Emotions as Heuristics in Multi-Agent Systems

Bas R. Steunebrink, Mehdi Dastani, John-Jules Ch. Meyer

Utrecht University
Institute of Information and Computing Sciences
The Netherlands
{bass,mehdi,jj}@cs.uu.nl

Abstract. In this position paper we argue that a model inspired by human emotions will provide reasonable and useful heuristics for 1) reducing and controlling nondeterminism involved in an agent's decision making process, 2) flexible cooperation and coordination between agents, and 3) building efficacious human-agent interfaces to facilitate the interaction between human users and artificial agents in hybrid multi-agent systems.

1 Introduction

Many branches of computer science in general and multi-agent systems in particular involve problems that need heuristics to manage their complexities. In this position paper, we discuss three problems in multi-agent systems for which we propose heuristics inspired by human emotions. These problems are related to the autonomy of individual agents, the cooperation and coordination between agents, and the believability and efficacious interaction with agents.

In multi-agent systems, individual agents are assumed to be *autonomous*. This implies that they should have the capability to decide which actions or plans to perform in order to reach their objectives. For many practical applications, standard decision-theoretic concepts (e.g., probability and utility) and rules (e.g., maximizing expected utility) will leave an agent with multiple actions or plans with equal or near-equal preference. This gives rise to nondeterminism in an agent's decision making, which is a problem if the agent has to choose only one of the options. Moreover, individual agents need to *cooperate* and *coordinate* their actions in order to achieve their global (social) objectives, often called the objectives of a multi-agent system. Interaction protocols and mechanisms are the common means to specify the cooperation and coordination of individual agents to guarantee the achievement of their global objectives. However, the use of interaction protocols imposes restrictive constraints on the behavior of individual agents and thereby limits their autonomy. On the other hand, the absence of interaction protocols brings the problem of nondeterminism back to individual agents since there may be many options for interaction with other agents. It should be noted that an interaction protocol can in fact be considered as a way to solve the problem of nondeterminism. We believe that for many (non-critical) applications of multi-agent systems the restrictive interaction protocols can be replaced by more flexible and more general policies or heuristics. Finally, in hybrid multi-agent systems, where human users are also considered as constituting agents, efficacious *interfaces* are needed to enhance and optimize the complex interactions between humans and artificial agents. Also, as humans intuitively assign a personality, motives, and an affective state to any system displaying behavior, the challenge in human-agent interaction lies with living up to these expectations of the human user and thus creating *believable* agent behaviors.

In this position paper, we propose using heuristics inspired by human emotions to (partially) solve the aforementioned problems. In particular, we propose to formalize an (existing) model of human emotions and incorporate it in the existing models of multi-agent systems in order to reduce and control the involved nondeterminism and enhance the interaction between human and artificial agents. In section 2, we discuss the role and application of emotions to the three mentioned problems. In section 3, we discuss the current and future research activities as well as the results we expect to achieve.

2 Emotions in Multi-Agent Systems

We argue that as long as heuristics are needed for solving the three mentioned problems in multi-agent systems, a model of affect based on human-inspired emotions may provide reasonable and useful heuristics. Moreover, using a model of affect is especially beneficial because it is a single mechanism that functions as a heuristic for all three of these problems. Below we will discuss how we envisage the use of emotions in multi-agent systems.

Agent Autonomy For reducing nondeterminism in decision making, we propose using a heuristic based on secondary emotions [2, 11]. Secondary (also called deliberative) emotions are the biological heuristics for preventing excessive deliberation. As an example of how emotions might aid the deliberation and decision making process, suppose the affective mechanism of an agent is modeled after the cognitive model of Ortony, Clore & Collins ("OCC") [10], and that the agent generates a plan to achieve a goal. The agent now *hopes* to achieve the goal as long as it is committed to the plan. However, as soon as the execution of the plan fails (e.g., the execution of some action fails or the effect of some action is not perceived), the agent will also experience *fear* with respect to the plan and the goal. As soon as the intensity of the fear becomes greater than the intensity of the hope, the agent may drop the plan and start replanning to achieve the goal. Of course the agent will prefer new plans that have previously caused it to experience *satisfaction*, while avoiding those that have previously resulted in *disappointment*. In this way, the incorporated affect model indicates which plans need to be dropped or adopted and thereby helps to reduce the nondeterminism involved in an agent's decision making process.

Agent Cooperation and Coordination If agents, which have to cooperate with each other, know how they work internally, they can anticipate expected actions. Note that this assumption is realistic only for closed (not open) multi-agent systems where the design of agents is known. For example, if an agent predicts that a cooperating agent cannot perform its current plan which contributes to the achievement of a common goal and is about to drop it, then the agent may be able to take anticipatory actions in order to prevent the imminent failure of the other agent. We believe that, if all agents were using an affective mechanism as a heuristic for their decision making, this kind of cooperation can be achieved in much the same way as humans do in their interactions. Specifically, an agent could accomplish this by mapping the perceived state of another agent to its own affective mechanism and approximate the affective state of the other agent. It can then predict what actions the other agent is most likely to perform next by running its own affective decision making heuristic. This is what humans do in their interactions; building a model of another person's affective state and predicting how the other person might react. If agents were using the same affective mechanism, they could do the same type of (implicit) cooperation.

