Emotional Agents in a serious game of Cooperation and Competition

Bernardo Grilo

Técnico Lisboa

November 25, 2015

Abstract

Emotions influence the way humans think, make decisions and act, determining their beliefs, motivations and intentions. It is a generally accepted fact that these influences are, most of the time, beneficial. In this work, we developed agents capable of replicating these emotion-driven decisions and that are also able to recognize emotion in other agent's actions. These agents exist in the context of a serious multi-agent vs player game with cooperative and competitive components. In this game, agents are able to display behaviours associated with emotions of gratitude that can promote cooperation between the community the agents belong to. On the other hand, agents also display behaviours associated with anger that can help them avoid adverse behaviour from other agents.

In this work we also evaluate the developed agents according to believability of their behaviours, their accuracy to human behaviour emulation and if these agents can actually get an advantage over other agents through emotional action.

Key words: Agents, Emotion, Sentiment, Influence, Memory, Decision-making, Cooperation, Competition, Appraisal, Personality

1 Introduction

In recent years there has been a substantial increase in the study of *Affective Computing*, leading to very important breakthroughs relevant to this field. We call *Affective Computing* the field of study and development of systems that can recognize, interpret, process, and simulate human emotions. The field's increased relevance comes from a growing need from society to create technology capable of interacting on an emotional level with humans, whether by systems being able to recognise user-generated emotion or by those systems being able to express their own emotions. Our work's focus is to design an agent-driven solution for intelligent systems that enables emotional expression and decision making processes driven by emotional reasoning, in a way that is recognizable to a human audience.

The problem that motivated our work, was the development of a multi-agent system capable of generating and processing emotions, as well as, making decisions guided by those emotions. This work had two main objectives, the first one was to find evidence that affective agents (agents that can emulate emotional processes) can benefit from a decision-making process that is guided by emotions, enabling them to gain more utility when compared with agents without affective behaviour; the second objective was to test our solution against a human audience and also find evidence that our solution for affective behaviour is in fact more believable to that audience in terms of how affective agents interact, when compared against a solution for agents with non-affective behaviour. We developed this solution for a multi-agent system that is able to make emotion-based decisions and process emotions in the context of a Prisoner's Dilemma-type of scenario.

We concluded our work by gathering evidence from several scenarios in agent vs agent interactions, to test our developed solution for affective agent-driven systems' with different personalities and different configurations to cover all possible emotional actions and responses, and elaborated on our conclusions relating to our hypothesis which states that emotions provide a way for our affective agents to make decisions in a more efficient way, which will give them an edge against non-affective agents. We also validated with a human audience how our solution stands in terms of the affective and non-affective agent's behaviour believability and how they perceived the emotional interactions that took place.

2 State of the Art

Because the scope of our work is very interdisciplinary, it was developed with an extensive study of many theories from diverse fields like Game Theory, Affective Computing, Theories of Cooperation and Competition, Memory Mechanisms and Personality Theories.

Since these theories are themselves very extensive, having many aspects that are not relevant for this work, we only delved into the relevant theories and detailed the most important and pertinent aspects of each one. This research allowed the development of an architecture for a multi-agent system with the specifications needed for validating our hypothesis and take conclusions concerning how our emotion-driven agents can have an edge over non-emotional agents when influencing their decision-making processes with emotion, as well as validate if these mechanisms of emotion were possible to replicate in Agent-Oriented software system in a way that is believable to a human audience.

We researched theories which acted as background for introducing the problem that motivated the work we developed. The main theories that supported the motivation for our work were Game Theory[1], more specifically how to approach Prisioner's Dilemma-type of scenarios, theories of Competition and Cooperation[6] among societies of individuals which clarifies factors that could influence behaviours of cooperation and competition and we also investigated appraisal theories[18] specifically the OCC Model[19] that is represented in Figure 1. as revised by the work of Steunebrink[4].

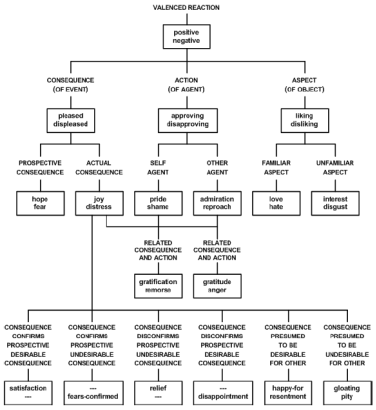

Figure 1: The Revised OCC Model

We also researched relevant works which acted as background for the solution we developed, in order to solve the problem of developing an architecture for emulating affective behaviour that can be found in human decision-making processes and replicate them for synthetic characters. Some of the more relevant works were the work of Pimentel [9] on the role of affective behaviours in sustained multi-agent interactions which set the stage for what we wanted to accomplish and also the work of Antunes [13] that consisted on the INVITE Game which allowed for the development and testing of our affective agents solution in the context of the INVITE Game's concept were both indispensable for the the work we developed.

