Unifying Emotion and Cognition:
The Flow Model Approach

Luís Morgado[1,2] and **Graça Gaspar**[2]

[1]Instituto Superior de Engenharia de Lisboa
Rua Conselheiro Emídio Navarro, 1949-014 Lisboa, Portugal
lm@isel.ipl.pt
[2]Faculdade de Ciências da Universidade de Lisboa
Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
gg@di.fc.ul.pt

Abstract

Growing experimental evidence shows the fundamental role that emotion plays in intelligent behavior as well as the close relationship between emotion and what is classically known as cognition. Due to this evidence different models have been proposed to support emotion in artificial agents. However, the classical separation between emotion and cognition still prevails, not favoring the definition of a general framework to support the implementation of emotional agents. In this paper we present a model that unifies emotion and cognition as two symbiotically integrated aspects of intelligent behavior. Despite being inspired by biological processes, the proposed model is independent of specific physiological or psychological details in order to support the implementation of concrete agents of different kinds and levels of complexity.

Introduction

In recent years emotion has received a growing attention concerning the modeling and implementation of intelligent agents. This trend was particularly notorious after relevant theoretical and experimental work, e.g. (Damásio 1994, LeDoux 1996), has demonstrated the close relationship between cognition and emotion. However, the focus of the work concerning emotions in intelligent agents has been on modeling and implementing emotional aspects and not much in combining emotion and cognition. We believe that this is a crucial aspect for the implementation of agents capable of intelligent behavior in dynamic and uncertain conditions, typical of real environments. Therefore, to effectively support emotion in intelligent agents it is necessary to address this problem and to define emotional models that can serve as a general framework for the implementation of agents of different kinds and levels of complexity, including agents capable of conceptual reasoning.

In this paper we present the overall structure of such a

model, which we call Flow Model of Emotion (Morgado & Gaspar 2003b), and discuss how it supports the symbiotic integration between emotion and cognition. The paper is organized as follows: in section 2, the motivation and overall structure of the Flow Model of Emotion is presented; in section 3, we build upon that foundational framework to define an emotional agent model; finally, in section 4 we discuss the relation between emotion and reasoning.

The Flow Model of Emotion

Two main approaches have characterized the development of emotion models for intelligent agents, physiologically inspired models, e.g. (Cañamero 2003, Lisetti 1999), and appraisal theories inspired models, e.g. (Bates 1994, Gratch & Marsella 2001). These models have allowed the definition of emotion like characteristics and behavior, however they also have some drawbacks. Physiologically inspired models are based on specific mechanisms of biological organisms, some of them still not well known, like hormonal mechanisms, resulting in highly specific implementations limited to relatively simple agents and contexts. On the other hand, appraisal theories don't say much about the processes underlying emotional phenomena (Scherer 2000) nor about other fundamental aspects like adaptive behavior.

The classical approach to emotional appraisal suggests the existence of base appraisal dimensions that individuals use to evaluate the significance of events, e.g. (Frijda 1986, Lazarus 1991, Ortony et al. 1988). Different authors have proposed different sets of appraisal dimensions. However, there are two dimensions that are common to the main models (even if appearing under different names): goal conduciveness and coping potential as named by Scherer (2000) or motive-consistency and control potential as named by Roseman et al. (1996), among others. From our point of view, these two dimensions are effectively base dimensions, while other proposed dimensions, such as pleasantness or norm compatibility are contextualized

variations of these two base dimensions, namely in cognitive and social contexts. However, even these two dimensions serve only to classify an emotional state from an external perspective, but are not sufficient to explain the process through which emotional states are attained and emotional behavior is generated. In fact this is one of the above-mentioned main problems of appraisal theories of emotion.

Another dimension that is normally included as an appraisal dimension is novelty (or other variations as unexpectedness or surprise). From our point of view, this dimension reflects essentially the level of change, therefore it results from the dynamics of the emotional processes - it can be needed to characterize an emotional state from an external point of view, but it is not explicitly represented in terms of internal structures.

