If I were you - Double appraisal in affective agents (Short Paper)

Ruth Aylett
MACS
Heriot-Watt University
Edinburgh, UK
+44 (0)1314514189
ruth@macs.hw.ac.uk

Sandy Louchart
MACS
Heriot-Watt University
Edinburgh, UK
+44 (0)1314513424
sandy@macs.hw.ac.uk

ABSTRACT
We report the implementation and evaluation of a Simulation Theory (ST) approach to the Theory of Mind in intelligent graphical agents driven by an affective agent architecture FAtiMA. The existing cognitive appraisal mechanism is adapted to produce a second appraisal cycle, a double appraisal, in order to evaluate the emotional impact of possible actions. The action with the greatest emotional impact is selected as a means of producing more interesting dramatic actions. A variant in which the actual minds of characters present are used is also implemented and evaluated. Results show that these mechanisms do produce more interesting stories.

Categories and Subject Descriptors
I.2.11 **[Distributed Artificial Intelligence Language]:** Intelligent agents

General Terms
Experimentation

Keywords
Theory of Mind, Cognitive Appraisal, Emergent Narrative.

1.INTRODUCTION
The work reported in this paper arose out of a desire to make autonomous graphical characters behave more like actors in real-world dramatic performance in order to support a generative approach to interactive narrative known as emergent narrative (EN)[1]. The EN approach postulates that rather than pre-scripting interactive narrative or using a top-down structural theory, one can generate interesting interactive narrative experiences bottom-up through interaction between intelligent graphical characters.

It is cleat in the real world that narrative does not spring into being just because people interact with each other. However, a significant difference between ordinary life and improvisatory drama is that in the latter, actors do not merely act in role, they also try to produce dramatically interesting actions. This raises the question of how one can define 'dramatically interesting'. The work reported here pursued the idea that a surrogate for dramatic interest might be the emotional impact of an action on the other characters present in the scene. So how would an autonomous agent assess this?

Most successful human interaction relies on correctly attributing beliefs, desires, goals and percepts to others using what are collectively known as Theory of Mind (ToM) skills [14]. These abilities involve the awareness that other people have different knowledge, beliefs and goals than one's own and have been extensively studied in developmental psychology. ToM skills had been thought to involve the explicit modelling of the mental states of others, extending an approach from other areas of cognitive science that supposed an internal knowledge representation – "a body of rules or principles or propositions" [13] – that would form the underlying theory for various human capabilities. This cognitivist approach has been widely applied in the agent research community, for example in relation to predicting the behaviour of other agents [5]. However in the later 1980s and early 90s, this approach was challenged by a number of researchers in philosophy [8] and psychology [13], who argued for a process-based approach to ToM skills. From this perspective, the mental states of others are captured by adopting their perspective: by tracking or matching their states with resonant states of one's own, *simulating* their mental processes rather than representing them. In the more recent period, the discovery of mirror neurons has lent some neuro-physiological support to this theory [7]. Here we apply Simulation Theory to cognitive appraisal-driven agents.

A Simulation Theory approach is attractive because of its parsimony. Rather than requiring an additional apparatus for modelling other minds, it suggests that the agent's own mind can be reused. In the work discussed here, characters had been implemented using a complex affective architecture., FatiMA [6], in which characters already assessed the emotional impact of events in the world around them as part of the process of deciding on their own actions.

Cognitive appraisal [12], the mechanism used in FatiMA, was anyway originally intended precisely for reasoning about the emotions of others rather than as a generative system. It therefore seemed entirely feasible to use the agent's mind to simulate what other characters might feel as a result of a projected action. This is different from forms of projection already tried in which an AI planner is run to predict future actions [10]. Rather than predicting actions, a character predicts the emotional responses of characters around it to the set of actions it could possibly take, allowing it to pick the action with greatest emotional impact. Dramatic effect rather than task efficiency is the objective here.

Cite as: If I were you: double appraisal in affective agents, Aylett, R.S & Louchart, S., *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Padgham, Parkes, Müller and Parsons (eds.), May, 12-16., 2008, Estoril, Portugal, pp. 1233-1236.

