Tutorial: Implementing Motivation and Emotion in AI Architectures

Joscha Bach
Harvard Program for Evolutionary Dynamics
bach@fas.harvard.edu

Keywords

Artificial Emotion, Motivation, Modulation, Autonomy, Cognitive Architectures, MicroPsi

Overview

Duration: ¼ day

Goal of the Tutorial

The tutorial addresses researchers that aim at implementing plausible and non-trivial models of emotion and motivation in AI architectures, or that want to gain a deeper understanding of how to address these topics in the context of computational cognitive models, human computer interaction, affective user modeling, computer games and models of personality.

The audience will gain a conceptual understanding of capturing emotion (differences in the *how* of cognition) and motivation (the *why* of cognition), the generation of autonomous goals, motivation-driven decision-making and the regulation of action. This understanding can be applied in many practical and theoretical contexts, especially for the implementation of concrete agent architectures.

Tutorial outline

Motivation can be modeled as a set of mechanisms that are directed on the regulation of needs of an agent, specifically the selection, creation and execution of behaviors to reach situations that afford the satisfaction and avoid the frustration of needs. Human behavior is directed by specific physiological, social and cognitive needs. Much of the variation in human personality can be understood as individual differences in the parameterization of these needs. Conversely, affective states can be modeled as the configuration of cognition, via a set of modulators that capture valence, arousal, dominance/submission and the direction, focus and depth of attention. The tutorial will address these factors in detail and give suggestions for implementation in computational models.

Previous tutorials

Earlier versions this tutorial have been given at the International Conference for Cognitive Modeling (ICCM) 2006, at the summer school of the Icelandic Institute for Intelligent Machines (IIIM) 2012, and the International Conference for Artificial General Intelligence (AGI) 2014.

Audience size

Based on previous tutorials, the estimated number of participants might be 10-30.

Prerequisites

There are no strict prerequisites, but knowledge of cognitive architectures is considered helpful.

Tutorial Content

"Intelligence can be understood as the ability to reach goals over wide range of environments. Intelligent behavior is to be repeatedly successful at satisfying one's [...] needs in diverse, observably different, situations on the basis of past experience" (van Heerden 1968). Understanding intelligent behavior does not just require modeling how an agent can reach a goal, but how goals are established in the first place.

From needs to intentions

While many AI systems can operate with a fixed set of goals or tasks, autonomous behavior in open-ended environments requires a motivational system that establishes goals via learning and anticipation, by associating situations and actions with needs of the system. The MicroPsi model of motivation (Bach 2009, 2012, 2015) is an adaptation of the Psi theory of the German psychologist Dietrich Dörner (1999, 2002), and describes this relationship in detail.

The basic needs of an agent are not *adaptive* in the sense that they are learned or otherwise acquired. Instead, they are *resistive*: they direct the agent's regulation to resist the

course of events that the environment would take without the agent's intervention.

Human needs are either physiological (such as nutrition, pain avoidance, rest, libido), social (e.g., affiliation, nurturing, dominance, legitimacy, romantic affect) or cognitive (competence, exploration, aesthetics). A need can be modeled as a target value, and the current deviation from that value. Needs vary in their relative strength (compared to each other), their response to satisfaction and frustration, and the way the deviation from the target value increases over time.

Needs are signaled to the cognitive system of an agent as *urges*, which are proportional to the target deviation, and several additional parameters that depend on the situation context. The satisfaction and frustration of a need lead to *pleasure* and *pain* signals, which act as reinforcement signals for learning, and are aggregated into the situational *valence*.

A *motive* can be understood as an urge, associated with a way to satisfy this urge, or avoid its frustration. This way is usually represented as a set of goal situations (appetitive or aversive) with a plan or policy to reach that goal. *Decision-making* may raise a motive to an *intention*, by committing to a policy or goal.

Emotion as a configuration of cognition

Emotion can be understood as *valenced reactions* (Orthony, Clore, Collins 1988) to changes in the external or internal environment of an agent. At the core of these reactions is a change in the *affective state* of the agent (Mehrabian 1980). An affective state is a *configuration* of the agent's cognition, which *modulates* the cognitive processing of the agent to adapt it to the demands of the situation at hand.

Examples for cognitive modulators are *arousal* (which controls action-readiness and intensity of behavior), *valence* and *dominance* (which controls whether objects are approached or avoided), *resolution level* (depth of attention), *focus* (width of attention), *securing rate* (direction of attention; perception vs. deliberation), and *suppression* (goal stability).

A modulator may be characterized by its *baseline*, its *range*, *volatility*, and the *duration* until its return to the baseline.

Emotions can be explained by a combination of an affect (i.e. a region in the configuration space of cognition given by the modulators) and their object (which has a relevance that is defined by motivation). Differences in the parameters of motivation may be understood as giving rise to *personality* properties, while differences in the parameters of modulation give rise to variations in *temperament*.

