Emotion-based Control Mechanisms for Agent Systems

Charles J. Hannon
Texas Christian University
TCU Box 298850, Fort Worth, TX 76129
CHannon@tcu.edu

Keywords: Agent-Based Systems and Designs, Cognitive-based Agent Architecture, Pervasive Computing

Abstract
Emotions in biological organisms are often related to control mechanisms which cause both task and system-level changes. This paper describes work-in-progress to improve the control mechanism of several of our cognitive-based agent models by using such an emotion-based approach. When combined with our current models for attention, arousal, vision processing and natural language use and learning, using our Goal Mind agent environment, early results demonstrate a marked improvement in the models’ ability to fuse language and vision input in a sensor rich environment.

INTRODUCTION
In biological systems, the existence of emotions can have both advantageous and detrimental effects on the cognitive process. This is especially true for the primary emotions (satisfaction-happiness, aversion-fear, assertion-anger, and disappointment-sadness [13]). While these primary emotions can have some detrimental effects on logical reasoning, the current psychological and neurophysiological evidence indicates that they are so closely associated with cognition that it is nearly impossible to isolate the two aspects within a complex biological system. We propose that emotions should not be ignored within a cognitive-based agent system since they can serve as a guide to how the overall control mechanism of the agent should function.

Here, we will discuss the neuroscience background and preliminary modeling of an agent control mechanism based on emotion. To avoid any unnecessary entanglement in philosophical controversy, we will assume for the current discussion that our model of emotional control is merely an artificial simulation of existing neurological models of emotions in biological systems, since an intelligent agent having a mechanism for an emotion does not necessarily imply that that agent has that emotion. If such a complex expression of consciousness as a true emotion can emerge from the software and hardware of an artificial agent, it will clearly be the result of a more ambitious effort than the one being reported here. However, the use of a flexible modeling environment (like our Goal Mind/Alchemy system) means that we can continue to build toward that level of model complexity and may someday be able to answer whether we are simply simulating or actual beginning the process of generating true emotion.

Our current model also avoids some of the psychological entanglements of using emotions to guide multiagent interaction. While not a major focus of our current work, our past research with learning and using natural language has addressed, at a very basic level, this ability and we will show how we plan to incorporate this previous research into future models.

RELATED WORK
A number of efforts to incorporate emotions as a control mechanism in an overall agent model are ongoing [2, 14, 17]. However, [15] points out that unless an approach is cognitively deep enough to be explanatory, they may do little overall good. Our approach is based on such explanatory mechanisms and differs from most other approaches in that we are looking not only at how emotions can determine the reasoning strategy being used by the Higher Order Processes (HOPs), but also how they can control the amount and type of information being presented to the HOPs by the periphery processing layer.

In [5], we proposed a cognitive unified model for language and visual processing used to improve an agent’s sensor fusion and communication ability. This model, built using our Adaptive Modeling environment for Explanatory Based Agents (AMEBA) system [6], used a control mechanism made up of a set of knowledge-base systems designed to emulate the HOP processes of the frontal cortex. It did not directly address the effect of emotion but did rely on AMEBA’s use of Stimuli Routing Networks which were designed to support the global effects of neurotransmitters, and thus, had the ability to emulate the effect of emotions on these neurotransmitters.

In [7], a second related model was built using the same basic control approach to examine a possible cognitive model for the Stroop Effect. The major purpose of this model was to demonstrate that different models could be built by reusing what were proposed to be explanatory components of the first model. While it was relatively successful at demonstrating the basic idea of reusable ‘brain-area’ components, it clearly pointed out that a generalized control mechanism needed to be more than a set of rule bases.

In [8], we demonstrated a biological inspired approach for improving an agent’s sensor fusion and input filtering by modeling animal attention and arousal mechanisms. This
model used our Alchemy/Goal Mind modeling system which is similar to AMBEA but allows models to be geographically distributed over the Internet [9]. Since we were focusing on the filter and fuser components, the control mechanism of this model was less focused on being purely explanatory, and thus, relied on a similar set of knowledge-base systems, designed to emulate HOPs, even though this approach was clearly starting to show signs of serious weakness as a generalized model of control. Despite this, the model was able to both filter and fuse the resulting data into a workable set of critical input, thus allowing a large number of simple sensors to provide constant input about the changing world around the agent.