Human-Agent Interaction Humans are experts in social reasoning and behavior. If artificial agents were to use the same affective mechanism as humans, then humans would have an intuitive and efficacious way of interacting with them. A reason is that such agents would be able to live up to the affective expectations of the human to a far greater extent. An affective agent could also present information to its user according to its affective state, further facilitating an intuitive mode of interaction. Consequently, this could greatly increase the *believability* of agents. Conversely, if an agent had affective mechanisms like a human, then it could also map the perceived affective state of a human to its own affective mechanism and choose more suitable and anticipatory actions, just like in multi-agent systems as described above.

3 Future Research and Conclusion

We are currently working on the formalization and implementation of the 22 deliberative emotions from the OCC model as heuristics for controlling nondeterminism in goal-directed agents. The OCC model has previously been used for emotion synthesis (often partially adapted to the research domain), for example for modeling personalities in social relationships [5], facial expressions for poker playing agents [6], and believable animated characters [1, 9]. We use the OCC model because of its suitability for formalization, but we plan to broaden our research with alternative theories of emotion. Currently, we are formalizing the entire set of 22 emotions in a modest extension of the KARO framework [8, 7]. Specifically, we are formalizing the eliciting conditions of these emotions in terms of beliefs, goals, actions, and plans. We plan to extend the formal model with a quantitative model for these emotions, capable of handling emotion potentials, thresholds, and intensities. Because of space limitations, we cannot present the formal details, but the reader can find some of the formal results in our previous publications [3, 8]. To further prevent unnecessary deliberation, we plan to extend the resulting affective model with real-time reasoning capabilities; that is, reactive behavior induced by primary emotions [2, 11]. To solve the issue of cooperation and coordination in multi-agent systems, we plan to further extend the affective model with social reasoning capabilities. We expect to accomplish this by allowing agents to map the perceived affective states of other agents to their own affective model, which will allow agents to reason about the goals and plans of other agents as they reason about their own. It is our intention to use the final resulting affective model in the new incarnation of 3APL [4, 3], which provides a multi-agent platform for goal-directed agents that use beliefs, actions, and plans to achieve their goals. As an example scenario to show the usefulness of affective agents, we envisage such agents in a dynamic environment in which external influences can change such that the agents' tasks can be obstructed. This could be a spatial environment characterized by parameters such as weather (sunny or rainy), time (day or night), and area (inside or outside). Agents have to cooperate to replace and/or paint objects which can be obstructed because of the location of the objects in combination with the time, weather, and the actions of other agents (e.g., an object located in an outside area cannot be painted if it rains).

To conclude, we have identified three problems in multi-agent systems, namely nondeterminism in autonomous decision making, cooperation and coordination in multi-agent systems, and believability in human-agent interaction. We expect that the described affective model based on human-inspired emotions provides a reasonable and useful solution to these problems.

References

1. Joseph Bates, A. Bryan Loyall, and W. Scott Reilly. An architecture for action, emotion, and social behaviour. In C. Castelfranchi and E. Wemer, editors, *Artificial Social Systems, MAAMAW-92 (LNAI Volume 830)*, pages 55–68, 1994.
2. Antonio R. Damasio. *Descartes' Error: Emotion, Reason and the Human Brain.* Grosset/Putnam, New York, 1994.
3. M. Dastani and J.-J. Meyer. Programming agents with emotions. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI'06).*, 2006.
4. M. Dastani, B. van Riemsdijk, F. Dignum, and J.-J. Meyer. A programming language for cognitive agents: Goal directed 3apl. In *Proceedings of the First Workshop on Programming Multiagent Systems: Languages, frameworks, techniques, and tools (ProMas03)*, volume LNAI 3067, pages 111–130, 2004.
5. Clark Elliott. I picked up catapia and other stories: A multimodal approach to expressivity for "emotionally intelligent" agents. In *Proceedings of the First International Conference on Autonomous Agents*, pages 451–457, 1997.
6. Tomoko Koda. Agents with faces: A study on the effects of personification of software agents. Master's thesis, MIT Media Laboratory, 1996.
7. J.-J. Meyer, W. van der Hoek, and B. van Linder. A logical approach to the dynamics of commitments. *Arificial Intelligence*, 113:1–40, 1999.
8. John-Jules Ch. Meyer. Reasoning about emotional agents. In Ramon López de Mántaras and Lorenza Saitta, editors, *Proceedings of the 16th Eureopean Conference on Artificial Intelligence (ECAI 2004)*, pages 129–133. IOS Press, 2004.
9. W. Scott Neal Reilly. *Believable Social and Emotional Agents.* PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA, 1996.
10. Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions.* Cambridge University Press, Cambridge, UK, 1988.
11. A. Sloman. Motives, mechanisms, and emotions. In The Philosophy of Artificial Intelligence, pages 231–247. Oxford University Press, Oxford, 1990.