Improving Adaptiveness in Autonomous Characters

Mei Yii Lim[1], João Dias[2], Ruth Aylett[1], and Ana Paiva[2]

[1] School of Mathematical and Computer Sciences,
Heriot Watt University,
Edinburgh, EH14 4AS, Scotland
{myl, ruth}@macs.hw.ac.uk
[2] INESC-ID, IST, Taguspark,
Av. Prof. Dr. Cavaco Silva,
2744-016 Porto Salvo, Portugal
{joao.dias, ana.paiva}@gaips.inesc-id.pt

Abstract. Much research has been carried out to build emotion regulation models for autonomous agents that can create suspension of disbelief in human audiences or users. However, most models up-to-date concentrate either on the physiological aspect or the cognitive aspect of emotion. In this paper, an architecture to balance the Physiological vs Cognitive dimensions for creation of life-like autonomous agents is proposed. The resulting architecture will be employed in ORIENT which is part of the EU-FP6 project eCircus[3]. An explanation of the existing architecture, FAtiMA focusing on its benefits and flaws is provided. This is followed by a description of the proposed architecture that combines FAtiMA and the PSI motivational system. Some inspiring work is also reviewed. Finally, a conclusion and directions for future work are given.

1 Introduction

The population of autonomous characters in games, interactive systems, and virtual world is rapidly increasing. The survival of an autonomous character requires that its systems produce actions that adapt to its environmental niche. At the same time, the character must appear to be able to 'think', have desires, motivations and goals of its own. A truely autonomous character will be able to react to unanticipated situations and perform life-like improvisational actions. This character will need a human-like regulation system that integrates motivation, emotion and cognition to generate behavioural alternatives. Damasio [1] proposes the existence of a body-mind loop in emotional situations and provides neurological support for the idea that there is no 'pure reason' in a healthy human brain. Furthermore, embodied cognition theory suggests that cognitive processes involve perceptual, somatovisceral, and motoric reexperiencing of the relevant emotion in one's self [2]. Supporting these views, we propose an emotion model that includes a body-mind link - the link between physiological processes and cognitive processes for effective action regulation in autonomous characters.

[3] http://www.e-circus.org/

2 ORIENT

ORIENT (Overcoming Refugee Integration with Empathic Novel Technology) aims at creating an innovative architecture to enable educational role-play for social and emotional learning in virtual environments. Its focus is on evoking inter-cultural empathy with the virtual characters through conflict resolution and narrative interaction where the users have to cooperate with the alien inhabitants to save their planet from an imminent catastrophe. Each character must be able to establish a social relationship with other characters and users to ensure successful collaboration. Subtle differences across cultures may result in varying emotional reactions that can create a challenge for effective social communication [3]. Hence, ORIENT characters must be able to recognise cultural differences and use this information to adapt to other cultures dynamically. The ability to empathise - being able to detect the internal states of others and to share their experience is vital to engage the characters in long-term relationships. Both cognitive [4] and affective [5] empathy are relevant since enhancement of integration in a group of culture relies both on the understanding of internal states of the persons involved and their affective engagement. Additionally, former experience is crucial in maintaining long-term social relationships, which means the existence of autobiographic memory [6] is inevitable. By being able to retrieve previous experience from its autobiographic memory, a character will be able to know how to react sensibly to a similar future situation. In short, ORIENT characters have to be autonomous agents with autobiographical memory, individual personalities, show empathy, adaptive and improvisational capabilities.

3 Architectures

3.1 FAtiMA

The ORIENT software is being built upon the FearNot! Affective Mind Architecture (FAtiMA) [7]. FAtiMA is an extension of the BDI (Beliefs, Desires, Intentions) deliberative architecture [8] in that it incorporates a reactive component mainly responsible for emotional expressivity and it employs the OCC [9] emotional influences on the agent's decision making processes. The reactive appraisal process matches events with a set of predefined emotional reaction rules while the deliberative appraisal layer generates emotions by looking at the state of current intentions, more concretely whether an intention was achieved or failed, or the likelihood of success or failure. After the appraisal phase, both reactive and deliberative components perform practical reasoning. The reactive layer uses simple and fast action rules that trigger action tendencies while the deliberative layer uses the strength of emotional appraisal that relies on importance of success and failure of goals for intention selection. The means-ends reasoning phase is then carried out by a continuous planner [10] that is capable of partial order planning and includes emotion-focused coping [11].

