Österreichisches Forschungsinstitut für / Austrian Research Institute for / Artificial Intelligence

TR–2006–04

Stefan Rank, Paolo Petta

Comparability is Key to Assess Affective Architectures

• Freyung 6/6 • A-1010 Vienna • Austria •
• Phone: +43-1-5336112 •
• mailto:sec@ofai.at •
• http://www.ofai.at/ •

Österreichisches Forschungsinstitut für / Austrian Research Institute for / Artificial Intelligence

TR–2006–04

Stefan Rank, Paolo Petta

Comparability is Key to Assess Affective Architectures

The Austrian Research Institute for Artificial Intelligence is supported by the Federal Ministry of Education, Science and Culture.

Citation: Rank S., Petta P. (2006): Comparability is Key to Assess Affective Architectures. In Trappl R. (ed.), Cybernetics and Systems 2006, Austrian Society for Cybernetic Studies, Vienna, pp.643-648.

Report: Rank S., Petta P. (2006): Comparability is Key to Assess Affective Architectures. Technical Report, Österreichisches Forschungsinstitut für Artificial Intelligence, Wien, TR-2006-04.

Comparability is Key to Assess Affective Architectures

Stefan Rank[1]
[1]Austrian Research Institute
for Artificial Intelligence
Freyung 6/6
A-1010 Vienna
Austria
stefan.rank@ofai.at

Paolo Petta[1,2]
[2]Institute of Medical Cybernetics
and Artificial Intelligence
Center for Brain Research
Medical University of Vienna
Freyung 6/2, A-1010 Vienna, Austria
Paolo.Petta@MedUniWien.ac.at

Abstract

Current research on affective architectures for situated agents is fragmented into modelling aspects of emotion for specific purposes. To improve comparability of different architectures, we propose an approach that comprises the analysis of both, the *niche space* of target scenarios and the *design space* of architectures for autonomous agents. In this paper we focus on the *niche space*, and illustrate how the emotional phenomena that may occur in scenarios can be used to contextualise the investigation of functional roles of emotion within and across agents.

1 Introduction

Our research focuses on the identification of requirements that warrant emotional mechanisms in cognitive architectures [Sloman, 1999], how these mechanisms are involved in aspects of cognition and action in situated agents [Smith, 1999], and how they can play a role in bridging sensing and motor action on the one hand, and reflective and communicative faculties on the other. The variety of posited functionalities of emotional processes in humans contributes to the fragmentation of the field of affective architectures, as do the varied aims of their modelling and application. These range from simulations for psychological experiments (e.g. [Wehrle and Scherer, 2001]); over improving human-machine interaction (e.g. with character-based interfaces [Egges *et al.*, 2004; Prendinger and Ishizuka, 2005]); to fundamental improvements in the coordination of (single and multi-) agent behaviour in dynamic and unpredictable environments [Staller and Petta, 2001; Oliveira and Sarmento, 2003; Delgado-Mata and Aylett, 2005].

In the attempt to operationalise emotion mechanisms, the design of affective architectures in principle can feed back to the underlying psychological theory by revealing implicit assumptions that need to be explicated for implementation. For many of these projects, it is, however, hard to determine what specific roles of affect are actually targeted, and to discriminate their contributions from side-effects of the implementation and the particular physical embodiment. This in turn makes it difficult to compare specific designs to other approaches in similar application domains.

For this reason, we relate the functionality of emotional mechanisms in situated agent architectures to classes of application scenarios: while certain requirements are applicable to all affective architectures (e.g., design parsimony, or the requirements of flexible combination and integration of functional competences), most of them are rooted in the target scenarios of use. A scenario-based characterisation thus improves comparability by making applicable criteria explicit.

2 Comparability of Architectures

To frame our discussion of application scenarios for affective agents, we briefly introduce two quite different examples. [Baldes *et al.*, 2002; Gebhard *et al.*, 2004] describe the interactive CrossTalk installation, which is to engage passing visitors in a staged presentation of information. Two animated characters talk about cars, and sometimes switch to a meta-discussion about their current job as actors in this sales dialogue. They are presented as elaborate virtual characters that use emotional expression, but their interaction with the environment is mostly limited to their direct exchange of dialogue acts (annotated with symbolic emotion eliciting event information used in appraisal according to the OCC model [Ortony, 2003]) and registering three types of reactions from users: positive or negative feedback, or requests for help.