To address the generalized model of control, it was clear that something more than a set of logical rules had to be used. Early research had indicated the usefulness of wide-area brain affectors like neurotransmitters, but it was unclear how such a mechanism should be controlled. The need for this kind of control drew our attention to the effect of emotion on cognition, and thus, to the current model. In this model, we limit the type of sensory input being used by the agent to that similar in complexity to a biological system. This is necessary to support both the existing attention and arousal model being incorporated into the current model and to provide an environment for the control mechanisms that closely approximates the biological system being used as a pattern. As with our other current models, this model is implemented using our Goal Mind/Alchemy environment.

EMOTION AS A CONTROL MECHANISM

While the philosophical debate over the relationship of cognition and emotion rages, most evidence from nature indicates that regardless of what ontological significance you give what psychologists and animal behaviorists call emotion, the state of being happy, fearful, angry, or sad does indeed effect the cognitive function of a biological system. Of these primary emotions, the easiest on which to learn is fear, since it occurs in a specific context. However, the phrase ‘frozen with fear’ is far more than a simple metaphor. Animals are often so torn between a desire to fight or flight, that they are incapable of doing either. Of first concern in designing a control mechanism based on fear is to try to get the benefits of such a strong protection mechanism without incurring the risk that it will generally make the agent’s actions less beneficial.

To do this, we need to first understand the location and process of fear generation, learning and effects in a biological system. We will concentrate here on humans, with the understanding that much of the research on fear has by necessity been conducted on other animals.

The actual starting location of a fear stimuli is clearly modal but otherwise still unclear. Sensory input routed through the sensory cortex or the sensory thalamus and arriving at the amygdala appears to drive most fear responses [4]. However, other HOPs and even memory (via the hippocampus) can trigger a fear response [10]. The amygdala appears to have different neural structures for different sensor modalities which means that it can both recognize the type of modalities triggering the response and process different modalities in parallel.

In the amygdala, a fear stimuli can trigger a number of basic responses. How these are divided and categorized is subject to some debate [10, 15]. Here, we will assume four basic responses: 1) directed behavior, 2) brain process modification, 3) autonomic and 4) endocrine. Directed behavior works through the normal neural pathways of the brain to stimulate areas in the neocortex. While some special neural mechanisms, like the gated dipole [11], appear to be in play in this type of response, it is by far the most reason-based response giving room for other needs and desires to interact with the response. Brain process modification for all emotions is facilitated by the release of neurotransmitters and neuroactive peptides. While we currently can not completely categorize the parts of the brain can release these, it is clear that midbrain systems directly driven from the amygdala have this capacity. Once released, these chemicals quickly modify the way neural activities occur, increasing the rate of learning, and memory and sensor processing. Autonomic responses pass from the amygdala through the hypothalamus to the body where they generate increased respiration, heartbeat and muscle tension in what can best be described as ‘a get ready to do something now’ reaction. An endocrine response does for the whole body (including the brain) what a brain process modification response does solely for the brain. Hormones and peptides are released into the bloodstream where they further prepare the body to act on a fight or flight command from the brain.

The association of fear with a set of input stimuli is learned in much the same way as other pieces of brain information. Since this learning primarily happens in midbrain, once an association is learned, it can be very difficult (if not impossible) to unlearn it. The major complexity in learning a process to alleviate fear is that the goal state is achieved by finding a state were the input stimuli are reduced. Such reliance on the absence of input requires a special learning mechanism, like the gated dipole, and is highly susceptible to error.
Anger

Anger is closely related to stress, and as such, can be a very damaging process in a biological system [13]. However, anger also serves as an important control mechanism when an animal decides it is necessary to fight rather than escape in a dangerous situation. The locus of generation and the basic neural effects of the anger emotion are similar to that of fear [15]. The basic difference between anger and fear is the overall effect it has on directed behavior and brain process modification. Fear tends to improve general sensory attention, while anger tends to reduce it. Anger also tends to reinforce actuation processing and certain automatic response mechanisms like hand-eye coordination, which makes it easier for the biological system to act using very directed automatic tasks. This is one of the reasons that militaries have throughout history stressed combat training.

