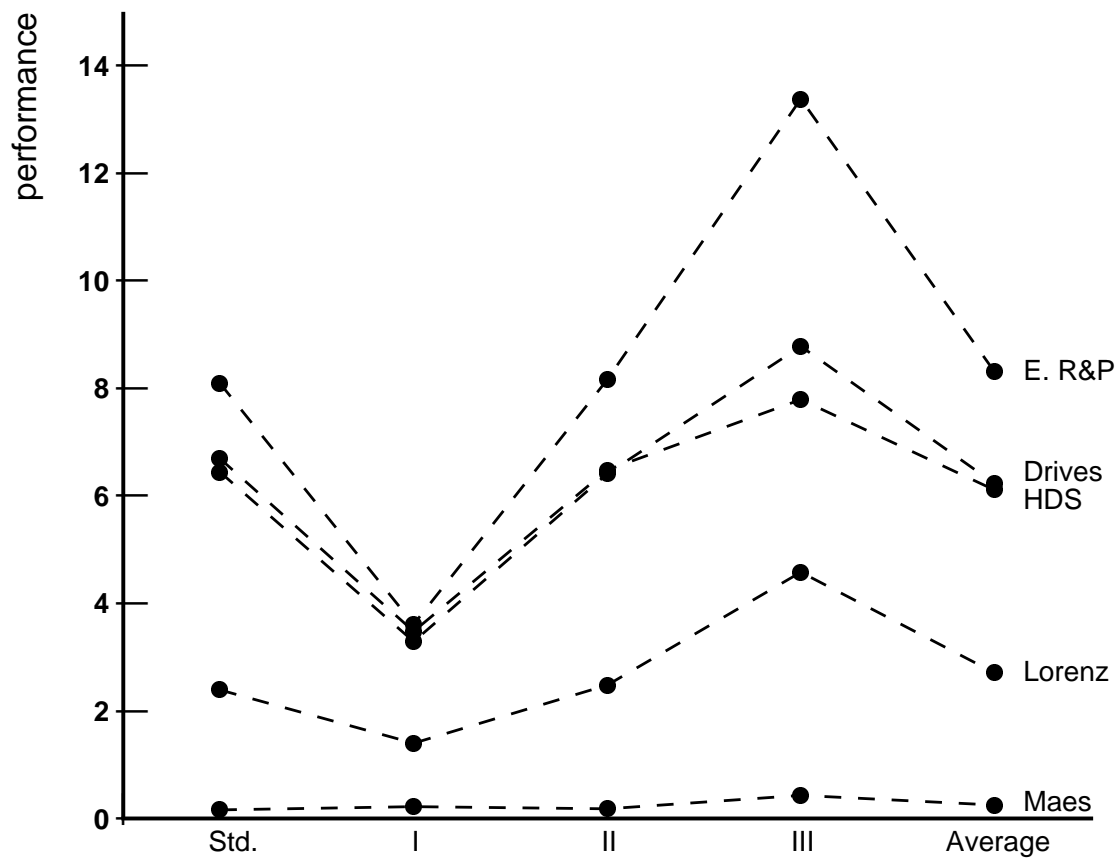


Figure 20: A graph showing the performance measurements (genetic fitnesses) of the final implementations of some ASMs that were tested in the SE. **Std** stands for the normal SE, **I** stands for the first altered version of the SE, **II** stands for the second altered version of the SE and **III** stands for third altered version of the SE.

interrupt an ongoing sequence if necessary (point 5). The division rules proposed do not, in their stated form, allow a node contributing to more than one goal to accrue more activation (point 6). To summarise, MASM satisfies some, but not all, of the desiderata for an ASM listed in section 2.1.

7. Discussion.

The previous section has outlined some drawbacks with the way that MASM calculates AS, but how can actions be selected more effectively? Although it isn't possible to describe other mechanisms in detail here, or to give a full and fair comparison between MASM and these other ASMs, in this section some aspects of the other ASMs are compared to the corresponding parts of MASM, and some suggestions are made as to how MASM might be improved upon. For an extended comparison of the different ASMs see [Tyrrell, 93a].

7.1. How Should Motivational Information be Propagated?

This question was discussed in the previous section. It was described how information from goals (motivational stimuli) can be propagated to consummatory and appetitive nodes in different ways. As shown in figure 15a, in MASM goal information is fed directly to the consummatory node only, and may then be passed 'down the chain' to the appetitive nodes. In some other mechanisms [e.g. Beer, 90; Halperin, 90; Tyrrell, 93a] the goal information is passed directly to all of the appetitive as well as the consummatory nodes.

