On the functional contributions of emotion mechanisms to (artificial) cognition and intelligence

Serge Thill and Robert Lowe

Interaction Lab
School of Humanities and Informatics
University of Skövde, PO Box 408, 54 128 Skövde, Sweden
{serge.thill,robert.lowe}@his.se

Abstract. We argue that emotions play a central role in human cognition. It is therefore of interest to researchers with an aim to create artificial systems with human-level intelligence (or indeed beyond) to consider the functions of emotions in the human cognition whose complexity they aim to recreate. To this end, we review here several functional roles of emotions in human cognition at different levels, for instance in behavioural regulation and reinforcement learning. We discuss some of the neuroscientific and bodily underpinnings of emotions and conclude with a discussion of possible approaches, including existing efforts, to endow artificial systems with mechanisms providing some of the functions of human emotions.

1 Introduction

Any endeavour to construct machines with human-level intelligence (and beyond) cannot proceed without considering, at least to some extent, our understanding of human cognition and intelligence in the first place¹. At a minimum, this is required to provide an understanding of what "human-level intelligence" actually is but it may also facilitate insights into what particular mechanisms are either required or very desirable in the creation of intelligent machines.

Given the wealth of recent evidence (e.g. Stapleton, 2011; Damasio, 2010; Lowe and Ziemke, 2011; Ziemke and Lowe, 2009; Ziemke, 2008; Pessoa, 2008) that emotion and cognition are closely intertwined, we argue that one such mechanism is given by emotions. It is important to note at the outset that human emotions have evolved to meet the specific requirements of the human body. Therefore, one cannot simply "copy" features of human cognition into machines without considering what effect the difference in embodiment might have (as also previously argued by Thill, 2011). Does the human body provide mechanisms that are essential for emotions yet not realisable in machines? Do emotions provide functions that are simply irrelevant to machines? There would be little point in building an emotional machine if this doesn't somehow result in a significant advantage. While it is clear that emotions (and affect) play a central role in human cognition (see the above references), it is an open question whether a machine would

¹ For the purposes of the present paper, we follow Thill (2011) and distinguish between intelligence and cognition by defining intelligence as a *metric* of the cognitive abilities of an agent

require emotions to reach human-level intelligence. Further, it is important to note that endowing machines with emotions may be a different issue from endowing machines with the ability to *recognise* emotions, although both abilities may be important for artificially intelligent agents.

In this paper, we review research on emotions from a perspective relevant to the creation of artificially intelligent systems. We focus on functional aspects, highlighting some of the most important hypothesised roles in human cognition and interaction. We argue that, although the functions of emotions are built upon the features of human existence, implementing both equivalent artificial mechanisms and the ability to recognise human emotions are desirable features in the design of artificial intelligent machines.

2 Functions of emotions

Here, we define emotions as a functional subset of all affective phenomena. The exact function of emotions remain a topic of debate in the literature. Briefly, three major positions can be identified (Keltner and Gross, 1999):

- 1. Emotions have no functions
- 2. Emotions once served functions that are no longer necessarily appropriate
- 3. Emotions serve important functions now

Adherents of the first position generally see emotions not only as useless, but actually as a nuisance, as "disorganising forces in human behaviour" (Keltner and Gross, 1999). The second view essentially sets emotions on a par with the appendix but, in contrast with the first position, does not necessarily imply that emotions are entirely useless. Rather, it argues that whatever functions emotions serve today are not the reason emotions evolved in the first place and are probably not very important.

However, both the above views are in conflict with a large body of recent research which illustrates that emotions indeed have important roles in human cognition (Pessoa, 2008; Damasio, 2010; Lowe and Ziemke, 2011; Stapleton, 2011); as contended in the third view. For the present purposes, we can distinguish between two functional categories: *intrapersonal* (those that relate to an individual agent) and *interpersonal* (those that relate to interactions between two or more agents).

