

Computational Models of Emotion for Autonomy and Reasoning

Darryl N. Davis and Suzanne C. Lewis

Department of Computer Science, University of Hull, Cottingham Road, Kingston-upon-Hull, UK, HU6 7RX.

[D. N. Davis@dcs.hull.ac.uk](mailto:D.N.Davis@dcs.hull.ac.uk), <http://www2.dcs.hull.ac.uk/NEAT/dnd/index.htm>

[S. C. Lewis@dcs.hull.ac.uk](mailto:S.C.Lewis@dcs.hull.ac.uk), <http://www2.dcs.hull.ac.uk/NEAT/index.html>

Keywords: Emotion, Computational Models, Autonomy, Perception, Reasoning

Recent evidence suggests that the emotions play a crucial role in perception, learning and rational decision making. Despite arguments to the contrary, all artificial intelligent systems are, to some extent, autonomous. This research investigates how emotion can be used as the basis for autonomy. We propose the use of an emotion-based control language that maps over all layers of a computational architecture. We report on how theoretical work and both design and computational experiments with this concept are being used to direct perception, behavior selection and reasoning in cognitive agents.

1 Introduction

Definitions of intelligence in artificial systems have involved the use of many concepts including deliberation, reasoning [19] and the selection of appropriate behaviors for any given situation [4, 17]. In reasoning, information is granted belief status, either partially in probabilistic systems or wholly in logic based reasoning systems, and used as the basis for further deliberation. This deliberation may give rise to altered belief states and leads to the selection of goals, plans and behavior. Typically an agent chooses from alternative responses because of design decisions or learning. The choice is made on the basis of information or control metrics or a-priori ranking of alternative behaviors.

Underlying these perceptual and reasoning processes is the concept of autonomy. Alterman [1] suggests task-effective artificially intelligent systems need not be designed in terms of autonomy. Intelligence arises out of the interaction of the system and the user. However when system goals and resource allocation are in conflict, considerable interaction is required with a user. This interaction, itself a system goal (whether implicit or explicit), may never be satisfied unless the system can decide to perform the appropriate actions. To decide between actions, given that not all preconditions for any action will be specified at design time, requires the system to be, in some sense, autonomous.

Emotions and their nature have been studied for a considerable time, with many contrasting theories and views of emotion being formed. A traditional perspective of emotion is of something that is irrational and detracts from reasoning. However, recent evidence [8] suggests that emotions are an essential part of human intelligence, and play a crucial role in perception, rational decision-making and learning.

Most major current theories of emotion agree that emotions constitute a very powerful motivational system that influences perception and cognition in many important ways. For example neurons in the amygdala

are driven particularly strongly by stimuli with emotional significance, indicating an important role in the coding of the emotional significance of sensory data. Further research suggests that motivation and emotion serve as filters that guide perception and affect the evaluation of perceptual information [3]. This view is supported by Izard [16] who argues that emotion is a guiding force for perception. If emotion is a primary source of motivation, it must play a significant role in both initiating and providing descriptors for the types of disequilibria described by Mearleu-Ponty [18] as underlying behavior in biological agents.

From a computational perspective, Sloman considers that intelligent machines will necessarily experience emotion (-like) states [24]. Following on from Simon [23], this developing theory of mind considers how perturbant control states ensue from attempting to achieve multiple goals, or goals at odds with resource availability and environment affordances. Perturbant states will arise in any information processing infrastructure where there are insufficient resources to satisfy all current and prospective goals. This will occur not only at the deliberative belief and goal management levels but over all layers of the architecture as goals are mapped onto internal or external behaviors and actions. An agent must be able to recognise and regulate these emotion-like states or compromise its task effectiveness.

The aim of this research is to investigate theories of emotion and understand how they can be used to underpin computational autonomy, to direct and inform perception and behavior selection and to form a better model of computational reasoning. This paper describes this ongoing research and the integration of an emotional model into two different types of computational architectures.

2 Psychology and Emotions

Research has shown that emotion affects many different aspects of cognitive function including memory [5], reasoning and social interaction [15]. There has never been any doubt that emotion disrupts reasoning under certain circumstances and that misdirected or uncontrolled emotion can lead to irrational behavior. However, evidence from Damasio and other sources also suggests the contrary and that emotions play a fundamental role in rational and intelligent behavior such as decision-making and reasoning.

