Engineering Emotionally Intelligent Agents

Penny Baillie-de Byl
University of Southern Queensland, Australia

Mark Toleman
University of Southern Queensland, Australia

INTRODUCTION: EVOLUTION OF EMOTIONAL AGENTS

Traditionally, philosophers have defined emotions to be interruptions to otherwise logical states of being (Smith & Kirby, 2000). The recent resurgence of research in the emotional realm in both psychology and cognitive science agrees with the view developed by the late Charles Darwin who, in the late 1800s, conceived that emotions play an important part in our cognition and serve to provide us with the mechanisms for adaptive behaviour in a dynamically complex world (Smith & Ellsworth, 1985).

One relevant line of research is the realm of affective computing. This relatively new domain examines the effect that emotions have on human intelligence and endeavors to use this to further enhance the field of artificial intelligence (AI). How the concept of emotions might heighten the intelligent functioning of artificial beings is still unclear, but through the variety of research programs that currently exist, areas that might benefit are being identified. In this domain, much work is being done to develop artificial intelligences capable of identifying, processing and synthesizing emotions. Picard (1997) suggests that emotions are an integral part and a natural progression for AI. She further states that: "...the inability of today's computers to recognize, express, and have emotions, severely limits their ability to act intelligently and interact naturally with us".

Emotional decision making provides a good solution for computer systems that face the problem of enumerating and evaluating multitudinous choices within an acceptable time frame. One application of AI that is benefiting by integrating emotional decision making mechanisms is that of intelligent agents. The word *agent* is used within the AI domain to refer to a number of different applications. The most popular use of the term pertains to an autonomous artificial being that has the ability to interact intelligently within a temporally dynamic environment.

BACKGROUND

Much of the current research for achieving these types of results with artificial agents is based on appraisal theories. Models such as the Affective Reasoner (Elliot, 1992, 1998), the Oz project (Bates, Loyall & Reilly, 1992), PETEII (El-Nasr, 1998), Silas (Blumberg, Todd & Maes, 1996), Yuppy (Velasquez, 1999), extensions of the INES intelligent tutoring system (Hospers, Kroezen, op den Akker, Nijholt, & Keylen, 2003; Heylen, Nijholt, op den Akker, & Vissers, 2003) and systems based on PMFserv (Silverman, 2001; Silverman, Cornwell, & O'Brien, 2003) are just a few examples of how appraisals (cognitive and non-cognitive) are used to generate emotional states and behaviors within an artificial intelligence.

Cognitive appraisal theory and the OCC Model developed by Ortony, Clore and Collins (1988), although not originally intended for synthetic generation of emotions, is a basis for several of these projects. Table 1 displays the values used in the OCC model and subsequent AI applications to appraise an event and determine an appropriate emotion. The table is a subset of the OCC appraisals, displaying eight of the possible 22 emotional states mentioned in the theory.

Heylen et al. (2003) report a proposed extension to INES to incorporate social and emotional intelligence skills of human tutors using emotional axes defined in Kort, Reily & Picard (2001). Implementation is still at the conceptual stage. Baillie (2002) developed an emotionally motivated artificial intelligence (EMAI) architecture that has the capacity for decision making influenced by simulated human emotional intelligence. The architecture consisted of several major processing and knowledge representation areas working together in a complex network of information gathering, manipulation and update. A focal point of the architecture was an "affective space" which acted as an emotional filter influencing an EMAI agent's perception of its beliefs about the environment and, as a consequence, how it behaved. EMAI agents have been

Table 1. A subset of the OCC cognitive structure of emotion

Consequential Disposition	Focus Object (self or other)	Consequences for Focus Object	Emotion
pleased	other	desirable	Happy For
pleased	other	undesirable	Pity
displeased	other	desirable	Resentment
displeased	other	undesirable	Gloating
pleased	self	relevant	Hope
pleased	self	irrelevant	*Joy*
displeased	self	relevant	Fear
displeased	self	irrelevant	Distress

implemented and evaluated for reasonableness in their simulation of emotions. Silverman, Cornwell & O'Brien (2003) report on progress towards an open agent architecture (PMFserv) to allow exploration of a range of human attributes, including emotions. The architecture has many similar structural components providing similar functionality to the EMAI architecture, but its application domains and integration efforts hinder experimentation and simulated scenarios have not been evaluated.

AN EMOTIONALLY INTELLIGENT AGENT ARCHITECTURE

The EMAI architecture (shown in Figure 1) is one example of recent emotional agent architectures. It is a complex set of mechanisms that process emotional concepts for their use in affective decision-making and reasoning. There are two types of emotion mechanisms integrated in the EMAI architecture. The first mechanism emulates fast primary emotions (Koestler, 1967) otherwise known as motivational drives. These drives can be classified according to their source, pleasure rating and strength. In an EMAI agent, these drives are used to initiate behaviour in the agent. They can include concepts such as hunger, fatigue or arousal. The strength of the drives is temporally dynamic and at particular threshold levels the agent will set goals, that when successfully achieved, will pacify the drives. For example, over time the strength of the hunger drive will increase. At a certain point, the agent will become so hungry that it will set appropriate goals to ensure it obtains food. On the consuming of food, the strength of the agent's hunger drive will decrease.

The agent's goals are generated by a motivational drive generator which consists of drive mechanisms and a set of internal state registers representing the primary emotions. Each register is represented by a single gauge and stores the value for a particular drive, for example hunger. The number of internal state registers implemented depends on the application for which the EMAI agent is being used.

The second type of emotion implemented in the EMAI architecture is secondary emotion. This category of emotion refers to the resultant mental (and in turn physical) states generated by attempts to satisfy the goals. These emotions include feelings such as happiness, anger, sorrow, guilt and boredom. Secondary emotions are represented in the EMAI architecture as values in an affective space.

The affective space is a six-dimensional space defined by six appraisal dimensions. The affective space, based on the psychological model of Smith and Ellsworth (1985), defines 15 emotions (happiness, sadness, anger, boredom, challenge, hope, fear, interest, contempt, disgust, frustration, surprise, pride, shame and guilt) with respect to the six dimensions of pleasantness, responsibility, effort, certainty, attention and control.

In addition to representing the agent's emotional state, the agent uses the affective space to associate emotions with all stimuli both internal to the agent (as internal sensory data) and within its environment (as external sensory data). The stimuli are perceived by the agent as part of an event. An event is a behavioral episode executed by the agent. Stimuli can be any tangible element in the agent's environment including the actions being performed, smells, objects, other agents, the time of day or even the weather.

Figure 1. Summary illustration of the EMAI architecture