

Emotion-Based Attention Shift in Autonomous Agents

Luis Miguel Botelho*

Helder Coelho†

* Department of Computer Science of ISCTE, Av. das Forças Armadas, Edifício ISCTE, 1600 Lisbon, Portugal
luis@iscte.pt

† Department of Computing Science, Faculty of Sciences of the University of Lisbon, Campo Grande Bloco 5 Piso 1, 1700 Lisbon, Portugal
Helder.Coelho@di.fc.ul.pt

Abstract. We present three mechanisms of attention shift for autonomous agents within the framework of the SALT model of memory [5]: activation-based attention shift, attention shift by event-driven emotion and attention shift by anticipation-driven emotion. The three mechanisms rely on automatically computed properties of memory (as opposed to deliberative processes).

Activation-based attention shift is based on the activation of cognitive structures stored in long term memory. The other two are two-step emotion-based mechanisms of attention shift: first emotion interrupts the agent's current cognitive process (as in activation-based attention shift), then the processing of the emotion directs the agent's attention to the external environment. In attention shift by event-driven emotion, the ongoing cognitive process may be interrupted when an external event causes the agent to experience an emotion. In the attention shift by anticipation-driven emotion, the ongoing cognitive process may be interrupted when the agent anticipates affective states attributed to the external environment.

1 - Introduction

Attention shift consists of interrupting or suspending the current cognitive task and attend to something else. In this paper we consider the shift of attention to the external environment. The global research goal of our work has been to investigate the adequacy of using cognitive models as a basis for an architecture for effective autonomous agents. Although we recognize the fundamental role of deliberative processes we want to deepen our understanding of the extent to which automatic processes may determine the agent's behavior. Besides other reasons (e.g., efficiency) this is an important issue when talking about architectures because automatic mechanisms may be built into them.

Some proposals regarding autonomous agents [28] don't address to the problem of control of attention. However, the ability to shift attention from the current cognitive task to the external world is an essential property of autonomous agents designed to operate in dynamic environments (e.g., a mobile robot picking up pieces of trash in a populated laboratory, or a softbot attending to stock rates, buying and selling stocks in a dealing room). Some approaches have dealt with the control of attention the same way they deal with regular action selection [11]: attention shift doesn't have any special status. However, since both artificial and natural agents have limited

rationality, we need a special policy to prevent the current cognitive process from being constantly interrupted by each and every external event. In [20] and [1] attention is controlled by production rules, but this requires a great amount of explicit representations. At the architecture level, attention shift has been addressed along two ways: the commitment-based approach [26][23]; and the insistence-based approach [4][16].

In the commitment-based approach the agent blocks external interruptions in certain phases of its activity, regardless of the importance of external events. When an external event (or other stimulus) occurs, the agent checks if the goals generated by the external event are compatible with its current plan. Plan-compatible goals are adopted by the agent (without disturbing its current plan). A plan-incompatible goal is adopted only if external interruptions are allowed. In the insistence-based approach the agent performs a fast and heuristic evaluation of the external event and decides if it should be attended, interrupting its current cognitive process. External stimuli are evaluated according to their insistence, importance and urgency.

In this paper we present three new attention shift mechanisms: activation-based attention shift, attention shift by event-driven emotion and attention shift by anticipation-driven emotion.

Activation-based attention shift represents a particular instance of the general ideas of the insistence-based approach within the framework of the SALT model of memory. Attention shift by event-driven emotion and attention shift by anticipation-driven emotion constitute two of the mechanisms by which emotion influences cognitive processes. In both of them, an emotion interrupts the agent's current cognitive process the same way an external stimulus does in activation-based attention shift. Afterwards, the processing of the emotion directs the agent's attention to the external environment. With respect to the two emotion-based mechanisms of attention shift, the paper concentrates mainly on issues regarding the ways emotion interrupts the agent's current cognitive process. The generation of emotions and the contents of their cognitive representation will not concern us much here (for more details on these, see [7]).

In the case of attention shift by event-driven emotion, the attention shift mechanism may interrupt the agent's current cognitive task if the agent experiences an emotion in response to an external event. In the case of attention shift by anticipation-driven emotion, the interruption may occur, not as a result of an emotion directly produced by the external event, but as a consequence of an emotion that results of anticipating an affective state attributed to the external event.

We don't regard our proposals as the best attention-shift mechanisms of all, much less the only ones. They are just other useful approaches. With respect to the broad agents view [13][16] we think an agent should be equipped with more than one mechanism for the same purpose. In the specialized view of agents [22] our contribution is to widen the palette of the designer. In general terms our work contributes in three ways: since it is specially concerned with the automatic components of processes (i) it reduces the task of agent specification, (ii) it may represent gains in efficiency, and since it uses ideas from cognitive modeling (iii) it may provide a sound basis for more comprehensive theories of such phenomena as mood and emotion both in human beings and in artificial agents.