The association of anger with a set of input stimuli is learned in much the same way as fear. As with fear, the midbrain location of learning makes it very difficult to unlearn an anger-stimuli association. However, some have claimed there is a greater association between things hated by a society than those things feared by a society [3]. If this is true, HOP processing may have a greater effect on the control of anger than the control of fear, thus making an anger response more easy to ‘rationally’ control.

Happiness and Sadness

Neither the emotion of happiness nor sadness are as well understood as fear and anger. A heightened state of sadness can cause serious negative effects on a biological system’s well being, but this does not mean that it does not also serve useful purposes. Without it, a biological system would not be driven to take steps to avoid certain actions which create a state sadness. It would also not understand the need to improve aspects about itself that cause it sadness. On the other hand, without happiness, a biological system would have trouble understanding when it has arrived at a sufficient solution to a goal.

Happiness appears to reside as both a limbic system process and as a generalized effect of neurotransmitters, while sadness appears to be almost totally related to neurotransmitters [13]. As control mechanisms, both seem to be primarily related to the overall performance of the brain, but sadness also tends to allow a greater activation of memory and other cognitive processes which cause introspection. The mechanism for overall brain performance is clearly related to neurotransmitter control, but the other effects of sadness on brain processes can not be clearly explained by current neuroscience research.

While happiness and sadness are almost universally recognized as separate emotions, the association of happiness and sadness with a set of input stimuli is often inversely related. This can be seen not only from introspection, but also in the way certain theories, such as the gated dipole, deal with them [11]. The actual mechanism for learning happiness and sadness is again unclear and somewhat controversial.

EMOTION CONTROL MODEL

While some interesting applications of happiness and sadness as control mechanisms seem clear, we currently do not have the existing stock of model components necessary to build a unified model with the complexity that could gain from such a mechanism. Using anger as a control mechanism also presents a similar problem since our existing models have focused mostly on sensory input and communication, and not actuation control. Therefore, we focus the model discussion here on the use of fear.

Based on the neurological research into emotion of fear presented above, we propose a computational model that attempts to emulate how fear can constructively control both attention and arousal in a dangerous situation for the agent. Since the existing model of attention and arousal has not yet been extended to generate anything more than simulated actuator responses, there seemed to be little point in attempting to address either the autonomic or endocrine responses, since these would have to be carefully tested on real actuators to determine any useful gain. However, the learning of associations between fear and a set of input stimuli can be addressed in the existing model.

Details of our attention and arousal model can be found in [8]. In summary, it is based on a hierarchical collection of three types of modules: 1) Time Filter/Fuser (t), 2) Space Filter/Fuser (s), and 3) Mixed Modality Filter/Fuser (m). Each module accepts sensor input, an arousal threshold, and output state directives. Output is either directly related to input (attention driven) or threshold driven (arousal) based on state. Under directions of the sensor processing element the attention control mechanism sets most modules to an arousal state so the sensor processing element could attend to a single input stream. When a threshold was breached the arousal control informed the attention control which then switched attention to that input stream. The sensor processing element could then either continue using the new input stream or inform the attention control to switch it back to the previous stream. If a number of unwanted switches occurred, the attention control mechanism could ask the arousal control to adjust its thresholds to avoid further such switching. The threshold could also be adjusted down to improve the chance of switching.

To add a fear-based control mechanism (FBCM), we need to simulate both directed behavior and brain process modification. Thus, the FBCM taps into the arousal control’s threshold stream and sensor processing element’s in-put data stream. If the FBCM detects a possible fear association from the input data stream, it takes over control of the attention control mechanism by inhibiting the sensor processing element’s path to this element. It then redirects attention to all input data streams it has associated with the given fear and directs the sensor processing element and
other HOPs to examine for the associated safety conditions. If the condition is not found, the HOPs order the FBCM to release the data stream control. If it is found, the FBCM sends an inhibit signal to block all unrelated stimuli coming to the other HOPs.

If a dangerous condition is detected by another HOP and that HOP informs the FBCM, the FBCM captures the characteristics of the current input stream to create a new fear association and associates this association with the given dangerous condition. Like the sensor processing element, the FBCM has the ability to order an adjustment to thresholds. If it detects a threshold or a partial association within the data stream, it inhibits the attention control mechanism’s path to the arousal control mechanism and adjusts the thresholds to search for the complete association. If it finds it, it takes over complete control of the attention control mechanism as described above, otherwise it resets the thresholds and deinhibits the attention control mechanism’s path to the arousal control mechanism.

