An Emotion-Based “Conscious” Software Agent Architecture

Lee McCauley[1], Stan Franklin[1] and Myles Bogner[1,2]

Institute for Intelligent Systems and
Department of Mathematical Sciences
The University of Memphis

Abstract Evidence of the role of emotions in the action selection processes of environmentally situated agents continues to mount. This is no less true for autonomous software agents. Here we are concerned with such software agents that model a psychological theory of consciousness, global workspace theory. We briefly describe the architecture of two such agents, CMattie and IDA, and the role emotions play in each. Both agents communicate with humans in natural language, the first about seminars and the like, the second about job possibilities. IDA must also deliberate on various scenarios and negotiate with humans. In CMattie emotions occur in response to incoming stimuli from the environment and affect behavior indirectly by strengthening or weakening drives. In IDA the emotions are integrated with the “consciousness” mechanism, and bidirectionally connected with all the major parts of the architecture. Thus, emotions will affect, and be affected by, essentially all of the agent’s disparate cognitive processes. They will thus play a role in essentially all cognitive activity including perception, memory, “consciousness,” action selection, learning, and metacognition. These emotional connections will provide a common currency among the several modules of the agent architecture. These connections will also allow for the learning of complex emotions. The emotions serve to tell the agent how well it’s doing.

Introduction

Evidence continues to mount for the role of emotions in environmentally situated human agents (LeDoux and Hirst 1986,Damasio 1994, Rolls 1999). The verdict is still out on exactly how the emotions affect cognition in humans, but there does seem to be general agreement that they play a more important role than previously believed. Neuroscientist Walter Freeman sees emotions as “essential aspects of, but subordinate to, intentionality, perception and the construction of meaning.” (1999, p. 96) CMattie and IDA are environmentally situated software agents that will utilize emotional mechanisms to help direct attention, learn from situations that generate “good” or “bad” emotional states, and make evaluations of external stimuli and internal states. Including emotional capabilities in non-biological autonomous agents is not a new idea (Bates et al. 1991, Sloman & Poli 1996, Picard 1997). However, as more information is learned regarding the functions and architecture of emotions in humans, we must constantly reevaluate our models in light of these emerging ideas. One such idea is the notion that emotions are closely linked to almost all aspects of cognition and that they may provide the initial building blocks for conscious thought in infants (Watt 1998).

Partially motivated by recent discoveries of the role emotions play in human cognition, we are reevaluating their role in the cognitive processes of our two software agents, CMattie and IDA. In CMattie emotions affect action selection by strengthening or weakening her primary motivators, her drives. They don’t directly affect perception, behaviors, the process of memory, or metacognition. What’s now known about the role of emotions in human cognition suggests that they should. Also, though emotions affect CMattie’s actions, those actions don’t affect her emotions. Not so with humans. With these issues in mind, we intend to significantly broaden the role of emotions in our more complex agent, IDA. We also intend to see that that role is bi-directional so that each of IDA’s cognitive modules also affects her emotions. In this paper we will describe the high level design for this reevaluation and re-implementation of the role of emotions in “conscious” software agents.

[1] Supported in part by ONR grant N00014-98-1-0332
[2] With essential contributions from the Conscious Software Research Group

CMattie

"Conscious" Mattie (CMattie) (Franklin 1997) is the next incarnation of Virtual Mattie (VMattie), an intelligent clerical agent (Franklin et al. 1996, Song & Franklin forthcoming, Zhang & Franklin forthcoming). CMattie's task is to prepare and distribute announcements for weekly seminars that occur throughout a semester in the Mathematical Sciences Department at the University of Memphis. She communicates with seminar organizers and announcement recipients via email in natural language, and maintains a list of email addresses for each. CMattie is completely autonomous, actively requesting information that has not been forthcoming, and deciding when to send announcements, reminders, and acknowledgements without external intervention. No format has been prescribed for any type of email message sent to her. CMattie implements a version of Hebbian type (temporal proximity) learning, and includes modules for perception (natural language understanding), action selection, metacognition, associative memory, "consciousness" and emotions. Unfortunately, her domain (seminar announcements) with respect to the emotion component may not be rich enough to require the emergence of complex emotions.