A very different affective application is the operationalisation of hormonal modulation to solve the two-resource problem for robots [Avila-Garcia *et al.*, 2003; Avila-Garcia and Canamero, 2005]. This environment is motivated by the behavioural problems of starvation, dithering (the rapid switching between behaviours), and competition. The robots' interaction with their environment comprises movement, collision detection (bumpers), and resource detection and consumption; following the Embodied AI approach, the overall environment is kept simple, but then as little detail as possible is abstracted away.

Scenarios as far apart as these two motivate different architectural designs, and we contend that allowing for a variety of approaches for different purposes should be preferred to searching for an only solution; and similarly, the search for the "real" answer to the question "how does cognition work" ought to be contextualised to read "how does brain-based cognition work in humans in a specific sociocultural environment". Even though in this paper we exclude the issues involved in justifying the use of the terms cognition and emotion to describe such artificial systems, we still propose that scenarios can help clarify the relation between human cognition and emotion on the one hand and artificial cognitive architectures and a possible analogue of emotion on the other.

To characterise individual affective architectures, the distinctions commonly used with models of cognition and action can be used. These include the dichotomy of symbolic and sub-symbolic reasoning to capture differences in implicit and explicit knowledge and learning [Sun and Naveh, 2004]; bottom-up vs. top-down modelling; reactivity vs. deliberation; embodied vs. reasoning-based, and online vs. offline cognition. The related hypothesis that understanding computational systems requires descriptions at different levels also has a long tradition [Marr, 1982; Pylyshyn, 1984].

To be able to relate different approaches that meet specific sets of requirements, we next develop a context-dependent analysis based on *scenario descriptions*. These descriptions explicate purpose of and motivation for building a system, detailing the desired interactions while maintaining a clear separation from implementation aspects. They form the context to evaluate applicability of particular distinctions and arguments for or against an architecture type. [Sloman, 2005] makes a similar argument concerning planning and evaluation of research[1], arguing for a focus on robotic scenarios. Scenarios thus are points in the *niche space*[2] for affective agents: the possible purposes and environments of use. The design of an agent architecture should be tailored to its intended niche.

To ultimately reveal the systematic relations between the scenarios and the architectures that implement warranted affective functionalities, a corresponding characterisation of the *design space* of architectures is also needed. The characterisation of design space is outside the scope of this paper (see [Rank *et al.*, 2005] for first terminological clarifications), but in section 4 we consider the relation of emotion to an agent's autonomy in order to arrive at a characterisation of the functionalities of emotion that affective architectures can aim for.

3 Scenarios for Affect

Scenario-based design and evaluation is an established concept in usability [Nielsen, 1993; Cooper, 1999]: specific user stories allow for iterative refinement of the actual requirements of a system. A scenario is a prototype that handles one user achieving one specific goal and thereby focuses on specific functionalities and a certain depth of the system. The analogy we use for affective architectures consists in regarding human affective behaviour as the system—a scenario is then the part that an agent architecture should be able to reproduce. We thus use scenarios to capture the *emotional potential* of an envisioned use of a system: those characteristics that help constrain which kinds of emotional phenomena can occur, and which cannot—and therefore can only be simulated or portrayed. Note that in most of today's applications of character-based interfaces, the expressed emotions are only portrayed, e.g. the simplification of interaction in the CrossTalk installation could not support a deeper model of emotion. This section provides details about the characteristics we propose.

The basic characteristics of an application scenario are the *motivation* for building the system, its *purpose*, and the *details* of a possible deployment. *Motivations* may range from specific data about humans or animals that the architecture should model (e.g. [Gratch and Marsella, 2004]) to explicit hypotheses and open empirical questions that need to be tested, and might also include specific engineering goals, e.g. the improvement of behaviour selection in robots [Moshkina and Arkin, 2003]. A characterisation of the *purpose* of the system positions it along a spectrum between real-world applications and the creation of virtual entities that can be used in controlled experiments for the scientific validation of (e.g. psychological) theories. This empirical and scientific context is closely connected to the envisioned mode of evaluation, possibly including explicit performance functions, but also less concrete design criteria that the system should meet at a social level. As to the details of deployment, a crucial point is the characterisation of the system's *interaction qualities*: this includes the *user interface* as well as the interaction between the agents and their environment. The user interface can be regarded as a special case on the spectrum of agent-environment interaction that ranges from sequenced binary decisions and sensations (e.g., in the Prisoner's Dilemma) to the complexity of human interaction in the real world. User interfaces can be as reduced as in the CrossTalk installation or as complex as in robotic applications. In the case of affective interactions, ontologies derived from folk psychology or a particular emotion theory are needed to describe the interactions possible in a scenario. We emphasise that these characterisations are mostly formulated from an *external* perspective: the objects; properties; relations; and processes described do not imply their actual use or reification in architectures tackling the scenario (symbolic models that directly operationalise folk psychological terms are one possibility, but usually lack grounding [Norling, 2004]). Interactions can be described informally as typical scenario scripts that illustrate the possible activities, including tool use and social relations, as well as the utilisation of second-order resources, complemented by negative scripts that explicate interac-

