

Progress Appraisal as a Challenging Element of Coordination in Human and Machine Joint Activity

Paul J. Feltovich, Jeffrey M. Bradshaw, William J. Clancey, Matthew Johnson,
and Larry Bunch

Florida Institute for Human & Machine Cognition (IHMC)
40 South Alcaniz Street, Pensacola, FL 32502 USA
{pfeltovich, jbradshaw, wclancey, mjohnson, lbunch}@ihmc.us

Abstract. Joint activity, as we define it, is a mutually interdependent social endeavor that requires sufficient predictability among participating parties to enable coordination. Coordination, in turn, sometimes requires the parties to appraise the state of progress of their activities so that, if necessary, they can adjust their actions to meet coordination needs and communicate their status to others as appropriate. A significant impediment as yet precluding the full participation of automation in joint activity with people is its inability to sense and communicate aspects of its state that would allow other participants to meaningfully assess progress toward (or anticipate failure with respect to) mutual objectives. In the current article, we address various issues associated with “progress appraisal” and the challenges it poses for human-machine systems. We point to promising directions for future work.

Keywords: Coordination, culture, human-agent-robotic systems, joint activity, ontology, policy, predictability, regulation, teamwork, progress appraisal, common ground.

1 Introduction

“How are things going?” or “How’re you doing?” or even just, “Hello,” may seem to be meaningless, perfunctory exchanges among friends and strangers as they pass on the street. But in reality they are more than what they seem. In our ongoing study of interaction among groups of humans and machines, we have come to realize that these deeply-reinforced conventions of decorum serve as simple probes to test the friendliness and predictability of the people in the environment, functioning as rough gauges to the safety of the moment or, on the other hand, to possible cooperative opportunity [23; 43]. With such a perspective in mind, think of what it indicates when, instead of the usual reply of “Fine, thanks,” one receives an unaccommodating rebuff, “That’s none of your business!”

In situations where people actually *do* join to engage in cooperative work, more complex forms of mutual probes can take on real importance as participants try to assess periodically how the shared work is progressing, especially with regard to interdependent aspects of that work. When machines join people in such endeavors, it

is important that they also be able to participate in these ongoing, social processes of probing and appraisal [7, 11, 26, 27].

As automation becomes increasingly pervasive in the ordinary social situations of life, the need for devices capable of such “mixed-initiative” interaction will become ever greater. Although the pursuit of fully autonomous systems is a worthy goal, in many situations coordination with people can improve performance (e.g., [25]). In other situations, agents¹ cannot yet be trusted to perform critical tasks on their own and must be teleoperated by one or more people [12]. Despite these realizations, today’s automation is too often implemented as what Sarter, Woods, and Billings call “strong, silent systems” [41] with only two modes: fully automatic and fully manual. In practice this can lead to situations of human “underload,” with the humans having very little to do when things are going along as planned, followed by situations of human “overload,” when extreme demands may be placed on the human in the case of sudden, unanticipated failure.

This essay explores some of the challenges that *progress appraisal* poses for joint activity involving humans and machines. We first introduce some of the characteristics of social activity and joint activity. Then, in Section 3, we discuss “progress” within human activity. In Section 4, we take up the special role of time, both as a way to specify and to indicate progress. In Section 5, we reach the heart of the matter, as we examine the many factors involved in progress appraisal for humans and machines. These aspirations for machines should be considered *desiderata* rather than current accomplishments. In Section 6, we describe some of our ongoing explorations of these ideas within mixed human-agent-robotic systems, and offer some tentative conclusions.

2 Activity and Joint Activity

Human beings partition the flux of the world into personally meaningful units that researchers have referred to as “activities” [14, 16, 36]. The nature of these activities can usually be revealed by asking a person at any moment, “What are you doing?” The answer will likely be something like “I’m shopping” for a person in a department store, or, from a college professor, “I am giving my lecture in my class,” or “I’m preparing my lecture for tomorrow’s class.” The response will not likely be “I am breathing” or “I am standing on my feet.” For observers of conventional scenes involving people doing things, there is usually pretty good agreement about what activity is going on and when one activity has ended and another one has started [37]. Despite such rough consensus in everyday life, what constitutes an activity (as well as its beginning and ending points) is not an objective “thing in the world,” but is, rather, a context-sensitive “conceptualization” [14] or “construal” [31, 32] generated by the participants involved. If our professor’s attention is absorbed by the painful physical effects he is suffering as a result of his long lecture, he may respond: “I have been standing on my feet too long.”