CMattie is designed to model the global workspace theory of consciousness (Baars 1988, 1997). Baars' processes correspond to what we call codelets, a name borrowed from Hofstadter and Mitchell's Copycat system (Hofstadter et al. 1994). A codelet is a small piece of code capable of performing some specific action under appropriate conditions. Many codelets can be thought of as autonomous agents, making CMattie a multi-agent system in the sense of Minsky's *Society of Minds* (1986). Almost all actions are taken at the codelet level. Her action selection mechanism chooses the next behavior, which is then implemented by lower-level codelets. These higher level behaviors correspond to goal contexts in global workspace theory. Emotion codelets influence not only other codelets, but also indirectly influence behaviors through drives.

CMattie also has an associative memory based on a sparse distributed memory (Kanerva 1988). A new percept, her understanding of an incoming email message, associates with past experiences including actions and emotions. These remembered emotions, and the percept itself, activate emotion codelets that, in turn, influence current action selection. Thus, CMattie will produce actions, at least partially based on emotional content, and appropriate for the active goal context.

Emotions in CMattie

In humans, emotions seem to play the role of the evaluation network. As well as affecting our choice of actions, they evaluate the results of these actions so that we may learn (Rolls 1999 Chapter 3). Emotions in an agent architecture should serve this same evaluative purpose. In CMattie (McCauley & Franklin 1998) emotion codelets update the gain value, which is used to vary learning rates and valence, and send activation to the drives, which effects the behaviors that the system performs. The gain in CMattie and IDA is not a single value; instead, it is a vector of four real numbers that measure anger, sadness, happiness, and fear. We may later add disgust and surprise (Ekman 1992, Izard 1993), although, for our current purposes the current four seem to suffice. CMattie's domain is narrow enough so that surprise and disgust would not be of great benefit. This may not be the case for IDA, who may well need them added to her repertoire.

The agent's overall emotional state at any one time is the result of a combination of the four (or six) emotions. A particular emotion may have an extremely high value as compared to the other emotions, and, consequently, dominate the agent's emotional state. For example, if a train has blocked the route to your favorite restaurant and you are hungry and in a hurry, your emotional state may be dominated by anger even though many other more subtle emotions may be active at the time. The same type of thing can occur in CMattie and IDA. In such a case the agent can be said to be angry. Do note that the agent will always have some emotional state, whether it is an easily definable one such as anger, or a less definable aggregation of emotions. No combination of emotions are preprogrammed; therefore, any recognizable complex emotions that occur will be emergent.

How exactly does CMattie determine emotional levels? The value of an individual element (emotion) in the gain can be modified when an emotion codelet fires. Emotion codelets have preconditions based on the particular state or percept the codelet is designed to recognize. For example, a message being from the system administration may be a precondition for a codelet that adds to the fear component of the emotion vector producing anxiety in the agent. How does this happen? When an emotion codelet's preconditions are met it fires, modifying the value of a global variable representing the portion of the emotion vector associated with the codelet's preconditions and may send activation to an associated drive. A two step process determines the actual value of an emotion at any one time. First, the initial intensity of the emotion codelet is adjusted to include valence, saturation, and repetition via the formula

$$a = v\frac{1}{1+e^{(-vx+x_0)/1.5}}$$

where a is adjusted intensity at creation time, x is the initial intensity of the emotion, v is the valence {1,-1}, and x_0 is the habituation factor, which shifts the function to the left or right. The x_0 parameter models the short-term habituation of repeated emotional stimuli. Its value is increased when the same stimulus is received repeatedly within a short period of time.

During the second step in the process each emotion codelet that has fired creates an instantiation of itself with the current value for adjusted intensity and a time stamp. This new codelet will add its adjusted intensity value (not to be confused with activation) to the global variable representing its particular emotion based on the formula (modified from Picard 1997)

$$y = ae^{-b(t-t_0)}$$

Where a is the adjusted intensity at creation time, b is the decay rate, t the current time, and t_0 the time of creation of the codelet.

Since the emotion vector is not a single value, a single codelet will only effect one component of the vector, anger, for instance. The overall anger value for the agent would, therefore, be a summation of all of the y values for codelets that fire and that effect the anger component of the emotion vector. In this way each active emotion codelets makes its contribution to the overall emotional state. The emotional state of the agent is written to associative memory with each incoming percept. During the recall process these emotions are remembered and re-effect the emotional state of the agent by instantiating a new codelet in much the same way as an original emotion codelet would. In such a circumstance, the codelet will affect the emotional state of the agent using the previous formula adjusted for the new time of activation and with a degraded initial intensity.