In section 2, we discuss the insistence-based and the commitment-based approaches. Then we present the activation-based attention shift mechanism and show how it relates to the commitment and insistence-based approaches, avoiding some of their potential drawbacks. In section 3, we present the attention shift by event-driven emotion and the attention shift by anticipation-driven emotion mechanisms. In section 4, we analyze our proposals in the context of an example. Section 5 presents some conclusions.

2 - Activation and Attention Shift

In the commitment-based approach to attention shift, planning phases in which interruptions are allowed alternate with phases in which interruptions are not allowed regardless of what external events or stimuli take place.

The architectures IRMA [10][25] and PRS [19][18] provide a filtering mechanism that enables the agent to ignore external events. This filter can be overridden by external events in particular conditions. [26] in the case of IRMA and [23] in the case of PRS report experimentation with the commitment-based approach using two different filter-overriding policies. In [26], the agent blocks external interruptions when its activity has much value (as anticipated by the designer, not by the agent). In [23], the agent blocks external interruptions until a fixed amount of planning steps are performed, then interruptions become allowed and then again interruptions become blocked and so on. Both mechanisms may be acceptable for relatively benevolent environments where the stakes are not very high. However, in more challenging worlds, an event that represents an important opportunity or risk may occur while the agent has blocked external interruptions. In such environments, the commitment-based attention shift mechanism is not a good design choice.

The insistence-based approach does not suffer from this important drawback for it evaluates each external event in order to decide if it should be attended. According to [4], insistence is a heuristic measure of the importance and urgency of external stimuli. If the insistence of the external event surpasses a certain threshold the event is further evaluated regarding its importance and urgency and possibly gains the agent's attention, interrupting its former processing. Otherwise, the external event is ignored and the current cognitive task proceeds. The exact nature and computation of insistence are important issues. On one hand, if insistence is a prefixed value of each event it does not require any additional computation when the event occurs, yielding a very fast attention shift decision, but the event will be treated the same way regardless of the context in which it occurs. On the other hand, if the insistence of an event is computed when needed, the decision may be made context-dependent, but it will be more time consuming. Aaron Sloman [32] presents a proposal that avoids this problem. According to his view, insistence should emerge automatically (as opposed to thoughtfully or deliberately). However, Sloman does not point a specific architecture with that automatic feature.

In both the commitment-based and the insistence-based approaches, if the architecture does not allow some form of parallel processing, any external event will always interrupt the current task, even if only during the evaluation of its compatibility with the current plan or the evaluation of its insistence.

In the remaining of this section we show that the concept of activation of the SALT model of memory [5] provides an automatic suitable equivalent of insistence. We present an attention shift mechanism based on the activation of cognitive structures stored in memory and show that both the commitment-based approach and the insistence-based approach to attention shift are special cases of our proposal.

2.1 - Activation of Cognitive Structures Stored in Memory

SALT is a model of memory for autonomous agents in the same tradition of a class of spreading activation cognitive models [14][9][2][1]. The purpose of this paper is not to explain SALT in great detail nor to discuss arguments in favor or against the model. Instead, we present a brief description of it just to allow the reader to better understand the concepts relevant to this paper.

SALT views long term memory as an associative network represented by a directed labeled graph in which nodes contain cognitive structures, and arcs represent associations between nodes with certain strengths. Nodes may contain symbolic representations of a variety of entities, like beliefs, plans, procedures, goals, desires, moods and emotions.

Whenever the agent is faced with an external stimulus the node representing it receives a certain amount of activation per time period. The activation received by a node spreads to the rest of the network through the arcs emanating from it. The amount of activation that flows through a given association depends on its strength. The stronger the association the more activation will flow through it to the next node. In the SALT model, the contents of any two nodes don't have to overlap in order for an association to exist between the two. Instead, the strength of an association from node N_i to node N_j reflects the ratio between (i) the number of times N_j is selected to working memory immediately after N_i , and (ii) the number of times N_i is selected.

According to SALT, more activated nodes are more likely to be selected to working memory where their contents will be processed, therefore the activation of a node reflects its accessibility in long term memory. The activation of all nodes in the network decays exponentially with time.

The automatic functioning of long term memory exhibits context dependent behavior. That is, the selection of nodes to be processed in working memory depends on the pattern of activation of the whole network that, in turn, depends on the history of previous interactions of the agent.

2.2 - Activation-Based Attention Shift Mechanism

According to SALT when a node is being processed by the agent, its contents are copied from long term to working memory, but the node is not removed from long term memory. It stays there subject to the dynamics of long term memory as described in subsection 2.1. The activation-based attention shift mechanism proposed here relies on such dynamics.

At a given instant of time, the agent pays attention only to the cognitive structures being processed in working memory at that instant. This means that, in order for a given cognitive structure to be attended, it must be copied to working memory. When this happens, attention shifts from the previous contents of working memory to its new contents.