¹ We use the term “agent” in an unqualified way to refer to autonomous or semi-autonomous entities such as software agents, robots, and similar technology.

Previously, we have discussed the nature of *joint activity* as a generalization of that concept proposed by Herbert Clark [20, 34, 35]. Joint activity occurs when people come together to try to accomplish something as a group. We have argued that “The essence of joint activity is interdependence—that what party “A” does depends on what party “B” does, and vice versa (e.g., “One if by land, two if by sea” in Longfellow’s account of Paul Revere’s famous ride). As soon as there is interdependence, there is often need for coordination in time (e.g., timing a live, multi-party phone call) and/or space (e.g., designating a drop-off point), which in turn requires some amount of predictability and order” [24].

We have also proposed that predictability and order derive from highly diverse regulatory systems that play an essential role in human cultures [23]. These systems can range widely, from devices such as formal law to ethnic folkways to organizational and work norms and practices to informal codes of courtesy and good manners that are meant to govern many aspects of everyday interactions [24]. These can apply to activities (e.g., driving), products (e.g., codes for household electrical wiring), or the social roles (e.g., squad leader) assigned to or adopted by actors.

3 Aspects of Progress

The concept of “progress” itself is, of course, not objective but must be related to the context and aims of individuals. For instance, members of winning and losing baseball teams within a game will have different appraisals of their progress toward victory. On the other hand, they will be more likely to share similar views with respect to clocks and landmarks (e.g., innings) that serve as indicators of the remaining time left to play. Two members of the same team may, in turn, appraise progress differently because they are concerned with their own piece of the overall joint activity in addition to the more global perspective. A good example of discrepancy in appraisal comes from a commentary on the ethnic turmoil in the Balkans, showing how different stakeholders can judge progress differently—in this case by attending to different components:

The omens before the March pogrom did not all auger ill. Political officers in the UN mission who had been monitoring inter-ethnic tolerance were seeing progress. Returns of displaced Serbs had increased... Talks between Belgrade and Pristina about a variety of practical issues of mutual interest had recently begun... Minority representation in the Kosovo Police Service had improved... But there were also signs that tensions were reaching breaking point... The Kosovo Assembly, the territory’s elected parliament, had marginalized the significant number of minority members... The official opening of the Assembly’s refurbished hall was marred by Kosovo Serb’s understandable complaints about murals depicting scenes that reflected only the Albanian’s view of history... [33, p. 7]

If the world could be completely objective and predictable, there might be little need to monitor the progress of the multiple efforts involved in a joint activity. Consider, for instance, two individuals who are involved in independent activities for the

afternoon (estimating the time needed for all of them) and then plan to meet at a certain place and time that evening. If everything goes as expected, they will simply meet at the appointed place and time, without having to consider their intermediate progress. But disruptions of various kinds often impinge in such situations (e.g., one of the parties is delayed by traffic or an emergency at home), and so, to successfully accomplish their coordination aim, they will need to communicate their status at critical junctures and adjust accordingly.

We note that even assessing one's own status regarding progress can be challenging. In the case of machines, though they may have excruciatingly detailed access to aspects of their internal functions, they are typically oblivious to the future implications of their current state or to important aspects of the world around them that may affect their ability to perform. It has long been a complaint about automation that it cannot make or broadcast these kinds of high level, reflective judgments and that this limitation often causes situations to unravel helplessly, as machines fail without warning or report trouble too late to enable corrective measures [5, 13, 39, 41]. This article addresses the kinds of requirements that must be met if machines were to have significant capability for *judging their own progress and, perhaps, being able to offer reliable warnings regarding deviations from expected progress or, in the worst case, an advance notice of impending failure*. Time is an important dimension for these kinds of assessments—a topic we take up next.