We propose that the two appraisal dimensions, goal conduciveness and coping potential, reflect the existence of base attributes that support the elicitation of emotional states and behavior. We call these base attributes *achievement potential* (P) and *achievement conductance* (C). P represents the potential of change that an agent is able to produce in the environment to achieve the intended state-of-affairs. C represents the degree of the environment's conduciveness or resistance to that change, which can also mean the degree of environment change that is conducive, or not, to the agent intended state-of-affairs. Therefore, achievement potential and achievement conductance represent two complementary aspects in the relation agent/environment, which some authors recognize as one of the main aspects underlying emotion (Lazarus 1991). These attributes are distinct from the appraisal dimensions in the sense that they are objective attributes that characterize the cognitive structure elements of an agent and not subjective psychological dimensions. As we will see, they are rooted on the dynamics of energy interchange between the agent and the environment. This is possible due to a conception of an agent as an open system that maintains itself in a state far from equilibrium, yet keeping an internally stable overall structure. This kind of systems is known as dissipative structures (Kondepudi & Prigogine 1998).

In this kind of systems the achievement potential can be viewed as a force (P) and the achievement conductance as a transport property (C). The dynamics of the emotional process can therefore be characterized as a relation corresponding to a flow, called achievement flow, defined by:

$$F = C \cdot P \qquad (1)$$

That means the achievement flow is proportional to the achievement potential, where the proportionality factor C is a transport property that reflects the degree of conduciveness or resistance to the achievement flow.

Emotional Dynamics

The behavioral forces that arise from the dynamic relation between achievement potential and achievement conductance, expressed as energy flows, generate behavioral dynamics that underlie the cognitive processes. To understand the nature of those forces let us consider a function *hs* with arguments P and C that characterizes the hedonic state of an agent, representing the level of well-being of the agent in relation to the achievement of its motivators. The evolution of *hs* can be related to the variation of the achievement potential (P) and the achievement conductance (C), as follows:

$$\frac{dhs}{dt} \equiv f(\delta P, \delta C) \quad \text{where} \quad \delta P = \frac{dP}{dt} \quad \text{and} \quad \delta C = \frac{dC}{dt} \qquad (2)$$

Independently of the nature of function *f*, it can be viewed as representing a motivating force leading to the change of the agent's hedonic state and therefore underlying the corresponding emotional state. Since we don't know the exact nature of that force we propose that it can be represented as a vectorial function *ED*, called *emotional disposition*, defined as follows:

$$ED \equiv (\delta P, \delta C) \qquad (3)$$

Functions P and C represent two independent and complementary aspects, therefore, at a given instant $t = \tau$, δP and δC can be represented as two orthogonal dimensions of variation of *ED*.
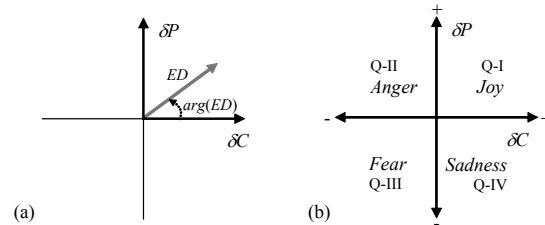

Fig. 1. Vector *ED* as a function of δP and δC (a); relation between *ED* quadrants and emotional quality tendency (b).

Since we consider P and C the main attributes upon which the elicitation of an emotional state is based, it is important to be able to relate their variation with the emotional tendency that may be underlying them, as characterized from experimental evidence. In figure 1.b we identify four quadrants of evolution of *ED* as a function of δP and δC: Q–I ($\delta P > 0$ and $\delta C > 0$) "everything goes well", directly related to joy; Q–II ($\delta P > 0$ and $\delta C < 0$) the agent has potential to handle the situation but there are "adversities", typical of anger; Q–III ($\delta P < 0$ and $\delta C < 0$) the agent does not have capacity to handle the "adversities", typical of fear situations; Q–IV ($\delta P < 0$ and $\delta C > 0$) corresponds to a situation where despite no significant opposition to the achievement of the agent's motivators, there is a negative variation of the agent's potential; this kind of situation can be directly related to sadness, which according to experimental studies, e.g. (Lazarus 1991), often results from sudden losses (a sudden decrease of P produces a strong negative variation of δP).

Therefore, an emotional disposition vector can be directly related to a specific emotional quality, defined by its

orientation (or argument) and an intensity defined by its module. That is:

$$Quality(ED) \equiv \arg(ED) \qquad (4)$$

$$Intensity(ED) \equiv |ED| \qquad (5)$$

It is important to note that the emotional tendency associated to each quadrant (joy, anger, fear, sadness) is only indicative of its main nature, since the quality of the emotional disposition is continuous. This is consistent with phenomenological well-known emotion blends. In this way it is possible to account for the existence of basic emotional patterns without compromising the rich and continuous nature of emotional phenomena.