3 Solution

The solution we developed focused on responding to the objectives that we set out to achieve and that were enumerated in the beginning of this document. This solution is tightly related with the inevitable problem posed by the prisoner's dilemma-type of scenarios, as it is described in Game Theory [1] that in these scenarios, although agents have the choice to cooperate or defect, will ultimately defect because, according to Game Theory, logic dictates that assuming that the other participant of a Prisoner's Dilemma-type of scenario will defect then we must also defect to guarantee a small amount of utility. With our solution for affective agents we aim to break this inevitability by allowing agents to add emotion to their decision-making process and enable them to consider not only logic but also emotion when deciding to either cooperate or defect. This should result in our emotional agents being able to increase their gain of utility either by influencing other agents to act in a more beneficial way towards the affective agents or simply dissuade them from continued adverse behaviour that is harmful for the affective agents' utility.

These actions to either motivate or dissuade through emotional action other agents, comes with an added risk in the form of spent utility for the affective agent. However if the scenario plays out in the agent's favour, that should translate into more utility, turning an irrational decision into a rational one, when considering the long run. With this objective in consideration, we developed an algorithm for agent behaviour based on the work of Pimentel [9] on the role of affective behaviours in sustained multi-Agent interactions. Our goal, by developing this solution for agent's that display af-

fective behaviour, is to find evidence that in a prisoner's dilemma-type of scenario agents who can display affective behaviour and also can detect those same emotional behaviours in others, will increase their utility and overcome agents whose behaviours do not have these affective characteristics.

Our affective agents were developed to be able to interact with each other as well as with human players, in the context of the INVITE game. We developed an algorithm which instantiates agents that can assume different personalities, are able to develop emotions, display affective behaviour and also detect affective behaviour in other agents. For our specific scenario, we will only develop affective behaviours related with the emotions of Anger and Gratitude as specified in [9]. Also to conform to the scenarios set up by Pimentel, our solution was developed taking into account that our agents would have to interact with each other on an emotional level as well as be able to identify emotional behaviour on others. Since the original INVITE Game [13] does not support this type of interaction that allows for displays of emotion, we had to extend it and develop mechanisms of our own that would enable our agents to engage each other in affective behaviour. In Figure 2. is a design model for the developed architecture for our affective agents.

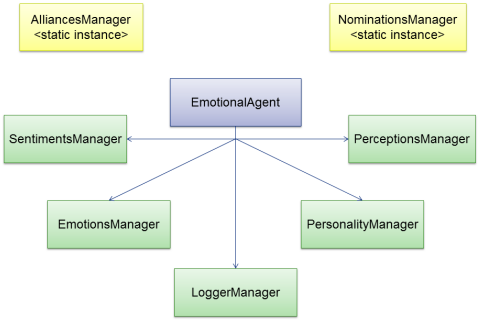

Figure 2: Affective Agent's Architecture Design

The main specification of the scenarios of emotional behaviour found in Pimentel's work [9] which played a great part in the design choices of the emotional interactions we developed, was that the participants of the scenarios had mechanisms of emotional interaction, both for grateful and grudgy behaviours, that allowed each participant to direct their emotions towards specific individuals. This design limited our use of the core INVITE mechanics to replicate these emotional actions. For example, if an agent was holding a grudge towards another and because of that specific grudge, that agent would stop collecting wood, this agent would not only be punishing the targeted agent of it's anger but would also be punishing all agents that belong to the same team. This would render the detection of emotional behaviour from others impossible since there would be no criteria for perceiving who is the target of a grudgy action that took place.

The mechanics that were developed for our solution take this very important requirement into account so that agents can act on their sentiments towards other agents in a targeted way. Two mechanics were developed to do this, one for agents to act on their grudges and another for them to act on their gratefulness.

The first mechanic is related with displaying grudges towards other agents. This mechanic is based on a common nominations system where agents vote on others with the objective of showing their displeasure towards them. An agent A takes an action that is evaluated by agent B as being too adverse to A's objectives, agent A will show it's displeasure towards agent B's action by nominating it. A example of adverse behaviour would be for example if an agent is indulging in individualistic behaviour by collecting too much gold. These agents are much more likely to get nominated by other agents whose behaviours are more focused on the common good than their own.