There can be multiple emotion codelets, each with its own preconditions that cause it to fire. The system may fire more than one emotion codelet at a time. The resulting emotional state of the agent, represented by the gain vector, is, therefore, a combination of the recent firings of various emotion codelets. Also, multiple emotion codelets can be included in concept (chunked) codelets (Jackson 1987, Bogner 1999), thereby learning complex emotions that are associated with a higher level concept.

So how does the emotional mechanism affect the behavior of CMattie? First, a message would come to the system and be understood by the perception module. For the purpose of this example, let us say that the message says something to the effect of, "the network will be shut down in two minutes due to a power problem." This type of message would probably come to CMattie as a system alert, but its contents are processed in the same way as the normal email messages. After perception has categorized the message, its processed form is put into the perception registers and then moved (after some higher level processing that takes into account remembered states) to the Focus. A particular emotion codelet will then see the word "shutdown" in the Focus (it does not matter where in the Focus it is), and will increase the fear feature of the emotion vector. At the same time, this same codelet will send activation to the self-preservation drive. Several other things may be occurring simultaneously within the system, including the reading from the different types of memories and the activation of various "conscious" codelets that respond to other items in the Focus. These other codelets may spawn or activate specific behaviors in the behavior net. Along with the previous activations of behaviors in the behavior net and the newly jazzed drive, which will, through spreading activation, increase the likelihood of a self-preservation action being chosen, one behavior will get picked to fire. The probability of a self-preservation action occurring here is influenced by the amount of fear that was generated by the original emotion codelet. After a behavior is chosen and executed, there is a write back to the various memories containing all the elements of the Focus along with the state of the emotion vector and the behavior that was chosen for execution. As previously stated, this will facilitate the association of the emotional state of the system with the current state of the perceptions and the action executed. Learning can then take place at some later time that will take into account the emotions of the system as a guiding factor to how the system was doing at that time and how the associated perceptions might have affected the actions of the agent.

Another short example of how the emotions might affect CMattie's behavior would be to note how her correspondence with a particular person could change over time. In this case, CMattie has not received a message from a particular seminar organizer letting her know who will speak at the "Computer Science Seminar" that week and it is getting close to the time when she must send out that week's list of seminars. She has already sent this organizer two previous reminders that week and not gotten a reply. The fact that there is missing information in the weekly seminar list, and that the time is growing short, has caused her emotional state to be high in the areas of fear (that she will not get the announcement out in time) and anger (that the seminar organizer has not sent her the necessary information). This combination of emotions might, in humans, include anxiety. The high level of anxiety might influence the template that CMattie chooses to send to this organizer. Instead of the normally "polite" version of the information request template, she might choose a more forceful version that more accurately expresses the urgency of getting the organizer's information. Over time, CMattie could recognize that she always gets anxious because of missing information from this same organizer. Her response to this could be to choose to use the more forceful form of the information request template earlier than normal when corresponding with this organizer.

IDA

IDA is an Intelligent Distribution Agent for the U.S. Navy. Like CMattie, she implements global workspace theory (Baars 1988, 1997). At the end of each sailor's tour of duty, he or she is assigned to a new billet. This assignment process is called distribution. The Navy employs some 200 people, called detailers, full time to effect these new assignments. IDA's task is to facilitate this process, by playing the role of one detailer as best she can. Designing IDA presents both communication problems and constraint satisfaction problems. She must communicate with sailors via email in natural language, understanding the content. She must access a number of databases, again understanding the content. She must see that the Navy's needs are satisfied, for example, that the required number of sonar technicians are on a destroyer with the required types of training. She must adhere to Navy policies, for example, holding down moving costs. And, she must cater to the needs and desires of the sailor as well as is possible.

IDA will sense her world using three different sensory modalities. She will receive email messages, read database screens and sense operating system commands and messages. Each sensory mode will require at least one knowledge base and a workspace. The mechanism here will be based loosely on the Copycat Architecture (Hofstadter et al. 1994, Zhang et al. 1998). Each knowledge base will be a slipnet, a fluid semantic net which operates with the help of a workspace (a working memory) to allow perception (comprehension) to occur. The perception process is constructive. Each mode, other than the email and operating systems commands and messages, will understand material from a particular database, for example personnel records, a list of job openings, or a list of sailors to be assigned.