To conclude the presentation of the emotional disposition notion, we must consider the following aspects: (i) emotional disposition is defined as an action regulatory disposition or tendency, but it does not constitute in itself an emotion; (ii) emotions are considered emergent phenomena that result from the dynamics of agent/environment interaction. Their existence is ascribed by an external observer or by self-reflective processes of an agent, through the observation of specific behavioral patterns and states (either external or internal).

From Emotional Dispositions to Emotional States

In the proposed model the distinct emotional phenomena, like emotion episodes and emotions, derive from the base emotional disposition forces and result in a continuum of manifestations, such as observable behavior. From the cumulative effects of these emotional disposition forces structural/behavioral congruencies emerge that, from an external observer point of view, can be characterized as emotional states.

In the Agent Flow Model the cognitive structure of an agent is composed by cognitive elements, that will be defined in the next section, which can form chains and clusters as a result of cognitive activity. In agents with conceptual abilities, like social or self-reflective abilities, these clusters of coupled cognitive elements form cognitive contexts (*eliciting contexts*). Complex emotional states, like the ones involving social relations, are elicited from such contextualized emotional forces. This is consistent with the diversity of appraisal dimensions in different appraisal models. However, an important difference exists. In the proposed model there are no specific appraisal dimensions, nor are emotional phenomena dependent on a specific subsystem. On the contrary, emotion is an intrinsic aspect of all cognitive activity.

A Motivating Example

As a motivating example we will consider a personal assistant agent that supervises the financial investments on behalf of its human user.

In this kind of situations emotion could play an important role both as motivator of the agent behavior as well as a highly effective tool for human-agent interaction.

The experimental framework consists of a simulation where the agent experiments different degrees of achievement conductance, from all-conducive financial situations ($C = 1$), corresponding to high profits, to totally non-conducive financial situations ($C = 0$), corresponding to high losses. Two base scenarios were simulated, as illustrated in figure 2.a: (i) limited adversity; (ii) full adversity.
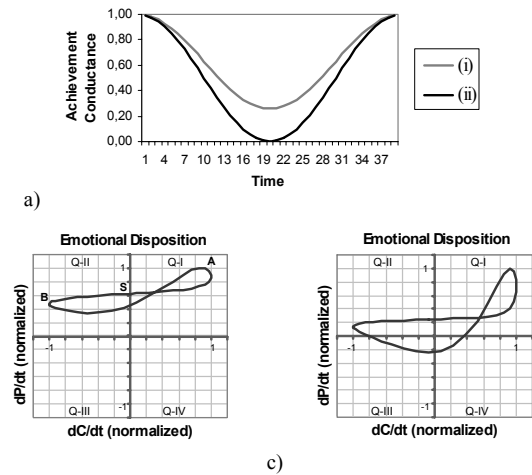

Fig. 2. Variation of achievement conductance with time (a). Resulting evolution of the emotional disposition (the time axis is omitted – successive *ED* vectors are projected on the emotional disposition plane – evolution direction is S-B-A) (b) and (c).

At each time step the global emotional disposition of the agent was evaluated. The results for the two base scenarios (i) and (ii) are shown, respectively, in figures 2.b and 2.c. In experiment (i), as the conductance varies from high values to low values, the emotional disposition gains a clear anger tendency. When the conductance becomes high again, the agent turns to a joy tendency (satisfaction). It is possible to observe two distinct peaks that can be interpreted as reflecting an anger peak (B), resulting from the experienced adversities, followed by an elation/satisfaction peak (A) resulting from the end of the adversities. However, if the conductance is further decreased, as is the case in experiment (ii), a fear tendency appears and, before the agent turns to joy, the quadrant Q-IV is crossed, corresponding to a relief like emotional tendency.

These experiments showed some other interesting results (more details can be found in (Morgado & Gaspar 2003b)), which are comparable to biological emotional phenomena, namely, sudden shifts in emotional disposition, dependence of the departure point, emotional decay, and a continuous nature allowing for complex blends or mixtures of emotions.

Modeling Emotional Agents:
The Agent Flow Model

The foundations of the Flow Model of Emotion just presented constitute a conceptual framework for modeling emotional agents, however, to support its concrete implementation and to define complex emotional phenomena, like emotional states or mood, an operational model is needed where the base notions of potential and flow can be rendered concrete. We called that model *Agent Flow Model* (Morgado & Gaspar 2003a). The Agent Flow Model defines two main aspects of agent construction: cognitive structure and dynamics; and architectural patterns and mechanisms. In this paper we will focus on the base cognitive structure and dynamics.

