A Proposal for a Study of Motive Processing

Luc P. Beaudoin
and
Aaron Sloman

School of Computer Science
and
Centre for Cognitive Science
University of Birmingham

October 1991

Note

This paper was mostly written by the first author, although it is based on and develops ideas of the second author. The nursemaid scenario was first described by the second author (Sloman, 1986). The first author is in the process of implementing the model described in the paper.

Thanks to Aluizio Arujo, Peter Greenfield, Inman Harvey, Tim Read, Edmund Shing, Sharon Wood, and Shiu Yuen for reading and commenting on drafts. The first author was supported by the Commonwealth Scholarship Commission in the United Kingdom, and FCAR of Quebec.

A Proposal for a Study of Motive Processing

1. Introduction

In this paper we discuss some of the essential features and context of human motive processing, and we characterize some of the state transitions of motives. We then describe in detail a domain for designing an agent exhibiting some of these features. Recent related work is briefly reviewed to demonstrate the need for extending theories to account for the complexities of motive processing described here.

2. Context and Features of Motivation

The fundamental features and context of motive processing postulated by Simon (1967) and Sloman (1978 ch 6, 1985, 1987; Sloman and Croucher, 1981) are not difficult to understand. They are nonetheless tied to important design constraints and it is therefore important (and difficult) to keep them in mind.

1. Human beings have multiple independent motivators. Their behaviour cannot, therefore, be explained in terms of a top level goal from which all other goals are derived. In Sloman's (1987) terminology people have non-derivative motivators, where by motivators he means "mechanisms and representations that tend to produce or modify or select between actions in the light of beliefs", and which include motive generators and motive comparators.

2. People have limited resources to deal with these motives. The limits are both internal and external. An instance of internal limits is that a person can't simultaneously plan and assess a large set of motives. This is commonly recognized in cognitive psychology. As an example of the external limits, she can't be in two places at the same time. (Technology may help to decrease these limits, but they can't be removed altogether.) These limits imply a need for (a) deciding which among incompatible motives to adopt as intentions and (b) "deliberation scheduling", i.e., deciding in which order the goals will be assessed and planned for, and "meta-planning", i.e., deciding in which order activities will be performed.

3. Human internal and external activity is interruptable. Explaining how this can happen will require accounting for the fact that people can't function properly whilst being continually interrupted; yet sometimes they allow their attention to be diverted, and sometimes their attention is diverted whether they allow it or not. Thus there are internal criteria for deciding whether or not interruption is justified. Interrupting an activity may lead to its modification, temporary dispension, or complete abandonment.

4. There are various criteria for halting an activity. I.e., such processes don't necessarily terminate with a perfect solution. Sometimes 'satisficing' criteria must be accepted (Simon, 1967), i.e., criteria which are 'good enough'. Different kinds of incomplete plans are discussed below. (See the section entitled "The Processing of Motives".)

5. Actions are often chosen so as to serve multiple purposes. This helps to compensate for limited resources (including time).

6. Monitoring may reveal various facts related to current (or dormant) plans and motives, which may need to be re-assessed. Dealing with this may require an action-motive index. (Compare Sloman, 1978 ch 6). For example, an agent may discover that a fact implies that her goal, G, is no longer necessary. (E.g., the goal

of saving a baby who has just died.) She should therefore consider deleting this goal and its actions from her intentions. To do this she has to know which actions serve that end (she could examine something like an action-motive index). She may need to re-assess some of her other goals as well. For some of them may have been dependent upon the disactivated motive or (the effects of) its actions. This is more complicated in the case of a system with multiple independent motivators than it is for those with one top level goal only, although some of the same complexity is needed to deal with interactions between subgoals of one goal, especially interactions due to unexpected events.

Monitoring may also reveal that an action didn't have its intended effect, that a goal has already been achieved, or that some new opportunity or danger is present.

Relevant facts about the context of motivation include the following:

7. An agent's knowledge of the world is always incomplete (and hence her ability to predict is limited).

8. Events can happen which are outside of the agent's control, and they sometimes happen too fast for rational consideration of implications, alternative courses of action, etc.

As discussed by Sloman (1978, 1985), a system exhibiting these features requires coarse grained parallelism: An ability to perform concurrent monitoring of the internal and external environment, whilst performing physical actions and evaluating information produced by such monitoring (including the elicitation of memories, etc.) and retrieving or creating plans. The machinery for dealing with ongoing plans has to deal with the fact that plans cannot be completely specified. The monitoring will therefore not only serve to detect opportunities and dangers relating to motives which aren't currently being actively pursued, but also to guide action which is only vaguely specified by a plan. (Sloman 1978, Agre and Chapman 1990).

3. The Processing of Motives

A system exhibiting these features will allow motives to pass through various states involving various processes, a partial specification of which is proposed here.

Motivators can have as their object an action, an outcome of an action, and/or a consequence of an outcome. (See Heckhausen and Kuhl, 1985.) Motivators usually involve the detection of a problem or opportunity, and/or the generation of a goal. For brevity, let us mostly speak in terms of goals. An initial step in motive processing is the production of a goal. According to Sloman (1987) a goal may or may not even be considered. This is because considering goals requires resources which may be currently needed for other purposes. The tendency of a goal to be considered is a function of its "insistence", a heuristic measure of its importance and urgency, which is assumed to be computed very quickly. Sloman (Sloman, 1987; Sloman and Humphreys, 1990) assumes that fallible variable threshold motive filtering mechanism may serve to block out less insistent motives and protect limited resources from diversion when important and urgent goals are being pursued.

3.1 The assessment of motives.

Goals may need to be assessed in terms of importance, urgency, opportunity, and costs. This assessment is usually relative to other goals with which the current goal may be compared.

Determining the importance of a goal requires some rules for ordering goals in terms of desirability. In this sense, goals may have derivative and/or non-derivative importance. A goal has derivative importance by virtue of it being recognized as being a means to an end which itself is important. This importance may be a function of the importance of the supergoal, and of whether there are other alternative goals which could achieve the supergoal. (This implies that a goal can be a subgoal to a very important goal, whilst not being very important if other means can be used to attain its supergoal.) Non-derivative importance is related to things that matter intrinsically to the agent, and not simply because they subserve some other goal or purpose. Non-derivative importance can be related to aspects of the agent's character and personality such as aesthetic or moral preferences, attitudes to other people, political ideals, etc.

Urgency relates time to measures of costs or benefits of the goals. Terminal urgency refers to how little time there is before it is too late to satisfy the motive at all. E.g., the goal of catching the 5 o'clock train may have a terminal urgency of fifteen minutes at a quarter to five. Some goals can be more or less well achieved (in terms of cost and/or benefit), and they may have a gradual urgency, which relates time to the degree to which they can be attained (or the degree of damage made, etc.) E.g., a student may lose 5 points for each day that a paper is late. In general urgency will not increase or decrease monotonically. E.g., some goals such as planting seeds successfully are related to cyclic seasonal changes. A notion related to urgency is the likelihood that something goes wrong or well as a function of time. E.g., the longer you leave your bicycle unlocked outside, the more likely it is to get stolen (but you won't worry about it being more or less stolen as a function of time).

Goals may need to be assessed in terms of the internal and external resources which would be required to attain them. This is their cost. E.g., one may reject, or prefer an alternative to (or otherwise think or act differentially towards) a goal if it is assumed that a lot of time would be required to construct a plan for it, or if it can be achieved only if many other goals are abandonned. (Prolog fails on this count, since it blindly tries to satisfy goals on the basis of their order. It cannot decide not to try to prove a goal which would be difficult to prove, unless it is augmented by some kind of meta-level control mechanisms. Hayes, 1985 makes a similar point.) One may also recognize that a goal would require more resources than are currently available (or worth using for it). This may be recognized before or while planning for a goal.

