Computational Accountability and Responsibility in the MAS Domain

Stefano Tedeschi

Università degli Studi di Torino - Dipartimento di Informatica, Torino, Italy
stefano.tedeschi@unito.it

Abstract. The aim of my PhD is to investigate the notions of *computational accountability* and *responsibility* in multiagent systems, and in particular in multiagent organizations. The main objective is to develop both a conceptual framework and a programming platform providing an explicit representation of the two concepts and allowing their usage to simplify the design and development of such systems.

Advisor: Matteo Baldoni

Keywords: Accountability · Responsibility · Multiagent Systems.

1 Introduction

Multiagent Systems (MAS) provide useful abstractions for the development of complex (distributed) systems characterized by multiple individuals that execute, interact and coordinate with each other to achieve individual or collective goals. In these kinds of systems software components are represented as *goal-oriented autonomous agents*, which act in a shared and possibly distributed *environment* and need to *coordinate* between each other. *Organizations* represent strategies to decompose complex goals into simpler sub-goals and to allocate them to roles which are then played by agents. Furthermore, *norms* allow to enrich this structure with a notion of *social expectation*. However, in open systems, each autonomous agent has its own goals and capabilities to accommodate in the global organizational picture. This result can be achieved by establishing explicit relationships, directed from one agent to another, that reflect the legitimate expectations the second principal has of the first. Such relationships can be then used as a basis (i) to effectively decompose the organizational goals and assign tasks to agents and (ii) to determine who should answer when an unexpected outcome is detected.

On this foundation, wishing to voice my own contribution, the aim of my PhD is to investigate the notions of *computational accountability* (see [1]) and *responsibility* as tools to represent and model social relationships in multiagent systems, and especially in multiagent organizations. The overall objective is to develop both a conceptual framework and a programming platform to simplify the design and the development of accountability and responsibility supporting MAS organizations.

2 Motivations and context

To face the inherent need of coordination among agents, the organizational metaphor [11] has been used for a long time in MAS to define a structure of roles and relations through which tasks are distributed among agents. A second generation of organization models (see, e.g., [12, 14]) introduced norms in the structure of roles and tasks, enriching it with the notion of *social expectations*. However, while addressing the requirement of assigning duties and rights to agents, agent organizations have been criticized (see e.g., [8, 10]) because they obfuscate accountability. As an illustration let's consider a simple organization that generates obligations to perform the actions necessary to build a house to some participating agents. If the goal of building a wall were wrongly assigned to a painter agent, would it be licit to sanction him/her for not having fulfilled the obligation? Maybe the agent is not even capable of building walls, or this is not in agreement with his/her own internal goals. Agent autonomy, indeed, demands a different way of conceptualizing coordination in terms of responsibilities that are explicitly taken on by them, and by establishing directed accountability relationships, from one agent to another, that reflect the legitimate expectations the second principal has of the first, and that are grounded in the notion of control the first agent has over a situation.

In general, one might see accountability as the assumption of responsibility for decisions and actions that a principal has towards another party. Different research communities have dealt with the topic of accountability in software systems. More specifically, in the MAS community, Chopra and Singh see accountability as an explicitly established context-specific relationship between two parties [9]. Burgermeestre and Hulstijin, in turn, focus on the entire process of accountability determination [7]. Nevertheless, a complete model of accountability and of how accountability relationships are created and evolve is still missing. Concerning responsibility, [13] sees it as a unique charge assigned to an agent. Accountability, in turn, specifies the proper way to discharge the given responsbility. This view is compatible with the triangle model of responsibility [16] according to which the term bears two main understandings: one amounting to causation, the other to answerability. In this view, accountability is a pyramid, that comes into being when an accredited public observes a responsibility triangle. However, none of the currently proposed approaches provide a precise information model that links the concepts to how the environment evolves.

3 Contribution

The aim of my PhD will be to characterize the notions of accountability and responsibility in software systems, especially in multiagent organizations. By these terms I mean the realization via software of the abilities to trace, evaluate, and communicate accountability and responsibility. In particular, the problem will be addressed (i) by supplying a formal model and definition of computational accountability, clarifying its relation with the sibling notion of responsibility,

and (ii) by providing both modeling and computational tools which simplify the realization of accountability and responsibility supporting agent organizations. The purpose is to come up with a formalization of the two concepts as first-class entities which can be used both by the designer to describe the expected behavior of the system and by the actual agents to direct their own conduct.

