Computational Mechanisms for Action Selection

Toby Tyrrell

PhD
University of Edinburgh
1993

Abstract

Imagine a zebra in the African savannah. At each moment in time this zebra has to weigh up alternative courses of action before deciding which will be most beneficial to it. For instance, it may want to graze because it is short of food, or it may want to head towards a water hole because it is short of water, or it may want to remain motionless in order to avoid detection by the predator it can see lurking nearby. This is an example of the problem of *action selection*: how to choose, at each moment in time, the most appropriate out of a repertoire of possible actions.

This thesis investigates action selection in a novel way and makes three main contributions. Firstly, a description is given of a simulated environment which is an extensive and detailed simulation of the problem of action selection for animals. Secondly, this simulated environment is used to investigate the adequacy of several theories of action selection such as the drive model, Lorenz's hydraulic model and Maes' spreading activation network. Thirdly, a new approach to action selection is developed which determines the most appropriate action in a principled way, and which does not suffer from the inherent shortcomings found in other methods.

Contents

Prologue

This thesis is an investigation of both the problem of *action selection* and the computational mechanisms for providing solutions to that problem.

The problem of action selection is that of choosing at each moment in time the most appropriate action out of a repertoire of possible actions (e.g. drinking, sleeping, moving, grooming). The problem can also be thought of as one of time-allocation; the animal has to decide how to apportion its available time so as to simultaneously satisfy many needs. Action selection is separate from perception (calculation and classification of stimuli) and motor control (execution of actions). The process of making a decision as to which action is most appropriate will need to take into account many stimuli, some of which will be external (e.g. nearby predators, food sources, mates) and some of which will be internal (e.g. blood sugar levels, body temperature). An example situation in which several different stimuli are important is shown in figure 0.1. Different action selection mechanisms take these stimuli and compute a preferred action in different ways. The simulated environment to be described in the first part of this thesis provides a method for testing and comparing these different computational approaches to action selection.

The action selection mechanisms considered here are somewhat similar to planning systems, in that they produce sequences of actions which achieve useful objectives. However, the mechanisms considered here are more reactive (responding to the immediate state of the environment) than 'deliberative' (following long-term plans formulated some time in the past). It is worth noting that while the mechanisms considered here do not involve explicit plans, they are to some extent *goal-driven*, in that the behaviour will tend to be oriented towards correcting disturbances in homeostatic variables or rectifying undesirable external situations. They are, in the main, designed to account for how animals can produce appropriate behaviour in natural environments, rather than to explain how people can solve traditional planning problems such as how to make coffee or how to stack blocks on top of one another in a certain way.

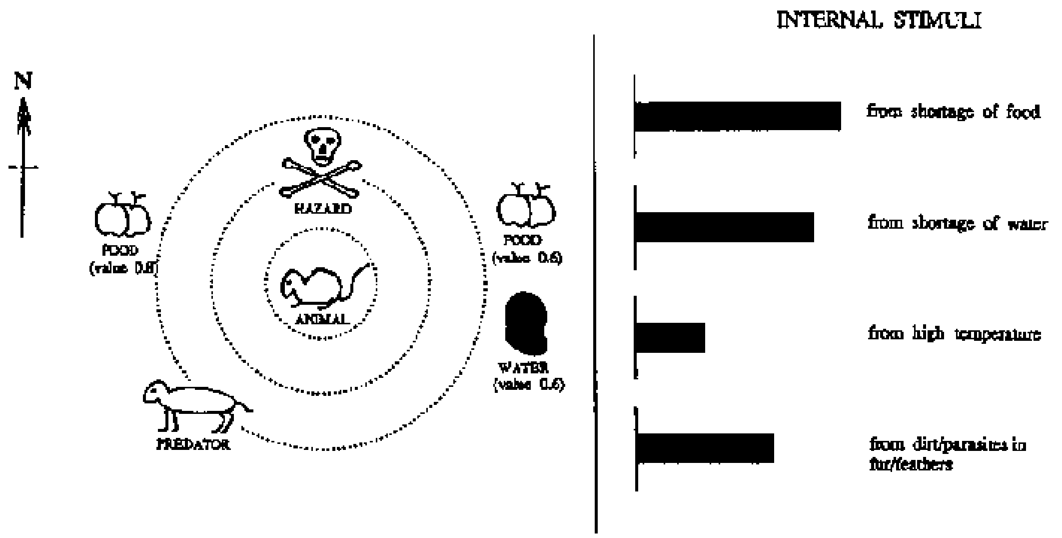


Figure 0.1: An example situation in which many different stimuli need to be taken into account in order to decide what the best action is. In this case the animal has several feasible actions to choose from: should it move south directly away from the hazard, move north-east directly away from the predator, 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.

The mechanisms considered here do *not* incorporate learning. The aim here is to examine the sorts of algorithms or computational methods which can be used to select actions, rather than to examine the ways in which the selection of actions can be improved by experience. For this reason, methods such as temporal difference learning [Sutton, 88] and Q-learning [Watkins, 89] are not covered in any depth here.

Terminology

Some terms which will be used extensively throughout this thesis are now defined in conjunction with figures 0.2, 0.3 and 0.4:

1. **internal stimuli** (also known as endogenous stimuli) — these are perceptions of the state of the animal's internal environment, such as perceived internal water deficit, perceived body temperature, etc.