Opportunities are situations in which the cost of attaining a goal (or more than one goal[1]) is reduced. An opportune goal may temporarily take precedence over other current candidates for deliberation or action. Conversely, a goal may be rejected or postponed because it is "unopportune", i.e., because in the present situation it is too costly, or because some other goal has had its cost reduced.

The assessment of goals along these dimensions will have to be computed by rule of thumb measures when information is not available or when time is short. (See Sloman, 1985.) These rules are not to be confused with the reflex prioritisation of motives on the basis of their insistence (a different kind of rule of thumb). It may

[1]E.g., when one can "kill two birds with one stone".

suffice to recognize that a goal is very opportune, or very urgent, for it to be considered before another. Furthermore, the depth of the assessment varies according to the situation. Assessment along these dimensions may occur either before or after a goal becomes an intention and has been planned for. And the results of assessment may independently affect deliberation scheduling, planning and metaplanning. It is also possible for insistent motives to gain attention by a planner even without being assessed.

It is often assumed that any computationally efficacious dimension along which motives vary and which is relevant to their assessment is quantitative, implying the use of total orderings with interval scales (e.g., Bindra, 1978; Heckhausen and Kuhl, 1985; Kuhl, 1982; Ortony, 1988). But this is not an accurate assumption for two related reasons. The first one is that qualitative (non-quantitative) information can be acted upon by a process determining the course of a motive. For example, an administrator on a committee can have the norm that he will not tolerate discrimination on irrelevant grounds. If he recognizes that a proposal to this committee implies an unjustified discrimination he may oppose or try to modify it. Of course, one can argue that his norm in fact has a numeric value associated to it--perhaps it is infinite or almost so. But note that it need not be numeric. Here's another example. When computing the gradual urgency of a motive, one can anticipate the things which will go wrong the longer it is not satisfied. This is qualitative information which is part of the assessment of the motive. More generally, factual information may be produced, stored, and used in the form of reasons for an action. A special yet common use of factual information is in the control of behaviour. E.g., when a person decides to do X unless Y.

The second reason is implied by the first but also stands on its own. It is that often only a partial ordering of motives is possible, and even that may be highly context sensitive. (Compare McCulloch, 1945.) The partial ordering is sometimes due to the fact that two goals measure up to two independent (non-commensurable) higher order motivators. (It can also be due to different dimensions playing a role in assessment, e.g., urgency and importance.) E.g., one goal may measure up well with regards to your disposition, D1, to maintain your physical health, whereas another measures up well with regard to your disposition, D2, to maintain good social relations. Even if we were to assume that any goal which has implications for D1 may have its importance quantitatively evaluated along D1, to yield a value d1, and any goal which is relevant to D2 may have its importance quantitatively evaluated along D2, to yield a value d2, d1 and d2 cannot be compared along a common interval scale of goodness. If your goals could be assessed in terms of a common interval scale of goodness, their ordering would be transitive, which is contrary to the facts.

3.2 Scheduling deliberation.

Assessment, planning, and metaplanning are deliberation processes. Deliberation scheduling is needed because deliberation can itself divert precious resources from important actions. It involves deciding which deliberation process will run now, and on what arguments.

3.3 Planning

Planning may involve retrieving an existing (more or less detailed) plan or constructing a new one. Constructing a new plan may involve producing a very elaborate specification of the actions to be performed, or producing a specification of the first few actions to be performed, or a solution to some of the most difficult problems to be solved. The processes which decide on courses of action are interruptable. In this case a good-enough solution may be proposed and/or planning may be postponed.

It is noteworthy that people are often able to produce plans the quality of which is a function of the time spent planning. Dean and Boddy (1988) refer to the procedures underlying this particular ability as "anytime algorithms". An agent of this sort needs flexible plan-execution mechanisms able to deal with more or less complete and more or less good plans. These procedures contrast with traditional computer science methods and most AI planning mechanisms which usually produce a solution only if they are allowed to terminate.

A planning process may fail altogether, in which case completely improvised activity or rejection or postponement of the goal may ensue. In any case, there will always be some degree of decision making at run time, and/or reliance on procedures which determine action on the basis of perceptual information and previously established habits (patterns of response to situations). Often goals which have been rejected may be posted again, and go through more or less different state transitions.

These are but some of the processes involving motives. More details can be found in Heckhausen and Kuhl (1985), and Sloman (1987; Sloman and Humphreys, 1990).

4. <u>The Robot Nursemaid Domain</u>

Taking a design stance to the study of motive processing, we propose to design a system which exhibits some of the key features of motive processing. We believe that adopting the design stance will help us to test, elaborate, clarify, and debug our understanding of motive processing, and help us to more rigourously assess our own and other theories. See Boden (1989) and Sloman (1978) for justifications of computational approaches to cognitive science.

The domain in question is simplified compared to the real world. Since this will be the first attempt to incorporate such a variety of relevant processes for the study of motive processing, this simplicity is warranted. (See Pfeifer, 1988 for a review of AI studies of emotions.) Once this program has been designed, implemented and sufficiently analyzed and criticized, it will be possible to envisage a more comprehensive scenario, in the light of a refined and extended theory, to repeat the cycle again.

However, although they are simple in comparison with those of real human life, the requirements of this domain go beyond a relevant portion of those of the state of the art of Artificial Intelligence. In particular the domain incorporates the requirements for coarse-grained parallelism, multiple independent sources of motivation, attention filtering, and multi-stage motive processing as sketched above. It thereby provides a stimulating challenge for theorists of motive processing.

In this section we describe the domain and requirements for an agent exhibiting some of these features.

4.1 Ontology of the domain

The domain involves an intelligent "nursemaid" looking after a collection of robot babies who arrive in a nursery in which they can move about. They need to be kept out of danger by robot nannies which have very limited information processing abilities. Nannies respond to problems they detect by sending signals to the nursemaid, an immobile machine outside of the nursery, who tends to respond to such motive signals by sending out instructions to nannies. There is also an administrator which merely informs the nursemaid of incoming babies some time before their arrival.

The role of the nursemaid is to assure the well-being of as many babies as possible. Other things being equal, she[2] prefers older, faster, and healthier babies. Although the nursemaid is more intelligent than the nannies, she is assumed to have a limited ability to process information. She will therefore not be able to respond to all of the motive signals that she receives or produces. Since the goal of this project is to model *one*, and not a collection, of intelligent agents, the nannies are kept relatively simple. As will be apparent, the nannies can be seen as part of the nursemaid's perceptual and motor systems.

The nursery is divided into play areas which are connected by gates and separated by fences and by opaque walls (See Figure 1). Two areas on the same vertical axis are said to be in the same "section" (e.g., area 1 and area 2 are in the same section, but not area 1 and area 3). The nursery is bounded by hazardous ditches. It contains a battery recharge point (for the robots) in Area 4 (i.e., A4), an infirmary in A2, an arrival point in A1, and a departure point in A1.