At perhaps the highest level of functional abstraction, emotions have been implicated, above all, in modulating learning (LeDoux, 1996; Rolls, 1999) and in guiding action selection and planning (Damasio, 2010; Frijda, 2010; Lowe and Ziemke, 2011). At a lower level of abstraction, emotions have been repeatedly implicated in:

- 1. homeostatic regulation: both behaviourally and internally (Sterling, 2004);
- 2. 'cognitive override' in goal-directed behaviour (Oately and Johnson-Laird, 1987; Rolls, 1999; Boureau and Dayan, 2010);
- 3. *behavioural adaptation* (*e.g.* Rolls, 1999) fundamentally concerned with the effects of emotions on learning;
- 4. *communication*: the highly influential cross cultural studies of Darwin (1872) and Ekman (2003) have suggested that expression of emotion is high on informational content

5. *social transaction*: emotional expression may generally provide a sort of social glue during agent interactions (Griffiths and Scarantino, 2009).

Intrapersonally, emotions combine 1. to 3. of the above. However, interpersonally, social interaction (4. and 5.) must also be seen as constrained or even motivated by goals and basic homeostatic needs. This intrapersonal 'grounding' is critical to understanding to what extent, and how, social interaction provides a key role of emotions. The next section is dedicated to discussing these functions in more detail.

2.1 Intrapersonal functions

Homeostatic regulation Although the human body is very adaptive to the external environment, it is also very sensitive to internal changes and can only function if internal parameters (blood pressure, levels of minerals and vitamins and so on) are kept within a very narrow range. This is ensured through homeostasis. Levenson (1999) reasons that it may occasionally be worthwhile to temporarily override this 'basal' homeostasis. For instance in a case of danger, it may be helpful to increase blood pressure, oxygen levels in the muscles of the leg and adrenaline levels to ensure a quick getaway. Thus a function of some (but certainly not all) emotions may be to override homeostasis. Levenson (1999) sees fear, anger and disgust as clear providers of such an 'emergency' function. This perspective has been echoed according to the notion of allostasis (c.f. Sterling, 2004). In Sterling's account, by overriding the (more or less) basic set points of 'essential' physiological control variables, the organism is empowered with a degree of predictive regulation. Through transiently modulating the control variables' sensitivity regime, organisms are equipped with the metabolic resources to deal with emergency situations characteristic of emotional activity (c.f. Damasio, 2010), even though a prolonged departure from the 'normal' state of the body is clearly noxious.

"Cognitive override" Levenson (1999) points out, that emotions are sometimes thought to be a 'disorganiser' of rational thoughts. However, he argues, they can in fact be understood as 'organisers'. This view is reflected in the notion of emotions serving as 'alarms' (Sloman, 2001) that "detect situations where rapid global redirecting of processing is required". The perspective of emotions as attention orienters and biasers of action selection, as well as path search, is popular both in the fields of neuroscience and artificial intelligence (c.f. Simon, 1967; Oately and Johnson-Laird, 1987; Frijda, 2010). Simon's 'interrupts' and Oatley and Johnson-Laird's 'goal juncture' redirection postulation provide purported computational functions to emotions. Oately and Johnson-Laird suggested that the "basic" emotions ('sadness', 'happiness', 'disgust', 'anxiety', 'anger'²) are elicited following perceived junctures to a plan presently enacted. These emotions serve to reconfigure the plan according to the new ('emergency') circumstances. In a weaker form, this can also manifest as a cognitive bias (Damasio, 2003).

² 'Surprise' was not considered a basic emotion.

Behavioural adaptation The above-mentioned functions of emotions as biasers, redirectors or interrupts of ongoing behaviour can be understood from a neuroscientific perspective that evidences a strong link between behaviour selection and behavioural adaptation (learning). Rolls (1999), for example, has viewed emotions as being triggered following detected 'reinforcement contingencies' in relation to learned stimulusreinforcer associations. These contingencies allude to violations of expectations concerning reward- or punishment-based returns. These violations concern: 1) immediate consequences: direct contact with a rewarding or punishing stimuli ('happiness', and 'fear', respectively); 2) anticipated consequences: unexpected presence or omission of obstacles to rewarding or punishing stimuli (precipitating 'anger' and 'relief', respectively). In this manner, Rolls has emphasized the interdependence of learning and biasing of action selection since emotion elicitation is triggered consequent to learned stimulus-reinforcer expectations. Where Rolls emphasized the neuroanatomic substrate of the reward-punishment systems constitutive of emotions (the interplay between orbitofrontal cortex and amygdala being key), Boureau and Dayan (2010) focused on the brainstem neuromodulator implementations of such systems. The implicated neuromodulators of dopamine and serotonin have been particularly linked to 'opponent process' reward and punishment based reinforcement learning. Their model, similar to Rolls, has a two dimensional flavour that links behavioural selection with adaptation. In this case, dopamine is suggested to encode for reward signals (utilized for learning) and active behavioural responding while serotonin encodes for punishment learning signals and inhibitory behavioural responding.