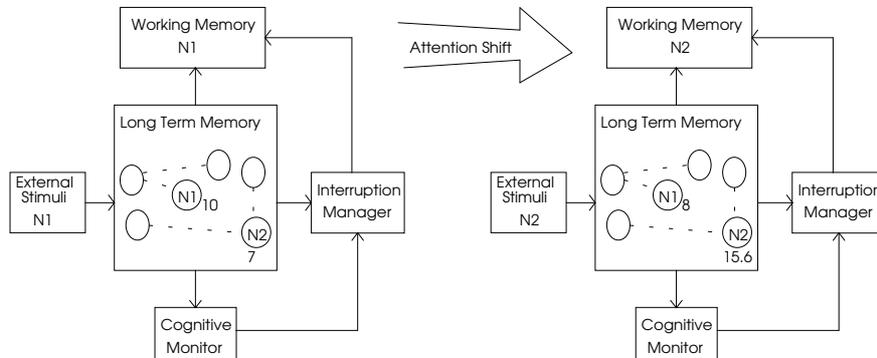


Figure 2.1 - Activation-based attention shift

In the framework just presented, attention shift may take place when the activation of a particular node in long term memory becomes greater than the activation of the node containing the cognitive structures being processed in working memory. Fig. 2.1 shows an abstract situation in which the attention of the agent shifts from node N_1 (in the left) to node N_2 (in the right). The underlying architecture is described in [7].

In the left side of fig. 2.1, the contents of the node N_1 are being processed in working memory. In that situation the activation of N_1 in long term memory has a certain value, say 10, and no other long term memory node has a greater activation. In the right side of fig. 2.1, after a while (and a certain activation decay) a new stimulus is presented to the agent changing the whole pattern of activation of the network. We assume that, as a consequence of this change, node N_2 has now a greater activation than N_1 ($16.6 > 7$). We propose that, under these circumstances, the cognitive monitor signals the interruption manager that N_2 has become more activated than the node being processed in working memory. The interruption manager performs some deliberation¹ in order to decide whether or not to send a message to working memory saying node N_2 should be attended. Of course, it may have happened that the change of activation resulting of the presentation of a new stimulus would not have lead N_2 to become more activated than N_1 . In this case no attention shift would have occurred. It is also possible that in spite node N_2 becomes more activated than N_1 , the interruption manager does not decide to interrupt the cognitive processing currently going on in working memory.

The attention shift mechanism just presented does the same job as the insistence-based mechanism, but the dynamic adjustment of the filter threshold as well as the automatic computation of the insistence measure, as suggested in [32], are no longer problems. If the activation of cognitive structures in memory plays the role of insistence, we know exactly how to compute insistence and what it means within our architecture: (i) insistence is computed automatically (as opposed to deliberately), and (ii) insistence has an adaptive nature because it is learned through the agent's

¹ The deliberation made by the interruption manager enables us to implement different attention shift policies. For instance, the interruption manager may check if interruptions are allowed.

experience. Besides, in a spreading activation model of memory like SALT, importance and urgency are determined relative only to the most activated criteria in long term memory (e.g., beliefs). As a final comment, the activation-based attention shift mechanism does not suspend the current cognitive task of the agent in order to decide whether or not attention should shift because the underlying architecture [7] has some forms of parallel computation (e.g., the cognitive monitor and working memory work in parallel, fig. 2.1).

[24] also contends that the mechanism to control attention should be as automatic as possible in order to avoid consuming cognitive resources. It presents such a mechanism (i.e., Alarms) to automatically determine which of the agent's goals is attended in each instant of time. However, unlike SALT, Alarms is not a completely automatic mechanism. The complete definition of the alarm intensity function requires the agent to do some *a priori* planning. [24] assumes the agent has always time to perform this *a priori* planning when each goal is detected in the environment. This, of course, might not be the case in many applications. In contrast, activation-based attention shift just requires that the activation of the node representing the external stimulus be compared with the activation of the node whose contents drive the current cognitive processing. Nevertheless Alarms exhibits some advantages relative to the SALT model².

If we want to build an agent capable of blocking interruptions like in the commitment-based approach to attention shift, the agent's plans just have to set or unset an interrupt-enable flag. When the interruption manager receives a signal specifying a node to be considered for being attended, it checks the status of the interrupt-enable flag. If it is unset, the new stimulus is ignored; if it is set, the new stimulus is attended.

In conclusion, we stress that activation-based attention shift plays the same role of the insistence-based approach avoiding the problems related to the determination of the insistence of external stimuli. The same can be said with respect to Alarms [24]: we don't have to build the alarm intensity function when a new goal is detected. Since activation-based attention shift is an automatic process it doesn't have to be explicitly specified by the agent designer. Finally, we have also seen that the activation-based attention shift can block external interruptions like in the commitment-based approach.