Insert Figure 1 about here

Each robot has a number of features, some of which may change. The features which every robot has are: a horizontal position, a goal position, a current speed, a maximum speed, an identification code, a vertical position (either picked-up or not), a battery charge, and a flag indicating whether it is dead or alive. In addition there are features pertaining only to babies: age, number of injuries sustained, illnesses, and tendency to strike other babies (this may be zero). And nannies have peculiar fields too: a plan (an ordered collection of instructions), an indication of the robots they are carrying, and an indication of whether they are strong or not. Strong nannies can carry two babies at a time, or one other nanny at a time. The potential speed of a baby is a function of its age.

Babies are in potentially fatal danger of falling into ditches, having their batteries run out, sustaining injuries from other babies who react to overcrowding by developing a tendency to hit one another (babies therefore shouldn't all be kept in one area), and falling ill. Babies can be transported by nannies. Babies who have developed a thug like tendency need to be isolated so as to lose that tendency. Both babies and nannies periodically need to have their batteries recharged at the recharge point. Babies can only sustain a certain number of injuries. They can recover from their injuries and illnesses by spending time in the infirmary. They should therefore be transported there if they are injured. There must be a nanny in the infirmary for the curative effect to take place. The infirmary can only reasonably accommodate three babies at a time. The speed of recovery decreases

[2]The feminine rather than the neutral gender is used to refer to the nursemaid for the sake of facilitating pronoun reference.

as a function of the number of babies above three in the infirmary. A dead baby must be dismissed from the nursery. The presence of a corpse in the same section increases the likelihood that babies get sick. Eventually, the presence of a corpse can take its toll on babies in other areas too.

Most nannies can only carry one baby at a time; strong ones can carry two. The nannies can see through fences but not through walls. They can figure out how to get from one place to another and generate a series of instructions to move themselves from one area to another. Nannies can only change sections on demand of the nursemaid. If a nanny is asked by the nursemaid to fetch a baby which is not in her section, she will report a failure to the nursemaid. Nannies cannot communicate between each other. Nannies can, however, inspect the characteristics of babies. For example, they can detect their charge and their number of injuries. Apart from sending motive signals they can also send perceptual information to the nursemaid. E.g., they can report the positions and identity of every robot in their visual field. The kinds of signals which can be sent to the nursemaid are discussed below.

Nannies are also in danger of having their batteries run out. Like those of babies, their batteries also run down as a function of the distance they travel and the speed at which they travel. Carrying robots wears down batteries faster. When a nanny's charge is very low, it will need to be carried to the recharge point. It takes two nannies (or one strong nanny) to transport another nanny. No nanny can transport both a nanny and another robot at the same time.

The nursemaid receives information in the form of motive signals and factual signals from the nannies. (Motive signals also have a factual component, and factual information can also modulate action. This is discussed below.) She does not receive signals from babies. She processes some of the information and decides upon courses of action to keep babies out of danger. She sends instructions out to nannies. The nannies use these instructions to come up with more detailed ones. E.g., the nursemaid can order a nanny to go to A6, without specifying which route to take. The nanny would later give itself the instruction in the form of a sequence of commands specifying the route.

4.2 Signals in the nursemaid domain

The nannies can send motive signals or perceptual feedback signals to the nursemaid. Motive signals are messages sent to the nursemaid, which signify that a problem has been detected or that an action has been successful. Perceptual feedback signals contain facts about (including relations between) robots. They can be expressed using the same syntax.

4.2.1 Motive signals

Motive signals contain information about the sender and the situation eliciting the signals. They can be expressed in a Prolog like syntax. They have the form:

signal(Intensity, Sender, Descriptor)

The <u>intensity</u> of the signal is computed by the sender according to simple rules. This will be the basis for insistence. The <u>sender</u> of a signal is either a nanny, the administrator or the nursemaid (i.e., she can send signals to herself). (Some problems can't be detected by the nannies and hence must be inferred by the nursemaid. E.g., no nanny can have the knowledge that a baby is in an unsupervised section.) The <u>descriptor</u> is an expression describing the situation. It has the form:

predicate(argument1, ..., argumentn).

There is therefore at least one argument, and sometimes only one. Some descriptors describe relations between objects.

Here follows a list and description of the motive signals which may arise. In the following description the specification of the proper response is ideal. Sometimes the nursemaid will not be able to perform the ideal response, but she may propose partial solutions. Alternative solutions are possible.

From a nanny: thug(Baby, Position). This means that Baby is a thug, and that it is at Position. (The position includes the area.) This signal is only emitted if Baby is not in an area by itself. The intensity of this signal is a function of the number of babies in the area. The action tendency is to isolate the thug.

From a nanny: lowcharge(Robot, Position, Charge). Robot (either a nanny or a baby) has a charge equal to Charge which is below its threshold. The robot may still be alive and mobile. The intensity of this signal is an inverse function of its charge. The action tendency is to transport or send Robot to the recharge point and have it recharged there. A nanny with a low charge may be told to transport itself, provided that its charge isn't completely exhausted.

From a nanny: old(Baby, Position, Age). Baby is over the threshold age and can therefore be dismissed. This merely involves bringing Baby to the departure point in A1. The intensity of the signal is a function of the baby's age.

From a nanny: dead(Robot, Position). Robot is dead. The intensity of this signal will be fixed and high. Robot must be brought to the departure point and dismissed.

From a nanny: injured(Baby, Position). Baby has been injured. This signal is sent if a baby is injured and not in the infirmary. The intensity of this signal is a function of the number of injuries and the number of babies in the area. The nursemaid's action tendency is to send the baby to the infirmary whilst assuring that a nanny is in that area. If the infirmary is full, then the nursemaid has to decide whether or not to take out a baby which is already there.

From a nanny: closetoditch(Baby, Position, Distance). Baby is at Position, which is Distance from the ditch. (The nursemaid can infer to which ditch the baby is close from its position.) This signal is sent by a nanny if Baby is within a threshold distance from a ditch provided Baby is not already picked up. The intensity of the signal is inversely proportional to the distance between Baby and the ditch. The action tendency is to order a nanny to pick up the baby and transport it to the centre of the area.

From a nanny: failed(Action). This signal is sent by a nanny when one of her behaviours has failed (because a condition wasn't met). For example, if a nanny was asked to pick up a baby and the baby wasn't in its area (a precondition for picking up a baby) then she sends this signal.

From a nanny: successful(action). This message is sent when a series of commands has been successfully executed. The nursemaid thereby knows that this nanny is free to do other things.

From the nursemaid: overpopulated(Area, Population). The nursemaid finds that there are more babies in Area than is safe. The intensity of the signal is a function of the number of babies in Area. The nursemaid responds by trying to distribute the population more evenly.

From the nursemaid: missing(Baby). This signal is emitted by the nursemaid when a baby which is supposed to be in the nursery is no longer contained in the nursemaid's model of the nursery. Baby has probably wandered off to an unsupervised area. The intensity of the signal is a function of the time since the nursemaid has last received information about Baby and of the age of the baby. She responds by trying to find the baby, which involves requesting perceptual information from nannies. If there is an unsupervised section, then she may send a nanny there.

From the administrator: incoming(Integer, Time). There are Integer babies about to arrive in A1 of the nursery in about Time time units. The intensity of this signal is a function of the number of babies about to arrive. Depending on the number of babies already in A1 and the number of babies about to arrive, A1 should be cleared, unless all areas are filled above capacity.

4.2.2 Perceptual feedback signals

Nannies spontaneously send basic perceptual information to update the world model. This contains information about the position and identification of babies in an area. More detailed information can be requested by the nursemaid, as discussed in the next section.

