Planning and Acting Together *

Barbara J. Grosz
Luke Hunsberger
Division of Engineering and Applied Sciences
Harvard University, Cambridge, MA 02138 USA
{grosz,luke}@eecs.harvard.edu

Sarit Kraus
Department of Mathematics and Computer Science
Bar-Ilan University, Ramat Gan 52900 Israel, and
Institute for Advanced Computer Studies
University of Maryland, College Park, MD 20712 USA
sarit@cs.biu.ac.il

May 14, 1999

Abstract

People often act together with a shared purpose; they collaborate on a group action or activity. An increasing number of computer applications also require collaboration among various systems and users. The plans for such collaborative activities must be formed with others, not in isolation. Groups may persist over long periods of time (as do orchestras, sports teams, and systems administration groups), form spontaneously for a single group activity (as when a group forms for a programming project), or come together repeatedly (as do surgical teams and airline crews).

A major challenge for researchers in Artificial Intelligence is to determine ways to construct computer systems that are able to act effectively as collaborative team members. Collaborative activities require more than the sum of individual plans. Participants must form commitments not only to the group action itself, but also to the activities of other participants that are in service of this group activity.

*This research has been supported in part by National Science Foundation grants IRI 95-25915, IRI-9618848 and CDA-94-01024. The authors thank David Sullivan for his helpful comments on earlier drafts and for his assistance with the section on empirical studies.

Group decision-making processes are required to expand partial plans to more complete ones. Furthermore, when conflicts arise from resource bounds, participants must weigh commitments to group activities against those for individual activities. This article briefly reviews the major features of one model of collaborative planning, SharedPlans (Grosz and Kraus, 1996, 1999), describes several current efforts to develop collaborative planning agents and systems for human–computer communication based on this model, and discusses empirical research aimed at determining commitment strategies in the SharedPlans context.

Keywords: Collaboration, Multi-agent planning, Intention-conflict resolution, SharedPlans.

Introduction

Pollack (Pollack, 1992, 1998) has argued that "there's more to life than making plans." Planning agents inhabit worlds that are constantly changing, worlds in which actions have uncertain outcomes. They must be able to adapt to these changes and uncertainties. Planning most often must be done incrementally and must be interleaved with execution. Thus, planning is only one of many "plan-management processes"; agents must be able to elaborate partial plans into more complete ones, manage commitments and monitor their environments effectively, assess alternative plans and possibilities, and coordinate with other agents (Pollack, 1998; Pollack and Horty, this issue).

There's not only more to life than making plans, there's more to managing plans than managing them alone. In a multi-agent context, agents do more than coordinate individual plans. Group activities require a variety of group decision-making and planning processes. Not only must the plan-management processes agents use for their individual plans operate in multi-agent contexts, but also agents must have group plan-management processes for plan evaluation, plan elaboration, decision making and commitment monitoring. Although these group processes have individual-plan analogues, their operation is necessarily complicated by the multi-agent context.

This article describes research focused on a particular kind of distributed, continual plan management—namely, plan management in the context of multi-agent collaboration. Collaboration is distinguished from other kinds of multi-agent interaction (e.g., cooperation, coordination, and competition) by the existence of a shared goal and the agents' commitments to that shared goal. Our work attempts to identify the capabilities needed by agents if they are to plan and act collaboratively. We have developed a formalization of collaboration, called SharedPlans (Grosz and Kraus, 1996, 1999), which provides both a specification for the design of collaboration-capable agents and a framework for research. A recent paper (Kraus and Hadad, 1999) presents several examples illustrating that the use of SharedPlans leads to better utilization of resources, better coordination of tasks, and higher chances of fulfilling agents' goals. The particular aspects of SharedPlans that contribute to these improvements include their provision for agents to interleave planning and acting, the inclusion of commitments that can lead agents to behave helpfully, and constraints that prohibit adoption of conflicting intentions.

We are currently developing three systems based on the SharedPlans theory: the GigAgent multi-agent system for collaborations of hetereogeneous groups of people and computer systems, the WebTrader system for electronic commerce, and the DIAL system that provides a collaborative interface for distance learning.

The GigAgent system has a general architecture that supports participation in several different kinds of collaborative activities involving teams consisting of both people and computer systems. The name "GigAgent" was motivated by a sample application involving a group of human musicians who form SharedPlans to perform musical gigs. Figure 1 provides a schematic of a GigAgent system that might support this kind of collaboration, i.e., one in which the main collaborative activity (playing the gig) is an action to be done by a group of people, but computer systems are involved in various supporting activities. In such settings, each person involved in the collaboration has a dedicated computer assistant. When information about a potential gig arises (e.g., over the internet or email), the group of computer agents could handle much of the work of determining whether or not the gig might be feasible and, if so, establishing the framework of a SharedPlan for doing that gig. For instance, computer GigAgents might handle such tasks as making arrangements for the group to rent a van or sending email confirming the gig contract.

Another type of GigAgent system is one in which a combined team of people and computer systems performs the overall collaborative activity. We are currently implementing the GigAgent system in a computer systems administration domain that requires this kind of organization (Figure 2). In this domain people perform such actions as installing new software, and systems perform such actions as running backup systems.