The advantage of using the OCC model for ORIENT characters is that empathy can be modelled easily because it is the appraisals of events regarding the consequences for others. It is - as far as we know - the only model that provides a formal description of non-parallel affective empathic outcomes. Moreover, since the OCC model includes emotions that concern behavioural standards and social relationships based on like/dislike, praiseworthiness and desirability for others, it will allow appraisal processes that take into consideration cultural and social aspects, two important requirements for ORIENT characters. However, empathic processes can have more emotional outcomes than those described in OCC: happy-for, resentment, gloating and pity. In reality, an individual may feel sad just by perceiving another sad individual. By contrast in FatiMA, an agent experiences empathy only if it is the direct object of an event, leading to a limited psychological plausibility. Moreover, FAtiMA does not take physiological aspect of emotion into account. The character's goals, emotional reactions, actions and effects, and action tendencies were scripted. As a result, the agents do not learn from experience, which is a common problem of BDI agents.

3.2 PSI

PSI [12] is a psychologically-founded theory that incorporates all basic components of human action regulation such as perception, motivation, cognition, memory, learning and emotions in one model of the human psyche. It allows for modelling autonomous agents that adapt their internal representations to a dynamic environment. PSI agents derive their goals from a set of basic drives that guide their actions. These drives include: existence-preserving needs; species-preserving need; need for affiliation; need for certainty and need for competence. A deviation from a set point constitutes the strength of each need. Needs can emerge depending on activities of the agent or grow over time. To be able to produce actions that are able to satisfy needs in a certain situation, the agent builds up intentions that are stored in memory and are - when selected - the basis of a plan. An intention is selected based on strength of activated needs, success probability and urgency.

Once an intention is selected, three levels of goal-oriented action execution can be distinguished. First, the agent tries to recall an automatic, highly ritualised reaction to handle the intention. If this is not possible, an action sequence may be constructed (planning). If planning also fails, particularly when the agent is in a completely new and unknown environment, it acts according to the principle of trial and error. While doing this, the PSI agent learns: after having experienced successful operations, the corresponding relations are consolidated, serving as indicators for the success probability of satisfying a specific need. Based on the knowledge stored in memory, abstractions of objects or events can be built. Moreover, PSI agents forget content with time and lack of use.

Emotions within the PSI theory are conceptualised as specific modulations of cognitive and motivational processes. These modulations are realised by so called emotional parameters including: *arousal* which is the preparedness for perception and reaction; *resolution level* that determines the accuracy of cognitive processes; and *selection threshold* that prevents oscillation of behaviour by giving the current intention priority. Different combinations of parameter values lead to different physiological changes that resemble emotional experiences in biological agents. Hence, a PSI agent does not require any executive structure that conducts behaviour, rather, processes are self-regulatory and parallel driven by needs, and rely on memory as a central basis for coordination. The motivational system serves as a quick adaptation mechanism of the agent to a specific situation and may lead to a change of belief about another agent as shown in [13], important for conflict resolution among ORIENT characters. Thus, PSI permits more flexibility in the characters' behaviour that FAtiMA lacks. Unfortunately, this also means an effective control over the agents' expected behaviour is missing, a limitation for applications where agents need to behave in certain ways, such as in ORIENT where the characters have a common goal to achieve.