2. **external stimuli** (also known as exogenous stimuli) — these are perceptions of the state of the external world, such as perceived food sources, perceived predators, etc.

3. **indeterminate stimuli** — some stimuli which are relevant to the animal are neither obviously internal nor obviously external. For instance, circadian rhythms in many animals are based on an internal clock which gets continually reset by the external day/ night cycle [e.g. Marler & Hamilton, 66]. Similarly, in order for the animal not to get lost, it must have access to some measurement of how well it knows where it is. That is to say, the animal must respond to a stimulus which is related to the variance of its estimate of its own position (see section 3.4). This stimulus (the variance of the animal's estimate of its position) can not be directly obtained from the current state of the external environment, and so it is not an obvious external stimulus. However, it *is* dependent on the external environment (encountering recognised features) rather than the internal environment. Any stimuli, such as the two just described, which do not fall neatly into either of the first two categories are classed here as indeterminate stimuli.

4. **action** — this is the name given to an entity at the lowest level, such as moving in a certain direction, sleeping or freezing (remaining motionless to avoid detection by predators). It is assumed that actions lie at the level of the 'behavioural final common path' [e.g. McFarland & Sibly, 75]. That is to say, the animal is forced to choose only one action to be executed at any moment in time. Different actions are incompatible in terms of their demands on the motor system of the animal, and are therefore mutually exclusive. Actions and *fixed action patterns* (rigid sequences of actions) are shown on the right-hand side of figure 0.2 (e.g. 'looking around', 'trampling' and 'pecking').

5. **action selection (AS)** — this is the process of choosing, at each moment in time, the most appropriate action with regard to all types of stimuli.

6. **action selection problem** — this is the overall problem for an animal of how to select its actions so as to maximise its future expected genetic fitness by surviving and reproducing, and by helping close genetic relatives to do likewise.

7. **action selection mechanism (ASM)** — this is a computational mechanism which can produce a selected action as output when given different stimuli as inputs. The problem of action selection is the problem of *what* action the animal should choose at each moment in time; an action selection mechanism specifies *how* those actions are chosen.

8. **simulated environment (SE)** — this is a computer program which models the action selection problem for an animal and thereby allows the implementation and evaluation of action selection mechanisms. The abbreviation SE will used to refer to simulated environments in general. 'The SE' will be used to denote the simulated environment proposed in part I of this thesis.

9. **sub-problem** — the overall action selection problem creates other subsidiary problems. For instance, the need to survive creates subsidiary problems of maintaining food, water and temperature levels within certain bounds, avoiding predators, keeping fur/ feathers in good condition, etc.

10. **system** — each sub-problem of the overall action selection problem will have a separate system of the action selection mechanism devoted to it. Systems frequently overlap though. For instance, two systems might both make use of the action 'rest'. System-level nodes are shown on the left-hand side of figure 0.2 (e.g. 'nesting', 'escape' and 'preening'). A system comprises not just the system-level node, but also all other nodes 'beneath' it. It is equivalent to a 'sub-tree', not a node in a tree.

11. **sub-system** — this is intermediate in size and abstraction between an action and a system. Some sub-systems in figure 0.2 are 'settling', 'locomotion' and 'trimming'.

The terms system, sub-system and action, as used here to describe entities on the 'solution side', are similar to terms used by Baerends [1976], but rather dissimilar to the terminology proposed in [McFarland & Sibly, 75]. These differences are discussed further in section 5.4.

Overview

Part I of the thesis describes a simulated environment that was developed to enable the experimental testing of action selection mechanisms. Part II considers mechanisms for action selection and contains several contributions:

1. the simulated environment is used to test and compare several different action selection mechanisms. Shortcomings of various mechanisms are described, and theoretical analyses explaining each inadequacy are presented. Similarly, differences in performance between mechanisms are described and explained.

2. a list of requirements for action selection mechanisms is given. Some examples of these requirements are: mechanisms must take account of the hierarchical nature of action selection; *decisions* must only be made at the level of the behavioural final common path; stimuli must be represented as real-valued variables; etc.

3. a novel mechanism for action selection is presented.