Each of IDA's senses is an active sense, like human vision rather than human hearing. They require actions on IDA's part before sensing can take place, for example reading email or accessing a database. One component of IDA's action selection is an enhanced version of a behavior net (Maes 1990, Negatu & Franklin 1999, Song & Franklin forthcoming). The behavior net is a directed graph with behaviors as vertices and three different kinds of links along which activation spreads. Activation originates from internal, explicitly represented drives, from IDA's understanding of the external world, and from internal states. The behavior whose activation is above some threshold value and is the highest among those with all preconditions satisfied becomes the next goal context as specified in global workspace theory. The several small actions typically needed to complete a behavior are performed by codelets. IDA's behaviors are partitioned into streams, loosely corresponding to the connected components of the digraph, each in the service of one or more drives. Streams of behaviors are like plans, except that they may not be linear. Behavior streams might be interrupted during their execution or possibly not completed. Examples of IDA's streams include Access Personnel Record, Send Acknowledgement, Offer Assignments, Produce Orders.

IDA, like CMattie, is very much a multi-agent system in the Minsky sense (1986), the agents being the codelets that underlie all the higher level constructs and that ultimately perform almost all of IDA's actions. We've mentioned the codelets that underlie behaviors. Others underlie slipnet nodes and perform actions necessary for constructing IDA's understanding of an email message or of a database screen (Zhang et al. 1998). Still other codelets play a vital role in the "consciousness" mechanism.

Having gathered all relevant information, IDA must somehow select the assignments she'll offer a given sailor. Being a constraint satisfaction problem, considerable knowledge is required for making these selections. Much of this knowledge is housed in an operations research type linear functional that measures the suitability of a particular billet for a given sailor. The rest of this knowledge is found in codelets that effect the deliberation process. This process creates and evaluates a scenario to check the temporal fit of a transfer of the given sailor to a particular new billet.

IDA employs a number of different memories. The offer memory is a traditional database that keeps track of the assignments IDA has offered various sailors. For cognitive modeling purposes this memory can be considered to be external, comparable to a human keeping notes. IDA's intermediate term memory acts as an episodic memory, providing context for email messages and for the contents of database screens. It'll be implemented as a case-based memory to facilitate case-based learning. IDA's associative memory associates memories, emotions and actions with incoming percepts as well as with internal events such as deliberations. It is implemented by an extension of sparse distributed memory (Kanerva 1988). Some of IDA's action selection codelets act as a kind of template memory holding the various small text scripts that IDA uses to compose commands to access databases or issue orders, and to compose messages to sailors.

Global workspace theory postulates the contents of "consciousness" to be coalitions of codelets shined on by a spotlight. Imagine a codelet workspace populated by many active codelets each working, unconsciously and in parallel, on its own agenda. Coalitions of codelets that arise from novel or problematic situations seek out the spotlight. The information contained in this coalition is broadcast to all the other codelets, active or not. The idea is to recruit resources in the form of relevant codelets to help in dealing with the novel or problematic situation. It seems that in humans almost any resource may be relevant depending on the situation. Global workspace theory asserts that consciousness attacks the problem of finding the relevant resources by brute force. Broadcast to all processes. IDA uses this method.

To do so, she needs a coalition manager, a spotlight controller, a broadcast manager and "consciousness" codelets (Bogner et al. in press). The coalition manager groups active codelets into coalitions according to the strength of the associations between them, and keeps track of them. If a collection of codelets is associated above a certain threshold level, these codelets are considered to be in a coalition. The spotlight controller determines which coalition becomes "conscious" at a particular time. It calculates the average activation level of each of the coalitions by averaging the activation levels of the coalition's codelets. The spotlight then shines on the coalition with the highest average activation level. Once the spotlight controller has determined a "conscious" coalition, it notifies the broadcast manager that is responsible for gathering information from the "conscious" coalition, and sending it to all of IDA's codelets. As prescribed by global workspace theory, messages are small and understood by only some of the agent's codelets. Specifically, the broadcast manager gathers objects labeled for broadcast from the codelets in the "conscious" coalition. These objects contain information needed for specifying the current novelty or problem. This information is then broadcast to all of IDA's codelets.