3 - Emotion and Attention Shift

Commitment-based attention shift, insistence-based attention shift and activation-based attention shift all provide ways of controlling the conscious cognitive processing of an agent but none of them acknowledges the special role of emotion. However, there is both psychological [21] and neurological [15] evidence and theory suggesting that emotion (as well as mood and other motivational controls) plays a special role in cognition. First, Herbert Simon [29] and later Aaron Sloman [30][32] recognized an intimate relationship between emotion and attention shift. Both of these researchers have proposed that emotion arises as a result of interrupting the agent's

² Unlike SALT, alarms can represent the effect of the passage of time in the "intensity" of each goal.

current cognitive task, but none of them considered it to be the initiating event of the interruption process. Here we propose that the feeling of an emotion may also cause attention to shift. This is a powerful mechanism since it may lead the agent to attend relevant external events that would have passed unnoticed otherwise.

In this section we present two emotion-based mechanisms of attention shift and show how they are designed within the framework of the SALT model of memory. In the first of these mechanisms (subsection 3.1) the emotional state that results of the occurrence of an external event triggers an attention shift process. The second mechanism (subsection 3.2) is based on the anticipation of affective states.

Before proceeding, a word about mood and emotion is in order. We've been talking about emotion but we haven't said what we mean by it yet. Since mood and emotion are very ambiguous possibly overlapping concepts we will try to make a clear distinction between them without giving comprehensive and definite definitions. First we try to informally distinguish mood from emotion. Then, following Sloman's idea concerning the definition of concepts such as consciousness and perception [31] we propose a more rigorous distinction based on the roles these concepts play within our architecture for artificial agents.

Informally, both mood and emotion have an evaluative component. Like in [12] and [17], when someone is feeling good we say he or she is feeling a positive mood or is in a good mood and conversely for bad mood. Emotion refers to more specific things such as sadness, fear, anger, happiness, fascination and excitement. Emotions may have a positive or a negative valence. Fear and anger have negative valence, happiness and fascination have positive valence, and different instances of excitement may have opposite valences. It is easy to confound some emotions with moods. For instance, happiness may be confounded with positive mood, and sadness with negative mood. However they are not the same: happiness is usually associated with laughter, smiling and other external signs and behaviors while positive mood isn't necessarily so. Besides, someone may be feeling good (i.e., positive mood) and yet be nostalgic, not happy.

Moods are more general than emotions, are more stable in time, and contrary to emotions can't be attributed to concrete identified causes. However people often try to make such attributions.

The distinction between mood and emotion is much easier to understand in terms of the roles they play in our architecture. Both moods and emotions have a representational component encoded as cognitive structures stored in long term memory. Both moods and emotions change the accessibility of cognitive structures in memory, biasing judgment and recall. The major difference between mood and emotion concerns the way they determine the agent's behavior.

Mood together with other factors condition the extent to which information is processed [6]; mood indirectly determines the global behavior of the agent due to the general mechanisms of mood regulation [8]; mood (specially good mood) can be used as a heuristic in judgment and decision making (affect as information).

Emotion has a more direct impact on behavior (intentional component of emotions), generating reflexive actions, generating goals and other motives, and generating behavioral features (e.g., aggressive, relaxed). While mood conditions the global behavior of the agent due to the general mechanisms of mood regulation, each

emotion directs behavior in specific ways dependent on the more accessible cognitive structure representing the emotion.

Finally, although mood and emotion *per se* don't integrate mental states, the conscious feeling of mood or emotion is a component of the agent's mental state.

Sometimes we use the expressions affect or affective state to refer both to mood and to emotion.

In the following subsections, we present two mechanisms by which emotion may control attention in autonomous agents.

3.1 - Attention Shift by Event-Driven Emotion

In [7] we present an architecture for autonomous agents in which some of the ingredients needed to generate emotions from special sensory information are described. These ingredients include an emotion generator and monitor. Since there is no space to provide full details we just present the more relevant information here. The emotion generator and monitor evaluates the current (internal and external) state of affairs using special purpose machinery (e.g., special purpose sensors, special purpose access methods, compiled representations). In certain conditions, as a result of this evaluation, a cognitive structure representing a particular emotion is activated. This cognitive structure is contained in a node stored in long term memory like any other cognitive structure. Just to be a little more informative, it is worth saying the process of activating an emotion may be much faster than the activation of cognitive structures in the usual way, because it involves special purpose sensors, evaluators and activators. The general ideas underlying the generation/activation of emotions are based on the theories of Sloman [30] and Bates [3]. In short, the emotion generated/activated in a particular situation depends on the instincts/needs/motives/goals/values whose satisfaction is threatened or made possible, the cause of that threat or opportunity and the motivators generated in response to it. Emotion is also the result of unusual or unexpected events or stimuli.

