An Artificial Life Approach to the Study of Basic Emotions

Matthias Scheutz (mscheutz@cse.nd.edu)
Department of Computer Science and Engineering
Notre Dame, ND 46556 USA

Abstract

We propose a methodological framework for the study of emotional control based on extensive computer simulations with arti cial agents implementing emotional control mechanisms and demonstrate the methodology with simulations experiments in an arti cial environment. Speci cally, a biologically plausible schema-based model of basic forms of fear and anger is proposed and tested with respect to a variety of parameter ranges.

Introduction

Emotions are an integrative part of our mentality. At the level of the functional architecture they serve several crucial roles, from fast perceptions of threats, to focusing and redirecting attention, to in uencing memory storage and retrieval, to social regulation through expression and perception of emotions, and many more (Derryberry & Tucker, 1994; Fredrickson, 1998; Bless, Schwarz, & Wieland, 1996; Schwarz, 1990; Blaney, 1986; Kahneman, Wakker, & Sarin, 1997; Clore, Gasper, & Conway, 2001; Frijda, 2000; Cosmides & Tooby, 2000). Several circuits have been hypothesized to be involved in emotional processing in mammalian brains, yet only a few computational models (mostly of fear mechanisms) have been proposed and implemented in an effort to test theoretical predictions about emotion processes and mechanisms. Moreover, these models are limited to very specific processes (e.g., Pavlovian fear conditioning) and do not specify other parts of an architecture that are required for a complete, functional control system (e.g., homeostatic control mechanisms, various forms of perceptual processing, action selection mechanisms, etc.). Hence, they leave out and cannot address many other emotional states that essentially depend on additional processing components (e.g., such as social emotional states that depend on the expression and perception of emotions).

One way to study the effects of emotional control circuits for individual agents as well as groups of agents is to conduct simulations with artificial agents that are controlled by architectures that define emotion models. Such simulation studies have the advantage that the role of emotions and the consequences of emotional disturbances can be analyzed at several different levels at the same time: the *mechanistic level* of the implementation of the model (e.g., a neuronal level), the *individualistic level* (e.g., the control loops between emotion circuits and the agent body), and the *social level* (e.g., the effects of emotional signaling for the well-being or functioning of a group).

In this paper we will (1) propose a methodological framework for the study of emotional control based on extensive computer simulations with artificial agents implementing emotional control mechanisms and (2) demonstrate the methodology with simulations experiments in an artificial environment. Specifically, a biologically plausible schema-based model of basic forms of fear and anger is proposed and tested with respect to a variety of parameter ranges. The results show where emotional control is successful and better than non-emotional strategies, but also where it fails.

Background on Computational Models of Emotions

While several suggestions about the neural and functional organization of emotional circuits exist in the literature, there are currently only a few proposals for computational models that implement and test them. The existing computational models can be categorized into two main classes, based on whether they are aimed at explaining low-level neurological structures and mechanisms, or whether they are intended to model higher-level emotional processes. The low-level models can further be divided into general processing models of brain mechanisms and specific emotion models of particular brain structures.

The most extensively developed low-level models among the first kind are Grossberg's *CogEM* models (e.g., (Grossberg & Schmajuk, 1987)), which are models of learning cognitive, emotional, and motor properties. *CogEM* models can account for several effects in Pavlovian fear conditioning (e.g., secondary conditioning or attentional blocking), but have not been directly applied to empirical data (e.g., data from fear conditioning studies with rats).

Another class of low-level neural models is targeted specifically at modeling the amygdala, which performs several functions in emotion processing (LeDoux, 1996; Rolls, 1995). The lateral amygdala, for example, has been shown to exhibit associative plasticity during fear learning (Blair, Tinkelman, Moita, & LeDoux, 2003) and a preliminary computational model of associative learning in the amygdala has been developed and tested in three associative learning tasks (Balkenius, 2000).

Moreover, recent evidence from studies with rats suggests that the amygdala, in particular, the frontotemporal amygdala, which is taken to integrate sensory information, encodes hedonic values of an unconditioned stimulus as part of the fear memory (Fanselow & Gale, 2003). LeDoux and colleagues have hypothesized a dual pathway model of emotional processing in the amygdala, which they tested in auditory fear conditioning studies (Armony, Servan-Schreiber, Cohen, & LeDoux, 1995). These models have been also used in simulated lesion studies and successfully compared to data from actual lesion studies with rats.

