The role of emotion in adaptive behaviour and cognitive robotics

SAB '08 workshop, Osaka, Japan, 11-12 July 2008

Organizers

Dr. Robert Lowe, (Robert.Lowe@his.se)
Dr. Anthony Morse, (Anthony.Morse@his.se)
Prof. Tom Ziemke, (Tom.Ziemke@his.se)
University of Skövde, Sweden

Scope of the workshop

Affect and emotion have recently become a hot topic in the study of adaptive behaviour and embodied cognition in both natural and artificial systems. However, the regulatory role of affect/emotion, the underlying mechanisms, and the interaction between affective/emotional and cognitive processes are still not well understood. In order to develop a better understanding of the role of affect/emotion in adaptive behaviour and cognitive robotics, this workshop will bring together research on the following themes:

- ***Affective Mechanisms* –** This includes a range of mechanisms and concepts such as drives, motivation, reward, metabolic/homeostatic/allostatic regulation, appraisal, etc. We are particularly interested in computational/robotic models of such mechanisms, but also in neuroscience research and theoretical work that can inform such models. We are also interested in neurobiologically inspired models of emotion elicitation and regulation that model relevant embodied neuroanatomical structures, e.g. amygdala, hippocampus, prefrontal cortex.

- ***Emotional Agents* –** The integration of affective mechanisms in situated and embodied agents (e.g. robots) provides a crucial testing ground not only for producing emotional artefacts but also for comparing and contrasting the hypotheses and results of various emotion theories. We are interested in understanding better the relationship between emotion constituents/affective mechanisms and emotional-cognitive behaviour, how these may be measured/analyzed and formalized.

- ***Social Interaction and Human-Robot Interaction* –** The display of affective and emotional states is crucial to social interactions between people and can similarly benefit robot-robot and human-robot interactions. We are interested in models displaying inter-robot or human-robot interactions being co-ordinated or modulated by emotional expression or displays. How (proto-) affective phenomena and its expression can serve to coordinate social behaviour in more minimalist agents is also of interest.

SAB 2008 Emotion Workshop – Programme

Day 1 – July 11th

0900-0915 – **Anthony Morse**: **Introduction from organizers**

0915-1025 – **Invited Speaker: Lola Cañamero** (University of Hertfordshire)
"Emotions in Autonomous and Social Robots: Four Perspectives"

1025-1045 – Coffee break 1

1045-1125 – **Vadim Bulitko, Steven Solomon, Jonathan Gratch, Michael van Lent**
"Modeling Culturally and Emotionally Affected Behavior"

1125-1205 – **John C. Murray, Lola Cañamero**
"A Hormone Modulated Network for Influencing the Emotion Expression for a Socially Interactive Robot Head"

1205-1310 – Lunch break

1310-1350 – **Christoph Bartneck, Michael Lyons, Martin Saerbeck**
"The Relationship Between Emotion Models and Artificial Intelligence"

1350-1500 – **Invited Speaker: Ron Chrisley** (University of Sussex)
"An expectation based robotic-model of affective experience"

1500-1520 – Coffee break 2

1520-1600 – **Ernesto Burattini, Silvia Rossi**
"Periodic Activations of Behaviors and Motivational States"

Day 2 – July 12th

0900-1010 – **Invited Speaker: Marc D. Lewis** (University of Toronto)
"Neural self-organization and processes of emotion"

1010-1050 – **Francesco Mannella**, **Marco Mirolli, Gianluca Baldassarre**
"Computational Principles Underlying the Functioning of Amygdala in the Affective Regulation of Behaviour"

1050-1105 – Coffee break 3

1105-1145 – **Rob Lowe**, **Pierre Philippe, Alberto Montebelli, Tony Morse, Tom Ziemke**
"Affective Modulation of Embodied Dynamics"

1145-1200 – **Tom Ziemke**: **Summary presentation**

1200-1240 – **Discussion session**

Neural self-organization and processes of emotion

Marc D. Lewis

Department of Human Development and Applied Psychology,
University of Toronto,
Canada

Abstract. One of the principal aims of emotion theory is to model the relations between appraisals (cognitive or perceptual evaluations) and the emotions with which they correspond. However, the cognitivist approach portrays appraisals as causal antecedents of emotions in a oneway progression. The neurobiology of emotion suggests, in contrast, that appraisals and emotions emerge concurrently as self-organizing gestalts, resulting from the spontaneous coordination of multiple regions of the brain. The brain does not separate the cognitive and emotional components of these gestalts, but psychologists find these distinctions useful. In this talk, I present principles of nested feedback loops, neuromodulator activity, vertical integration of the neuroaxis, phase synchrony, and circular causality, as essential ingredients for understanding self-organizing emotional gestalts in real, living brains. I then discuss some of the implications of these principles for the role of emotion in human development.

Marc D. Lewis – Biographical Sketch

Marc Lewis is a Professor of Human Development and Applied Psychology at the University of Toronto. He specializes in the study of personality development as it relates to emotion and emotion regulation. His work is informed by developmental psychology, affective neuroscience, and a dynamic systems perspective on brain and behavior. His research has focused on transitions in emotional development and, in collaboration with Isabela Granic, he has developed a state space grid methodology for analyzing socioemotional behavior as a dynamic system. More recent work utilizes EEG methods for identifying the neural underpinnings of emotion regulation in normal and clinically-referred children and for assessing neural changes corresponding with successful treatment. His papers on the contribution of dynamic systems theory and affective neuroscience to understanding human development have appeared in high-profile journals such as *Child Development, Behavioral and Brain Sciences, Journal of Abnormal Child Psychology,* and *Development and Psychopathology.*

An expectation-based robotic model of affective experience

Ron Chrisley

COGS/Department of Informatics, University of Sussex, England

Abstract. Previous work in the SEER-3 project has focussed on exploring ways that one might use the states of a robotic model of a simple agent to serve as depictive specifications of the content of particular experiential episodes of the modeled agent. In particular, an expectation-based theory of the non-conceptual content of visual experience has been assumed as a working hypothesis, in order to investigate the depictive specifications that can be constructed as a result. An extension of this theory from non-affective to affective visual experience is proposed. As in the non-affective model, the content of the experience that the robot models at a given time is taken to be the superposition of the expected states of the sensory manifold were the robot to perform one of a specified set of actions at that time. However, in the proposed extension, the expected sensory manifold includes not just the usual (non-affective) anticipated visual sensory values, but also anticipated affective states. The dynamical notion of an affective state (with positive or negative polarity) deployed here is adapted from (Sloman, Chrisley and Scheutz 2006), roughly: positive affect states are sinks, negative affective states are sources, but the dynamic space is such that the variables that define the space are not directly in the agent's control, but require indirect control via intervention in, and interaction with, the environment. Integrating such states into the SEER-3 expectation-based architecture allows the depiction of an "affective field" to be superimposed on the non-affective visual field, thus specifying affective aspects of visual experience. This suggests a way of overcoming one of the common critiques of representationalism: that it can only handle the factual, dispassionate aspects of cognition, and must be silent concerning meaning and significance of the more engaged, affective variety.

Ron Chrisley – Biographical Sketch

Ron Chrisley is the Director of COGS, the Centre for Research in Cognitive Science at the University of Sussex, where he holds a Readership in Philosophy in the Department of Informatics. He has held various research positions in Artificial Intelligence, including a Leverhulme Research Fellowship at the University of Birmingham and a Fulbright Scholarship at the Helsinki University of Technology, as well as brief positions at NASA-Ames, Xerox PARC, the Stanford Knowledge Systems Laboratory and ATR Laboratories in Kyoto. For the past 15 years he has also been a visiting lecturer and researcher at the University of Skövde in Sweden. He was awarded his doctorate by the University of Oxford in 1997.

Modelling Culturally and Emotionally Affected Behavior

Vadim Bulitko, Steven Solomon, Jonathan Gratch, Michael Van Lent

[1] Dept. of Computing Science, University of Alberta, Edmonton, AB, T6G2E8, CANADA
bulitko@ualberta.ca
[2] USC Institute for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA, 90292, USA
{solomon,gratch}@ict.usc.edu
[3] Soar Technology, Inc., 3600 Green Court, Suite 600, Ann Arbor, MI, 48105, USA
vanlent@soartech.com

Abstract. Culture and emotions have a profound impact on human behavior. Consequently, high-fidelity simulated interactive environments (e.g. trainers and computer games) that involve virtual humans must model socio-cultural and emotional effects on agent behavior. In this paper we discuss two recently fielded systems that do so independently: Culturally Affected Behavior (CAB) and EMotion and Adapation (EMA). We then propose a simple language that combines the two systems in a natural way thereby enabling simultaneous simulation of culturally and emotionally affected behavior. The proposed language is based on matrix algebra and can be easily implemented on single- or multi-core hardware with an off-the-shelf matrix package (e.g., MATLAB or a C++ library). We then show how to extend the combined culture and emotion model with an explicit representation of religion and personality profiles.

Towards a Hormone-Modulated Model for Emotion Expression in a Socially Interactive Robot Head

J.C. Murray and L. Cañamero

Adaptive Systems Group
University of Hertfordshire
STRI & School of Computer Science
College Lane, Hatfield, Hertfordshire
AL10 9AB, UK
john@jcmurray.com, l.canamero@herts.ac.uk

Abstract. In this paper[1] we present a robot head ERWIN capable of human-robot interaction, endowed with interactive mechanisms for allowing the emotional state and expression of the robot to be directly influenced by the social interaction process. Allowing the interaction process to influence the expression of the robot head can in turn influence the way the user interacts with the robot, in addition to allowing the user to better understand the intentions of the robot during this process. We discuss some of the interactions that are possible with ERWIN and how this can affect the responce of the system. We show an example scenario where the interaction process makes the robot go through several different emotions.

1 Introduction

The concept of imbibing robots with emotional functionality, either via internal structures and models or with the ability to express a particular emotion is a concept that has been pondered for many years but only something that has been researched in greater detail in recent years. With the desire to bring the interaction between Robots and Humans one step closer to that of Human-Human interaction then the interaction process needs to be as natural as possible [8]. In order to allow for better human-robot interaction, it is not only necessary for robots to communicate but also to automatically adapt their behaviour based on feedback from the human, and the visual and auditory modalities. This automatic adaption and learning is particularly important as it allows the human to feel that the interaction with the robot is being conducted in a more natural manner.

[1] This research is supported by the European Commission as part of the FEELIX GROWING project (http://www.feelix-growing.org) under contract FP6 IST-045169. The views expressed in this paper are those of the authors, and not necessarily those of the consortium.

2 ERWIN's Capabilities

ERWIN stands for **E**motional **R**obot **w**ith **I**ntelligent **N**etworks, and is a robot head capable of basic social interaction. ERWIN draws on several modalities in order to provide a rich interactive system. These modalities include the visual and acoustic domains. In addition to this, ERWIN is capable of expressing several emotions as described in section 2.3. The head is built from a pan and tilt mechanism, with two 1.3Mpix CCD cameras used as eyes for the system, and two microphones used as ears. There are four servos controlling the eyebrows and mouth (a top and bottom lip), each with 1 DOF.

2.1 Visual Tracking

Using two cameras it is possible to use many of the available visual cues [13] to improve the social interaction ability of ERWIN. During 'local proximity' social interaction, vision plays an important role, providing much-needed information on whom we are speaking to, directing our attention towards whom we are interacting with, and allowing for the use of gestures [17]. When we communicate with someone, we usually expect them to give us their attention by looking at us. Using the OpenCV library it is possible to detect faces within a visual scene; any faces within the visual field of the robots cameras can thus be detected and their position within the image determined. This allows for the robot to stay focused on the person keeping the feeling of one-to-one attention, as the person moves to the left or right of ERWIN's Field-of-View the head will maintain a visual track on them.

ERWIN can analyse a frame from the cameras approximately every $40 - 80$ms in order to determine if a face can be detected within the scene. Once a positive match for a face is detected, the center point of the face is calculated and the Euclidean distance from the center point of the camera image taken. Using this the pan and tilt of the robot head is adjusted accordingly.

Face Recognition ERWIN's visual system is also capable of basic facial recognition allowing previously seen faces to be recognised and new faces to be learnt by the system. This adds an extra dimension to the interaction process as it allows for the robot to react differently depending on if it recognises a face or not, and therefore allowing 'familiarity' to be coded.

The first stage of the face recognition process uses the OpenCV library and is based on Viola's [20] rapid object detection process with improvements made by Lienhart [12] based on Haar-like features for object detection [10]. Using an intermediate representation of the image, termed the integral, the features are rapidly detected. This integral contains the sums of the pixels' intensity values located directly to the left and above the pixel at (x,y) inclusive.

The object classifiers for face detection have been trained using the FERET [18] database. Once a face is detected in the scene it is processed to extract certain features allowing for future recognition. For this we use the relative positions of significant elements present on all faces; the eyes, nose and mouth. Therefore, we train several independent classifiers to detect these elements. Again the images used for training the classifiers are taken from the FERET database. The extracted face image is normalised to 250px in width, maintaining a resizing ratio of 1:1 this ensures that the distances between features remain stable even when the distance from ERWIN changes. The second stage of processing involves splitting the face image into three sections and applying the appropriate classifier on each part. Figure 1 A) shows how the detected face is split for the relevant features.

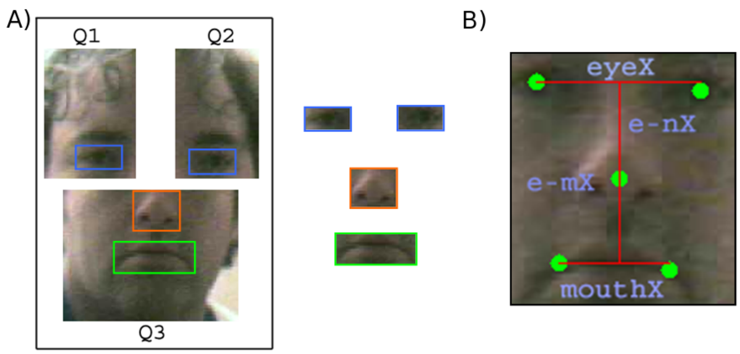

Fig. 1. A) The detected face split into three sections for further processing, B) The metrics measured for facial recognition.

Once the features are detected, their positions are used to calculate a distance metric in order to recognise a particular face. The metrics recorded are: distance between eyes (eyeX), distance from eye level to nose (e-nX), width of mouth (mouthX) and distance from eye level to mouth level (e-mX). These values are given in number of pixels, and they are normalised using Eq. 1 to bring the values to between 0 and 1.

$$I_n = \left(\frac{1 + (x - A)}{B - A} - 1 \right) \tag{1}$$

These values are presented to a neural network which determines if the face has been seen before. If not, the network retrains adding in this additional information to the already recognised set of faces. Fig. 1 B) shows the face feature metrics used for recognition. Network training is performed using back-propagation as shown by equations 2 and 3, with an empirically determined sum-squared-error tolerance of 0.02. On presenting a face for recognition, a value below 0.9 is classed as non-recognition.

$$\Delta w_{ij} = \eta \delta_j x_i \tag{2}$$

ΔW_{ij} is the update for the weight between units i and j, η is the learning rate set to 0.2, δ_j is the error of unit j and x_i the activation of unit i. For back-propagation the weight change for pattern $n+1$ is dependant on the weight change for n.

$$\Delta w_{ij}(n+1) = \eta(\delta_j o_i) + \alpha \Delta w_{ij}(n) \tag{3}$$

α is the learning rate, η is introduced to help prevent weight oscillations and o_i is the output of unit i.

2.2 Sound-Source Localisation

Acoustics is also an important modality for social interaction, especially for humans, as our predominant form of communication is that of words and sounds [14] [19] [1]. ERWIN is therefore equipted with stereo microphones as ears to allow for speech processing and sound-source localisation to take place. This dramatically improves the social interaction process by allowing for auditory cues and abilities to be exploited by the system.

The sound-localisation capabilities of ERWIN are implemented using the Interaural Phase Difference cue [11], that is, the phase difference of the signals arriving at the two microphones. The received sound signals are processed for phase by cross correlation. Cross correlation processes the two signals *g(t)* and *h(t)* recorded by the two microphones by incrementally sliding them across each other to determine the offset between points of maximum similarity. Equation 4 shows the formulae for computing the cross correlation of two data series.

$$Corr(g,h)_j \equiv \sum_{k=0}^{N-1} g_{j+k} k_k \tag{4}$$

The largest value within this vector corresponds to the point of maximum similarity between the two signals, i.e. the two signals being in phase with each other. The position of this value within the vector is then used to determine the Interaural Time Difference of the signals and ultimately the azimuth angle. Figure 2 shows the process of cross correlation on two signals with a resultant correlation vector. Finally, to determine the azimuth of the sound-source Eq. 5 is used.

$$\Theta = Sin^{-1}\frac{a}{c} = Sin^{-1}\frac{(\Delta \times \sigma) \times c_{air}}{c} \tag{5}$$

where σ is the lag (number of offsets) between the signals *g(t)* and *h(t)* determined from the correlation vector $\boldsymbol{C}$, Δ is the sample time increment as determined by the sample rate, i.e. $1/44100 = 22.7\mu s$ and Θ is the angle of incidence and the speed of sound taken to be $348m/s$ at $24C$.

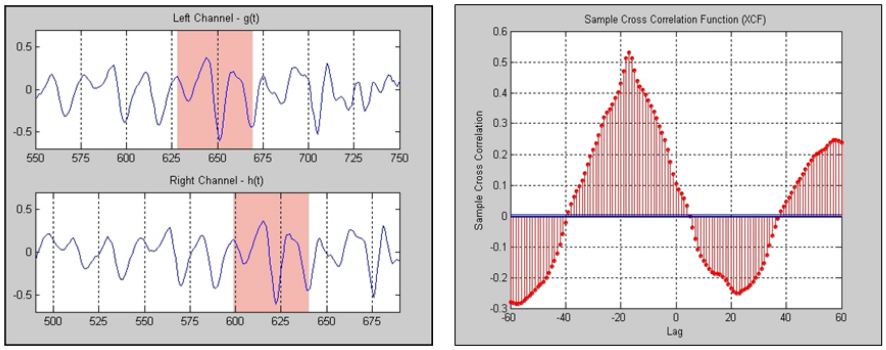

Fig. 2. Cross correlation of the signals *g(t)* and *h(t)* with the resultant cross correlation vector showing the maximum point of similarity.

2.3 Emotional Expression

Emotion in robotics may be a an emerging research area [2, 3] [4] [6] However, philosophers and researchers in biology, psychology and cognative science have debated its meaning and underlying mechanisms for many years. For this reason, we refer to our emotional expressions in terms of Ekman's simple emotions [5]. From these set of emotions, the ones that can be expressed by ERWIN are: Happy, Sad, Fear, Surprised and Angry. Fig. 3 shows how each of the five different emotional expressions look on ERWIN.

These emotions can only be expressed one at a time and independently of each other, and are used to provide additional visual feedback about the interaction process to the user. In our previous models these emotional expressions have been coded directly into the system and set to activate at specific points within an interaction determined by the code and not directly by the interaction process per se. However, the model we present in this paper makes the system much more dynamic allowing for the interaction process and its responses to have a more direct influence over the emotion expression.