The Somatic Marker Hypothesis [8], for instance, states that decisions, made in circumstances whose outcome could be potentially either harmful or advantageous, and which are similar to previous experience, induce a somatic response used to mark future outcomes. When the situation arises again the somatic marker will signal the danger or advantage. Thus, when a negative somatic marker is linked to a particular future outcome it serves as an alarm signal to be wary of that particular course of action. If instead, a positive somatic marker is linked it becomes an incentive to make that particular choice.

The appraisal approach to emotion has cognition as the core element in emotion. The OCC (Ortony, Clore and Collins) model [21] synthesises emotions as outcomes to situations. Emotions arise out of a valenced reaction to situations consisting of events, objects and agents. The emotion type elicited is dependent upon appraisals made at each branch of the model. The OCC model allows for an emotional state to be a situation itself, so emotions can trigger additional emotions or the same one repeatedly. The OCC model is well suited to computational modeling as shown in the work of Elliot [11].

The goal-oriented approach suggests that emotions arise from evaluations of events relevant to goals. Again cognition is central to the elicitation of emotion. Oatley and colleagues [20] argue that emotions are caused by cognitive evaluations that may be conscious or unconscious. Each kind of evaluation gives rise to a distinct signal that reflects the priority of the goal, which then influences resultant behaviors. Frijda [13] uses a similar definition of emotion and states that certain stimuli elicit certain emotional phenomena because of the individual's concerns and the relevance of the stimuli to the satisfaction of these concerns.

Duffy posed the question, at which particular degree does a characteristic become an 'emotion' or at which degree is it a 'non-emotion' [10]. For example, a raised heartbeat is characteristic of both emotional and non-emotional behavior. When does the difference in the characteristic occur? Is emotion a distinguishable state or a difference in the degree certain response characteristics exhibit?

According to Duffy, the phenomena that are described as emotions occur in a continuum or a number of continua. The responses called 'emotional' do not appear to follow different principles of action to other responses of the individual. She states that all responses, 'emotional' or 'non-emotional', are reactions of an organism as it adapts

to a situation. Emotion represents a change in the energy level, or the degree of reactivity of an individual. For example, situations, which are interpreted as threatening, are characteristically responded to with increased energy. Small changes in energy level may occur during 'interest' or 'boredom', whereas 'anger' is associated with a more extreme change.

Duffy supports the goal-oriented view that emotions are only experienced in situations of significance to the individual. The intensity of the 'emotion' is proportional to the degree of importance associated with a particular goal and the degree of threat or promise the situation bears for that goal. The emotion experienced is also affected by the background and information that the individual has about the particular situation.

Many theories use the concept of basic emotions. For example, the OCC model contains twenty-two different emotion types. Oatley and Johnson-Laird cite four basic emotions derived from evolutionary origins: happiness, sadness, fear and anger. A further five are derived from innate biological substrates: attachment, parental love, sexual attraction, disgust and interpersonal rejection. However, other theorists question the notion of basic emotions. Scherer and Duffy oppose the view of basic emotions and examine evidence that emotions are patterns of interrelated changes. Using basic emotions in a theory can lead to what Scherer calls 'bunching' of the different emotional states around a limited number of types. Conversely, the scope of emotional states in both Duffy's and Scherer's theories is considerably broader.

Scherer points towards the existence of a large number of universal 'response elements' as opposed to basic emotions [22]. His concept of modal emotions attempts to address many questions. For example, why does the same situation not necessarily provoke the same emotional expression nor the use of the same label in two individuals? Like Duffy, Scherer sees emotion as a number of changes that occur over time in response to an event. He defines emotion as "*a sequence of interrelated synchronised changes in the states of all organismic subsystems in response to the evaluation of an external or internal stimulus event that is relevant to central concerns of the organism*". The emotional state results from the cumulative evaluation of relevant changes in internal or external stimulation. Scherer proposes that such organisms make five types of checks: novelty, intrinsic pleasantness; relevance to meeting plans; ability to cope with the perceived event; and compatibility of the event with self-concept and social norms. An appraisal according to these checks is carried out which elicits an emotional response. Scherer believes that the information from these checks is needed in order to choose how to respond. Some combinations of evaluation checks would be frequently encountered, giving rise to the same recurring patterns of state changes. The term 'modal emotions' refers to states resulting from these recurring stimulus evaluation check patterns. Although some patterns occur more frequently, the number of potential emotional states is virtually infinite.

