Chapter XXIII
Modelling Hardwired Synthetic Emotions: TPR 2.0

Jordi Vallverdú
Universitat Autònoma de Barcelona, Spain

David Casacuberta
Universitat Autònoma de Barcelona, Spain

ABSTRACT

During the previous stage of our research we developed a computer simulation (called ‘The Panic Room’ or, more simply, ‘TPR’) dealing with synthetic emotions. TPR was developed with Python code and led us to interesting results. With TPR, we were merely trying to design an artificial device able to learn from, and interact with, the world by using two basic information types: positive and negative. We were developing the first steps towards an evolutionary machine, defining the key elements involved in the development of complex actions (that is, creating a physical intuitive ontology, from a bottom-up approach). After the successful initial results of TPR, we considered that it would be necessary to develop a new simulation (which we will call “TPR 2.0.”), more complex and with better visualisation characteristics. We have now developed a second version, TPR 2.0., using the programming language Processing, with new improvements such as: a better visual interface, a database which can record and also recall easily the information on all the paths inside the simulation (human and automatically generated ones) and, finally, a small memory capacity which is a next step in the evolution from simple hard-wired activities to self-learning by simple experience.
INTRODUCTION

This is an update of a former project about creating a simulation of an ambient intelligence device which could display some sort of protoemotion adapted to solve a very simple task. In the next section we’ll describe the first version of the project, and then we’ll deal with the changes and evolution in a third second version. But first let us introduce the main ideas that are the backbone of our research.

Bottom Up Approach

AI and robotics have tried intensively to develop intelligent machines over the last 50 years. Meanwhile, two different approaches to research into AI have appeared, which we can summarise as top down and bottom up approaches:

i. **Top Down**: Symbol system hypothesis (Douglas Lenat, Herbert Simon). The top down approach constitutes the classical model. It works with symbol systems, which represent entities in the world. A reasoning engine operates in a domain independent way on the symbols. SHRDLU (Winograd), Cyc (Douglas Lenat) or expert systems are examples of it.

ii. **Bottom Up**: physical grounding hypothesis (situated activity, situated embodiment, connectionism ↔ veritat? No seria connectionism?). On the other hand, the bottom up approach (led by Rodney Brooks), is based on the physical grounding hypothesis. Here, the system is connected to the world via a set of sensors and the engine extracts all its knowledge from these physical sensors. Brooks talks about “intelligence without representation”: complex intelligent systems will emerge as a result of (or o of?) complex interactive and independent machines. (Vallverdú, 2006)

Although we consider that the top-down approach was really successful on several levels (cf. excellent expert systems like the chess master Deep Blue), we consider that the approaches to emotions made from this perspective cannot embrace or reproduce the nature of an emotion. Like Brooks (1991), we consider that intelligence is an emergent property of systems and that in that process, emotions play a fundamental role (Sloman & Croucher, 1981; DeLancey, 2001). In order to achieve an “artificial self” we must not only develop the intelligent characteristics of human beings but also their emotional disposition towards the world. We put the artificial mind back into its (evolutionary) artificial nature.

Protoemotions and Action

As we have published extensively elsewhere (Casacuberta 2000, 2004; Vallverdú 2007), on how emotions play a fundamental role in rational processes and the development of complex behaviour, including decision making (Schwarz 2000). There is a huge body of literature on these ideas which we will not analyse here but which can be consulted (Damasio 1994, Edelman 2000, Denton 2006, Ramachandran 2004).

After describing emotions as alarm systems that activate specific responses (Vallverdú & Casacuberta, 2008a), we considered it necessary to minimise the number of basic emotions and choose two: pain and pleasure, considered as negative and positive inputs, respectively. We called them protoemotions, because they are the two basic regulators of activity. In this sense, we considered synthetic emotions as “an independently embedded (or hard-wired) self-regulating system that reacts to the diverse inputs that the system can collect (internal or external).” (op. cit, 105). From this point of view the cybernetics concept of feedback, as a property of biological entities is added to our conceptual model.

We must also take into account another use of this term, protoemotions, by clinicians who