While most research on emotional modeling in low-level models is focused on Pavlovian conditioning and targeted at neural structures and processing mechanisms, higher-level models of emotions are intended to capture the processing sequence involved in emotion processes and are typically concerned with a wider range of emotions. While all low-level models are neural network models, higher-level models comprise both connectionist and symbolic approaches.

An example of a high-level connectionist approach is the ITERA model (Nerb & Sperba, 2001), which is intended to study how media information about environmental problems in uences cognition, emotion, and behavior. Facts, input types, emotions, and behavioral intentions are all represented in terms of individual neural units that are connected via excitatory and inhibitory links and compete for activation.

Most attempts to model emotions at higher levels, especially in artificial intelligence research, are however based on symbolic architectures (e.g., Soar (Newell, 1990) or ACT (Anderson, 1993)). They typically focus on the OCC model (Ortony, Clore, & Collins, 1988), which hypothesizes prototypical "update rules" for changes in emotional state that can be directly implemented in rule-based systems (e.g., (Marsella & Gratch, 2002).

What is common to all the above emotion models is that they have been implemented and tested in isolation from any *body model.* Consequently, it is difficult if not impossible to investigate crucial aspects of emotion processing that need a body for control and thus go beyond functional properties (like the effects of Pavlovian conditioning), which can be tested in stand-alone models (e.g., by applying a stimulus and measuring the output).

While some attempts have been made to implement connectionist emotion models on robots, where different emotions types are represented as connectionist units that compete for activation, which in turn cause the robot to exhibit a particular behavior (e.g., (Michaud & Audet, 2001; Breazeal, 2002; Arkin, Fujita, Takagi, & Hasegawa, 2003)), these architectures do not attempt to model any specific psychological or neurobiological theory of emotions (e.g., in an effort to verify or falsify its predictions). Rather, they are mainly concerned with the applicability of a particular control mechanism from an engineering perspective. Moreover, these models typically lack a systematic evaluation of their performance (an exception is (Breazeal, 2002)). Finally, no experiments have been performed with these robotic architectures to investigate the effects of "emotional malfunctioning".

Probably the most significant restriction of current efforts to model emotions is that they have not been extended to multi-agent environments. Yet, social aspects of emotions (such as signaling emotional states through facial expressions, prosody, gestures, etc.) and the resultant effects at the group level cannot be studied in a single, isolated agent. Rather, multiple interacting agents with emotional control systems are required, especially for arguments about the adaptive role of emotions (e.g., (Cosmides & Tooby, 2000)). To our knowledge only one project (Dulk, Heerebout, & Phaf, 2003) uses an artificial life simulation to study some evolutionary aspects related to emotional processing, specifically, the evolutionary justification for LeDoux's dual-route fear processing proposal (LeDoux, 1996). However, the employed neural network does not and is not intended to implement emotions or model emotional circuits. And while the employed neural network suggests some interesting conclusions about the circumstances under which dual processing routes might be beneficial, it does not capture emotional circuits, and is, therefore, silent about emotional phenomena.

Simulations of Emotional Agents

Over the last few years we have developed an agent-based simulation environment SWAGES to investigate different agent architectures and architectural mechanisms. In particular, two main roles of emotions in agent control systems have been studied in extensive simulations in an effort to evaluate the utility of emotional control (compared to other non-emotional control strategies): *the role of emotions for individual agents* (e.g., the selection of actions) and *the role of emotion for social groups* (e.g., in con icts with conspecifics and individuals from other species).

Results from simulation experiments with agents performing foraging tasks, for example, show that *action selection* based on emotional states can be very effective in the competition for resources in hostile multi-agent environments (e.g., (Scheutz, 2001) and that motivational "hunger" and "thirst" states as well as emotional "fear" and "anger" states are likely to evolve in a variety of competitive multi-agent environments (Scheutz & Sloman, 2001)).