4.3. Nursemaid commands

There is a number of commands which she the nursemaid can send to nannies to achieve her ends. She can combine the commands by sending a series of commands. Her commands are of the form

command(Argument1, ..., Argumentn)

where "command" is a command name, and every argument is a Prolog expression. Argument1 specifies the nanny to which the command is sent. Commands are inserted in a nanny's goals queue. The nursemaid can also delete commands from a nanny's queue.

Here is a list of the commands available to the nursemaid.

go_to(Nanny, Position|Area)

This tells Nanny to move to Position or Area. Nanny figures out a route to get there.

superintend(Nanny, Area)

This tells Nanny to wander around Area and look out for problems. If Area is not in Nanny's current area, it instructs itself to go_to(Nanny, Area). If problems are encountered, they are reported. Note that this is every nanny's default behaviour, i.e., the one in which it engages if it has nothing else to do.

fetch(Nanny, Robot)

This tells Nanny to go pick up Robot. If Robot is not in Nanny's section, then Nanny reports an error to the nursemaid. Otherwise, Nanny instructs itself to go_to(position(Robot)) and then it tries to pick Robot up. If Robot is another nanny, then unless Nanny is strong, it must wait for another nanny's presence before picking it up, for two non-strong nannies are required to transport another nanny.

deposit(Nanny, Robot)

This tells Nanny to deposit Robot, which she must already be carrying. If she is not carrying robot she reports an error to the nursemaid.

take(Nanny, Robot, Position|Area)

This tells Nanny to take robot to Position or Area. If Nanny is not already carrying Robot, it instructs itself to fetch(Robot) and then to go_to(Area). If Robot is not in the same section as Nanny, then Nanny should report an error.

recharge(Nanny, Robot)

This tells Nanny to recharge Robot. If it is not next to the recharge point with Robot, it first instructs itself to take(Nanny, Robot, recharge_point).

dismiss(Nanny, Baby)

If Baby is in a different section, Nanny reports an error. Otherwise, Nanny must fetch the baby (if it is not already carrying it), and (if it is not already there) go_to(Nanny, departure_point). Then Nanny can dismiss Baby. The effect of dismissing Baby is that Baby is no longer in the nursery.

request(Nanny, ID, Fields)

This asks Nanny to return information about Fields of the robots in its section. Fields is either "all" or a collection of robot fields (e.g., position, age). If Fields is "all" then the nanny returns all of the information associated with each robot. (As in the first feedback signal exemplified above.) Or else it returns information for each of the fields.

4.3.1 The ability of nannies to expand commands.

When the nursemaid is rushed she can delegate some simple decision making to nannies. The description above already implies that nannies are capable of obeying instructions that require complex actions. Another facility which will is presently envisaged involves letting a nanny decide to which object it ought to apply a command. E.g., in the scenario to be described below, rather than giving the following commands to nanny D

fetch Bl	(i.e., fetch baby "l")
go_to A1	(i.e., go to area 1)
deposit Bl	(i.e., deposit baby "l")

the nursemaid could tell nanny D to pick up the baby which is closest to the gate between Area 2 and Area 1. The nannies could be designed to respond to commands such as:

take(nearest(baby, (A2, 50, 0)), A1)

In this example, the nanny has to find the baby in A2, which is the closest to the position (A2 50 0).

We might allow for an even greater ability, where the nannies understand commands like "reduce density A1 to A2", and "take n babies from A1 to A3". This will depend on whether this can be achieved without requiring intelligence. (Nannies are supposed to be kept simple.)

The problem with letting the nannies fill in major portions of plans is that they do not have as much knowledge as the nursemaid, and hence might come up with less efficient plans. For example, whereas the nursemaid can figure out a route through less populated areas, the nanny does not know about population densities. Hence if a nanny is itself to decide on the route to take with a baby, she may go through a high density area, with its associated risk, even if a less hazardous route were possible. But this is a price which is sometimes worth paying to free the nursemaid's thinking processes. For in any case the nursemaid will often not have the time to discover or use the best solution.

4.4 A Scenario in the nursemaid domain

Here's a little scenario to illustrate the domain. Suppose that the positions of the robots are the same as those in Figure 2. Note that the real situation (partially depicted in Figure 2) differs from the nursemaid's model of the situation. E.g., the nursemaid may not know the position about Bc. The real situation also includes information about each baby's fields (top speed, current speed, number of injuries, etc.)

Insert Figure 2 about here

Suppose the nursemaid receives motive signals from two different sources. (See Figure 3.) One from the nursemaid to herself to the effect that the population of babies in A2 is superior to the threshold (T=3) above which babies may get violent. The other is from Nc to the effect that Bi has a low battery charge.

Insert Figure 3 about here

The nursemaid considers the "overpopulated" signal first and, in order to decrease the population of babies in A2, she tells Nd to bring Bl into A1. Then she starts to consider the signal about Bi; but as she is doing so, she is interrupted because as she is concurrently monitoring the environment she suddenly infers that Bc is nowhere to be seen. (This is an illustration of her coarse grained parallelism.) Bc is an old baby so this signal is intense. (See Figure 4)

Insert Figure 4 about here

She responds to this signal by asking every nanny to send a view of its area back to her. She programs herself to send a nanny out to look for Bc in a few time units if it is still missing. Meanwhile, she can resume her enterprise of trying to figure out what to do about Bi. She decides to tell Nc to bring Bi to the recharge point, which is in A4, and, in order to avoid a population overflow in A4, she tells Nb to bring a baby out of A4 into A3. (During this time, she is not paying attention to other problems, like the fact that Ba's charge is low too.) Still no news of Bc. She therefore sends Na to Section 4, and tells it to bring Bc back into A4, because she infers that Bc may be there, since that is the only section which she hasn't monitored.

4.5 A specification of the nursemaid's abilities

Here we describe at greater length the abilities which the nursemaid should have. These requirements were selected on the basis of the analysis of the features of motive processing, presented above.

The nursemaid should have the ability to ignore (i.e., filter out) less insistent motive signals when she is dealing with an urgent and important task.

She should be able to respond to some of the signals by producing goals.

She should be able to assess signals in terms of importance, urgency, opportunity, planning time required and cost. The depth of this assessment will be a function of time available and number of motives to be handled. (Sometimes it will not be possible to assess the situation.) The nursemaid should be endowed with rule of thumb measures for performing this assessment. (These are different measures from the determination of insistence, which determine whether or not a signal will actually be considered.) A more sophisticated model would allow the nursemaid to <u>learn</u> the rules.

It may not be reasonable to assume that the nursemaid is capable of planning and performing complex evaluations of goals at the same time. Hence, planning and assessment will not usually be permitted to occur concurrently.

The nursemaid should be able to decide the order in which she will plan and assess the current goals. I.e., she should be capable of deliberation scheduling. Her deliberation scheduling should be based on the assessment of current motives. As supposed by Dean and Boddy (1988), deliberation scheduling should be fairly quick.

The nursemaid's deliberation scheduling should be able to decide to only partially evaluate a signal. For example, if a signal is highly urgent and important, then it may not be necessary to compute the planning time required (this would take time better spent actually planning). Note that this implies that deliberation scheduling should go on concurrently with goal evaluation. (For a partial result of an evaluation process may affect the performance of a deliberation scheduling process.)

Deliberation scheduling should be capable of interrupting current planning processes, if new goals are sufficiently more urgent, opportune, or important. (See below.)