4 Time as a Special Dimension

Time serves many roles as part of progress appraisal. For instance, it can function as constraint (“I have—or estimate that I have—thirty minutes to complete the task”), a planning factor (“Given the time I have, I will try doing the task this particular way”), or more purely as an indicator of progress (“This is taking way too long, I’m not going to make it in time”).

Time, along with the degree of tightness in the coupling of the interdependent activities, is often a central factor in judging progress. In general, shorter time and tighter coupling reduce the margins of error in joint activities and allow less time for appraisal and the making of any necessary adjustments. In any form, progress appraisals will need to be made in different kinds of contexts that involve timing:

- **Deterministic: Designed, Fixed Timing.** These situations involve fixed deadlines that a process must meet, such as a final date for submitting a research proposal or making a reservation. These deadlines are set in order to coordinate with the processes of others (e.g., the review and handling procedures by the respective organizations involved) in a fashion that has been specifically designed to minimize surprises and to allow the deadlines (coordination points) to be planned, stable, and public. In some instances adjustments are possible, as in extensions, but there is often fixed procedure (and set times) for making these, too. A key characteristic of such situations is the desire for predictability (and relative stability) of both the processes involved in the coordination and the timing of the interdependencies among the parties. For example, sub-process relations, causal influences, and causal effects are deterministic.

- **Emergent Internal Relations: Contingent Inter-Process, Fluid Timing.** In this kind of coordination, a deadline is not fixed as a time, but, rather, is dependent on the progress of other interdependent processes (see section on “Other activities” below). For example, a coauthor on a scientific article cannot make progress on her part of the writing until some data are analyzed by another colleague (with both authors, perhaps, subject to a fixed deadline for completing their overall project). In such cases, both interdependent processes must somehow be made aware of (or estimate) progress within the other. When all goes “as planned,” they may coordinate through a pre-set schedule or plan. But perhaps more often, they will need to communicate their progress to one another as they move forward.
- **Emergent External Influences: Timing Imposed from Indirect Influences.** Many influences external to the main joint activity affect the efficiency of its completion. For example, available resources may be germane, as when a pilot is waved off from a scheduled landing site and needs to find an emergency site that is reachable with his available fuel. Loss of primary communication modes slows (or prevents) interdependent progress. Local regulations may impinge, unanticipated obstacles (or affordances) pop up, the weather change, and so forth. *Surprises*, in general, seem particularly relevant in this category.
- **Emergent External Effects: Timing Imposed by the Half-life (Perishability) of the Usefulness of Components.** Elements may only “work” for some purpose for a limited window of time. For example, a photograph of an active battlefield, provided by a higher command to an officer on the ground, will likely have a shorter span of usefulness than a weather report pertaining to the same locale, or an even more stable geological survey. The value of work performed, or information provided, may plummet to zero if it does not appear until after the “usefulness deadline” has passed. More positively, partial products will sometimes suffice for the coordinating party to be able to make *some* progress, even if less than the full product would allow. The parties involved have a number of options if they determine that a usefulness deadline may be missed. For example, they can intervene by providing more resources to the activity, or they can seek the product from an alternative source.

To summarize, the necessary timing of a process (always subject to need for adjustment) is determined in advance (a schedule) or emergent. Emergent timing is contingent on sub-process relations of variable duration and on external factors. External factors include 1) *causal influences on the process* that may change before the process is complete and 2) *causal effects of the process* (e.g., functional value of its products) whose timing affects the external factors. Because timing is inherent in causal processes—whether parallel or sequential—the temporal nature of the process provides information for specifying or indicating how the process is progressing.

5 Factors Affecting Progress and Progress Appraisal

Many other factors affect progress and its appraisal in interdependent human-machine systems. We will now discuss some of these and give examples to clarify the nature of the factors and to provide ideas about how they might be addressed. While the particular factors presented have been stimulated by some of our own experiences in

developing human-agent-robotic systems over many years, as well as those reported by others (e.g., the criticality of communications), we make no claim to exhaustiveness. In this sense, the factors discussed can function as a stimulus to further investigation and reporting.

Communication. Coordination without some form of communication is challenging for humans and agents alike. In critical situations with a high degree of mutual interdependence, a loss of communication can spell danger. This has been a major lesson from almost all recent international responses to disaster, as reflected in pleas, contained in many after-action reports, for “*bullet-proof*” and *interoperable* communication devices and networks (e.g. [22]).