4 FAtiMA-PSI: A Body-mind Architecture

We have seen that despite having several advantages over FAtiMA, the PSI model suffers from a lack of control. Thus, the ideal would be to integrate key components of both architectures to build a body-mind architecture where goals are originated from drives, but at the same time gives authors some control over the agents' learning and expected behaviour. The rationale is to get a system between PSI and FAtiMA in the Physiological vs Cognitive dimension. In the new architecture shown in Figure 1, goals are driven by needs. Five basic drives from PSI are modeled: Energy, Integrity, Affiliation, Certainty and Competence. Energy represents an overall need to preserve the existence of the agent (food + water). Integrity represents well being, i.e. the agent avoids pain or physical damage while affiliation is useful for empathic processes and social relationships. On the other hand, certainty and competence influence cognitive processes.

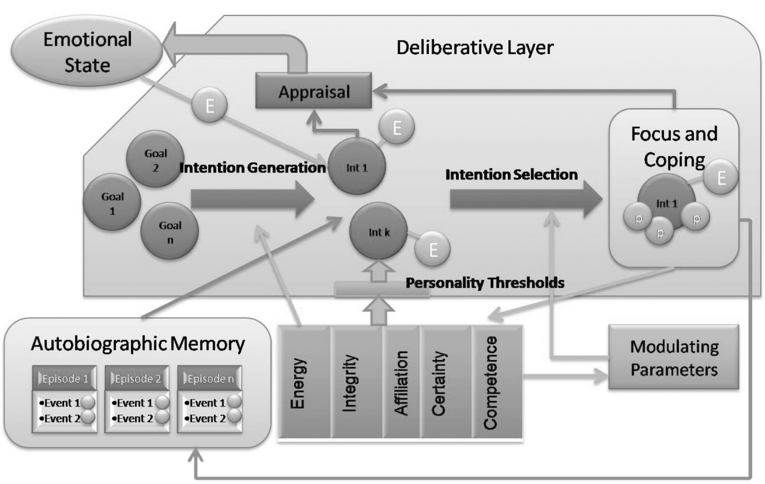

Fig. 1. FAtiMA-PSI architecture

Each need has value ranging from 0 to 10 where 0 means complete deprivation while 10 means complete satisfaction. A weight ranging from 0 to 1 underlines its importance to an agent. In order to function properly, an agent has to reduce a need's deviation from a fixed threshold as much as possible at all time. The strength of a need (*Strength(d)*) depends on its current strength plus the amount of deviation from the set point and the specific weight of the need. By assigning different weights for different needs to different agents, characters with different personalities can be produced, fulfilling one of the requirements of ORIENT. For example, if agent A is a friendly character, affiliation would be an important factor in its social relations, say weight 0.7 while a hostile agent B would have a low importance for affiliation, say weight 0.3. Now, if both agents have a current affiliation value of 2 and if the deviation from set point is 4, agent A's need for affiliation would be 4.8 while agent B's need for affiliation would be 3.2 based on Equation 1. This means that agent A will work harder to satisfy its need for affiliation than agent B.

$$Strength(d) = Strength(d) + (Deviation(d) * Weight(d)) \tag{1}$$

The inclusion of needs requires a change to FAtiMA's existing goal structure. Needs are also affected by events taking place in the environment and actions the agent performs. Since each agent has a different personality, the effect of an event may differ from one agent to another, which in turn affects their emotional and behavioural responses. In the new architecture, each goal will contain information about expected contributions of the goal to energy, integrity and affiliation needs, that is, how much the needs may be deviated or satisfied if the goal is performed. Likewise, the existing structure of events in FAtiMA has to be extended to include its contributions on needs. As for certainty and competence, no explicit specification of contributions is necessary because they are cognitive needs and their values can be calculated automatically as described below.

Whenever an expected event fails to turn up or an unknown object appears, the agent's certainty drops. Certainty is achieved by exploration of new strategies (trial and error), which leads to the construction of more complete hypotheses. If trial and error is too dangerous, developments in the environment are observed in order to collect more information. Please note that the character does not learn by forming new goals because this will lead to a lack of control on its behaviour. Instead, it learns by trying out different actions from a pre-specified set of actions and remembering which actions helped it to tackle a situation best. This information is stored in its autobiographic memory and serves as an indicator to the success probability of satisfying a specific need in future.