The WebTrader system (Kraus and Hadad, 1999) was designed to oper-

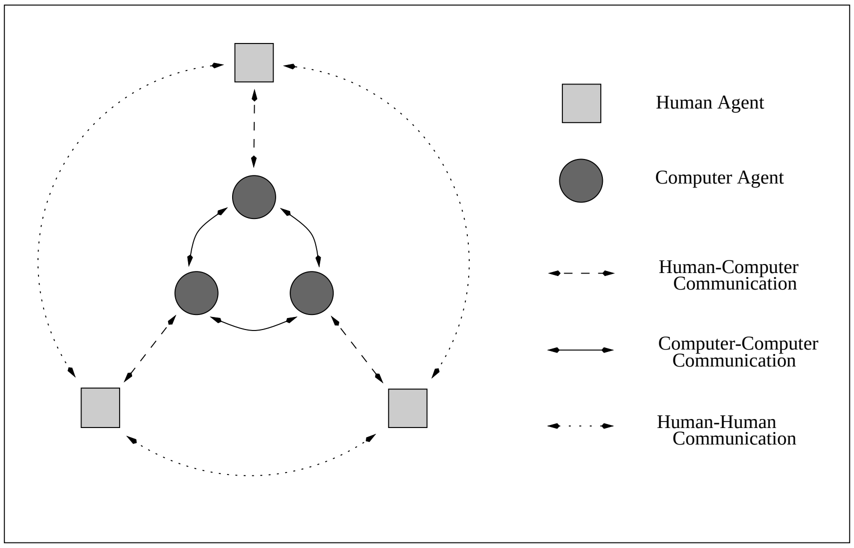

Figure 1: GigAgents for the Musical Gig domain

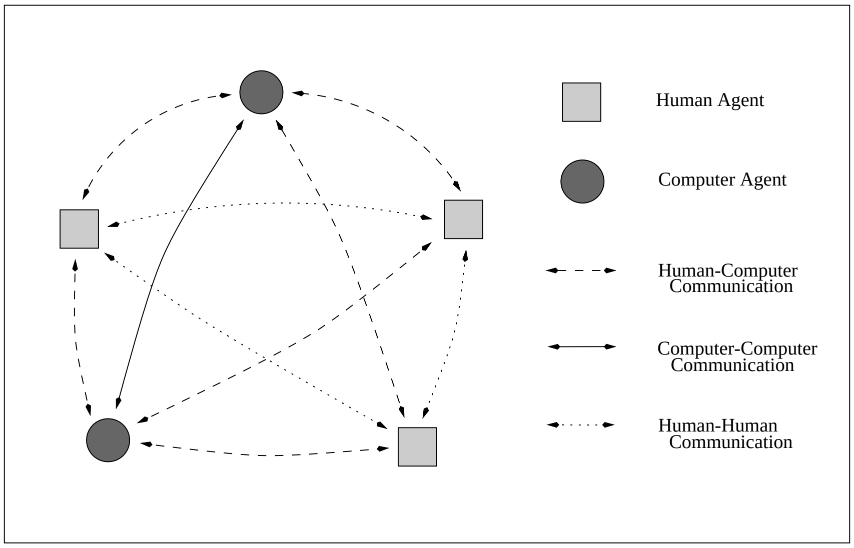

Figure 2: GigAgents for the Systems Administration domain

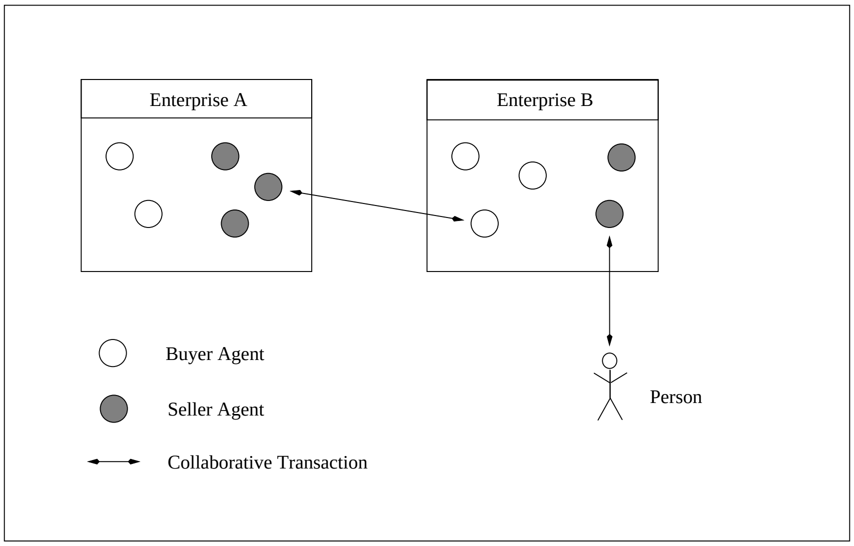

Figure 3: WebTrader: SharedPlans in an Electronic Commerce domain

ate in the Electronic Commerce domain. Figure 3 gives a schematic view of a typical WebTrader configuration. In this domain, there are several enterprises, each containing several buyer and seller agents. In addition, persons may interact with the seller agents of an enterprise. A collaborative transaction occurs when a buyer agent (or a person) and a seller agent form a SharedPlan for the buyer to purchase some item from the seller.