In general, we found that agents with emotional control mechanisms performed much better in a variety of foraging and survival tasks in environments with little to no structure than agents with much more sophisticated cognitive control systems if the "cost of deliberation" is taken into account (e.g., (Scheutz & Schermerhorn, 2002)).

On the social side, we found that expressing emotions and being able to react to emotional expressions of others can have a beneficial regulatory effect in social groups and lead to superior con ict resolution strategies (e.g., (Scheutz & Schermerhorn, 2004)).

In all these studies, we construed emotions as control processes that initiate, interrupt, suppress, reprioritize, or in general modify behavior or behavioral dispositions. Emotions are implemented in terms of control components (typically, in neural networks) that are connected in appropriate ways to sensors and effectors of agent body models. The underlying assumption is that the level of control components is appropriate for analyzing and understanding the functional organization of emotion mechanisms. In the following we brie y outline our architectural approach to the study of emotions and present some experimental results.

Basic Motivations and Emotions as Control Processes

Motivations may be considered *desire-like states* in that they in uence and bias an agent's behavioral dispositions in such a way as to contribute to the realization of a desired change in the environment and/or agent. We use the term "basic motivations" to refer to motivations that have little to no cognitive involvement and are primarily linked to "basic needs" of an agent (e.g., to maintain a certain energy level). For some of these, the familiar term "drive" is appropriate, namely if the agent is driven in a mostly reactive way to act so as to eliminate the disparity between a desired and an actual state that was the cause for the motivation. For example, a state of an agent's control system qualifies as a "hunger" state if it is caused by lack of energy and results in food-seeking behavior (McFarland, 1981).

It is possible to use control components, whose outputs control gain values of motor controllers, to implement the kind of control system that will be able to instantiate basic motivations. For example, "hunger" could be instantiated by a proportional controller P (Özbay, 2000) such that input to P comes from an internal sensor S that measures the current energy level. P compares a desired equilibrium energy level (i.e. set point), e_{des}, to the actual energy level e_{act} and scales the difference by a gain factor g_e: $P = g_e \cdot (e_{des} \quad e_{act})$. The output then is a measure of the urgency with which the system requires energy. Hence, the intensity of basic motivations is modeled by the magnitude of the control circuits' outputs that can in turn modulate behavior.

Emotions may also be considered to be desire-like in that they in uence and bias an agent's behavior. Again, we use "basic emotion" to refer to states with little or no cognitive involvement. For our purposes, we distinguish *basic emotions* from *basic motivations* in that basic emotions need not be related to a perceived difference between an actual and a desired state. Furthermore, basic emotions themselves can be states that the agent does or does not desire whereas basic motivations are directed towards or away from what the agent desires. "Fear", for example, in and of itself is an undesirable state of an agent in that it is indicative of danger. As such, it causes the agent to behave in such a way as to be prepared for or avoid danger. Hence, while "fear" can be also motivational in the sense that it may move the agent away from the cause of fear it is also emotional as it itself is not a desired state. A fear state with no clearly discernible danger present, which causes an agent to be more cautious and alert, may itself not instantiate a motivational state that is connected to a particular goal such as running away from a particular threat.

"Fear", as discussed above, can be instantiated by a controller C, which integrates over time the frequency of occurrence of fear triggering conditions. Input to C comes from an internal sensor S that is activated by a fear triggering condition. C integrates these inputs over time and outputs a signal that corresponds to the intensity of "fear" and modulates behavior to be more alert and ready for sudden activity. A neural control circuit implementing an appropriate response characteristic (similar to that given by $g(t) = e^{-t}$ to a unit impulse, which is generated by the sensor or the perceptual system detecting a dangerous stimulus), could use an *interactive activation and competition* (IAC) unit (McClelland & Rumelhart, 1988), whose change in activation is given by $\Delta act = S \cdot g_S \cdot act + decay \cdot act$, where act is the current activation level of the control system, g_S is the gain for the sensor input and $decay$ is the discount value for past activations.