3 Social Interaction

One of the goals of the robot head presented in this paper is to allow the interaction process between the robot and human to feel dynamic making the human user feel that they are dealing with a system that is fluidly reacting to their interaction [2]. This is achieved with the use of visual and acoustic cues in addition to the displaying of various emotions by ERWIN. We hope that with the inclusion of these cues and direct emotional feedback that the process of social interaction will become more natural, without the need of prerequisite knowledge on the part of the human. We also feel that this will allow the human partner to gauge how their interaction and responses are being interpreted by

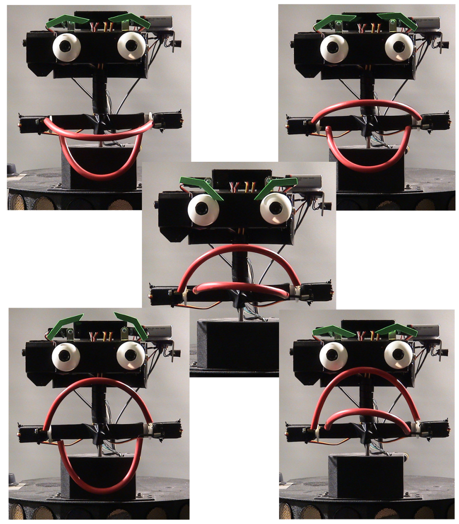

Fig. 3. The emotions expressed by ERWIN in terms of Ekman's basic emotions.

the robot, ultimately allowing them to adapt their behaviour to accommodate for that of the robots. For example, they may find that their speech is being missinterpreted and thus the robot begins to express Anger. Humans can therefore actively change how they are communicating with ERWIN and change their responses accordingly.

3.1 Parameters for Emotional Expression

As previously discussed in section 1, the robot head presented within this paper is capable of expressing several basic emotions. The main focus of this paper is to present our model of dynamic control of these emotions and expression via hormonal like parameters, allowing for the interaction process to change the levels controlling each of the emotions and thus invoking a specific interactively determined response. We choose to develop our model using hormonal like parameters as it allows us to model the dynamics of emotional expressions. Although the current parameters are not based on a biological model of hormonal control for emotional expression we hope that they can prove a tool to gain initial insight into the relationship between emotional expression generation via social interaction, and underlying biological emotional mechanisms.

The goal here is to be able to modify the functioning of the system by modulating parameters within our emotional activation model that control each of the different emotions [7] [16]. When interacting with people, the responses we receive via this interaction process determine or affect our current emotional state, for example we may get frustrated if we can not understand someone, or become sad or angry if we are ignored. To improve the interaction process, bringing it one step closer to that of human-human interaction, the emotional expression of the robot should be a dynamic process based on this interaction.

Each of the five emotional expressions have a particular "hormonal" value associated with them. The system initialises with each of the five values set to a baseline of 0.5 with a minimum of 0.0 and maximum of 1.0. We initially attach several pseudo emotional meanings to the possible expressions in order to better understand what is transpiring during the interaction process. The following scenarios are employed: Sad - when no interaction has taken place for x amount of time or when a response is given that warrants the Sad expression. Happy - when there is a face within the visual field of the robot, and the interaction process goes without incident. Angry - is assigned to express when during the interaction the responses from the user can not be interpreted. Surprised - this emotional expression is used when ERWIN detects a face in the visual scene that is then recognised as a face previously seen.

Activating an Emotion Currently there are six forms of interaction that can influence the emotional expression of the system. Table 1 details the possible interactions and their influence over the hormonal levels. As can be seen from table 1 each of the various interaction functions affect different emotional expression hormonal parameter levels either by increasing them (excitory +) or decreasing (inhibitory -).

Table 1. The hormonal parameter levels affected by the varying interaction modalities.

Hormonal Control				
Modality	Function	Affects	Active	Inactive
Vision	Face Detection V_{FD}	$\sigma_H \sigma_{SA}$	+0.4, -0.2	-0.05, +0.01
	Face Tracking V_{FT}	$\sigma_A \sigma_H \sigma_{SA}$	-0.04, +0.02, -0.01	+0.01, -0.01, +0.01
	Face Recognition V_{FR}	$\sigma_{SU} \sigma_{SA}$	+0.8, -0.5	-0.05, +0.0
Speech	Orientation S_O	$\sigma_H \sigma_{SU}$	+0.4 +0.5	-0.02, -0.05
	Tracking S_T	$\sigma_H \sigma_{SU}$	+0.1, +0.2	-0.005, -0.01
	Communication S_C	$\sigma_A \sigma_H$	-0.1, +0.2	+0.1, -0.04
Silence	Non-Interaction S_{NI}	σ_{SA}	+0.05	-0.1

Currently ERWIN is only capable of expressing one emotion at a time. For this reason, the emotion expression mechanisms is based on a winner-takes-all approach. That is, the hormonal parameter levels with the highest value will take control of the expression. In addition, there is a threshold that must be reached in order to express the emotion - for initial tests we used 0.8. Each of the various levels are in a constant flux, either being inhibited or excited by the various interaction modalities being used. As we attribute different 'arbitrary' levels of importance to each of the interactive methods, their activation increases or decreases the hormonal levels differently. Table 1 shows the effect each of the social interaction methods have on the hormonal levels.

4 Results and Discussion

As each interactive scenario with the robot and a user will be different, either through the responses given in answer to questions or the general interactive environment, the modulation of the various hormonal levels will vary accordingly. Each of these levels have a decay rate as shown in table 1 within the 'inactive' column; this shows how they decay at each time step. In addition, the increase in activity for the a number of the hormones takes place during each time step, if that particular interaction method is active, other interactive methods only affect their respective hormonal levels once per interaction, that is when they are presented. These hormonal activations will then be unable to increase until they have dropped below a particular threshold; one example of this is σ_{SU}, when a familiar face is detected will increase dramatically, but needs to drop below a lower threshold before it can activate again.

Figure 4 shows an interactive process between ERWIN and a human, and the resulting emotion level responses and their effect. The interaction experiment shown in Fig. 4 begins with no user interaction with the head and therefore increasing the 'Sad' levels σ_{SA}. After five time steps the robot expresses its Sad emotion as the level passes the 0.8 threshold. We then introduce a new face to the system, which is detected and therefore increases the 'Happy' levels σ_H enough to express the emotion allowing the user to know that they have been seen by ERWIN and allowing the interaction to begin. Once a face is detected, it is continuously tracked. ERWIN then asks a question and the first time a response is expected the user stays silent, this occurs at between 17 and 21 time steps. As can be seen, the level of σ_A increases but not enough to invoke an expression response; then the question is repeated and again the user remains quiet. This time the levels of σ_A rises above the threshold at time step 29 and the 'Angry' emotion is expressed.

As is shown by this scenario, it is possible to see how dynamically changing the emotion expressions of ERWIN informs the user as to how their interaction is being taken. With the use of dynamically modulating the control of emotional expression, the different interactive modalities can inhibit and excite certain parameter levels. This can prevent the expression of an emotion, depending on how the previous and current aspects of the interaction are proceeding, whereas this would not be possible by statically encoding a particular interactive response with a particular emotion.

5 Future Work

As previously mentioned, the development of the emotional expression model will look more into the biological aspects of hormone modulation for emotions, looking in greater detail at the underlying mechanisms that control the expression of emotion. For example, one such model using arousal levels that vary as a function of tactile stimulation patterns has been used in [4] for robotic emotion expression. Further to this, currently the model presented in this paper is only

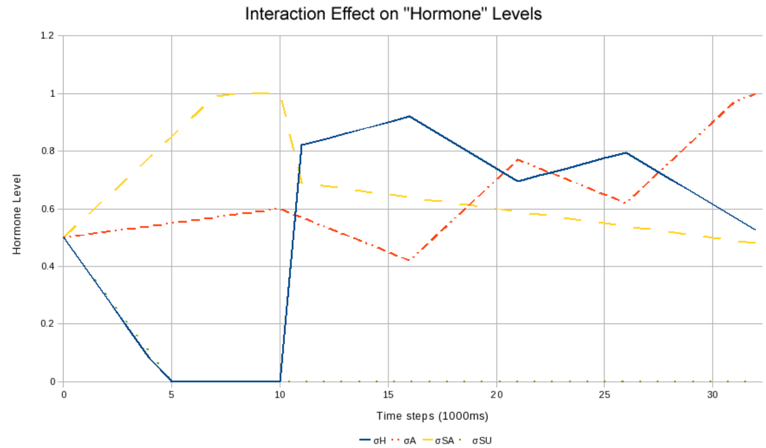

Fig. 4. Shows the response of the hormone levels during an interaction phase.

capable of expressing a single emotion at a time, and lacks an expressional model of transitioning between expressions.

The model presented also adapts its emotional expression too quickly during social interaction. That is, even if the current emotional level for 'Anger' is very high, simply conforming to the interaction process can quickly bring back a balance for 'Happy'. Therefore, in further developing our model we hope to introduce some form of hysteresis into the system; this would allow the expression levels to change not only based on the social interaction, but also depending on previous interactions. With the use of face recognition this could allow ERWIN to change hormone levels not only based on the social interaction but also on the interaction partner.

References

1. M. Bennewitz, F. Faber, D. Joho and S. Behnke, *Fritz - A Humanoid Communication Robot*, 16thd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Jeju Island, Korea, August 2007.
2. C. Breazeal, *Toward sociable robots*, Robotics and Autonomous Systems, Vol. 42, Issues 3-4, pp. 167-175, 2003.
3. L. Cañamero, *PLAYING THE EMOTION GAME WITH FEELIX - What Can a LEGO Robot Tell Us about Emotion?*, In: SOCIALLY INTELLIGENT AGENTS - Creating Relationships with Computers and Robots. Eds. K. Dautenhahn, A.H. Bond, L. Cañamero and B. Edmonds. Kluwer Academic Publishers, ISBN: 1402070578, Pages.69-76.
4. L. Cañamero and J. Fredslund, *I Show You How I Like You-Can You Read it in My Face?*, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Vol. 31, No. 5, Pages 454-459, Sep. 2001.

5. P. Ekman, *Facial Expression and Emotion*, American Psychologist, Vol. 48, Issue 4, Pages 384-392, April 1993.
6. T. Fong, I. Nourbakhsh and K. Dautenhahn, *A survey of socially interactive robots*, Robotics and Autonomous Systems, Vol. 42, Issues 3-4, Pages 143-166, March 2003.
7. O. Avila-García and L. Cañamero, *Hormonal Modulation of Perception in Motivation-Based Action Selection Architectures*. In L. Cañamero (Ed.), Proc. Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action, Symposium of the AISB'05 Convention, University of Hertfordshire, UK, April 14-15, 2005.
8. A. Kerepesia, E. Kubinyib, G.K. Jonssonc, M.S. Magnussonc and A. Miklosib, *Behavioural comparison of human-animal (dog) and human-robot (AIBO) interactions*, Behavioural Processes, Vol. 73, Issue 1, Pages 92-99, July 2006.
9. C.G. Kohler, T. Turner, N.M. Stolar, W.B. Bilker, C.M. Brensinger, R.E. Gur and R.C. Gur, *Differences in facial expressions of four universial emotions*, Psychiatry Research, Vol. 128, Issue 3, Pages 235-244, October 2004.
10. A. Kuranov, R. Lienhart and V. Pisarevsky, *An Empirical Analysis of Boosting Algorithms for Rapid Objects With an Extended Set of Haar-like Features* Intel Technical Report MRL-TR, July, 2002.
11. J.C.R. Licklider *Three auditory theories*, In: Psychology: A Study of a Science, ed Koch, E. S., Study 1, Vol. 1, New York: McGraw-Hill, pp. 41-144, 1956.
12. R. Lienhart and J. Maydt, *An Extended Set of Haar-like Features for Rapid Object Detection*, IEEE ICIP, Vol. 1, pp. 900-903. 2002
13. G. Medionia, A.R.J. Françoisa, M. Siddiquia, K. Kima and H. Yoonb, *Robust real-time vision for a personal service robot*, Computer Vision and Image Understanding, Vol. 108, Issues 1-2, 2007.
14. J. Murray, H. Erwin and S. Wermter. *Auditory Robotic Tracking of Sound Sources using Hybrid Cross-Correlation and Recurrent Network*, International Conference on Intelligent Robots and Systems, Edmonton, Canada, 2005.
15. J. Murray, C. Rowan, A. Yau, M. Elshaw, S. Wermter, *Sound localisation and emotional language communication in the Mira robot head*. Proceedings of the AI Video Competition at 22nd AAAI Conference on Artificial Intelligence, 2007.
16. M.J. Neal, *An Artificial Immune System for Continuous Analysis of Time-varying data*, In proceedings of ICARIS 2002, Canterbury, U.K.
17. K. Nickel and R. Stiefelhagena, *Visual recognition of pointing gestures for human-robot interaction*, Image and Vision Computing,Vol. 25, Issue 12, Pages 1875-1884, December 2007.
18. P.J. Phillips, H. Moon, P.J. Rauss and S. Rizvi, *The FERET evaluation methodology for face recognition algorithms*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, October 2000.
19. Y. Shaoa,S. Srinivasanb, Z. Jina and D. Wanga, *A computational auditory scene analysisnext term system for speech segregation and robust speech recognition*, Computer Speech and Language. March 2008.
20. P. Viola and M.J. Jones, *Rapid Object Detection using a Boosted Cascade of Simple Features*, IEEE CVPR, 2001.

The Relationship Between Emotion Models and Artificial Intelligence

Christoph Bartneck[1], Michael Lyons[2], Martin Saerbeck[1,3]

[1] Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2, 5600MB Eindhoven, The Netherlands

[2] College of Image Arts and Sciences
Ritsumeikan University
56-1 Tojiin, Kitamachi, Kita-ku, Kyoto, Japan, 603-8577

[3]Philips Research
Connected Consumer Services Group
High Tech Campus 34, 5656AE Eindhoven, The Netherlands

c.bartneck@tue.nl, michael.lyons@gmail.com, martin. saerbeck@philips.com

Abstract. Emotions play a central role in most forms of natural human interaction so we may expect that computational methods for the processing and expression of emotions will play a growing role in human-computer interaction. The OCC model has established itself as the standard model for emotion synthesis. A large number of studies employed the OCC model to generate emotions for their embodied characters. Many developers of such characters believe that the OCC model will be all they ever need to equip their character with emotions. This study reflects on the limitations of the OCC model specifically, and on the emotion models in general due to their dependency on artificial intelligence.

Keywords: emotion, model, OCC, artificial intelligence

1 Introduction

Marvin Minsky boldly stated that "The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions" [1]. In this study, I will reflect on the relationship between emotion modeling and artificial intelligence and show that Minsky's question is still open. Emotions are an essential part of the believability of embodied characters that interact with humans [2-4]. Characters need an emotion model to synthesize emotions and express them. The emotion model should enable the character to argue about emotions the way humans do. An event that upsets humans, for example the loss of money, should also upset the character. The emotion model must be able to evaluate all situations that the character might encounter and must also provide a structure for variables influencing the intensity of an emotion. Such an emotion model enables the character to show the right emotion with the right intensity at the right time, which is necessary for the convincingness of its emotional expressions [5]. Creating such an emotion model is a daring task and in this section I will outline some of its problems. In particular, I will argue for the importance of the context in which the emotion model operates.

Emotions are particularly important for conversational embodied characters, because they are an essential part of the self-revelation feature of messages. The messages of human communication consist of four features: facts, relationship, appeal and self-revelation [6]. The inability of a conversational character to reveal its emotional state would possibly be interpreted by the user as missing sympathy. It would sound strange if the character, for example, opened the front door of the house for the user to enter and spoke with an absolute monotonous voice: ”Welcome home”.

The OCC Model

From a practical point of view, the developer of a screen character of robot is wise to build upon existing models to avoid reinvent the wheel. Several emotion models are available [7, 8]. However, Ortony, Clore and Collins [9] developed a computational emotion model, that is often referred to as the OCC model, which has established itself as the standard model for emotion synthesis. A large number of studies employed the OCC model to generate emotions [2-4, 10, 11]. This model specifies 22 emotion categories based on valenced reactions to situations constructed either as being goal relevant events, as acts of an accountable agent (including itself), or as attractive or unattractive objects (see Figure 1). It also offers a structure for the variables, such as likelihood of an event or the familiarity of an object, which determines the intensity of the emotion types. It contains a sufficient level of complexity and detail to cover most situations an emotional interface character might have to deal with.

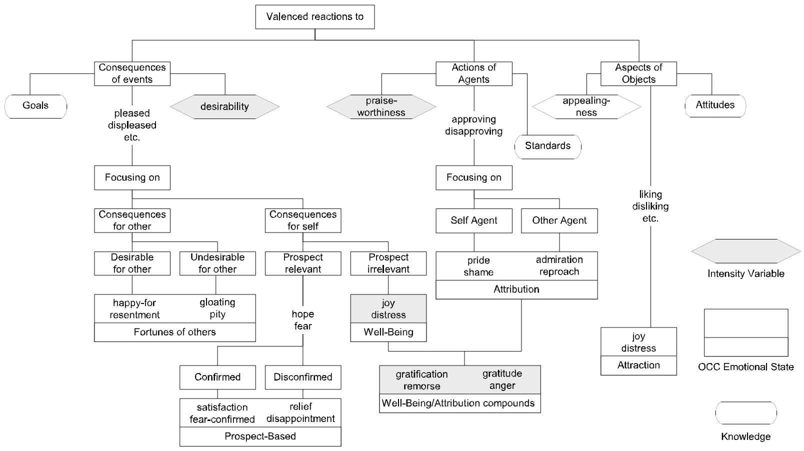

Fig. 1. The OCC model of emotions.

When confronted with the complexity of the OCC model many developers of characters believe that this model will be all they ever need to add emotions to their character. Only during the development process the missing features of the model and the problem of the context become apparent. These missing features and the context in which emotions arise are often underestimated and have the potential to turn the character into an unconvincing clown. I will point out what the OCC model is able to do for an embodied emotional character and what it does not.

The OCC model is complex and this paper discusses its features in terms of the process that characters follow from the initial categorization of an event to the resulting behavior of the character. The process can be split into four phases:

1. Categorization - In the categorization phase the character evaluates an event, action or object, resulting in information on what emotional categories are affected.
2. Quantification - In the quantification phase, the character calculates the intensities of the affected emotional categories.
3. Interaction - The classification and quantification define the emotional value of a certain event, action or object. This emotional value will interact with the current emotional categories of the character.
4. Mapping - The OCC model distinguishes 22 emotional categories. These need to be mapped to a possibly lower number of different emotional expressions.

Categorization

In the categorization phase an event, action or object is evaluated by the character, which results in information on what emotional categories are affected. This categorization requires the character to know the relation of a particular object, for example, to its attitudes. Depending on this evaluation either the “love” or “hate” emotional category will be affected by the object.

Consider this example: a character likes bananas and the user gives him a whole bunch. The character will evaluate the consequences of the event for the user, which results in pity, since the user has a whole bunch of bananas less. It will also evaluate the consequences of the event for itself, which results in satisfaction because it received a bunch of bananas. Next, it evaluates the action of the user, which results in admiration and finally the aspect of the object, which results in love. It appears that ironic that the category “love” is being used in the OCC model only for objects, since the more important usage for this word is certainly found in human-human relationships.