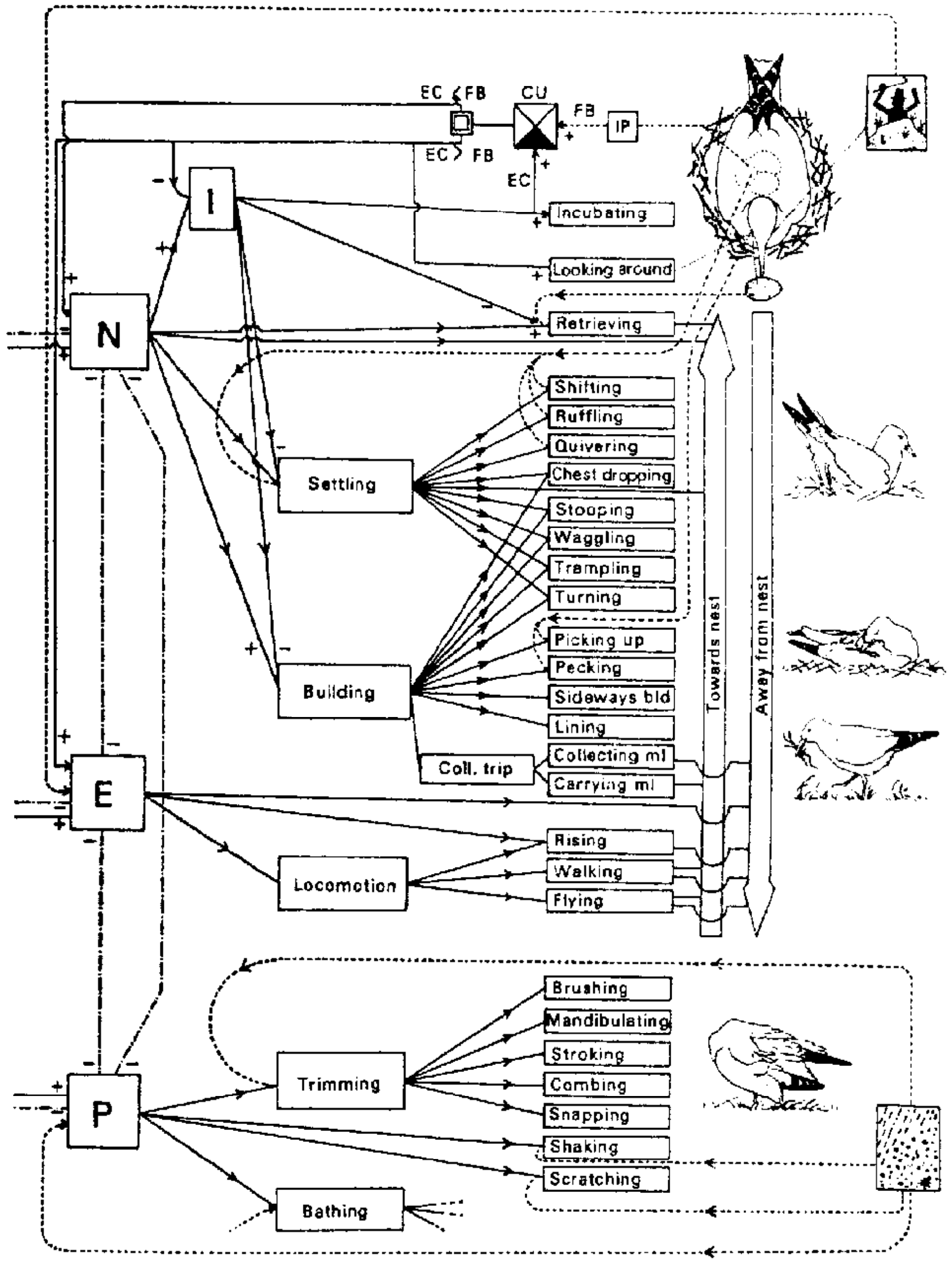


Figure 0.2: Baerends' model to account for incubation behaviour of the herring gull. Actions or fixed action patterns are in the far right column. "Superimposed control systems" of higher order (sub-systems and systems) are to the middle and left. **N** = incubation system, **E** = escape system and **P** = preening system. Taken from [Baerends, 76].