"Consciousness" codelets play a critical role in this process. A "consciousness" codelet is one whose function is to bring specific information to "consciousness." Specific "consciousness" codelets spring into action when the information from perception is relevant to them. Some "consciousness" codelets check for conflicts among the relevant items returned from the percept and the memory. "Consciousness" codelets are designed to recognize and act on the kinds of novel or problematic situations that should be brought to "consciousness."

Emotions in IDA

IDA's emotion module, like a human's, provides a multi-dimensional method for ascertaining how well she is doing. We will experiment with mechanisms for emotions. These may include anxiety at not understanding a message, guilt at not responding to a sailor in a timely fashion, and annoyance at an unreasonable request from a sailor. Emotions in humans and in IDA influence all decisions as to action (Damasio 1994).

IDA's emotional system will need to be a good bit more robust than CMattie's. In addition, IDA's emotions will be more tightly integrated with her "consciousness" mechanisms. This implementation attempts to model an emerging perspective on the relationship between emotions and the rest of cognition, in particular, the "consciousness" module. Arguments have been made that the concept of a "limbic system" that is largely separate from the rest of the brain is at best misleading. It has been suggested that the distinction between areas of the brain that are considered limbic and non-limbic cannot be made due to the incredible interconnectedness and pervasiveness of the limbic system (Watt 1998). In other words, it is not clear what aspects of cognition are being conducted with or without the aid of emotions. It seems that each time we learn a bit more about how the brain processes emotions, we are forced to reevaluate our notions of what makes emotional cognition different from all other cognitions. As Aaron Sloman has pointed out on numerous occasions, the functions that emotions seem to play can be accomplished with a complex pattern filter and alarm system (1987). CMattie is an example of just this sort of system.

What happens in humans, however, seems to be *much* more complex than what an alarm type system could produce. The first step in trying to model this complexity in IDA will be to meld portions of the emotion and "consciousness" mechanisms borrowed from CMattie. As described above, IDA's "consciousness" module depends on codelets that each look for some particular event. These codelets, upon recognizing their preprogrammed event, activate themselves and "helper" codelets and attempt to get into the spotlight of "consciousness". With the exception of actively trying to reach "consciousness" and recruiting other codelets, emotion codelets look very much like "consciousness" codelets, often even looking for the same event. To meld these two tasks we simply add to some of the "consciousness" codelets the ability to change the emotion vector, and link their activation to the amount of emotional change produced by that codelet.

The next step attempts to provide massive interconnectedness between the emotional mechanisms and the rest of the major cognitive areas of the system. A network is built up by connecting the "consciousness"/emotion codelets to key behaviors, goals, drives, perception codelets, etc. The links of this network are to have weights and carry activation. Weights will decay with disuse. Each use tends to decrease the decay rate. Weights will increase according to a sigmoidal function of any activation carried over the link, allowing for Hebbian style learning. The product of weight and carried activation is added to the activation already present at the head of the link. Spreading activation then becomes the common currency that integrates the separate modules that use these constructs.

This formulation provides several benefits that conform to current evidence regarding emotions. First, emotions seem to be pervasive in the human brain, linking activities that occur in relatively distant areas with a consistent value judgement. By connecting all of the major systems of IDA, we allow activities that occur in one area to effect what happens in other areas even though they seem quite disparate. Second, complex emotions can be learned over time based on their temporal cooccurrence with a situation that results in a particular emotional response. By maintaining the functionality of the emotional mechanism in CMattie, IDA will still be able to learn complex emotions through remembered emotional states and through association with other codelets that are in "consciousness" at the same time (Bogner et al. in press). Also, the activation network as described above will allow for the learning of link weights thereby affecting the influence of emotions on actions, and vice versa. Finally, emotions can both trigger certain reactions and can be triggered by actions. In humans, for example, the physical act of singing a joyful song, even forced, can make a person temporarily feel better when they are sad. In IDA, the spreading of activation in the emotional network can occur in any direction. Thus a "consciousness"/emotion codelet can trigger an active response. In the other direction an action can trigger an emotional response or even instigate a group of codelets coming into "consciousness" that would otherwise not have been triggered.