3 Autonomy, Goals and Emotion

Many frameworks are used for thinking about, designing and building intelligent systems. The use of rational BDI (Belief-Desire-Intention) models [6] is understandable, as they provide formal systems with well-defined properties. The limitations of such systems, e.g. logical omniscience and resource constraints, are known. Goal competition due to incompatible goals, insufficient resources or skills is a major research issue. Ferber [12] categorizes goal interaction in multi-agent systems as one of three categories: indifference, cooperative and antagonistic. Certainly in the latter case, and even for cooperative agents or goal interaction, perturbant states can arise. Such agent societies and intelligent systems need some means to manage these states or risk compromising their autonomy and reactivity, and hence their task effectiveness. Even the most rational agent architecture will be compromised if it lacks the mechanisms to cope with the emergent effects of antagonistic goal conflicts.

One stance is to place a computational analogue to emotion at the core of an agent. This provides an agent with an internal model that maps across different levels and types of processing. Emotion provides an internal basis for autonomy and a means of valencing information processing events. It provides an internal model of use in ordering motivation and goals, and the means for choosing actions and regulating behavior. This emotional core can be used to recognise and categorise transient, episodic, trajectory and persistent control states.

Sloman [25] also differentiates between episodic and persistent mental phenomena. His architectures for functioning minds include primary, secondary and tertiary emotions. Primary emotions are analogous to arousal processes in the emotion theories introduced above and have a reactive basis. Secondary emotions are those initiated by appraisal mechanisms and have a deliberative basis. Tertiary emotions are cognitive perturbances – typically negatively valenced emergent states – such as arising from goal or motivator conflicts in an information processing architecture, for example a multi agent society. Any agent architecture that supports multiple motivations or goals is liable to this type of dysfunction. Perturbant states can arise through resource inadequacy or mismanagement while pursuing multiple and not necessarily incompatible goals. Most computational systems face this type of problem even if their underlying theory does not. Possible solutions are particularly relevant to the design of goal-oriented and agent systems.

4 Four Layer Computational Architecture

Earlier research on agents focused on an architecture that supports motivation [9]. The current framework builds on that architecture. It is used to pursue alternative computational perspectives on architectures of mind. Here the interplay of cognition and emotion is emphasized through mechanisms that support appraisal, motivation, tasks and roles. Emotions are accepted to be

part mental (appraisal) states with descriptive (valencing) and causal (arousal) processes. This concept is used to provide a control or regulatory framework to model the different forms of emotion inducing events. The fundamental tenet of this work is that all agent events and actions, internal and external, can be described in terms of this model of emotion.

A salient feature of many definitions of emotion is that they are described in terms of goals, roles (or norms) and responsive behaviors. This enables different aspects of motivational behavior to be consistently defined over different levels of the architecture in terms of an emotion-based control language. Global drives are those associated with the agent's overall and persistent purpose. Temporally-local drives are related to ephemeral states or events within the agent's environment or itself. Emotional autonomy allows an agent to select and attempt to maintain an ongoing globally-temporal *disposition* towards its roles. The nature of this is temporarily affected and perhaps modified through current goals and motivations. Over time events occur that modify, stall, negate or satisfy goals. Such events can be described within a model of emotion. An emotion-based control language can therefore be used to mediate the interaction of global roles and the temporally-local drives that reflect the current focus of the agent.

An agent's internal environment can be defined in terms of its perception of external events, objects and agents and the behaviors (whether internal or external) they afford. Such descriptions can be organised according to control state theory [9]. The control language used to navigate this internal environment needs to be consistent across many levels and types of control state from autonomous reflexes to extensive deliberation associated with goal satisfaction or belief management. Various combinations of qualitatively different behavior are required of an agent as it attempts to achieve different categories of goals associated with a role. Different problem-solving trajectories, described in terms of goal-achieving behaviors, exist for any one role. A greater range exist where an agent has multiple and not necessarily contingent roles. Some trajectories while impossible are supported or attended to for any number of reasons; for example, the motivational intensity associated with a preferred goal or role. The possible trajectories depend on an agent's design. An agent is autonomous to the extent that it can choose to pursue specific motivational trajectories. An agent is rational to the extent that it follows feasible or achievable trajectories.

