A Functional Approach to Emotion in Autonomous Systems

Ricardo Sanz, Carlos Hernández, Jaime Gómez, and Adolfo Hernando

Abstract The construction of fully effective systems seems to pass through the proper exploitation of goal-centric self-evaluative capabilities that let the system teleologically self-manage. Emotions seem to provide this kind of functionality to biological systems and hence the interest in emotion for function sustainment in artificial systems performing in changing and uncertain environments; far beyond the media hullabaloo of displaying human-like emotion-laden faces in robots. This chapter provides a brief analysis of the scientific theories of emotion and presents an engineering approach for developing technology for robust autonomy by implementing functionality inspired in that of biological emotions.

1 Introduction

The central tenet of engineering research is the development of technology for achieving some desired level of performance in artificial systems: 220 volt in a wall socket, 240 k/m in a car, 80 Hz beat in a pacemaker, etc.

Once the development of the base technology—electrical engineering, mechanical engineering, embedded electronics—lets reach this performance level, a second aspect gains in importance: maintaining this performance. The maintenance of a certain level of performance is obvious in the pacemaker or wall socket, where pumping rate and voltage must be maintained at certain values; it may be less obvious in the car with its continuously varying road conditions, but is clear for the speed which in modern vehicles is to be maintained by the cruise control system. The need for performance keeping may be not so easy to pinpoint; e.g. for the brakes, where it is not so clear what is the variable to be kept within a certain range, although, if carefully analysed, you may find one: the braking power.

R. Sanz, C. Hernández, J. Gómez, and A. Hernando
Autonomous Systems Laboratory (ASLab-UPM), Universidad Politécnica de Madrid,
José Gutiérrez Abascal 2 28045 Madrid
e-mail: ricardo.sanz@upm.es; carlos.hernandez@upm.es; jd.gomez@upm.es;
adolfo.hernando@upm.es

A. Hussain et al. (eds.), *Brain Inspired Cognitive Systems 2008*, Advances in Experimental Medicine and Biology 657, DOI 10.1007/978-0-387-79100-5_14,

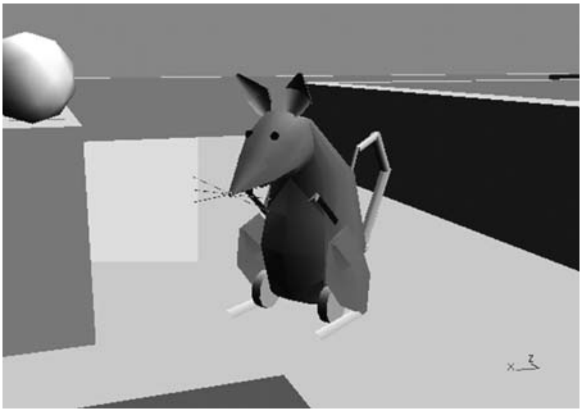

Fig. 1 The IST ICEA project is focused in the extraction of integration patterns among the cognitive, emotional and autonomic systems of the rat. These are evaluated on technical systems including physical and simulated rats

1.1 Sustainable Performance

The preservation of performance levels shall be understood in relation with the concepts of *resilience* and ultimately *robust autonomy*, which relates to the capability of systems to keep their proper functioning despite the uncertain and sometimes hampering and even hazardous dynamics of the environment where they perform. This idea of sustaining performance obviously maps in different ways to different kinds of applications that have more or less resilient structures in changing environments.

In our research of robust autonomy, we focus on two domains of artificial systems: large-scale industrial plants and mobile robotics. While in the case of industrial plants sustaining performance maps crystal clear to maintaining the production rate and keep it within a quality range, this mapping in a mobile robot scenario is far from evident.

In order to achieve *robust autonomy*, one approach could be to build systems physically robust, so as to minimise the effect of external disturbances from the environment. However this approach is always cost prohibitive—i.e. let us think of building a bulldozer-like field robot so that it does not need to avoid obstacles—not to say usually unrealisable—consider a thermodynamically isolated kiln with no need for a temperature control. Discarded this somewhat outdated "brute force" stratègy, engineers now must look for a more "intelligent" approach, where such a term does not only refers to smartness by their part, but also to the capability of the built systems to take advantage of information (Sanz et al. 2000).

To effectively and efficiently address the problem of robust autonomy, the material and energy flows from/to the environment ought to be accompanied by the corresponding flows of information that enable the system to manage the uncertainty in the operational conditions. Therefore artificial systems must build informational structures which implement relevant information about the environment, themselves and the interaction between both. Consequently, the systems are operationalised;

they can cope successfully with the inherently dynamic environment, maximising their effectivity when they pursue the performance goal they were designed for. The construction of maximally effective systems seems therefore to pass through the proper exploitation of goal–centric self-evaluative capabilities that let the systems teleologically self-manage.