Beyond the mere availability of communication, the quality and timeliness of progress reports play a key role. Reports that arrive later than planned because of a downed system can handicap the activities of others who are dependent on them. The negative effect often ripples outward in a variety of ways, impacting available time to adjust or to consult with others to get assistance. Messages that deceptively or inaccurately report progress have similar effects. Over-messaging can also have deleterious effects on progress and its appraisal [40].

Human Example: Joe and a colleague, who will arrive first at the airport, agree to meet at the AJAX car rental desk where they will share a ride. Upon arriving at “the” desk, Joe discovers that this particular airport has four AJAX desks, one in each of three terminals and one outside the terminal. His partner is not there. He tries to call his partner from his cell phone but finds that he has no signal at this spot. After some time, he decides to ask the agent at the nearest AJAX desk to call the other desks to look for his partner. They finally connect and successfully co-locate.

Agent Teamwork Example: A human-agent team could incorporate a policy that causes regular checks on all agents. If they do not respond, one knows that either the communication has failed—or worse. At a specified heartbeat rate, a probe assigned to each agent could automatically test capability without interfering with the ongoing work of the agents (cf. [2]). Also desirable would be a system of back-up communication that could be engaged when first-line devices fail or are suspect, and an agreed upon “safe” mode or default behavior to which each agent would revert if communication becomes impossible.

Landmarks. These are recognizable entities, including partial products, that should be seen or produced if acceptable progress is being made.

Human Example: Knowing the general distance from the airport to his hotel and the current flow rate of traffic, Joe believes that he should have seen the bridge on his map by now. He becomes concerned that he may have made a wrong turn somewhere and starts comparing the street signs he can see around him with street names on the map. Having determined his current location, he adjusts his route accordingly.

Agent Teamwork Example: Landmarks encountered, goals achieved, or results produced during a process can be indicators of current state in comparison to observables, that is, to detectables known to reflect degree of progress toward goal states. Physical metrics, such as distance from a physical destination, are one kind of indicator. But it is also important to be able to track more abstract indicators. For

example, if there is a simple plan that contains two steps, it would count as progress when the first step was completed.

Successful use of landmarks clearly depends on familiarity with an activity. One needs to have expectations of what kinds of things appear or are revealed, get produced, or are consumed, as a process develops to completion. This can be supported by actual experience and learning, or by various sorts of props that represent vicarious experience (e.g., maps, plans, guides, checklists). Examples of landmarks include goals achieved or steps completed in a plan, entities encountered or revealed in relation to those expected, computational results produced in comparison to those required, resource use in relation to average consumption, and actual time intervals between two events in comparison with “normal” intervals.

Helps, Obstacles, and Affordances. By these we mean all the diverse elements that aid, hinder, or merely allow progress. Some of these may be of long standing, while others pop-up unexpectedly (e.g., summer-long bridge construction vs. a traffic accident).

Human Example: On his drive from the hotel to the airport, Joe discovers that his intended exit from the highway has been closed that day for repairs. He checks his map to see how far he is from the next exit.

Agent Teamwork Example: Policy can require agents to notify other team members of helps, obstacles, and potential affordances that will affect their individual performance—or that of the team. Such an obligation recalls the teamwork heuristic, developed by Cohen and Levesque, that required agents to tell team members when a team goal was achieved or became impossible or irrelevant [21]. This generic approach saved developers from having to write numerous special purpose exception handling procedures for specific situations [44]. Degree of anticipation is an important factor in dealing with helps, obstacles, and affordances. Anticipated obstacles, e.g., increased crowds at restaurants around lunch hour, can be planned for (block out more time) or worked around (go earlier or later). Accuracy in anticipation, again, depends on familiarity and learning.

Resources. Resources range from such things as *energy and bandwidth* to necessary *artifacts and tools* (e.g., a hammer, a car [and energy for the car], a map). What stands as a resource is relational, that is, “This A serves as a resource for this B in context C.” Thus, just about anything can serve as a resource for something else in the right context.

Human Example: Joe runs out of gasoline in his car and stalls. He tries to use his cell phone to call for help, but he discovers his phone battery is dead.