2.THE FAtiMA APPRAISAL MECHANISM

FatiMA was developed as an agent architecture specifically to drive characters acting in strongly emotional situations. As seen in Figure 1, the emotional state of the character plays a central role, and is generated by the Appraisal boxes in the Reactive and the Deliberative Layer components. Cognitive appraisal posits that humans continuously evaluate sensory data for its significance in relation to their own goals. In the taxonomy of emotions proposed by Ortony, Clore and Collins (OCC) [12] 22 emotions are each associated with a type of appraisal rule linking them to events, objects or other characters. The Reactive Layer executes an OCC appraisal to give reactive action tendencies, for example, a character crying if its level of distress is high. To select one action from what are usually several that could be executed in a given context, that with the highest associated emotional intensity wins.

The Deliberative Layer (DL) component carries out appraisal as part of its planning process, using the emotions hope and fear, related in the OCC taxonomy to *prospective* or future events. This component implements *coping behaviour* [11], as the means through which planned actions are related to emotional state. Problem-focused coping is where the character tries to deal with its emotions by planning actions to be carried out in the world. Thus if a character was insulted, it might deal with the resulting anger by planning to punch the character that insulted them. Emotion-focused coping instead adjusts internal state and thus goals One might also deal with anger by denying that the action that caused it was important – 'oh, that character is crazy and doesn't know what they are saying, I'll take no notice' for example.

The DL distinguishes between generic goals, and intentions that concretely instantiate goals. Goals are on the left of the DL in Figure 1 with intentions on their right. The generic goal of eating might be activated when the character is hungry and sees something it can eat. If what it sees is an apple, a specific intention to eat that and not something else will be created. It is only when a goal becomes an intention that planning can take place. Initial hope and fear emotions are created and stored with the intention, based on its probability of success and the goal's importance. These are used as part of the mechanism for selecting between competing intentions and possible coping strategies. Planning is carried out for the intention with the most intense associated emotions, but the process of planning itself has the effect of updating hope and fear emotions depending on the probability of success or failure of the alternative plans relating to the intention. As planning takes time, only one planning action is carried out in each agent-mind cycle When actions become available for completed plans, they are sent for execution.

3. DOUBLE APPRAISAL

FatiMA's action-selection mechanisms are based on emotions generated within the agent. In the reactive case, this is directly related to the event that was appraised. In the deliberative case, an internal reappraisal has taken place over a number of planning cycles as the agent considers the hope and fear aroused in itself by the actions suggested in its plan. What is now required is appraisal of the agent's projected actions wrt their emotional impact on *other* characters around it. If a character is to perform more like an actor, it should execute the action that has the greatest emotional impact on others, not the action that results from the highest-intensity emotional state in itself.

The Simulation Theory suggests that a way to do this is to run the agent mind again, but this time using the projected action as if it were an event. In this way the agent considers the emotional impact of its action as if it has happened to them: an action to punch another character would be appraised as if another character hit them. We call this *double appraisal(DA)*.

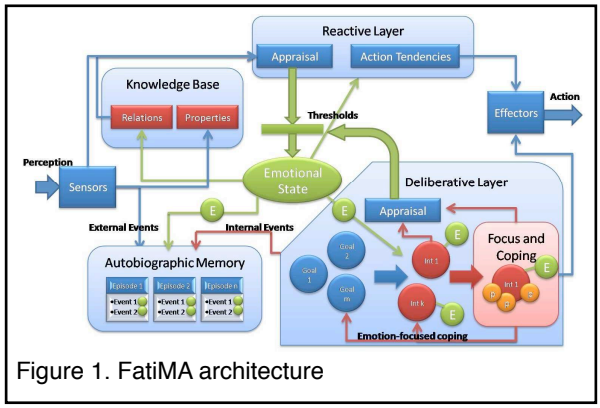

Figure 1. FatiMA architecture

3. 1. Implementing double appraisal

Actors do not pick the most dramatically interesting action they can think of: they pick the most dramatically interesting action compatible with the role they are playing. The emotional intensities used by FatiMA for action-selection are indeed what allow characters to act 'in role'. Thus in implementing double appraisal, one must balance the external emotional impact just mentioned with the internal emotions already generated. For this reason, DA reappraises a set of valid and eligible elements selected by the first appraisal cycle, ranked from highest emotional intensity for the character downwards in a valued-action array. In the reactive case, this is a set of possible actions, in the deliberative case, a set of intentions, re-appraised based on the plans to achieve them.