When designing the second mechanic for grateful behaviour between agents, the first instinct was to enable agents act gratefully towards another team mate by collecting more wood pieces to progress further the raft completion, since this is the concept of being cooperative in the INVITE game. However this approach was not adequate to the requirements scoped for our work since this way, every time an agent collected more wood to reward a certain agent, it would actually improve every agent's utility regardless of all agents having earned that improvement or not and this fact did not conform to the specifications of scenarios for grateful behaviour introduced by the work of Pimentel [9]. Because of this fact, we developed a different approach to this mechanic that involves the resource of gold pieces instead of wood pieces, to try and design a reward system that would allow targeted displays of grateful behaviour by donation of gold pieces. We enabled our agents with grateful behaviour by allowing them to create alliances among each other. If an agent is enabled to display gratitude towards others it will try to find

other agents who display this same grateful behaviour with whom they can create alliances with. This process takes place at the beginning of each day, where agents who are able to display gratitude start to probe other agents by **offering them pieces of gold as a gesture that represents a cooperative action**. This is supposed to motivate the targeted agents to also take cooperative action by returning that gold offer and through that action, create an alliance with the other agents. Note that after an agent tries to create an alliance with another agent, that agent becomes checked and won't be probed again with cooperative actions in the future to create alliances for the rest of the game. The following diagram represented in Figure 3. displays how the agent's routine flow takes place and the order in which every key moment happens. Note that this diagram includes our own changes to the INVITE game so we can clearly understand when nominations related actions and alliances related actions occur.

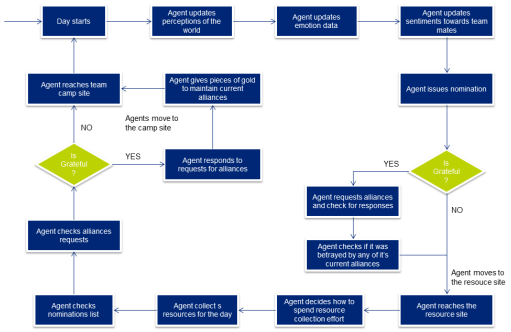

Figure 3: Emotional agent's day-to-day cycle diagram

4 Data Analysis

We validated two hypothesis through the data that we were able to gather. First we wanted to show evidence that, with our solution for affective behaviour, agents that are able to act emotionally towards others can benefit from an advantage against agents who are not enabled to act emotionally. The second hypothesis could only be validated through user tests and consisted on getting data which supports the agent behaviour believability from a human player's perspective increases when an agent acts emotionally towards them. We started by testing scenarios composed only of agents running our solution for affective behaviour. For each scenario we detailed the affective configuration of every agent that was going to participate.

The evidence gathered from the execution of these scenarios lead us to the conclusion that these mechanics of emotional behaviour are useful but their use seems to have a circumstantial effect in the sense that affective behaviour is not always a bullet-proof strategy to improve utility since we were not able to see this certainty in the results that we extracted and analysed from our agent-only test scenarios.

We see a clear advantage of affective-driven strategies against non-affective behaviours. We also had evidence of these advantages when the affective agents are interacting with a team whose member's personalities allow the affective agent to leverage on it's emotional actions and influence others to behave in a way that is most beneficial. However if the agent belonged to a team in which it's personality traits do not allow it to thrive and influence the other team mates, we have evidence that this advantage of affective behaviour becomes a harmful handicap.

Concerning this conclusion, the work of Wilson on Group Selection gives great insight by stating that in social scenarios where individuals belong to certain groups, these individuals should be allowed to maximize the utility that they can achieve with their behaviour by enabling them to act on multiple levels. This consists on individuals being able to interact between themselves in the same social group as well as between groups, allowing agents to bond not only with their own groups but also with other individuals outside of the group with the objective of being able to select which group they wish to interact with in order to maximize their utility. Throughout most of the scenarios that we tested, the problem that kept persisting was that when agents were performing badly because of how their personalities were not able to influence their team mates, they would become doomed to failure because agents would be limited to interact with the same team mates throughout the whole game. We believe that allowing agents to interact on multiple levels that do not limit the agents to bond only towards their team mates but allow them to also bond with agents from other teams, could allow them to find individuals that would possibly maximize the utility that they can achieve by surrounding themselves with agents that they could influence to act in their best interests through both their personality traits and affective behaviours. Another important source of

data for the analysis of our second hypothesis was to validate with a human audience the believability of our affective agents' behaviours. This user feedback was gathered from a sample of fifteen users that participated in our user testing sessions. To gather this feedback from the users we resorted to the methodology for evaluating believability of synthetic characters through quantifiable data as defined by the work of Gomes [10] on metrics for character believability in interactive narrative. We developed a survey with questions related with the users experience and their interactions towards our affective agents. We also approached questions concerning how those interactions translated to them in terms of believability in the way they bonded with the user and the way they acted emotionally towards them.