To do this classification the character needs an extensive amount of knowledge. First, it needs to know its relationship to the user, which was assumed to be good. Hence, pity is triggered and not resentment. Moreover, it needs to know what this event means to the user. Otherwise the character’s happy-for category might be triggered (User Model). Second, it needs to have a goal “staying alive” to which the bananas contribute (Goals). Third, it needs to know what to expect from the user. Only knowing that the user does not have to hand out bananas every other minute the character will feel admiration (Standards). Last, it needs to know that it likes bananas (Attitudes).

The standards, goals and attitudes of the character that the OCC model requires need to be specified, organized and stored by the designer of the character. A new character knows even less than a newborn baby. It does not even have basic instincts. One way to store this knowledge could be an exhaustive table in which all possible events, actions and objects that the character might encounter are listed together with information on which emotional categories they affect and how their intensity may be calculated. This approach is well suited for characters that act in a limited world. However, it would be rather difficult, for example, to create such an exhaustive list for all the events, actions and objects that the character might encounter at the home of the user. With an increasing number of events, actions and objects, it becomes necessary to define abstractions. The bananas could be abstracted to food, to which also bread and coconuts belong. The categorization for the event of receiving food will be the same for all types of food. Only their intensity might be different, since a certain food could be more nutritious or tasty. However, even this approach is inherently limited. The world is highly complex and this approach can only function in very limited “cube” worlds.

This world model is not only necessary for the emotion model, but also for other components of the character. If, for example, the character uses the popular Belief, Desires and Intention (BDI) architecture [12], then the desires correspond to the goals of the emotion model. The structure of the goals is shared knowledge. So are the standards and attitudes. The complexity of the OCC model has a direct influence on the size of the required world model. However, the AI community has long given up the hope to be able to create extensive world models, such as the widely known Cyc database. The amount of information and its organization appears overwhelming. Only within the tight constraints of limited worlds was it possible so far to create operational world models.

As mentioned above, the OCC model distinguishes 22 emotional categories (see Figure 1). This rather cumbersome and to some degree arbitrary model appears to be too complex for the development of believable characters [13]. The OCC model was created to model human emotions. However, it is not necessary to model a precise human emotion system to develop a believable character. A "Black Box" approach [14] appears to be sufficient. The purpose of this approach is to produce outcomes or decisions that are similar to those resulting from humans, disregarding both the processes whereby these outcomes are attained as well as the structures involved. Such a "Black Box" approach is more suitable, particularly since the sensory, motoric and cognitive abilities of artificial characters are still far behind the ones of humans. The characters emotion system should be in balance with its abilities. Several reason speak for a simplification of the OCC model.

First, only those emotional categories of the OCC model should be used that the character can actually use. If a character uses the emotional model only to change its facial expression then its emotion categories should be limited to the ones it can express. Elliot [2] implemented all 22 emotional categories in his agents because they were able to communicate each and every one to each other. This is of course only possible for character-character interaction in a virtual world. It would be impossible for characters that interact with humans, since characters are not able to express 22 different emotional categories on their face. Ekman, Friesen and Ellsworth [15] proposed six basic emotions that can be communicated efficiently and across cultures through facial expressions.

Second, some emotional categories of the OCC model appear to be very closely related to others, such as gratitude and gratification, even thought the conditions that trigger them are different. Gratification results from a praiseworthy action the character did itself and gratitude from an action another character did. It is not clear if such a fine grained distinction has any practical advantages for the believability of characters.

Last, if the character does not have a user model then it will by definition not be able to evaluate the consequences of an event for the user. In this case, the "fortunes of others" emotional categories would need to be excluded. Ortony acknowledged that the OCC model might be too complex for the development of believable characters [13]. He proposed to use five positive categories (joy, hope, relief, pride, gratitude and love) and five negative categories (distress, fear, disappointment remorse, anger and hate). Interestingly, he excluded the emotional categories that require a user model. These ten emotional categories might still be too much for a character that only uses facial expressions. Several studies simplified the emotional model even further to allow a one-to-one mapping of the emotion model to the expressions of the character [3, 16].

Quantification

The intensity of an emotional category is defined separately for events, actions and objects. The intensity of the emotional categories resulting from an event is defined as the desirability and for actions and objects praiseworthiness and appealingness respectively (see Figure 1). One of the variables that is necessary to calculate desirability is the hierarchy of the character's goals. A certain goal, such as downloading a certain music album from the internet, would have several sub goals, such as download a specific song of that album. The completed goal of downloading of a whole album will evoke a higher desirability than the completed goal of downloading of a certain song, because it is positioned higher in the hierarchy. However, events might also happen outside of the character's current goal structure. The character needs to be able to evaluate such events as well. Besides the goal hierarchy, the emotion model also needs to keep a history of events, actions and objects. If the user, for example, gives the character one banana after the other in a short interval then the desirability of each of these events must decrease over time. The character needs to be less and less enthusiastic about each new banana. This history function is not described in the original OCC model, but plays an important role for the believability of the character. The history function has another important advantage. According to the OCC model, the likelihood of an event needs to be considered to calculate its desirability. The history function can help calculating this likelihood. Lets use the banana example again: The first time the character receives a banana, it will use its default likelihood to calculate the desirability of the event. When the character receives the next banana, it will look at the history and calculate how often it received a banana in the last moments. The more often it received a banana in the past the higher is the likelihood of this event and hence the lower is its desirability. After a certain period of not receiving any bananas the likelihood will fall back to its original default value. This value should not be decreased below its default value, because otherwise the character might experience an overdose of desirability the next time it receives a banana. Another benefit of the history function is the possibility to monitor the progress the character makes trying to achieve a certain goal. According to the OCC model, the effort and realization of an event needs to be considered to calculate its desirability. The history function can keep track of what the character has done and hence be the base for the calculation of effort and realization.

Mapping

If the emotion model has more categories than the character has abilities to express them, the emotional categories need to be mapped to the available expressions. If the character, for example, uses only facial expression then it may focus on the six basic emotions of happiness, sadness, anger, disgust, fear and surprise [15]. Interestingly, there is only one positive facial expression to which all 11 positive OCC categories need to be mapped to: the smile. Ekman [17] identified several different types of smiles but their mapping to the positive OCC categories remains unclear. The 11 negative OCC categories need to be mapped to four negative expressions: Anger, Sadness, Disgust and Fear. The facial expression of surprise cannot be linked to any OCC categories, since surprise is not considered to be an emotion in the OCC model. Even though the character might only be able to show six emotional expressions on its face, the user might very well be able to distinguish between the expression of love and pride with the help of context information. Each expression appears in a certain context that provides further information to the viewer. The user might interpret the smile of a mother next to her son receiving an academic degree as pride, but exactly the same smile towards her husband as love.

Reflection

The main limitation of the OCC model is its reliance on world model. Such models have only been successfully used in very limited worlds, such as pure virtual worlds in which only virtual characters operate. Furthermore, the OCC model will most likely only be one part of a larger system architecture that controls the character or robot. The emotional states of the OCC model must interact with the other states. Not only the face of the character is influenced by the emotional state of the character, but also its actions. It would be unbelievable if the character showed an angry expression on its face, but acted cooperatively. The mapping of the emotional state should be based on strong theoretical foundations. Such theoretical foundations might not be available for every action that a character might be able to execute and thus force the developer of the character to invent these mappings. This procedure has the intrinsic disadvantage that the developer might introduce an uncontrolled bias based on his or her own experiences and opinions.

Besides the actions of the character, the emotional state may also influence the attention and evaluation of events, actions and objects. In stress situations, for example, humans tend to focus their attention on the problem up to the point of "tunnel vision". [13] categorized the behavioral changes of the character through its emotional state in self-regulation (such as calming down), other-modulation (punish the other to feel better) and problem solving (try to avoid repetition). The latter will require the history function mentioned above. The emotional state of the character might even create new goals, such as calming down, which would result in actions like meditation.

Facial Expression Synthesis

There is a long tradition within the Human-Computer Interaction (HCI) community of investigating and building screen based characters that communicate with users [18]. Recently, robots have also been introduced to communicate with the users and this area has progressed sufficiently that some review articles are available [19, 20]. The main advantage that robots have over screen based agents is that they are able to directly manipulate the world. They not only converse with users, but also perform embodied physical actions.

Nevertheless, screen based characters and robots share an overlap in motivations for and problems with communicating with users. Bartneck et al. [21] has shown, for example, that there is no significant difference in the users' perception of emotions as expressed by a robot or a screen based character. The main motivation for using facial expressions to communicate with a user is that it is, in fact, impossible not to communicate. If the face of a character or robot remains inert, it communicates indifference. To put it another way, since humans are trained to recognize and interpret facial expressions it would be wasteful to ignore this rich communication channel.

Compared to the state of the art in screen-based characters, such as Embodied Conversational Agents [18], however, the field of robot's facial expressions is underdeveloped. Much attention has been paid to robot motor skills, such as locomotion and gesturing, but relatively little work has been done on their facial expression. Two main approaches can be observed in the field of robotics and screen based characters. In one camp are researchers and engineers who work on the generation of highly realistic faces. A recent example of a highly realistic robot is the Geminoid H1, which has 13 degrees of freedom (DOF) in its face alone. The annual Miss Digital award [22] may be thought of as a benchmark for the development of this kind of realistic computer generated face. While significant progress has been made in these areas, I have not yet reached human-like detail and realism, and this is acutely true for the animation of facial expressions. Hence, many highly realistic robots and character currently struggle with the phenomena of the "Uncanny Valley" [23], with users experiencing these artificial beings to be spooky or unnerving. Even the Repliee Q1Expo is only able to convince humans of the naturalness of its expressions for at best a few seconds [24]. In summary, natural robotic expressions remain in their infancy [20].

Major obstacles to the development of realistic robots lie with the actuators and the skin. At least 25 muscles are involved in the expression in the human face. These muscles are flexible, small and can be activated very quickly. Electric motors emit noise while pneumatic actuators are difficult to control. These problems often result in robotic heads that either have a small number of actuators or a somewhat larger-than-normal head. The Geminoid H1 robot, for example, is approximately five percent larger than its human counterpart. It also remains difficult to attach skin, which is often made of latex, to the head. This results in unnatural and non-human looking wrinkles and folds in the face.

At the other end of the spectrum, there are many researchers who are developing more iconic faces. Bartneck [25] showed that a robot with only two DOF in the face can produce a considerable repertoire of emotional expressions that make the interaction with the robot more enjoyable. Many popular robots, such as Asimo, Aibo and PaPeRo have only a schematic face with few or no actuators. Some of these only feature LEDs for creating facial expressions. The recently developed iCat robot is a good example of an iconic robot that has a simple physically-animated face. The eyebrows and lips of this robot move and this allows synthesis of a wide range of expressions.

Another important issue that needs to be considered when designing the facial expression of the character is that they need to be convincing and distinct at low intensity levels. Most events that a character encounters will not trigger an ecstatic state of happiness. The evaluation of a certain event should be roughly the same as could be expected of a human and most events that humans encounter in everyday life do unfortunately not result in ecstasy. If the character managed to download a complete album of music it still did not save the world from global warming. Hence, it should only show an appropriate level of happiness.

While there is progress in the facial expressions of robot faces, we are sill facing several conceptional problems that stem from the field of Artificial Intelligence. Lets take the example of emotions that I discussed in detailed above. The emotional state of the character is defined through values for each of its emotional categories. This emotional state needs to be expressed through all available channels. A conversational embodied character, for example, needs to express its emotional state through its speech and facial expressions. It would be unconvincing if the character would smile, but speak with a monotonous voice. However, the systematic manipulation of speech to express emotions remains a challenge for the research community. Emotional facial expressions are understood better, but a fundamental questions remains. Shall the character only express the most dominant emotional category, or shall it express every category at the same time and hence show a blend of emotions. The blending of emotional expression requires a sophisticated face, such as Baldi from the CSLU Toolkit. Cartoon like characters, such as eMuu [16] or Koda's Poker Playing Agent [3] are not able to show blends and therefore they can only express the most dominant emotional category.

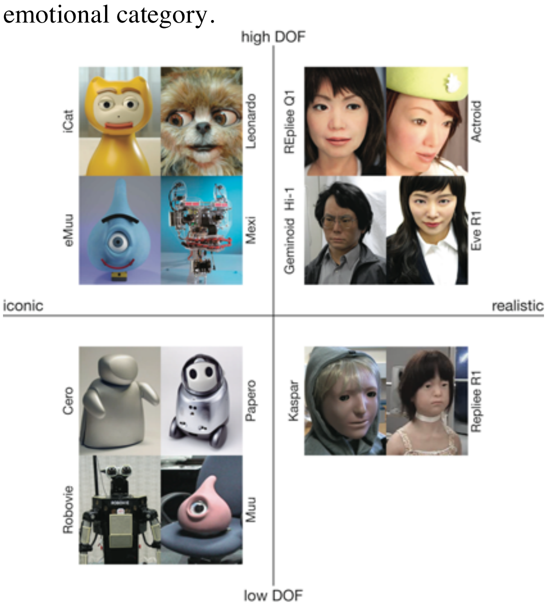

Fig. 2. Robots with animated faces

It becomes obvious that the problems inherited by human-robot interaction (HRI) researchers from the field of AI can be severe. Even if we neglect philosophical aspects of the AI problem and are satisfied with a computer that passes the Turing test, independently of how it achieves this, we will still encounter many practical problems. This leads us to the so-called "weak AI" position, namely claims of achieving human cognitive abilities are abandoned. Instead, this approach focuses on specific problem solving or reasoning tasks.

There has certainly been progress in weak AI, but this has not yet matured sufficiently to support artificial entities. Indeed, at present, developers of artificial entities must to resort to scripting behaviors. Clearly, the scripting approach has its limits and even the most advanced common sense database, Cyc, is largely incomplete. Emotion modeling should therefore not bet on the arrival of strong AI solutions, but focus on what weak AI solutions can offer today. Of course there is still hope that eventually also strong AI applications will become possible, but this may take a long time.

When we look at what types of HRI solutions are currently being built, we see that a large number of them do barely have any facial features at all. Qrio, Asimo and Hoap-2, for example, are only able to turn their heads with 2 degrees of freedom (DOF). Other robots, such as Aibo, are able to move their head, but have only LEDs to express their inner states in an abstract way. While these robots are intended to interact with humans, they certainly avoid facial expression synthesis. When we look at robots that have truly animated faces, we can distinguish between two dimensions: DOF and iconic/realistic appearance (see Figure 2).

Robots in the High DOF/Realistic quadrant not only have to fight with the uncannieness [26, 27] they also may raise user expectations of a strong AI which they are not able to fulfill. By contrast, the low DOF/Iconic quadrant includes robots that are extremely simple and perform well in their limited application domain. These robots lie well within the domain of the soluble. The most interesting quadrant is the High DOF/Iconic quadrant. These robots have rich facial expressions but avoid evoking associations with a strong AI through their iconic appearance. I propose that research on such robots has the greatest potential for significant advances in the use of emotions in HRI.

Conclusion

A problem that all these artificial entities have to deal with is, that while their expression processing has reached an almost sufficient maturity, their intelligence has not. This is especially problematic, since the mere presence of an animated face raises the expectation levels of its user. An entity that is able to express emotions is also expected to recognize and understand them. The same holds true for speech. If an artificial entity talks then we also expect it to listen and understand. As we all know, no artificial entity has yet passed the Turing test or claimed the Loebner Prize. All of the examples given in Table 1 presuppose the existence of a strong AI as described by John Searle [28].

The reasons why strong AI has not yet been achieved are manifold and the topic of lengthy discussion. Briefly then, there are, from the outset, conceptual problems. John Searle [28] pointed out that digital computers alone can never truly understand reality because it only manipulates syntactical symbols that do not contain semantics. The famous ‘Chinese room’ example points out some conceptual constraints in the development of strong AIs. According to his line of arguments, IBM’s chess playing computer “Deep Blue” does not actually understand chess. It may have beaten Kasparov, but it does so only by manipulating meaningless symbols. The creator of Deep Blue, Drew McDermott [29], replied to this criticism: "Saying Deep Blue doesn't really think about chess is like saying an airplane doesn't really fly because it doesn't flap its wings." This debate reflects different philosophical viewpoints on what it means to think and understand. For centuries philosophers have thought about such questions and perhaps the most important conclusion is that there is no conclusion at this point in time. Similarly, the possibility of developing a strong AI remains an open question. All the same, it must be admitted that some kind of progress has been made. In the past, a chess-playing machine would have been regarded as intelligent. But now it is regarded as the feat of a calculating machine – our criteria for what constitutes an intelligent machine has shifted.

In any case, suffice it to say that no sufficiently intelligent machine has yet emerged that would provide a foundation for many of the advanced application scenarios that have been imagined for emotional agents and robots. The point I hope to have made with the digression into AI is that the application dreams of researchers sometimes conceal rather unrealistic assumptions about what is possible to achieve with current technology. Emotion models heavily rely on the progress made in artificial intelligence and hence I would like to reply to Minsky's statement with a question: "Will emotional machines have intelligence?"