It was demonstrated in the previous section that the indirect method of passing motivational information possesses hidden difficulties because of the impossibility of a totally satisfactory set of general rules for calculating the weights on the predecessor connections (due to a dilemma over whether any two predecessor-type inputs are due to the same goal or different goals). The direct method of passing motivational information to consummatory and appetitive nodes is thought to be preferable because it is simpler and does not give rise to the same dilemma [Tyrrell, 93a].

7.2. How Should Consummatory and Appetitive Alternatives be Arbitrated Between?

As mentioned earlier, consummatory actions should be preferred to appetitive ones because they bring about a concrete benefit to the animal more reliably and more immediately (with less eventual expenditure of time). Similarly, 'closer' appetitive options in the chain are preferable to ones which are 'further away' from the consummatory option (e.g. approaching water, if a reasonably good source of water is perceived, is preferable to exploring for it).

Three different ways of ensuring that the consummatory node is chosen (if relevant to the current state of the environment), or else the closest relevant appetitive node is chosen, are shown in figure 21. A 'relevant' node is one which is appropriate to the situation which the animal is in at that moment in time (and so the node "drink" is only relevant when the animal is at a water source, and "run away from predator" is only relevant when the animal can perceive a predator).

MASM ensures that the earliest node in the sequence that is appropriate to the current situation is chosen by passing motivational stimuli 'down the chain' only so far as the first relevant node (figure 21a). Beer [90] and Halperin [91] have used a system in which motivational stimuli are fed directly to the consummatory and appetitive

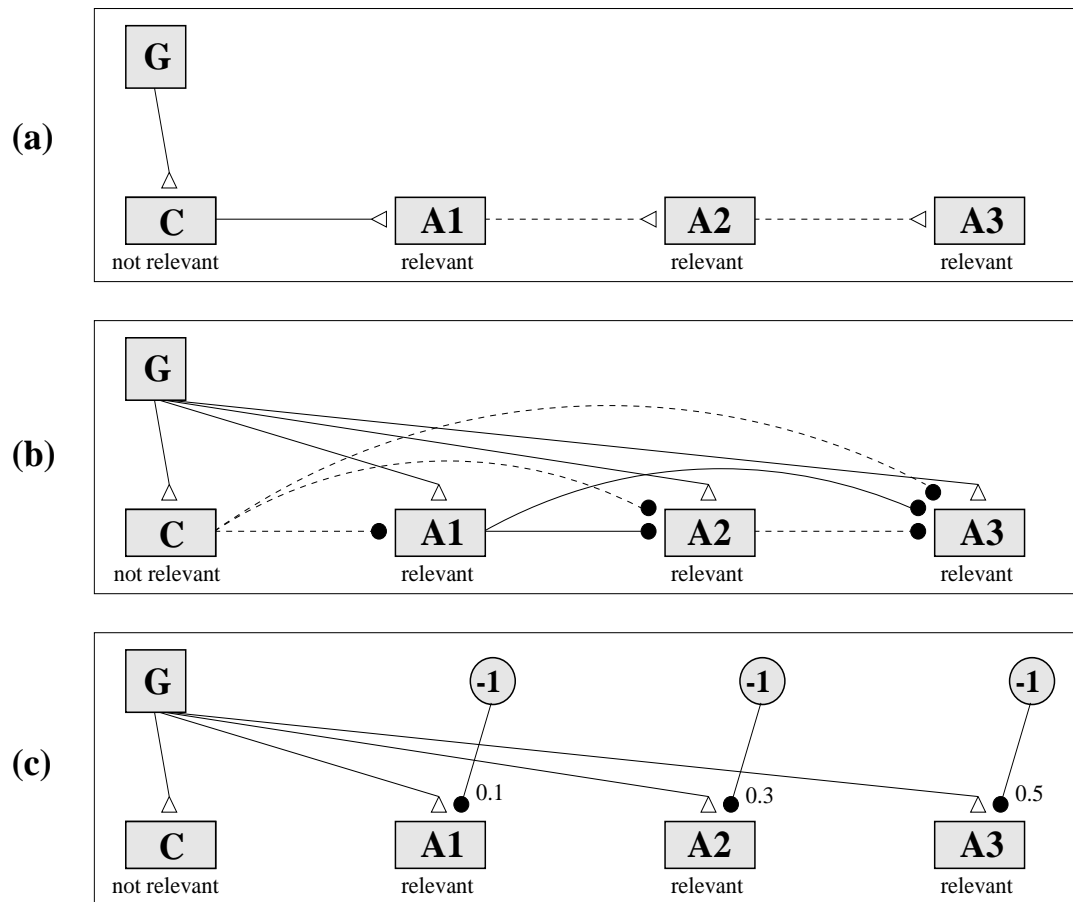


Figure 21: Three different ways of using excitation and inhibition to arbitrate between consummatory and appetitive nodes. A consummatory node should be preferred over an appetitive one, if both are equally relevant to the state of the environment. Empty triangles denote excitatory connections, solid circles denote inhibitory connections, and dashed lines denote links that are currently inactive.