[1]See [Sloman and Wyatt, 2005] for a proposed scenario template.

[2]In evolutionary biology the role of an organism in an ecological system to which it is adapted is called its ecological niche.

tions that fall outside a given scenario.[3] A more formal description of interaction qualities could also include all agent tasks possible; agent-local performance measures (e.g., the amount of collected resources per time in a simulation); the average number of conflicting long-term or short-term tasks; and further qualitative behavioural criteria such as coherence, variedness, or believability in virtual character applications. Even though hard to quantify, the latter often form an essential part of scenario descriptions.

Another part of a scenario description is the characterisation of the environment as presented to the agent. This comprises the intrinsic limitations; dynamics; and regularities of the interactions. For simulations, this includes properties such as being time-stepped or asynchronous, with the implied differences for the possible interactions and mechanisms (cf. the well-known PEAS characterisation in [Russell and Norvig, 2003]). It is apparent that the interface to its world differs substantially from a robotic agent to a virtual one, but even in simulated environments a range of sensorimotor interactions is possible, including simple choices; artificial life simulations; and simulated physics. The scenario should also specify the numbers of agents and agent types (including interacting humans) in terms of typical and hard or practical limits[4]. For practical reasons, references to related scenarios are also helpful.

Our overall tenet is that architectures targeting related scenarios will benefit from analyses framed by scenario characteristics. This is especially true for target scenarios calling for several emotional competencies, such as interactive story-worlds [Rank and Petta, 2005a]. Our ongoing collection of scenarios in which a social lifeworld [Rank and Petta, 2005] is important provides a basis for comparative research on architectures integrating multiple affective functionalities [Rank *et al.*, 2005].

4 What Architectures Aim For

The questions that scenario descriptions are to frame are: What are different affective architectures aiming for when they implement emotional mechanisms? What is the intended role of mechanisms designed to capture (parts of) the phenomena called emotion for a system in a given scenario? In this section, we give an overview of our conceptualisation of the possible roles emotion can play in an affective architecture.

Issues that are related to problems of shallow vs. deep modelling of emotion need to be considered: Is it sensible to attempt to model only a single mechanism purported to be part of emotional phenomena? We investigate the roles of emotion as a whole as well as the roles single affective mechanisms can play in scenarios from a computer science and AI perspective, striving to distinguish ultimate functions from proximate effects [Ketelaar and Clore, 1997]. As noted in section 2, we are deferring the question of the legitimacy to refer to phenomena involving these artificial mechanisms as emotion. Further, it remains unclear whether not only the extensional but also the intensional definition of emotion varies with scenarios. It is our hope that a scenario-based survey of emotion-related functionalities in affective architectures can contribute to the clarification of elements of emotion that are constant. The functional roles attributed to emotion include not only informing an entity's decision making system about the significance of situation changes [Frijda, 1986; Sander *et al.*, 2003], but also integrating disparate architectural mechanisms, and a strategic role in structuring situated behaviour—especially in socially situated contexts. [Scheutz, 2004] presents a list of potential roles of emotion in artificial agents and [Arzi-Gonczarowski, 2002] proposes upgrades to basic reactive mechanisms that are related to emotional functionality. These different functional roles are called for in different subsets out of all possible scenarios. The generation and recognition of expressive behaviour are prominent examples of emotional functionality, as they are the focus of most work in the realm of Embodied Conversational Agents, but are less important for research on behaviour arbitration in robots.