The tutorial will discuss the dynamics of motivation and emotion, their interaction, the different effects of these mechanisms on cognitive processing, and ways of implementing them in a variety of agent models and cognitive architectures.

References

Bach, J. 2009. Principles of Synthetic Intelligence. An architecture for motivated cognition. Oxford University Press, New York

Bach, J. 2012. A Framework for Emergent Emotions, Based on Motivation and Cognitive Modulators. International Journal of Synthetic Emotions (IJSE), 3(1), 43-63

Bach, J. 2015. Modeling Motivation in the MicroPsi Architecture. Proceedings of AGI 2015, Berlin

Dörner, D. 1999. *Bauplan für eine Seele*. Reinbeck

Dörner, D., Bartl, C., Detje, F., Gerdes, J., and Halcour, D. 2002. Die Mechanik des Seelenwagens. Handlungsregulation. Verlag Hans Huber, Bern

van Heerden, P. V. 1968. The foundation of empirical knowledge: with a theory of artificial intelligence. Wistik, The Netherlands

Mehrabian, A. 1980. Basic dimensions for a general psychological theory. Oelgeschlager, Gunn & Hain Publishers, 39–53

Orthony, G., Clore, A., and Collins, A. 1988. The cognitive structure of emotions, Cambridge University Press

About the Presenter

Joscha Bach, PhD.
Research scientist

Address

Harvard Program for Evolutionary Dynamics
One Brattle Square #6
Cambridge MA 02138
USA

Phone: +1 617 669 0337
Email: bach@fas.harvard.edu

Academic Background

Diploma (MSc. equiv.) in Computer Science at Humboldt University of Berlin, 2000
PhD in Cognitive Science at University of Osnabrück 2007

2000-2003 Research assistant, department for Artificial Intelligence, Humboldt University of Berlin (Socionics, MicroPsi project, RoboCup)
2003-2005 Research assistant, lecturer, Institute for Cognitive Science, University of Osnabrück
2011-2012 Postdoctoral Fellow, Berlin School of Mind and Brain
2013-2014 Research Fellowship, Humanity Plus foundation
2014-2016 Research Scientist, MIT Media Lab

Teaching Experience:

MIT Media Lab:
Course MAS.S66 "Future Destination of Artificial Intelligence", Fall 2015
Course MAS.S63 "Cognitive Integration", Spring 2016

Humboldt University of Berlin, Artificial Intelligence:
Seminar "Socionics and Cognition", Autumn/Winter 2000/2001
Seminar "Emotional Agents", Spring/Summer 2001
Seminar "The Workshop of Emotions", Autumn/Winter 2001/2002
Seminar "Perception and Action Control of Virtual Cognitive Agents", Autumn/Winter 02/03
Seminar "Introduction to Mindbuilding", Autumn/Winter 03/04
with Prof. Dr. H.-D. Burkhard: Course "Cognitive Robotics", Academic Year 01/02, 03/04

University of Osnabrück, Cognitive Science:
Seminar "Introduction to Mindbuilding", Autumn/Winter 2003/2004
with Prof. Dr. K. Kühnberger: Course "Methods of Artificial Intelligence", Autumn/Winter 2003/2004
Seminar "Mindbuilding: Cognition and Representation", Spring/Summer 2004
with Prof. Dr. K. Kühnberger: Course "AI Perspectives on Learning", Spring/Summer 2004
Course "Cognitive HCI", Autumn/Winter 2004/2005 •
Seminar "Mindbuilding: Models of Perception in a Cognitive Architecture", Autumn/Winter 2004/2005
Course "Multi-Agent Systems", Spring/Summer 2005
Seminar "Mindbuilding: The Symbol Grounding Problem", Spring/Summer 2005
Course "Cognitive HCI", Autumn/Winter 2005/2006 •
Seminar "Mindbuilding: Topics in Cognitive Architectures", Autumn/Winter 2005/2006
Course "L'Homme Machine: An AI Perspective on Cognitive Science", Summer 2008

Humboldt University of Berlin, Institute for Philosophy:
Seminar "Computation and the Mind", Summer 2012

Previous Tutorials on Modeling Emotion and Motivation:

Bach, J., and Vuine, R. 2014. The MicroPsi Cognitive Architecture. Tutorial at the Seventh Conference on Artificial General Intelligence (AGI 2014), Quebec, CA

Bach, J. 2014. The Machine of Desire. Tutorial presenation at Google, Mountain View, September 2nd, 2014

Bach, J. 2013. Autonomy and the Computational Mind. Presentation at Icelandic Institute for Intelligent Machines, Reykjavik, Iceland, May 15th, 2013

Bach, J. 2011. Modeling Emotion as an Interaction between Motivation and Modulated Cognition. Lorentz Center, Leiden. August 19th, 2011

Bach, J., Dörner, D., Vuine, V. 2006: Psi and MicroPsi. A Novel Approach to Modeling Emotion and Cognition in a Cognitive Architecture. Tutorial at ICCM 2006, Stresa, Italy