nodes, and in which the first node that is relevant to the current situation can exert a severe inhibition on all later nodes (figure 21b). The author has suggested a third scheme [Tyrrell, 93a], in which motivational stimuli are fed directly to all of the consummatory and appetitive nodes, and in which there are constant inhibitory penalties on appetitive nodes only (figure 21c). These penalties increase as the appetitive nodes are further removed from the consummatory ones. These penalties give the earlier nodes in the sequence an advantage over later ones (in the competition to accumulate the most activation).

All three means of discriminating in favour of earlier nodes in the sequence work when only one system is considered. However, it was shown in the previous section that MASM is not guaranteed to arbitrate properly between consummatory and appetitive nodes in separate systems, due to the occurrence of a mutually-reinforcing feedback loop between an unexecutable node and an executable node one along in the sequence. There is also nothing in the second scheme to make consummatory nodes in one system be preferred over appetitive nodes in another system (an appetitive node in one system can amass similar activation to a consummatory node in another system), whereas the third scheme is able to bring about proper arbitration between nodes from different systems, as well as between those in the same system.

Another consideration with action sequences is the desirability of continuing a sequence to completion once it has been started, assuming that the sequence will have to be started again from scratch if it is abandoned half-way through. Because an executable node in MASM passes activation to the next node in the sequence, and the activation is partially carried over to the next timestep, a tendency for contiguous sequences can be encoded with this scheme. It is less obvious how this would be done with the second or third schemes in figure 21.

7.3. A Hierarchical or Non-Hierarchical Structure?

Maes pointed out several inadequacies with traditional, rigid-switching type hierarchical mechanisms, such as the top-down, rigid sort of control they exert and the consequent lack of robustness [Maes, 91a]. In response to these problems with the traditional type of hierarchy, Maes designed her mechanism to be non-hierarchical. However, other work [Rosenblatt & Payton, 89; Tyrrell, 93b] has looked at a different way of using a hierarchical architecture, and it was argued that ‘free-flow hierarchies’, which do not have a switching process between nodes at the same level and which allow excitation to flow unhindered through the hierarchy, do not suffer from the deficits associated with the traditional type of hierarchy. In [Tyrrell, 93b] it is argued that these free-flow hierarchies are preferable to non-hierarchical mechanisms because they allow for the expression of preferences between layers, and the choice of compromise candidate actions. MASM does not allow for the choice of compromise actions because consummatory and appetitive nodes are not allowed to express preferences (i.e. to ‘vote’) for different actions such as “move south”. Instead, a consummatory or appetitive node is selected and then the ‘executable code’ is executed, with the effect that the best action for that node is selected, without simultaneously taking into account any preferences from other nodes.

An example of the usefulness of expressing preferences, i.e. voting for different actions,