References

1. Minsky, M.L.: The society of mind. Simon and Schuster, New York, N.Y. (1986)
2. Elliott, C.D.: The Affective Reasoner: A Process model of emotions in a multi-agent system. The Institute for the Learning Sciences, Northwestern University, Evanston, Illinois (1992)
3. Koda, T.: Agents with Faces: A Study on the Effect of Personification of Software Agents. MIT Media Lab, Cambridge (1996)
4. O'Reilly, W.S.N.: Believable Social and Emotional Agents. Carnegie Mellon University, Pittsburgh, PA (1996)
5. Bartneck, C.: How convincing is Mr. Data's smile: Affective expressions of machines. User Modeling and User-Adapted Interaction **11** (2001) 279-295
6. Schulz, F.v.T.: Miteinander Reden - Stoerungen und Klaerungen. Rowolth Taschenbuch Verlag GmbH, Reinbeck bei Hamburg (1981)
7. Roseman, I.J., Antoniou, A.A., Jose, P.E.: Appraisal Determinants of Emotions: Constructing a More Accurate and Comprehensive Theory. Cognition and emotion **10** (1996) 241-277
8. Sloman, A.: Architectural requirements for human-like agents both natural and artificial. In: Dautenhahn, K. (ed.): Human Cognition And Social Agent Technology, Advances in Consciousness Research. John Benjamins Publishing Company, Amsterdam (1999)
9. Ortony, A., Clore, G., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, Cambridge (1988)
10. Bondarev, A.: Design of an Emotion Management System for a Home Robot. Industrial Design. Eindhoven University of Technology, Eindhoven (2002) 63
11. Studdard, P.: Representing Human Emotions in Intelligent Agents. The American University, Washington DC (1995)
12. Bratman, M.E., Israel, D.J., Pollack, M.E.: Plans and Resource-Bounded Practical Reasoning. Computational Intelligence **4** (1988) 349-355
13. Ortony, A.: On making believable emotional agents believable. In: Trapple, R.P., P. (ed.): Emotions in humans and artefacts. MIT Press, Cambridge (2003)
14. Wehrle, T.: Motivations behind modeling emotional agents: Whose emotion does your robot have? In: Numaoka, C., Cañamero, L.D., Petta, P. (eds.): Grounding Emotions in Adaptive Systems. 5th International Conference of the Society for Adaptive Behavior Workshop Notes (SAB'98), Zurich (1998)

15.Ekman, P., Friesen, W.V., Ellsworth, P.: Emotion in the human face : guidelines for research and an integration of findings. Pergamon Press, New York (1972)
16.Bartneck, C.: eMuu - an embodied emotional character for the ambient intelligent home. Eindhoven University of Technology, Eindhoven (2002)
17.Ekman, P.: Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage. W.W. Norton, New York (1985)
18.Cassell, J., Sullivan, J., Prevost, S., Churchill, E.: Embodied Conversational Agents. MIT Press, Cambridge (2000)
19.Bartneck, C., Okada, M.: Robotic User Interfaces. Human and Computer Conference (HC2001), Aizu (2001) 130-140
20.Fong, T., Nourbakhsh, I., Dautenhahn, K.: A survey of socially interactive robots. Robotics and Autonomous Systems **42** (2003) 143-166
21.Bartneck, C., Reichenbach, J., Breemen, A.: In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. In: Kurtgözü, A. (ed.): Design and Emotion 2004, Ankara (2004)
22.Cerami, F.: Miss Digital World. (2006)
23.Mori, M.: The Uncanny Valley. Energy **7** (1970) 33-35
24.Ishiguro, H.: Android Science - Towards a new cross-interdisciplinary framework. CogSci Workshop Towards social Mechanisms of android science. Cognitive Science Society, Stresa (2005) 1-6
25.Bartneck, C.: Interacting with an Embodied Emotional Character. In: Forzlizzi, J., Hamington, B., Jordan, P.W. (eds.): Design for Pleasurable Products Conference (DPPI2004). ACM Press, Pittsburgh (2003) 55-60
26.MacDorman, K.F.: Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, Vancouver (2006)
27.Bartneck, C., Kanda, T., Ishiguro, H., Hagita, N.: Is the Uncanny Valley an Uncanny Cliff? : 16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2007. IEEE, Jeju, Korea (2007) 368-373
28.Searle, J.R.: Minds, brains and programs. Behavioral and Brain Sciences **3** (1980) 417-457
29.McDermott, D.: Yes, Computers Can Think. New York Times (1997)

Periodic Activations of Behaviors and Motivational States

Ernesto Burattini and Silvia Rossi

Dipartimento di Scienze Fisiche
Università degli Studi di Napoli "Federico II" – Napoli, Italy
ernb@na.infn.it, silrossi@unina.it

Abstract. The possible modulatory influence of motivations and emotions is fundamental in designing robotic adaptive systems. In this paper, we will try to connect the concept of periodic behavior activations to emotions, in order to link the variability of behaviors to the circumstances in which they are activated. We will study the impact of emotion, described as timed controlled structures, on simple reactive behaviors. We will show, through this approach, that the emergent behaviors of a simple robot designed with a parallel or hierarchical architecture are comparable. Finally, we will see that conflicts in behaviors may be solved without an explicit action selection mechanism.

1 Introduction

In Robotics one of the main issues in designing a control system is to enable an autonomous robot to react and adapt in useful time to environmental changes [1]. This reaction depends on the correct identification of objects and their properties by appropriate sensor devices, with a strong emphasis on the concept of the stimuli-response loop. Moreover, the robotic community, started to pay attention not only to the robot-environment interactions, but also, so to speak, to the interactions that may arise within the robots itself [2] and how these latter (for example its emotional states) may influence the emergent behavior of the robot.

In these last years some researchers [2–7] started to pay attention to the role of emotional and motivational states in order to achieve an adaptive emergent behavior of robotics systems. In particular, the role of emotions has been introduced for behavior modulations [3, 4], to provide adaptivity to environmental changes. Moreover, cognitive psychology considers thinking, learning and memory activities as a problem of information processing. However, the description of motivational issues and emotional states as a processing problem is not an obvious task [3]. The interest for such "internal mechanisms" comes within the robotic community taking inspiration from ethological, biological and neuroscience studies. In our opinion, in order to model different and new architectures for controlling the robot behavior, both these aspects (the interaction with the surrounding world and the internal states) have to be considered, since they influence each other. For example, the simple perception-action response to an external stimulus may produce different patterns of actions consequently to a different internal state of the robot. This internal state may change according to his emotional or motivational state or following its past perceptions and it will tune and adapt both the executions of different behaviors for the robot and the processing frequency of the sensors' inputs. Our working hypothesis is that such adaptive behaviors can be achieved in the control activity of a robot starting from self-regulated periodic mechanisms.

In previous papers [8, 9] we highlighted the opportunity of managing the frequency of processing the sensors inputs in an efficient way, because it may have negative effects on the robot behavior. This kind of problems leads us to find a solution for the efficient use of the Robotic sensor apparatus. Therefore, we moved to study how rhythmic computations may be introduced in a control mechanism for robotics systems and how such introduction may lead to a framework that will cope with some of the common problems in designing control systems for robots. In this paper we will analyze our architecture in terms of emergent behavior driven by motivational and emotional states, and we will describe how our architecture may deal with conflicting behaviors (for example, predator avoidance and food acquisition) starting from the concepts of periodical adaptive activations of behaviors. We present a robotic architecture that has the capability of adapting its behavior to the rate of change of a dynamic environment - e.g. of tuning the velocity of reaction to the external stimuli coherently to the changes occurring in the environment. On the other hand, we want our model to take into account that such stimuli may come not only from the external environment (as a bottom-up process), but they can be generated by the robot itself (top-down) [2] - e.g. the robot has to adapt its perceptual system according to its "needs".

2 AIRMs

A motivation-based architecture should be able to integrate a combination of internal and external factors to select the appropriate behavior. However, these architectures are not always sufficiently adaptive to unpredictable environmental changes [4]. What we want to achieve is the ability, for a robotic system, of adapting its emergent behavior to the surrounding environment and to its internal state. At the same time we want the robot to opportunely react according to environmental changes and to efficiently spend the resources necessary to monitor the surrounding environment. To achieve this goal we started from the consideration that a wide type of behaviors are generated by the so called central pattern generators [10], i.e., central oscillators whose output modulates rhythmic movements. The role of such oscillator in coordination of motor patterns [11], such as breathing and walking, is well accepted in neuroscience [12].

So, we would like to have a control system for the percept inputs that performs a quasi-periodic activity (i.e. it has at least an active and inactive phase) and should be flexible (i.e. dynamically adapt its period and amplitude to external and internal constraints). In particular we would like to associate a periodic control system to the activation of each single behavior. Lorentz [13] and Tinbergen [14] identified in many animals an innate releasing or inhibiting mechanism (IRM) able to control and coordinate behaviors. An IRM presupposes a specific stimulus that releases a pattern of actions. For example, a prey animal may have, as an IRM, the stimulus coming from the view of the predator, which activates the escape behavior. IRMs were included in the schema representation of behaviors [15] in the form of releasers, controlling when behaviors must be activated or deactivated. A releaser may be an activation mechanism that depends both by exogenous factors, that trigger an emotional state (a particular environmental condition – for example a prey that detects the presence of the predator), and by endogenous factors (a motivational state – for example hunger).

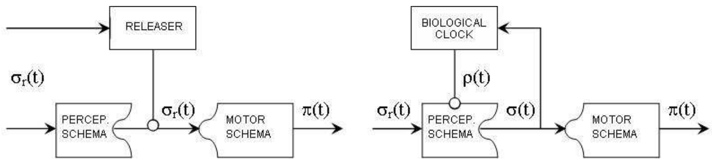

Fig. 1. A schema representation [15] of releasers and biological clocks. The function $\sigma_r(t)$ represents the input coming from sensors at each time interval; $\pi(t)$ is the command sent to actuators; $\sigma(t)$ represents the inputs elaborated by the perceptual schema and sampled by the function $\rho(t)$.

The releaser's function, somehow, recalls the notion of "internal clock", already introduced in some approaches [7, 16] in order to activate motivational states for a robot (for example, hunger or sleep). In fact, an internal clock, as a releaser, is a mechanism which regulates the behavior of living organisms. Moreover, this may depend on endogenous factors (i.e., independent from external environment) or exogenous factors. Starting from this analogy, we try, in some sense, to abstract the concept of internal clocks and to connect it to periodic activations of behaviors in a robotic architecture, in a similar way as a mechanism for releasing is related to a behavior using any representative models well studied and used in the Robotic community [15]. There are, however, substantial differences between the two concepts. An internal clock is responsible for the activation of a particular behavior, but has something more than a releaser (see Fig.1). First of all, the releaser acts as a control signal for the whole behavior and it, somehow, may involve an elaboration of the input (for example a releaser may be the presence of a predator). An AIRM (Adaptive Innate Releasing Mechanism), instead, works only on the perceptual schema and has an active (or inactive) state that depends also on endogenous factors (the perceptual schema elaborates the input when the AIRM is active). Furthermore, internal clock may imply a regular and periodic activation of the perceptual schema of a behavior, whose activations in time may be predicted – and so, also the amount of resources spent for the elaboration of inputs. Instead, the activity of a releaser depends only on contingent factors. In [8, 9] we connected the concept of IRM to the concept of a periodical activation of behaviors (AIRM). In this way, no computational resources are spent to elaborate not needed stimuli, because the corresponding control systems are kept "inactive" until a new periodical activation takes place, and at the same time we are able to control the amount of resources spent in the elaboration of the sensors' inputs. Moreover, the introduction of internal clocks, within a robotic architecture, has also the effect of controlling behaviors that may require a fixed pattern of activation in time. This activation of behavior may be interpreted as large time scale activities, for example the activation of macro-behavior like feeding or sleeping, or as short time scale activities, in the sense of central-pattern generators in controlling rhythmic movements of a robot as walking, but also as a general mechanism for controlling activation of simple behaviors. Finally, we foresee that the introduction of such asynchronism in the robot control system may lead to an emergent behavior that is able to change and adapt according to its context without having an explicit action selection mechanism.

We assume the hypothesis of an architecture with some periodic releasing mechanisms of activation of behaviors. Such mechanisms, according to the environment, speed up or gradually slow down the period of behaviors activation and thereby the reading frequency of the sensors. In this system, however, the feedback does not come only from the outside, but can also be generated by the robot itself [2], allowing the robot to adapt itself also according to its emotional or motivational state. An emotional state, in our work, has to be interpreted, following the Damasio definition [17], as an unconscious and automatic response in reaction to a stimulus that involves an adjustment in homeostatic balance as well as the enhancement of specific behaviors. Moreover, in neuroscience, while classical theories of sensory processing view the brain as a passive, stimulus-driven device, more recent approaches [18] view the perception as an active and highly selective process.

Our working hypothesis is that each behavior of a Robotic System (RS) may be provided with clocks that control the periodic activation of behaviors. We may think that each of the releasers, that manage the various micro/macro-behaviors, is activated by an individual clock with a variable period p_β, as it will be explained in the following, depending on the purpose of the behavior and on the sensors data involved in the behavior. Timed releasing functions take data from a perceptual schema and from the internal state of the robot and return enabling/disabling signals to the perceptual schema itself. That is, the perceptual schema of a behavior is regulated by an internal clock that says how frequently the inputs have to be elaborated. For example, if the initial value of the period of a clock is of four time units, it means that the input from the sensors for this particular behavior will be processed only every four time units. In the other cases, during the inactivity state of the perceptual schema, no new commands will be send to the motor schema and so no new actions will be produced.

In our architecture, motivational behaviors have an impact on the value of the period of each behavior and so they can regulate and modify the perceptual abilities of the robot as well as its actions in time. Motivational and emotional behaviors may be induced both by internal states of the robot (for example, hunger), coded as linear time-dependent functions, and as an emergent process from the interaction with the environment (for example, fear).

3 Implementation and Testing

Let us assume that each behavior of the robotic system has a variable period initially equal to a preferred value. In our experiments, we assumed that these periods are proportional to powers of two. We designed a system whose behavior is mainly guided by the visual information in a 3D environment. In particular, according to [5], the reaction of the robot may be driven by moving objects. In order to achieve the proper reaction of the robot in respect to a moving object, we implemented a control schema to change the period of the clock based on the Weber law of perception. We already discussed in section 2 the perceptual schema modulation according to a periodic releasing function. In particular we noticed that the robot can evaluate the perceptual inputs only when the releaser/clock is on. While the reaction of the robot depends on the perceptual inputs (for example, the robot that sees a predator will produce an action to escape) the self-regulation mechanism, encoded in the internal clock, will confront the current percept with the last available percept, stored in the temporal or working memory of the robot. The change in its emotional state (encoded as a change in the releasing period) depends on how much this value has fluctuated. However, in order to set the appropriate thresholds for evaluating this change, we cannot refer to absolute values. In this sense the Weber law allows us to compute the relative change in the percept input ($\sigma(t)$) as $\frac{\Delta\sigma}{\sigma}$. In figure 2 the percept input of an experiment is plotted. The dotted lines represent the thresholds we use to adapt the period of the releasing function according to the values of the input percept. In fact, in order to make the robot able to react in time, for an increasing percept we want the period to decrease according to the input changing rate. Let us remark that, in our approach, the period of a behavior may change its value, varying among power of two values. Moreover, also the selected thresholds for changing the period are proportional to a power law. For example, if the percept exceeds the first threshold, the period will be halved; instead, if the percept exceeds the second threshold, the current value of the period will be reduced to a quarter, and so on. On the contrary, when we have a decreasing function, we want the process for coming back to the maximum value of the period to be slow.

In order to test our working hypotheses we used a PIONEER 3DX provided with a blob camera (see Fig.3). The robot architecture is constituted by four simple behaviors: `WANDER`, `FIND_FOOD`, `EAT` and `ESCAPE`. In particular we were interested in observing only the `FIND_FOOD` and `ESCAPE`, whose perceptual schemas were controlled by AIRMs and whose behaviors may be in conflict requiring an action selection mechanism. We implemented both a subsumption architecture (see figure 4(a)) with `ESCAPE` subsuming `FIND_FOOD` and a parallel architecture (see figure 4(b)) whose output was the sum of the outputs of the two behav-

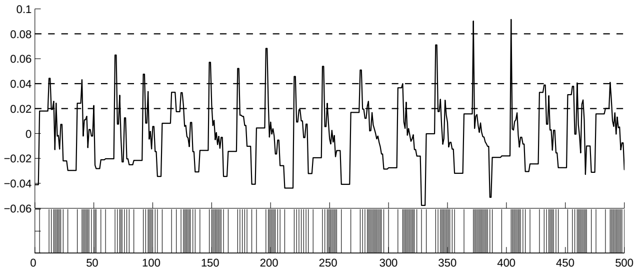

Fig. 2. Changing of the percept input plotted as $\frac{\Delta\sigma}{\sigma}$. The dotted lines represent thresholds for the adaptation of the period process.

iors. The output (π_i) of these two behaviors consists in a predefined velocity and direction, except for the ESCAPE behavior whose output velocity depends on the internal clock. In particular, if the value of the clock is equal to the initial maximum value, the module of the velocity will be equal to the velocity set by FIND_FOOD. If this value is equal to the minimum value, the velocity will be set to a much higher value in order to escape. In all the other cases, velocity will be a constant value in between the maximum and the minimum values.

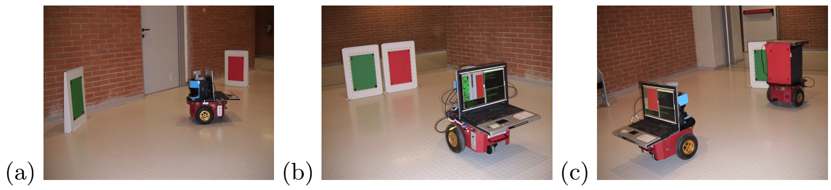

Fig. 3. Snapshots form the case study: (a) the robot wanders looking for food; (b) food and predator are in the same direction; (c) the predator moves toward the prey.

Let us suppose that a red object represents a predator and a green object represents food (see Fig.3(a)). What if the system is in the case of having in the same direction both the food and the predator (see Fig.3(b))? In this situation the emergent behavior will depend on motivational states and will be influenced by their impact on the activations of behaviors. The FIND_FOOD behavior has an internal clock with a period whose value depends on the motivational state of “hunger”. This state is regulated by a linear time-dependent function, and this means that at the beginning, when the value of the hunger is low, the FIND_FOOD behavior is released with a predefined period that depends on the life cycle of the robot. During the simulation, the hunger value will grow in time and, accordingly, also the period of the clock of the corresponding behavior will be reduced. When the behavior is enabled and the robot senses a green object, the output of the FIND_FOOD behavior will set the direction of movements towards the food. The ESCAPE has an internal clock that simulates the state of “fear”. At the beginning of the simulation the value of the period is set in order to safely check the presence of a predator. If the robot senses a red object (the predator) and the behavior is enabled, the output of the behavior will be a movement in the opposite direction of the predator. The period of this clock does not depend on an internal variable (like in the case of `FIND_FOOD`), but on the changing of the value of the percept itself according to the Weber law. This means that the "fear" of the robot will increase if the predator is moving toward the prey (i.e., the period will be reduced). Moreover, let us highlight that this process will have, as a consequence, an adaptation of the behavior of the robot if the predator is not moving. However, in the case of both the food and the predator in face of the robot, while approaching the food the movement of the robot itself may induce a change in the perception of the dimension of the red blob.

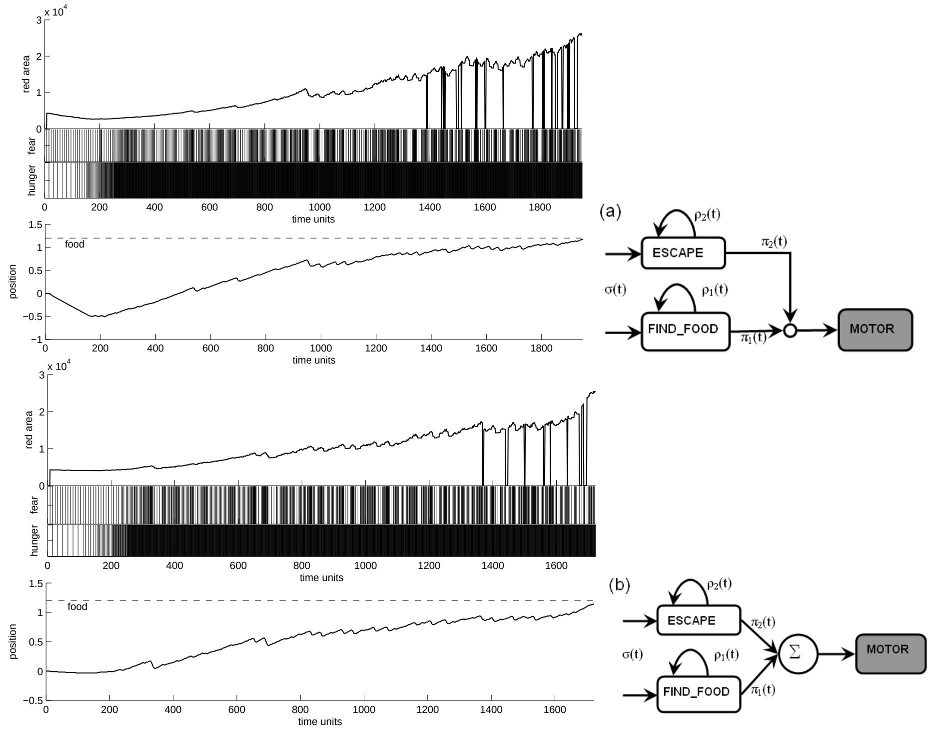

Fig. 4. Subsumption (a) and parallel (b) implementations of behaviors. For each case the plots show the changing of the input (red area), of the two clocks (fear and hunger) and of the position of the robot towards the food at each time unit.