If the node containing the representation of the generated emotion becomes more activated than the node containing the information being currently processed in working memory, a signal is sent to the interruption manager as described in subsection 2.2 and the process proceeds the same way as in activation-based attention shift. In this case, the signal is generated by the emotion monitor and not by the cognitive monitor. Once again, the operation of this special purpose monitor provides much faster responses than the cognitive monitor. It should be noticed that in this process, the current cognitive task is interrupted by an emotion, not by the external stimulus. Hence, if attention is to be shifted to the external stimulus, the selected cognitive representation of that emotion must include a procedure whose execution directs the agent's attention to the external environment. Such a procedure is shown in subsection 3.2, fig. 3.2 and in [7].

In summary, there are two main differences between this attention shift mechanism and the activation-based attention shift described in section 2. First, attention shift by event-driven emotion represents some gains in efficiency provided by special purpose machinery. Second, it is a two step process in which the current cognitive task is interrupted by emotion and only then is the agent's attention directed to the external environment.

In situations that represent immediate risks or opportunities regarding the agents motives (e.g., needs, goals) but in which activation-based attention shift wouldn't have occurred the agent may still be led to attend the external environment due to the mediating role of emotion, even before the conscious evaluation of the situation.

Attention shift by event-driven emotion may happen in different kinds of situations according to the mediating emotion. If the mediating emotion arouse of the evaluation of a threatening situation, attention shifts due to danger; if the emotion arouse because the satisfaction of a motive was made possible, attention shifts due to detected opportunity; finally if the emotion appeared as a consequence of an unusual or unexpected stimuli, attention shifts due to surprise.

3.2 - Attention Shift by Anticipation-Driven Emotion

It is widely accepted that emotions have an important evaluative component, thus if in the current situation a negative affective state is anticipated a negatively valenced emotion will arise. Similarly, when a positive affective state is anticipated, a positively valenced emotion will take place. Hence when an emotion arises of the anticipation of an affective state, and the anticipated affective state is attributed to an external stimulus, the agent should pay attention to external information in order to try to avoid the experience of a negative state or to ensure the occurrence of a positive state. The main idea is to find a fast way of anticipating the possible occurrence of affective states and determine if they are attributable to the agent's current thinking or to an external event. In the remaining of this section we show that the SALT model of memory provides the means for such anticipation and attribution processes.

We start by the attribution process. Although it doesn't mean to imply that there aren't attribution processes due to deliberative reasoning (e.g., causal explanation), in what follows we present an automatic mechanism. Imagine the situation depicted in fig. 3.1 in which the node whose contents are driving the agent's current thinking (N) is associated to a node representing an affective state (S). Suppose also the node representing the occurrence of an

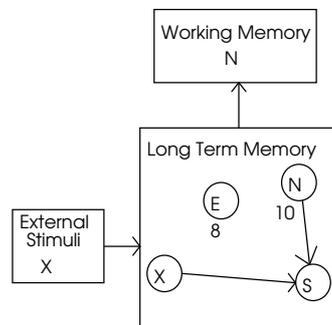


Figure 3.1 - Attribution and anticipation

external event (X) is also associated to the same affective state S. Since the flow of activation from one node to another is proportional to the strength of the association between them [5], we may assume that a sudden large variation of the activation of a node is more likely to be caused by a node strongly associated to it than by a node weakly associated to it. Then, in fig. 3.1 if we observe a large variation of the activation of S and the association from X to S is stronger than the association from N to S, we have a good reason to attribute the variation of the activation of S to X (and conversely). Of course this process is subject to errors (see [27] for attribution errors

in persons), but it provides an automatic and efficient way of making attributions, avoiding conscious and lengthy deliberations.

Concerning the anticipation process, suppose the activation of *S* suffered a great increment as a result of the occurrence of *X*. Since the strength of the association from any node (N_1) to any other (N_2) represents the relative frequency of selection of N_2 to working memory following the selection of N_1 , and *X* is strongly associated to *S*, there is a good reason to think that *S* is going to be selected soon (i.e., to anticipate the conscious feeling of the affective state *S*). Once more, the anticipation process could have been the result of deliberative reasoning. Nevertheless the automatic mechanism described can be of much use, specially when the agent can't find any other way of doing things, i.e., automatic mechanisms can be used by default.

The mechanism of attention shift by anticipation-driven emotion relies on the previous arguments. If there is a sudden large variation of the activation of a node (*S*) representing an affective state, and the association from the node representing an external event (*X*) to *S* is stronger than the association from the node driving the agent's current thinking (*N*) to *S*, the emotion monitor and generator produces an externally-driven emotion of expectation activating the node representing that emotion (*E*). The remaining of the attention shift process is analogous to the previous cases (subsections 2.2 and 3.1): if the activation of *E* is greater than the activation of *N*, the emotion monitor sends a signal to the interruption manager saying that *E* should be considered to be attended. As before, the interruption manager decides whether or not the emotion *E* should interrupt the current cognitive task and informs working memory accordingly.

```
procedure expectation_of_affect(node:X)
begin
  If the contents of X do not drive the current cognitive process Then
    If the buffer for external stimuli contains a codification represented in X
      Then attend X
end
```

Figure 3.2 - Simplified procedure for the emotion "expectation caused by something"

As in the case of event-driven emotion, in attention shift by anticipation-driven emotion, the agent's attention is shifted to the node representing the emotion, not to the external event. As before (subsection 3.1) the agent will attend the external environment only if the selected node representing the emotion contains a procedure whose execution forces it to do so. Fig 3.2 shows a simplified version of such a procedure for the emotion of expectation of an affective state attributed to something or some event represented in node *X*.