2.2 Interpersonal functions

Emotion expression as communication An essential question is to what extent emotional expression is of 'communicative' value. Hauser (1996) puts it thus: "in a majority of species, affective states are responsible for the production of communicative signals". However, communication implies an information exchange which implies that both expressor and perceiver gain some advantage from the communicative encounter.

An evolutionary mechanism of such communication (information exchange) has been posited. Darwin's (1872) 'principle of antithesis' proposed that emotional expressions in animals and humans have become, over evolutionary time, disambiguated for the purpose of communication: Orthogonal emotional states (e.g. fear vs anger) will be similarly expressed in a contrary manner. He took the specific example of dogs expressing anger and submission. Keltner et al. (2003) also suggest that facial emotional expressions may have evolved into disambiguated discrete forms for the benefit of communication.

In the spirit of Darwin, Ekman (2003) has accumulated much cross cultural evidence for the existence of unambiguously perceived "basic" emotions. His research indicates that a function of emotion may indeed be for communicative purposes. However, it is acknowledged that whereas perceptions of expressions may be universal, they may also be deceptive regarding the underlying emotional state; for example, a social smile is often hard to detect, by the untrained eye, relative to the 'natural', or 'duchenne' smile.

Emotion expression as social exchange A different perspective to the above holds that emotion expression, more generically, provides a sort of social glue: It disambiguates the respective roles and needs of conspecifics though not necessarily according to the equal benefit of all interacting parties.

The expressive component of emotion has been viewed by Griffiths and Scarantino (2009) as linked more to a social transaction than to pure communication. Here, the expression is not communicative in the sense of objectively expressing a cognitively held belief or intention. Instead, emotion expression facilitates a social harmony. Griffiths and Scarantino give the example of 'guilt', citing a study of Kroon (1988) in which only 28 percent of experimental subjects reporting this emotion attributed to themselves blame for the particular guilt-evoking event. In the social transaction view, guilt may be seen as promoting 'social engagement aimed at reconciliation'. This may provide a net benefit to the interactants but may not be considered pure information exchange benefitting all equally.

The means of communicative emotion expression may also be disputed. Perhaps contrary to the hitherto purported role of disambiguation, Snowdon (2003) suggests that for affect/emotion to have evolved a communicative function expression should not be particularly stereotyped or elaborate. Emphasis should be rather placed largely on perceivers discerning the relevance of expressions in a given context. He comments "we should expect little plasticity in the production and usage of calls. At the same time we can expect that it will be important to read signals accurately ... so plasticity in the development of responding to signals might be useful".

Camras (2011) has suggested that affective/emotional expression is developed according to the learned association of coordinated (facial) motor primitives. In this view, the expressed "basic" emotions as identified by Ekman may just be the products of early developmental exposure. On this basis, disambiguated 'antithetical' emotional expressions may even imply a shift in a perspective of the role of emotions in expression. Rather than being *for* information exchange regarding objectively appraised events, the primary role of expressing emotions is for manipulating the perceiver to the benefit of the expressor's bodily desires and needs.

3 Emotion components for artificial systems

As we have seen in the above, human emotions are intrinsically tied to the requirements and constraints of the body. By itself, this may be a strong indication that artificial systems may never possess emotions in the human sense of the term. However, this does not exclude the possibility of creating machine equivalents thereof; processes that mimic the functionality of human emotions to the extent that this is relevant to an artificial agent (AA). In relation to the previous section, we now evaluate the extent to which AAs may be imbued with emotions and what functional role this may serve, both for the agent itself and for human-AA interactions.

3.1 The role of the embodiment

Loosely, we may understand embodiment in terms of the intrapersonal quality of the AA: (1) what are the AA's homeostatically regulated needs; (2) what are the AA's goals

and how should it respond to goal junctures; and (3) how should the AA behaviourally adapt to unanticipated change and when?