In figure 4 we plotted some results for the case study described above. The first plot refers to the subsumption implementation of the behaviors, while the second one refers to a parallel architecture. The first plot of each of the two cases represents the percept (i.e., red blob area for the `ESCAPE` behavior). Such percept is sampled according to the corresponding internal clock that simulates fear. Let us notice that while the internal clock is inactive the robot does not update its perceptual input, which remains constant until the next activation of the clock. Moreover, let us notice how the frequency of activations of the clock is modified following the input percept. The last part of each plot represents the internal clock of the `FIND_FOOD` behavior, that depends on time. As soon as this value increases more than the value of the `ESCAPE` behavior, the robot will start to move toward the food with an oscillating behavior that will lead to reach the position of the food. Let us notice that the emergent behavior of both the two approaches, represented by the changing of the position of the robot towards the food, is comparable, in the sense that both the approaches, if the behaviors are controlled by internal clocks, will lead to the same oscillating pattern towards the food. The only substantial difference between the two approaches happens when the hunger is low: in fact, while in the subsumption architecture the robot will move in the opposite direction of the food (and the predator), in the parallel architecture the robot is not moving.

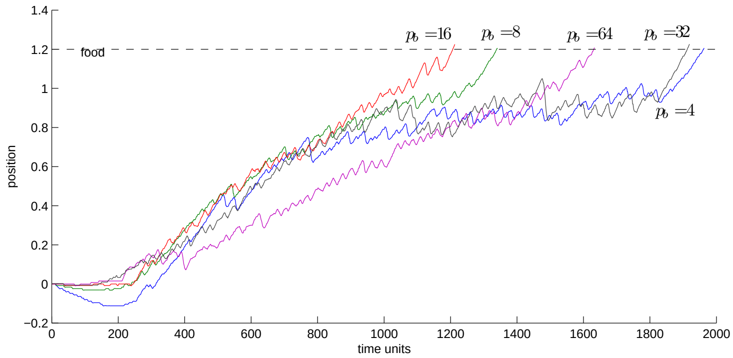

Fig. 5. The robot emergent behavior with different initial values of the `ESCAPE` clock.

In figure 5 we compared the emergent behavior of the same robot with a parallel architecture, changing the initial maximum value for the period of the `ESCAPE` behavior. First of all, let us highlight that the emergent behaviors of the robot seem not to depend on this initial value. The explanation of this situation, in this particular case study, is that while approaching the food the clock of the `ESCAPE` behavior frequently changes its value, also for the presence of the predator. This oscillation pattern makes the robot not able to return the initial value of the `ESCAPE` period that keeps oscillating between the minimum value and a constant average value. However, while in this case it seems that the maximum value of the `ESCAPE` period does not have any impact on the emergent behavior, we want our robot to react in useful time to moving obstacles (i.e., the predator). In Fig.6 we plotted the changing of the red area and, accordingly, the changing of the period of the clock of the `ESCAPE` behavior and the changing in the velocity of the robot in the case that the red object starts to move (see

Fig.3(c)). Let us notice that when the period reaches its minimum value the module of the velocity is bigger in order to escape.

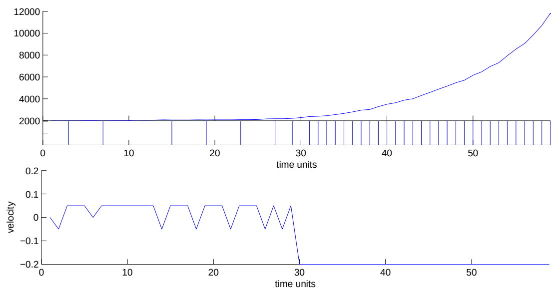

Fig. 6. Robot reaction to a moving obstacle.

4 Discussion

In this paper, we started to explore the feasibility of designing robotic architectures based on motivational modulation of behavior activities by means of periodic releasers. The embedding of such controlled rhythms within a RS behavior allows the realization of flexible/adaptive behavior which can realize timed activation of the behavior itself as well as modulation of its performance according to its internal state and sensorial information. Other authors dealt with this kind of problems. For example, in [7] the authors presented a parallel architecture focused on the concept of activity level of each schema which determines the priority of its thread of execution. A more active perceptual schema can process the visual input more quickly and a more active motor schema can send more commands to the motor controller. However, while in our approach such effects are obtained through rhythmic activation of behaviors, in [7] the variables are elaborated through a fuzzy based command fusion mechanism.

Moreover, behavior based robotic usually resolves conflicts by using a subsumption architecture or by implementing some control mechanisms in order to switch between tasks, selecting the action. For example, in [5] the authors presented a schema theoretic model for a praying mantis which behaviors are driven by motivational variables such as fear, hunger and sex-drive. In this approach, the action selection module takes into account only the motivational variable with the high value. As in our case, when the hunger is too high the robot will move toward the food even through there is a predator in sight. Moreover, fear depends of the view of the predator, but when the predator is in the field of view of the prey this variable is only set to a predefined high value.

Let notice that while a motivational behavior may have a linear model of development, emotions are not a linear succession of events [6]. In our approach we presented a model of an emotional behavior that does not depend only on linear time dependent functions, but it is directly connected to the changing rate of the surrounding environment. However, while for a simple case study our architecture was able, by means of asynchronous computation, to act like an action selection mechanism and to adapt to its context, one of the problems of more complex parallels architectures comes from the possibility of arising interferences between different processes. Since emotions and motivation are not independent processes, as future work we will move forward in the direction of studying how these adaptive periodical activations of behaviors may influence and constrain each other.

References

1. Wasson, G., Kortenkamp, D., Huber, E.: Integrating active perception with an autonomous robot architecture. In: Proceedings of the second international conference on Autonomous agents, New York, NY, USA, ACM Press (1998) 325–331
2. Parisi, D.: Internal robotics. Connect. Sci. **16**(4) (2004) 325–338
3. Doerner, D., Hille, K.: Artificial souls: Motivated emotional robots. In: International Conference on Systems, Man and Cybernetics. (1995) 3828–3832
4. Cañamero, L., Avila-García, O.: A bottom-up investigation of emotional modulation in competitive scenarios. In: Affective Computing and Intelligent Interaction, Second International Conference. Volume 4738 of LNCS., Springer (2007) 398–409
5. Arkin, R.C., Ali, K., Weitzenfeld, A., Cervantes-Pérez, F.: Behavioral models of the praying mantis as a basis for robotic behavior. Robotics and Autonomous Systems **32**(1) (2000) 39–60
6. Pérez, C.H., Moffat, D.C., Ziemke, T.: Emotions as a bridge to the environment: On the role of body in organisms and robots. In: 9th International Conference on Simulation of Adaptive Behavior. Volume 4095 of LNCS., Springer (2006) 3–16
7. Pezzulo, G., Calvi, G.: A schema based model of the praying mantis. In: SAB – 9th International Conference on Simulation of Adaptive Behavior. Volume 4095 of Lecture Notes in Computer Science., Springer (2006) 211–223
8. Burattini, E., Rossi, S.: A robotic architecture with innate releasing mechanism. In: 2nd International Symposium on Brain, Vision and Artificial Intelligence. Volume 4729 of Lecture Notes in Computer Science., Springer (2007) 576–585
9. Burattini, E., Rossi, S.: Periodic adaptive activation of behaviors in robotic systems. Inter. Journal of Pattern Recognition and Artificial Intelligence (in press)
10. Marder, E., Calabrese, R.L.: Principles of rhythmic motor pattern generation. Physiol. Rev. **76**(3) (1996) 687–717
11. Cruse, H.: The functional sense of central oscillations in walking. Biological Cybernetics **86**(4) (2002) 271–280
12. Buzsaki, G.: Rhythms of the Brain. Oxford University Press, USA (2006)
13. Lorenz, K.: King solomon's ring. Penguin (1991)
14. Tinbergen, N.: The study of instinct. Oxford University press (1951)
15. Arbib, M.A.: Schema theory. In: The handbook of brain theory and neural networks. MIT Press, Cambridge, MA, USA (1998) 830–834
16. Stoytchev, A., Arkin, R.C.: Incorporating motivation in a hybrid robot architecture. JACIII **8**(3) (2004) 269–274
17. Damasio, A.R.: The Feeling of What Happens: body and emotion in the making of consciousness. Harcourt (1999)
18. Engel, A.K., Fries, P., Singer, W.: Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci **2**(10) (2001) 704–716

Computational Principles Underlying the Functioning of Amygdala in the Affective Regulation of Behaviour

Francesco Mannella, Marco Mirolli and Gianluca Baldassarre

Istituto di Scienze e Tecnologie della Cognizione,
Consiglio Nazionale delle Ricerche (ISTC-CNR)
Via San Martino della Battaglia 44, I-00185 Roma, Italy
{francesco.mannella, marco.mirolli, gianluca.baldassare}@istc.cnr.it*

Abstract. This paper presents a short review, compiled with a computational perspective, of the empirical neuroscientific evidence related to amygdala, a brain complex situated at the core of various brain systems underlying motivations and emotions. The functions of amygdala are fundamental for organisms' adaptive behaviour as they allow them to assign subjective saliency and value to experienced world states, so enhancing the adaptive power of their cognitive processes. In this respect, the major goal of the review is outlining the main computational functionalities of amygdala emerging from the neuroscientific investigations on affective processes so as to contribute to highlight the general architectural and functioning mechanisms underlying organisms' emotional processes. This effort is also expected to fertilise the design of robot controllers exhibiting a flexibility and autonomy comparable to that of real organisms.

1 Introduction: Exploiting the Synergies Between the Neuroscientific Research on Amygdala and Embodied Artificial Intelligence

In decades of research, neuroscience has produced a large amount of data and insights relative to the neural substrates underlying emotions. These are now seen as a fundamental product of evolution that allows organisms to suitably regulate and flexibly modify behaviours on the basis of their survival and reproduction needs. Emotions play a central role in the behavioural flexibility exhibited by real organisms, and for this reason their study is important not only for the advancement of their overall scientific understanding but also for

* This research was supported by the EU Project *ICEA - Integrating Cognition, Emotion and Autonomy*, contract no. FP6-IST-IP-027819.

the design of autonomous robots capable of tackling unpredictable and non-stationary environments with a versatility similar to that of organisms. These principles have been investigated in depth in some models developed within the embodied artificial-intelligence community, for example see [1–3]. These works have the merit of outlining the general principles underlying emotions and of giving a general account of them in terms of embodiment and dynamic coupling with the environment (see [4] for a review). However, they usually present models that are only weakly related to the aforementioned empirical data. This implies missing important synergies in the study of emotions that might stem from the integration of the two approaches.

This paper introduces the first results of the theoretical and reviewing efforts of a research agenda directed to contribute to build those synergies and to lead the two research threads to have a stronger integration. In particular, the paper introduces relevant empirical evidence related to *amygdala* (Amg), probably the most important brain system integrating processes involving *external stimuli*, internal *cognitive processes*, and *internal states* related to organism's needs and homeostatic regulations. In doing so, the focus will be on the neuroscientific research showing the core functionality implemented by Amg. In this respect, we anticipate that the general function of Amg is to associate "unlearned behaviours", internal body and brain states, and internal body and brain modulations, to neutral stimuli coming from the external world so that they can acquire a biological salience and play a role in the regulation of various behaviours and cognitive processes. (note that, in the following, the expression "unlearned behaviours" will be used to refer to behaviours that might be either innate or developed during the first phases of life under strong genetic pressures and general environmental constraints, cf. [5]).

As mentioned above, the review of Amg's properties will be done with a computational perspective in mind (adaptive functions, neural mechanisms, etc.) and with the aim of isolating the fundamental principles underlying the functioning of the main brain systems involved in the regulation of emotions, motivations and learning. This effort is expected to produce insights that should be useful as a general framework for designing and implementing detailed computational embodied models, as it already happened in three of our previous works [6–8].

2 The Amygdala Anatomy and Core Functions

The Amg is an almond-shaped group of nuclei located within each medial temporal lobe of the brain. Figure 1 illustrates the broad anatomical organisation of Amg. In particular the figure shows that Amg is formed by three major sets of nuclei each playing a major distinct functional role: lateral amygdala (LA), basolateral amygdala (BLA) and central nucleus of amygdala (CeA). The graph also shows the main connections of these nuclei with other brain areas with which the Amg's nuclei form various brain sub-systems implementing several functions related to affective regulation of behaviour (discussed in Section 3).

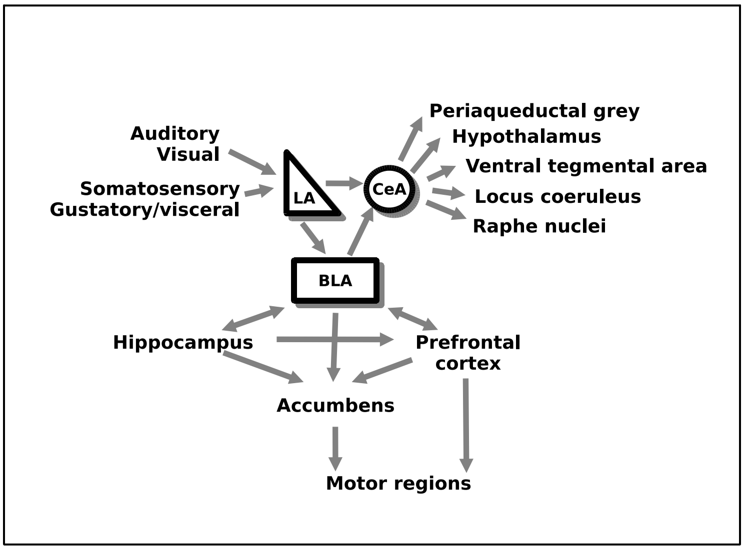

Fig. 1. Major connections between the main nuclei of Amg and between these nuclei and other brain districts with which it forms important brain sub-systems underlying various affective regulations of behaviour. LA receives afferent connections from many cortical and subcortical areas and projects mainly to other nuclei of Amg. CeA receives afferent connections from other nuclei of Amg and projects efferent connections to many subcortical systems. BLA has complex reciprocal connections with prefrontal cortex, hippocampus and nucleus accumbens.

The role that Amg plays in such affective regulations relies upon three functions (Figure 2). The first function is based on unlearned associations existing between a number of *biologically-salient stimuli* with the direct triggering of various appetitive and aversive unlearned responses directed to the environment, the body and the brain itself. In particular, some kinds of tastes and olfactory stimuli, as well as nociceptive stimuli [9], can, via unlearned Amg's pathways, *directly* contribute to trigger unlearned behaviours (e.g., salivation, freezing, startling, and approaching), to regulate emotional body states (e.g. heart rate and blood pressure), to broadly activate whole brain areas and regulate learning processes (e.g., via the neuromodulation processes performed by the nuclei of the reticular formation).

The second Amg's function is based on the strengthening of the neural pathways which allow *neutral stimuli* from the environment to trigger the aforementioned unlearned reactions. Amg can implement this process on the basis of two associative mechanisms.

The first associative mechanism is based on the creation of *direct* neural associations between the representations of *neutral* stimuli and the aforementioned unlearned reactions (these are S-R types of associations). S-R learning occurs at the level of the LA-CeA pathway, via connections that depart from LA units

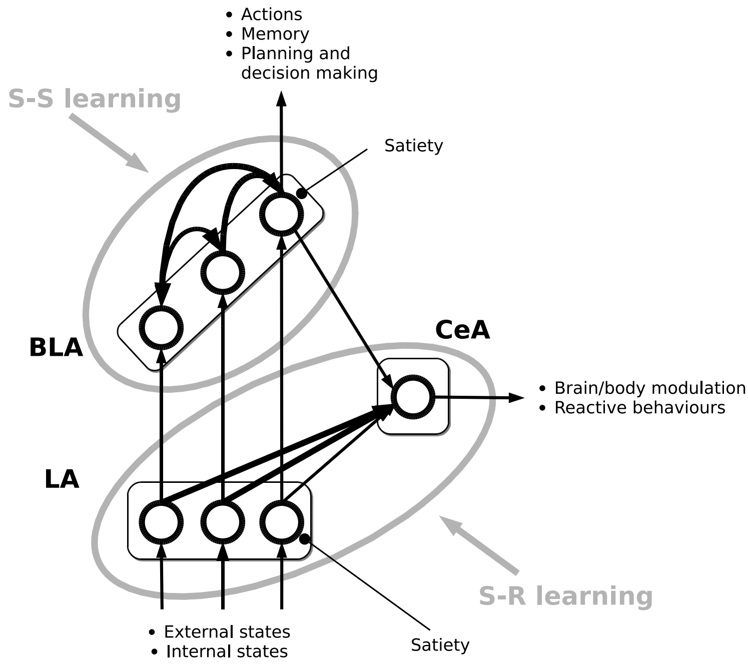

Fig. 2. Major learning processes involving the three main nuclei of Amg. Circles indicate clusters of neurons representing stimuli or reactions received from or directed to the environment, the body or the nervous system itself (for simplicity, the graph represents only few units). Bold connections represent associations formed during learning, whereas plain connections represent unlearned associations. S-R learning is implemented by the LA-CeA pathway: this allows external stimuli activating LA to directly trigger the unlearned reactions of CeA. S-S learning is implemented within BLA. Only few BLA units are associated with the CeA units: other BLA units representing external stimuli can trigger CeA reactions only by forming lateral associations with those units. Importantly, internal states, such as satiety, can modulate *on the fly* the triggering of Amg's responses by acting on the representations of the unconditioned stimuli, e.g. by inhibiting them (connections with a dot head).

representing stimuli from the world and converge to CeA which triggers the unlearned reactions. With learning, each LA unit can become directly associated with CeA reaction units.

The second associative mechanism is based on the formation of neural associations between internal representations of neutral stimuli and the internal representations of the aforementioned salient stimuli (these are S-S types of associations): the activation of these representations can then trigger the unlearned responses. S-S learning occurs within BLA. Only few BLA units, representing biologically salient stimuli, are associated with the CeA units. Other BLA units, representing stimuli from the environment, can have access to CeA reactions only by forming lateral associations with the BLA units representing salient stimuli.

A last important function of Amg relies on its capacity to modulate the effects of the associations that it forms in the ways just described *on the fly* (i.e., without the need of re-learning) on the basis of current homeostatic body states and overall brain states. For example the Amg is capable of avoiding to trigger approaching behaviours towards a source of food if this has been temporarily or permanently devalued through satiation or poisoning.

3 The Functions that Amygdala Plays in Different Brain Sub-Systems

Amygdala has been associated with a wide range of cognitive functions, including emotional regulation, learning, action selection, memory, attention and perception. In particular, a large amount of studies have now firmly established the involvement of Amg in aversive behaviours such as those involved in fear conditioning and taste aversion experiments [10–12]. Recently, an increasing amount of behavioral evidence has started to reveal an Amg's involvement also in appetitive behaviours [13–16]. This is also being corroborated by anatomical investigations that indicate the existence of Amg's afferent neural pathways which carry information related to both aversive and appetitive events [17, 18].

The Amg plays a function in these aversive and appetitive behaviours as it is an important component of several brain sub-systems involving the hypothalamus, insular cortex, brain stem (in particular the reticular formation), hippocampus, basal ganglia, and prefrontal cortex. In general, the role that the Amg plays in all these sub-systems relies on its capacity to use input information related to internal body states to assign positive and negative emotional valence to stimuli from the environment on the basis of the associative mechanism described in Section 2.