In both the GigAgent and WebTrader systems, we use SharedPlans not only explicitly to model the collaborations among the various agents, but also implicitly in the design of the requisite human–computer interfaces. The goal is to make the interfaces collaborative and, hence, more effective. In this effort, we are drawing on our experience in designing a collaborative interface, called DIAL, for a distance learning system.

The next sections of this article present brief overviews of the SharedPlans formalization and the basic processes required in collaborative systems based on it. Subsequent sections provide additional details of the GigAgent, WebTrader, and DIAL systems. The SharedPlans formalization also provides a framework for identifying and investigating fundamental questions about collaboration. The final section of this article describes empirical research on the problem of intention-reconciliation in a group context; this work was prompted by both the formalization and the implementation efforts.

The SharedPlans Formalization of Collaboration

Pollack (1990) argued that agents typically have different beliefs about the ways to do an action or to achieve some desired state and that, as a result, it was necessary to move away from a view of plans as primarily data-structure-like to a view of plans in terms of mental state. In this view, an agent has a plan to do some action if it has a certain set of beliefs and intentions pertaining to that action. Bratman (1987) also argued for a mental-state view of plans, emphasizing the roles of intentions in focusing means–ends reasoning, in constraining the other intentions an agent might subsequently adopt, and in guiding replanning.

The SharedPlans formalization of collaborative planning is based on a mental-state view of plans. It was originally conceived to provide the basis for modeling the intentional structure of discourse (Grosz and Sidner, 1990). We subsequently generalized the formalization to accommodate more than two participating agents and to support the construction of teams of collaboration-capable agents (Grosz and Kraus, 1996, 1999; Hunsberger, 1999). Lochbaum (1995, 1998) demonstrated the power of using SharedPlans as the basis for intentional structure in dialogue. Others adapted this use of SharedPlans to the design of graphically-based human–computer interfaces (Rich and Sidner, 1998; Ortiz et al., 1999). Tambe (1997) has incorporated some elements of SharedPlans in STEAM (a Shell for Teamwork) which he has used to build systems for Robocup Soccer tournaments and for military simulations and exercises. A comparison of the SharedPlans formalization with alternative computational accounts (Cohen, Levesque, and Nunes, 1990; Cohen and Levesque, 1991; Kinny et al., 1994) and philosophical theories (Searle, 1990; Bratman, 1992) may be found in our earlier papers (Grosz and Kraus, 1996, 1999).

The SharedPlans formalization treats complex actions and partial plans, models the possibility of agents contracting out actions to other agents, provides for the interleaving of planning and execution, and distinguishes the information needed by agents directly involved in the planning of some constituent action from the information the group at large must have about that action. The model presumes general procedures for reconciling conflicts among intentions and requires a variety of group decision-making processes.

The formalization distinguishes two kinds of intention: (1) intention *to* do an action and (2) intention *that* a proposition hold. An agent intending to do some action must not only be committed to doing that action, it must also have appropriate beliefs about its ability to do that action, and it must have some knowledge about how to do it or how to figure out how to do it. An agent intending that some proposition hold must be committed to doing what it can, if anything, to bring about the proposition, but it need not necessarily be able to do anything. Both forms of intention preclude agents (knowingly) adopting conflicting intentions. In SharedPlans, intentions-that are used to represent agents' commitments to their group activity, to specify the collaborative support participants offer one another, and to provide a basis for agents to form the mutual beliefs necessary for collaboration (Grosz and Kraus, 1999).

The mental-state view of plans is crucial to the SharedPlans formalization. It enables the coordination and integration of SharedPlans for group actions with individual plans for single-agent constituent actions of those plans, because the same kinds of beliefs and intentions play a role in each. Furthermore, in a multi-agent context, plans must be ascribable to agents and groups of agents—either theoretically, from an omniscient-observer perspective, or from the perspective of a fellow collaborator. Because ascribing a plan to an agent or group of agents depends on a determination that the relevant beliefs and intentions hold, specifying the conditions under which a plan may be so ascribed requires a mental-state view of plans. In addition, adopting the mental-state view has enabled the proof of several theorems relating agent knowledge and beliefs to the existence of plans (Hunsberger, 1999).

Plan Definitions

The SharedPlans formalization distinguishes partial plans and complete plans. A *Full SharedPlan* (FSP) is a complete plan in which agents have fully determined how they will perform an action. Because most of the time that agents are involved in a collaborative activity their plans are partial, the FSP definition serves as a target for their planning activity—agents know they must reach the mental states specified by the FSP definition to have a complete plan. The *Partial SharedPlan* (PSP) definition provides a specification of the minimal mental-state requirements for collaboration to exist and gives criteria governing the process of completing the plan.

Agents involved in a SharedPlan to do some action are required to have certain individual and mutual beliefs about how that action and its constituent subactions are to be done. These beliefs are specified in terms of *recipes.* A *(full) recipe* for doing some action A_α is a set of actions and constraints such that the doing of those actions under those constraints constitutes the doing of A_α. A *partial recipe* is a set of actions and constraints that may be extended to a full recipe.