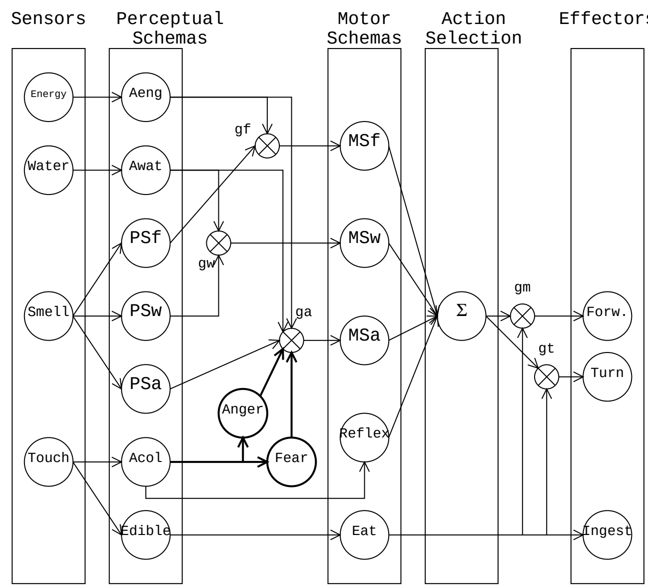

Figure 1: The schema-based architecture for the simulated emotional agents (see the text for details).

To be able to instantiate a fear state, the above controller needs to control the agent's effectors in a way that the positive output from the controller can in uence and bias the agent's behavior towards avoiding or attempting to avoid dangerous objects. As such, the intensity with which the agent avoids or attempts to avoid these objects depends on the magnitude of the output of the controller: the agent's behavior is modulated by its level of fear.

A Schema-Based Agent Architecture

Using the above control elements to implement basic motivations and emotions, we have compared the performance of agents with mechanisms to implement *fear* and *anger* to that of agents without these mechanisms in a hostile multi-agent environment, where agents need to forage for resources in order to survive and procreate. The employed architecture is a biologically plausible schema-based architecture (Arbib, 1992) for both agent kinds, which allows the agents to forage for food and water. In this architecture, the behavior of an agent depends at any given time on the relative contributions from a variety of motor schemas. While non-emotional agents have fixed behavioral dispositions to deal with competitors for resources, emotional agents use their emotional control circuits to adapt their behaviors based on past encounters.

Figure 1 shows the architecture for the emotional agents (their emotional subsystem is an implementation of the higher-level functional organization of the basic mammalian "fear/anger system" in the terms of the above suggested control units, e.g., (Berkowitz, 2003)). Schemas are depicted by large circles where the names indicated their function.[1] Small crossed circles indicate gains of schemas (i.e., behavioral dispositions) that are taken as architectural parameters to be varied in the experiments: the degree to which an agent is attracted to food (g_f), to water (g_w), and to other agents (g_a). The bold-face circles labeled "Fear" and "Anger" represent the "fear schema" and "anger schema", respectively. They are only present in the architecture of emotional agents. Both emotion schemas are connected to an "alarm schema" (*Acol*), which is triggered if an agent touches other agents. This mechanism changes the agent's propensity to fight other agents or to ee: the higher the output of a controller, the more stronger the behavioral disposition (i.e., to fight for anger, or ee for fear).

More formally, let $Ent = \{f, w, a\}$ be an index set of the three types of objects in the simulation environment: *food*, *water*, and *agents*. For each object type in Ent, a force vector F_i is computed, which is the sum, scaled by $1/|v|^2$, of all vectors v from the agent to the objects of type i within the respective sensory range, where $|v|$ is the length of vector v. These *perceptual schemas* are mapped into motor space by the transformation function $T(x) = \sum_{i \in Ent} g_i \cdot F_i(x)$, where the g_i are the respective gain values of the perceptual schemes. The gain values simply scale the effect of sensory input, providing a means by which to prioritize certain inputs (e.g., if food is especially important, its gain value could be higher than the other gain values, so that sensing food has a greater impact on the direction chosen than sensing other entities).

All feedback controllers are implemented in a feed-forward three-layer interactive activation and competition neural network (with three input units *in*, three hidden units *hid*, and three output units *out*). The input units receive their activations (via appropriate scaling functions) from the $Water$ (in_w) and $Energy$ level sensors (in_f) via the perceptual $Awat$ and $Aeng$ schemas as well as from the $Touch$ sensor via the $Acol$ schema (in_a), respectively.