Figure 1 shows emotion used as the core to a motivation based model of agenthood. This architecture emphasizes four distinct processing layers: a reflexive layer analogous to the autonomic systems in biological agents, a reactive or behavioral layer, a deliberative layer and a regulating reflective layer. The broad picture is of high and low level processes co-existing and interacting in an asynchronous, parallel and holistic manner. The majority of the higher level processes tend to remain dormant and

state persistent; activated only when sufficiently required. The agent's processing exists in relation to the agent's environmental stance; i.e. what roles the agent has adopted, what objects, agents and events occur in the environment and how they affect the logistics of goal and role satisfaction. Motivator processing, planning and other cognitive processes are not merely abstract, nor just reactions to the current state of an agent's external environment but exist in relation to an agent's long term goals. Motivations, goals and the behaviors they subsume are all influenced by components of the emotion engine.

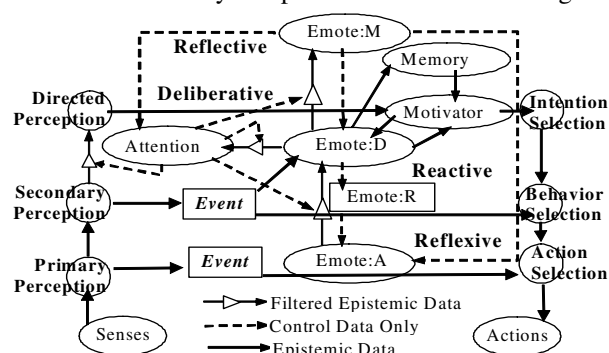


Figure 1. Sketch of the emotion engine based four-layer architecture. Overall the process is of information assimilation and synthesis, and information generation that typically map onto internal and external behaviors

If emergent behaviors are to be recognized and described in terms the emotion based control language and then managed, there must be a design synergy across the different layers of the architecture. Processes at the deliberative level (for example *Emote:D*, *Attention* and *Motivator* in Figure 1) can reason about emergent states arising from anywhere in the architecture using explicit (motivator or goal) representations (see [9]) and the internally consistent control language. In an earlier architecture, the reflective processes (*Emote:M*) were used to classify the processing patterns of the agent in terms of combinations of a set of basic emotions and favored emotional responses (or disposition). Subsequent rejection of the concept of basic emotions, for theoretical and computational reasons, required a redesign of this component. The emotion-changing reactive behaviors (*Emote:R*) are used to pursue a change in disposition through changing the functional behavior of the lowest-level autonomous processes (*Emote:A*). This module is modeled using multiple communities of cellular automata (or *hives*). The behaviors associated with this module, and set by the *Emote:R* module, are those that govern the internal behavior of single cells, the communication between adjoining cells in communities and inter-community communication. Emotion is discretely valenced at the cell level as positive-neutral-negative. Ordinal measures across the valence of all the cells at the community level provide the basis for ascending control signals. Various threshold models have been used to determine if arousal occurs; for example a community of cells with a high aggregate valence, or a high degree of valence contrast across the cell community.

Emotions can be instantiated by events both internal and external at a number of levels, whether primary, e.g. ecological drives, or by events that require substantive cognitive processing. Emotions can be invoked through cognitive appraisal of agent, object or event related scenarios, including for example the unwanted postponement or abandonment of a goal. To move to a preferred aspect of the possible emotional landscape, an agent may need to instantiate other motivators and accept temporarily unwanted dispositions. An agent with emotional autonomy needs to accept temporary emotional perturbation if it facilitates goal satisfaction at some future time.

In the model shown in Figure 1, intense emotions or arousal events effectively override the reactive-level filters, activating the deliberative components of the emotion engine. Deliberative appraisal of an emotion inducing event can initiate lateral activation at the deliberative layer, affecting memory, attention and motivator management. Memory responds to emotional context as an aid to the storage and recall of memories about external events, objects and agents. Attention management makes use of the emotional state of *Emote:D-R-A* complexes to provide a semantic context for motivator filters, and set the quantitative emotion filters. The intensity levels of these filters are set in response to the *Emote:D* mechanisms and the reflective component (*Emote:M*) of the emotion engine.