Ultimately, actions must be performed by individual agents. Thus, SharedPlans contain individual plans of the participating agents as constituents. Extending the work of Pollack (1990), the SharedPlans formalization specifies that for an agent to have a *full individual plan* to do some (single-agent) action, it must have knowledge of a recipe for that action, believe that it will be able to perform the actions in that recipe, and have intentions to do those constituent actions. Because individual plans, like SharedPlans, are partial much of the time, the formalization also defines a *partial individual plan* in which an agent may have only a partial recipe for doing the action and may have only partial plans for doing the actions in that recipe.

For both individual and group plans, *elaborating* a partial plan into a full plan involves choosing a full recipe or extending (perhaps incrementally) a partial recipe. It also involves establishing subordinate plans for each subaction in the recipe. In a full plan, the group action has been fully decomposed into *basic actions* (i.e., actions that an individual agent may execute at will under appropriate conditions).

The Full SharedPlan (FSP) and Partial SharedPlan (PSP) definitions are stated informally in Figures 4 and 5.[1] The numbers in the figures indicate some of the essential elements of a SharedPlan: (1) commitment to the group activity, (2) commitment to the actions of others in service of the group activity, (3) the existence of and commitment to a group process for selecting a recipe for the group activity, and (4) the existence of and commitment to a group process for assigning agents or subgroups to constituent actions. As the definitions make evident, planning collaboratively introduces several complexities not present in the single-agent case. First, collaborating agents typically have access to different information and thus hold different sets of beliefs about actions. For instance, only the agents directly involved in the planning and execution of some subaction typically need to know the details of how that subaction will be done. Second, plan elaboration in a multi-agent context requires the assignment of agents or subgroups to the constituent individual and multi-agent actions (respectively) in the recipe. Third, agents must form commitments not only to doing their own actions, but also to supporting the actions of others and to supporting the group activity as a

[1] To simplify the presentation, single-agent groups are permitted. Formal definitions may be found elsewhere (Hunsberger, 1999).

For a group GR to have a Full SharedPlan (FSP) to do a complex action A_α requires that:

- GR mutually believe that each member agent intends that the group do A_α, (1)
- GR mutually believe that they have a full recipe for doing A_α, and
- each action in that recipe be **fully resolved**.

A basic action A_β is **fully resolved** if

- some agent G_β in GR intends to do A_β,
- GR mutually believe that G_β intends to do A_β and is able to do A_β, and
- GR mutually believe that each member agent intends that G_β be able to execute A_β. (2)

Similarly, a complex action A_κ is **fully resolved** if

- some subgroup GR_κ in GR have a full plan to do A_κ,
- GR mutually believe that GR_κ have a full plan to do A_κ and are able to do A_κ, and
- GR mutually believe that each member agent intends that GR_κ be able to execute A_κ. (2)

Figure 4: Informal definition of a Full SharedPlan (FSP)

For a group GR to have a Partial SharedPlan (PSP) to do a complex action A_α requires that:

- GR mutually believe that each member agent intends that the group do A_α, (1)
- either GR mutually believe that they have a full recipe for doing A_α or they mutually believe that they have a partial recipe that may be extended into a full recipe they can use to do A_α and they have a full plan to select such a recipe, and
- each action in the (possibly partial) recipe be either **at-least-partially-resolved** or **unresolved**.

A basic action A_β is **at-least-partially-resolved** if

- some agent G_β in GR intends to do A_β,
- GR mutually believe that G_β intends to do A_β, and
- GR mutually believe that each member agent intends that G_β be able to execute A_β and each agent in GR intends that GR mutually believe that G_β is able to execute A_β. (2)

Similarly, a complex action A_κ is **at-least-partially-resolved** if

- some subgroup GR_κ in GR have a plan to do A_κ,
- GR mutually believe that GR_κ have a plan to do A_κ, and
- GR mutually believe that each member agent intends that GR_κ be able to execute A_κ and each agent in GR intends that GR mutually believe that GR_κ are able to execute A_κ. (2)

A basic action A_ϵ is **unresolved** if

- GR mutually believe that *some* member of GR could do A_ϵ, and
- GR have a full plan to select such an agent. (3)

Similarly, a complex action A_μ is **unresolved** if

- GR mutually believe that *some* subgroup of GR could do A_μ, and
- GR have a full plan to select such a subgroup. (4)

Figure 5: Informal definition of a Partial SharedPlan (PSP)

whole. Finally, the decision making required for plan elaboration is a group process and thus requires negotiation and group decision-making protocols.

Plan-Management Processes

A group of collaborating agents employ a variety of group processes to manage their multi-agent plan. As a group, they must elaborate their partial plan into a full plan which in turn requires that they complete the recipe if it is partial, select agents or subgroups to do constituent actions, and establish a variety of individual commitments and mutual beliefs. Although the general structure of these group processes is similar to their individual-plan counterparts, they differ in their fundamental reliance on group decision making. Taken together, these planning processes are the means by which partial plans evolve into full plans; as such, they constitute necessary elements of the plan-management repertoire of any group of collaboration-capable agents.

The SharedPlans definitions constrain the group decision-making processes in many ways. For example, a group decision to extend a partial recipe is constrained by the requirement that the new subactions in the recipe extension meet, at a minimum, the specifications for unresolved subactions in the PSP definition (see Figure 5). Similarly, a group decision to assign an agent or subgroup to some constituent action is constrained by the specifications for at-least-partially-resolved subactions in the PSP definition. Although the SharedPlans definitions constrain these group planning processes, they do not fully determine them. Exploring candidate algorithms for these group processes is one of the ways we are using the GigAgent system, as discussed in the next section.