The first part of my project is, and will be, focused on the development of a methodology and framework for the design and development such organizations. The construction of a comprehensive system requires many different elements: a formal model of accountability and responsibility, an engine able to properly distribute responsibilities, an automated forum able to discern the accountability of all the involved agents, and a mechanism that keeps track of who could be accountable for what and in which situation. The second part of the project will be devoted to the development of an actual programming platform implementing the previously mentioned framework. The idea is to develop a library to be used to integrate accountability and responsibility support in some of the most widely used frameworks for programming MAS organizations. One platform that seems particularly promising is JaCaMo [6], a conceptual model and platform built on the top of three existing platforms, namely Jason, CArtAgO, and Moise. The choice of JaCaMo is due to the fact that it provides a very good integration of the concepts which carachterize agents, environments and organizations, both from a conceptual and from a computational point of view.

4 Preliminary results

The project has found a first realization in the ADOPT (Accountability-Driven Organization Programming Technique) protocol for creating and manipulating accountability relationships. Technically, the core of the proposal builds upon the notion of role and the action of *role adoption*, on one side, and on the concept of *social commitment* [18], on the other side. An early version of ADOPT has been presented in [3, 2]. The main intuition is that, when an agent wants to play a role in an organization, it has to explicitly accept all the accountability requirements associated with the role itself, expressed as social commitments. The protocol specifies the shapes of these commitments and controls their creation. ADOPT has been, then, refined in [4] by clearly separating the *role adoption* phase from the *goal agreement* phase. Moreover, as a case study, an extension of the JaCaMo framework that implements the protocol has been developed.

At the same time, in [5], an information model that describes which data should be available, together with their relationships, in order to identify accountabilities in a group of interacting parties has been proposed. The model refines the characterization of accountability traced with ADOPT, by identifying the main concepts that come into play in the process of accountability determination, such as *mutually held expectation* and *control*. The model is provided in Object-Role Modeling[1] (ORM) due to the relational nature of the represented concepts.

[1] http://www.orm.net/pdf/ORMwhitePaper.pdf

5 Evaluation plan

The conceptual model and platform described above will be evaluated in two ways, as follows. Regarding the formal model of accountability and responsibility, the proposal will be compared to the main approaches to accountability from other areas, such as social sciences, economics, public administration, psychology, and so on. Accountability is considered a central concept in many fields that study human interaction and the very same abstractions used in these fields could be effective in the context of intelligent systems. Accountability is rather a complex concept, with many different declinations; the aim is to propose a characterization of it able to capture as many of them as possible. Moreover, the proposal will be compared with other widely accepted models of coordination, such as the ones based on norms and the ones based on social commitments.

For what concerns the programming platform, the plan is to evaluate it inside the JaCaMo framework. I plan to use the platform to model real-world scenarios in order to show the benefits coming from an explicit representation of accountability and responsibility in both the design and execution phases of the systems. A particularly promising source of practical use cases is the business processes area. Actually, multiagent organizations can be effectively used to represent complex business processes with several interacting components. Concretely, the plan is to show how an explicit representation of accountability and responsibility relationships, with a well-defined semantics, can simplify both design execution, by addressing some weaknesses that emerge in scenarios which proved to be particularly challenging to model. One emblematic case is the one-to-many pattern of interaction as proposed by Silver [17], in the Hiring Process scenario.

6 Conclusions and future work

Agents in a social state can influence the environment and the agents around them. Moreover, in open systems, agents in an organization consider it as a means for reaching their own goals. These goals, however, may not coincide with the organizational ones. The explicit representation of the relationships created between the agents and between agents and organizational goals can strongly increase the robustness of such systems (i) by supporting the decomposition and distribution of the organizational goal(s) and (ii) by providing a mechanism to find who should answer in a given situation of interest. My aim, with my PhD project, is to model these social relationships relying on the notions of accountability and responsibility and use them to simplify the design and development of the resulting systems.