In the context of the EU-funded project ICEA[1], this search for resilience is focused on the way that cognitive, emotional and autonomic subsystems are integrated into a single control architecture: that of the mammal brain. The special focus on emotions is due to the fact that emotions, conceived here as internal states and processes, seem to play this kind of role of valence–centric modification in biological systems. Several brain subsystems are being investigated in this direction, especially basal ganglia, amygdala and hippocampus, because of their involvement in basic emotional and cognitive processes: the hippocampus plays a key role in spatial cognition and memory, the amygdala is a main centre for emotion and basal ganglia are involved in valenced decision making.

1.2 Emotion for Engineering Sustainable Performance

Most research in applying emotion to technical systems has focused on the involvement of the display of emotions (Darwin, 1872) in social interaction and communication, and its application for the improvement of human-artificial systems interfaces (see Figure 2).

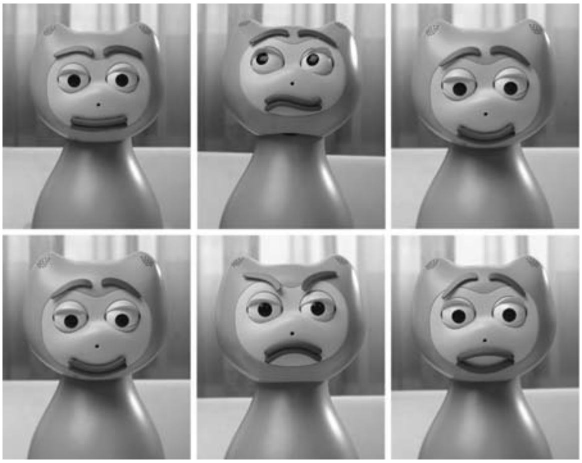

Fig. 2 iCat is the Philips research platform for studying human-robot interaction topics, intended to stimulate research in this area by building a research community through supporting a common hardware and software platform (from philips.com)

[1] www.iceaproject.eu

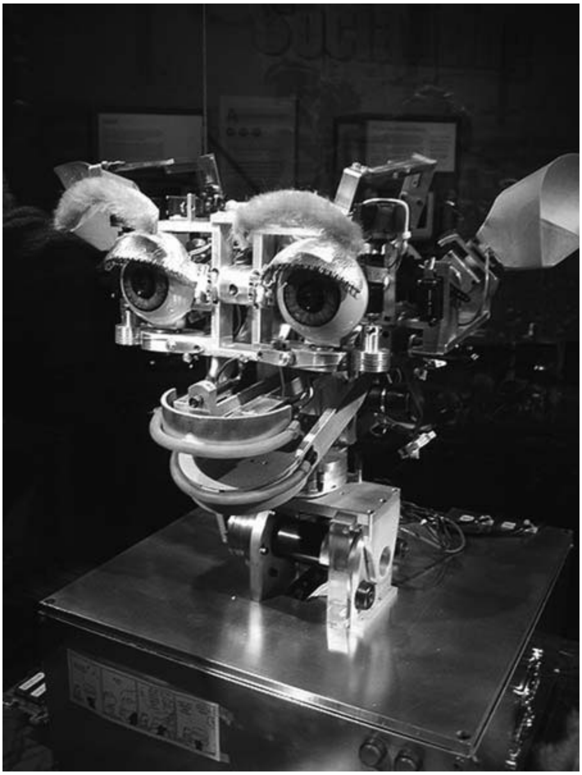

Fig. 3 The MIT robot Kismet is a paradigmatic example of research in the display of emotions by robots (Breazeal, 2000)

However, beyond the obvious capability for displaying human-like, emotion-laden faces in robots (see Figure 3), emotional mechanisms may play a *critical role in the sustainment of function in changing environments*. From a layman's understanding of this hypothesis, we could take the example of the fear one would experience upon discovering an intruder in one's home. That strong emotion elicits a different state in humans: we enhance our attention to sensory systems (i.e. hearing), the discharge of adrenaline prepares our motor system for faster responses, increases our heart rate, etc. Everything is aimed towards the maintenance of a high-level function which we could label as "survival". From a more formal, scientific perspective, it is commonly agreed that biological emotions, as considered in the previous example, are an evolved mechanism for adaptation related to a certain appraisal of internal and external events (Botelho, 2001). This appraisal seems to assign a certain value to stimuli, external or internal, related to the goal tree of the system (i.e. survival, reproduce, search food, avoid predators) and helps reconfigure the system accordingly. Let us take a National Geographic example. When in the mating epoch, a gazelle may feel hungry. That feeling (a bodily emotion) helps reconfiguring its behaviour toward a goal that has gain priority. Going further with this example, suppose that, when searching for food, the gazelle hears the possible presence of a lion or other predator: now the emotion of fear enhances her sensory system, to confirm the presence of the predator and locate it, and shifts her brain into survival mode after the decision-reconfiguration typically called *fight or flight*.