In addition to group plan-management processes, the SharedPlans formalization requires each agent to have various individual plan-management processes. Some of these individual plan-management processes are associated with intentions-that arising from group activities. The role of these *intention-cultivation processes* is to determine what is necessary to monitor the intended proposition, to determine possible courses of action to make the proposition hold, and to decide which courses of action, if any, should eventually be carried out (Grosz and Kraus, 1999). Other individual plan-management processes are associated with intention conflict resolution. We are examining the effects of various intention conflict-resolution strategies in our empirical research, as discussed later in this article. In addition, a major emphasis of our work in constructing collaboration-capable systems is the investigation of candidate algorithms for intention-cultivation processes.

Manifesting Collaboration

The SharedPlans formalization places numerous constraints on the design of a collaborative agent. First, an agent intending to do some action must commit to means–ends reasoning (as represented by a plan to elaborate a partial plan). Second, for a group to have a SharedPlan, they must have agreed about certain decision-making procedures (e.g., to select recipes, to assign agents and subgroups to subactions). Third, since a rational agent may not simultaneously hold conflicting intentions, a collaboration-capable agent must have some means of detecting conflicting intentions and of reconciling conflicts when they occur. Fourth, as noted above, an agent participating in collaborative activity must have processes with which to manage its various intentions-that—from intentions that certain mutual beliefs are established to intentions that various subgroups be able to carry out their assigned tasks.

In this section, we describe three systems currently under development: WebTrader, a system for electronic commerce (Kraus and Hadad, 1999), GigAgents, a collaboration-capable multi-agent system, and DIAL, a system for collaborative distance learning (Ortiz et al., 1999). In the first two systems, elements of the SharedPlans specifications are directly reflected in the system architectures. In the third system, the SharedPlans formalization has been used to inform the system design, but the architecture does not directly reflect the SharedPlans specifications.

WebTrader: Electronic Commerce Application

The goal of the WebTrader application is to incorporate collaborative features into a computerized system for buying and selling items such as books, clothes and furniture on the Internet. In the WebTrader environment, several enterprises, each with several kinds of goods to sell to each other or to individual (human) buyers, have intelligent buyer and seller agents. The job of buyer agents is to purchase goods that are missing from the stocks of their enterprises. The job of seller agents is to sell their enterprises' goods.

In the WebTrader environment, agents' benefits are measured in terms of income. SharedPlans may be formed between agents belonging to the same enterprise who want to maximize their enterprise's income. They may also be formed among agents interested in collaboration only as a means of maximizing their individual benefit. For instance, suppose a buyer agent wants to purchase an item from a seller agent of another enterprise. Although each agent wants to maximize its own enterprise's benefit, and the agents may have certain conflicting interests, they share the goal of successfully completing the transaction. By forming a SharedPlan, they increase their chances of completing a mutually beneficial transaction. For example, having a SharedPlan may motivate the seller to notify the buyer should delivery problems arise.

The use of SharedPlans in the system design may also prove mutually beneficial when an automated seller interacts with a human buyer. For example, tracking the intentional context of an interaction (as we also do in the DIAL system, described later) may make it easier for the seller to help the buyer identify an item to purchase. Thus, if a buyer has selected an out-of-stock CD and the intentional context is that the buyer wanted that CD because it contained a particular song, then the seller might suggest another CD containing a different version of that same song. In contrast, if the intentional context is that the buyer wanted the CD because it contained songs recorded by a particular singer, then the seller might suggest some other CD recorded by that singer.

In the Electronic Commerce domain, agents begin with partial plans and develop them over time into complete plans. For example, at the beginning of an interaction a buyer may only be able to characterize her needs approximately; later on, with the help of a seller agent—who may, for example, present information about available items and their properties—the agents may be able to find a specific item to satisfy the buyer's needs. The Electronic Commerce domain is also dynamic: things change, possibly in the middle of the planning process, while the plan is still partial. In addition, an agent's beliefs about the results of its own actions may be uncertain or faulty and its knowledge about the world or other agents may be incomplete. For example, a buyer and seller may have agreed upon the date of payment for a particular item but, as the date approaches, the buyer may realize that she is unable to pay for the item as originally planned. Thus, while an agent is planning or acting on the basis of a partial plan, the plan may have to be revised. By providing for both individual and collaborative plans to be partial and by specifying the processes by which they may be elaborated, the SharedPlans formalization enables us to develop agents that are able to act in the dynamic and uncertain Electronic Commerce environment.

In contrast to other agent-based markets or retail outlets on the Web (Takahashi et al., 1996; Doorenbos et al., 1997; Schrooten, 1996; Chavez and Maes, 1996; Chavez et al., 1997; Wurman et al., 1998; Klaus et al., 1996; Albayrak et al., 1996) which focus on either extraction of information or negotiation strategies, our system supports cooperative interactions of buying and selling goods. The current prototype of the WebTrader agent is able to construct and carry out simple plans for buying and selling items in a simulated environment (which we implemented), but is not yet able to buy and sell on the Web. It is able to cooperate simultaneously with more than one agent. To enable such cooperation, we developed and implemented data structures and control mechanisms that enable the agent to handle several complex actions at a time. For example, a WebTrader agent is able to interrupt negotiations with one agent, start new negotiations with another agent, and then interrupt these negotiations to return to the first agent.