Attention shift by anticipation-driven emotion presents the same advantages as attention shift by emotion-driven emotion (subsection 3.1).

Emotion-based attention shift (either due to the current external stimuli or due to anticipation) is not an entirely automatic process since it depends on the contents of the nodes representing the cognitive component of emotions. The agent designer has to specify such contents (e.g., procedures like fig. 3.2, or goal creation procedures),

therefore it is possible that feeling an emotion doesn't cause attention to shift to the external environment.

4 - Attention Shift by Anticipation-Driven Emotion: An Example

This section discusses an example of the operation of attention shift by anticipation-driven emotion. We have chosen this mechanism because it involves features shared by all three mechanisms presented in the paper.

In the first situation (situation 4.1 a), a man is trying to rest in his couch in a Sunday afternoon, while his wife and his children went to visit his mother in law. The man is thinking he has not been playing with his small children lately because he has been so busy. Meanwhile a small dog yelped outside but the man didn't notice it and kept with his thoughts.

In the second situation (situation 4.1 b), the man is also trying to rest in his couch while his wife and his children went to visit his mother in law. The man is thinking he has to get up and study those files he brought from work. Meanwhile his wife and his children have just arrived from their visit. After a little while, the man realizes his family has returned and their children have been talking outside.

Fig. 4.1 (a) exhibits the cognitive structures relevant to situation 4.1 (a), and fig. 4.1 (b) shows the cognitive structures relevant to situation 4.1 (b).

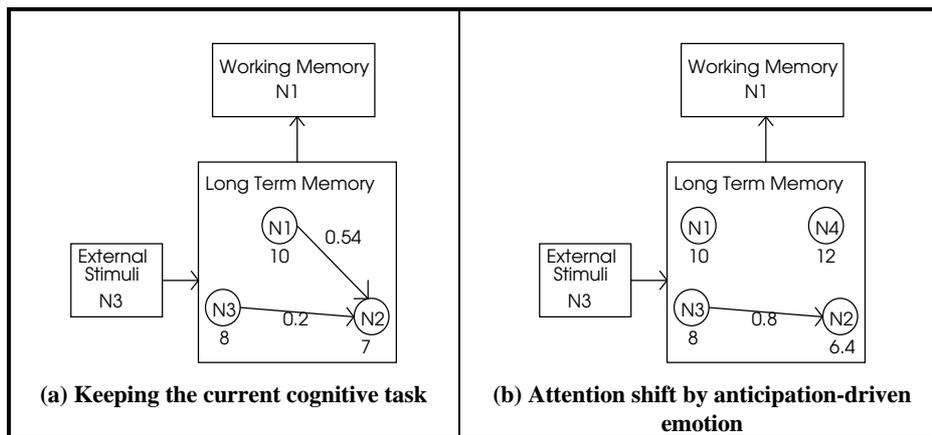


Figure 4.1 - Focus of attention

In what follows we assume a simplifying hypothesis: nodes containing the representation of encodings stored in the buffer for external stimuli receive 8 units of activation (pre-attentive activation parameter), nodes whose contents drive the agent's current cognitive process receive 10 units of activation (attentive activation parameter), and nodes activated by the emotion generator receive 12 units of activation (affective activation parameter). Accordingly, time is not considered.

In fig. 4.1 (a), node N_1 represents the man's current thinking "not playing with children", node N_2 represents bad mood, and node N_3 represents yelping dogs. Before the dog has yelped, the activation of N_1 was 10, the activation of N_2 was 5.4 (due to the association from N_1 to N_2) and the activation of N_3 was 0. When the dog yelped

outside, the activation of node N_3 became 8. Since the activation of N_1 is 10, N_3 does not interrupt the man's current thinking. As yelping dogs (N_3) are also associated to negative mood, N_2 gets an additional amount of activation, becoming 7. Since the activation of N_2 did not suffer a great change (only from 5.4 to 7) there is no attention shift. That is, the man proceeds his current thinking and doesn't notice a dog is yelping outside.