It should be possible to suspend a planning process which has been interrupted, and resume it at a later time. The ability to resume a planning process avoids having to start planning from scratch: this saves time, but requires suitable mechanisms for recording details of suspended processes. It will not always be useful, however, to resume a planning process. For it will happen that some of the assumptions on which a planning process was based will have changed. The nursemaid should therefore be able to decide whether or not to resume a planning process. This is related to the frame problem. Among the things that can follow from a change in the world is a change in the relevance, urgency, cost or importance of a goal.

The nursemaid should be able to plan for a goal. This will involve constructing or retrieving solutions to problems. It can also require obtaining more information. Planning will also involve selecting resources (nannies) to perform the task. It will often happen that some of the resources required to solve a problem are currently being used. In this case, the nursemaid will have to decide whether to interrupt the nannies, or whether to postpone the goals. This will involve comparing the motives and plans involved. If the nursemaid is to be able to decide to postpone a goal until suitable conditions are met, she ought to be able to recognize when those conditions are met. (For example, she could set up temporary goal generators. See Kuhl and Kazen-Saad (1988), on context sensitive intentions.) In humans this ability is fallible, and it will be in the nursemaid.

Beside allowing for new goals which have been assessed to interrupt planning, there are two other noteworthy conditions under which planning may need to be interrupted. First, the nursemaid should be able to infer whether new information implies that the problem for which she is currently planning is no longer an issue. For example, if she is planning to bring a baby away from a ditch, and the baby is no longer close to it, then she can kill this planning process. Second, the deliberation scheduling process should allow some very intense signals which have short-circuited normal processing to quickly interrupt planning. E.g., if a nanny suddenly discovers that a baby is right by the edge of a ditch, then she will send a highly intense signal to the nursemaid. This signal may have the ability to interrupt current planning. That is, some reflex consideration of signals should be possible. However, the threshold of such signals should vary as a function of the urgency and importance of the goal being planned for. The filtering mechanism and the process controlling its thresholds will by design be quick and therefore fallible.

If the nursemaid's planning for a goal were sensitive to an estimate of the time available for planning for that particular goal, this would decrease the occurrence of interrupts. I.e., the nursemaid could be able to choose her planning strategies so as to minimize the likelihood of being interrupted. This may also permit for more rapid planning than usual when the opportunity or problem is very urgent.

The nursemaid when interrupted, or when the urgency is extreme, should be able to output an incomplete solution (e.g., to churn out a few commands to get a nanny moving).

The specification does not commit one severely to a particular form of planning (though there are some constraints, e.g., on time). For example, the nursemaid's planning may to a large extent depend on retrieving existing solutions.

The nursemaid will have to keep track of the status of motives she has processed. This will help her deal with the fact that motive signals may be generated many times and by different agents, even if a solution is being developed or proposed or has been adopted or rejected. This need not imply that the nursemaid can't reconsider goals which she has already rejected (for example, the situation may have changed).

5. The Relevance of the Nursemaid Domain and Design for the Study of Motivation

In the first part of this paper, important features of motive processing were discussed. It should be clear that the design specification was moulded to require these features. In this section we express some of the nursemaid's intended features in terms of the ones which were posited above as important for human motive processing. That is, the features described below are shared in some respects by people.

1. The nursemaid has multiple independent motivators. A disposition to take care of each baby is a motivator. It causes the generation of non-derivative goals. There are many babies, therefore many motivators. And for the same baby many (sometimes conflicting) motives can be generated. E.g., the goal to recharge a baby and the goal to bring it to the nursery. And there are motivators such that the nursemaid tries to keep the nannies from dying.

2. There are also meta-motivators influencing the kind and course of the deliberation. E.g., to solve urgent problems quickly, to avoid high cost plans.

3. Processing of motives (including planning) occurs while things are happening in the world. Therefore, planning is not assumed to be instantaneous. This implies the need for extra interactions between planning and perception.

4. Processing is modular and parallel at a coarse grained level. Monitoring and evaluation of information go on in parallel with deliberation scheduling and planning and reasoning about known facts. (This will raise a lot of questions regarding the coordination of processes.)

5. Planning and other processing of motives are interruptable. So are actions.

6. She can do multidimensional evaluation of goals.

7. Given a problem, she is able to propose a solution the quality of which is a function of time available for planning.

8. The nursemaid has to process her motives with incomplete knowledge of the world. Lack of knowledge is due to: (1) inability to know every thing about the present state of the environment; (2) an inability to perfectly predict future states, including the effects of her actions. (She doesn't know for example what individual babies are going to do next.)

9. The macro-structure of her abilities is hierarchical, with nannies being her sensors and effectors, and her having an attention filtering mechanisms to protect the high level processes of assessment and planning.

10. She will display a variety of paths from goal generation to action, some very direct, some very indirect.

6. Related Work

Artificial Intelligence research on Planning and recent trends in the cognitive psychology are relevant to the study of motive processing.

6.1 AI research on planning and activity

"Planning" is a problem with which AI researchers have been dealing for many years. (See Chapman, 1985 and Steel, 1987 for recent overviews.) Some of the literature on this problem is of concern to us. Planning can be grossly characterized as the mental processes required to achieve ends. More precisely, planning refers to abilities and processes used to generate a collection of instructions which are meant to transform the present state of the world to a desired state (or an approximation of the desired state), or prevent the occurrence of an unwanted state. The instructions may be fairly abstract, and instantiatable in a wide variety of ways. Many a planning system has an "executor" which uses the instructions it receives from the planner as a basis for computing more detailed commands which it sends to efferent links with the world. Some systems are capable of detecting problems in the execution of a plan, and causing partial or complete replanning (e.g., Wilkins, 1985).

Since 1986 or so many AI theorists of activity have expressed dissatisfaction with the methods and assumptions of most planning research. (E.g., Agre and Chapman, 1987, 1990; Lyons and Hendriks, 1991; Sanborn and Hendler, 1988; Spector and Hendler, 1990; Wood, 1990). They argue that "traditional" planning (e.g., Fikes and Nilsson, 1971; McDermott, 1978), cannot effectively generate action in the real world, where many things are unknown, the effects of actions (or operators) cannot be perfectly predicted, and things happen that are outside of the agent's control. They emphasize the need for an agent to be reactive to new contingencies, within time constraints on reasoning and action.

In fairness to research pre-dating "reactive planning", it should be noted that many of the problems they focus on were previously recognized, if not solved. For example, as discussed throughout this paper, Simon and Sloman emphasize the need for concurrent monitoring (to react to new contingencies by positing motives or generating reflexes, to pick up information to complete partial plans, etc.), and the need to output solutions within the time available.

Sanborn and Hendler (1988) present a model, CROS, which deals with the pedestrian's task of crossing a four lane road (a top level goal) with automobile traffic. They focus on reaction and don't address plan generation: for "reaction is inherently plan following: the system cannot be expected to synthesize new reactions in real time" (p. 99). Their system has a set of possible actions, each of which has enabling, inhibitory and constraining conditions. An action can be taken if an enabling and no inhibitory condition is matched. CROS has monitors, some of which check for the appearance of cars, the others predict whether the car may collide with the agent and they fire only if so (i.e., they cause reaction, to the extent that an alternative behaviour is possible). The implication of this is that only relevant new data affect behaviour. Actions which will cause the robot to approach its goal of crossing the street are taken if they are enabled. However, once an action is taken, if the system predicts a collision, it will have to constrain some actions (i.e., inhibit some and enable others). This gives priority to the new goal of avoiding the car. The authors claim that the goal of getting to the other side of the street still may influence behaviour, if moving right is safely possible the system will move right.