The output units are connected to the gain values in the motor scheme via individual scaling functions $f_i(x) = x \cdot c_i + b_i$ (where b_i is the *base gain value* and c_i the scaling factor for the activation of out_i).

The activation value $act_i(t)$ of an IAC unit i at time t is defined by

$$act_i(t) = \begin{cases} (max \quad act_i(t \quad 1)) \cdot net_i(t) \quad decay, \\ \hfill net_i(t) \geq 0 \\ (act_i(t \quad 1) \quad min) \cdot net_i(t) \quad decay, \\ \hfill net_i(t) < 0 \end{cases}$$

where min and max are the minimum and maximum activation level, respectively, $decay$ is a decay factor defined by $d \cdot (act_i(t) \quad rest)$ (where d is a constant), $rest$ the rest level, and $net_i(t)$ the weighted sum of all inputs to unit i at time t.

The choice of IAC units over standard perceptrons is based on their update rule, which is particularly suited to implement important temporal features of some emotional states in that it (1) takes into account the *previous activation* (hence, can be used to implement "inner states"), and (2) incorporates a *decay term* to raise or lower the activation to a predetermined *base level*.

Non-emotional agents have a constant g_a gain (i.e., their $c_i = 0$), hence their behavioral dispositions towards other agents are fixed. Emotional agents, on the other hand, can adapt their behavior dispositions, i.e., their g_a gain, by virtue of the feedback controllers implemented in the neural net (their $c_i \neq 0$). Depending on whether g_a is positive or negative, they can implement basic "anger" or "fear" states (as argued in (Scheutz, 2001)).

The Utility of Anger and the Limits of Fear

We report results from two classes of experiments studying the role of emotions in foraging and survival tasks.[2] In the first class, the gain g_a is set to a negative value for both agent kinds, thus making them disposed to avoid other agents. For the second class, g_a is positive for both agent kinds, thus making them disposed to be aggressive towards other agents. Performance was measured in terms of the number of surviving agents after 10000 simulation cycles averaged over 40 runs with random initial conditions. The upper and lower parts of Figure 2 show the results from both classes of experiments for both agent kinds for two architectural variations: agent gain and water gain (i.e., 25 sets of 40 experimental runs each). All runs started with 10 agents of each of the two kinds placed at random location in the environment together with 20 randomly placed food and 20 randomly placed water items; new food and water items are generated on every 4 and 6 cycles in random locations, respectively.

While emotional agents in the first set have a performance peak (of 23.625) that is slightly higher than that of non-emotional agents (of 23.35), the difference is not

[1] For space reasons we cannot describe all the details of the architecture here.

[2] For more details about simulation setup and simulation parameters see (Scheutz, 2001).

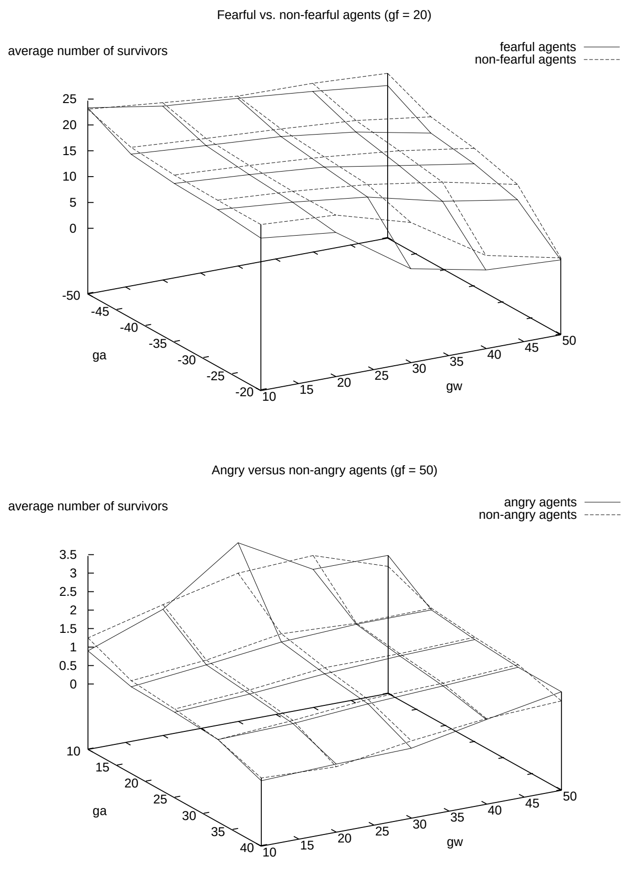

Figure 2: The performance space of the emotional vs. and non-emotional agents (fearful top, angry bottom) based on variations along two architectural dimensions.