Agent Teamwork Example: Levels of necessary consumables and achievability of enabling conditions for actions can be monitored. For example, batteries, communications, and other resources can often be monitored—and perhaps reasoned about (e.g., “Do I have enough gasoline to make it to my destination?”). Enabling conditions for the execution of actions can also be investigated (e.g., “Do I have a hammer if I need to pound a nail, a vehicle if I need to make a trip?”).

Knowledge. Included here are basic knowledge of requirements for an activity and also the means for addressing these, including alternative means, routes, geography, places to acquire consumable resources (e.g., gas and food), people who can help, the

roles of team members, knowledge of pertinent regulations, access routes, and so forth. Experience with a joint activity aids progress appraisal in at least two ways. First, with time the parties involved build up norms for how much time component activities usually take and the likely impediments that may arise. This helps detection of unusual time delays and allows preplanning and workarounds for many known potential impediments. Second, with experience, team members get to know each other's roles, habits and manners, for example, leadership, timeliness, trustworthiness, work-habits, and degree of communication availability. If a team facing a hard deadline needs a deliverable from a team member by a certain time, they will appraise the situation differently if the partner is reliable and always comes through, versus another colleague whose delivery patterns are spotty.

Human Example: Joe, a visitor, runs out of cash in Japan and, to no avail, tries to find ATM machines or banks to help him. He does not know that in Japan many of the functions carried out by banks in other countries are, instead, handled by post offices. A local resident tries to explain this situation to Joe, but Joe cannot understand Japanese.

Agent Teamwork Example: Agent planning capabilities generally address problems concerning what knowledge, actions, and resources are necessary to complete a given task. Typically, however, such planners are limited in their ability to be self reflective, let alone being able to reason effectively about coordination issues involving other human and agent team members. Our research group is working on collaborative planning approaches that take as a premise that people are working in parallel alongside (more or less) autonomous systems, and, hence, adopt the stance that the processes of understanding, problem solving, and task execution are necessarily incremental, subject to negotiation, and always tentative [1, 6]. In this way, agents and robots can work with people in a mixed-initiative mode—doing as much as they can autonomously based on their own knowledge, but also being aware of when they need to take direction or ask for help from others.

Mistakes. These are actions other than those intended, including “slips” [15, 38]. When recognized, mistakes should trigger attempts to recoup and re-estimate progress (e.g., how much time, relative to the previously expected time, the process may now take for completion, given the mistake).

Human Example: Joe arrives at his departure gate for a connecting flight. Just as the final call is being made, he considers whether he has time to visit the restroom. He decides he has just enough time. Upon exiting the restroom, he turns down the hall and starts walking. Suddenly, he realizes he went the wrong direction. Since he caught the mistake quickly, he turns around and walks at a regular pace. Had he walked farther before noticing the error, he may have had to quicken his step.

Agent Teamwork Example: Unlike obstacles, mistakes are the result of the agent's own choices. Effects of these choices can be monitored. When a mistake is detected, the agent can abandon the plan, retrace back to the intended path and then continue, ask for help, or construct a new path. Teng (IHMC) has developed an initial version of *Kab* (KAoS abstract backup), a new special-purpose planner that works in conjunction with the KAoS *Kaa* component (see below) to help agents and agent teams formulate

and select appropriate generic backup plans for such situations. Unlike typical approaches for this problem, *Kab* relies on a small number of human-compatible strategies for plan repair based on our observations of teamwork-in-practice, rather than a collection of general-purpose formalisms grounded in logic alone.

Regulatory Devices. These include any rules, regulations, customs, or other constraints (and affordances) that apply to the activity at hand (e.g., speed limits, rights of way, policies) or to the roles of team members (e.g., associated restrictions, rights, obligations).

Human Example: Not having planned well, Joe finds himself with less than ample time to get to the airport. On the map, he sees an alternative route that looks shorter and more direct. He takes this route and soon discovers that the speed limit on this road is low. He also reads on a sign that the fines for speeding at this particular time are doubled because there are workers doing maintenance on the road. He considers tracing back to his original route or trying to find yet another one.