The Amg exploits these associative processes to play several important affective-regulation functions within various brain sub-systems (Figure 3). The detailed investigation and modeling of these functions, only broadly described here, form the main research objectives of the research agenda mentioned in Section 1.

Three of these functions involve the affective regulation of behaviours directed to the external environment:

1. *Selection and triggering of unlearned behaviours.*
   The Amg plays an important role in triggering unlearned behaviours (e.g., the "unconditioned responses", or "URs", used in conditioning experiments). on the basis of biologically salient stimuli (e.g., the "unconditioned stimuli", or "USs", used in conditioning experiments). In particular, studies about both appetitive and aversive Pavlovian conditioning focused on the triggering of unlearned behaviours such as visceral responses [19], freezing [10, 20, 21], startle [22], and orienting responses [23, 24]. Behavioural and anatomical evidence indicates that these kinds of reactions are triggered by CeA activations [23, 24, 19].
   Also approach and avoidance behaviours, the conditioning of which depends on the BLA functioning [24, 19], can be included in the category of URs that

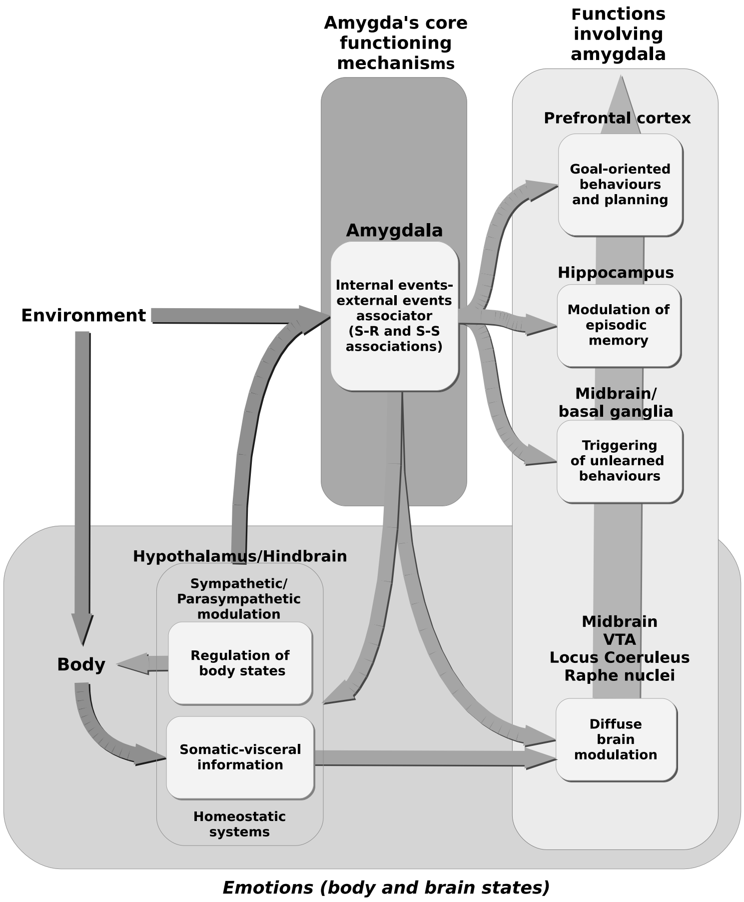

Fig. 3. A scheme indicating the main functions played by the Amg within some of the main affective regulatory systems of brain. Notice the core associative mechanisms implemented by the Amg, which subserve all such functions, and the role that Amg plays in the modulation of emotions in terms of the regulation of diffused brain states and body homeostatic states.

animals produce in the presence of USs via unlearned neural connections existing between their neural representations.

2. *Furnishing emotional states for the generation of fast-forming episodic memories.*

   The BLA's massive reciprocal connections with hippocampus might allow Amg to influence multi-modal fast-associative episodic memory processes taking places in it on the basis of current emotional states. In particular, as Amg is one of the main brain loci where the information on internal states and on the value of external stimuli is integrated, its input to the hippocampus might furnish the emotional context to memory formation and consolidation processes that it supports [25, 26].

3. *Emotional evaluation of stimuli for goal-oriented behaviour.*
   The BLA-prefrontal cortex reciprocal connections play an important role in modulating the cognitive processes behind goal-oriented behaviours and planning, as shown by the seminal works of Balleine and Dickinson on rats [27] (see also [28]).
   Within the neuropsychological literature, it has been proposed [29, 30] that the essential contribution of the Amg to decision making processes consists in evoking the emotions (the "somatic states") that are appropriate to rewards and punishments. The idea is that orbitofrontal cortex, part of prefrontal cortex, elaborates the emotional value of action outcomes on the basis of Amg's activation. Decision making processes, having prefrontal cortex as a principal actor, can then use this information for selecting actions with uncertain payoffs.

The last two functions of Amg involve the regulation of body states, diffused brain states and learning:

1. *Diffused modulation of brain functioning and regulation of learning processes.*
   Efferent connections from CeA project to ventral tegmental area, locus coeruleus and Raphe nuclei, the three main systems of departure of respectively dopaminergic, noradrenergic and serotoninergic innervations directed to virtually all districts of brain.
   Phasic dopaminergic responses at the timescale of milliseconds might underly synaptic reward-based modifications, whereas tonic dopaminergic activation at the timescale of minutes or hours might regulate the intensity of production of neural responses of the target areas [31–33].
   Also norepinephrine operates at different timescales. However, differently from dopamine, its phasic activation does not depend on the rewarding or aversive value of the stimuli, but only on its properties as a signal of novelty [34].
2. *Regulation of body homoeostatic states.*
   Animals reactions to events include unlearned patterns of modulation of homoeostatic body parameters such as blood pressure, heart rate, gastric and intestinal motility, and others. Efferent connections from Amg can control these regulatory processes (URs) depending on particular biologically-salient stimuli (USs) or neutral stimuli associated with them (e.g., the "conditioned stimuli", or "CSs", used in conditioning experiments).
   This kind of modulation passes through the activation of the CeA and its connections to the hypothalamus and autonomic centers of brainstem, including the vagal nuclei and the sympathetic system [18, 19, 21].

4 Conclusions

This paper reviewed empirical neuroscientific evidence with the goal of showing that amygdala, at the core of various brain systems underlying the affective regulation of behaviour in organisms, can be viewed as an interface between organisms' cognitive processes and body homeostatic regulations. In particular, the review showed how amygdala implements some important mechanisms that allow the association of various environmental stimuli and context with the triggering of "behaviours" directed to regulate organisms' body states, their interactions with the outer environment, and the general functioning of brain itself. These associative functions are fundamental for adaptive behaviour as they allow organisms to assign subjective saliency and value to experienced world states, so enhancing the adaptive power of their cognitive processes.

We believe that the re-organisation of empirical knowledge and data on emotions within a computational perspective, as done here, will help to both highlight the general principles underlying emotional regulation of behaviour in organisms and to design robots' controllers endowed with a flexibility and autonomy comparable to that of organisms.

References

1. Parussel, K., Cañamero, L.: Biasing neural networks towards exploration or exploitation using neuromodulation. In Marques de Sá, J., Alexandre, L.A., Duch, W., Mandic, D.P., eds.: Proceedings of the 17th International Conference on Artificial Neural Networks, Part II (ICANN2007). Volume 4669 of Lecture Notes in Computer Science., Berlin, Springer-Verlag (2007) 889–898 Porto, Portugal, 9-13 September 2007.
2. Balkenius, C., Morèn, J.: Dynamics of a classical conditioning model. Autonomous robots **7**(1) (1999) 41–56
3. Montebelli, A., Herrera Pérez, C., Ziemke, T.: An analysis of behavioral attractor dynamics. In Almeida e Costa, F., Mateus Rocha, L., Costa, E., Harvey, I., Coutinho, A., eds.: Proceedings of the 9th European Conference on Artificial Life (ECAL2007). Volume 4648 of Lecture Notes in Computer Science., Berlin, Springer-Verlag (2007) 213–222 Lisbon, Portugal, 10-14 September 2007.
4. Cañamero, L.: Emotion understanding from the perspective of autonomous robots research. Neural networks **18**(4) (2005) 445–455
5. Arias, C., Chotro, M.G.: Amniotic fluid can act as an appetitive unconditioned stimulus in preweanling rats. Developmental psychobiology **49**(2) (2007) 139–149
6. Mannella, F., Mirolli, M., Baldassarre, G.: The role of amygdala in devaluation: a model tested with a simulated robot. In Berthouze, L., Prince, C.G., Littman, M., Kozima, H., Balkenius, C., eds.: Proceedings of the Seventh International Conference on Epigenetic Robotics, Lund University Cognitive Studies (2007) 77–84
7. Mannella, F., Zappacosta, S., Mirolli, M., Baldassarre, G.: A computational model of the amygdala nuclei's role in second order conditioning. In: From Animals to Animats 10: Proceedings of the 10th International Conference on the Simulation of Adaptive Behaviour. (2008)
8. Fiore, V., Mannella, F., Mirolli, M., Gurney, K., Baldassarre, G.: Instrumental conditioning driven by neutral stimuli: a model tested with a simulated robotic rat. In Berthouze, L., Schlesinger, M., Balkenius, C., eds.: Proceedings of the Eighth International Conference on Epigenetic Robotics, Lund, Lund University Cognitive Studies (2008) submitted

9. Lanuza, E., Nader, K., LeDoux, J.E.: Unconditioned stimulus pathways to the amygdala: effects of posterior thalamic and cortical lesions on fear conditioning. Neuroscience **125**(2) (2004) 305–315
10. Blair, H.T., Sotres-Bayon, F., Moita, M.A.P., LeDoux, J.E.: The lateral amygdala processes the value of conditioned and unconditioned aversive stimuli. Neuroscience **133**(2) (2005) 561–569
11. Knight, D.C., Nguyen, H.T., Bandettini, P.A.: The role of the human amygdala in the production of conditioned fear responses. Neuroimage **26**(4) (2005) 1193–1200
12. Maren, S.: Building and burying fear memories in the brain. Neuroscientist **11**(1) (2005) 89–99
13. Baxter, M.G., Murray, E.A.: The amygdala and reward. Nature reviews neuroscience **3**(7) (2002) 563–573
14. Cardinal, R.N., Parkinson, J.A., Hall, J., Everitt, B.J.: Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neuroscience and biobehavioral reviews **26**(3) (2002) 321–352
15. Balleine, B.W., Killcross, S.A., Dickinson, A.: The effect of lesions of the basolateral amygdala on instrumental conditioning. The Journal of neuroscience **23**(2) (2003) 666–675
16. Balleine, B.W., Killcross, S.: Parallel incentive processing: an integrated view of amygdala function. Trends in neurosciences **29**(5) (2006) 272–279
17. McDonald, A.J.: Cortical pathways to the mammalian amygdala. Progress in neurobiology **55**(3) (1998) 257–332
18. Pitkänen, A., Jolkkonen, E., Kemppainen, S.: Anatomic heterogeneity of the rat amygdaloid complex. Folia morphologica **59**(1) (2000) 1–23
19. Davis, M., Whalen, P.J.: The amygdala: vigilance and emotion. Molecular psychiatry **6**(1) (2001) 13–34
20. Hobin, J.A., Goosens, K.A., Maren, S.: Context-dependent neuronal activity in the lateral amygdala represents fear memories after extinction. The Journal of neuroscience **23**(23) (2003) 8410–8416
21. Kalin, N.H., Shelton, S.E., Davidson, R.J.: The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. The Journal of neuroscience **24**(24) (2004) 5506–5515
22. Shi, C., Davis, M.: Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. The Journal of neuroscience **19**(1) (1999) 420–430
23. Gallagher, M., Graham, P.W., Holland, P.C.: The amygdala central nucleus and appetitive Pavlovian conditioning: lesions impair one class of conditioned behavior. The Journal of neuroscience **10**(6) (1990) 1906–1911
24. Hatfield, T., Han, J.S., Conley, M., Gallagher, M., Holland, P.: Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. The Journal of neuroscience **16**(16) (1996) 5256–5265
25. LaBar, K.S., Cabeza, R.: Cognitive neuroscience of emotional memory. Nature reviews neuroscience **7**(1) (2006) 54–64
26. Phelps, E.A.: Human emotion and memory: interactions of the amygdala and hippocampal complex. Current opinion in neurobiology **14**(2) (2004) 198–202
27. Balleine, B.W., Dickinson, A.: Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology **37**(4-5) (1998) 407–419
28. Yin, H.H., Knowlton, B.J.: The role of the basal ganglia in habit formation. Nature reviews neuroscience **7** (2006) 464–476

29. Bechara, A., Tranel, D., Damasio, H., Damasio, A.R.: Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. Cerebral cortex **6**(2) (1996) 215–225
30. Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W.: Insensitivity to future consequences following damage to human prefrontal cortex. Cognition **50**(1-3) (1994) 7–15
31. Schultz, W.: Multiple dopamine functions at different time courses. Annu Rev Neurosci **30** (2007) 259–288
32. Dreher, J.C., Burnod, Y.: An integrative theory of the phasic and tonic modes of dopamine modulation in the prefrontal cortex. Neural networks **15**(4-6) (2002) 583–602
33. Beninger, R.J., Gerdjikov, T.: The role of signaling molecules in reward-related incentive learning. Neurotoxicity research **6**(1) (2004) 91–104
34. Berridge, C.W., Waterhouse, B.D.: The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. Brain research. Brain research reviews **42**(1) (2003) 33–84

Affective Modulation of Embodied Dynamics

Robert Lowe, Pierre Philippe, Alberto Montebelli, Anthony Morse, Tom Ziemke

University of Skövde, School of Humanities and Informatics,
Skövde, Sweden
Robert.Lowe,@his.se

Abstract. The coupling of non-neural internal states essential to an agent's survival with artificial nervous systems can increase adaptivity in terms of (1) exploitation of sensorimotor possibilities, (2) regulation of internal and behavioural activity and (3) behavioural emergence via complex network dynamics that enable the agent to contend with a challenging and unpredictable world. This paper provides a review of recent research on the relevance of non-neural internal states to adaptive behaviour in the field of adaptive robotics. The paper derives a methodological approach that promises to further extend our understanding of how non-neural internal states can increase adaptivity in robots as relevant to the proposed core benefits extracted.

1 Introduction

The relative importance to adaptive behaviour of the dynamics of non-neural variables considered essential to organismic viability (such as systolic blood pressure or glucose level) has tended to be neglected by approaches to computational modelling that have considered neural dynamics and non-neural dynamics to be largely separable. The dominance of this position has, perhaps, weakened in parallel with an increased interest in the role embodiment plays in cognition as compared to the more traditional cognitivist and connectionist approaches. Nevertheless, even within the embodied AI/cognitive science research program emphasis has been placed on the link between behavioural activity and neural dynamics with the justification being that a degree of separability between nervous system activity and non-neural activity (i.e. governed by core metabolic processes) exists such that the activity of the former often cannot be reliably determined by that of the latter ([1]; [2]). However, in recent years research linking non-neural internal mechanisms to neural dynamics and behaviour suggests that far from providing an unnecessary burden on the adaptive capabilities of agents endogenously generated modulation of an agent's internal states can improve adaptivity ([1]; [3]) in robotic and simulated agents.

In a very general sense, intelligent/adaptive behaviour (behaviour beneficial to the agent's ongoing needs) may be best served by the mutual modulation of neural and non-neural states loosely coupled according to situated (evolutionary) demands.

In this paper, we provide evidence to suggest that relative to agents whose artificial nervous system dynamics are unconstrained by non-neural variables, agent design that is more 'neurophysiologically' motivated can allow for improved adaptivity with respect to:

1) *exploitation of sensorimotor possibilities,*
2) *efficient regulation of internal and sensorimotor activity,*
3) *behavioural emergence via complex network dynamics.*

where internal here refers to both neural and non-neural states.
The remainder of the paper is dedicated to demonstrating how these three points apply and subsequently a methodological approach is derived and described. This approach involves the use of robotic agents required to regulate behaviour and internal states according to a two-resource problem ([4]) necessitating an appropriate trade-off between task performance and maintenance of a non-neural viability constraint. This constraint, in directly regulating neural activity and indirectly behaviour, and in being both directly (neurally) and non-directly (behaviourally) regulated, is expected to further elucidate the extent to which points 1-3 above pertain to the generation of adaptive behaviour.

The breakdown of the paper is as follows: Section 2 focuses on the relevance of non-neural internal states to adaptive behaviour. How such brain-body interplay has been modelled in embodied agents in recent years is then evaluated with respect to the three points outlined above. This section provides the motivation for the agent architectural choice the details of which are expounded in Section 3. This section also provides a brief description of an appropriate task relevant to the development of embodied architectures that benefit in the sense of points 1-3 above. Finally, Section 4 briefly relates the methodological approach, essentially centred on affective mechanisms, to emotional regulation with respect to brain-body-behaviour dynamics.

2 The Relevance of Non-neural Bodily States

2.1 Essential variables in organisms – glucose as a paradigmatic example

The ways in which variables essential to an organism's viability affect and modulate nervous system dynamics are complex. An essential variable may be defined as one that must be maintained within certain limits and that if its value goes beyond such limits it will precipitate changes to the organism that will be seriously and sometimes irreversibly damaging to its prospects of survival ([5]). Variables such as systolic blood pressure, glucose and water levels in the blood must be maintained within fairly rigid limits for an organism of a given species to be able to survive. Following [6], Ashby proposed that adaptive behaviour entails the homeostatic maintenance of essential variables, i.e. "a form of behaviour is adaptive if it maintains the essential variables within physiological limits" ([5], p.58).

A paradigmatic example of an essential variable related to the energetic capabilities of organisms whose values must be maintained within certain limits is glucose. Levels of glucose in the blood must remain in the region of 5mmol/l in humans so that the organism may continue to function (lower than 3mmol/l and higher than 11mmol/l can lead to a cascade of harmful effects including diabetes – see, for example, [7], for further details).

Glucose levels are regulated through the interaction of physiological, mental/psychological and behavioural mechanisms/processes. Physiological mechanisms exist to regulate blood levels of glucose to ensure a relatively stable level that does not impinge dramatically on other processes (e.g. brain activity). While glucose, once released into the blood stream, dissipates in the absence of external replenishment (i.e. through intestinal absorption into the blood), much glucose is regulated through delivery to muscle and other peripheral tissues and then ‘recycled’ in the form of lactate and alanine returning to the liver as substrate for gluconeogenesis. Glucose is also released into the bloodstream via a process of glycogen breakdown, which is stimulated, as for gluconeogenesis, by a decreased insulin/glucagon ratio (see [7]).

As a form of energy, glucose is also oxidized in the brain, and thereafter irreversibly ‘lost’ (i.e. non-recycled), and this plays a vital role in cognition. The human brain is thought to oxidize around 120g of glucose per day which amounts to roughly 20% of the energy expenditure of the whole body (see [8]; [7]). Research in psychology indicates that there is a reciprocal relationship between mental ability and blood glucose regulation. There is evidence that self-control, for example, is correlated with glucose blood level. Self-control is impaired by low blood levels of glucose ([9]; [10]; [11]) whilst glucose administration to subjects has been associated with increased performance on a number of cognitively demanding tasks (e.g. [12]) such as word retrieval tasks. Kennedy and Scholey [12] observe that “Brain imaging studies demonstrate that both the rate of blood-to-brain glucose transport ... and glucose metabolism ... are stimulated in task specific areas of the brain. It follows that any fluctuations in the availability of blood-borne metabolic substrates may modulate brain metabolism and thereby cognitive function”, p.63. Glucose levels may also constrain/be regulated by mental effort, via the elevation of heart-rate ([12]), and by hormones such as cortisol.