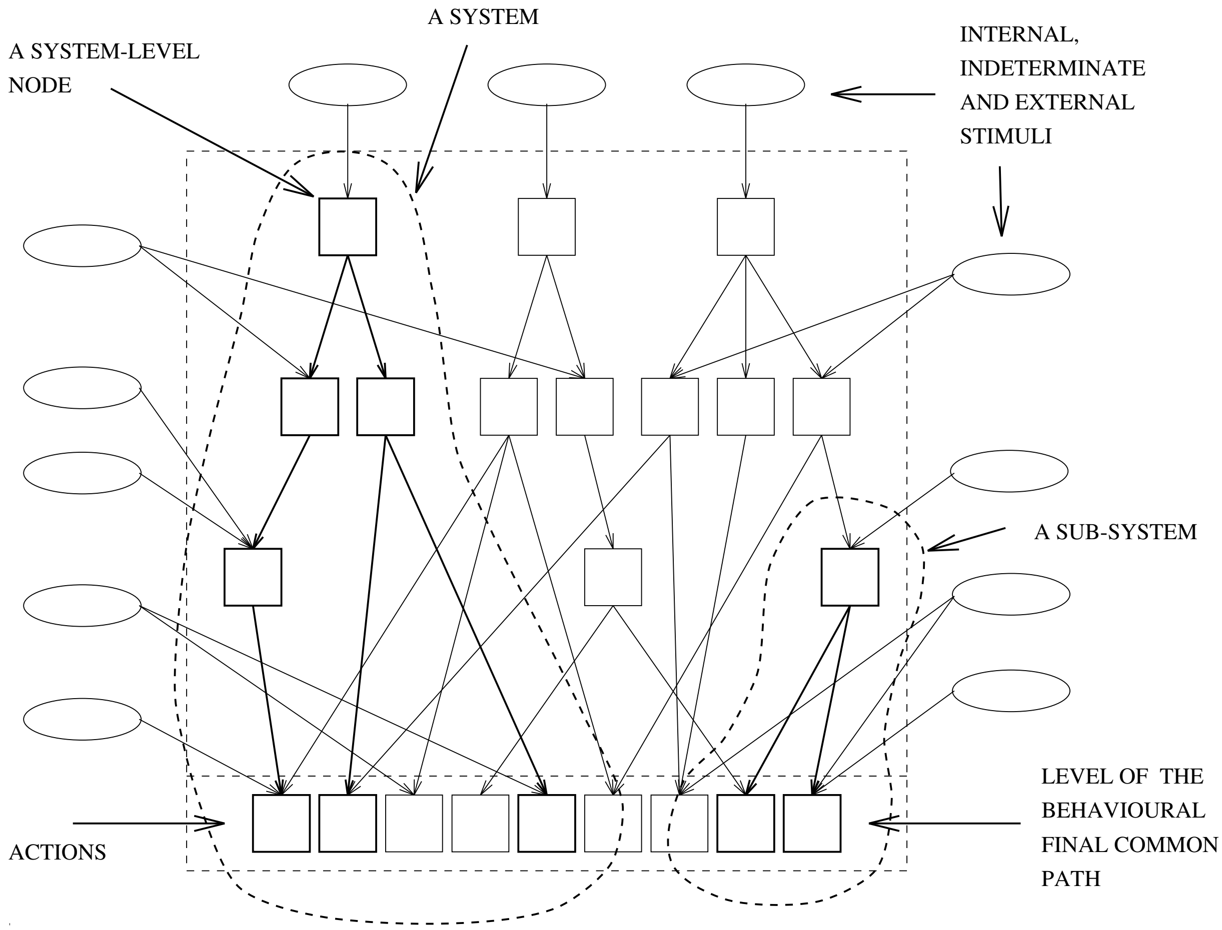


Figure 0.3: An example of the structure of a hierarchical action selection mechanism. The terms system, sub-system and action are illustrated.

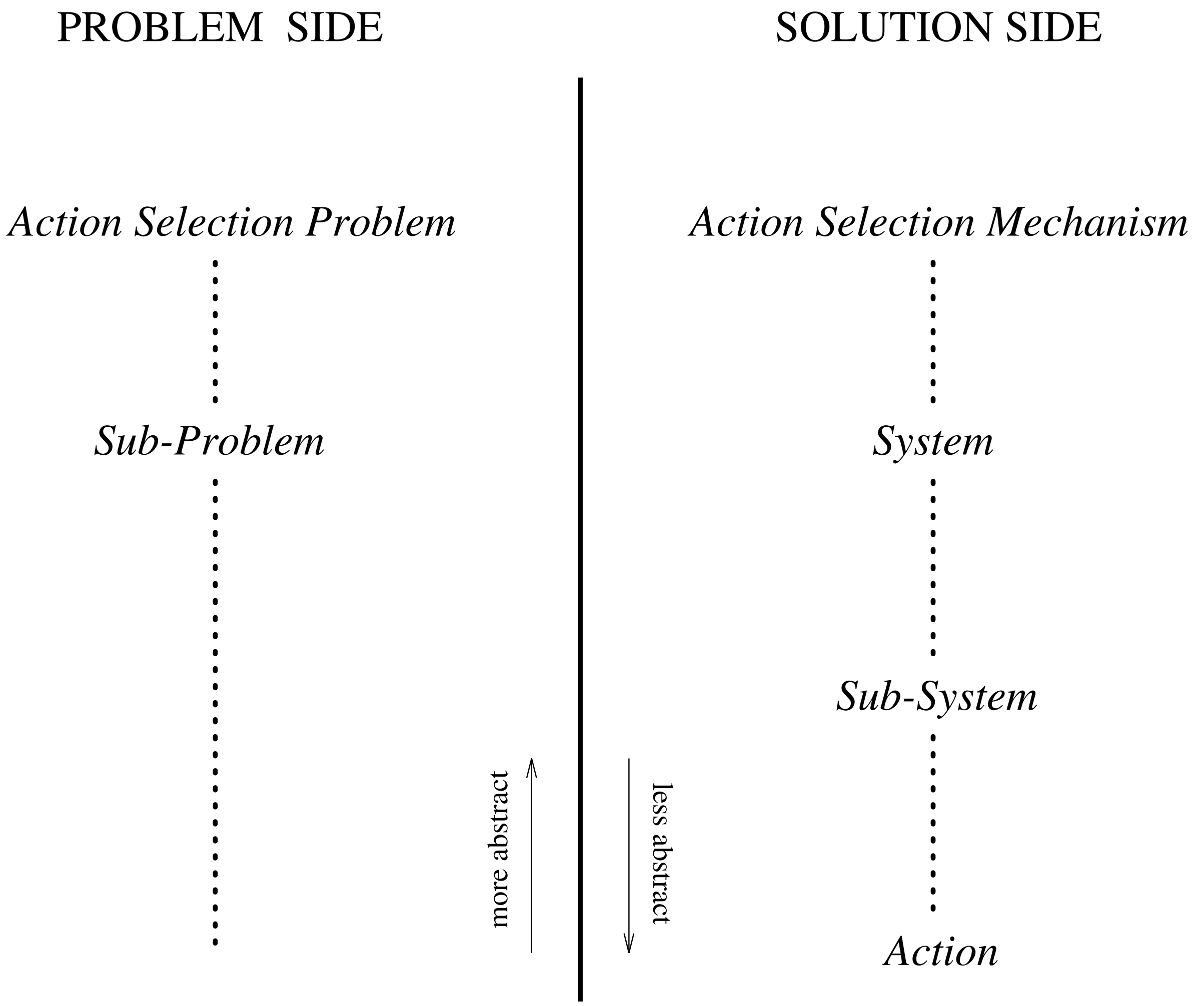


Figure 0.4: Labels for different entities in the description of the problem for the animal, and in the description of the solutions to that problem.