Cognitive Elements

One of the main ideas underlying the Agent Flow Model is that interactions between cognitive elements occur as forces resulting from the flow of energy, acting as behavioral driving forces and forming the basis of emotional phenomena. On the other hand, there is much experimental and theoretical evidence that biological cognitive activity is based on the composition of basic components, e.g. (Levine 2000, Gärdenfors 2000).

Based on these evidences, cognitive potentials and flows are characterized as energetic signals composed by a base signal with a specific frequency (its *quality*), and an amplitude (its *intensity*) varying in time. Through superposition, aggregates of potentials can be formed. These aggregates of potentials, which we call *cognitive elements*, are the base of the cognitive structure and activity. Formally, a cognitive potential is defined as:

$$p(t) = a(t).\varphi(\omega,t) \qquad (6)$$

where $a(t)$ is the amplitude of the potential and $\varphi(\omega,t)$ is a *base signal* that represents some aspect of the agent's cognitive "reality" with a specific quality ω. The base signals that characterize the cognitive potentials are orthogonal among each other, which implies superposition of energy. Therefore a cognitive element $\sigma(t)$ is defined as a superposition of cognitive potentials. That is:

$$\sigma(t) = \sum_{i=1}^{N} p_i(t) \qquad (7)$$

where N is the number of potentials in the aggregate.

The interaction between cognitive elements is determined by their quality. Elements with the same qualities have strong interactions, reinforcing each other. Elements with different qualities interact according to their similarity.

Cognitive Space

The base signals that compose potentials and cognitive elements form a signal space underlying the cognitive structure of an agent, which we call a *cognitive space*. A cognitive space is characterized as a multi-dimensional signal space where each base signal defines a dimension. Formally, a cognitive space CS_K is defined by a set of K orthonormal basis functions $\Phi = \{\varphi_i : i = 1, 2, \ldots, K\}$ with $K \in \aleph$. Each basis function φ_i corresponds to a base signal $\varphi(\omega_i,t)$ with a specific quality ω_i.

Cognitive elements correspond to specific positions in the cognitive space. Since cognitive elements change with time, at successive time instants they occupy different positions, describing trajectories that reflect the behavior of the agent. At some instant $t = \tau$, a cognitive element $\sigma(t)$ is represented in a cognitive space CS_K as a vector σ, defined as:

$$\sigma = (a_0, a_1, \ldots, a_k) \qquad (8)$$

where the dimensional factors a_i represent the intensity of quality ω_i in the cognitive element.

Cognitive Dynamics

One of the main characteristics of intelligent behavior is the orientation towards the achievement of motivations. These motivations can take various forms according to the cognitive context (e.g. drives, desires), but they all share two fundamental characteristics: (i) they represent an intended situation; (ii) they act as a motivating force driving agent's behavior.

A cognitive element that represents an intended situation is called a *motivator*. On the other hand, an intended situation is relative to a current situation, represented by cognitive elements called *observations*. These observations result from inward flows associated to activities like perception. It is the difference between observations and motivators that produce the main forces underlying agent's cognitive dynamics. The cognitive activity is consequently guided by the maximization of the flows that lead to the reduction of the distance between observations and motivators. For an agent to succeed in reducing that distance, it needs some form of mediation that supports the transformation of the motivating force into applied force to act in order to change the current situation. When a cognitive element plays this role we call it a *mediator*. The commitment to apply a motivating potential through a specific mediator results in a new element called an *achiever*, which support the concrete action that leads to the change of the current situation and to the corresponding movement of the observations in the cognitive space, as illustrated in figure 3.
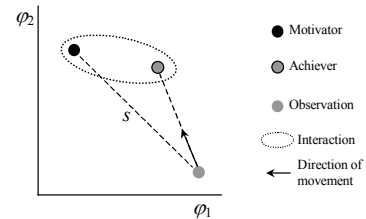

Fig. 3. Elements participating in the achievement of a motivator (two-dimensional cognitive space with dimensions φ_1 and φ_2).

Thus, cognitive elements play different roles according to their characteristics and the cognitive context where they are involved. Four main roles are identified: (i) *motivators*, elements representing an intended situation and composed by motivating potentials, produce a motivating force for an agent to act; (ii) *mediators*, elements that mediate the transformation of the motivating force into applied force to act; (iii) *achievers*, elements resulting from the commitment to apply a motivating potential through a specific mediator, forming the concrete support for action; (iv) *observations*, elements resulting from inward flows associated to activities like perception.