In fig. 4.1 (b), node N_1 represents the man's current thinking "study files", node N_2 represents good mood, node N_3 represents his children talking, and node N_4 represents the emotion "positive externally driven expectation". In the initial conditions, the activation of N_1 was 10 and the activation of all other nodes was 0. When the children started talking outside, the activation of N_3 became 8 which is less than the activation of N_1 (the node whose contents are driving the man's current thinking). Since there is a strong association in the man's mind from "talking children" (N_3) to the node representing good mood, the activation of N_2 suffers a great change (from 0 to 6.4). The sudden variation of the activation of N_2 is detected by the emotion generator and monitor. Since the association from N_3 to N_2 (0.8) is stronger than the association from N_1 to N_2 (0) the emotion generator attributes the anticipated positive mood to the external stimulus and produces the emotion "positive externally driven expectation" activating N_4 . Since the activation of N_4 becomes 12 (greater than the activation of N_1), the emotion monitor sends a signal to the interruption manager informing it that N_4 should be considered to gain the man's attention. Assuming the interruption manager decides to interrupt the man's current thinking, N_4 is copied to working memory and its contents are processed. If N_4 contains a procedure whose execution results in the man's information processing copying to working memory the node that matches the sensory information still present in the buffer for external stimuli (see fig 3.2, subsection 3.2), N_3 gains the man's attention and he suddenly realizes his children have been talking. That is, the man's current thinking ("study files") is interrupted and his attention shifts to "talking children".

In the above examples we have shown that our proposals can explain human-like behavior involving attention shift. Unfortunately we didn't build an implementation of the full architecture due to hardware and software constraints. Therefore we couldn't test our ideas in artificial agents. Nevertheless the examples discussed show that attention shifts are difficult if not impossible to predict since they rely on the automatic functioning of memory and depend on the memory state prior to the occurrence of external events. It is also obvious that the exact behavior of such an agent depends of the parameter setting used (e.g., pre-attentive activation, attentive activation, and affective activation parameters).

5 - Final Remarks

This paper presents three mechanisms of attention shift: activation-based attention shift, attention shift by event-driven emotion, and attention shift by anticipation-driven emotion. We have shown that the commitment-based approach and the insistence-based approach are particular cases of the activation-based attention shift. All three mechanisms of attention shift presented rely heavily on automatically computed properties of memory, specifically activation of nodes and strength of associations. This feature is extremely important because it avoids the use of lengthy deliberative

processes to decide if an external stimulus is or is not attended. Our two emotion-based mechanisms of attention shift constitute precise interpretations of general ideas concerning the special role of emotion in cognition. Not only is emotion a consequence of the interruption of current goals [29][30][32], but it can also interrupt and direct cognitive processes. That is, in line with [15] and [33], besides the evaluative aspects of emotions, our approach preserves their intentional aspects.

As presented here, attention shift mechanisms provide the means for an agent to have self-oriented (volitional or goal-driven) thoughts which may be interrupted by particular external events (data-driven). Furthermore, since activation is a context-dependent property, the conditions enabling the current cognitive task to be interrupted depend on the particular context in which the event occurs.

Emotion-based attention shift is an essential feature in achieving successful behavior in coping with rapidly changing high stakes situations since the agent is led to attend to its environment even before the relevant event has been consciously evaluated.

It is worth noting the proposals presented here and the architecture underlying them exhibit all the requirements postulated by Simon [29] for intelligent agents, that is, interrupting and terminating mechanisms. The main topic of the present paper has been the definition of interruption mechanisms that rely on properties of the SALT model of memory. However, SALT was extended to accommodate the influence of motivation in information processing. The extended model is called COMINT [6]. If the present analysis had been made substituting COMINT for SALT we would have obtained the four terminating conditions proposed by Simon: goals should be terminated (1) when they become achieved; (2) when they become achieved well enough (satisfied); (3) when motivation or time is run out; and (4) when they become believed to be impossible to achieve.

Acknowledgments

The authors are indebted to Pedro Ramos for constant comments and suggestions about the ideas presented in this paper, specially in what respects the way any node in long term memory may become associated to cognitive representations of moods and emotions. We also want to thank the anonymous reviewers for their helpful comments.

References

1. J.R. Anderson. ACT: A Simple Theory of Complex Cognition. *American Psychologist*, 51:355-365, 1996
2. J.R. Anderson and P.L. Pirolli. Spread of Activation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10:791-798, 1984
3. J. Bates, A.B. Loyall and W.S. Reilly. An Architecture for Action, Emotion and Social Behavior. *Proc. of the 4th European Workshop on Modeling Autonomous Agents in a Multi-Agent World, (MAAMAW'92)*, 1992
4. L.P. Beaudoin and A. Sloman. A Study of Motive Processing and Attention. In Sloman, A., Hogg, D., Humphreys, G., Partridge, D.