This description of core functionality of emotions in biological systems parallels the functional needs in artificial systems—presented in the previous section—in the pursue of robustness and sustainable performance. In this chapter we will analyse some of the most relevant theories on emotions and coalesce them into a theoretical framework—the *ASys Framework*—for addressing the technical issues of building similar mechanisms into artificial systems to achieve greater levels of resilience and performance.

2 A Review of Emotion

A common, somewhat folk, theory of human emotions say that they are, to a large extent, subjective and non deterministic. This accounts for the obvious fact that apparently identical stimuli may raise different emotions in different humans, and the same individual may experiment different emotions in response to similar stimuli. Obviously, from a purely systemic perspective, there must be some disparity between systems if the behaviour differs, but the ascription of the variety in emotional response by dilettantes have gone from subtle details in otherwise apparently identical stimulus to the very possibility of human freedom and non-determinism.

Starting at the last decades of 19th century, with the work of William James (James, 1884), and through the 20th century, the scientific community has managed to turn the ancestral and pre-scientific ideas about emotions into theory-laden models, were the sustaining theories formal or not. Nowadays emotions are studied just as another natural phenomenon of living systems, such as digestion or homeostasis, and, despite the diversity of approaches and theories, there is a consensus that emotions are an evolved adaptivity mechanism related to situation assessment and decision making (Panksepp, 1998).

Some researchers (Ventura and Pinto-Ferreira, 1999) have also pointed out a division regarding the scientific analysis of emotion which comprises an external, social perspective of emotion as it is involved in communication between individuals through the display by them of emotional states and attitudes, and on the other hand a internal point of view considering how emotion is involved in decision making processes.

We can summarise the usual understanding of emotions into several related aspects (Bermejo Alonso, 2006):

A1 – how emotional behaviour is *triggered* by event surrounding the agent;
A2 – how emotion is *manifested* (displayed) by/within the agent; and
A3 – how emotion is *felt* by the agent.

This heterogeneity notwithstanding, it is necessary to think about basic physiological principles going down to neural-hormonal mechanisms that make a particular event "emotional". Through this section we will summarily analyse the scientific principles and theories about emotions abstracted so far.

2.1 Classical Models

The classical theories of James-Lange or Cannon-Bard address the causality of the relation between A1-A2-A3. The model of James-Lange states that in animals the triggering (A1) are experiences in the world, the autonomic nervous system then creates physiological events such as increased heart rate or muscular tension (A2), and then emotions come as conscious feelings (A3) which come about as a result of these physiological changes (rather than being their cause as the Cannon-Bard model postulates).

Plenty of models can be found in the literature among which we would like to distinguish—for their closeness to the ASys emotion model—the model of Arnold (1960), due to its relation with the shaping of action tendencies; the model of Frijda (1987), due to its perspective of emotions constituting forms of action readiness; and the models of Plutchik and Kellerman (1980) and James (1884) in relation with the bodily basis of adaptive mechanisms.

2.2 Damasio's Model

When relating emotion to the rest of the evolved adaptive mechanisms in humans, from the hormonal system to the more cognitive ones such as consciousness, one of the most complete models is that of Damasio. The somatic marker hypothesis (Damasio, 1999) and related machinery can be used to provide a deeper understanding of emotional system organisation (see Figure 4), and accounts for all A1-A2-A3, with special relevance for the last two aspects. This structure is among the architectures explored in the beforehand mentioned ICEA project in order to provide a coherent picture of the integration of cognitive, emotional and autonomic aspects in mammals.

Damasio's is a concrete proposal in line with what is needed from a scientific and technical approach to emotion and cognition (Ortony et al. 1988). Architectural approaches go beyond the three behavioural and phenomenal aspects mentioned before (triggering, display and feeling) addressing what are the physiological mechanics for all this functioning.