Note that it is not the mechanisms used here to model emotions, codelets, spreading activation, etc., that are unique, but the *way* that these mechanisms are being used, and the functions that they perform within the larger system. The point is that the mechanisms involved in emotions in humans or otherwise are not "special." The same can be said about most any other part of the human brain. The visual cortex, for instance, uses the same basic mechanisms as the temporal lobe yet play different roles within the system. It is not their substance but their organization and interconnections that set them apart. On the boundaries between areas of the brain, the naming of one neuron as being part of a module and not part of another is not only counter productive but impossible. We give names to different parts or functions of areas of the brain because it is easier to comprehend and analyze them that way, but we must keep in mind that they are all still part of the same highly parallel, highly integrated system. Emotions are much like any other cognitive function and should not be viewed as magical or special – they just perform a different function.

Related Work

Several other authors have proposed systems that are comparable (Canamero 1997; Velasquez 1997, 1998). The most similar of these methods is Velasquez's Cathexis, which has been used in several agent systems.

As with Cathexis, the emotion mechanism in IDA is tightly interwoven into the codelets (nodes for Velasquez) of the various systems. This interconnectedness affects the behavior of these systems differently, however, because of the nature of the implementations. In IDA, for example, we use a different mechanism for perception than we do for action selection. This means that the two modules cannot directly "talk" to each other and must pass information through the focus. The emotion mechanism provides a direct way for separate modules to affect each other's operation. For CMattie, emotions affect the behavior of the system through the activation of appropriate drives which, in turn, effect the activation level of behaviors in the behavior net.

Another difference between these systems involves the way that emotions are remembered. In IDA, the emotion vector can be remembered in associative (Sparse Distributed) and case based memory. This stores the current state of the vector along with the state of the focus and the behavior chosen for activation. It is important to note that this associates the whole vector with the state of the system at a particular time, it does not associate the change in the emotion vector with the element that triggered that change. IDA can also remember emotional activation when individual emotion triggers become associated with their effects over time via the links between emotion codelets and other codelets that tend to be active at the same time.

Both Canamero and Velasquez's work shares with our mechanism the aspect of using emotions to bias the other cognitive elements of the system. This function serves two purposes; the first being the ability to assist cognitive judgements by providing a priming for paths that lead to emotionally desirable outcomes. The second purpose is the ability to facilitate default responses to perceptions or stimuli to which the cognitive mechanisms have not yet been able to react. In both cases the emotions can be viewed as telling the agent how well it's doing.

Conclusion

In light of continuing developments in the field of emotion research and functional gaps in the emotional architecture for CMattie, modifications to this architecture have been described for use in IDA. It is not yet known, since IDA is not yet fully implemented, how these enhancements will ultimately affect the operation of the system. This new architecture will provide IDA with an emotional mechanism that more closely resembles that of humans. These mechanisms are more closely integrated with the "consciousness" mechanism, and connected to all of the major parts of the system in a bi-directional manner. Thus, emotions will affect, and be affected by, essentially all of the agent's disparate cognitive processes. For this reason, emotions will play a role in essentially all cognitive activity including perception, memory, "consciousness," action selection, learning, and metacognition and will provide a common currency among the several modules of the agent architecture. These connections will also allow for the learning of complex emotions through the refinement over time of the weights on the links.

Finally, must a system have emotions for intelligence? The same question could be asked of almost any other function of the brain. Must a system have sight, particular reactive modules, or metacognition? There is probably no requirement for any one particular piece or function at this specific level. Although a system may exhibit *more* intelligence with the addition of emotions, it is probably not a necessary component. With that said, we still believe that the addition of emotions to a system can be very useful both for interacting with emotional humans and for the enhancement of cognitive abilities.

References

Baars, B. J. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baars, B. J. 1997. *In the Theater of Consciousness*. Oxford: Oxford University Press.

Bates, J. Loyall., B. A., and Reilly W. S. 1991. *Broad Agents. In: Proceedings of the AAAI Spring Symposium on Integrated Intelligent Architectures. In: SIGART Bulletin, Volume 2, Number 4, August 1992*. Stanford University.

