Collaborative Adaptive Autonomous Agents

Doctoral Consortium

Mirgita Frasheri
Mälardalen University

KEYWORDS

Adaptive Autonomy; Multi-agent Systems; Collaborative Agents

ACM Reference Format:
Mirgita Frasheri. 2018. Collaborative Adaptive Autonomous Agents. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 3 pages.

1 INTRODUCTION

Autonomous agents are usually defined as software entities capable of perceiving and acting on their environment, with various degrees of learning and adaptation skills, which may or not be embodied, i.e. have a physical body through which to act. Autonomy itself has been defined in several ways during these past decades. On one hand, it can be expressed in two dimensions [9], that of self-directedness – the agent can choose its own goals, and self-sufficiency – the agent can achieve a goal by itself without help. On the other hand, autonomy represents a relational concept: (i) autonomy with respect to the environment – how independent an agent is toward environmental stimuli, and (ii) with respect to other agents – how independent an agent is with respect to influences coming from other agents or social autonomy. Furthermore, the autonomy of agents – and different levels of it – can be defined in the context of dependence theory [2]. An agent A that lacks either a goal, resource, plan, action, know-how, or any means to achieve a task, needs to depend on another agent B which can be a provider for them. In this case A is not autonomous from B with respect to the means it is lacking for completing its task. As a result, the level of autonomy of an agent can be regulated by the granularity of dependencies it can have toward other agents.

In present days, the field of autonomous agents continues to be relevant in research, whether from an academic or industrial perspective. A considerate body of work in the literature tackles the definition of autonomy, what it means for autonomy itself to change, what role should the human have in the whole process, and recently ethical considerations are gaining attention as well. The 10-levels of autonomy scheme proposed by Parasuraman et. al [11], represents one of the earlier attempts to provide a guideline for understanding different levels of autonomy. In the first level, the machine has no decision-making authority and does everything the human says. Going up in these levels, the machine can either provide options to the human, choose for the human, or even be completely independent from the human in the tenth level. Many other theories on changing autonomy levels have been developed such as: adjustable autonomy, adaptive autonomy, mixed-initiative interaction, sliding autonomy, collaborative control, dynamic autonomy among many others [1] [3] [8] [9] [10]. Moreover, there is not a unified theory of autonomy. These theories usually distinguish from each other in the actor that performs the change in autonomy level (e.g. the agent itself, human operator, or combined decision-making in some way), but their ambition remains similar, i.e. to make the transition of control smooth between agents themselves, agents and humans, and achieve good teamwork and collaboration between the different actors involved. Moreover, human-machine interaction is the focus of many works in the field. On the other hand, the state-of-practice is being shaped by big companies such as Google, and Tesla among others, which have each introduced their own prototypes of self-driving cars. At the heart of such products is the intelligent software which should continuously make decisions about what to do next while operating in a dynamic environment. Such decisions are taken autonomously, or with some degree of autonomy, that depends on the influence/feedback from human operators or other software.

This research focuses on how (software) agents can change their own autonomy given particular circumstances that can arise during their run-time. This particular flavour of changing one's own autonomy is referred to as adaptive autonomy (AA). The main assumption in this work is that agents change their own autonomy during their operation, when the dependence relations between them change. Therefore, AA agents should make their own decisions on whether they will accomplish a goal or task themselves or by depending on other agents, and whether to let others depend on them at any point in time. The problem tackled is that of determining when an agent should change its autonomy and enter into collaboration with other agents. To this end, firstly, a high-level agent architecture is proposed which models the agent's internal operation. Secondly, the adaptive autonomous behaviour is achieved by introducing the willingness to interact, a composite concept, composed of the willingness to ask for and give assistance. The willingness to interact defines both aspects of interactive behaviour. A mathematical framework is being developed for the calculation of its components. Potential application domains relevant to this work can be search and rescue, or agriculture solutions, in which it might not always be possible to rely on a central coordinator located remotely, due to unreliable communication channels. Thus, AA can be used to add a layer of flexibility that allows agents to attempt task completion by collaborating with one another.