Part I

A Simulated Environment

Chapter 1

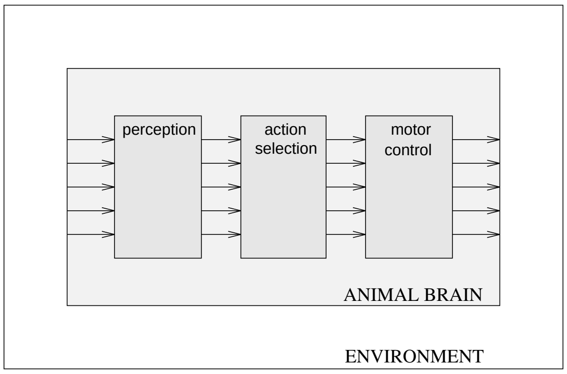
Introduction

Since the time of Charles Darwin it has been accepted that the physical characteristics of animals are the result of evolution due to the processes of natural selection. It is now accepted that the behaviour of animals has also been formed by natural selection.

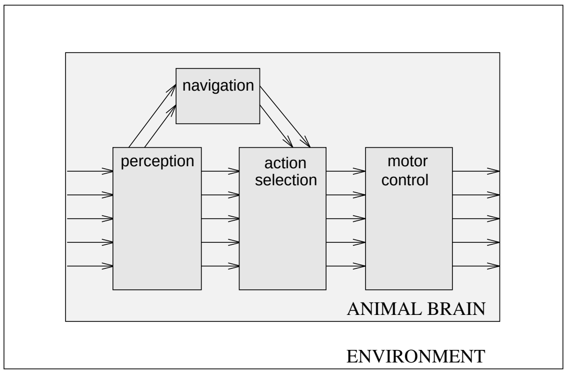
The part of animals that is responsible for controlling behaviour is the central nervous system. As Albus [81] stated "All brains, even those of the tiniest insects, generate and control behaviour". Although brains can be responsible for other functions (e.g. control of the autonomic system), their predominant function is the generation of appropriate behaviour.

This basic task of animal brains has often been split into three sub-tasks as shown in figure 1.1a. A slightly modified system has been assumed here with the addition of a fourth sub-task, navigation, as shown in figure 1.1b. The four sub-tasks are therefore:

1. Sensing of the environment and interpretation of the sensory signals to provide a high-level description of the internal and external environment at each moment in time (**perception**).
2. Keeping track of where the animal is in its environment, and remembering the positions of important features in that environment (**navigation**).
3. Using the perceptual and navigational inputs to decide which of the animal's repertoire of actions is most appropriate at that moment in time (**action selection**).
4. Transforming the chosen action into a pattern of contractions and relaxations of muscles so as to produce movements and rotations of parts of the body (**motor control**).



(a)



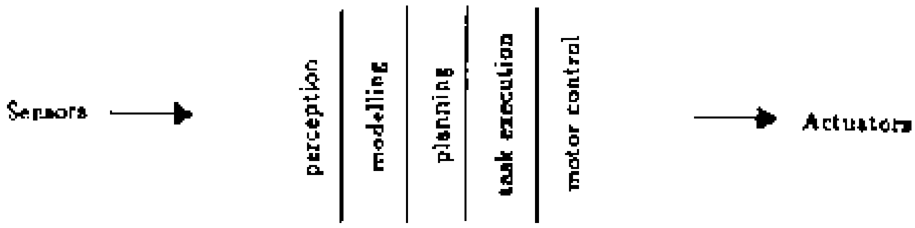
(b)

Figure 1.1: Two decompositions of the function of an animal's brain: **(a)** the more traditional decomposition, and **(b)** the decomposition used here.

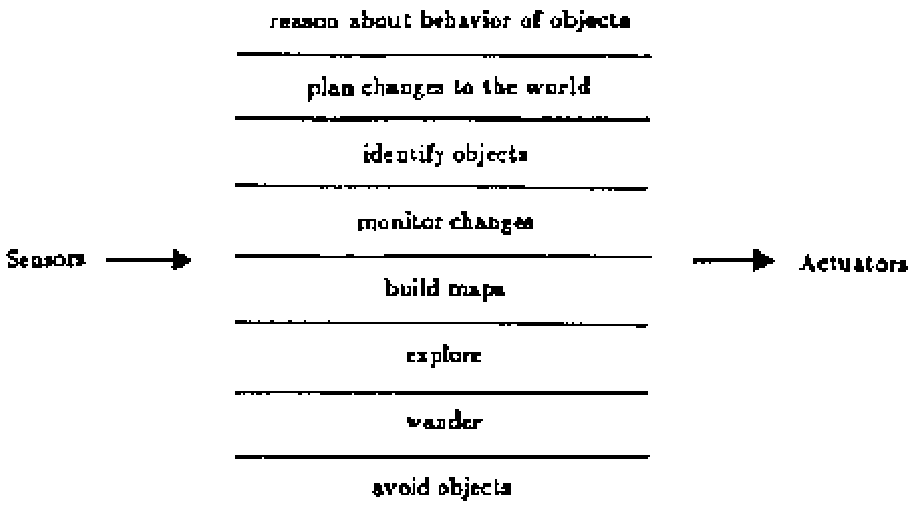
Brooks [1986] has argued against a 'horizontal decomposition' of this sort (figure 1.2a), in which there is a sequential, multi-step process in which all sensor information gets fed to one global modelling unit, then on to one global planning unit, and so on. Instead, he argues for a 'vertical decomposition' (figure 1.2b) in which each of a number of task-achieving processes receive and process their own perceptual sensor information, and then calculate their own independent commands for the actuators (motor effectors).