significant (t-test, $p > 0.1$). Consequently, being fearful in addition to having the behavioral disposition of avoiding other agents does not increase the overall performance, it may in fact reduce it for some settings of the gain values (e.g., $g_a =$ 20 and $g_w = 30$). For emotional agents in the second class of experiments, however, we find a marginally significant global maximum at $g_a = 10$ and $g_w = 30$. Consequently, in the kinds of environments studied, "anger" does sometimes prove useful for survival.

Discussion

The results reported here are only a very small part of a large set of experiments, in which up to five architectural dimensions were varied in an effort to determine the circumstances in which emotional control is beneficial and where it might be detrimental. The methodology on which they are based consists of a four step process: (1) emotion concepts are analyzed and defined in terms of architectural capacities of agent architectures (Sloman, 2002). (2) Agent architectures with particular emotional control mechanisms (as defined in (1)) are constructed for a given task, for which also a performance measure is defined. (3) Simulations experiments are carried out with the so-defined emotional agents and their performance is determined for a predetermined set of architectural and environmental variations. The outcome then is a *performance space* that corresponds to the varied parameters. The last two steps are repeated with agents implementing non-emotional (or, in general, other) architectures. (4) All resulting performance spaces are then compared with respect to the agents' *performance-cost tradeoffs*, i.e., their performance taken relative to the (computational) cost necessary to maintain and run the instantiated architecture (in the reported experiments the cost for both architectures was taken to be the same). The last point is crucial as it may well be that emotional agents do not perform better than non-emotional ones on a given task in absolute terms, but that they do much better in relative terms, i.e., with fewer resources (which is usually believed to be the case by emotion researchers). Especially from an evolutionary perspective relative performance is the relevant measure.

We believe that the proposed methodology to experiment with agent architectures in an artificial life environment cannot only form the basis for a thorough comparison of the different emotion models that can otherwise not be studied easily (e.g., social emotions and their role in the control of agents), but can also inform emotion researchers interested in clinical aspects of emotions by performing simulated *lesion studies*, where parameters of functional agents are modified or components of the architecture are removed. This, in turn, might help us isolate not only the functional roles of emotions in the control of creatures, but also the ways in which emotional control can fail and how it might be possible to reestablish normal functioning in dysfunctional systems.

References

Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Arbib, M. (1992). Schema theory. In S. Shapiro (Ed.), *The handbook of brain theory and neural networks* (pp. 830–833). MIT Press.

Arkin, R., Fujita, M., Takagi, T., & Hasegawa, R. (2003, March). An ethological and emotional basis for human-robot interaction. *Robotics and Autonomous Systems*, *42*, 3-4.

Armony, J., Servan-Schreiber, D., Cohen, J., & LeDoux, J. (1995). An anatomically constrained neural network model of fear conditioning. *Behavioral Neurosciences*, *109*(2), 256–257.

Balkenius, J. M. C. (2000). A computational model of emotional learning in the amygdala. *Cybernetics and Systems*, *32*(6), 611–636.

Berkowitz, L. (2003). Affect, aggression, and antisocial behavior. In (Davidson, Scherer, & Goldsmith, 2003) (pp. 804–823).

Blair, H., Tinkelman, A., Moita, M., & LeDoux, J. (2003). Associative plasticity in neurons of the lateral amygdala during auditory fear conditioning. *Annals of the New York Academy of Sciences*, *985*, 485–487.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, *99*(2), 229–246.

Bless, H., Schwarz, N., & Wieland, R. (1996). Mood and the impact of category membership and individuating information. *European Journal of Social Psychology*, *26*, 935-959.