THE GOAL MIND ARCHITECTURE

The Goal Mind system is the next generation of our AMEBA architecture [6] and part of the Gold Seekers project, a set of tools for both general distributed application and multiagent intelligent system design and implementation. The first of these tools is Alchemy [9], a distributed processing environment which supports: 1) the asynchronous processing model needed by our cognitive-based approach, 2) a GUI-driven dynamic generation, operation and testing environment, and 3) a multi-level security facility for safe operation over the Internet or other public networks. The second tool is Goal Mind, a redesign of AMEBA to run on top of Alchemy.

Goal Mind attempts to capture the explanatory force of a connectionist neural model while allowing the use of the better-understood representation and reasoning methods of symbolic AI. From a system perspective, it provides processor transparency within a parallel system and a flexible method of process and knowledge management. The key element that supports these requirements is the etheron which supports: 1) a standard way to load and store knowledge, 2) interfaces to a set of predefined management tools and 3) a generalized set of communication channels for talking with other etherons.

Goal Mind models draw their explanatory depth from the environment’s ability to support hierarchical cognitive processing. Using adaptive distributed processing and generalized inter-process communication, cognitive functions can be modeled at different levels of abstractions without changing the logical relationship between these functions. Thus, a function like the conceptual reasoning about the world and self can be simulated with a reasoning and knowledge storage system which has far less capacity than that of a real human. This allows us to preserve the overall model’s explanatory depth, as long as we preserve explanatory relationships between cognitive components.

THE GOAL MIND IMPLEMENTATION

Goal Mind supports a set of Representation and Inference Mechanisms (RIMs), Coded Response Mechanisms (CRMs) and InterFace Nodes (IFNs) which make the design, implementation and testing of a model easier. These components also support reuse. Figure 1 presents a high-level view of one agent in our current test model. In this agent, the IFNs can either emulate sensor input or condition real sensor input used by the agent. The filter/fuser components are implemented as CRMs. The Attention Control, Arousal Control, Sensor Processing, FBCM and HOPs are implemented as RIMs using either a knowledge-base or a semantic-network reasoner.

At start-up, a set of input-fear associations are loaded in the FBCM, the Attention system sets the input stream to simple temperature monitoring and all baseline threshold values are set by the Arousal Control component.

![Figure 1. Implementation of the Current Test Model](image-url)
For testing of the FBCM, we are using the basic configuration of sensors reported in [7]. The model divides sensor input into five rooms, with two temperature and two light sensors per room. In the sensor processing rule-bases, the second sensor per modality per room is normally viewed as redundant to the first and provides a certain level of fault tolerance regarding the sensor data coming from each room. This model allows us to study the agent’s ability to reason about the cause of temporal, spatial and dual-modality temperature and light relationships from the environment. The pre-FBCM model has been used to examine much greater sensor input complexity, but this basic model provides a simpler domain for testing new components like the FBCM.

The hardware testing environment for the FBCM is the same as for all of our other models. The main processing element of this testbed is a 26 node non-heterogeneous Beowulf processing cluster with a combined processing speed of 26.3 GHz using a channel-bonded network with a throughput of 600 MB/s. This cluster is connected to via a VPN to a smart home lab consisting of a number of X-10 controllers and a wireless network to a set of mobile robots including a Pioneer 2AT and six-leg walker robot of our own design (still under development).

GOAL AND INITIAL RESULTS

Upon completion of testing of the basic model of attention and arousal reported in [8] within a controlled simulation environment, we had expected the next step to be the application of this model against a real-world smart home environment. However, we found that to handle real-world data, the existing model needed a far more sophisticated control mechanism to improve its adaptive nature. Therefore, we are now using our original attention and arousal model to drive the FBCM.

The ultimate test of the FBCM and the rest of the current model is its ability to deal with real-world data and protect the agent and others from real-world dangers. The test platform for testing is our smart home lab which is connected via the Internet with our Beowulf cluster. In initial test, this new FBCM augmented model has greatly improved our existing model’s ability to process real-world data within our testbed by adding both a fine-tuned attention spotlight control and a realistic associative learning mechanism. It also has demonstrated that the additional processing complexity of the new RIM and IFN elements needed to support it are far outweighed by the overall performance of the model. (After all, there is little point to a fast system that does not work.) However, since neither the original or the extended model have sufficiently taxed the processing capacity of our Beowulf cluster, we have not completed any quantitative study of the two models’ relative speed or speedup performance.