Competence represents the efficiency of an agent in reaching its goals and fulfilling its demands. Success increases competence while failure decreases it. The agent's autobiographic memory provides a history of previous interactions, which records the agent's experience in a specific task useful for calculation of goal competence (Equation 2). Since there is no distinction in competence in terms of achieving an important goal and a less important one, one can assume that all goals have the same contribution to the success rate. If the agent cannot remember previous activations of the goal, then it ignores the goal competence and increases the goal's contribution to certainty. The autobiographic memory also stores information about the agent's overall performance useful for calculation of overall competence (Equation 3). The expected competence (Equation 4) of the agent will then be a sum of its overall competence and its competence in performing a current goal. A low competence indicates that the agent should avoid taking risks and choose options that have worked well in the past. A high competence means that the agent can actively seek difficulties by experimenting new courses of action less likely to succeed. During this learning process, the agent also remembers a specific emotional expression of another agent in a certain situation. It continuously updates this information and changes its belief about the agent enabling it to be engaged in empathic interaction in future.

$$Comp(goal) = NoOfSuccess(goal)/NoOfTries(goal) \tag{2}$$

$$OverallComp = NoOfSuccess/NoOfGoalsPerformed \tag{3}$$

$$ExpComp(goal) = OverallComp + Comp(goal) \tag{4}$$

During the start of an interaction, each agent will have a set of initial values for needs. Based on the level of its current needs, the agent generates intentions, that is, it activates goal(s) that are relevant to the perceived circumstances. A need may have several goals that satisfy it (e.g. I can gain energy by eating, or by resting) and a goal can also affect more than one need (e.g. eating food offered by another agent satisfies the need for energy as well as affiliation). So, when determining a goal's strength (Equation 5), all drives that it satisfies are taken into account.

$$Strength(goal) = \sum Strength(d) \tag{5}$$

In terms of a particular need, the more a goal reduces its deviation, the more important is the goal (e.g. eating a full carbohydrate meal when you're starving satisfies you better than eating a vegetarian salad). By looking at the contribution of the goal to overall needs and to a particular need, goals that satisfy the same need can be compared so that success rate in tackling the current circumstances can be maximised. So, the utility value of a goal can be determined taking into consideration overall goal strength on needs, contribution of the goal to a particular need (*ExpCont(goal, d)*) and the expected competence of the agent as shown in Equation 6.

$$EU(goal) = ExpComp(goal) * Strength(goal) * ExpCont(goal, d) \tag{6}$$

As in PSI, needs generate modulating parameters - *arousal, resolution level* and *selection threshold* that enable ORIENT characters to adapt their behaviour dynamically to different interaction circumstances. There may be more than one intention that is activated at any time instance. One of these intentions will be selected for execution based on the *selection threshold* value. After an intention is selected, the agent proceeds to generate plan(s) to achieve it. Emotions emerge as each event affects the character's needs level and hence modulates its planning behaviour. The level of deliberation that the character allocates to actions selection will be proportional to its *resolution level*. For example, if an event leads to a drop in the character's certainty, then its *arousal* level increases causing a decrease in the *resolution level*. In such situation, quick reaction is required hence forbidding time consuming search. The character will concentrate on the task to recover the deviated need(s) and hence may choose to carry out the first action that it found feasible. This physiological changes and behaviour may be diagnosed as anxiety.

5 Related Work

Some examples of existing physiological architectures are those by Cañamero [14] and Velásquez's [15]. These architectures are useful for developing agents that have only existential needs but are insuffcient for controlling autonomous agents where intellectual needs are more important. Another problem of these architectures is that all behaviours are hard-coded. On the other hand, the BDI architecture [8] is the core of deliberative agent architecture. The ways BDI agents take their decisions, and the reason why they discard some options to focus on others, are questions that stretch well beyond artificial intelligence and nurture endless debates in philosophy and psychology. Furthermore, BDI agents do not learn from errors and experience. These problems are associated with the BDI architecture itself and not from a particular instantiation. Fortunately, these questions are addressed by the FAtiMA-PSI architecture where intentions are selected based on strength of activated needs and success probability. Additionally, the motivational system will provide ORIENT characters with a basis for selective attention, critical for learning and memory processes. The resulting agents learn through trial and error, allowing more efficient adaptation and empathic engagement in different social circumstances.