This SE simulates the perceptual and navigational stimulus inputs to an ASM, and models the effects of the motor commands that are required to execute individual actions, but does not make any assumptions about whether the inputs are computed globally or locally, or whether the motor commands are calculated by global or local entities. That is to say, the SE does not assume either a horizontal or vertical decomposition. There has to be some arbitration between different actions, since otherwise many incompatible actions will be attempted simultaneously and the resultant set of motor commands will be chaotic, but there is no assumption of any other communication between, or global control over, the different systems.

Action selection, along with navigation, is perhaps the least well understood of the four different parts portrayed in figure 1.1b due to the fact that the processes involved are internal and further removed from the outside world than the processes of perception and motor control. Whereas perception can be studied by presenting known stimuli and measuring neural firing rates, and motor control can be studied by exciting neurons and observing motor responses, it is not so easy to apply either of these techniques to the behavioural parts of the brain. Although areas of the brain governing behavioural responses will respond to perceptual stimuli in certain cases, and produce motor responses, the relationships are more complex because the inputs and outputs only interact with the outside world via other interfacing systems [Halliday, 83].



(a)



(b)

Figure 1.2: Two possible decompositions of a mobile robot control system, according to Brooks. **(a)** shows a traditional, horizontal decomposition; **(b)** shows Brooks' preferred vertical decomposition. Taken from [Brooks, 86].

Because of these difficulties in understanding the physiology of AS, theories of how AS takes place in animals have generally existed 'in a vacuum'. Many theories have been proposed, but, because there are no means of validating or invalidating them, little progress has been made. The theories cannot be compared to what occurs in animals because the physiology of the appropriate parts of animals' brains is little understood. The theories have not been able to be implemented as mechanisms whose performance can be observed (to see if they can reproduce the behavioural phenomena produced by animals) because no satisfactory means of doing so has been available.

1.1 Testing Theories of Action Selection Using Robots

One response to this problem of how to test the validity of ASMs has involved the construction of robots which can 'behave' in (usually highly-constrained) artificial environments [e.g. Peabody, 91 or Rosenblatt & Payton, 89]. A particular ASM can be implemented and allowed to control the behaviour of a robot and the efficacy of the resulting choices of actions can then be observed. While this approach does allow for some testing and comparison between ASMs, there are some limitations to it.

It is now generally accepted that the problems of getting a robot to perceive and manipulate its world satisfactorily are enormously difficult [Brooks, 87 & 90]. It used to be the case that AI researchers concentrated on the more 'intellectual' aspects of thought, assuming that perception and motor control were trivial and could easily be solved later. Nowadays, however, perception and motor control are recognised as extremely hard problems. Little progress has been made in getting robots to achieve them effectively, at least when compared to the abilities of higher animals.

Because robots are not very good at interpreting complex environments, they need to be placed in environments where there are few features, each of which is easily distinguishable. Because robots are generally rather bad (compared to animals) at coordinating complex motor responses, they will usually possess only a limited repertoire of possible actions, most of which revolve around locomotion. A typical research robot can move effectively around a flat, fairly uncluttered environment but can do little else in terms of manipulating its environment.

Because of these problems with robots, using them to investigate AS in complex AS problems is difficult. While recognising this, it is also worth pointing out that the use of robots has a methodological advantage in that they are grounded in the real world, whereas simulations are not. This is a benefit because it means that people using robots are less likely to make incorrect assumptions about the inputs and outputs of an ASM. Roboticists *know* that perception of the environment is error-prone, that intended actions are not always accomplished as desired, and so on.

In summary, while the use of robots to investigate ASMs is valid, they are at present limited in their usefulness by the simplicity of the AS problems that they can be presented with. This will become less of a drawback in the future as sensor and effector mechanisms become more advanced. At the current time though it is hard to present a robot with a complex AS problem composed of many different sub-problems which vary in many different ways, and for which there is a large repertoire of possible actions to choose from.