GigAgents: Domain-Independent Collaborative Multi-Agent System

The GigAgent system is a multi-agent system that supports the collaboration of any number of human and computer agents, each of whom may participate simultaneously in activities involving several different groups. Even within a single group activity, hierarchical task expansion may lead to a single agent being involved simultaneously in several active, subordinate SharedPlans. The participation of human agents in the GigAgent system is facilitated by agent *wrapper* software that handles most of the routine chores required of collaborating agents (e.g., sending and receiving messages and keeping track of commitments, proposals, pending decisions, and their interrelationships).

The purpose of the GigAgent system is to provide a platform for studying multi-agent collaboration. We are using GigAgents to test and evaluate various strategies and algorithms for team formation, commitment generation, plan elaboration and conflict resolution (especially conflicts arising from the interaction of commitments to individual and group activities). In the current implementation, we are simulating a computer systems administration domain in which heterogeneous groups of computer and human agents work together on tasks such as upgrading software and hardware, restoring deleted files from backups, keeping Web pages up to date, and checking system security.

Group decision making is a significant component of GigAgent activity. Agents must agree on how they will select recipes, bind various parameters of their group actions, and assign agents or subgroups to various constituent tasks. Furthermore, this group decision-making activity typically happens incrementally. The GigAgent architecture supports the incremental execution of planning actions (and domain actions, if desired) thereby enabling each agent to participate in several ongoing collaborative negotiations pertaining to different decisions required by various group activities. The system neither imposes nor presumes any particular organizational or management structure among collaborating agents. A range of possibilities—from completely democratic protocols (one agent, one vote) to autocratic ones—may be incorporated and tested. An initial voting-based protocol has been implemented to enable us to test the overall systems architecture and infrastructure.

We plan to use the GigAgent system to test and evaluate strategies for conflict resolution coming out of our empirically-based SPIRE research project (discussed later). We are also beginning to investigate algorithms for the single-agent *cultivation* processes associated with intentions-that held by an agent. These cultivation processes, implemented as incrementally-executable actions, are responsible for monitoring the status of their associated intentions-that, for considering ways of bringing about the intended propositions, and for determining any potential conflicts. Although we expect the cultivation processes to rely in part on domain-dependent reasoning, our current research aims to uncover the domain-independent portions of such processes.

Finally, in addition to the explicit collaboration among agents in the GigAgent system, there is implicitly modeled collaboration in the implementation of the agent wrapper software that facilitates the participation of human agents in the system. To make this interface more effective, we are investigating how to increase its collaborative nature by basing its operation on general principles gleaned from our work on the SharedPlans formalization and the human–computer interface systems we have built based on SharedPlans. The next section discusses this use of SharedPlans in more detail.

SharedPlans for Interface Design

We have used SharedPlans in the design of a collaborative human–computer interface system. The Distributed Information Access for Learning (DIAL) system (Ortiz et al., 1999) provides for multi-media interactions with a complex information system. Our goal in building DIAL was to demonstrate the efficacy of the SharedPlans model, and collaboration more generally, in providing people with natural, flexible, and effective means of communicating with computer systems. A central aim of the work has been to develop a system that enables users to obtain the information they need without having to specify the way in which the system should find it. DIAL is active in working with users to identify information relevant to their needs or tasks, rather than providing the narrow input–output window of current-generation interfaces.

The current DIAL implementation operates in the domain of information support for distance learning by students in an introductory programming class. It allows students to access course notes, videos, and assignments, as well as reference materials and teaching fellows, when undertaking such tasks as studying for exams, reviewing lecture materials, and working on assignments. In this domain, the collaborative plans formed by DIAL and its users are "information-locating" plans—that is, plans to locate information relevant to some task of the user (e.g. studying for an exam, preparing a lecture). These collaborative plans are in service of the user's individual intentions and plan to carry out the task. Typically, the initial SharedPlan will spawn many subsidiary SharedPlans (Lochbaum 1995, 1998) which will also be information-locating plans. For example, a user's individual intention to study for a midterm exam might lead her to engage DIAL to form a SharedPlan with the system to find information relevant to the exam. In such a case, DIAL would present a menu indicating the types of information available (e.g., assignments and lecture notes). At that point, were the user to select "review an assignment", she and DIAL would then form a subsidiary SharedPlan to locate assignments and assignment information relevant to the exam.

DIAL builds a representation of context based on the subsidiary relationships between the various SharedPlans it forms with the user. It employs this contextual information to reduce the user's communication burden. In particular, information requests are interpreted relative to the graphically displayed (and hence mutually believed) prevailing intentional context which may be manipulated by the user as well as the system. As a result, requests may be expressed more succinctly. For instance, if a user asks to see videos in the context of reviewing some topic (say, sorting algorithms), DIAL would interpret the request as a request to see videos on that topic. DIAL has been designed to use information about the intentional context also to respond and act collaboratively, rather than in the master–slave style typical of most current human–computer interfaces. For instance, it presents choices that the user might be unaware of as requests are pursued. The system has also been designed to use intentional context to plan alternative courses of action, but such plan-revision capabilities are not yet integrated into the implementation.