In relation to 1, there is for instance no point in increasing the oxygen levels in the leg muscles if the machine doesn't have legs, or muscles for that matter. Nevertheless, an AA may have what Ashby (1960) identified as 'essential variables' (EVs). EVs serve as effective control variables that are required to operate within a homeostatic regime. Examples of EVs include blood glucose levels in a human, battery level in a robot, or perhaps system designer-specified performance variables in an AA. The AA is required to make a trade off so as to satisfy the set of concerns whilst not falling into irrecoverable deficits. Avila-Garcia and Cañamero (2005) have applied the idea of EVs as homeostatic control variables to robots demonstrating the potential for 'emotional' agents to produce autonomous and sustainable behaviour.

In relation to 2 and 3, a system may have many goals or aims, and may utilize principles of reinforcement learning in order to successfully arrive at them. Artificial systems that purport to explore emotional learning have often focused on neural circuitry that qualitatively replicate neurobehavioural characteristic profiles of emotional activity (Armony, 2005; Balkenius et al., 2009; Lowe et al., 2009; Ziemke and Lowe, 2009; Roesch et al., 2010). These systems are somewhat divorced from homeostatic concerns though the reward-punishment systems that they model abstractly capture such intrapersonal concerns. It is also unclear to what extent these non-homeostatically regulated learning systems gain added value from being labelled "emotional".

Nonetheless, one of the major aspects of an artificial system that can be said to be intelligent in a general sense is the ability to perform a (general) range of tasks autonomously and adaptively. Such a system will necessarily be confronted with multiple possible actions at a given time. The system will therefore need the ability to select amongst these actions. For this, homeostatic processes and mechanisms to regulate them (as given by one function of emotions discussed here) might provide a significant advantage. In addition, as discussed above, such processes may play an important role in reinforcement learning mechanisms and may therefore be equally important for an AA's learning abilities.

3.2 Expressing and recognising emotions

One has to expect that an artificial system with human-level intelligence would need to interact and communicate proficiently with humans, thus implying both a need to be able to express emotions when relevant and to be able to recognise the emotions of humans. It should of course be kept in mind that we refer to an expression of internal states here; emotion expression may thus be of no value to artificial systems if they have no internal states and no sensorimotor autonomy.

Assuming the existence of relevant internal states, the problem for 'disembodied' artificial systems is not the fact that computer hardware is not able to somehow convey emotion-like states. After all, it would be relatively easy, for example, to make a computer screen go red if the computer is angry or blue if it is sad. The problem rather concerns believability: would a red computer screen still convey the emotion as effectively as an angry face (robotic or human)? While machines may be able to express an emotion, whomever it is directed at may fail to be moved by its message. Consequently,

the function of the expression is lost. Of course people could (and probably would) learn to recognise 'computer emotions', but emotional computers should be as forthcoming to human expectations as possible in their design. In a first instance, this could well be confined to artificial 3D models of a human face or entire body displayed on the screen, which could then be animated. Much work using virtual characters of this type has been undertaken based on these principles (*c.f.* Becker-Asano and Wachsmuth, 2010).

Perhaps the most famous emotionally expressive artificial system or agent is Kismet developed at MIT by Breazeal (2003). This robot can be said to express disambiguated, but hardwired, 'basic' emotions of Ekman (2003). Kismet, therefore, is not so much a cognitive appraiser but an expressor of desires. Use of such a robot in human-robotic interactions is functional insofar as the facial expressions of Kismet are readily recognizable to human interactants. Not only does Kismet express stereotyped emotions, but is able to express according to: degree (e.g. across dimensions of 'arousal' and 'valence'); internal drives and goals mediated by an abstact homeostatic system; to cognitive appraisal of the social context; tone of voice.

As has been mentioned before, there are two sides to a communication. Thus an emotional agent would not only have to be able to convey its emotions, it would also have to be able to perceive the emotions of other persons. This is an area in which machines do rather well (see for instance work as far back as Picard, 1996, who describes a number of ways in which a computer can accurately determine a person's emotion). The aforementioned Kismet is also able to use a visual recognition system to similarly recognize such emotions in humans.