To our mind, an overarching characterisation of emotion can be reached by considering the relation to an agent's autonomy. Autonomy is a matter of degree (cf. the important role of automaticity in human behaviour [Gollwitzer and Bargh, 2005]), and an agent needs to determine when and how it can (or needs to) exert its autonomy with respect to its environment. To rephrase, a basic need of an autonomous agent is the active construal of the subjective meaning of changes (including the passing of time) in relation to its concerns, where a concern indicates the disposition to desire occurrence or non-occurrence of a given kind of situation [Frijda, 1986], and the reconfiguration of the agent to meet the identified challenge or opportunity. This is increasingly relevant in scenarios that call for the repeated re-evaluation of construals because of ambiguities; the absence of clear-cut distinctions; or sustained real-time interaction with humans. Functionalities of single mechanisms can be regarded as subserving the appraising and coping processes. We want to highlight aspects related to the control of behaviour; perception; and the social embedding of an agent, i.e., the functionalities that become important once an agent is interacting with a group of individuals that are behaviourally similar to itself. One behavioural functionality associated with emotional processes (and an agent's personality) is the long-term persistence and coherence of behaviour [Ortony, 2003].

Process component models of emotion elaborate in detail on perceptual functions of emotional processes, i.e., the detection of constituents of emotional experience, such as novelty; pleasantness; and goal rele-

[3] A negative script for the robotic competitive two-resource problem could for example include recognising that the other robot is currently consuming, or maybe even sensing the other robot at all. A cooperative solution would then fall outside this specific scenario, but could be part of an extended version.

[4] Consider e.g. the possibility of shame in a simulation with only two agents.

vance. The more the interaction with an agent's environment is unconstrained, the more complex are principled assessments of these components. It then becomes worthwhile to provide operationalisations even of single mechanisms. Formation and retrieval of memories are also influenced by emotion, although this role may be considered a proximate effect and its ultimate functions are varied.

The social lifeworld of an agent [Rank and Petta, 2005] needs to be continually re-enacted, with support from emotional competences such as the detection of norm violations and from shared activities, i.e., routines of structured agent interactions. The knowledge about regularities and their active maintenance is as crucial in the social context as at the physical and behavioural levels. A social aspect of emotional competence therefore is the ability to exploit these regularities as a special case of (external and internal) cognitive artifacts [Hutchins, 1999]. A different perspective on the social functionality of emotion is provided by transactional accounts that describe emotional behaviour as skillful engagement with the world for relationship reconfiguration [Griffiths and Scarantino, to appear]. The strategic role of emotional behaviour, on these accounts, is to influence the behaviour of other organisms. A similar conceptualisation is at the heart of Affect Control Theory [Heise, 2002; 2004], which highlights the confirmation of expectations related to one's social role. Evidence sustaining these views includes the importance of social context, such as the effects of audiences on the likelihood of display of emotion. Examples of scenarios that warrant several of these functionalities can be found in robotic applications integrating human-robot interaction, and in virtual worlds that allow relatively unconstrained real-time interaction with multiple agents and human users. While robotic scenarios often need to be preoccupied with the details of real-world interaction, virtual worlds allow to create scenarios of different degrees of complexity. A more constrained example of the latter type of scenario are virtual story-worlds. If an interesting amount of detailed interaction should be possible without preventing dramatically interesting coherent activities, then different functionalities of emotion are needed. The coherent expression of emotion and its long-term influences, as well as the embedding in a social context are crucial for dramatic experiences.

In relatively unconstrained environments, the problem arises of relating the rich and direct interactions on a small time-scale to the variety of activities on a larger time-scale. In current agent architectures these different aspects are separated as they involve different representations and levels of abstraction: direct coupling on the one hand and symbol manipulation on the other. Comparing scenarios and investigating those that warrant both levels should help to bridge this current gap.

5 Related Work

The transactional account of emotion in [Griffiths, 2004] elaborates on the dynamic coupling of agents situated in an environment that is part of emotional phenomena. The situated perspective on emotions considers the context-dependence of any situated behaviour as a central aspect. The difficulty of explaining the social phenomenon of sulking without a transactional analysis is given as an example.

[Sloman, 2005] proposes the use of detailed scenarios to overcome difficulties in planning and focusing research on cognitive architectures. He relates these problems to *ontological blindness*: the failure of noticing the actual functions of cognition in a specific environment. The same paper also states the need for modelled competences to *scale out*, i.e., to be amenable to flexible combination and integration. In [Sloman, 2004] the same author argues for constructing a basis of concepts in emotion theories using architectural mechanisms. This research program is complemented by a description of possible mechanisms and the outline of the architecture blueprint COGAFF as an approach towards design space [Sloman and Scheutz, 2002].