Agent Teamwork Example: This involves the detection of impedance to progress due to the enforcement of a regulatory mechanism (e.g., a policy, rule, or role responsibility/obligation). To the extent these situations can be represented by (computational) policy, this is straightforward (e.g., having a policy to prevent entry to a restricted area). *Kaa* is an adjustable autonomy capability that uses decision-theoretic reasoning to determine how to adjust policies, resources, or circumstances to assist in the achievement of team objectives [9].

Conflicts. Because resources such as time and attention are finite, conflicts regarding their allocation sometimes occur. These include conflicts among alternative activities, goals, obligations, and allegiances.

Human Example: Not only must Joe board his flight in time, he also needs to finish off a piece of writing before boarding (finishing it after the flight would be too late) to send to a colleague back home who is completing submission of a grant proposal under deadline. He decides to move close to the airplane entrance door, not board when first called, and work on the writing up until the last moment before the airplane door is closed.

Agent Teamwork Example: Conflicts can be handled by agents in several ways. One is to split the team, if possible, to cover the different duties (differential reassignment). For example one agent covers one task, and another covers the other. Their work can also be prioritized—a kind of triaging. Alternatively, they might do a merely adequate, rather than a superb, job on each task so they create time for both. They might just speed up. Policies can specify priorities for competing demands and can also constrain what alternatives are available for re-tasking and delegation.

Changes of Plans. Sometimes in the midst of the conduct of a planned joint activity, events evolve that cause overall aims to change. This may happen for many reasons, for example, a more pressing need/objective has emerged, elements of the current operation have failed so badly that the original aim must simply be aborted, or an opportunity has arisen that enables achievement of a higher priority goal. The effect on progress will depend on how discrepant the new objective and plan are from the original one.

Human Example: Joe is in a distant town for a three-day academic workshop. At the end of the first day, he receives a call from his wife back home reporting that there has been a family emergency. He immediately starts making arrangements to return home, recruiting a colleague to deliver his paper at the workshop.

Agent Teamwork Example: The topic of this essay, progress appraisal, is pertinent to judging whether an activity is progressing acceptably, with change of plan being a recurrent theme (e.g., as the result of a surprise). In that sense, the whole article involves suggestions for building this capability in agents. One approach utilizes “back-up” plans, both contingency (worked out in advance) plans and dynamic re-planning (see brief discussion of the development of *Kab*, above).

Other Activities. This category is particularly pertinent to interdependent activities. We are interested in *joint* activity, so how well one activity “is going” is not independent to that activity, but is contextually related to how other processes are progressing, especially regarding their points of interdependency (e.g. how fast one participant must complete its activity depends on how fast some other process needs the output to accomplish *its* part of the joint activity).

Human Example: At their destination airport, Joe and a colleague, who is coming from a different city but is scheduled to arrive about the same time as Joe, are to join each other and to share a rental car to their hotel. Joe’s flight experiences a moderate delay in departure, and he calls his friend to tell him his flight will be late. Because the delay will be relatively short, the friend says he will get something to eat at the airport and wait for Joe. Had the estimated delay been long, the friends might have decided just to make their own ways to the hotel.