1.2 The Value of Computer Simulations

The human brain has a limited ability to conceptualise complex dynamic processes such as the interactions between a complicated ASM and the rapidly changing internal and external environment provided by the complex simulated environment described here. It is not possible for people to predict accurately how such processes will change with time. Other examples of systems which are too complex to be fully predictable by introspection alone are the patterns of flow of air over various shapes (e.g. aerofoils) and the behaviour of complex industrial installations such as nuclear power stations under varying conditions. In cases such as these there are a few techniques that can be applied to help with the visualisation or calculation of the behaviour of the system: *(i)* mathematical analysis – often intractable for more complex systems, *(ii)* construction and observation of a simplified, smaller-scale physical model – not always feasible or realistic enough, and *(iii)* design and observation of a computer simulation. Mathematical analysis is not feasible for the interactions between a reasonably complex AS problem and a non-trivial action selection mechanism (see section 5.5). A simplified physical model in this case corresponds to a robot.

Although there are drawbacks to the use of a simulation to investigate action selection (see chapter 5), there are also three important benefits:

1. using an SE allows the 'bypassing' of the sensory perception and navigation processes (the relevant information is calculated by the SE and passed to the ASM), and also of the motor control process (the effects of the animal's actions on the environment are calculated by the SE computer program without concern as to how the actions are achieved). The perception, navigation and motor control systems are all part of the SE. They only have to be programmed once, after which it is possible to just slot in different ASMs without adjusting any of the rest of the programming (see figure 1.3).

2. In order to simulate an ASM, it must be implemented, which requires an explicit and complete specification of the mechanism. The process of implementation brings to the fore any vaguenesses, ambiguities and inconsistencies in the description of the mechanism. Theories of action selection can be described at a fairly high level, and may sound perfectly sensible on paper, but turn out in practice to be rather vague and completely unimplementable as mechanisms.

3. The process of simulating an animal environment, and then simulating a mechanism which selects actions in it, can, if the simulation captures the important characteristics of real animal environments, give some idea of how the mechanism could perform in the real world. Most importantly, gross defects in performance, such as completely inappropriate selection of actions in this case, will show up in the simulation.

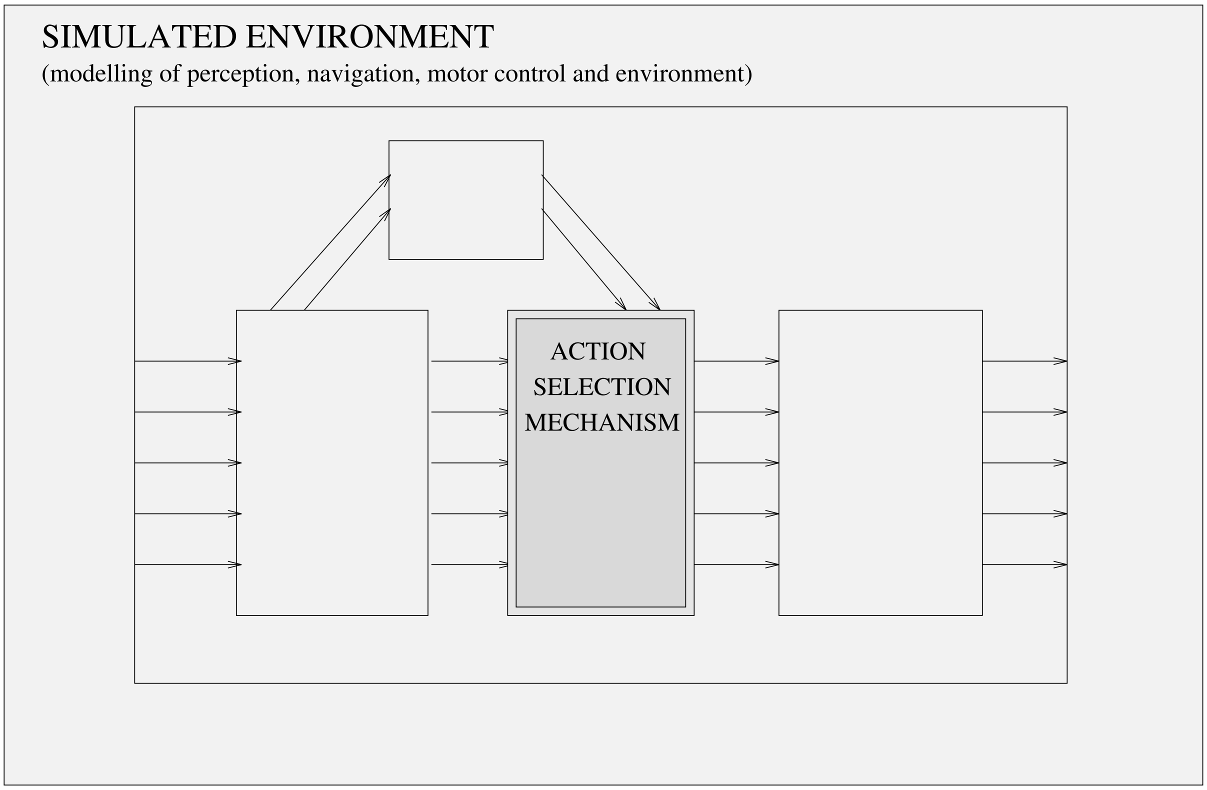


Figure 1.3: Relationship between the simulated environment and the action selection mechanisms. The environment (static and dynamic features, weather, day/ night cycle, etc) and perception, navigation and motor control are all calculated in the simulated environment. Different action selection mechanisms can be slotted in to the simulated environment.