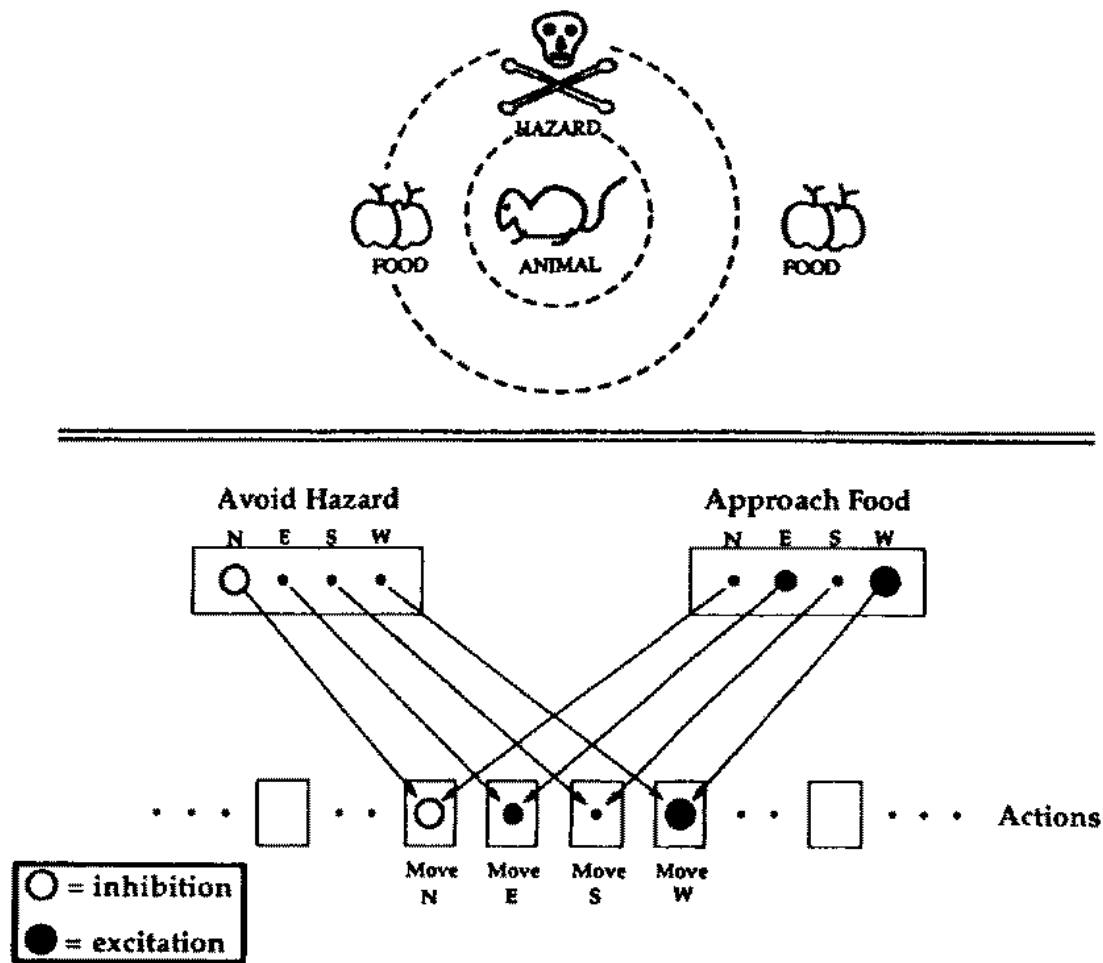


Figure 22: Choosing a compromise candidate by combining preferences from proscriptive and prescriptive systems (chosen action is MOVE_SOUTH if the needs of only one system can be considered at one time, and is MOVE_WEST if combination of preferences is allowed). Larger circles denote greater levels of excitation or inhibition.

is shown in figure 22, in which a compromise candidate action is chosen by combining the preferences from a proscriptive and prescriptive node [Tyrrell, 92].

7.4. A Digital or an Analogue World?

The sensors of the environment (stimuli) for MASM are constrained to take binary values. This results in a loss of information because many properties of a real-world environment are continuous (for instance, the *amount* or *quality* of water at a water source, the *degree* of protection offered by different instances of vegetation, or the *distance* to a predator). For these cases the sensors in MASM will be TRUE if the values of the continuous variables are above a certain threshold and will be FALSE otherwise (and so, for instance, “water in square” might only be TRUE if the level of the water source in the animal’s square is greater than a certain value). The sensors

will return the same readings for large variations in the values of the continuous variables, and there is an inevitable loss of potentially useful information.

In the initial specifications of MASM [Maes, 89 and 90] goals were also binary (ON/OFF), but in the later version [Maes, 91a] they were allowed to take real values.

Similarly, the relationships between nodes and goals in MASM are of a true/false or on/off nature. A node is assumed to either fully achieve a goal or not, and to either completely counter-act a goal or not. There is no way of expressing the fact that the action/behaviour generated by a node can achieve or undo a goal to some degree, or with some probability. For instance, strenuous actions such as “mate” may tend to increase body temperature by a small amount, thus slightly contributing to a goal “keep warm”, and slightly counteracting a goal “keep cool”. But the action is unlikely to achieve the one goal or undo the other by itself. To give another example, because of the error-prone nature of the animal’s navigation, the execution of the node “approach remembered food” cannot be guaranteed to take the animal to a food source. There is a probability that this will occur, but this probability is less than 1.

In general, the use of graded or continuous variables and relationships would be preferable to the binary ones used here.

8. Conclusions.

Two particular inadequacies in the performance of Maes' "bottom-up mechanism for behaviour selection" have been demonstrated. Firstly, it was shown that the form of the division rules that was originally proposed by Maes is unsatisfactory, and also that no satisfactory set of division rules exists for the predecessor and conflictor connections, unless extra information is passed through the mechanism. Secondly, it was shown that a mutually-reinforcing, exponentially-increasing feedback loop can exist between an unexecutable consummatory node and an executable appetitive node, thus leading to inappropriate selections of actions.