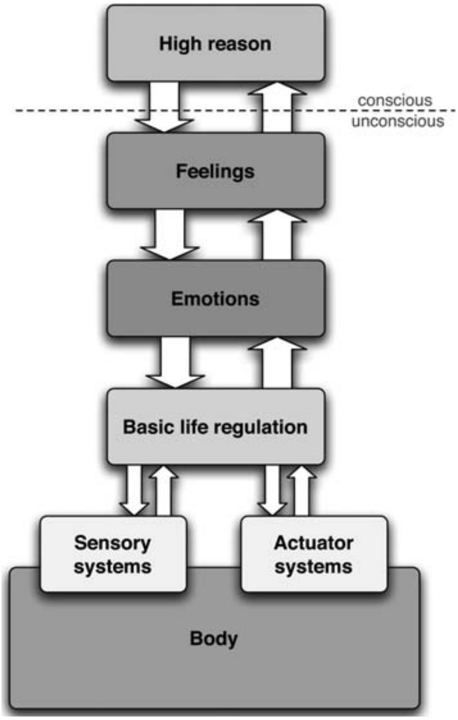

Fig. 4 Damasio proposes a hierarchical process for emotion–raising distinguishing between basic autonomic mechanisms, emotion, feeling and conscious awareness of emotion

3 Assessing Models

Most data concerning emotion comes from experimentation on animals and humans. Data coming from these sources are widely heterogeneous; from single neuron firing patterns, columnar behaviours or activation levels of a whole brain area to hormone concentrations or verbally reported psychological data. This heterogeneity in the level of resolution and abstraction of the data is surely a factor for the difficulty of building up a unified theory of emotion which could address such a broad variety of data. For example, the neurophysiological and hormonal response to fear in mammals is quite well known (Fendt and Fanselow, 1999) together with the behavioural response of rats in experiments of fear conditioning (Hatfield et al. 1996). However, there still remains missing a unified theory covering the gap between the low level, populated with neurons and hormonal mechanisms, and the higher levels of elicited behavioral responses.

Some theoretical models of emotion and associated computational implementations are being explored as a promising tool for integrative understanding of this emotive-cognitive mechanisms. The main value this approach offers is the possibility of having a precise, more rigourous methodology to grasp the core concepts and architecture.

In this sense we can cite Botelho (2001):

> We present a preliminary definition and theory of artificial emotion viewed as a sequential process comprising the appraisal of the agent global state, the generation of an emotion-signal, and an emotion-response. This theory distinguishes cognitive from affective appraisal on an architecture-grounded basis. Affective appraisal is performed by the affective component of the architecture; cognitive appraisal is performed by its cognitive component.

Emotions *affect all levels of operation in a system*, from basic life regulation to conscious, cognitive processes. We use the term *transversal* to indicate this fact. The concrete way in which system operation is affected is specific to each level. Within each of these levels, it is specific to each organ, component and process.

In other words, emotions provide a common *control broadcast* infrastructure which may be used differently by each of the processes in the system. In natural systems, emotions may be conveyed by neural firings and hormones (i.e. the bodily signal broadcasting mechanisms). These mechanisms must be shared by many organs and processes in the system, which will interpret signals according to their purposes and architectures. For instance, a cognitive process may interpret hormonal levels to obtain auxiliary information for making a decision regarding what the system must do next. The same hormones may be interpreted concurrently by other processes in order to detect danger, risk, or a need to obtain food, for example.

This global and multi-level character of emotions explains some distinctions of emotion-relative phenomena present in the literature, such as Damasio's (Damasio, 1999, 2004):

- state of emotion,
- state of feeling an emotion,
- state of a feeling of an emotion made conscious.

One particular way in which emotions are transversal is by broadcasting a summarised picture of the system state to many of its components and processes. This means not only a summary of how its components find themselves, but also a certain sense of affordance of the current scenario relative to the current system situation, processes and objectives.

This is useful to the system in order to adapt to its scenario of operation, mainly for three reasons:

- Emotions are fast, and are available before other more cognitive information.
- Emotions, being to some extent global, contribute to co-ordination and focus of large quantities of system processes and components, which is a factor for preserving system cohesion (López, 2007).
- Emotions can be externalised and hence used for behavioural organisation (co-operation and competition are examples) in multi agent environments (societal behaviour being the clearest example).

3.1 The 'Emotional' Facial Mimicking

This last aspect, that of externalisation of emotional states, has rendered emotional expression one of the main topics of emotion research (Figure 5) (Darwin, 1872; Ekman, 1982).

There have been plenty of efforts made towards the implementation of emotion in machines as inspired by biosystems (Trapl et al. 2002) but in most of the cases they have neglected addressing the core issues and have instead just focused on mimicking shallow, observable manifestations of emotion (e.g. making robot faces *a la* Ekman). But this work, outside the psychological arena, is irrelevant in theoretical and operational terms. All that is expected is some improvement in social capability by facial displaying for human emotions. This is hopeless because the functional value of the display of an emotional state in a social interaction is based on the

Fig. 5 A big amount of research on emotion has been focused on the expression of facial emotion neglecting the inner functional aspects of it

activation in the receptor of the display of a behavioural model of the displayer so as to maximise effectiveness of interaction—it is more a question of better exploiting the human capabilities than of concrete robot competencies.