Nonetheless, a significant need for future research in expressing and recognising emotions by AAs remains. Kismet, although it can be understood as expressive emotional system, ultimately remains merely a robotic head. Producing a fully mobile robot able to detect emotional states in itself and others not limited to facial expressions (e.g. posture, gait) and elaborated social context (e.g. 'transactional', or involving aspects of deception) promise significantly greater challenges to naturalize the emotional range of an AA. Such challenges may need to be met in order to enable AAs to seamlessly integrate into human environments. Grounding higher cognitive capacities according to integrated and synchronized sensor-motoric and internal homeostatic activation patterns might be requisite to such further development.

3.3 "Feeling" emotions and higher-level cognitive functions

The state of the art concerning AAs' higher level cognitive functions, whether they concern appraisals of social/non-social events in relation to planning and action selection, has thus far tended to neglect grounding such functionality in basic bodily requirements. These requirements involve levels of internal monitoring, e.g. regarding essential variables, so as to prioritize behaviours that meet current needs and goals. Such monitoring, however, also requires an apprehension of the social and bodily context of the present—grabbing the last piece of cake at a formal meal in order to satisfy a glucose deficit is a socially inappropriate act with potential long term detrimental consequences. The artificial (and biological) agent is thus required to continuously evaluate the appropriateness and feasibility of selecting particular behaviours (and planning for such). Such mon-

itoring requires the integration and synchronization of many different inputs, external and internal to the agent, over many different timescales.

As previously argued (Lowe and Ziemke, 2011), emotional feelings, or the neural-dynamic substrate upon which they exist, may provide the means for such high level monitoring. The "feeling" of emotion, rather than being merely epiphenomenal, may provide a powerful way for learning and adapting to the outside world. Levenson (1999) argues that what we 'feel' are essentially the sensations from the physiological changes that accompany an emotion, like a particular heartbeat or breathing pattern. Damasio (2010) suggests that neural maps of the body provide inputs into multimodal maps - 'zones of convergence' - that integrate and synchronize external and internal activation patterns. Many candidates exist for sub-symbolically representing such information in artificial systems, e.g. dynamic neural fields (cf. (Lowe and Ziemke, 2011), self-organizing maps, hierarchical neural networks).

It has been mentioned above that one of the functions of "feeling" emotions may be to make us aware of our emotional state. But the fact that we are able to realise that we are in a certain state appears to presuppose a certain sense of self. Sloman (2000) argues that, if one is truly experiencing emotions, then one finds it very hard to ignore them. One cannot stop thinking about them and they may return at any point. One is thus losing control over one's thought processes. But one cannot lose what one doesn't have and thus Sloman reasons that we are both able to a) control our thought processes and b) lose control over them. AAs, similarly imbued, would have systems that note discrepancies in expected or desired states at different levels of homeostaticallostatic regulation, providing meta-levels of regulation in relation to a nested hierarchy of embodied states (Damasio, 2003, 2010, *c.f.*). Signals from reward and punishment systems, and their combined gestalt, for example, provide such information.

4 Conclusion

In the present paper, we have illustrated several functions of emotions in human cognition, showing that emotions form a central part of (human) cognition. We posit that an artificial intelligent agent cannot be expected to attain "human-like intelligence" if it does not possess at least a subset of the functional abilities provided by emotions, for instance concerning learning and adaption or behaviour selection. We have discussed some of the mechanisms underlying human emotions as well as existing work in endowing AAs with at least a rudimentary system providing emotional functionality. To conclude, we suggest that further research into the functional contributions of emotions to an (artificial) agent's cognition (and therefore intelligence), including ways of providing these functions in non-human systems, will play an important role in the creation of (generally) intelligent artificial systems.

5 Acknowledgments

This work was supported by the European FP7 project *NeuralDynamics*, (A neurodynamic framework for cognitive robotics: scene representations, behavioral sequences, and learning), Grant agreement no. 270247.