What can we draw from CROS for the design of the nursemaid? The world of CROS is a simple one in which there are only three basic operators (move left, stop, move right), one long term goal and one possible short term goal. The nursemaid's multiple motives and primitive operators offer a contrast. CROS doesn't plan, it retrieves plans; whereas the nursemaid is intended to deal with problems of deliberation. But even if one only retrieves plans, that doesn't remove the problem of multidimensional evaluation of goals. If deliberation is kept simple, there is still the problem of meta-planning. It may be possible for the nursemaid to have reactive responding in highly urgent situations. These would be reflex responses (some of which are innate, some learnt). And she could synthesize plans if time is available. A generalized notion of monitors firing may be incorporated into the goal generators of the nursemaid and the nannies.

Firby (1987) presents a model of reactive planning based on the notion of "reactive action packages" (RAPs). A RAP is an independent process which aims at achieving an end. A RAP may be run, suspended and killed. Many RAPs can exist in the execution queue simultaneously. Like Sanborn and Hendler's model, Firby's model doesn't create plans. Associated with each RAP, however, is a set of solutions ("task nets"). Solutions have conditions of applicability. Even if one solution doesn't achieve its end, it may still be useful in changing the state of the world such that another solution may be tried out (for the same or a different RAP). A RAP is killed only if either (1) its goal is achieved, (2) every possible solution has been tried, or (3) its conditions of applicability no longer hold. RAPs can cause other RAPs to appear on the execution queue.

Firby seems to have in mind the model of an operating system for dealing with the task of choosing which RAPs will run. He doesn't address the problem of having to assess, reject or postpone goals. His model nonetheless illustrates the idea of having multiple dependent and independent planning processes.

Agre and Chapman (Agre, 1988; Agre and Chapman, 1987, 1990), believe that the world is too complex, uncertain, and that it requires responses in too short a time for plans "as programs" to be of great use. They prefer to view a plan as a resource to be used to guide, but not determine, behaviour. They claim that almost all human activity is inherently "improvised" (Agre, 1988), i.e., that it is continually dependent on the present circumstances. People nonetheless are capable of goal directed behaviour. They propose that people are capable of reacting to current events by postulating a set of candidate responses, and that some mechanism, called "action arbitration", must select between the candidate responses.

Agre and Chapman (1987) propose a model of an agent, Pengi, which plays the arcade game Pengo, which basically requires that Pengi rapidly fight and escape from bees, by manoeuvring (around) blocks. This game was chosen because things happen too rapidly for the construction of plans. Pengi plays this game by responding to the current situation by proposing actions which compete for control of behaviour. It selects them by action arbitration. Their model doesn't act on a model of the world activated from long-term memory (see Agre, 1990). Rather it responds to its immediate perceptual knowledge of the world (though it maintains some state).

Both the Pengo and the nursemaid domain are fast and complex. The focus of Pengi (and its theory of activity) is more on the selection of action than the selection of motives. The nursemaid will actually have to cope with more knowledge than Pengi, since Pengi only considers actions which are local to its surround, whilst the nursemaid will receive signals from distal locations, and hence will have to cope with more alternative motives. The nursemaid will do something like action arbitration, except that the decision rules will be more diversified (there will be filtering of motives and multidimensional evaluation of motives), and her focus is on motives, which may nonetheless require specific actions in given situations.

In Agre and Chapman (1987), Firby (1988), and Sanborn and Hendler (1988) action is reactive: The systems don't construct new plans. (Kaelbling, 1988, sees this as a weakness and proposes a language for building reactive plans.) The behaviour of the system is determined by rules which contain in their antecedents relevant aspects of the current state of the world. Plan selection mechanisms may be worth investigating for the design of the nursemaid. However, one musn't assume that people always (even in the realm of fast changing environments like Pengo) are capable of instantly solving problems. Some measure of deliberation is often required. Therefore, the need for assessing motives and scheduling deliberation musn't be ignored, especially in worlds in which many possible independent motives may accrue, and need to be ordered, since they can't all be attended to at once. As was noted above, this is a focal point of our research. In recent work in cognitive psychology, emotions have been postulated to handle some of this selective attention.

6.2 Emotion and Motive processing

Oatley and Johnson-Laird (1987) and Frijda (1986) argue that emotions serve the purpose of focussing attention on a limited number of action alternatives. These are functionalist theories of emotions, in as much as they suppose that emotions have an important function. (Contrast Sloman, 1987, Pfeifer, 1988).

6.2.1 Oatley and Johnson-Laird's theory of emotions

Oatley and Johnson-Laird (Johnson-Laird, 1988; Oatley, 1988; Oatley and Johnson-Laird, 1987) propose that the mind is modular, and that this modularity requires special mechanisms for communication between modules. Modules may communicate between each other with propositional or non-propositional signals. Emotions are based on non-propositional signals called "emotion signals". Emotion signals are elicited when the evaluation of the outcome of a plan changes very much. Their effect is to elicit an "emotion mode", i.e., to "lock" the operating system and its submodules into a state in which they focus on a limited set of goals, events, and kind of possibilities. There is a small set of emotion modes, underpining the (five or so) basic emotions: happiness, sadness, anxiety, anger and disgust. Each emotion mode tends to inhibit the others and favour certain courses of action. The action may be caused by internal reflex responses, or it may require planning and other partly conscious processes. The following quotation summarizes their thesis.

> We assume that these ["significant"] junctures are both distinctive and recurring, so that the emotional system in mammals has evolved to recognize them and to establish distinctive responses to them. Indeed the function of these modes is to organise a transition to a new phase of planned activity directed to the priorities of the mode with associated goals and certain stored plans for dealing with what has happened [Oatley and Johnson-Laird, 1987, p. 35. Italics ours]

Thus they assume an emotional system which recognizes and responds to junctures by organizing a transition. They therefore take Simon's emphasis on interrupts in emotion to the point of positing a special emotional system for dealing with certain kinds of transition.

The main problems with this thesis have to do with its lack of parsimony in dealing with phenomena which meet their specification of emotion modes without being emotional. That is, it is difficult for them to parsimoniously account for the fact that the kind of transitions they are talking about often occur in the absence of an emotional state! One's everyday life is shot through with junctures in which the evaluation of the likely success of one's plans changes. Motivation is the story of a system repeatedly changing its strategies and revising its objectives in the light of new information. Yet, only few of these junctures are emotional. Moreover, many of these shifts are very complex (and no doubt use complex control mechanisms) and they also move the system into a state in which it is focussing on a limited range of behavioural options (hence the focus of attention, which they view as characteristic of emotion, is not peculiar to emotional transitions).

Oatley and Johnson-Laird could handle the criticism by ascertaining that a transition isn't enough: "the distinctive phenomenological tone, the somatic changes, the behavioural expressions, and courses of action [are necessary for an emotion]" (Oatley and Johnson-Laird, 1987, p. 34). Of course the question would still remain: What determines the appearance of these features? To this they could respond like Ortony (1988) and suggest that there is some measure of importance associated with the goals involved in the transitions and which is required for an emotional state to arise.[3] After all, they do speak of "significant" junctures of plans. But this is not satisfactory, since people can attentively yet "coldly" replan for very important goals. (Perhaps people classified as having "anti-social personality disorder" under the Diagnostic and Statistical Manual of Mental Disorders-Revised (1987) do this more than others.) (Compare the discussion of anger in Sloman, 1987.)