Mimicking faces is thus useless unless the operational state of the displayer is what is captured in the model going to be activated in the receptor. Clearly this is not the case of human vs. robot architectures (i.e. the mental model of the receptor will be a model of a human whereas the robot is not a human at all from an architectural nor functional point of view). This issue has been widely addressed in the field of human-computer interaction and the mental models community (Gentner and Stevens, 1983). What is important then is the raising of mental states in the receptor (i.e. the activation of mental models) that are relevant for the interaction. This can only be done if emotion is tightly tied to the inner operational mechanisms of the agent displaying the emotion (Conde, 2005).

To conclude with this subject, we shall summarise that facial expression of emotion is an externalisation of an emotional state to help reconfigure a multi-agent organisation taking into account individual agent operational states.

3.2 *What Emotions Are and Are Not: Ideas to Solve the Puzzle*

After this succinct analysis, the puzzle of emotion, from an engineering point of view, can be reduced to three core aspects:

Three questions:

- **What is the function that emotional mechanisms do play?**
- **What is the general form of an emotional mechanism?**
- **What is the best strategy for emotion implementation?**

The analysis of the several models of emotion produce some conclusions regarding *what emotions are not*:

- Emotions are not just sophisticated input handling, i.e. *not just reacting to bears*.
- Emotions are not just sophisticated action generation for social affective behaviour, i.e. *not just showing embarrassment*.
- Emotions are not just mechanisms for re-goaling, i.e. *not just deciding to change from eating to doing sex*.

A deeper analysis abstracting from the biological mechanics into the functional structure renders some conclusions about *how emotions work*:

- Emotions do generate *synthetic compact states* (performing state space reduction) for the effective tuning and use of evolutionary meta-controllers.

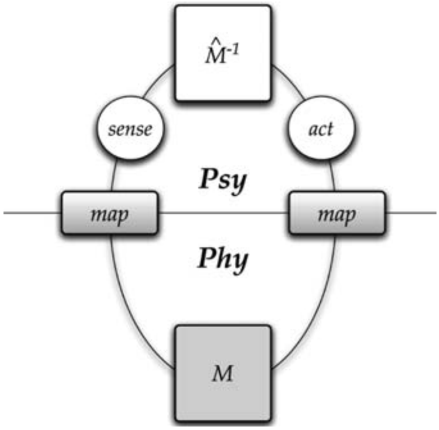

Fig. 6 The mind as model-based controller vision is a central concept in the ASys Framework. It helps bind the dynamics of the mental (Psy) to the dynamics of the physical (Phy). This should not be understood as a dualistic position, it is not. The mind itself is part of the phenomena of the physical (Landauer, 1992) and we just focus on two aspects of this physicality, stressing the informational nature of minds

- Emotions do change the *control structures* at the component functional level (patterns and roles) of subsystems.
- Emotions operate in a global controller configuration approach rendering a *transversal structural feedback* architecture.

In the following section we will develop these three core ideas about emotions in the context of a technical framework intended for the engineering of maximally autonomous systems by applying bioinspired functional concepts.

4 The ASys Framework

The ASLab ASys Project is a long-term research project focused in the development of technology for the construction of autonomous systems. What makes ASys different from other projects in this field is the extremely ambitious objective of addressing *all the domain of autonomy*. We capture this purpose in the motto "engineering any-x autonomous systems". The *ASys Framework* is both a theoretical framework for understanding all the relevant issues and a software-intensive technological framework that enables the technically sound creation of autonomous systems, where autonomy is understood in its broadest sense and not in the severely restricted sense of the term *autonomous intelligent systems* that is usually equated to *mobile reactive robots*.

One of the central topics in the ASys Framework is the pervasive model-based approach. A truly autonomous system will be continuously using models to perform its activity. An ASys system will be built using models of it. An ASys can exploit its own very models for driving its behaviour. Model-based engineering and model-based behaviour then merge into a single phenomenon: *model-based autonomy*. We equate this conceptualisation with *cognition*.

The ASys Framework hence establishes that a system is said to be cognitive if it exploits models of other systems in its interaction with them. Models and knowledge are then equated and the ASys Framework provides a link between the ontological and epistemological aspects of mind.

On the technical side, the ASys Framework follows a principled approach to autonomous system mind construction, the *cognition as model-based behaviour* being the first principle, so as to ground a systematic engineering approach that shall end in rendering machine consciousness (Sanz et al. 2007).