Bibliography

- Armony, J. L. (2005). Computational models of emotion. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pages 1598–1602.
- Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behaviour*. Chapman and Hall.
- Avila-Garcìa, O. and Cañamero, L. (2005). Hormonal modulation of perception in motivation-based action selection architectures. In *Proceedings of Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action, Symposium of the AISB05 Convention*.
- Balkenius, C., Morén, J., and Winberg, S. (2009). Interactions between motivation, emotion and attention: From biology to robotics. In *Proceedings of the Ninth International Conference on Epigenetic Robotics*.
- Becker-Asano, C. and Wachsmuth, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems*, 20:32–49.
- Boureau, Y. and Dayan, P. (2010). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology Reviews*, 1:1–24.
- Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human Computer Interaction*, 59:119–155.
- Camras, L. A. (2011). Differentiation, dynamical integration and functional emotional development. *Emotion Review*, 3:138–146.
- Damasio, A. (2003). *Looking for Spinoza: Joy, Sorrow and the Feeling Brain*. Harcourt, Orlando, FL.
- Damasio, A. (2010). *Self Comes to Mind Constructing the Conscious Brain*. Pantheon Books, New York.
- Darwin, C. (1872). *The Expression of Emotions in Man and Animals*. Reprinted by Julian Friedmann Publishers, London.
- Ekman, P. (2003). Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Times Books, New York.
- Frijda, N. H. (2010). Impulsive action and motivation. *Biological Psychology*, 84:570–579.
- Griffiths, P. E. and Scarantino, A. (2009). *Cambridge Handbook of Situated Cognition*, chapter Emotions in the wild: The situated perspective on emotion.
- Hauser, M. D. (1996). *The Evolution of Communication*. MIT Press/Bradford Books, Cambridge, MA.
- Keltner, D., Ekman, P., Gonzaga, G. C., and Beer, J. (2003). *Handbook of Affective Sciences*, chapter Facial Expression of Emotion, pages 415–433. Oxford University Press.
- Keltner, D. and Gross, J. J. (1999). Functional accounts of emotions. *Journal of Cognition and Emotion*, 13(5):467–489.
- Kroon, R. M. (1988). Aanleidingen en structuur van schuld gevoel. Master's thesis, University of Amsterdam, Amsterdam.
- LeDoux, J. E. (1996). The emotional brain. Simon & Schuster.

- Levenson, R. W. (1999). The intrapersonal functions of emotions. *Journal of Cognition and Emotion*, 13(5):481–504.
- Lowe, R., Humphries, M., and Ziemke, T. (2009). The dual-route hypothesis: Evaluating a neurocomputational model of fear conditioning in rats. *Connection Science*, 21(1):15–37.
- Lowe, R. and Ziemke, T. (2011). The feeling of action tendencies: on the emotional regulation of goal-directed behavior. *Frontiers in Psychology*, 2(346):1–24.
- Oately, K. and Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion*, 1:29–50.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9:148–158.
- Picard, R. (1996). *HAL's Legacy: 2001's Computer as Dream and Reality*, chapter Does HAL cry digital tears? MIT Press, Cambridge, MA.
- Roesch, E. B., Korsten, N., Fragopanagos, N., and Taylor, J. (2010). *Blueprint for Affective Computing*, chapter Emotion in artificial neural networks", pages 194–212. Oxford University Press.
- Rolls, E. (1999). The brain and emotion. Oxford University Press.
- Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, 74:29–39.
- Sloman, A. (2000). Human Cognition and Social Agent Technology, chapter Architectural Requirements for Human-Like Agents both Natural and Artificial (What Sort Of Machines Can Love?).
- Sloman, A. (2001). Beyond shallow models of emotion. *Cognitive Processing: International Quarterly of Cognitive Science*, 2(1):177–198.
- Snowdon, C. T. (2003). *Handbook of Affective Sciences*, chapter Expression of emotion in non-human animals, pages 457–481. Oxford University Press.
- Stapleton, M. (2011). *Proper embodiment: the role of the body in affect and cognition*. PhD thesis, The University of Edinburgh, UK.
- Sterling, P. (2004). *Allostasis, Homeostasis, and the Costs of Adaptation*, chapter Principles of allostasis: optimal design, predictive regulation, pathophysiology and rational therapeutics. Cambridge University Press, Cambridge.
- Thill, S. (2011). Considerations for a neuroscience-inspired approach to the design of artificial intelligent systems. In Schmidhuber, J., Thórisson, K. R., and Looks, M., editors, *Proceedings of the Fourth Conference on Artificial General Intelligence, LNAI 6830*, pages 247–254, Heidelberg. Springer.
- Ziemke, T. (2008). On the role of emotion in biological and robotic autonomy. *BioSystems*, 91:401–408.
- Ziemke, T. and Lowe, R. (2009). On the role of emotion in embodied cognitive architectures: from organisms to robots. *Cognitive Computation*, 1:104–117.