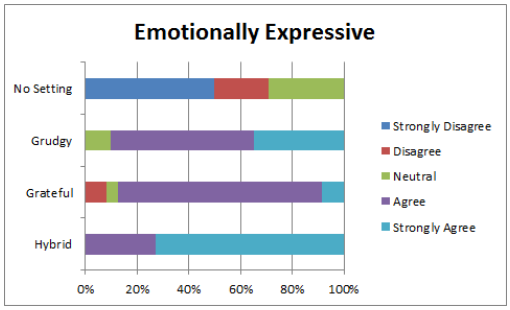

Figure 4: Emotional Expression User Feedback

These user testing scenarios provided us with evidence, according to Figure 4. leading us to conclude that our developed solution for emulating emotional behaviour is able to do it in a believable way that allows for emotional interaction. We analysed the agents believability feedback given by our testers, which allowed us to conclude that some disapproval was noted towards the way the agent's personality shaped their resource collection efforts which made them look too predictable in the sense that agents were too consistent in their behaviour. However in a general way our agents were classified as believable to a human audience in the way the affective agents behaved in their actions that resulted from emotional responses. Actions like nominations issuing and alliances building, according to the mainly positive results from the survey, were classified as believable by our testers when comparing with results concerning the absence of these interactions from non-affective agents.

5 Conclusion

We believe we were able collect enough evidence from the data analysis towards validating our hypothesis. According to the data related with the agent-only scenarios we were able to confirm how the affective agents had a clear advantage over the non-affective agents which confirms that emotional judgement can be more beneficial against pure logical decision making. However we were also able to conclude that when setting affective agents against each other, these agents were not always able to benefit from these emotional behaviours. Those benefits occurred in a circumstantial manner depending on the affective and personality characteristics/configurations of each team mate. This leads us to infer that emotionally driven behaviour does not lead to guaranteed advantage, this is actually determined by the other agent's affective configurations and how those configurations react to specific emotional behaviours and if they allow the an agent's emotional action to enable it to either succeed or fail.

Regarding user tests, we were able to gather more evidence leading us to conclude that our developed solution for emulating emotional behaviour is able to do it in a believable way that allows for emotional interaction between human players and agents running our solution for affective behaviour. We witnessed similar interactions on scenarios of agents vs agents and agents vs human players. However, much like in scenarios composed only of agents, once again any advantages from emotional behaviour proved to be circumstantial and only took place depending on the type of player that was interacting with the agents, as well as the agent configurations at a given scenario. We witnessed scenarios where humans players were able to exploit the emotional agents when their configurations were set in ways that allowed for the human players to thrive, as well as scenarios where the affective agents were not compatible with the human player and did not allow themselves to be influenced by the human player into behaving in a more beneficial way. The fact that these similarities took place for most of the scenarios, when compared with the agent vs agents scenarios, gives us evidence that our solution for affective agents is able to emulate emotional behaviours that occur in human individuals.
Finally we analysed the agents believability feedback given by our testers, which allowed us to conclude that some disapproval was noted towards the way the agent's personality shaped their re-

source collection efforts which made them look too predictable in the sense that agents were too consistent in their behaviour. However in a general way our agents were classified as believable to a human audience in the way the affective agents behaved in their actions that resulted from emotional responses. Actions like nominations issuing and alliances building, according to the mainly positive results from the survey, were classified as believable by our testers when comparing with results concerning the absence of these interactions.

6 Future Work

An interesting future work would be based in the work of Wilson on the theory of Group Selection [16], and use this limitation from our work as a stepping stone in allowing the agents to not conform to one group of agents to bond with throughout the full duration of the scenario they belong to. Instead allow them to bond with all the agents within the scenario that they belong to but exist on other social groups and with the information gathered from those bonding interactions, enable the understanding of which agents would be the ideal ones to interact with, and leverage on their particular set of affective behaviour configurations, allowing them to maximize their utility when interacting with these agents. This would enable the them to not become doomed from day one to interact with agents that are not influenced by their emotional actions and allow them to choose the agents they wish to interact with in order to make the most out of their time interacting with other agents that could allow themselves to be influenced by emotional action to become motivated to act in a more beneficial way, as well as avoid adverse behaviour.