In the reactive case, the second appraisal process reruns the original action selection system. Each possible action in the set is recast as an imagined event, where the target of the event is the agent itself, and appraised using a copy of the agent's current emotional state. Re-appraisal uses a duplicated emotional state because otherwise the mere act of imagining how another character would react would change the agent's actual emotional state. As a result of this re-appraisal, emotions are generated and the value of the strongest emotion generated determines the value of the emotional impact for the re-appraised event. The instantiated emotional state and event pool are then reset for the re-appraisal of the next selected action. The cycle is run until all actions in the valued action array list have been re-appraised. The system then selects the action whose emotional impact is the strongest.

In the deliberative case, emotional intensity is being used to select the intention for which to plan. An intention cannot be recast as an event, but the actions in the plan the agent creates to achieve the intention can. One possibility would be to look at *all* the actions in the generated plan, but for multiple intentions it was felt this might lead to a combinatorial explosion. It was therefore decided to use the action in the plan that finally meets the intention – the action for which the intention is a post-condition or effect - as the event to be reappraised. This is an arguable choice given that one can imagine action sequences in which the emotional impact is not all derived from the last action. However in the applications so far considered, plans in fact usually involve a set of movement actions to get the character to a position where the action meeting the intention is to be executed. Pragmatically therefore this seemed a reasonable starting point for evaluation.

3.2 Extended double appraisal

In the DA approach of the previous section, the agent assumes that other characters are exactly the same as they are. However, Simulation Theory allows a human simulating another to modify their simulation according to known differences from their target. This might simply relate to known differences in beliefs, which from a cognitivist standpoint involves reasoning about the beliefs of others, a long-standing branch of rational agency theory [4]. However it might also involve assessment of the personality or the emotions of the other. In this case affective empathy, feeling the feelings of the other, might be involved. This is clearly particularly relevant to the case discussed here, where it is the emotional impact of actions on other characters that is being assessed.

The DA approach was therefore extended to take this into account, and to see if it made any difference in generating more interesting interactive narratives. There are a number of sophisticated ways in which this might be done: one character might learn the characteristics of others from observing their reactions to events and comparing this with their own emotional responses. However investing the extra effort in such an approach without knowing if it was likely to produce any positive results seemed risky. Thus the simplest possible approach was taken, using the actual emotional state of characters present in a scene. Clearly this does not have any human parallel since the minds of others are not accessible like this in the real world. In this variant of DA, called DAM, rather than assessing an action or intention with regard to the agent's own set of emotional reactions and goals, it is carried out for all the agents present in the scenario. An action is now assessed on the single highest emotional impact generated for any of the agents involved in the scenario.

4.EVALUATION

The evaluation focused on the double appraisal mechanism as a means of generating dramatic interest for both users (the interactive case) and spectators (the non-interactive case) of interactive narratives. The underlying hypothesis was that stories in which characters were using the double appraisal mechanism would be perceived as being more interesting than the ones in which they were not

FAtiMA was developed in the context of an educational anti-bullying application for children, FearNot! [2]. An entirely new scenario was created for this work, involving a group of five characters with sharply conflicting goals exploring an Egyptian pyramid that turns out eventually to contain an alien spaceship. A graphical visualisation system can be linked to FAtiMA characters, as in FearNot!, but in this instance a text-based visualiser was used to avoid the considerable effort involved in generating graphical assets as well as to neutralise the influence graphics could have on the participants' appreciation of the stories presented to them.

The system generates different stories using this scenario (identical set and initial character definitions) on different occasions, depending on which FAtiMA variant is implemented. Original FAtiMA, the DA variant and the DAM variant may result in different actions being selected by characters. Moreover, some physical actions have a stochastic outcome: for example a character who is shot may be killed or only wounded, adding a further element of indeterminacy. To take these variations into account, and also in order to prevent the user interface design from affecting results, the system was used to exhaustively generate all possible versions of a single short scene to be used for evaluation. Five distinct stories in the form of sets of language and non-language actions resulted. The language actions were turned into text by hand, manually applying a standard templating approach already used within FearNot!