[Arbib and Fellous, 2004] discern two general kinds of emotional functionality in use in robotic architectures: emotional expression for social coordination and communication and the role of emotion for the organisation of behaviour. The paper also stresses the importance of a functional view of emotions and the usefulness of the concept of an ecological niche for robotics.

6 Conclusion and Further Work

An anticipated criticism of the scenario-based approach is that scenarios as outlined here are not detailed enough to allow productive comparisons. We contend, however, that they are intended as a coordination tool for research efforts: they should not be too detailed, but iteratively adapted. This approach also allows to relate abstract sets of scenarios.

The niche space for affective architectures is vast and the shared endeavour to structure it can enhance the field even if it is not (yet) a strictly systematic effort. All scientific publications on affective architectures already provide (part of) the motivations for certain designs, and the relation of scenarios can be documented by references to related work. An ongoing survey of application scenarios of current research is to provide the context for assessing affective architectures. Especially for targeting complex scenarios that need the integration of multiple functionalities, the characterisation of architectures in relation to the scenario improves the comparability of operationalisations of emotion theories.

In the European Network of Excellence Humaine[5], our research is focused on scenarios that ask for several emotional competences that have traditionally been targeted separately and that require integration of the

[5] http://emotion-research.net/ (2006-01-28)

current state of the art of embodied approaches to emotion with symbolic models of reflection in cognition and action [Canamero, 2005].

Acknowledgments

The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry for Education, Science and Culture and by the Austrian Federal Ministry for Transport, Innovation and Technology. This research is carried out within the Network of Excellence Humaine (Contract No. 507422) that is funded by the European Union's Sixth Framework Programme with support from the Austrian Funds for Research and Technology Promotion for Industry (FFF 808818/2970 KA/SA). This publication reflects only the authors' views. The European Union is not liable for any use that may be made of the information contained herein.

References

[Arbib and Fellous, 2004] Michael A. Arbib and Jean-Marc Fellous. Emotions: from brain to robot. *Trends in Cognitive Sciences*, 8(12):554–561, 2004.

[Arzi-Gonczarowski, 2002] Zippora Arzi-Gonczarowski. AI Emotions: Will One Know Them When One Sees Them? In Trappl R., ed., *Cybernetics and Systems 2002*, Austrian Society for Cybernetic Studies, pp.739–744, 2002.

[Avila-Garcia *et al.*, 2003] Orlando Avila-Garcia, Lola Canamero and René te Boekhorst. Analyzing the performance of "winner-take-all" and "voting-based" action selection policies within the two-resource-problem. In *Proc. of the Seventh European Conference in Artificial Life (ECAL03)*, pp.733–742, Springer Verlag, LNAI 2801, 2003.

[Avila-Garcia and Canamero, 2005] Orlando Avila-Garcia and Lola Canamero. Hormonal Modulation of Perception in Motivation-Based Action Selection Architectures. In *Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action*, AISB, University of Sussex, Falmer, Brighton, UK, EU, pp.9–16, 2005.

[Baldes *et al.*, 2002] Stephan Baldes *et al.* The Interactive CrossTalk Installation: Meta-theater with Animated Presentation Agents. In Prendinger H., ed., *International Workshop on Lifelike Animated Agents: Tools, Affective Functions, and Applications*, 7th Pacific Rim International Conference on Artificial Intelligence, National Center of Sciences, Tokyo, Japan, pp.9–15, 2002.

[Brooks, 1991] Rodney A. Brooks. Intelligence without Representation. *Artificial Intelligence*, Special Volume: Foundations of Artificial Intelligence, 47(1–3):139–159, 1991.

[Canamero, 2005] Lola Canamero. *Proposal for exemplars and work towards them: Emotion in Cognition and Action* Deliverable D7d, Humaine (EU FP6 Network of Excellence, No.507422), 2005.

[Cooper, 1999] Alan Cooper. *The Inmates Are Running The Asylum.* SAMS publishing, 1999.

[Delgado-Mata and Aylett, 2005] Carlos Delgado-Mata and Ruth Aylett. Having it both ways—the impact of fear on eating and fleeing in virtual flocking animals. In Bryson J.J. et al., eds., *Modelling Natural Action Selection*, AISB, 2005.

[Egges *et al.*, 2004] Arjan Egges, Sumedha Kshirsagar, Nadia Magnenat-Thalmann. Generic personality and emotion simulation for conversational agents. *Computer Animation and Virtual Worlds*, 15(1):1–13, 2004.

