



Critically Evaluating the Design of an Emotional Agent

Robert Marsh & Lola Cañamero
University of Hertfordshire
HUMAINE – WP 7.1



Purpose

- To present a live critical evaluation of 4 emotion architectures that are currently being proposed at UH for development as part of HUMAINE WP 7.1
- To argue for the necessity of such critical evaluations throughout the development process of all emotion architectures.



The Problem

- HUMAINE WP7 concerns itself with 'Emotion in Cognition and Action'.
- As such, the systems that we develop must be defensible in terms of the definitions of Emotion, Cognition and Action that we ultimately feel are correct.
- The problem is how to guarantee this to be the case.



The Solution

- The solution is to develop metrics, measures, and general criteria that are rooted in the definitions of Emotion, Cognition, and Action that we personally adhere to.
- This critical framework must then be used to assess our systems throughout the development process.



The Payoff

- This approach delivers an enhanced probability that the systems we develop will be genuinely Emotional, Cognitive and deeply rooted in Action, in addition to being technically accomplished and of genuine scientific or engineering interest.



Live Evaluation – Choices

- At UH we have developed four potential emotion architectures for WP 7.1.
- One of these systems must be put forward for further development. Which one should that be? Which system shows the most promise in terms of the project deliverables?
- Stage 1; Develop the critical framework.
- Stage 2; Evaluate the architectures.
- Stage 3; Make the choice.

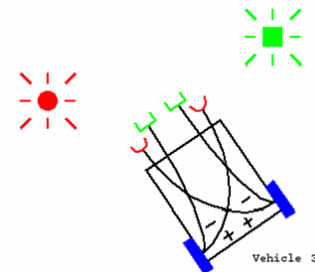
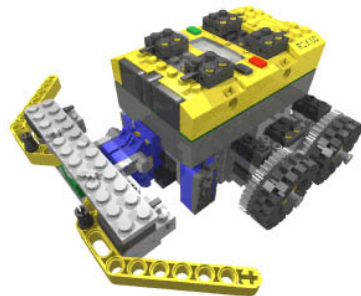


Live Evaluation – Robotic Scenario

- It is our belief that Emotion plays a central role in behaviour control. As such, the 'problem of behaviour control' is the experimental test bed that we use to examine our Emotional agents.
- The problem of behaviour control is the problem any agent faces when it has to decide which behaviour to be performing at any given time.
- This is a problem because survival and success in a complex and dynamic environment depends critically upon the agent engaging in the right behaviours at the right times.

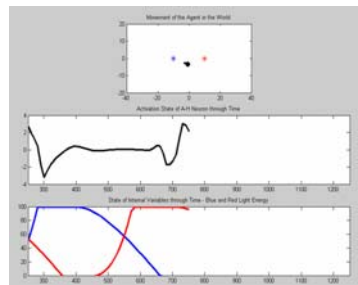
Live Evaluation – The Architectures

- Motivation Based Architectures
- Avila-Garcia & Cañamero, (2001, 2002, 2003), designed and tested various control architectures that used competition and arbitration between various motivational states in order to decide upon a final behaviour for a robotic agent.
- Neuromodulatory Braitenberg Vehicles
- Considering neuromodulation to be an integral part of emotional and motivational systems, French & Cañamero (2005) investigated the performance of simple Braitenberg vehicle architectures under neuromodulatory control.



Live Evaluation – The Architectures

- Hybrid Dynamical System Model of Action-Selection
- Inspired by the work of Avila-Garcia & Cañamero, (2001, 2002, 2003), Marsh and Cañamero are currently investigating an emotional and motivation based behaviour control architecture that has it's roots in a dynamical systems viewpoint of agent-environment interaction.
- Evolved Neuromodulatory Action-Selection
- The work of French & Cañamero (2005), and Seth (1998) both suggest that a large range of psychologically relevant phenomena can be supported by a minimal neural framework under neuromodulatory control.





Live Evaluation – Core Criteria

- Emotional.
- Is the system capable of supporting emotional content?
- Following Fellous (2004), my working definition of emotion is as follows;

Emotions are to be understood as 'dynamical patterns of neuromodulations', whose function is to achieve a 'multi-level communication of simplified but high impact information' throughout the nervous system / control system of an agent, the aim of which is to facilitate adaptive behaviour and therefore maintain the viability of the agent.



Live Evaluation – Core Criteria

- Cognitive Component.
- Does the system have a cognitive component to it?
- Following the definition provided by Pfeifer & Scheier (1999), my working definition of cognition is;

'A descriptive term for the large class of so-called higher-level processes, that is, processes not directly driven by the sensory and motor systems'. Examples include perceiving, recognizing, conceiving, judging, reasoning, and imagining.



Live Evaluation – Core Criteria

- Grounded Emotional Content.
- Does the emotional content of the agent have a semantic interpretation or meaning that is intrinsic to the system, rather than simply parasitic on the meanings in the designer's head?
- Following Harnad (1990), this criteria examines whether the emotions of the agent are grounded (and therefore have intrinsic meaning for the agent) or whether the emotions have been imposed upon the system by the designer (therefore giving the emotional content extrinsic meaning; meaningful for the designer but not for the agent itself).

Live Evaluation – Core Criteria

- Distinct Actions.
- Does the system select distinct actions, or does the agent's behaviour emerge from a structured agent-environment interaction? (WP7 considers the role of emotion in low-level cognition and action)
- Following Pfeifer & Scheier (1999), an 'action' is something that is synonymous with a behaviour; More specifically, behaviours that are goal-oriented (in other words, behaviours selected on the basis of some anticipated result). Examples of actions an agent might perform include searching for resources, consuming resources, and avoiding obstacles / other threats to its safety.

Live Evaluation – Results

Architecture / Dimension	Emotional	Motivation-Based	Cognitive Component	Dynamical Systems Component	Neuromodulation	Maintenance of viability	Grounded Emotional Content	Distinct Actions
Motivation Based Architecture	No	Yes	Yes	No	Minimal (not a robust mechanism)	Yes	No emotional content	Yes
Neuromodulatory Braitenberg Vehicles	Yes	No	No	Minimal	Yes	Yes	Emotional content has extrinsic meaning	No
Hybrid Dynamical Systems Model of Action Selection	Yes	Arguably	Yes	Yes	Yes	Yes	Emotional content has extrinsic meaning	Yes
Evolved Neuromodulatory Action Selection	Yes	No.	No	Minimal	Yes	Yes	Emotional content has intrinsic meaning	No

Live Evaluation - Conclusions

- Full and detailed evaluation is available at;

<http://emotion-research.net/ws/wp7/>

- It must be stressed that the evaluation just performed is not based on a supposedly definitive critical framework. The critical framework applicable to YOUR system all depends upon the particular set of working definitions that YOU employ for the concepts of Emotion, Cognition and Action. There is no definitive framework.
- However, *that* one should critically examine one's architectures using such a critical framework is *vital* throughout the development process.
- It is arguably the only way to guarantee that the systems that we end up with are thoroughly rooted in the definitions of Emotion, Cognition and Action that we personally adhere to.