Emotional Driving Forces

As we can observe in figure 3, the direction of the observation's movement may not be the exact direction of the motivator. Besides that, the velocity of the movement can also change during the achievement process. Even if the achiever is aligned with the motivator, the dynamics of the environment (either internal or external) can influence the movement of the observation. This means that the agent must continuously adjust its behavior in order to succeed in the achievement of its motivators, especially in uncertain and dynamic environments.
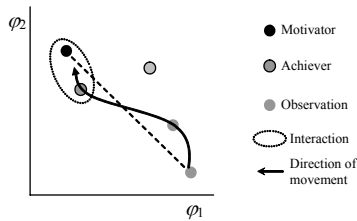

Fig. 4. Continuous adjustment of agent's behavior.

Figure 4 shows a possible trajectory resulting from the adjustment of agent's behavior by switching to a different achiever. This second achiever can be a new option or a refinement of the previous achiever through a planning process.
Independently of the specific processes that generated the new achiever, the forces that led to that change underlie all the cognitive dynamics of the agent. In the Flow Model the emotional phenomena are considered the expression of those forces, which can be characterized as *emotional dispositions*, as presented before.
In the cognitive space, the cognitive dynamics can be described by the movements of cognitive elements, and the associated emotional disposition by the evolution of the distance and velocity relative to the motivators. Considering two generic cognitive vectors σ_1 and σ_2 at some instant $t = \tau$, distance s and velocity v are defined as follows:

$$s = \|\sigma_1 - \sigma_2\| \quad \text{with} \quad \|\sigma\|^2 = \langle \sigma, \sigma \rangle \qquad (9)$$

$$v = \frac{ds}{dt} \qquad (10)$$

where $\langle x, y \rangle$ represents the scalar product of vectors x and y.
The variations of potentials and flows underlying the movement of an observation in the cognitive space relative to a motivator are the base of the emotional disposition forces. Therefore, in the cognitive space the emotional disposition is defined as follows:

$$ED \equiv (\delta s, \delta v) \quad \text{where} \quad \delta s = -\frac{ds}{dt} \quad \text{and} \quad \delta v = \frac{dv}{dt} \qquad (11)$$

In this way it is possible to characterize the cognitive dynamics as emotional phenomena in a concise way, as proposed initially. However, we must note that this characterization is made from an external observer point of view. From an operational point of view, the forces underlying the emotional disposition tendencies are behavioral forces that constrain the cognitive processes of an agent. Therefore, the dynamics resulting from these forces are, at the same time, a result of the cognitive activity and a constraint that influences it.

Emotion and Reasoning

The behavior of a Flow Model agent is determined by its cognitive structure and by the associated cognitive processes. Therefore, different types of cognitive structures and processes produce different types of agents as illustrated in the following figure.
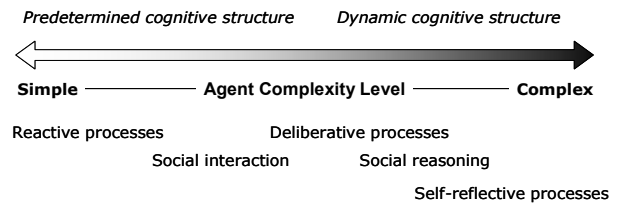

Fig. 5. Agent types.

Simple agents have predetermined cognitive structures and very simple cognitive processes with few or no manipulation of the cognitive elements information contents. Their behavior is guided by the cognitive forces resulting from the energetic potentials and flows.
The *kineses* of some organisms are an example of this type of behavior (e.g. bacterial *chemotaxis*) (Staddon 2001, Webb & Consi 2001). For example, bacteria and other types of unicellular organisms possess no nervous system and only the most rudimentary sensors, however, despite their apparent simplicity they are able to avoid predators and find sources of food in an efficient way. In some organisms, like the *Stentor roeselii* we can even observe remarkable behaviors for a unicellular organism without nervous system, which from an observer point of view are easily classified as emotional (Staddon 2001). These examples are illustrative not only of the power of these apparently simple mechanisms, but also of their intrinsic

adaptive behavior and emotional content, even if in its most simple form.

Reasoning. As the complexity of the cognitive structure and processes grow, a qualitative transition happens when the cognitive processes are able to explicitly manipulate the information content of the cognitive elements and, specially, when the cognitive processes are able to manipulate the cognitive structure, changing and generating cognitive elements in order to anticipate environmental situations. This anticipative behavior is clearly distinct from the adaptive behavior observable in simple organisms and together they form the basis for what is known as intelligent behavior.