DIAL does not explicitly reason about beliefs and intentions, but does incorporate the specifications of SharedPlans listed in Figure 6. To simplify the construction and use of DIAL, we stipulated which actions the user and the system could perform and fixed the subtask assignments rather than

- *Intention-that the group perform the group action:* DIAL's commitment to the collaborative information-locating activities it undertakes with a user are manifest in its tracking of the intentional context and its use of that context to guide the interpretation of user queries and selections, and to constrain its replies.
- *Constructing and extending (partial) recipes:* The user and DIAL incrementally compose recipes for their group activities throughout any given session.
- *Mutual belief of subtask assignments and action capabilities:* The visual display of intentional structure provides common ground for establishing these mutual beliefs.
- *Intentions-that other participants be able to do subactions in the recipe to which they are committed:* DIAL constrains user choices to those actions that can be performed in the current context.

Figure 6: Specifications of SharedPlans incorporated in the DIAL system

leaving them to a dynamic decision-making process.

Empirical Studies of Decision Making in a Collaborative Context

Agents involved in a collaborative activity may need to reconcile their intentions to do group-related actions with other, conflicting intentions. For example, an agent committed to perform a certain task as part of a group activity may be given the opportunity to do another, unrelated activity that occurs at the same time as the group-related task. Since agents are presumed to act in an individually rational manner, seeking to maximize their (or their owner's) utility rather than the utility of the group, they need to be able to weigh the costs and benefits of the two options in the context of their commitment to the group. Intention reconciliation in such contexts is complex because agents need to consider how other group members will react if they fail to honor their commitments. Since agents may want to collaborate in the future with other members of the group, there may be significant costs to defaulting on group-related tasks.

To study the problem of intention reconciliation in the context of collaboration, we have constructed the SPIRE (SharedPlans Intention-Reconciliation Experiments) simulation system (Sullivan et al., 1999). SPIRE is general enough to allow us to model agents from a large set of problem domains and enables us to consider the impact of group norms and environmental factors on agents faced with conflicting intentions, as well as the effectiveness of different intention-reconciliation strategies that agents may adopt. The group norms include penalties imposed when agents default on group-related tasks. The environmental factors include the size of the group and the number of tasks in each time period. We aim to provide insight into the types of factors that affect individual and group behavior and outcomes, and thus derive principles for designers of collaboration-capable agents.

Initial pilot studies in the SPIRE framework (Sullivan et al., 1999) have established that the number of times agents default on group tasks drops off as agents give more weight to group factors in their decision making, that the number of defaults decreases as the penalty for defaulting increases, and that individual income and group income increase as the number of defaults decreases. While these results may be unsurprising, they verify that the group factors and norms we have designed are reasonable and provide a concrete demonstration of how group-related factors can affect the decision making of self-interested agents. A study that examined various aspects of "social consciousness" (agents being willing to sacrifice short-term personal gain for longer-term group good) has shown that as agents become more socially conscious, defaults decrease and more group-related tasks are completed, but that a moderate degree of social consciousness functions better than an extreme degree when the combination of group-task and individual-task income is considered (Glass and Grosz, 1999).

Conclusions

The SharedPlans formalization, of which we provide an overview in this article, specifies the minimal mental state requirements for a group of agents to have a plan for collaborative activity, stopping conditions for planning processes, and constraints on agents' beliefs and intentions as they initiate and expand partial plans. It provides a means of representing the commitments of participants in collaborations to their group activities and treats the partiality that arises naturally in most planning situations.

We have designed an architecture for collaborative agents based on this formalization and have built two systems based on this architecture. One, called GigAgents, models and supports explicit collaboration in planning and acting among both human and computer agents. The second, a system for electronic commerce, called WebTrader, supports the cooperative processes of buying and selling goods on the Web. The formalization has also proved useful in the development of a collaborative interface to a system for distance learning, called DIAL.

The development of the SharedPlans formalization and of systems based on it have suggested several fundamental problems in multi-agent collaborative planning. In this article, we briefly describe one such problem area: the problem of reconciling intentions in the context of group activities. We have developed SPIRE, an empirical framework that enables different decision-making processes to be simulated and studied. The article presents results of early studies using this framework.

Our current efforts are focused on continued empirical studies of intention-reconciliation, incorporation of the results of these studies into the collaboration-capable systems we are developing, and development of more sophisticated group decision-making algorithms for use in the systems.

References

[1] Sahin Albayrak, Ulrich Meyer, Bernhard Bamberg, Stefan Fricke, and Hermann Tobben. Intelligent agents for the realization of electronic market services. *The First International Conference on the Practical Application of Intelligent Agents and Multi Agents Technology*, pages 11–23, April 1996.

[2] Michael E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1987.

[3] Michael E. Bratman. Shared cooperative activity. *The Philosophical Review*, 101(2):327–341, 1992.

[4] Anthony Chavez, D. Dreilinger, R. Guttman, and P. Maes. A real-life experiment in creating an agent marketplace. *The second international conference on the practical Application of Intelligent Agents and Multi Agents Technology*, April 1997.