Finally, glucose regulation may occur through behavioural processes, e.g. energy invested in, and recouped, through successful foraging. When glucose levels increase through feeding, the body stores excess glucose as glycogen in the liver and muscle in order to maintain glucose levels in the blood at optimal functional levels. Furthermore, low glucose levels may stimulate feelings of hunger which can induce increased behavioural activity perhaps as an anticipatory response to ensuing low glucose levels.

Therefore, the regulation of the essential variable glucose, as a form of energy, is achieved through a complex combination of physiological, psychological and behavioural mechanisms. Whilst the focus of the rest of the paper is not on a strict modelling of a particular essential variable such as glucose but rather on modelling the general effects of mutual regulation and modulation of activity of neural states and generic non-neural essential variables, nevertheless, the example of glucose as a form of energy that serves to constrain brain activity and is regulated within certain bounds according to physiological and behavioural processes is a source of inspiration for our modelling approach.

2.2 Exploitation of sensorimotor possibilities

The effects of non-neural constraints on the sensorimotor activity of embodied agents has been tested in a number of recent experiments at different levels of abstraction in the form of simulated energy values. McHale and Husbands [13], for example, demonstrated that by accounting for energy within the coupled simulated robot-environment system via the imposition of an additional cost carried into the fitness function the evolved robot tended to be more energy-efficient with respect to carrying out a simple sensorimotor task. It was found that the robot made better use of its sensors, rather than relying heavily on energy-expensive motor activity as was the case for robots in the control condition. Essentially, the robots exhibited an ability to more easily exploit sensorimotor possibilities as a consequence of the energy constraint. This particular model simulated the kinetic and potential energy cost of the entire system which, effectively, expressed itself in the form of level of motor activity. The energy cost did not take into account the effects of deceleration, sensor activity, or nervous system activity, in this case based on an arbitrarily recurrent artificial neural network – a so-called GasNet (see [14], and also section 3 in this paper). However, the results offered a simple proof of concept that the inclusion of energy constraints in a dynamical system need not be a hindrance but in fact a help to the evolvability of adaptive behaviour. While the energy cost in this particular experiment was restricted to motor activity, an energy constraint could easily be applied to nervous system activity in a similarly abstract manner.

Melhuish and colleagues, ([15]; [16]), similarly adopting a heavily bio-physico-chemically inspired approach, have developed a robot running on microbial fuel cells that can effectively convert fuel into biochemical energy through a digestive 'bacterial sludge' in an anode department that filters electrons through a proton exchange membrane into a cathode department. When the threshold of activity in a bank of capacitors in the cathode is exceeded, effectors are utilized that allow the robot to perform elementary phototaxis and light sensing behaviours in its environment and allow it to continue to 'devour' applicable substrate. The delay between 'digestion' and behavioural output could be said to represent a sort of storing of energy that leads to 'pulses of behaviour' ([16]). This might be an example of a particular non-neural constraint that could potentially be exploited by emergent behaviours in a robot endowed with a more complex nervous system. For example, the robot might use more energy-efficient behaviours (including, for example, simulated behaviours - 'thoughts') during the necessary charging period at lower thresholds of energy accumulation. The lack of 'metabolic self-production' ([17]) involved in Melhuish and colleagues' use of pre-designed microbial fuel cells to generate energy might be viewed as a constraint on the extent to which the robots can produce adaptive behaviour, however, we feel that such a complex relation between non-neural metabolic processes, neural and behavioural processes is a step in the right direction - it can be argued that such a designed imposition of complex biochemical (metabolic) constraints may engender the types of flexible behavioural strategies not readily available to robots whose 'energy' levels are simply recharged via a battery.

The point that these studies offer is that modulation of nervous system dynamics via some form of energy-relevant non-neural constraints need not restrict the behavioural possibilities of the organism but, if anything, may serve to increase adaptive capabilities.

2.3 Efficient regulation of sensorimotor activity

The importance to adaptive behaviour in artificial mobile agents of linking non-neural internal states to neural states and sensorimotor activity has been noted by Parisi [18] who suggested that emphasis in the field of evolutionary robotics had been focused too narrowly on external behaviour. He used the term 'internal robotics' to emphasize the need to account for non-neural states and listed a number of key points that might constitute this sub-discipline of robotics[1] which we will refer to throughout different sections of this paper. Work carried out by Parisi and his co-workers (e.g. [19]; [20]) consistent with the internal robotics approach has highlighted the extent to which regulation of the activity of an artificial nervous system by a non-neural state may be essential to adaptive behaviour. For example, Mirolli & Parisi [19] investigated the importance of combining a biological clock mechanism with a light sensor to regulate agent activity so as to maximize utilization of resources through the regulation of behavioural activity. A reason posited, by the researchers, as to why biological organisms sleep is that relative immobility in dark conditions reduces energy loss through ineffective foraging behaviour. Similarly, robots might do well to dampen their motor activity when sensor activity is noisy due to poor lighting conditions, e.g. when it is dark. However, it is undesirable to have robots (or biological organisms) remain inactive indefinitely due to constant poor lighting conditions – as might apply if the agent happened to dwell for a period in a poorly lit area such as a cave. To cater for the need to cycle activity appropriately Mirolli and Parisi [19] produced experiments linking a light sensor to a simulated biological clock; thus, allowance was made for an energy-efficient 'sleep-wake' cycle that was nevertheless non-prohibitive to adaptive behaviour in conditions of more or less constant poor lighting. They tested a number of configurations regarding the links between the light sensor, biological clock, and simple feed forward artificial neural network. They found that independence of the biological clock mechanism and the light sensor with respect to their influence on the artificial neural network's (ANN) motor output units produced the worst results in terms of foraging behaviour. The best results involved the use of the configuration whereby only the biological clock modulated the motor output units while the light sensor instead modulated the biological clock mechanism. Therefore, there is a sense in which, a non-neural bodily mechanism modulated the activity of the ANN through altering the mapping between sensory input and motor output.

Essential variables that may serve to modulate neural and behavioural activity in agents were discussed in the previous section with respect to Ashby [5] who also has been a source of inspiration for a number of experiments carried out by Di Paolo (e.g. [21]; [22]; [23]). His research emphasis has been primarily on the modelling of homeostasis of neural, as opposed to non-neural, 'essential variables' with the rationale that neurons must have activity levels maintained within certain limits in

[1] This is centred on differences between how neural systems interact with non-neural systems and how neural systems interact with the external environment.

order to remain viable (see, e.g., [24]). The aim of Di Paolo's work has been to establish ways by which homeostatically regulated neural and behavioural activity could be achieved on different tasks.

Di Paolo [22] suggested that while *living* agents may require metabolism, a more obviously realizable task for cognitive science and AI modellers is the producing of agents that conserve a *way of life* based on habit formation maintained in the face of perturbations via the use of adaptive mechanisms. In this sense, it may be the case that "adaptation occurs in the animal world ... not because the organismic survival is challenged directly but because the circular process generating a habit is" ([22], p.13). Di Paolo [22] also acknowledged that in evolved organisms the *way of life* that tends to be adhered to is one that matches internal stability with behavioural stability; that is, internal stability promotes behavioural stability and vice-versa. He provides an example of this phenomenon through the description of a simple experiment where an agent is required to regulate a non-neural variable - battery level - between upper and lower bounds which the agent is able to achieve via simple phototactic behaviour around an 'energizing' light source in the face of sensory perturbations – sensor inversion. In this particular example, deviations from the 'homeostatic' battery level led to modifications of the sensor-motor ouput mapping rather than affecting a non-trivial internal artificial nervous system.

McFarland and Spier [4] proposed that self-sufficient robots are required to perform basic cycles of behaviour, like biological organisms, that involves trading off energy and work. Energy may relate to the regulation of a battery level, or in the case of a living organism some particular nutrient, perhaps glucose, while work might relate to some pre-designed task for the robot, or in the case of the biological organism it might relate to reproductive activity or investment in offspring. This trade-off between energy regulation and task performance (work) was labelled the 'two-resource problem' and McFarland and Spier provided a simple *Cue-Deficit Model* that enabled action selection as a function of extent of deficit of a single essential variable and proximity of resource. It was suggested that the model provided a flexible alternative to Brooks' subsumption architecture ([25]) since it allowed for 'opportunistic behaviour', i.e. a particular motivated behaviour could be interrupted before consummation in light of a more promising alternative. The adaptive capabilities of McFarland and Spier's cue-deficit model have been further verified by Avila-García and Cañamero ([26]; [27]) who focused on *non-neural* essential variables the values of which were replenished via external resources. Avíla-Garcia and Cañamero [26] found that robots were able to deal with the 'two-resource problem' through homeostatic regulation of essential variables manifesting in increasingly efficient behavioural activity cycles (see [4]), even in the absence of neural plasticity.

The above work by Di Paolo and Iizuka as well as Avila-García and Cañamero (also see [28]; [29] for highly biologically inspired approaches) might loosely be labelled *homeostatic internal robotics* or perhaps more accurately *homeostatic interactive robotics* given the focus on the interaction between sensorimotor activity and internal dynamics to produce regulated body-nervous system-behavioural activity. These approaches also demonstrate how modulation of at least behaviour, if not nervous system activity, by non-neural internal variables can promote adaptive regulatory behaviour.

2.4. Behavioural emergence via complex network dynamics

The production of emergent behaviours via a minimalist coupled embodied agent-environment system through the use of a non-neural essential (energy) variable was demonstrated by Montebelli et al., [30]. The authors, having "systematically selected the simplest available [agent body, internal dynamics, artificial nervous system, environment]" (p.191), found that the input of a single linearly decrementing 'energy' variable to a fully connected feed forward reactive artificial neural network (ANN) could modulate sensorimotor activity in such a way as to give rise to a number of behavioural attractor states commensurate with dynamically changing energy requirements. Individual agents exhibited degrees of phototactic and photophobic behaviour with respect to spatially displaced rewarding and non-rewarding lights, respectively. The slow dynamics of the decrementing energy value with respect to the reactive nature of the ANN, evaluated via spectral and quantitative spatio-temporal analyses, allowed for a complex dynamic producing "an effective non-deterministic action selection mechanism" (p.187) over sets of emergent behavioural attractors. Essentially, the 'brain-body-environment' self-organized relation allowed for complex cognitive ('intelligent') behaviour in spite of the intentional simplicity of the experimental set-up.

The approach of Montebelli et al., [30] accords with the 4th principle of *Internal Robotics* as postulated by Parisi [18]: "For the nervous system the rest of the body is something that is always present and always more or less the same, whereas the external environment can be present or absent and it can be very different at different times", p. 331. The *relative* stability of bodily states – that is, the latency of response to external events of non-neural states (e.g. visceral, metabolic) – as compared to fast acting neural activity provides the embodied agent with an internal dynamic that intrinsically integrates processes operating on different time scales. Montebelli et al., [30] emphasize the importance of the interplay between coupled dynamic systems working on such different time-scales in their work: "[O]ur major concern as designers should be to incorporate in our models the critical level of complexity sufficient to trigger and sustain the process [of 'self-organized coupling of internal dynamics, neurocontrollers, bodies, and environments, all inter- and intra-operating on different timescales']" ([30], p.192) in order for 'autonomous' and 'meaningful' cognition to emerge.

Interestingly, regarding the issue of complexity, Di Paolo [22] similarly acknowledged the requirement of a degree of complexity in order to equip agents with behaviour that might be considered cognitive or emotional, but suggested that in order to understand truly intelligent and adaptive agents complexity alone is insufficient: "The solution ... will require more complex designs for robot bodies and controllers, more complex tasks, and more complex methods of synthesis. But seeking such complexity blindly, by typically restricting the search to achieving more complex behaviours, does not accomplish much." (p.29). As referred to previously, Di Paolo [22] suggested that the 'solution' for modellers that circumvents the need to model living organisms from metabolism up is to model a *way of life* based on what we referred to in section 2.3 as *homeostatic interactive robotics*.

The approach that we propose is one that:

1. accounts for the need to imbue agents with the potential for homeostatic regulation of both internal states and behavioural activity cycles.
2. allows for increased complexity of the type that affords more *organismically inspired* (from [22]) behaviour.

This is achieved through providing an agent with a *mutually modulating* homeostatically regulated non-neural (essential) variable and artificial nervous system (ANN). The ANN abstractly models both fast-acting synaptic transmission and slower-acting neuromodulation. This thereby produces sensorimotor activity regulated according to complex dynamics, reflective of the internal dynamics, necessary to produce cycles of behavioural activity trading off designated task performance and essential variable replenishment in the face of unpredictable and challenging life-time events.

The next section describes the choice of ANN architecture, some details of implementation of the network, and the means by which the network is spatio-temporally coupled to non-neural essential variables.

3 Spatio-Temporal Coupling of the Essential Variable to a Neural Controller: A GasNet Approach

The previous section provided evidence for how non-neural states can constrain the behavioural performance of robots and simulated agents, via a coupling with aspects of the artificial neural network, in ways that potentially allow for more intelligent and adaptive behaviour. It did so by recourse to three aspects deemed to be facilitated by such a coupling: 1) *exploitation of sensorimotor possibilities,* 2) *efficient regulation of internal and sensorimotor activity,* 3) *behavioural emergence via complex network dynamics;* these three aspects were essentially addressed in sections 2.2, 2.3, and 2.4, respectively.

The methodological approach we propose that accounts for all three aspects mentioned above involves the spatio-temporal coupling of a homeostatically regulated non-neural essential variable to an ANN the activity of which is governed by synaptic transmission and neuromodulation. Consistent with the three respective identified aspects above, the embodied neurocontroller has the potential to provide agents with nervous system activity that 1) is modulated by the non-neural essential variable, and potentially constrained by 'energy' (see below), 2) regulates internal (neural/non-neural) and sensorimotor activity in a manner that is reflective of its intrinsic needs, 3) provides temporally complex cycles of behavioural activity as reflected in the different time scales at which the essential variable, synaptic transmission, and neuromodulation, and their interactions, operate.

The way in which we can model these three aspects is done with reference to a number of the seven key features of *Internal Robotics* identified by Parisi [18].

Specifically, the first three points posited by Parisi are relevant here:

1. "The nervous system's interactions with the external environment are predominantly physical, whereas those with the internal environment are predominantly chemical", p. 329.
2. "[There are] [t]wo kinds of influences on the nervous system [neuron-to-neuron, diffusive/neuromodulatory]", p.329.
3. "The circuit 'nervous system-rest of the body' is entirely evolved, whereas the circuit 'nervous system-external environment' is evolved only for the part of the circuit that is in the nervous system", p.330.

Points 1-3 above are of critical relevance as to how we model the coupling between the non-neural essential variable and the artificial nervous system. Points 1. and 2. intimate that activity internal to the organism involving interactions between non-neural bodily and neural states are mostly chemical as mediated by, for example, neurohormones. An organism's interaction with its environment is predominantly physical (though may involve contact with chemical molecules via odours, tastes etc.). Furthermore, point 2. suggests that activity internal to the organism is mediated by eletrical transmission in the case of non-bodily interference and can be referred to as *neuro-transmissory* while non-neural bodily – neural interactivity may be considered *neuro-modulatory*. Neurotransmission depends less on space and more on the particular topology of synaptic connections whereas neuromodulation is affected by space, i.e. the position of neurons in relation to one another. Point 3 indicates that the interaction between non-neural and neural states is 'entirely evolved' whereas the interaction between neural states and the external environment is evolved only in relation to relevant sensory circuitry in the nervous system – the environment itself has not been shaped by the dynamics of the nervous system.

As mentioned earlier, point 4. of the *Internal Robotics* 'manifesto' implies that non-neural states interacting with neural states allows for a shaping of behaviour that expresses itself according to the interactions of processes working on different time scales. Parisi suggests (in point 6) that affective/emotional components arise from the interaction between bodily states and the nervous system (whereas cognitive states emerge from the interaction of the nervous system with the external environment[2]). On this basis, we might suggest that such affective states tend to engender transient emotions, moods, personality traits that guide cognition over different time scales ranging from the agent's immediate present to its entire life-history[3].

There are many candidate approaches from which we may draw inspiration regarding the modelling of nervous system activity that afford complex neurodynamics. For example, Ziemke and Thieme [36] provided a mechanism by which synaptic plasticity could be modulated according to a context that dynamically altered mappings from sensory input to motor output. This embodied ANN successfully resolved time-delayed presentation of a target stimulus with respect to presentation of a light (a sort of conditioned stimulus) during the process of negotiating a number of T-Maze configurations. This form of neuromodulation

---

[2] The neat separation of cognition and emotion is however contested by some researchers (e.g. [31]; [32]; [33]; [34]).

[3] See [35] for a description of a perspective of self-organization of affective states over three such time-scales.

demonstrated how, in principle, synaptic plasticity might be relevant not just for long-term but for short-term memory.

Alexander and Sporns [37] adopted an approach to studying the effects of neuromodulation on time-varied reward prediction tasks. Their approach was heavily neuro-anatomically inspired in that it abstractly modelled neural structures such as the Prefrontal cortex, Ventral Tegmental Area (for dopamine transmission), Sensory Cortex, Olfactory Cortex, Motor Cortex. They were able to demonstrate the relevance of their model to time-varying reward conditioning through testing it in disembodied computational form and transferring the results to an embodied mobile robot (khepera) where sensorimotor activity, sensor range, and spatial configuration of reward-relevant resources were all accounted for.

These embodied neuro-controllers allow for conditioning to time-relevant tasks but may be limited with respect to their ability to deal with more complex dynamical conditioning tasks; that is, tasks that require a highly nuanced sensorimotor interaction over time.

Furthermore, in order to investigate complex basic cycles of activity ([4]), the manner in which the activity of the non-neural essential variable and artificial nervous system are reciprocally modulated should not be pre-designed. The aim here is not to impose a particular neuro-anatomically inspired structure on the network, nor to assume that the non-neural essential variable should interact with the ANN in any pre-defined manner. Rather, in order to adhere to a more situated, embodied approach, essential variable-ANN interaction and spatio-temporal structure should be constrained by the choice of evolutionary algorithm and fitness function.

On this basis, the GasNet ANN ([14]) is viewed as a suitable candidate for permitting a loose structural coupling between body, brain and environment. Interestingly, Parisi [18] himself makes reference to the work done with GasNets in point 2. of his key aspects to *Internal Robotics*. Adherence to Parisi's approach, therefore, compels us to view the manner in which the non-neural essential variable interacts with the ANN to be neuromodulatory (chemical), as opposed to neurotransmissory (electrical). GasNets fulfill this function, neuromodulation does not occur through neurohormonal interaction with the nervous system, but nevertheless interaction between activity of the non-neural essential variable and the (chemical) neuromodulatory activity of the GasNet allows us to follow the guidelines set out by Parisi in order to provide further insights into the three central points made in the opening section of this paper.