Another possible opportunity for improvement would be to develop a solution that takes even further the OCC Model [4] for more than only the two emotions of *Gratitude* and *Anger* which were the ones developed for this particular work. It would be interesting to increase the complexity of this model and see a system that could generate a wider range of emotions and actually apply this to a more generic and complex scenario for agents to interact in. Ultimately, this more generic solution should reflect similar conclusions to the ones we reached in our work which was to show how agents interacting on an emotional level would eventually be able to make better decisions than agents who are acting only on a logical level, while at the same time enforce this idea with the added feature that agents can profile the society they belong to and choose which agents to interact with, in order to improve their utility through continuous interaction.

References

[1] Wooldridge M.: "Multiagent Interactions", in *An Introduction to MultiAgent Systems*, John Wiley & Sons LTD (eds), pp 105-129, University of Liverpool, Liverpool, UK, 2009

[2] Carver J.M.: "Emotional Memory Management: Positive control over your memory", in *Burn Survivors Throughout the World Inc*, available at `http://www.burnsurvivorsttw.org/articles/memory.html`, 2005

[3] De Raad B.: "The Big Five Personality Factors", in *The psycholexical approach to personality*, Hogrefe & Huber Publishers (eds), Ashland, OH, US, 2000

[4] Steunebrink B.R., Dastani M. and Meyer J.C.: "The OCC Model Revisited", in *Proc. of the 4th Workshop on Emotion and Computing*, Utrecht University, Utrecht, NLD, 2009

[5] Lim M.Y., Aylett R. and Jones C.M.: "Emergent affective and personality model", in *Intelligent Virtual Agents*, Heidelberg, DEU: Springer Berlin Heidelberg, 2005

[6] Deutsch M., Coleman P.T. and Marcus E.: "Cooperation and Competition", in *The Handbook of Conflict Resolution Theory and Practice*, John Wiley & Sons (eds), San Francisco, CA, 2011

[7] Yong C.H. and Miikkulainen R.: "Cooperative Coevolution of Multi-Agent Systems", University of Texas Austin, TX, 2001

[8] Theraulaz G. and Bonabeau E.: "A Brief History of Stigmergy", in *Artificial Life*, pp 97-116, Institute of Technology, MA: MIT Press Journals, 1999

[9] Pimentel C.: "Grateful Agents and Agents that Hold a Grudge: The Role of Affective Behaviours in Sustained Multi-Agent Interactions", ICAART, Barcelona, 2013

[10] Gomes P., Paiva A. and Martinho C.: "Metrics for Character Believability in Interactive Narrative", Instituto Superior Técnico, Lisbon, PT: Springer International Publishing, 2013

[11] Oatley K.: "The Sentiments and Beliefs of Distributed Cognition", in *Emotions and Beliefs - How Feelings Influence Thoughts*, Frijda N.H., Manstead A.S.R., and Bem S. (eds), pp 78-107, Cambridge, UK: Cambridge University Press, 2000

[12] Minsky M.: "Emotion", in *The Society of Mind*, Simon & Schuster Paperbacks (eds), pp 162-172, New York, US, 1985

[13] Antunes B.V.: "Believable Synthetic Characters with Social Identity", MSc Thesis, Instituto Superior Técnico, Lisbon, PT, available at `http://gaips.inesc-id.pt:8081/invite/`, 2012

[14] Pimentel C.: "Emotional Reasoning in AI: Modelling Some of the Influences of Affects on Reasoning", PhD Thesis, Instituto Superior Técnico, Lisbon, PT, 2010

[15] Picard R. and Ahn H.: "Affective-cognitive learning and decision making: A motivational reward framework for affective agents", in *Affective Computing and Intelligent Interaction*, pp 866-873, Cambridge, MA: MIT Media Lab, 2005

[16] Wilson D.: "Multilevel selection theory and major evolutionary transitions implications for psychological science", in *Current Directions in Psychological Science*, Binghamton University, New York, US, 2008

[17] Brembs B.: "Kin Selection", in *Preliminary draft for the Encyclopedia of Genetics*, New York, US: Academic Press, available at `http://brembs.net/hamilton/`, 1996

[18] Clore G.L.: "Appraisal theories: How cognition shapes affect into emotion", in *Handbook of Emotions (3rd ed.)*, Lewis M., Jeannette M. and Barret L. (eds), pp. 628-642, New York, US: Guilford Press, 2010

[19] Ortony A., Clore G. and Collins A.: *The Cognitive Structure of Emotions*, Cambridge, UK: Cambridge University Press, 1990