[5] Anthony Chavez and Pattie Maes. Kasbah: An agent marketplace for buying and selling goods. *The first international conference on the practical Application of Intelligent Agents and Multi Agents Technology*, pages 75–90, April 1996.

[6] Philip R. Cohen and Hector J. Levesque. Teamwork. *Nôus*, 25:487–512, 1991.

[7] Philip R. Cohen, Hector J. Levesque, and Jose H.T. Nunes. On acting together. In *Proceedings of the Annual Conference of the American Association for Artificial Intelligence (AAAI-90)*, pages 94–99, 1990.

[8] Robert B. Doorenbos, Oren Etzioni, and Daniel S. Weld. A scalable comparision-shopping agent for the world-wide web. *Autonomous Agent*, 1997.

[9] Alyssa Glass and Barbara J. Grosz. Socially conscious decision making. In *Proceedings of the 6th Annual Bar Ilan Symposium on the Foundations of Artificial Intelligence (BISFAI-99).* To appear.

[10] Barbara J. Grosz and Sarit Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.

[11] Barbara J. Grosz and Sarit Kraus. The evolution of SharedPlans. In A. Rao and M. Wooldridge, editors, *Foundations of Rational Agency*, pages 227–262. Kluwer Academic Publishers, The Netherlands, 1999.

[12] Barbara J. Grosz and Candace Sidner. Plans for discourse. In P. Cohen, J. Morgan, and M. Pollack, editors, *Intentions in Communication*, pages 417–444. Bradford Books/MIT Press, Cambridge, MA, 1990.

[13] Luke Hunsberger. Making SharedPlans more concise and easier to reason about. In J.P. Muller, M.P. Singh, and A.S. Rao, editors, *Intelligent Agents V — Agent Theories, Architectures and Languages (ATAL-98)*, volume 1555 of *Lecture Notes in Artificial Intelligence*, pages 81–98. Springer-Verlag, Heidelberg, 1999.

[14] D. Kinny, M. Ljungberg, A. S. Rao, E. Sonenberg, G. Tidhar, and E. Werner. Planned team activity. In C. Castelfranchi and E. Werner, editors, *Artificial Social Systems, Lecture Notes in Artificial Intelligence (LNAI-830)*, pages 227–256, Amsterdam, The Netherlands, 1994. Springer Verlag.

[15] Fischer Klaus, Joerg P. Muller, Ingo Heimig, and August-Wilhelm Scheer. Intelligent agents in virtual enterprises. *The first international conference on the practical Application of Intelligent Agents and Multi Agents Technology*, pages 250–223, April 1996.

[16] Sarit Kraus and Meirav Hadad. SharedPlans in electronic commerce. In Matthias Klusch, editor, *Intelligent Information Agents.* To appear.

[17] Karen E. Lochbaum. The use of knowledge preconditions in language processing. In Chris S. Mellish, editor, *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-95)*, volume 2, pages 1260–1266, San Mateo, CA, 1995. Morgan Kaufmann Publishers, Inc.

[18] Karen E. Lochbaum. A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572, December 1998.

[19] Charles L. Ortiz, Barbara J. Grosz, and Nathan Scales. Interpreting information requests in context: a collaborative web interface for distance learning. Manuscript in preparation.

[20] Martha E. Pollack. Plans as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 77–103. MIT Press, 1990.

[21] Martha E. Pollack. The uses of plans. *Artificial Intelligence*, 57(1):43–68, 1992.

[22] Martha E. Pollack. Plan generation, plan management, and the design of computational agents (abstract). In *Third International Conference on Multi-Agent Systems (ICMAS-98)*, pages 8–9. IEEE Computer Society, 1998.

[23] Charles Rich and Candace L. Sidner. COLLAGEN: A collaboration manager for software interface agents. Technical Report 97-21a, Mitsubishi Electric Research Laboratories, Cambridge, MA, March 1998. (To appear in *User Modeling and User-Adapted Interaction, Special Issue on Computational Models for Mixed Initiative Interaction.*).

[24] Ronald Schrooten. Agent-based electronic consumer catalog. *The first international conference on the practical Application of Intelligent Agents and Multi Agents Technology*, pages 543–556, April 1996.

[25] John R. Searle. Collective intentions and actions. In *Intentions in Communication*, chapter 19. The MIT Press, 1990.

[26] David Sullivan, Alyssa Glass, Barbara J. Grosz, and Sarit Kraus. Intention reconciliation in the context of teamwork: An initial empirical

investigation. In M. Klusch, O. Shehory, and G. Weiss, editors, *Cooperative Information Agents III*, Lecture Notes in Artificial Intelligence. Springer Verlag. To appear.

[27] Katsumi Takahashi, Yoshiyasu Nishibe, Ichiro Morihara, and Fumio Hattori. Collecting shop and service information with software agents. *The first international conference on the practical Application of Intelligent Agents and Multi Agents Technology*, pages 587–595, April 1996.

[28] Milind Tambe. Agent architectures for flexible, practical teamwork. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*, pages 22–28. AAAI Press/The MIT Press, 1997.

[29] Peter R. Wurman, Michael P. Wellman, and William E. Walsh. The Michigan internet auctionbot: A configurable auction server for human and software agents. *The Second International Conference on Autonomous Agents*, May 1998.