3.1 The GasNet

The algorithmic details of the GasNet can be found in [14]; [38]; [39]. One of the key features of the GasNet model is that unlike ANNs that depend exclusively on point-to-point (synaptic) chemical transmissions GasNets model the gaseous diffusible modulatory effects of nitric oxide (NO) which permit a less localized influence on brain activity. The biochemical inspiration for the GasNet model can be followed in [14]. For the purposes of this paper, it is sufficient to understand that individual nodes in the network produce *both* localized (synaptic) *and* diffuse (neuromodulatory) transmission of information which affect the activity of the network over different time scales and render relevant the notion of spatially distinguished sub-structure – nodes are more or less affected by gas emission, or not at all, depending on their position in a two-dimensional space with respect to the gas emitting node.

Thus, rather than producing a 2-dimensional effect of activity in a network (for time and for a single spatial dimension) the GasNet produces a more complex spatio-temporal dimensional effect. This effect can be fine tuned according to the particular environmental problem task over evolutionary time modulating the interactions between nodes with respect to a number of properties. In this sense, the GasNet is arbitrarily recurrent as connections between nodes are selected for evolutionarily.

GasNets have been demonstrated to be particularly evolvable with respect to their performance on a number of tasks involving robots (and simulated embodied/robotic agents). Evolvability is enabled, according to Smith et al., [38]; [39], through the GasNet feature of 'temporal adaptivity'. This property entails the reflection of the complex temporal dynamics of the network in the sensorimotor activity of the embodied GasNet.

3.2 Coupling an Essential Variable with the GasNet

The non-neural essential variable used in the approach we are proposing consists of a single value that decrements as a function of discrete time. At a highly abstract level the essential variable may be considered an essential 'energy' variable.

The essential variable *E* and artificial nervous system *NS* can be configured in three ways:

a. Activity of *E* and *NS* are not linked,
b. *E* modulates *NS*,
c. *E* and *NS* modulate each other.

Figure 1 depicts the three possible *E-NS* configurations with respect to agent-environment interactions.

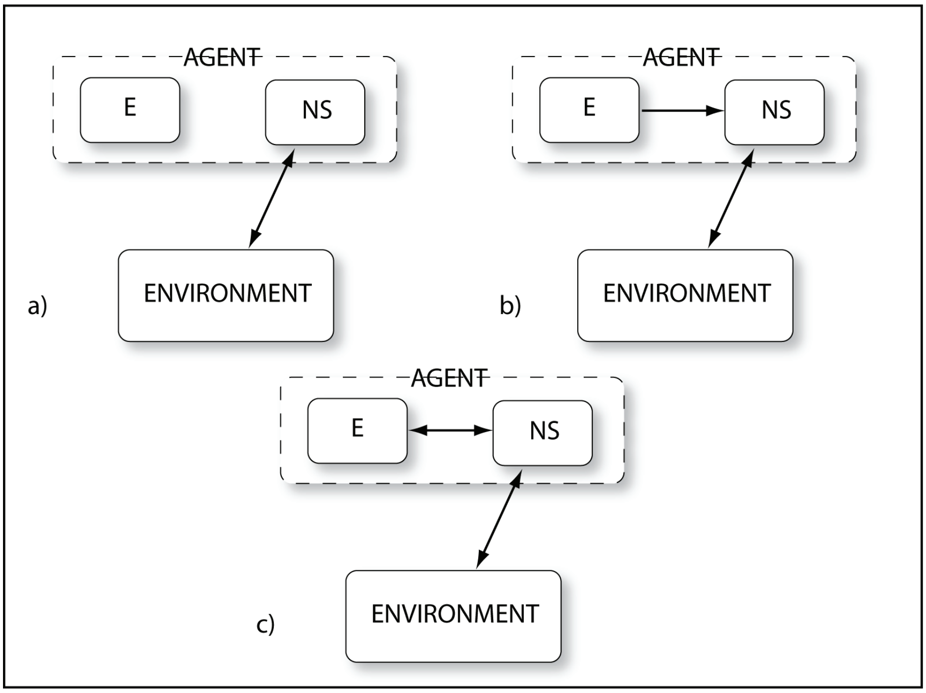

Fig. 1. Three configurations relating agent non-neural essential variable (E), nervous system (NS) and environment according to experimental condition.

This relationship in figure 1c) is essentially that depicted by [18] regarding the importance of body-*NS*-environment interplay for the emergence of cognition.

In the case of fig. 1b) and 1c) *E* serves as a direct constraint on the activity of *NS*. In fig. 1c) activity in the network involves an *E* cost and therefore *NS* will affect the level of *E* which in turn can be modulated by *E*. How and when *E* should affect *NS* and, conversely, be affected by *NS* is a matter of contention. Here, we take inspiration from the approach of Di Paolo [22]. *E* should affect *NS* conditional on whether its homeostatic bounds have been breached thereby influencing the *NS* in such a way that effectively communicates the critical state of *E*. There are, therefore, two aspects regarding the modulation of the activity of the ANN by the essential variable *E* in these experiments:

1) When? – at the point when the homeostatic boundaries of *E* are violated,
2) How? – via effects of the gas emission of *E*-connected nodes in the GasNet as determined evolutionarily.

With respect to 1) upper and lower homeostatic bounds are pre-designed. With respect to 2), given the minimalist nature of the model we are using we choose not to restrict the type of connectivity between *E* and *NS* to particular nodes. Instead, we seek to allow the connectivity between *E* and *NS* to be initially arbitrary but sculpted by the evolution of the GasNet. Therefore, the direction (excitatory/inhibitory) and connectivity of *E*-input to particular nodes is genetically determined. The strength of activity should be pre-designed to be a function of the weighted strength (evolutionarily determined) of deviation from the homeostatic bounds. Figure 2 gives an example *E-NS* evolved configuration.

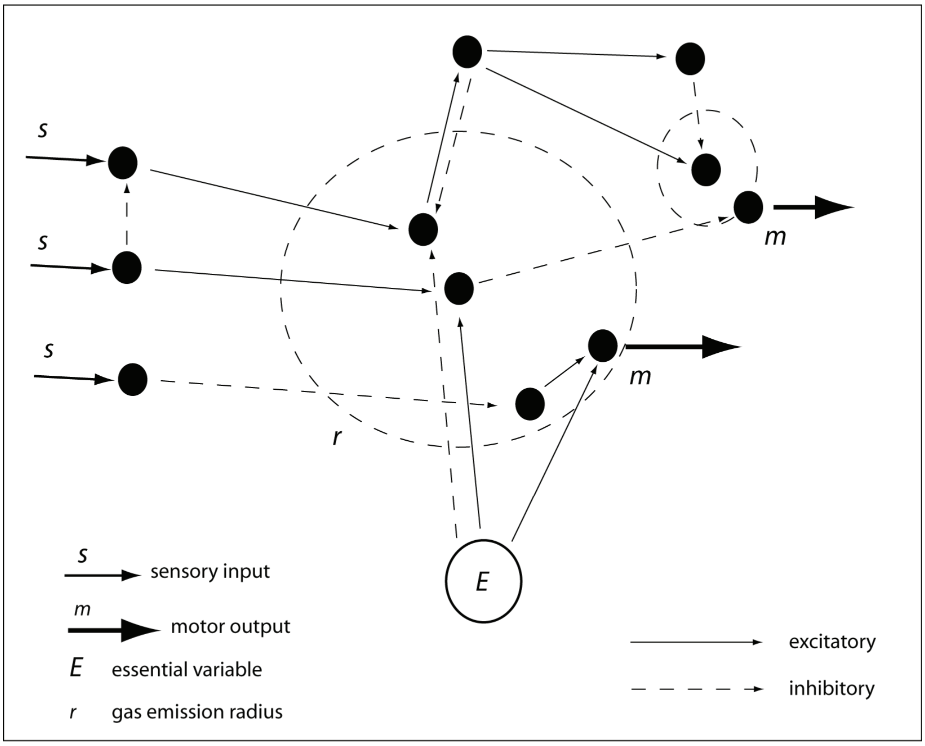

Fig. 2. An example of a possible evolved *E*-modulated GasNet. *E* can have excitatory or inhibitory weighted inputs to nodes within the network. If the strength of activation is sufficient *E* can induce gas emission (dashed-line circles) in the connected node and affect activity of the network in temporally complex ways (adapted from Husbands et al., [14]).

In figure 2 we have an example of how an *E*-modulated GasNet might evolve where *E* inputs are only made to potentially gas emitted nodes (note, nodes do not emit gas if the gas emitting threshold is not exceeded).

3.3. Basic Tasks – The Two-Resource Problem

In order for the three key benefits, as referred to in the introduction section, of a co-dependency between the non-neural essential variable *E* and the artificial neural network (GasNet) to manifest individual agents naturally need to be motile and need to carry out appropriate tasks. The two-resource problem [4] referred to in Section 2.3 is suggested to provide a suitable candidate experimental scenario given simple wheeled robots (e.g. Kheperas, E-pucks). The problem simply entails the robot (or living organism) having to fulfill a particular utility maximizing task (as evaluated in the fitness function) whilst ensuring that its essential variable level (e.g. energy or battery level) is appropriately (i.e. homeostatically) maintained (evaluated in the fitness function but its maintenance is a pre-requisite for completing tasks). Space precludes reference to a detailed experimental scenario and the purpose of this paper is to derive a general approach that could be used to explore the three key benefits of studying non-neural essential variable – nervous system interactive dynamics. Subsequently, we suggest that scenarios such as those featured in experiments by [26]; [27] offer a promising starting point. The approach of [26]; [27] involves mobile robots being required to produce behaviour that maximizes the balance and total level (according to a fitness function) of two essential variables, as replenished by two resources. Similarly, in the generic experimental scenario we propose individual agents inhabit environments with two resources of which one such resource would represent a task that requires exploitation of the temporal dynamic of the *E-NS* embodied system and for which the predictability level is varied over the agent's life-time. The challenge must be sufficient so that differences among the three configurations of the *E-NS* embodied system referred to in 3.2 can manifest providing scope for evolutionary *exploitation of sensorimotor possibilities* (point 1 of introduction). Testing the *efficiency of regulation of internal and sensorimotor activity* (point 2) can be gauged via a utility analysis of the type McFarland & Spier [4] used and literal cycles of behavioural activity facilitate ease of analysis for understanding the cycles of both internal and behavioural activity and their interaction. Finally, it is envisaged that *behavioural emergence via complex network dynamics* (point 3) can be afforded as a consequence of rendering tasks challenging in terms of predictability and the time-dependent component of their being optimally accomplished.

Ongoing work is intended to demonstrate these principles in a number of experiments with reference to the essential variable modulated GasNet testing the three *E-NS* configurations across a range of tasks of varying difficulty with an E-puck robot.

4 The Role of Emotions in Brain-Body-Behaviour Regulation

This paper has discussed the role of the interaction between neural and non-neural states in regulating and producing adaptive behaviour. In a more general sense this relationship can be considered emotional or at least proto-emotional. Parisi [18] himself suggests in point 6 of his guidelines towards achieving an *Internal Robotics* that emotions emerge from interactions between bodily/non-neural states and neural states can be considered emotional whereas cognition emerges from interactions between neural states and the environment. This was referred to briefly in Section 3. The emphasis on the role of embodiment in emotions has driven emotions theory into a domain that focuses on a more dynamical systems perspective, e.g. [40]; [41]; [35], while Damasio [42] views emotions as being rooted in constitutive processes (a nested hierarchy of homeostatic regulation) from metabolic processes 'up' to feeling states (registrations of bodily states in neural activity). We similarly view emotions as being rooted in non-neural variables essential to the organism that interact with neural states. The modelling of internal 'affective' states and their interaction with neural states at a minimalist level according to the principles we have outlined represents a starting point to understanding more complex emotional regulation. It might be suggested, for example, that the evolution of dynamic structure as evaluated via the emergence of stable spatio-temporal structures within the essential variable modulated GasNet can provide insights into the fundamental neurophysiological mechanisms that allow for emotional regulation.

References

1. Di Paolo, E.A, Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 2005. **4 (4)**: 429-452.
2. Barandiaran, X., and Moreno, A, On What Makes Certain Dynamical Systems Cognitive. *Adaptive Behavior*, 2006, **14(2)**: 171-185.
3. Ziemke, T, Adaptive Behavior in Autonomous Agents. *Presence*, 1998, **7(6)**: 564-587.
4. McFarland, D. and Spier, E, Basic cycles, utility and opportunism in self-sufficient robots. *Robot. Autonom. Sys.*, 1997, **20**:179–190.
5. Ashby, W.R, *Design for a brain: The origin of adaptive behaviour*. 1960: Chapman and Hall.
6. Cannon, W.B., *Bodily Changes in Pain, Hunger, Fear and Rage*. 1915: Appleton, New York.
7. Frayn, K.N., *Metabolic Regulation: A Human Perspective*. 1996: Portland Press.
8. Dunbar, R.I.M, The social brain hypothesis. *Evolutionary Anthropology,* 1998, **6**: 178–190.
9. Gailliot, M.T., Baumeister, R.F., DeWall, C.N., Maner, J.K., Plant, E.A., Tice, D.M., Brewer, L.E., and Schmeichel, B.J, Self-Control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, 2007, **92**: 325–336.
10. Gailliot, M.T., and Baumeister, R.F, The physiology of willpower: Linking blood glucose to self-control, *Personality and Social Psychology Review*, 2007, **11**: 303–327.
11. Fairclough, S. H., and Houston, K, A, metabolic measure of mental effort. *Biological Psychology*, 2004, **66**: 177–190.
12. Kennedy, D.O., and Scholey, A.B, Glucose administration, heart rate and cognitive performance: effects of increasing mental effort, *Psychopharmacology*, 2000, **149**:63-71.
13. McHale, G. and Husbands, P, Incorporating Energy Expenditure into Evolutionary Robotics Fitness Measures, In L. Rocha et al. (Eds), *Proc. Alife X*, 206-212, MIT Press, 2006.
14. Husbands, P., Smith, T., Jakobi, N., and O'Shea, M, Better Living through Chemistry: Evolving GasNets for Robot Control, *Connection Science*, 1998, **10(3/4)**: 185-210.
15. Melhuish, C., Ieropolous, I., Greenman, J., and Horsfield, I, Energetically autonomous robots: Food for thought, *Autonomous Robot*, 2006, **21**:187-198.
16. Ieropoulos, I., Melhuish, C., Greenman, J., and Horsfield, I, EcoBot-II: An artificial agent with a natural metabolism. *Int. J. Advanced Robotics Systems*, 2005, **2(4)**: 295-300.
17. Froese, T., and Ziemke, T, Enactive Artificial Intelligence, submitted.
18. Parisi, D, Internal Robotics, *Connection Science,* 2004, **16(4)**: 325-338.
19. Mirolli, M., and Parisi, D, Artificial organisms that sleep. In *Advances in Artificial Life* (New York: Springer), 2003, 377-386.

20. Acerbi, A., and Parisi, D, The evolution of pain, In F. Almeida e Costa (Ed), *Proc. ECAL 07*, Springer , LNCS **4648**, 816-824, 2007.
21. Di Paolo, E.A, Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. In J.-A. Meyer et al. (eds.), *From animals to animats 6: Proc. of the 6th Int. Conf. on Simulation of Adaptive Behavior*, Cambridge, MA: The MIT Press, 2000, 440-449.
22. Di Paolo, E.A, Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. In K. Murase & T. Asakura (eds.), *Dynamical Systems Approach to Embodiment and Sociality,* Adelaide, Australia: Advanced Knowledge International, 2003, 19-42.
23. Di Paolo, E.A., and Iizuka, H, How (not) to model autonomous behavior. *Biosystems*, 2008, **91(2)**, 409-423.
24. Turrigiano, G.G, Homeostatic plasticity in neuronal networks: The more things change, the more they stay the same. *Trends Neurosci*., 1999, **22**:221-227.
25. Brooks, R, A Robust Layered Control System for a Mobile Robot, *IEEE Journal of Robotics and Automation*, **RA-2**, 1986, 14-23.
26. Avíla-García, O., Cañamero, L., and te Boekhorst, R, Analyzing the Performance of "Winner-Take-All" and "Voting-Based" Action Selection Policies within the Two-Resource Problem. In *Advances in Artificial Life, 7th European Conference*, 2003, 733-742.
27. Avíla-García, O., and Cañamero, L, Using hormonal feedback to modulate action selection in a competitive scenario. In *From Animal to Animats 8: Proc. of the 8th International Conference on Simulation of Adaptive Behaviour*, Cambridge, MA: MIT Press, 2004, 243-252.
28. Williams, H, Homeostatic plasticity in recurrent neural networks. In Schaal, S., Ijspeert, A., Billard, A., Vijayakumar, S., & Meyer, J. (Eds.), *From Animals to Animats 8: Proceedings of the 8th International Conference on the Simulation of Adaptive Behavior*, Cambridge MA: MIT Press , 2004, 344-353.
29. Vargas, P., Moioli, R., de Castro, L.N., Timmis, J., Neal, M. and Von Zuben, F.J, Artificial Homeostatic System: A Novel Approach, *Advances in Artificial Life*, **3630**, 2005, 754-764.
30. Montebelli, A., Herrera, C., and Ziemke, T, On Cognition as Dynamical Coupling: An Analysis of Behavioral Attractor Dynamics, *Adaptive Behavior*, 2008, **16(2/3)**: 182-195.
31. Lewis, M.D, Bridging emotion theory and neurobiology through dynamic systems modeling, *Behavioral and Brain Sciences*, 2005, **28:** 169-245.
32. Lewis, M.D., and Todd, R.M, Getting Emotional. *Journal of Consciousness Studies*, 2005, **12(8-10)**: 210-235.
33. Colombetti, G., and Thompson, E, Enacting emotional interpretations with feeling. *Behavioural and Brain Sciences*, 2005, **28:** 200-201.
34. Colombetti, G., Enactive appraisal, *Phenom. Cogn. Sci*., 2007, **6**:527-546.
35. Lewis, M.D, Emotional Self-Organization at Three Time Scales. In Lewis, M.D, Granic, I., eds., *Emotion, Development, and Self-Organization*, 2000, Cambridge University Press, 36-69.
36. Ziemke, T., and Thieme, M, Neuromodulation of Reactive Sensorimotor Mappings as a Short-Term Memory Mechanism in Delayed Response Tasks. *Adaptive Behavior*, 2002, **10(3-4):**185-199.
37. Alexander, W.H., and Sporns, O, An Embodied Model of Learning, Plasticity, and Reward. *Adaptive Behavior*, 2002, **10(3-4)**:143-159.
38. Smith, T., Husbands, P., Philippides, A., and O'Shea, M, Neuronal Plasticity and Temporal Adaptivity: GasNet Robot Control Networks, *Adaptive Behavior*, 2002, **10(3-4)**:161-183.
39. Smith, T., Husbands, P., Philippides, A., and O'Shea, M, Temporally adaptive networks: Analysis of GasNet robot control. In K. Standish et al. (Eds), *Proc. Artificial Life VIII*, MIT Press, 2002, 274-282.

40. Freeman, W.J, Emotion is Essential to All Intentional Behaviors. In Lewis, M.D, Granic, I., eds., *Emotion, Development, and Self-Organization*, 2000, Cambridge University Press, 209-235.
41. Panksepp, J, *Affective Neuroscience: The Foundations of Human and Animal Emotions*. 1998: Oxford University Press.
42. Damasio, A, *looking for Spinoza: Joy, Sorrow and the Feeling Brain*. 2003: William Heinemann: London.