Bogner, M. 1999. Realizing "consciousness" in software agents. Ph.D. Dissertation. University of Memphis.
Bogner, M., U. Ramamurthy, and S. Franklin. in press. Consciousness" and Conceptual Learning in a Socially Situated Agent. In *Human Cognition and Social Agent Technology*, ed. K. Dautenhahn. Amsterdam: John Benjamins.
Canamero D. 1997. Modeling Motivations and Emotions as a Basis for Intelligent Behavior. In: *Proceedings of AGENTS '97*. New York: ACM Press.
Damasio, A. R. 1994. *Descartes' Error*. New York: Gosset; Putnam Press.
Ekman, P. 1992. An Argument for Basic Emotions. In *Basic Emotions*, ed. N. L. Stein, and K. Oatley. Hove, UK: Lawrence Erlbaum.
Franklin, S. 1997. Global Workspace Agents. *Journal of Consciousness Studies* 4:322–334.
Franklin, S., A. Graesser, B. Olde, Song H., and A. Negatu. 1996. *Virtual Mattie--an Intelligent Clerical Agent. AAAI Symposium on Embodied Cognition and Action, Cambridge MA*. : November.
Freeman, W. 1999. *How Brains Make Up Their Minds.* London: Weidenfeld and Nicolson.
Hofstadter, R. D., and Mitchell M. 1994. *The Copycat Project: A model of mental fluidity and analogy-making. In: Advances in connectionist and neural computation theory, Vol. 2: Analogical connections, eds. K. J. Holyoak & J. A. Barnden*. Norwood N.J.: Ablex.
Izard, C. 1993. Four Systems for Emotion Activation: Cognitive and Noncognitive Processes. *Psychological Review* 100:68–90.
Jackson, J. V. 1987. Idea for a Mind. *Siggart* Newsletter, 181:23–26.
Kanerva, P. 1988. *Sparse Distributed Memory*. Cambridge MA: The MIT Press.
LeDoux, J. E. andHirst, W. 1986. *Mind and Brain: Dialogues in Cognitive Neuroscience.* Cambridge: Cambridge University Press.
Maes, P. 1990. How to do the right thing. *Connection Science* 1:3.
Minsky, M. 1986. *The Society of Mind*. New York: Simon & Schuster.
McCauley, T. L., and S. Franklin; 1998 An Architecture for Emotion; AAAI Fall Symposium Emotional and Intelligent: The Tangled Knot of Cognition"; AAAI; Orlando, FL.
;Negatu, A., and S. Franklin; 1999 Behavioral learning for adaptive software agents. Intelligent Systems: ISCA 5th International Conference; International Society for Computers and Their Applications - ISCA; Denver, Colorado; June 1999.
Picard, R. 1997. *Affective Computing*. Cambridge MA: The MIT Press.
Rolls, E. T. 1999. *The Brain and Emotion*. Oxford: Oxford University Press.
Sloman, A. 1987. Motives Mechanisms Emotions. *Cognition and Emotion* 1:217–234.
Sloman, A., and Poli R... 1996. *SIM_AGENT: A toolkit for exploring agent designs in Intelligent Agents. In: Vol. II (ATAL-95), eds. M. Wooldridge, J. Mueller & M. Tambe*. : Springer-Verlag.
Song, H., and S. Franklin. forthcoming. A Behavior Instantiation Agent Architecture.
Velasqueze, J. 1997. Modeling Emotions and Other Motivations in Synthetic Agents. In: *Proceedings of AAAI-97*.
Velasqueze, J. 1998. Modeling Emotion-Based Decision-Making. In: *Proceeding of 1998 AAAI Fall Symposium*, Technical Report FS-98-03. Menlo Park, CA: AAAI Press.
Watt, D.; 1998 Emotion and Consciousness: Implications of Affective Neuroscience for Extended Reticular Thalamic Activating System Theories of Consciousness; http://www.phil.vt.edu/ASSC/esem4.html; On Line Web Seminar at the Association for the Scientific Study of Consciousness.
Zhang, Z., and S. Franklin. forthcoming. Metacognition in Software Agents Using Fuzzy Systems. .
Zhang, Z., S. Franklin, B. Olde, Y. Wan, and A. Graesser. 1998. Natural Language Sensing for Autonomous Agents. In *Proceedings of IEEE International Joint Symposia on Intellgence Systems 98*.