Prospects for E-Collaboration with Artificial Partners

Kathleen Keogh
The University of Ballarat, Australia

Liz Sonenberg
The University of Melbourne, Australia

INTRODUCTION

Recent work shows that there is interest in how individual artificial agents can work in successful competitive and collaborative teams including people and other agents. Applications involving competing agents include online auctions. Applications for collaborative teams include remote space missions, disaster recovery (e.g., to coordinate a rescue mission) and helping organize appointments for a team of people (Pynadath & Tambe, 2003); as an aid to independent living developing teams of health carers, including artificial carers (Wagner, Guralnik, & Phelps, 2002); in command and control as coordination and communication assistants (Fan et al., 2005); and pedagogical agents in teaching systems (e.g., Shaw, Ganeshan, Johnson & Millar, 1999; Feng, Shaw, Kim & Hovy, 2006).

There is considerable interest in developing multi-agent systems and teams involving people and/or artificial agents collaborating to achieve a common goal. This article will outline some current issues in software agent design with respect to team communication, coordination, and sharing of team situation awareness regarding the current state of a dynamic world. Topics covered include:

- How might artificial agents be defined?
- What mechanisms are helpful to enable agents to form and work in teams in a dynamic world?

Artificial agents are computer programs that have some inbuilt attributed human-like "intelligence" and that operate autonomously with some ability to choose whether or not to perform a task. Wooldridge defines an agent as "a computer system that is capable of independent action on behalf of its user or owner" (Wooldridge, 2002, p, 3). Such agents are able to reason dynamically and make decisions regarding tasks to be completed, so the solution is not preprogrammed in a deterministic way. Agents are usually programmed in terms of defining goals, tasks the agent is capable of performing, and how the agent should react or interpret the data about the world that is available to it. Multi-agent systems (MAS) comprise a set of agents interacting: involving cooperation, coordination and negotiation (Wooldridge, 2002)

BACKGROUND

How are Artificial Agents Defined?

Typical agent attributes to include and inform design include: beliefs, desires, and intentions (BDI). The BDI architecture (Rao & Georgeff, 1991) is not the only model for agency used, however due to space limitations, it is the only model discussed in this article. It is not uncommon to also find discussion and definition of agent roles, responsibilities, obligations, trust, commitment, and protocols for communication and negotiation in agent systems research literature.

When defining an agent, the following agent attributes are defined:

- **Beliefs** about the world encoded as a database of defining attributes and values that are accessible in some way to the agent
- **Desires** or goals that the agent is trying to satisfy
- **Intentions** adopted plans of action that have been chosen in order to achieve a goal.

The designer/programmer empowers the agent with a set of predetermined plans of action and the agent reasons based on predefined factors such as expected utility or specific preferences, to choose a plan to execute. The

plan is essentially a predefined script that outlines a series of actions or sub-goals to be performed in order to satisfy a goal. The programmer creates a library of potential plans based on the domain expertise and how human experts would behave. Roles can be predefined as a set of goals and objectives—responsibilities that must be met by the agent accepting that role. In some systems roles are more explicitly defined in terms of hierarchical positions and expected behavior (Zambonelli, Jennings, & Wooldridge, 2003).

Agent reasoning is often non-monotonic, that is, the world is dynamically changing, so beliefs currently held may not be true in the future. Agent programming languages are generally created to enable agents to dynamically change their plans—for example, if a goal is no longer relevant, it should be ignored. Allowing agents to collaborate requires a meta level of additional self-knowledge in the agent to enable agents to negotiate. Agents need to know and possibly negotiate around their adopted roles and what actions they are capable of performing. An agent role can be defined statically at design time—in terms of goals to be performed or the role might be more flexible and negotiated dynamically—to enable more flexible and adaptive team reorganization at run time. Providing the infrastructure to enable an agent to be more flexible and to enable the reorganization of teams requires a more sophisticated agent design than the BDI approach of itself provides and more resources. According to the domain and level of sophistication and reorganization needed, the decision to "keep it simple," or to include more complicated structures is a trade off between flexibility and extra resources and structure required.

Agent Models for Cooperation

An additional component in the agent model that is made explicit by Griffiths and colleagues is an agent's motivation (Griffiths, Luck, & d'Inverno, 2003). They describe motivation in terms of intensity, a threshold for when it applies and functions for goal generation. Agents can chose from a library of partial plans. The motivation component provides a utility measure that an agent can use to decide upon their adopted plans.

Agents can be self-interested or collaborative. Collaborative agents have *obligations* toward other agents in their collaboration team. Self-interested agents make decisions and take actions only in their own interest. When agents have goals that they cannot achieve alone, or that can be achieved more efficiently by sharing the workload, agents can be motivated to collaborate with other trusted agents to work together on a joint plan (Grosz & Kraus, 1999). When an agent receives conflicting requests from other agents, it has been suggested that these can be resolved by using explicit knowledge regarding relationships existing between agents and agent roles (McCallum, Norman, & Vasconcelos, 2004). Another approach to resolving conflict is to permit agents to engage only in group activities that don't conflict with personal goals (Findler & Malyankar, 2000).

There are different approaches to how agents may communicate and form into groups to work together: central command and control, dynamically allocated teams created "on the fly," negotiated team membership—by invitation and commitment. In some multiagent systems, the team membership is agreed upon prior to beginning a collaborative task, other models allow for the team to be formed at the time of need and commitment prior to that time is not required. Team formation and coordination issues are discussed further in the next section.

INFRASTRUCTURES FOR AGENT TEAMWORK

Formation of Agent Groups: Team Membership, Team Plans, Motivation to Join In

Institutional organizations are collections of (human) agents that have roles, rights, and obligations. Multiagent systems are collections of agents that interact within a dynamic situation. Agents are *situated* in the world and have available information via sensors or other input mechanisms to inform them about the current state of the world. When agents begin to collaborate and interact, they need an awareness of the limitations over time and space of their current knowledge, and an ability to share with other team agents possibly different information about the world.

One view of situation awareness (SA) is described with three main elements: (Endsley, Bolte, & Jones, 2003)

1. Perception of the environment in time and space