A Formal Model of Responsibilities in Agent-Based Teams

J. Colby Johnson, Thomas R. Ioerger
Department of Computer Science
Texas A&M University
College Station, Texas 77843-3112
{jcj0650, ioerger}@cs.tamu.edu

ABSTRACT

In this short paper, we describe the foundations of responsibility in MAS and present a formalism of responsibility for use in future work.

1. INTRODUCTION

Whether implicit or explicit, responsibility plays a large role in almost any teamwork environment in which humans participate. It is one of the foundational components of many teamwork and communication paradigms [2]. Thus, a complete understanding and formal model of responsibility will be of utmost importance in team structures composed of collaborative, rational agents. For this reason, we present a definition of responsibility based upon its two most overt aspects: persistence and dependence. We argue that responsibility can be reduced to constraints on agent beliefs and goals, though with different properties than intentions [1].

2. FOUNDATION

Responsibility is very similar to intention and, like intention, it can be decomposed into agent beliefs and goals. However, there are two important aspects of responsibility that set it apart from intention: persistence and dependence. The former defines the constraints dealing with the adoption and abandonment of goals, while the latter details not only the relationship between two individuals sharing a responsibility, but also the need for communication between them.

From the perspective of an individual agent, responsibility entails a certain level of persistence that is not present in intention. This persistence is the same as seen in human domains; that all plausible paths to satisfaction must be exhausted before an individual can concede that the responsibility cannot be fulfilled. However, this definition alone does not fully capture the notion of persistence as it pertains to responsibility. With responsibilities, individuals may transfer (delegate) their commitment to another; however, they must remain persistent in that they retain a passive form of their original commitment (one not requiring immediate action). This type of persistence is not seen in intention as the latter is only defined within the scope of a single agent and does not consider transference of commitment, which is common when extending to teams of agents.

Dependence is the second aspect of responsibility and it deals primarily with the relationships defined between agents participating in a responsibility. Fundamentally, responsibility is a relationship between agents indicating that one agent requires a certain condition to be satisfied, and that another agent is assisting by agreeing to help satisfy that condition. Such a relationship imposes certain constraints and necessities concerning the informational dependence upon one another or the need for communication between the participants in the responsibility. Yet while dependence serves a specific purpose with respect to agent communication, it also plays a large role in belief reasoning about teamwork. Knowledge of this dependence and belief about the responsibilities of others allows agents to abstract or narrow the scope of their concern. For example, an agent will generally not concern itself with a responsibility that even though it needs fulfilled, is being satisfied by a different agent. The dependence that responsibility imposes on the two agents guarantees that the agent will receive knowledge concerning the status of the other agent's progress. This allows teams of agents to continue to act in the face of uncertainty based on the belief that, in all probability, each required responsibility will be fulfilled by some agent that possesses it.

3. FORMALISM

3.1 Partial Order

In order to create a foundation with which to formalize the mental state of responsibility, we employ the use of modal logic utilizing possible world semantics. We extend the basic Kripke structure by adding a partial order for encoding delegation hierarchies. Using this, we present definitions for two types of responsibility that are useful for distinguishing meaning: indirect and direct responsibility.

For minimal agents in the order, those that have a superior but no subordinate, we introduce the concept of direct responsibility. All agents that have direct responsibility are required to act in order to satisfy that responsibility. Directly responsible agents can be viewed as the actuators whose responsibility demands action and whose actions attempt to somehow satisfy the responsibility.

For agents that participate in the responsibility but are not directly responsible, we term their commitment as being indirect responsibility. Agents with indirect responsibility were, at some earlier point in time, directly responsible and have since passed direct responsibility on to another agent. These agents, while not required to act, are not completely absolved of their commitment and may be called upon to act at some point in the future should the need arise. It is possible that through a series of failures, they may be called upon to be directly responsible again at some point in the future. This mechanism leads to robustness of the team behavior with respect to the fulfillment of the responsibility.

With a model defined, all that remains is to provide axioms defining the relationships between the agents participating in the responsibility, rules of conduct, and an operator with which to construct partial orders.