Computational experiments have used both sets of basic emotions and type-less emotion arousal models. Early experiments found that from any given state, a hive rapidly achieved a steady (continuous or oscillating) state. By changing the currently extant behavior set, or by communicating with another hive, transitions to the same or other steady states always occurred. Approximately 20,000 transition possibilities exist. Rules are used to select different hive dispositions and transitions. Similarly, through the modification of the internal state of a small number of cells, the emotion engine moves to a closely related but preferred state. This is analogous to the modal responses described in the Scherer model of emotion.

5 CRIBB and Emotion

5.1 The CRIBB Model

CRIBB (Children's Reasoning about Intentions, Beliefs and Behavior) is a computer model based upon a general sketch for belief-desire reasoning in children [2]. It simulates the knowledge and inference processes of a competent child solving false-belief tasks [27]. A simulation run in CRIBB starts by giving propositions containing facts and perceptions about some scenario in sequential steps according to the time interval in which the propositions arise. On the basis of the given propositions and the inferences drawn, CRIBB answers test questions about the cover story. The questions can be about its own beliefs or about the intentions, beliefs and behavior of another person in the scenario.

CRIBB represents propositions about physical states of a given situation and the intentions, beliefs, perceptions and behavior of others. Its knowledge base consists of four types of practical syllogisms and three other inference schemata, which represent the relations between these propositions. Practical Syllogisms denote knowledge about the relations between intentions, behavior and beliefs of another person. The three other classes of inference schemata relate perception-belief, belief-time and fact-time. These are split into primary and secondary representations. Primary representations are the system's own beliefs about the situation and the behavior of other people. Fact-time inferences, propositions about facts along a time scale, are classed as primary representations. Belief-time and perception-belief inference schemata are both types of secondary representation as they contain beliefs about the system's own and others' beliefs. A further element of CRIBB is a consistency mechanism that detects and resolves contradictions in belief sets. This is invoked each time a new proposition is added, in order to ensure the consistency of its knowledge base.

5.2 Extending CRIBB with Emotions

Bartsch and Wellman's model [2] for belief-desire reasoning includes an emotion element that CRIBB does not implement. Consequently, CRIBB can be extended to perform some experiments with different models of emotion. Certain theories of emotion are more suitable for implementation in CRIBB. Both the appraisal and the goal-oriented approach cite cognition as the core of emotions. The scenarios used in CRIBB are based around a goal-oriented structure. The existence of intentions in CRIBB is comparable to a goal state. Therefore, implementing a goal base and using tenets of the goal-oriented approach to emotion is a suitable foundation on which to base a model of emotions.

Gibson's theory of direct perception [14] can be used to extend CRIBB's perception-belief mechanism to incorporate emotional capabilities. Gibson describes how sensory data when perceived is given affordances and valences. An affordance is something that refers to both the environment and the perceiver in a way that no existing term does. They are properties taken with reference to the observer. Affordances of the environment are what it offers, what it provides, either for good or bad. For example, if a surface is horizontal, nearly flat and sufficiently extended and if its substance is rigid then the surface affords support. Affordances can also be valenced. The theory of affordances can be extended to allow emotion to exhibit an effect on perception of the environment according to the importance of needs, goals and plans to the individual. The following extension to CRIBB does this.

When CRIBB is given a proposition, a belief is inferred from this. The consistency of this belief is checked with the existing set of beliefs. If no contradiction is found then the new proposition is added to the belief set. If there is a contradiction then this is resolved and the most certain belief is added to the belief set. For example:

$$P := \{r, s, q, p\}$$

$$B := \{\neg p\}$$

$$P \otimes B \rightarrow B'$$

$$B' := \{r, s, q, p\}$$

B is the existing belief set and P is perception set. The new set B' contains the system's new belief set with all possible contradictions resolved (p is preferred to $\neg p$). In this scenario each perception of the world has equality.

Rather than attempt to completely and accurately model the agent's world, emotion can be used to guide attention so an agent is drawn to aspects of the environment deemed to be of importance. Assigning an emotional affordance will enable a process by which perceptions can be filtered according to their importance. Hence:

$$P := \{r, s, q, p\}$$

$$E := \{\text{importance}(\text{high}, p), \text{importance}(\text{low}, r)\}$$

$$B := \{\neg p\}$$

$$E \otimes P \rightarrow EP$$

$$EP := \{p, s, q, r\}$$

$$EP \otimes B \rightarrow B'$$

$$B' := \{p, s, q, r\}$$

The perception set, P, contains the same perceptions as before. However, the order in which the perceptions are processed can be changed according to the emotional affordance, E, attached to each one. The new belief set, B', contains the perceptions which have been processed in the order that accords with their emotional significance to that individual.