[Frijda, 1986] Nico H. Frijda. *The Emotions.* Cambridge University Press, Editions de la Maison des Sciences de l'Homme, Paris, 1986.

[Gebhard *et al.*, 2004] Patrick Gebhard, Martin Klesen and Thomas Rist. Coloring Multi-character Conversations through the Expression of Emotions. In André E. et al., eds., *Affective Dialogue Systems: Tutorial and Research Workshop (ADS 2004)*, June 14–16, 2004, Kloster Irrsee, Germany, Proc., Springer Verlag, Heidelberg, pp.128–141, LNCS 3068, 2004.

[Gollwitzer and Bargh, 2005] Peter M. Gollwitzer and John A. Bargh. Automaticity in Goal Pursuit. In Elliot A.J., Dweck C.S., eds., *Handbook of Competence and Motivation*, Guilford Press, New York/London, pp.624–646, 2005.

[Gratch and Marsella, 2004] Jonathan Gratch and Stacy Marsella. Evaluating a General Model of Emotional Appraisal and Coping. In Hudlicka E., Canamero L., eds., *Architectures for Modeling Emotion: Cross-Disciplinary Foundations*, Papers from the 2004 AAAI Spring Symposium, AAAI Press, Menlo Park, CA, USA, pp.52–59, 2004.

[Griffiths, 2004] Paul E. Griffiths. Towards a 'Machiavellian' Emotions of Emotional Appraisal. In Cruse P., Evans D., eds., *Emotion, Evolution and Rationality*, Oxford University Press, pp.89–105, 2004.

[Griffiths and Scarantino, to appear] Paul E. Griffiths and Andrea Scarantino. Emotions in the wild: The situated perspective on emotion. In Robbins P., Aydede M., eds., *Cambridge Handbook of Situated Cognition*, to appear.

[Heise, 2002] David R. Heise. Understanding Social Interaction with Affect Control Theory. In Berger J., Zelditch M.Jr., eds., *New Directions in Contemporary Sociological Theory*, Rowman & Littlefield Publishers, 2002.

[Heise, 2004] David R. Heise. Enculturating Agents With Expressive Role Behavior. In Payr S. and Trappl R., eds., *Agent Culture*, Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, London, pp.127–142, 2004.

[Hutchins, 1999] Edwin Hutchins. Cognitive Artifacts. In Wilson R.A., Keil F., eds., *MIT Encyclopedia of the Cognitive Sciences*, A Bradford Book, MIT Press, pp.126–128, 1999.

[Ketelaar and Clore, 1997] Timothy Ketelaar and Gerald L. Clore. Emotion and reason: The proximate effects and ultimate functions of emotions. In Matthews G., ed., *Personality, Emotion, and Cognitive Science*, Elsevier Science Publishers (North-Holland), Amsterdam, pp.355–396, 1997.

[Marr, 1982] David Marr. *Vision.* Freeman, Oxford/San Francisco, 1982.

[Moshkina and Arkin, 2003] Lilia Moshkina and Ronald C. Arkin. On TAMEing Robots. In *Proc. IEEE International Conference on Systems, Man and Cybernetics*, October 5–8, 2003, Hyatt Regency, Washington D.C., USA, 2003.

[Nielsen, 1993] Jakob Nielsen. *Usability Engineering.* Academic Press, 1993.

[Norling, 2004] Emma Norling. Folk Psychology for Human Modelling: Extending the BDI Paradigm. In Jennings N. et al., eds., *Proc. of the third International Joint conference on Autonomous agents and multiagent systems*, 19–23 July 2004, New York City, NY, USA, IEEE Computer Society Press, Washington D.C., USA, Vol.1, pp.202–209, 2004.

[Oliveira and Sarmento, 2003] Eugenio Oliveira and Luís Sarmento. Emotional advantage for adaptability and autonomy. In Rosenschein J.S. et al., eds., *Proc. of the second international joint conference on Autonomous agents and multiagent systems*, 14–18 July 2003, Melbourne, Australia, ACM Press, New York, NY, USA, pp.305–312, 2003.

[Ortony, 2003] Andrew Ortony. On making believable emotional agents believable. In Trappl R. et al., eds., *Emotions in Humans and Artifacts*, MIT Press, Cambridge/London, pp.189–212, 2003.