Breazeal, C. L. (2002). *Designing sociable robots*. MIT Press.

Clore, G., Gasper, K., & Conway, H. (2001). Affect as information. In J. Forgas (Ed.), *Handbook of affect and social cognition* (p. 121-144). Mahwah, NJ: Erlbaum.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). NY: Guilford.

Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (Eds.). (2003). *Handbook of affective sciences*. New York: Oxford University Press.

Derryberry, D., & Tucker, D. (1994). Motivating the focus of attention. In P. Neidenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influence in perception and attention* (pp. 67–96). San Diego, CA: Academic Press.

Dulk, P. den, Heerebout, B., & Phaf, R. (2003). A computational study into the evolution of dual-route dynamics for affective processing. *Journal of Cognitive Neuroscience*, *15*(2), 194–208.

Fanselow, M. S., & Gale, G. D. (2003). The amygdala, fear, and memory. *Annals of the New York Academy of Sciences*, *985*, 125–134.

Fredrickson, B. (1998). What good are positive emotions? *Review of General Psychology*, *2*, 300–319.

Frijda, N. H. (2000). The psychologists' point of view. In (Lewis & Haviland-Jones, 2000) (pp. 59–74).

Grossberg, S., & Schmajuk, N. (1987). Neural dynamics of attentionally-modulated pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. *Psychobiology*, *15*, 195–240.

Kahneman, D., Wakker, P., & Sarin, R. (1997). Back to bentham? explorations of experienced utility. *Quarterly Journal of Economics*, *112*, 375-405.

LeDoux, J. (1996). *The emotional brain*. New York: Simon & Schuster.

Lewis, M., & Haviland-Jones, J. M. (Eds.). (2000). *Handbook of emotions* (2nd ed.). New York: The Guilford Press.

Marsella, S., & Gratch, J. (2002, May). Modeling the in uence of emotion on belief for virtual training simulations. In *Proceedings of the 11th conference on computer-generated forces and behavior representation*. Orlando, FL.

McClelland, J. L., & Rumelhart, D. E. (1988). *Parallel distributed processing* (Vol. 1 and 2). Cambridge: MIT Press.

McFarland, D. (1981). *The oxford companion to animal behavior*. Oxford: Oxford University Press.

Michaud, F., & Audet, J. (2001). Using motives and artificial emotion for long-term activity of an autonomous robot. In *5th autonomous agents conference* (pp. 188–189). Montreal, Quebec: ACM Press.

Nerb, J., & Sperba, H. (2001). Evaluation of environmental problems: A coherence model of cognition and emotion. *Cognition and Emotion*, *4*(15), 521–551.

Newell, A. (1990). *Uni ed theories of cognition*. Cambridge, MA: Harvard University Press.

Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of the emotions*. New York: Cambridge University Press.

Özbay, H. (2000). *Introduction to feedback control theory*. London: CRC Press.

Rolls, E. T. (1995). A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1091–1106). Cambridge, MA: MIT Press.

Scheutz, M. (2001). The evolution of simple affective states in multi-agent environments. In D. Cañamero (Ed.), *Proceedings of aaai fall symposium* (pp. 123–128). Falmouth, MA: AAAI Press.

Scheutz, M., & Schermerhorn, P. (2002). Steps towards a theory of possible trajectories from reactive to deliberative control systems. In R. Standish (Ed.), *Proceedings of the 8th conference of arti cial life*. MIT Press.

Scheutz, M., & Schermerhorn, P. (2004). The role of signaling action tendencies in con ict resolution. *Journal of Arti cial Societies and Social Simulation*, *7*(1).

Scheutz, M., & Sloman, A. (2001). Affect and agent control: Experiments with simple affective states. In N. Zhong, J. Liu, S. Ohsuga, & J. Bradshaw (Eds.), *Intelligent agent technology: Research and development* (pp. 200–209). New Jersey: World Scientific Publisher.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. Higgins & R. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, p. 121-144). New York: Guilford Press.

Sloman, A. (2002). Architecture-based conceptions of mind. In *Proceedings 11th International Congress of Logic, Methodology and Philosophy of Science* (pp. 397–421). Dordrecht: Kluwer. ((Synthese Library Series))