The underlying processes responsible for anticipative behavior are characterized by the internal generation of cognitive elements that represent possible future situations or purely hypothetical situations, and they manipulate them in order to elaborate and select options to achieve the agent motivators. This kind of processes can be related to the "*as-if-circuits*" proposed by Damásio (1994).

Reasoning processes shape the overall form of a cognitive element by changing its dimensional factors, that is, by changing the amplitude of the potentials that compose the cognitive element. This change of energetic content results from the cognitive processes acting with hypothetical situations in the same way that in concrete situations, despite the fact that the intensity of the changes may be less. These changes also reflect the subjective evaluations made during the reasoning activities.

In this way, reasoning and emotion cannot be viewed as distinct functional systems since they are two aspects of the overall cognitive activity of an agent in a symbiotic integration. In the same way, emotion cannot be relegated to a peripheral sub-system because it underlies the cognitive activity, being intrinsic to adaptive behavior.

Conclusions and Future Work

Research on emotional phenomena progressively reveals a view of emotion as a dynamic process intrinsic to cognition, e.g. (Damásio 1994, Gray et al. 2002, Rolls 1999). However, to adopt this view we need to go beyond the classical separation between emotion and cognition. In this paper we tried to contribute to that aim by presenting an approach that unifies emotion and cognition as two symbiotically integrated aspects of intelligent behavior. The proposed models form a conceptual framework for cognitive modeling that is independent of the architectural organization of the agents, making possible to describe emotional phenomena in a concise way and to model agents of different types and levels of complexity,

Future research will aim at modeling different types of agents, namely, deliberative agents with social ability and to further refine the integration between emotional and reasoning aspects for the implementation of emotional deliberative agents.

References

Bates, J. 1994. The Role of Emotion in Believable Agents. *Communications of the ACM*, 37(7):122-125.

Cañamero, L. 2003. Designing Emotions for Activity Selection in Autonomous Agents. In R. Trappl, P. Petta, S. Payr, (Eds.), *Emotions in Humans and Artifacts*, 115-148, MIT Press.

Damásio, A. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam PG.

Frijda N. 1986. *The Emotions*. Cambridge Univ. Press.

Gärdenfors, P. 2000. *Conceptual Spaces: The Geometry of Thought*. MIT Press.

Gratch, J., Marsella, S. 2001. Tears and Fears: Modeling Emotions and Emotional Behaviors in Synthetic Agents. In Proc. of the 5th Int. Conf. on Autonomous Agents, 278-285.

Gray, J., Braver, T., Raichle, M. 2002. Integration of Emotion and Cognition in the Lateral Prefrontal Cortex. Proceedings of the National Academy of Sciences USA, 99:4115-4120.

Kondepudi, D., Prigogine, I. 1998. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*. J. Wiley & Sons.

Lazarus, R. 1991. *Emotion and Adaptation*. Oxford Univ. Press.

LeDoux., J. 1996. *The Emotional Brain*. Touchstone.

Levine, M. 2000. *Fundamentals of Sensation and Perception*. Oxford Univ. Press.

Lisetti, C. 1999. Emotion Generation for Artificial Agents via a Hybrid Architecture. In Proc. of the Autonomous Agents Workshop on Emotion-Based Agent Architectures at the 3rd Int. Conf. on Autonomous Agents.

Morgado, L., Gaspar, G. 2003a. Extending BDI with Emotional Dispositions for Adaptive Social Reasoning. In Proc. of the 2nd Int. Workshop on Regulated Agent-Based Social Systems.

Morgado, L., Gaspar, G. 2003b. Emotion in Intelligent Virtual Agents: The Flow Model of Emotion. In Rist, T. et al. (Eds.), *Intelligent Virtual Agents*, LNAI 2792, 31-38, Springer-Verlag.

Ortony, A., Clore, G., Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge Univ. Press.

Rolls, E. 1999. *The Brain and Emotion*, Oxford Univ. Press.

Roseman, I., Antoniou A., Jose, P. 1996. Appraisal Determinants of Discrete Emotions: Constructing a more accurate and comprehensive theory, *Cognition and Emotion*, 10(3):24-277.

Scherer, K. 2000. Emotions as Episodes of Subsystem Synchronization Driven by Nonlinear Appraisal Processes. In Lewis, M., Granic, I. (Eds.), *Emotion, Development, and Self-Organization*, 70-99, Cambridge Univ. Press.

Staddon, J. 2001. *Adaptive Dynamics: The Theoretical Analysis of Behavior*, MIT Press.

Webb, B., Consi, T. 2001. *Biorobotics: Methods and Aplications*, AAAI Press/MIT Press.