Future work includes, but is not limited to, a further refinement of the formalization of accountability and responsibility relationships for computational use. This includes, for instance, the integration of the model in the well-known organizational specification of the Moise [15] language. A big effort will be then devoted to the realization of the actual programming platform for computational accountability and responsibility support in MAS organizations.

References

1. Baldoni, M., Baroglio, C., May, K.M., Micalizio, R., Tedeschi, S.: Computational accountability. In: Proc. of URANIA 2016 workshop co-located with AI*IA 2016. CEUR Workshop Proceedings, vol. 1802, pp. 56–62 (2016)
2. Baldoni, M., Baroglio, C., May, K.M., Micalizio, R., Tedeschi, S.: ADOPT JaCaMo: Accountability-driven organization programming technique for JaCaMo. In: Proc . of PRIMA 2017. LNCS, vol. 10621, pp. 295–312. Springer (2017)
3. Baldoni, M., Baroglio, C., May, K.M., Micalizio, R., Tedeschi, S.: Supporting organizational accountability inside multiagent systems. In: Proc. of AI*IA 2017. LNCS, vol. 10640, pp. 403–417. Springer (2017)
4. Baldoni, M., Baroglio, C., May, K.M., Micalizio, R., Tedeschi, S.: Computational Accountability in MAS Organizations with ADOPT. Applied Sciences **8**(4) (2018)
5. Baldoni, M., Baroglio, C., May, K.M., Micalizio, R., Tedeschi, S.: An information model for computing accountabilities. In: Proc. of AI*IA 2018. LNCS, Springer (2018), To appear
6. Boissier, O., Bordini, R.H., Hübner, J.F., Ricci, A., Santi, A.: Multi-agent oriented programming with JaCaMo. Sci. of Computer Programming **78**(6), 747–761 (2013)
7. Burgemeestre, B., Hulstijn, J.: Handbook of Ethics, Values, and Technological Design: Sources, theory, values and application domains, chap. Designing for Accountability and Transparency: A value-based argumentation approach. Springer (2015)
8. Chopra, A.K., Dalpiaz, F., Aydemir, F.B., Giorgini, P., Mylopoulos, J., Singh, M.P.: Protos: Foundations for engineering innovative sociotechnical systems. In: IEEE 22nd Int. Req. Eng. Conf. pp. 53–62 (2014)
9. Chopra, A.K., Singh, M.P.: The thing itself speaks: Accountability as a foundation for requirements in sociotechnical systems. In: 7th IEEE Int. Workshop RELAW. pp. 22–22 (8 2014)
10. Chopra, A.K., Singh, M.P.: From social machines to social protocols: Software engineering foundations for sociotechnical systems. In: Proc. of the 25th Int. Conf. on WWW. pp. 903–914 (2016)
11. Corkill, D.D., Lesser, V.R.: The use of meta-level control for coordination in distributed problem solving network. In: Proc. of IJCAI 1983. pp. 748–756 (1983)
12. Dignum, V., Vázquez-Salceda, J., Dignum, F.: Omni: Introducing social structure, norms and ontologies into agent organizations. In: ProMAS. vol. 4, pp. 181–198. Springer (2004)
13. Feltus, C.: Aligning Access Rights to Governance Needs with the Responsability MetaModel (ReMMo) in the Frame of Enterprise Architecture. Ph.D. thesis, University of Namur, Belgium (2014)
14. Fornara, N., Viganò, F., Verdicchio, M., Colombetti, M.: Artificial institutions: a model of institutional reality for open multiagent systems. Artificial Intelligence and Law **16**(1), 89–105 (2008)
15. Hubner, J.F., Sichman, J.S., Boissier, O.: Developing organised multiagent systems using the MOISE+ model: Programming issues at the system and agent levels. Int. J. Agent-Oriented Softw. Eng. **1**(3/4), 370–395 (2007)
16. Schlenker, B.R., Britt, T.W., Pennington, J., Rodolfo, M., Doherty, K.: The triangle model of responsibility. Psychological Review **101**(4), 632–652 (October 1994)
17. Silver, B.: BPMN Method and Style, with BPMN Implementer's Guide. Cody-Cassidy Press, Aptos, CA, USA, second edn. (2012)
18. Singh, M.P.: An ontology for commitments in multiagent systems. Artificial Intelligence and Law **7**(1), 97–113 (1999)