These principles establish guidelines for the systematic, formally grounded development of a real-time control framework based on the control and software principles of the Integrated Control Architecture (Sanz et al. 1999), which will be furtherly discussed at the end of this section. This will render a methodology, a toolset and an execution framework for the engineering of robust autonomous systems based on the implementation of cognitive mechanisms up to the level of consciousness (Sanz et al. 2005).

4.1 Emotion, Consciousness and Control in ASys

The mechanisms of emotion impinge on the behavioural capability of the agent so as to prepare it for future action. This makes emotion a core capability for sophisticated self-management control architecture where outer control loops (emotion) determine the functioning of inner control loops (homeostasis) so as to maximise survivability. Damasio's model on consciousness lays out another control loop atop of these two (see Figure 4) rendering a high-level reasoning capability. Emotions realise meta-controllers.

As was the case with emotions, there are plenty of models of consciousness that try to address the relation of physiology and the three core aspects of consciousness: world-awareness, self-awareness and qualia. We can distinguish as maximally relevant for our work, due to their abstract, general nature, the Global Workspace model of Baars (1997) and the information integration model of Tononi (2004).

The integrated control model of consciousness (Sanz et al. 2007), also part of the ASys Framework, is based on the provision of self-awareness by means of model-based perceptual mechanics.

The ASys perspective on cognition/emotion goes beyond Damasio's approach of putting emotions/feelings as additional layers in hierarchical controllers. Emotion is no longer another layer in the architecture but a transversal mechanism that crosses across all layers. This is indeed a well known fact in the studies of emotion. Emotions do appear from the subconscious plane to the conscious surface, affecting all levels in the cognitive structure, from the physiological up to the cognitive, social, self-conscious level.

This implies (see Figure 7) that emotional mechanics are part of each level of the control hierarchy. The level of focus of the analysis is what determines the

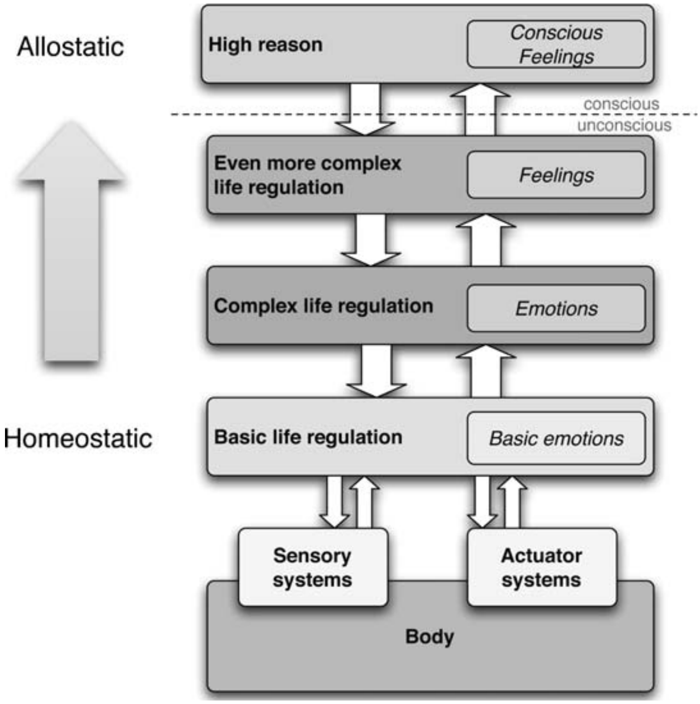

Fig. 7 Damasio's layering of emotion appears as labelling of the transversal emotional mechanisms across a layered architecture for control

labelling used for this mechanism: basic emotion, emotion, feeling, conscious feeling, etc. In the language of information technology we would say that emotion is an *aspect*—in the computer science interpretation (Filman et al. 2004)—of the different systems that constitute the body and mind of an autonomous agent. From a functional perspective we can also observe that the goals pursued by such control structures go from the purely homeostatic mechanisms for life survival to the higher-level, socially-originated allostatic mechanisms for social behaviour. From hunger to embarrassment, emotions do share the meta-control capabilities over basic behavioural structures.

Artificial implementations of emotions are not developed yet to the same degree as the natural. However, large, distributed, fault-tolerant systems include mechanisms which already play a similar role (Aström et al. 2001). Fault detection, damage confinement, error recovery and fault treatment are based on broadcast messages and other mechanisms shared and used by system components in analogous ways to the natural counterparts.

4.2 Emotion Mechanics in ASys

The ASys Framework for autonomous systems is based on an architecture for software-intensive, distributed, real-time control called the Integrated Control Architecture (ICa) (Sanz et al. 1999). This architecture is based on the implementation of patterns of activity across sets of distributed real-time agents. These patterns respond to the needs of the control task that can follow a multilayered, multi-objective control strategy (Alarcon et al. 1994).