The overall performance of FBCM is tied to the original model. In simulation, this original model has been shown to improve both the overall performance of an agent in a sensor-rich environment and an agent’s ability to detect and respond to critical input that is currently not in its attention spotlight. Without this working model of attention and arousal there would have been nothing for the FBCM to control. In retrospect, the original model’s inability to work with real-world data demonstrates the need for just such an explanatory control mechanism that the FBCM provides and demonstrates our basic assumption that for an agent system to truly support intelligent behaviour, a number of explanatory models must be the integration into a larger more capable agent model. Using the FBCM, our larger model is already showing the ability to learn new arousal strategies and adapt to changing threats in the environment, but this work still needs to be tested further.

Just as there is clearly a limit to the amount of sensory input a centralized reasoner can timely process, and thus, a need for arousal strategies that can be automatically switched between the input that needs processing, there is a limit to number of predefined arousal strategies one can incorporate into a rule-based system before the control mechanism is buried in possibilities. Therefore, we are currently focusing on the FBCM’s learning mechanism which has already shown the ability to add and delete strategies as required.

To summarize what our research with the FBCM model has to this date demonstrated, it has shown: 1) that some kind of adaptive control of attention and arousal is need to fully gain from the filtering provided by these, 2) that the FBCM does support adaptive control within an explanatory context, 3) that the Goal Mind implementation of the FBCM model has sufficiently good enough performance to not be process bound in a cluster environment.

FUTURE WORK

While our current research has shown that a sophisticated control mechanism is needed to support an attention and arousal system, it has not yet shown that the FBCM is the only or best solution. We are clearly dealing with the leading-edge of our understanding of cognition here and the way emotions are used as a model of control needs to be further explored. However, the need for something like the FBCM in our current model introduced an unexpected turn in our research since we had not expected to examine emotion for some time. It was only after discovering the usefulness of emotion as a mechanism for control, that we decided to take on what is arguably one of the most controversial aspect of current AI research.

Having ventured down this road, we now hope to incorporate the emotions of anger, happiness and sadness in an overall agent model to improve not only sensor input management but also the integration of our earlier models of language use and learning within the current sensor-rich environment of smart home. Since the emotion of anger seems best suited for controlling actuators, we are also examining how it could be used by mobile robots in both the smart home and other application domains.
CONCLUSION

While still in the early stages of testing, we are already seeing that adding a FBCM to our model of attention and arousal improves the performance of the existing model. It is also beginning to shed some light on the actual cognitive processes on which the model is based and opening the very exciting possibility of using emotions to improve the cognitive performance in other applications. With the addition of an explanatory FBCM to our attention and arousal model, we believe that the Gold Seekers project is, once again, taking yet another small step toward a unified understanding of cognition, emotion and consciousness.

We conclude with a natural question arising from our current research with the FBCM. That is, will the addition of other emotional models for such emotions as anger, happiness and sadness also positively impact our generalized models of cognition? This is a very telling question since to get directly to the heart of what the FBCM is actually doing in our current model. Is it just a clever control mechanism with some explanatory foundations or is it really a rudimentary emulation of the emotion of fear?

As alluded to in the introduction, the emulation of a signal emotional response in an agent should never be claimed to give that agent that emotion. There are just too many possible routes for mistakes in the process of capturing a simple emulation of a single emotion. But what if the model grows to the point that it can emulate a number of conscious activities like fear, anger, happiness and sadness at the same time using the same collection of cognitive and emotional components? If you accept such thought problem arguments as the ‘Chinese Room’ as proof that a non-biological agent cannot have consciousness, this addition of skills will not impress you, but why not? When will it be acceptable to stop believing in the skeptics’ theoretical arguments and start believing what a complex implementation of a non-biological agent’s cognition and emotion is suggesting is possible. While we would never suggest that our research will be the sole catalyst for forcing the issue, this is indeed a very interest question which we truly believe will one day have to be answered.

REFERENCES