2 AGENT MODEL

The agent proposed in this work has five states, which are: *idle*, *execute*, *interact*, *regenerate*, and *out of order* (Figure 1). Any agent is assumed to start its operation in the *idle* state, where it is not committed to any goal or task. While in *idle*, an agent can generate a task itself, or receive a request for a task from another agent. In

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

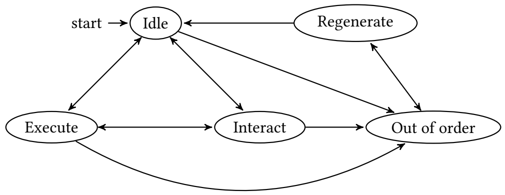

Figure 1: The agent's five states of operation [5].

the former case, the agent will shift to the *execute* state. In the latter case, the agent will shift to the *interact* state, where it will decide whether to accept or drop the request based on its willingness to give help. If the request is accepted, then the agent will shift to *execute*. Otherwise, the agent will return to *idle*. While in *execute*, the agent continuously reasons whether it needs assistance from other agents based on its willingness to ask for help. In case it does, it will make a request to another agent, otherwise it will continue on its own. An agent can also receive a help request from another agent whilst in *execute*. As a result, it will shift to *interact* where it will decide whether to accept the request. If it decides to accept, the agent will put its own task in a FIFO queue, and continue with the new task. The *regenerate* and *out of order* states are auxiliary. Different triggers can be used to make the transition from any other state to *out of order*. In this work, an agent's energy level is used as a trigger, i.e. if this level goes below a specified threshold then the agent will switch to *out of order*. Whilst in this state, the agent can shift to *regenerate* in which it attempts to regenerate itself. It is assumed that the agent always succeeds in *regenerate*, from which it shifts to *idle*. Nevertheless, in principle it can be possible for an agent to fail its regenerative routine and go back to *out of order*. As of now, the focus has been on the three main states, *idle*, *execute*, and *interact*, and the auxiliary states are defined in less detail.

3 WILLINGNESS TO INTERACT

The willingness to interact concept has been proposed in this work to shape an agent's own autonomy. It is composed of two components, the willingness to give and ask for help. The overall goal is to develop a mathematical framework that can be used to calculate the willingness to interact based on a set of factors that are considered relevant in such process. Previous work [5] has analysed possible factors that could be taken into account, such as: agent's energy level, knowledge, equipment, abilities, tools, perceived environment risk, perceived collaborator risk, task progress, task trade-off, and own performance.

4 EMPIRICAL RESULTS

The hypothesis under investigation in this research is: "Agents which display adaptive autonomous behaviour complete more tasks than agents with static autonomy". In order to evaluate this hypothesis, computer simulations have been run with a population of 20 agents. AA agents change their willingness during runtime, whereas agents with static autonomy do not. The initial simulations that were performed [6] gave the results depicted in Figure 2. The interpretation of the graph is as follows. On the x-axis there are different combinations of willingness to interact, given by the tuple $\langle willingness\ to\ give\ help\ (\delta),\ willingness\ to\ ask\ for\ help\ (\gamma)\rangle$ and chosen based on previous work [4], whereas on the y-axis there

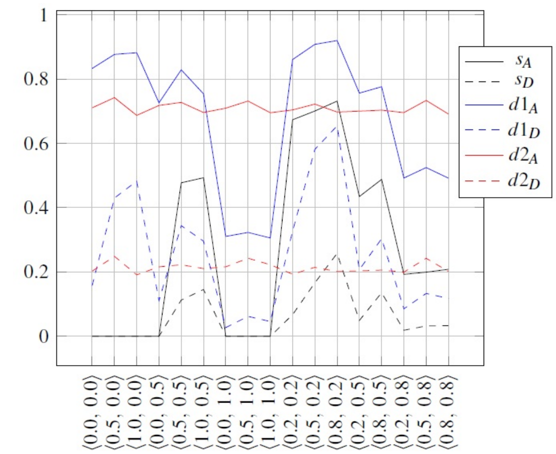

Figure 2: Comparison of performance between static and adaptive autonomy [6].