The implementation of a controller over ICa renders a collection of interacting software components that realise patterns of activity as sequences of service requests. The software component model is of extreme importance in the implementation of such controllers because it provides a common modelling framework for both the physical components of the system under control, which are but the organs in a biological system and constitute the fleshly infrastructure, and the mental components, which constitute the control superstructure. Figure 8 shows three such components in a simple, layered control structure: an organ, a controller and a meta-controller.

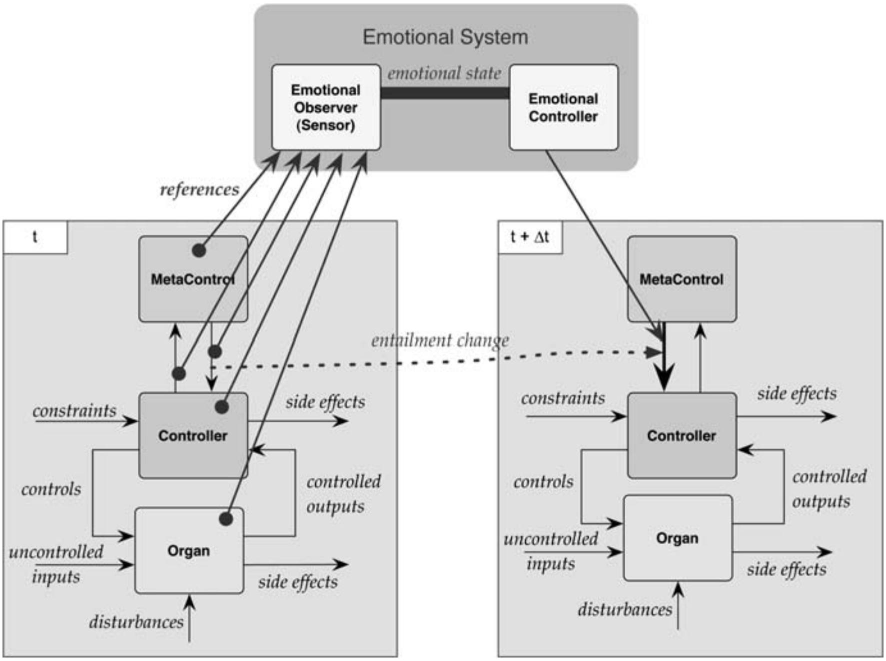

Fig. 8 The figure depicts how an emotional system following these ideas would perform reconfiguration in an modularised control architecture. The emotional system changes the functional organisation to adapt it to new operating conditions

An emotional system incorporated to ICa, according to the previous conceptualisation of emotion in the ASys Framework, would provide the structural mechanisms for control pattern adaptation to the current state of affairs. Examples of this kind of architecture are already available in the realm of control systems, for example the previously mentioned fault-tolerant controllers and sliding mode controllers.

The primary effect of the emotional system is the change in the functional organisation of the control system of the body. In Figure 8 the emotional system changes the functional organisation from time t to $t + \Delta t$, concretely in this example the output of the meta-controller to the controller—e.g. the goal or reference, in control jargon—. Both the emotional observation and control are done in terms of a value system for the agent. This happens in a multi-scale, multilayer organisation that constitutes the integrated global controller of the agent.

Now we can provide the ASys Framework answers to the three core aspects of emotion mentioned before:

Three answers:

- *What is the general form of an emotional mechanism?*
  A self-reorganising meta-controller.

- *What is the function that emotional mechanisms do play?*
  Provide value-centric functional reorganisation.

- *What is the best strategy for emotion implementation?*
  Functional modularisation of control functions and integration over a common infrastructure.

5 Summary

The chapter has reviewed some of the common approaches to emotion understanding, with an special emphasis on Damasio's model of emotion and feeling.

It has also been analysed the extended functional role that emotions can play in complex adaptive controllers and how the different aspects of emotion—triggering, emotional states, bodily effect—are addressed from this perspective.

This understanding has been put in the context of the ASys Framework, a theoretical and technical framework for the implementation of autonomous systems. This framework is based on the construction of modular, component–based control systems following the architectural guidelines of the Integrated Control Architecture (ICa)—a software architecture based on distributed real-time objects.

This framework is being applied to the modelling and understanding of autonomic-emotional-cognitive integration aspects in the rat brain and the implementation of embedded controllers in the context of the IST ICEA project.