6.2.2 Frijda's concern-based theory of emotions

Frijda's (1986; Frijda and Swagerman, 1987) theory of emotions is more explicitly akin to a theory of motivation than Oatley and Johnson-Laird's. He supposes that emotions involve concerns, which are states which the system wants to avoid, achieve, or maintain. The mind is unconsciously looking out for possible violations or satisfactions of concerns, and if it detects some, it may, after a few steps of processing, generate an "action tendency"--a tendency to satisfy (or avoid the violation of) a concern.

As a theory of emotions, Frijda's model is too general. His theory of emotion can therefore be read as a theory of motivation. "As regards motivation, a sharp distinction between emotion and motivation cannot be made. They are overlapping concepts" (77). Yet, again, for the same reasons as were noted in relation to Oatley & Johnson-Laird model, the distinction needs to be drawn.

6.2.3 Sloman's attentional theory of emotions

Sloman (1987) proposes a relationship between emotions and motive processing as partly described above. He characterizes emotional states as states in which some motives and thoughts have a subtle dispositional property--their insistence--to get through attention filters and divert limited limited resources. E.g., if X is romantically in love with Y, then thoughts about Y and desires regarding Y have the tendency to interrupt X's thinking. These thoughts need not actually disturb X's thinking. For example, if X is engaged in a highly attention demanding task, then they may not be able to get through attention filters.

Recall that motive filtering based on insistence is supposed to be designed for quick (and thereby fallible) responding. These thoughts and motives may therefore manage to divert attention even if the high order assessment process has already rejected them because they are incompatible with other motives.

This theory of emotions rests on the notion of motive processing. Emotions, however, are not equated with motive processing. They are seen as a special collection of substates into which a system with a certain kind of motive processing architecture can fall.[4]

---

[3]Note that they address the question of whether the transition will be euphoric or dysphoric, but not whether it will be emotional or not.

[4] There may be a subspace of possible architectures which permit emotional substates.

By extending our characterization of motive processing and developing a computational model, we may also shed some light on the issue of emotion. Since the model will incorporate mechanisms for filtering, assessment, and planning, it may be possible to study the effects of varying these processes. We will try to use the model of the nursemaid to perfect the attentional theory of emotion, and this will feed back into our assessment of our model of motive processing.

7. Conclusion

The current psychological models of motive processing aren't sufficiently elaborate or specified to account for the essential features of motive processing described above. None has been subject to a test as rigourous as that of performance in the nursemaid domain. Hopefully, focussing our attention on the simple yet rich domain of the nursemaid will prove useful for improving existing theories.

References

Agre, P. and Chapman, D. (1987). Pengi: An implementation of a theory of activity. *AAAI-87*, 268-272.

Agre, P. E.. (1988). *The dynamic structure of everyday life.* Technical Report, MIT AI 1087, Department of Comptuter Science, Massachusetts Institute of Technology.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (Third edition-Revised). Washington, D.C.: American Psychiatric Association.

Bindra, D. (1978). How adaptive behavior is produced: a perceptual-motivational alternative to response-reinforcement. *The Behavioral and Brain Sciences, 1*, 41-91.

Boden, (1988). Fashions of Mind. In M. Boden (1988) *Artificial Intelligence in Psychology.* Cambridge, MA: MIT Press.

Chapman, D. (1985). Planning for conjunctive goals. *Artificial Intelligence, 32,* 333-377.

Dean, T. and Boddy, M.. (1988). An analysis of time-dependent planning. In AAAI-88, 49-54.

Fikes, R. E. and Nilsson, N. J. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence.* 2, 189-208.

Firby, R. J. (1987). An investigation into reactive planning in complex domains. *AAAI-87, 1,* 202-206.

Frijda, N.H. (1986). *The Emotions.* Cambridge, Cambridge University Press.

Frijda, N.H. and Swagerman, J. (1987). Can computers feel? Theory and design of an emotional system. *Cognition and Emotion, 1,* 235-257.

Hayes, P.J. (1985). The second naive physics manifesto. In J.R. Hobbs and R.C. Moore (Eds) *Formal theories of the common sense world,* 1-36. Norwood, NJ: Ablex Publishing Corp.

Heckhausen, H. and Kuhl, J. (1985). From wishes to action: The dead ends and short cuts on the long way to action. In M. Frese and J. Sabini (Eds), *Goal directed behavior: the concept of action in psychology.* London: Lawrence Earlbaum Associates.

Kaelbling, L. P. (1988). Goals as parallel program specifications. *AAAI-88.* 60-65.

Kuhl, J. (1982). The expectancy-value approach within the theory of social motivation: Elaborations, extensions, critique. In Norman Feather (Ed), *Expectations and actions: expectancy-value models in psychology.* Hillsdale, NJ: Lawrence Earlbaum Associates.

Kuhl, J. (1988). A motivational approach to volition: activation and de-activation of memory representations related to uncompleted intentions. In V. Hamilton, J. Bower, and N. H. Frijda *Cognitive perspectives on emotion and motivation* (pp. 63-85).

Lyons, D.M. and Hendriks, A.J. (Unpublished Manuscript, 1991). *Reactive Planning.*

McCulloch, W. (1945). A Heterarchy of values determined by the topology of nervous nets. Reprinted in W. McCulloch (2nd Ed, 1989) *Embodiments of Minds.* Cambridge, MA: MIT Press.

McDermott, D. (1978). Planning and acting. *Cognitive Science, 2,* 71-109.

Oatley, K. (1988). Plans and the communicative function of emotions: a cognitive science theory. In V. Hamilton, Bower, and N. Frijda (Eds) *Cognitive perspectives on emotion and motivation* (pp. 321-343). Amsterdam: NATO.

Oatley, K. and Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion, 1,* 29-50.

Ortony, A. (1988). Subjective importance and computational models of emotions. In V. Hamilton, Bower, and N. Frijda (Eds) *Cognitive perspectives on emotion and motivation* (pp. 321-343). Amsterdam: NATO.

Pfeifer, R. (1988). Artificial Intelligence models of emotion. In V. Hamilton, Bower, and N. Frijda (Eds) *Cognitive perspectives on emotion and motivation* (287-320). Amsterdam: NATO.

Sanborn, J.C. and Hendler, J.A. (1988). A model of reaction for planning in dynamic environments. *Artificial Intelligence in Engineering, 3,* 95-102.

Simon, H.A. (1967). Motivational and emotional controls of cognition. *Psychological Review, 74,* 29-39.

Sloman, A. (1978). *The computer revolution in philosophy: philosophy, science, and models of mind.* Atlantic Highlands, NJ: Humanities Press.

Sloman, A. (1985b). Real time multiple-motive expert systems. In M. Merry (Ed.) *Expert systems 85.* Cambridge: Cambridge University Press.

Sloman, A. (1987). Motives mechanism and emotions. *Emotion and Cognition,* 1, 217-234. Reprinted in M.A. Boden (ed) *The Philosophy of artificial intelligence* "Oxford Readings in Philosophy Series" Oxford University Press, 1990.

Sloman, A. (Unpublished manuscript, 1986). Robot nursemaid scenario.

Sloman, A. and Humphreys, G. (Unpublished manuscript,1990). Towards a computational model of affective processes.

Sloman, A., and Croucher, M. (1981). Why robots will have emotions. Proceedings 7th International Joint Conference on A.I.

Spector, L. and Hendler, J.A. (1990). *Knowledge strata: reactive planning with a multi-level architecture.* CS-TR-2564. Department of Computer Science and Systems Research Center, University of Maryland, MD.