6 Conclusion and Future Work

This paper proposes a new emotion model that balances Physiological vs Cognitive dimensions to create autonomous characters that are biologically plausible and able to perform life-like improvisational actions. Combining FAtiMA and PSI, the problems of psychological plausibility and control are addressed, neither of which can be solved by either architecture alone. Cultural and social aspects of interaction can be modelled using FAtiMA while PSI provides an adaptive mechanism for action regulation and learning, fulfilling the requirements of ORIENT characters. This model also addresses the ambiguity of decision making process in BDI architecture in general. Currently, the motivational system has been integrated into FAtiMA and the next step is to apply the modulating parameters in the deliberative processes such as intention selection and planning. Further effort will also be allocated to include the cultural and social aspects into the architecture. Besides using the information in autobiographic memory solely to determine the need for certainty and competence, it would be desirable to utilise the information to guide the future actions of the characters.

Acknowledgements

This paper is supported by the eCIRCUS (Contract no. IST-4-027656-STP) project carried out with the provision of the European Community in the Framework VI Programme and by a scholarship (SFRH BD/19481/2004) granted by the Fundação para a Ciência e a Technologia.

References

[1] Damasio, A.: Descartes' Error: Emotion, Reason and the Human Brain. Gosset/Putnam Press, New York (1994)

[2] Niedenthal, P.M.: Embodying emotion. Science (316) (2007) 1002–1005

[3] Elfenbein, H.A., Ambady, N.: Universals and cultural differences in recognizing emotions of a different cultural group. Current Directions in Psychological Science **5**(12) (2003) 159–164

[4] Hogan, R.: Development of an empathy scale. Journal of Consulting and Clinical Psychology (35) (1977) 307–316

[5] Hoffman, M.L.: Empathy, its development and prosocial implications. Nebraska Symposium on Motivation **25** (1977) 169–217

[6] Ho, W.C., Dautenhahn, K., Nehaniv, C.L.: Computational memory architectures for autobiographic agents interacting in a complex virtual environment: A working model. Connection Science **20**(1) (2008) 21–65

[7] Dias, J., Paiva, A.: Feeling and reasoning: A computational model for emotional agents. In: 12th Portuguese Conference on Artificial Intelligence (EPIA 2005), Portugal, Springer (2005) 127–140

[8] Bratman, M.E.: Intention, Plans and Practical Reasoning. Harvard University Press, Cambridge, Massachusetts (1987)

[9] Ortony, A., Clore, G., Collins, A.: The cognitive structure of emotions. Cambridge University Press, Cambridge, UK (1988)

[10] Aylett, R., Dias, J., Paiva, A.: An affectively driven planner for synthetic characters. In: International Conference on Automated Planning and Scheduling (ICAPS2006), UK (2006)

[11] Marsella, S., Johnson, B., LaBore, C.: Interactive pedagogical drama. In: Fourth International Conference on Autonomous Agents (AAMAS), Bologna, Italy, ACM Press (2002) 301–308

[12] Dörner, D.: The mathematics of emotions. In Frank Detje, D.D., Schaub, H., eds.: Proceedings of the Fifth International Conference on Cognitive Modeling, Bamberg, Germany (Apr, 10–12 2003) 75–79

[13] Lim, M.Y.: Emotions, Behaviour and Belief Regulation in An Intelligent Guide with Attitude. PhD thesis, School of Mathematical and Computer Sciences, Heriot-Watt University, Ediburgh, Edinburgh (2007)

[14] Cañamero, D.: A hormonal model of emotions for behavior control. In: VUB AI-Lab Memo 97-06, Vrije Universiteit Brussel, Belgium (1997)

[15] Velásquez, J.D.: Modeling emotions and other motivations in synthetic agents. In: Proceeding AAAI 97, AAAI Press and The MIT Press (1997) 10–15