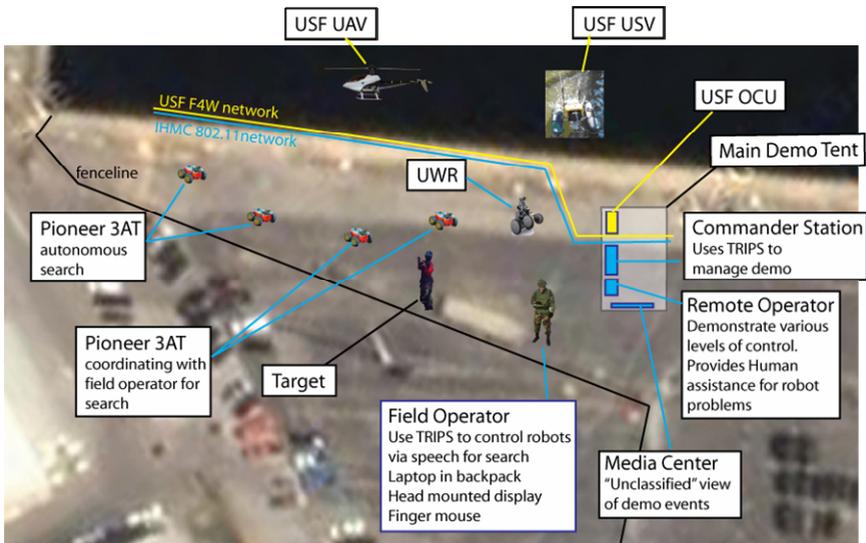


Fig. 1. Overview of Coordinated Operations Exercise Components

Agent Teamwork Example: Any evidence, reported by an agent or solicited from an agent, indicating that some process in which it is involved is progressing differently from expected, especially with reduced progress, is important. One approach is for agents involved in joint activity to monitor each other's progress and to make adjustments as needed along the way.

6 Applications to Human-Agent-Robotic Teamwork

Our group is applying the ideas presented in this article to facilitate joint activity in mixed human-agent-robot teams.² We have developed the KAoS HART (Human-Agent-Robot Teamwork) Services framework as a means to flexibly specify, analyze, enforce, and adapt policy constraints to facilitate team coordination [8, 45]. KAoS policies, represented in OWL (W3C's Web Ontology Language, <http://www.w3.org/2004/OWL>), are of two types: *authorizations* (constraints that permit or forbid some action in a given context) and *obligations* (constraints that require some action or waive a pre-existing requirement). More complex teamwork policies are built out of these primitive types. To more adequately represent some of the subtleties of joint action, we are augmenting the policy ontology with a broader Ontology of Regulation [24].

In the context of human-agent-robot teams, we have been involved in domains such as space exploration, disaster response, and military operations [9, 17, 18, 28, 42]. A recent field exercise, particularly germane to the present essay, involved complex coordinated operations of a team consisting of two humans and seven robots (Fig. 1) [30]. In this demonstration, the team had to perform reconnaissance of a port facility to determine the presence of underwater obstacles, explosives, structural soundness of pier facilities, and the nature and extent of any armed resistance. In the course of this surveying, an intruder was detected, and the team then needed to secure the boundaries to prevent escape, and to search the area to find and apprehend the intruder. The team consisted of two humans: a remote commander and a lieutenant in the field interacting locally with the robots. The robots included an unmanned air vehicle, an unmanned water surface vehicle, a highly mobile IHMC "tBot" robot, and four Pioneer 3AT ground vehicles, variously equipped with sonar, GPS, cameras, and SICK lasers.

6.1 Progress Appraisal in the Coordinated Operations Exercise

Members of the human-robot teams were strongly associated with roles. Roles can be thought of as ways of packaging rights and obligations that go along with the necessary parts that people play in joint activities [10, 29]. Knowing one's own role

² We realize that there are important differences between human teams and the mixed teams of which we write. Even the authors of this paper have had lively debate as to whether the use of the term "team" is appropriate in this context and whether machines and software can appropriately be classed as "team members." While recognizing the significant—and perhaps insurmountable—differences between the contributions that technology and people can make to joint activity, a large segment of the research community uses "team" as a rough way of characterizing the ideal forms of interaction to which we aspire. For snapshots of this ongoing debate, see [3, 4, 18].

and the roles of others in a joint activity establishes expectations about how others are likely to interact with us, and how we think we should interact with them. In addition, progress is often associated with the duties of a role. When roles are well understood and regulatory devices are performing their proper function, observers are likely to describe the activity as highly-coordinated. On the other hand, violations of the expectations associated with roles and regulatory structures can result in confusion, frustration, anger, and a breakdown in coordination.

Collections of roles are often grouped to form organizations such as teams. In addition to regulatory considerations at the level of individual roles, organizations themselves may also add their own rules, standards, traditions, and so forth, in order to establish a common culture that will smooth interaction among parties.

The lesson here for mixed human-agent-robot teams is that the various roles that team members assume in their work must include more than simple names for the role and algorithmic behavior to perform their individual tasks. They must also, to be successful, include regulatory structures that define the additional work of coordination associated with that role.