3.2 Delegation

In order to allow agents to transfer responsibility, we introduce the delegate operator. Delegation is the primary event that is derived from teamwork and responsibility in team structures [3]. This operator allows an agent to (1) pass direct responsibility to another agent and (2) modify its own responsibility to that of indirect. The implementation of delegation itself is simple; a new minimal element is inserted into the partial order, thus modifying the responsibility of certain agents within the order. This operator is significant due to the fact that it allows the fulfillment of the responsibility to be qualified over a set of agents (a team) rather than within the scope of a single agent. This allows for more alternatives with which an agent can fulfill its commitment as compared to intention, where a single agent can only satisfy a goal itself or alternatively determine that it is unachievable.

Delegation plays a large role in the representation of persistence in responsibility. It allows an agent an entirely new set of alternatives with which to satisfy its responsibility, and thus empowers agents to be more effective in a team environment by introducing the possibility for load balancing and capability maximization (the most qualified agents performing a task). We formalize the role of delegation in persistence as a responsible persistent goal (RPG) in the following manner:

DEFINITION: Responsible Persistent Goal

$$\mathrm{RPG}(A,\theta) \equiv \mathrm{PGoal}(A, \mathrm{Done}(A,\theta) \vee$$
$$[\exists B\ (\mathrm{canDelegate}(A,B,\theta) \wedge \mathrm{Done}(\mathrm{doDelegate}(A,B,\theta)))])$$

This definition states that an agent with a RPG will have a persistent goal (i.e. commitment) [1] to either satisfy the condition itself or delegate it to another agent. It is important to note that the definition for canDelegate (not shown here for brevity) includes reasoning about the capability of the potential recipient for satisfying the responsibility should it receive it. This prevents the introduction of “hot potatoe” into the team domain where agents simply off-load a responsibility to another agent as soon as they receive it, without any concern for its completion.

3.3 Mutual Belief

We note that the relationships in the partial order, while being authoritatively asymmetric, are communicatively symmetric. This means that while one agent may have authority over the other, both must actively participate in keeping the other informed of the status and any changes relating to the responsibility. This is achieved through the establishment of mutual belief between the two agents. Each agent maintains mutual belief with its subordinate and superior pertaining to the need for information pertinent to the commitment and the status of the goal. In this sense, a chain of mutual belief is constructed using the partial order, allowing facts and information to be propagated throughout the partial order through the maintenance of mutual belief. We incorporate the requirement for maintenance of mutual belief through the establishment of an accountable responsible persistent goal (A-RPG), which we define as follows:

DEFINITION: Accountable Responsible Persistent Goal

$$\mathrm{A\text{-}RPG}(A,B,\theta) \equiv \mathrm{RPG}(A,\theta) \wedge \mathrm{BEFORE}(\neg\mathrm{RPG}(A,\theta),$$
$$\mathrm{PGoal}(A, \mathrm{MB}(\{A,B\}, \neg\mathrm{RPG}(A,\theta) \wedge \mathrm{status}(\theta))))$$

This concept of maintained mutual belief is important to responsibility as it helps represent both the dependence and the persistence that are inherent in the mental state. Any agent who participates in the responsibility is aware of the dependence of its subordinates and superiors, and is also aware of the need to maintain mutual belief as to the status of the responsibility. This can be used to derive formal conditions under which agents must communicate; if communicative acts are the only way of affecting another agent’s beliefs, the agents become obligated to communicate whenever mutual beliefs become inconsistent.

4. FUTURE WORK

We are using this formal framework to provide the semantics behind responsibilities in our team-agent architecture, CAST [4], especially to derive methods for coordination and communication among agents to produce correct and robust team behavior.

5. REFERENCES

[1] Philip Cohen and Hector Levesque. Intention is Choice with Commitment. *Artificial Intelligence*, 42: 213–261, 1990.

[2] N.R. Jennings. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence*, 75: 195–240, 1995.

[3] T.J. Norman and C. Reed. Delegation and Responsibility. In *Proc., Seventh International Workshop on Agent Theories, Architectures, and Languages*, 2000.

[4] J. Yin and M. Miller and T. Ioerger and J. Yen and R. Volz. A Knowledge-Based Approach for Designing Intelligent Team Training Systems. In *Proc., Fourth International Conf. on Autonomous Agents*, pages 427–434, 2000.