Emotion can be used to extend the belief and perception mechanism of CRIBB further. Consider a perception received from one source and a further perception, from a different source, that contradicts this. If, through the contradiction mechanism in CRIBB, the first perception is found to be false then this may affect the truth value of any beliefs and perceptions from that particular source. In other words CRIBB will now be less inclined to believe information received from this source. Or conversely, the information from some source may now be considered more reliable than before. This situation can be represented in CRIBB by creating an emotional correspondence for each possible source. This would give an indication of the likelihood of information from this source being either true or false.

Ongoing work on this model is using various agent test-beds to gather metrics to inform our research. One particular experiment makes use of the fungus eater scenario [26, 28]. Early results suggest that the addition of emotion to CRIBB results in a more effective use of resources to achieve tasks, with a more efficient resolution of goal conflicts. For example in an environment evenly populated with fungus and ore, both CRIBB and ECRIBB agents achieve their task goals (the collection of ore). The ECRIBB agents however make more effective use of available energy sources (the fungus). In terms of the earlier arguments about goal conflicts, the emotion model augments the agent's

autonomy and facilitates the resolution of goal conflicts. Experimentation continues to determine how an agent can adapt to environment changes through the modification of the emotional valences associated with perceptual affordance and goal importance.

6 Future Directions and Discussion

The research described here reflects on two perspectives to the integration of emotion into cognitive agents. The architecture of Figure 1 has limited reasoning capabilities (more limited than the implementations described in [9]), but makes use of a coherent emotion-based control language. CRIBB on the other hand is a serial deliberative (BDI) model that does not try to provide a coherent story for all the types of control states identified in [23] and [9]. A complete architecture would subsume the use of emotion as a control language, and indeed the entire BDI reasoning processes of CRIBB. The current separation enables complimentary work to progress independently. As this research develops the two complementary architectures can be integrated.

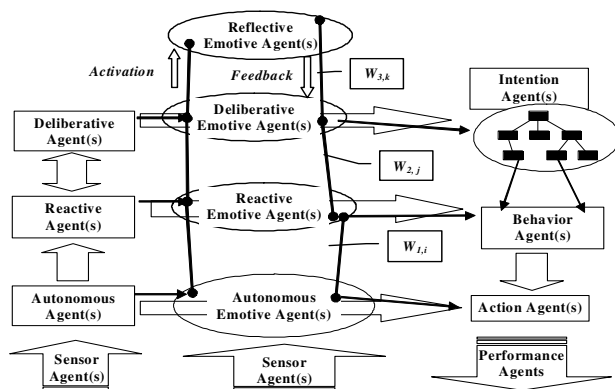


Figure 2. A distributed model that draws together the emotion engine of Figure 1, CRIBB and earlier work.

Duffy's theory that emotions occur in a continuum or a number of continua, which includes both 'emotional' and 'non-emotional' behavior, views emotion as a more integrated part of behavior rather than a separate element. This view is also supported by Scherer who argues the pattern of all synchronised changes in the different components over time constitutes an emotion. Both of these theories can be viewed as 'distributed' models. Using a distributed emotion model is problematic in CRIBB as CRIBB's serial reasoning model is not amenable to the asynchronous re-evaluation of plans and information processing that takes place within a distributed system. No such problem exists with the four layer architecture – it is designed to be asynchronous and distributed.

Oatley and Johnson-Laird propose that each goal and plan has a monitoring mechanism that evaluates events relevant to it. This mechanism broadcasts to the whole cognitive system, allowing it to respond to change as it occurs. For a distributed model of emotion, the monitoring system would need not only to communicate goals and plans but also respond to each sub-system. CRIBB can be readily extended with a central

monitoring system without jeopardising its reasoning model. This module already exists in the architecture of Figure 1 as the deliberative Motivator processes.

As a development of the work described here, distributed versions of the four-layer architecture are being investigated. The extended model includes those aspects described earlier in this paper and other work [9], and is being implemented as a multi-agent society (see Figure 2). In this architecture CRIBB is modeled as a deliberative perception agent and a separate deliberative reasoning agent. Changes to an agent's beliefs are possible through external influence, as in the Castelfranchi model of autonomy [7], using the mechanisms inherent in extended CRIBB (set E of affordances) mediated by ongoing emotional valences of the emotion engine. Exploratory implementations have made use of a simplified version of the motivator structures used in earlier work. Current work looks to formalise the control language based on a computational model of emotion that draws on the Oatley, Frijda and Sherer theories, i.e. a goal-based theory of emotion with modal responses and no basic emotions.