Steel, S. (1987). The bread and butter of planning. *Artificial Intelligence Review,* 1, 159-181.

Wilkins, D. E. (1985). Recovering from execution errors in SIPE. *Computational Intelligence, 1,* 33-45.

Wood, S. (1990) *Plan recognition, dynamic world modelling and plan elaboration in the autodrive system for rapidly changing, uncertain, multi-agent environments.* CSRP 170. School of Cognitive and Computing Sciences, University of Sussex, UK.

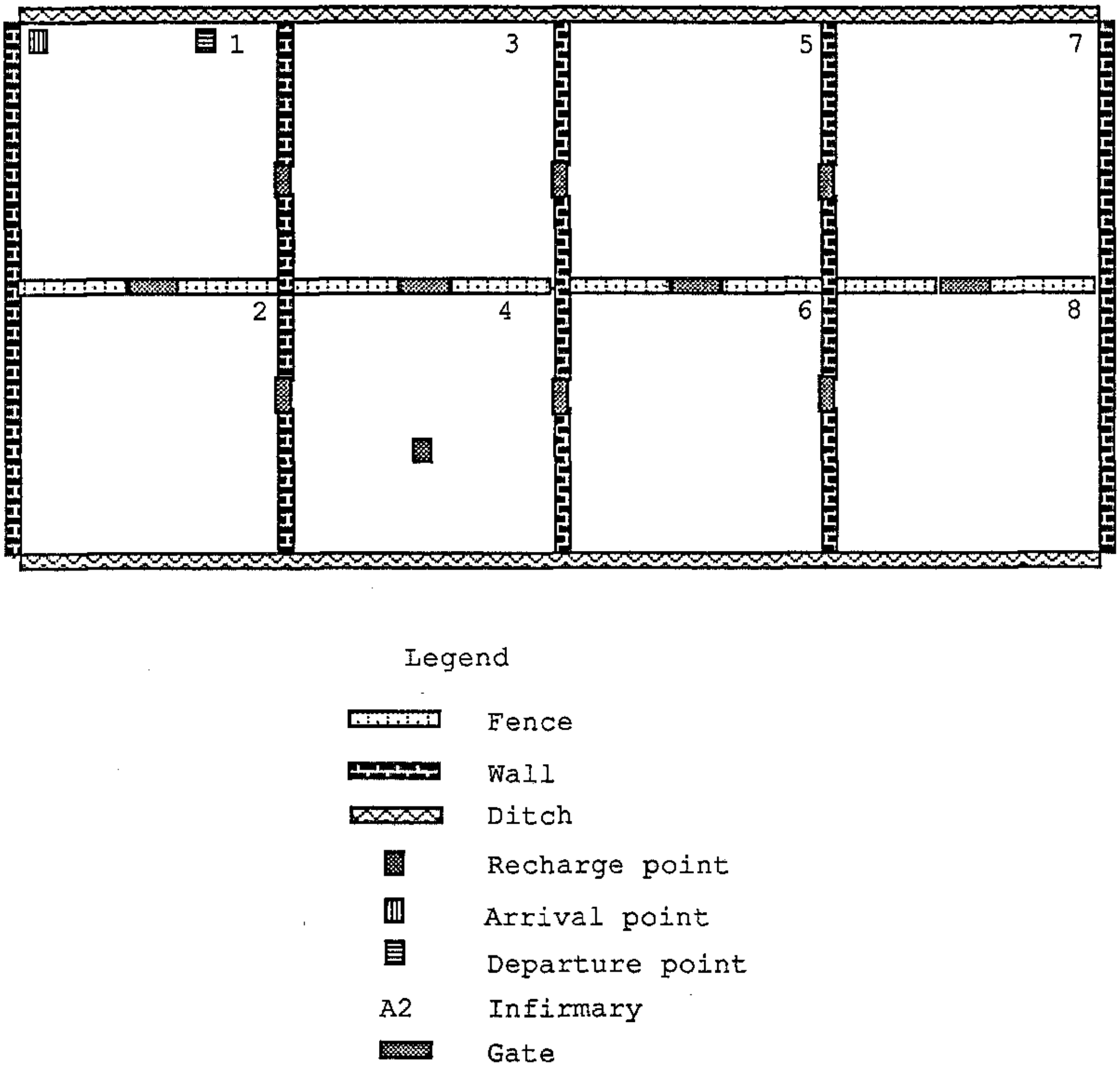

Figure 1. A pictorial representation of the nursery.
Areas are numbered in their upper right corner.

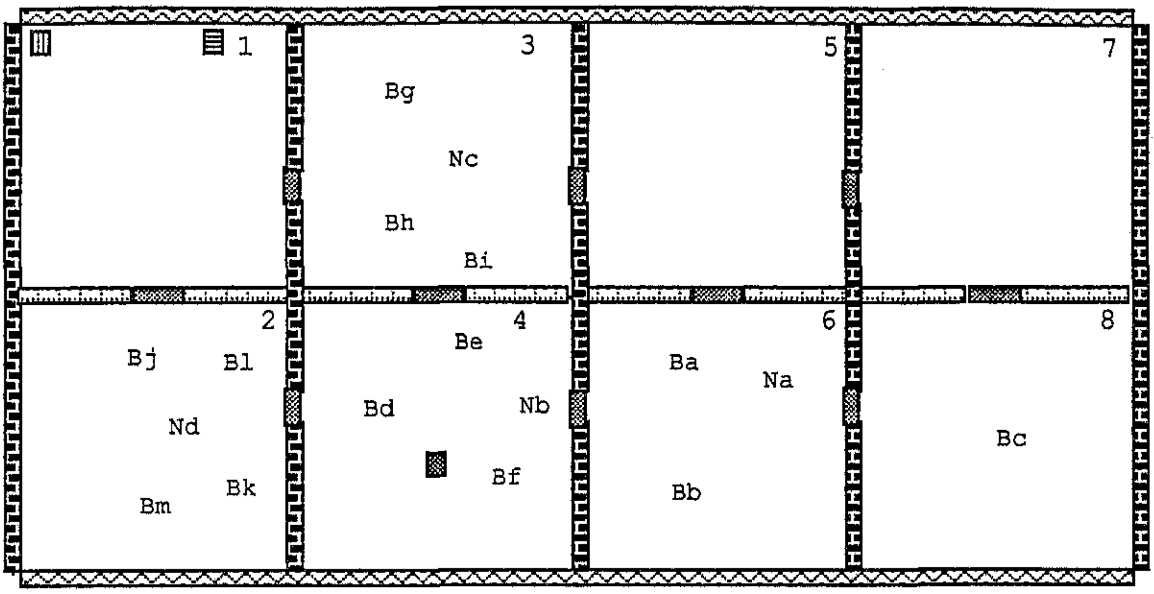

Figure 2. A pictorial representation of possible positions of robots in the nursery.

```
signal(4, Nc, lowcharge(Bi, [A1 50 95], 4)).
time = 14.

signal(3, nursemaid, overpopulated(A2, 4))
time = 12.
```

Figure 3. A representation of recent signals sent to the nursemaid by a nanny and the nursemaid

```
signal(5, nursemaid, missing(Bc)).
(time = 16).

signal(4, Nc, lowcharge(Bi, [A1 50 95], 4)).
time = 14.

signal(3, nursemaid, overpopulated(A2, 4))
(time = 12).
```

Figure 4. A representation of recent signals sent to the nursemaid by a nanny and the nursemaid.

Figures 2, 3, and 4