- and Ramsey, A. (eds) *Prospects for Artificial Intelligence*, p229-238. IOS Press, Amsterdam, 1993
5. L.M. Botelho and H. Coelho. A Schema-Associative Model of Memory. *Proc. of the 4th Golden West International Conference on Intelligent Systems (GWICS'95)*, p81-85, 1995
 6. L.M. Botelho and H. Coelho. Information Processing, Motivation and Decision Making. *Proc. of the 4th International Workshop on Artificial Intelligence in Economics and Management (AIEM'96)*, 1996
 7. L.M. Botelho and H. Coelho. An Agent Architecture for Attention Shift. *Working Paper*, 1996
 8. L.M. Botelho and H. Coelho. Autonomous Agents with Adaptive Behavior: Learning by Mood Regulation. *Working Paper*, 1996
 9. G.H. Bower. Mood and Memory. *American Psychologist*, 36:129-148, 1981
 10. M.E. Bratman, D. Israel and M.E. Pollack. (1988) Plans and Resource Bounded Practical Reasoning. *Computational Intelligence* 4:349-355, 1988
 11. R.D. Bulos. An Economic Approach to Reasoning About Next Action Selection in Intelligent Agents. *Unpublished document CSRP 433*. School of Cognitive and Computing Sciences, University of Sussex, UK, 1996
 12. M.S. Clark and A.M. Isen. Towards Understanding the Relationship Between Feeling States and Social Behavior. In Hastorf, A. and Isen, A. (eds) *Cognitive Social Psychology*, p73-108. Elsevier/North-Holland, N.Y., 1982
 13. H. Coelho. Facing Hard Problems in Multi-Agent Interactions. *Proc. of the Nato Advanced Research Workshop on the Future of Intelligent Systems*. Springer-Verlag, 1992
 14. A.M. Collins and E.F. Loftus. A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, 82:407-428, 1975
 15. A.R. Damásio. *Descartes' Error: Emotion, Reason and Human Brain*. Picador, London, 1994
 16. D.N. Davis. Reactive and Motivational Agents: Towards a Collective Minder. In Müller, G.P., Wooldridge, M.J. and Jennings, N.R. (eds) *Intelligent Agents III — Proc. of the Third International Workshop on Agent Theories, Architectures and Languages (ATAL'96)*. Springer-Verlag LNAI Series, 1996. In this volume.
 17. J.P. Forgas. The Role of Emotion in Social Judgments: an Introductory Review and an Affect Infusion Model (AIM). *European Journal of Social Psychology*, 24:1-24, 1994
 18. M.P. Georgeff and F.F. Ingrand. Decision Making in an Embedded Reasoning System. *IJCAI'89*, p972-978, 1989
 19. M.P. Georgeff and A.L. Lansky. Reactive Reasoning and Planing. *AAAI'87*, p677-682, 1987
 20. B. Hayes-Roth. An Architecture for Adaptive Intelligent Systems. *Artificial Intelligence*, 73:329-365, 1995
 21. I. Janis and L. Mann. *Decision Making. A Psychological Analysis of Conflict, Choice and Commitment*. The Free Press, a division of Macmillan Publishing Co. Inc., New York, 1977

22. H.A. Kautz., B. Selman and M. Coen. Bottom-up Design of Software Agents. *Communications of the ACM*, 37:143-146, 1994
23. D.N. Kinny and M.P. Georgeff. Commitment and Effectiveness of Situated Agents. *IJCAI'91*, p82-88, 1991
24. T.J. Norman and D.P. Long. Alarms: Heuristics for the Control of Reasoning Attention. *Proc. of the Annual Conference of the Cognitive Science Society (CCSS'95)*, 1995
25. M.E. Pollack. The Uses of Plans. *Artificial Intelligence*, 57:43-68, 1992
26. M.E. Pollack and M. Ringuette. Introducing the TILEWORLD: Experimentally Evaluating Agent Architectures. *AAAI'90*, p183-189, 1990
27. L. Ross. The Intuitive Psychologist and his Shortcomings: Distortion in the Attribution Process. *Advances in Experimental Social Psychology*, 10:174-221, 1977
28. M. Schroeder, I.A. Móra and L.M. Pereira. A Deliberative and Reactive Diagnosis Agent Based on Logic Programming. In Müller, G.P., Wooldridge, M.J. and Jennings, N.R. (eds) *Intelligent Agents III — Proc. of the Third International Workshop on Agent Theories, Architectures and Languages (ATAL' 96)*, Springer-Verlag LNAI Series, 1996. In this volume.
29. H.A. Simon. Motivational and Emotional Controls of Cognition. *Psychological Review*, 74:29-39, 1967
30. A. Sloman. Motives, Mechanisms and Emotions. *Cognition and Emotion*, 1:217-234, 1987
31. A. Sloman. Why Consciousness Is not Worth Talking About (Yet): Extended Abstract. *Proc. of the International Conference on Cognitive Science (ICCS' 91)*, 1991
32. A. Sloman. What Sort of Control System Is Able to Have a Personality? To appear in the *Proc. of the Workshop on Designing Personalities for Synthetic Actors*, 1995
33. I.P. Wright. Cognition and Currency Flow. Notes Towards a Circulation of Value Theory of Emotions. *Unpublished document*. Cognitive Science Research Center of the University of Birmingham, UK, 1995