Despite sharing many common elements with each other (predictable given their common scenario and action repertoire), particular stories related to particular implementations. Story 1 is generated by the original FAtiMa architecture, Story 3 by the DA extension, and Story 5 by the DAM extension. Stories 1 and 5 can be seen in [3], showing that the actions of the Doctor character change as a result of double appraisal, with a knock-on effect on other actions.

4.1Evaluation process

46 subjects (32 Males, 14 Females) carried out a number of ranking and marking exercises with the five stories. Participants were categorised into experts (10 Males, 1 Female) and non-experts (22 Males, 13 Females); their level of expertise was determined by their amount of experience with respect to storytelling or similar activities. One-way-ANOVA was applied and results are statistically significant to a 0.1 range. The probability of insignificance (p) and degree of significance (%R) are indicated for each result. Evaluation was composed of 5 different tests T1..T5.

T1 and T2 assessed stories from a non-interactive spectator perspective (Males 10 Females 6) by presenting a whole story, asking the user to mark them and also rank them in order of preference. T1 and T2 displayed the same stories, but in T2 dramatically neutral actions were added so that all stories contained the same number of actions, making them of equal length. This was to exclude the possibility that stories with more actions in them always seem more interesting. The added actions did not influence the appreciation of stories as none of them were reported as being either interesting or meaningful when subjects were asked to nominate such actions at the end of the tests.

T3, T4 and T5 assessed stories from an interactive participative perspective (Males 22 Females 8) and asked users to make decisions for one character in every cycle. The users' decisions determined the story they experienced. These stories were also marked by users. In all test cases, participant were asked to mark stories (1-5 scale) for dramatic interest and rank them. They were also instructed to mark actions for their meaningfulness and dramatic interest (1-10 scales).

4.2Results

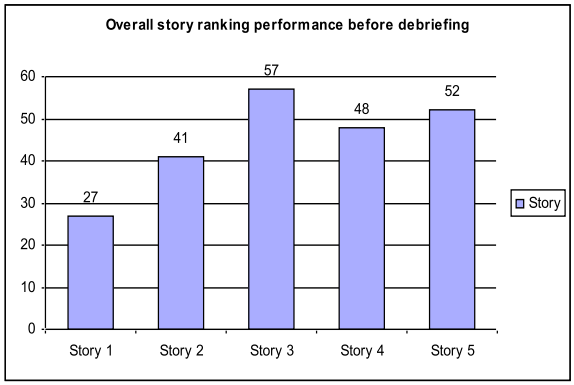

Figure 2. Overall story ranking (Population 15 – M(9)/F(6))

The results showed that the participants' perceptions of the story based on the original FAtiMA architecture (Story 1) were not as good as DA/DAM stories (Story 3, 4, 5), supporting the hypothesis that double appraisal produces more dramatically interesting stories. Figure 2 ($p = 0.00061$/ 99.39 %R) above shows

the overall story ranking for the complete set of stories before debriefing. Not only were Stories 3 and 5 perceived as the best stories, Story 1 was ranked lowest. A similar trend can also be observed when studying the overall story marking (1-5 scale) for each story in Figure 3 ($p = 0.0917$/ 90.83 %R). Here again, stories 3 and 4 score higher and story 1 scores the lowest average marking.

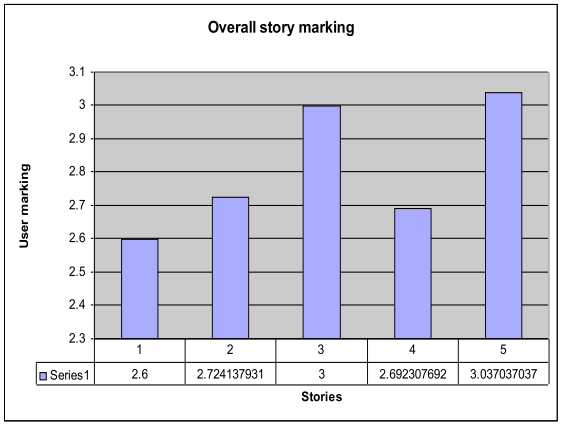

Figure 3. Overall story marking (Population 46 –M(32)/F(14))

While these results demonstrate the benefit of double appraisal they do not show any clear distinction between the two implementations tested, DA v DAM. The DA implementation feature marginally better than DAM in story ranking but marginally lower in story marking. The effect seen thus seems almost entirely due to considering the emotional impact of one's actions at all, with the accuracy of that impact in relation to specific individuals of much lesser importance. Of course this evaluation only considered one short scene, and it is possible that a small effect might take much more interaction to become obvious.