With respect to point 3, it will be argued (section 5.3) that defects or differences in the performances of ASMs in the SE are not important in themselves. They are only useful as indicators of phenomena which require further investigation. The defective performance of an ASM is only useful when it can be shown that it would also occur in the real world, and is a product of a fault in the way the mechanism selects actions rather than a product of some quirk of the programming of the SE. Similarly, differences in performance between mechanisms are only useful when the reasons behind them are understood.

1.3 Other Work on Simulated Environments

Other simulated environments have been written before, but up to now none with the complexity of that presented here. A brief review of these other environments is now given:

1. **RAM** [Taylor *et al*, 89]: this SE was designed for the purpose of examining population behaviour and dynamics, such as the variation in population sizes of predators and prey with different

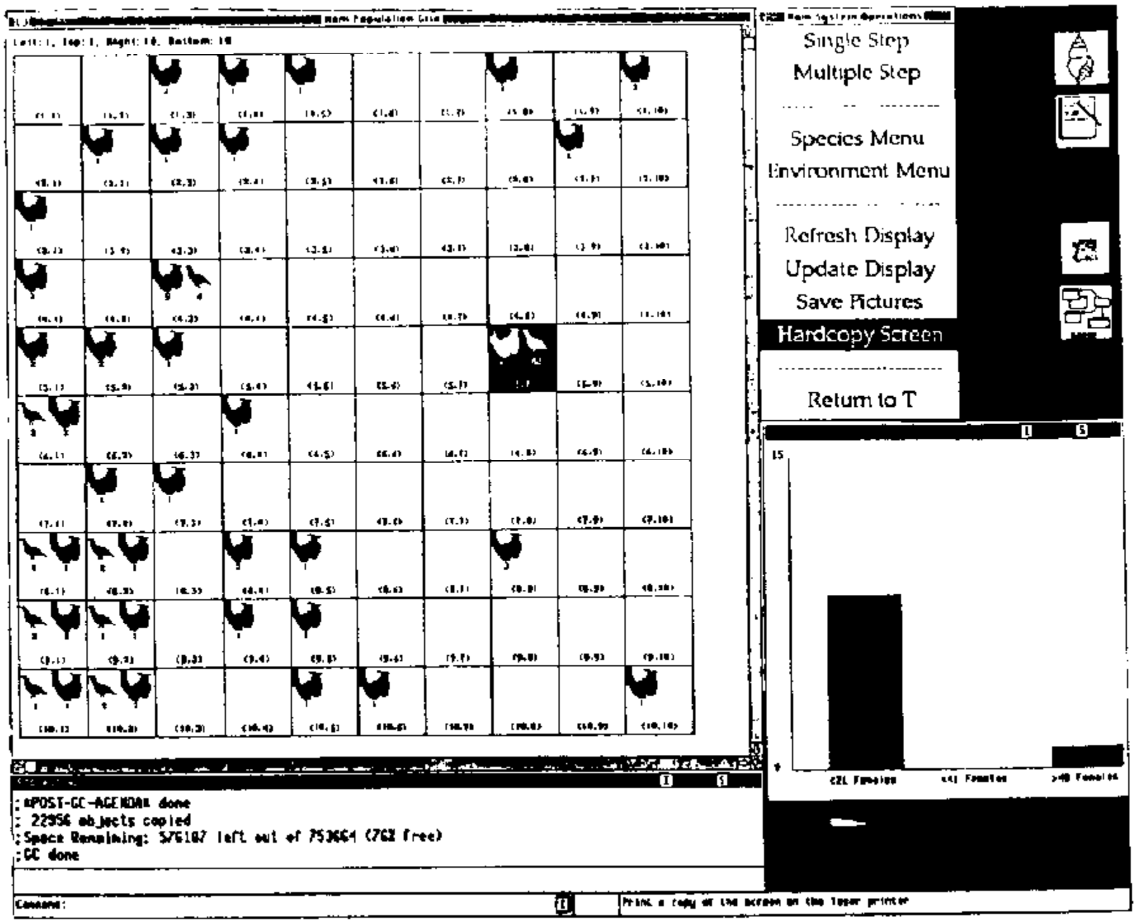


Figure 1.4: RAM. The numbers underneath the male and female birds indicate how many of them are in the square. A 'lek' has formed at (5,8). Taken from [Taylor *et al*, 89].

strategies, and the formation of 'leks' (arenas/ display grounds where males congregate and defend neighbouring patches of territory on which to display to potential mates) by birds such as the sage grouse (see figure 1.4). RAM was constructed with the intent of examining the population effects of various AS strategies, rather than evaluating the efficacy of an ASM for an individual animal. There are animals in RAM and it is possible to add other features if desired. The squares of the grid can have different attributes such as 'food value', or 'altitude', and there are global environment variables such as 'time of day'. The perception of the animals is local and noise-free. The only possible actions are movements.

2. **WOODS7** [Wilson, 85]: Wilson studied a simple method for associating perceptual situations and actions that could easily be optimised using classifier systems. The WOODS7 environment (see figure 1.5) does not provide a complex and interesting test of ASMs in general. There are only two types of feature (tree and food) in the environment and perception of these is perfect if they lie in neighbouring squares. There are 8 possible actions — a movement in each of 8