It was argued that these inadequacies could lead to a significantly lower level of performance, and that they are inevitably tied to central principles of the mechanism's design. It was also argued that the inadequacies are not specific to just the particular problem posed here, but instead that they will be more generally applicable to other complex, animal-like action selection problems. The two deficits will not be apparent in all tests of action selection though, just those in which there can be two or more goals that are active at the same moment in time, and that each have candidate actions which are appropriate given the current state of the environment.

In summary, Maes' mechanism copes satisfactorily with many aspects of the action selection problem, and is a distinct improvement over the AI mechanisms that it was designed to succeed. In addition, the mechanism is specified in a commendably formal and explicit manner, making it easier to analyse the mechanism and to determine its shortcomings. However, some deficiencies in the design of the mechanism mean that it will not perform well on many animal-like action selection problems.

References

- Agre, P.E. 1990. Book Review. *Artificial Intelligence*. 43:369–384.
- Agre, P.E. and Chapman, D.A. 1987. Pengi: An Implementation of a Theory of Activity. In *Proceedings of the Sixth National Conference on Artificial Intelligence [AAAI-87]*, Morgan Kaufmann.
- Baerends, G.P. 1976. The Functional Organisation of Behaviour. *Animal Behaviour*. 24:726–738.
- Beer, R.D. 1990. *Intelligence as Adaptive Behavior: an Experiment in Computational Neuroethology*. Academic Press.
- Dawkins, R. 1989. *The Selfish Gene* (2nd edition), Oxford University Press.
- Dewsbury, D.A. 1978. *Comparative Animal Behavior*. McGraw Hill.
- Halperin, J.R.P. 1990. *A Connectionist Neural Network Model of Aggression*. PhD thesis. Dept. of Zoology, University of Toronto.
- Halperin, J.R.P. 1991. Machine Motivation. In *Proceedings of the First International Conference on Simulation of Adaptive Behavior*, (JA Meyer & SW Wilson, eds). MIT Press/Bradford Books.
- Kaelbling, L. 1987. An Architecture for Intelligent Reactive Systems. In *Reasoning about Actions and Plans: Proceedings of the 1986 Workshop*, Morgan Kaufmann.
- Kilmer, W.L. and McCulloch, W.S. and Blum, J. 1969. A Model of the Vertebrate Central Command System. *International Journal of Man-Machine Studies*. 1:279–309.
- Lorenz, K. 1950. The Comparative Method in Studying Innate Behaviour Patterns. *Symposia of the Society for Experimental Biology*. 4:221–268.
- Lorenz, K. 1985. *Foundations of Ethology*. Heidelberg: Springer-Verlag.
- Maes, P. 1989. How to Do the Right Thing. *Connection Science*. 1(3):291–323.
- Maes, P. 1990. Situated Agents Can Have Goals. *Journal for Robotics and Autonomous Systems*. 6(1 &2).
- Maes, P. 1991a. Bottom-Up Mechanism for Behavior Selection In An Artificial Creature. In *Proceedings of the First International Conference on Simulation of Adaptive Behavior*, (JA Meyer & SW Wilson, eds). MIT Press/Bradford Books.
- Maes, P. 1991b. The Agent Network Architecture (ANA). In *Proceedings of the AAI Spring Symposium on Integrated Intelligent Architectures*. AAI Press.
- Maes, P. 1991c. Adaptive Action Selection. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*.
- McFarland, D.J. 1985. *Animal Behaviour*. Longman.
- McFarland, D.J. and Houston, A.I. 1981. *Quantitative Ethology: The State Space Approach*, Pitman.
- McFarland, D.J. and Sibly, R.M. 1975. The Behavioural Final Common Path. *Philo-*

- sophical Transactions of the Royal Society* (Series B). **270**:265–293.
- Morris, D. 1990. *Animal Watching*. Jonathan Cape.
- Pebody, M. 1991. *How to Make a Lego Vehicle Do the Right Thing*, MSc Thesis, Department of Artificial Intelligence, University of Edinburgh.
- Rosenblatt, K.J. & Payton, D.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*.
- Tinbergen, N. 1950. The Hierarchical Organisation of Nervous Mechanisms Underlying Instinctive Behaviour. *Symposia of the Society for Experimental Biology* 4:305–312.
- Tinbergen, N. 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.
- Tyrrell, T. 1993b. The Use of Hierarchies for Action Selection. *Journal of Adaptive Behavior*, **1**(4):387–419.