5.DISCUSSION AND FUTURE WORK

The results of the previous section indicate that adding double appraisal to FAtiMA has met its primary objective of more interesting narrative. It produced different stories from those generated by characters running original FatiMA; a characteristic of an emergent narrative is that once one character does something different, this may have a substantial effect on the direction of the story. The use of Simulation Theory has successfully allowed characters to anticipate the emotional reactions of other characters and to act upon them and in doing so to produce a more interesting story.

However, double-appraisal is a generic technique that could in principle be implemented in any agent architecture that uses cognitive appraisal as a basis for an affectively-determined action-selection system [8]. In the work reported here, it has been applied to character-character interaction, but it could also be included in characters whose prime function is to interact with human users, as an initial step towards modelling empathy. Because cognitive appraisal rules are contextually determined, double appraisal is likely to strengthen the contextual appropriateness of actions selected by an agent.

Two interesting extensions to the work carried out so far have already been mentioned. One would be to investigate further the relationship between double-appraisal and the action-sequences of plans generated during deliberation. The simplification adopted here of using the single action in the plan that actually satisfies the intention concerned ought to be investigated further, both empirically, using a number of different scenario domains as well as theoretically in terms of the complexity of the alternatives. For example, in some plans, the initial intention, if it is a complex one, might be satisfied by more than one action in the plan.

A second extension would be to incorporate learning of the variations in other characters through observing their behaviour, allowing a more principled DAM which does not require access to the actual running emotional states of other characters. This would make the DAM variant applicable to humans as well as to other agents, and would make an interesting addition to agents implemented in persuasive roles such as customer service representatives or sales agents.

In conclusion, we suggest that double appraisal is a novel and interesting implementation of the Simulation Theory in intelligent agents of potentially wide application

6.REFERENCES

[1] Aylett, R and Louchart, S. (2003) Towards a narrative theory of VR. Virtual Reality Journal Vol 7.(2003).pp2-9.

[2] Aylett, R.S; Louchart, S; Dias, J; Paiva, A; Vala, M; Woods, S. & Hall, L. (2006) Unscripted Narrative for affectively driven characters. IEEE Journal of Graphics and Animation. May/June 2006 (Vol. 26, No. 3) pp. 42-52

[3] Aylett, R.S and Louchart, S (2007) Being there: Participants and Spectators in Interactive Narrative M. Cavazza and S. Donikian (Eds.): ICVS 2007, LNCS 4871, pp. 116–127, 2007. Springer-Verlag2007

[4] Barwise, J. (1988) Three Views of Common Knowledge, Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge,

[5] Castelfranchi, C. (2005) Mind as an anticipatory device : For a theory of expectations. Proceedings, BVAI 2005. international symposium, vol. 3704, pp. 258-276,

[6] Dias, J and Paiva, A. (2005) Feeling and Reasoning: a Computational Model. 12th Portuguese Conference on Artificial Intelligence, EPIA (2005). Springer. pp 127-140.

[7] Gallese,V & Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences,

[8] Gordon, R. 1986: Folk Psychology as Simulation. Mind and Language, 1, 158-171

[9] Gratch,, J & Marsella,S (2005) Lessons From Emotion Psychology For The Design Of Lifelike Characters. Applied Artificial Intelligence 19(3-4): 215-233

[10] Laird, J. (2001) It knows what you are going to do: adding anticipation to a Quakebot. Autonomous Agents, 385-392, ACM, Montreal 2001

[11] Lazarus, R. (1991) Emotion and adaptation. *NY Oxford University Press(*1991).

[12] Ortony, A and Clore, G and Collins, A. (1988) The cognitive structure of emotions. *Cambridge University Press.(*1988).

[13] Stich, S. & Nichols, S., 1992, "Folk psychology: Simulation or tacit theory?" Mind and Language 7, 35-71.

[14] Whiten, A. (1991) Natural Theories of Mind. Basil Blackwell, Oxford, UK.(1991)