References

- [1] Alterman, R., (2000) Rethinking autonomy, *Minds and Machines*, 10(1):15-30.
- [2] Bartsch, K. & Wellman, H. (1989) Young children's attribution of action to beliefs and desires. *Child Development*, 60, 946-964.
- [3] Buck, R. (1986) The Psychology of Emotion. In J. LeDoux, W. Hirst (Ed.) *Mind and Brain: Dialogues in Cognitive Neuroscience*, Cambridge University Press. Cambridge.
- [4] Brooks, R. A. (1991) Intelligence without representation, *Artificial Intelligence*, 47:139-159.
- [5] Bower, G. (1994) Some Relations Between Emotions and Memory, In P. Ekman and R. Davidson (Ed.). *Nature of Emotion*, Oxford University Press. New York, 303-306.
- [6] Bratman, M.E., Israel, D.J. & Pollack, M.E., (1988) Plans and resource-bounded reasoning, *Computational Intelligence*, 4, 349-355.
- [7] Castelfranchi, C. (1995) Guarantees for autonomy in cognitive agent architectures. In: Wooldridge, M. and N. R. Jennings (Eds), *Intelligent Agents*. Springer-Verlag: 56-70
- [8] Damasio, A. (1994) *Descartes Error: Emotion, Reason and the Human Brain*, New York. Avon Books.
- [9] Davis, D. N. (2001) Control States and Complete Agent Architectures, *Computational Intelligence*, 17(4).
- [10] Duffy, E. (1941) An explanation of 'Emotional' phenomena without the use of the concept 'Emotion', *The Journal of General Psychology*, 25, 283-293
- [11] Elliott, C. (1992) *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System*. PhD thesis, Northwestern University.
- [12] Ferber, J. (1999) *Multi-Agent Systems*, Addison-Wesley.
- [13] Frijda, N. (1987) *The Emotions*, Cambridge University Press, Cambridge
- [14] Gibson, J. (1986) *The Ecological Approach to Visual Perception* Lawrence Erlbaum Associates, New Jersey
- [15] Goleman, D. (1998) *Working with Emotional Intelligence* Bloomsbury, London.

- [16] Izard, C. (1993) Four systems for emotion activation, *Psychological Review*, 100(1), 68-90.
- [17] Matari, M. J. (1997) Studying the role of embodiment in cognition, *Cybernetics and Systems*, 28(6), 457-470.
- [18] Merleau-Ponty, M. (1965) *The Structure of Behaviour*, Methuan, London.
- [19] Newell, A., (1990), *Unified Theories of Cognition*, Harvard University Press.
- [20] Oatley, K. (1992), *Best Laid Schemes*, Cambridge University Press, Cambridge.
- [21] Ortony, A., Clore, G. & Collins, A. (1988) *The Cognitive Structure of Emotions* Cambridge University Press, Cambridge.
- [22] Scherer, K. (1994) Toward a Concept of 'Modal Emotions', In *Nature of Emotion*, P. Ekman, R. Davidson (Eds.) Oxford University Press, New York.
- [23] Simon, H. A. (1979) Motivational and emotional controls of cognition, In: *Models of Thought*, Yale University Press.
- [24] Sloman, A. & Croucher, M. (1987). Why Robots will have emotions. *Proceedings of IJCAI-87*, 197-202.
- [25] Sloman, A. (1999) Architectural requirements for human-like agents both natural and artificial, In *Human Cognition and Social Agent Technology*, K. Dautenhahn (ed.), Benjamins.
- [26] Toda, M. (1962) The Design of a Fungus Eater, *Behavioural Science*, 7, 164-183.
- [27] Wahl, S., Spada, H. (2000) Children's Reasoning about intentions, beliefs and behavior, *Cognitive Science Quarterly*, 1, 5-34.
- [28] Wehrle, T. (1994) New Fungus Eater Experiments. In *P. Gaussier & J. D. Nicoud (Eds), From Perception to Action (400-403)*. Los Alamitos. IEEE Computer Society Press