Acknowledgements Authors would like to acknowledge the support coming from the European Community's *Seventh Framework Programme FP6/2004-2007* under grant agreement IST 027819 ICEA—*Integrating Cognition, Emotion and Autonomy*—and the Spanish *Plan Nacional de I+D* under grant agreement DPI-2006-11798 C3—*Control Cognitivo Consciente*.

References

Alarcon I, Rodriguez-Marin P, Almeida L, Sanz R, Fontaine L, Gomez P, Alaman X, Nordin P, Bejder H, de Pablo E (1994) Heterogeneous integration architecture for intelligent control systems. Intelligent Systems Engineering 3(3):138–152

Arnold M (1960) Emotions and Personality. Cambridge University Press

Aström K, Albertos P, Blanke M, Isidori A, Schaufelberger W, Sanz R (eds) (2001) Control of Complex Systems. Springer

Baars BJ (1997) In the theatre of consciousness. global workspace theory, a rigorous scientific theory of consciousness. Journal of Consciousness Studies 4:292–309

Bermejo Alonso J (2006) A state of the art on emotion. Tech. Rep. R-2006-002, UPM Autonomous Systems Laboratory, Universidad Politécnica de Madrid

Botelho L (2001) Machinery for artificial emotions. Cybernetics and Systems: An International Journal 32:465–506

Breazeal C (2000) Sociable machines: Expressive social exchange between humans and robots. PhD thesis, Department of Electrical Engineering and Computer Science, MIT

Conde RP (2005) Mapeo facial de emociones sintéticas. Master's thesis, Universidad Politécnica de Madrid

Damasio A (1999) The feeling of what happens: body and emotion in the making of consciousness. Harcourt Press, New York

Damasio A (2004) Emotions and feelings: a neurological perspective. In: Manstead A, Fridja N, Fischer A (eds) Feelings and emotions, Cambridge University Press, chap 4, pp 49–57

Darwin C (ed) (1872) The expression of the emotions in man and animals. D. Appleton and Company

Ekman P (ed) (1982) Emotion in the human face. Cambridge University Press

Fendt M, Fanselow MS (1999) The neuroanatomical and neurochemical basis of conditioned fear. Neuroscience & Biobehavioral Reviews 23(5):743–760

Filman RE, Elrad T, Clarke S, Aksit M (eds) (2004) Aspect Oriented Software Development. Addison Wesley

Frijda N (1987) The emotions. Cambridge University Press

Gentner D, Stevens AL (eds) (1983) Mental models. Lawrence Erlbaum Associates, Hillsdale, NJ

Hatfield T, Han JS, Conley M, Gallagher M, Holland P (1996) Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. J Neurosci 16(16):5256–5265

James W (1884) What is an emotion'. Mind 9:188–205

Landauer R (1992) Information is physical. In: Workshop on Physics and Computation, PhysComp '92., pp 1–4

López I (2007) A framework for perception in autonomous systems. PhD thesis, Departamento de Automática, Universidad Politécnica de Madrid

Ortony A, Clore GL, Collins A (1988) The Cognitive Structure of Emotions. Cambridge University Press

Panksepp J (1998) Affective neuroscience: the foundations of human and animal emotions. Oxford University Press, New York, URL http://www.loc.gov/catdir/enhancements/fy0635/98015955-d.html

Plutchik R, Kellerman H (eds) (1980) Emotion: Theory, research and experience. Volume I: Theories of emotion. Academic Press, New York

Sanz R, Matía F, Puente EA (1999) The ICa approach to intelligent autonomous systems. In: Tzafestas S (ed) Advances in Autonomous Intelligent Systems, Microprocessor-Based and Intelligent Systems Engineering, Kluwer Academic Publishers, Dordretch, NL, chap 4, pp 71–92

Sanz R, Matía F, Galán S (2000) Fridges, elephants and the meaning of autonomy and intelligence. In: IEEE International Symposium on Intelligent Control, ISIC'2000, Patras, Greece

Sanz R, López I, Bermejo-Alonso J, Chinchilla R, Conde R (2005) Self-X: The control within. In: Proceedings of IFAC World Congress 2005

Sanz R, López I, Rodríguez M, Hernández C (2007) Principles for consciousness in integrated cognitive control. Neural Networks 20(9):938–946, DOI http://dx.doi.org/10.1016/j.neunet.2007.09.012

Tononi G (2004) An information integration theory of consciousness. BMC Neuroscience 5:42

Trapl R, Petta P, Payr S (eds) (2002) Emotion in Humans and Artifacts. MIT Press

Ventura R, Pinto-Ferreira C (1999) Emotion–based agents: three approaches to implmentation. In: Velasquez JD (ed) Workshop on Emotion–based agent architectures, Seattle, U.S.A.