[Prendinger and Ishizuka, 2005] Helmut Prendinger and Mitsuru Ishizuka. The Empathic Companion: A Character-based Interface that Addresses Users' Affective States. *Applied Artificial Intelligence*, Educational Agents and (e-)Learning, 19(3–4):267–285, 2005.

[Pylyshyn, 1984] Zenon W. Pylyshyn. *Computation and Cognition, Toward a Foundation for Cognitive Science*, MIT Press/Bradford Books, Cambridge/London, 1984.

[Rank and Petta, 2005] Stefan Rank and Paolo Petta. Motivating Dramatic Interactions. In *Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action*, AISB, University of Sussex, Falmer, Brighton, UK, EU, pp.102–107, 2005.

[Rank and Petta, 2005a] Stefan Rank and Paolo Petta. Appraisal for a Character-based Story-World. In Panayiotopoulos T. et al. eds., *Intelligent Virtual Agents, 5th International Working Conference, IVA 2005*, Kos, Greece, September 2005, Proc., Springer Berlin Heidelberg, pp.495–496, LNAI 3661, 2005.

[Rank *et al.*, 2005] Stefan Rank, Pablo Lucas dos Anjos, Ruth Aylett, and Paolo Petta. What is In an Affective Architecture for Situated Agents? In Canamero L., ed., *Humaine WP7 Workshop Proceedings*, King's College, London, July 4–5, 2005.

[Russell and Norvig, 2003] Stuart Russell and Peter Norvig. *Artificial Intelligence—A Modern Approach.* Pearson Education, London[2], 2003.

[Sander *et al.*, 2003] David Sander, Jordan Grafman, Tiziana Zalla: The Human Amygdala: An Evolved System for Relevance Detection, *Reviews in the Neurosciences*, 14(4):303–316, 2003.

[Scheutz, 2004] Matthias Scheutz. Useful Roles of Emotions in Artificial Agents: A Case Study from Artificial Life. In McGuinness D.L., Ferguson G., eds., *Proc. of the 19th National Conf. on Artificial Intelligence*, July 25–29, 2004, San Jose, CA, AAAI Press, Menlo Park, CA, pp.42–48, 2004.

[Sloman, 2004] Aaron Sloman. What Are Emotion Theories About? In Hudlicka E., Canamero L., eds., *Architectures for Modeling Emotion: Cross-Disciplinary Foundations*, Papers from the 2004 AAAI Spring Symposium, March 22–24, 2004, AAAI Press, Menlo Park, CA, pp.128–134, 2004.

[Sloman, 2005] Aaron Sloman. AI in a New Millenium: Obstacles and Opportunities. http://www.cs.bham.ac.uk/research/cogaff/sloman-ijcai05-manifesto.pdf (2005–11–30), 2005.

[Sloman and Scheutz, 2002] Aaron Sloman and Matthias Scheutz. A Framework for Comparing Agent Architectures. In *Proc. UKCI 02: UK Workshop on Computational Intelligence*, Sept.2002, Birmingham, UK, 2002.

[Sloman and Wyatt, 2005] Aaron Sloman and Jeremy Wyatt. Scenario Template http://www.cs.bham.ac.uk/research/projects/cosy /scenarios/scenario-template.txt (2005–11–30), 2005.

[Sloman, 1999] Steven Sloman. Cognitive Architecture. In Wilson R.A., Keil F., eds., *MIT Encyclopedia of the Cognitive Sciences*, A Bradford Book, MIT Press, pp.124–126, 1999.

[Smith, 1999] Brian Cantwell Smith. Situatedness/Embeddedness. In Wilson R.A., Keil F., eds., *MIT Encyclopedia of the Cognitive Sciences*, A Bradford Book, MIT Press, pp.769–771, 1999.

[Staller and Petta, 2001] Alexander Staller and Paolo Petta. Introducing Emotions into the Computational Study of Social Norms: A First Evaluation. *Journal of Artificial Societies and Social Simulation* 4(1), 2001.

[Sun and Naveh, 2004] Ron Sun and Isaac Naveh. Simulating Organizational Decision-Making Using a Cognitively Realistic Agent Model. *Journal of Artificial Societies and Social Simulation*, 7(3), 2004.

[Wehrle and Scherer, 2001] Thomas Wehrle and Klaus R. Scherer. Toward Computational Modeling of Appraisal Theories. In Scherer K.R. et al., eds., *Appraisal Processes in Emotion: Theory, Methods, Research*, Oxford University Press, Oxford New York, pp.350–365, 2001.