is completion rates for the total amount of tasks completed, and the dependent tasks completed. Dependent tasks refer to tasks that need help from other agents. There is a probability of 0.2 that a task should become a dependent task. The legend is interpreted as follows: sA - completion of all tasks in the static case, sD - completion of dependent tasks in the static case, d1A - completion of all tasks in the dynamic case where the willingness to interact is always updated from the same initial values, d1D - completion of dependent tasks in the dynamic case where the willingness to interact is always updated from the same initial values,d2A - completion of all tasks in the dynamic case where the willingness to interact is updated based on the previous calculated values, d2D - completion of dependent tasks in the dynamic case where the willingness to interact is updated based on the previous calculated values. It can be observed that the best results with respect to dependent tasks completed are achieved in the dynamic case – where the willingness is always updated from the same initial values – for the initial configurations $\langle 0.8, 0.2\rangle$, $\langle 0.5, 0.2\rangle$, and $\langle 1.0, 0.0\rangle$ [1]. If more tasks are dependent then the performance will degrade in all cases, following similar patterns [6].

5 CONCLUSION

In this research, a high-level agent model and the concept of willingness to interact have been proposed in order to allow for adaptive autonomous behaviour in (software) agents. The focus of the work is to develop a mathematical framework for the calculation of the willingness to interact, based on potential relevant factors. This model targets explicitly interaction and cooperation mechanisms in multi-agent systems, as compared to models such as belief-desire-intention (BDI) [7]. Nevertheless, it needs to be investigated how it compares to these models in different scenarios. Thus, future work will focus on: (i) the further development of the model, (e.g. provide a finer specification of the auxiliary states), and (ii) comparisons with other agent models in appropriate settings.

[1] Note that, in the static case for $\gamma = 0.0,\ or\ 1.0$ and $\delta = 0.0$, no simulations were run. In these cases no dependent tasks are completed because either agents never ask or give help, or all of them are always asking for help without completing any task. In these cases, values for s_A are represented with 0.

REFERENCES

[1] Jonathan Brookshire, Sanjiv Singh, and Reid Simmons. 2004. Preliminary results in sliding autonomy for assembly by coordinated teams. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, Vol. 1. IEEE, 706–711.

[2] Cristiano Castelfranchi. 2000. Founding agent's 'autonomy' on dependence theory. In *Proceedings of the 14th European Conference on Artificial Intelligence*. IOS Press, 353–357.

[3] Terrence Fong, Charles Thorpe, and Charles Baur. 2001. *Collaborative control: A robot-centric model for vehicle teleoperation*. Vol. 1. Carnegie Mellon University, The Robotics Institute.

[4] Mirgita Frasheri, Baran Çürüklü, and Mikael Ekström. 2016. Towards Collaborative Adaptive Autonomous Agents. In *9th International Conference on Agents and Artificial Intelligence 2017 ICAART, 24 Feb 2017, Porto, Portugal*.

[5] Mirgita Frasheri, Baran Çürüklü, and Mikael Ekström. 2018. Analysis of Perceived Helpfulness in Adaptive Autonomous Agent Populations. *LNCS Transactions on Computational Collective Intelligence* (2018).

[6] Mirgita Frasheri, Baran Çürüklü, and Mikael Ekström. 2018. Comparison Between Static and Dynamic Willingness to Interact in Adaptive Autonomous Agents. In *Proceedings of the 10th International Conference on Agents and Artifical Intelligence 2018 ICAART*.

[7] Michael Georgeff, Barney Pell, Martha Pollack, Milind Tambe, and Michael Wooldridge. 1998. The belief-desire-intention model of agency. In *International Workshop on Agent Theories, Architectures, and Languages*. Springer, 1–10.

[8] Benjamin Hardin and Michael A Goodrich. 2009. On using mixed-initiative control: A perspective for managing large-scale robotic teams. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM, 165–172.

[9] Matthew Johnson, Jeffrey M Bradshaw, Paul J Feltovich, Catholijn M Jonker, Birna Van Riemsdijk, and Maarten Sierhuis. 2011. The fundamental principle of coactive design: Interdependence must shape autonomy. In *Coordination, organizations, institutions, and norms in agent systems VI*. Springer, 172–191.

[10] Cheryl Martin and K Suzanne Barber. 2006. Adaptive decision-making frameworks for dynamic multi-agent organizational change. *Autonomous Agents and Multi-Agent Systems* 13, 3 (2006), 391–428.

[11] Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans* 30, 3 (2000), 286–297.