Consistent with this role-based orientation, coordination of search and apprehension activities was facilitated by five sets of KAoS policies addressing chain of command, acknowledgement, and progress appraisal issues as they relate to the requestor, the team leader, and other team members. The particular policy sets chosen for initial implementation are related to core components of our developing Progress Appraisal investigations. Some are directly related to progress appraisal as indicated by their titles. Two others, Chain of Command and Acknowledgement, are important auxiliary components. The first puts into play a central set of regulatory constraints that can aid or impede progress in our particular domain of application, a military operation (see section on “Regulatory Devices” above). The latter is a basic set of communication policies for supporting the basic progress appraisal process (see section on “Communications” above). We note the similarities of some aspects of our work to Winograd’s Coordinator [46].

6.1.1 Chain of Command

This policy set enforces a hierarchical, military-style chain of command and consists of four policies:

- A Robot is not authorized to perform Action requests from just any Requestor EXCEPT
- A Robot is authorized to perform Actions requested by its Team Leader
- A Robot is authorized to Accept Actions requested by a higher authority
- A Robot is authorized to Accept Actions that are self-initiated

These policies support the norms of authority in a military operation. Here, we focus on the role of a leader in relationship to subordinates (e.g., the robots, above).

6.1.2 Acknowledgment

This policy set enforces acknowledgment of all commands, except those that are directly observable:

- A Robot is obligated to Acknowledge when the Robot Accepts an Action EXCEPT
- A Robot is not obligated to Acknowledge Teleoperation requests
- A Robot is not obligated to Acknowledge Query requests

“Acknowledgement” means, “I (e.g., a robot) got your message,” *prior to acting on it*. Simply acting on it, when the act’s execution is visible, usually obviates the need for pre-acknowledgement, as Clark has written about regarding his “Joint Action Ladder” [20].

6.1.3 Requested Action Progress Appraisal

This policy set enforces communication norms between a requestor and requestee, based on progress:

- For *Continuous Actions*, A Robot is obligated to notify the requestor when the Status of the requested Action *changes*
- A Robot is obligated to notify the Requestor when requested Action is Finished (includes statuses of Completed, Aborted, and Failure)

EXCEPT *Certain types of commands are directly observable and do not require feedback unless something goes wrong.*

- A Robot is not obligated to notify the Requestor when a requested Teleoperation Action is Completed successfully
- A Robot is not obligated to notify the Requestor when a requested Query Action is Completed successfully

These help provide the requestor with progress appraisal information *for execution of the specific action requested*, except when progress can be assessed more directly, e.g. can be seen.

6.1.4 Leader Progress Appraisal

This policy set enforces communication norms with a team leader based on progress:

- A Robot is obligated to notify its Team Leader when an Action is requested by a higher authority (higher than the team leader)
- A Robot is obligated to notify Its Team Leader when starting a self-initiated Action
- A Robot is obligated to notify its Team Leader when a self-initiated Action is Finished (includes statuses of Completed, Aborted, and Failure)

These policies serve much like those of 6.1.3, except that the role of “leader” demands some special kinds of notices, compared to a general Requester (role).

6.1.5 Group Peer Progress Appraisal

This policy set enforces communication norms between members on the same team based on progress:

- A Robot is obligated to notify other participants in a Joint Task when the Joint Task is Finished (includes statuses of Completed, Aborted, and Failure)

- A Robot is obligated to notify its Team Members if the Team Goal is Aborted and no longer applicable

Again, these policies support appraisal of progress among members of the working group, but in some ways differently from the like support provided to a “Leader (role).”

6.2 Further Policy Considerations

Each of the operative policies helps to maintain common ground (mutual understanding) by enabling the robotic agents to coordinate with their human counterparts in a manner consistent with human norms and expectations [19, 34, 35]. As a simple example, when asked to perform a task, the robots will acknowledge the request (simply saying that it has *received* the message; acknowledgement in this sense does not refer to its possible subsequent actions taken). However, for teleoperation, the requests are numerous and the effect, the resulting action, is directly observable, so acknowledgement would become annoying and detrimental. Therefore we waive this requirement. Similarly, if the lieutenant is in charge of a robot, and his commander overrides his authority and tasks the same robot, we ensure the lieutenant is informed, rather than leave the lieutenant in a state of confusion about the unexpected actions of the robot.³ Another example of progress appraisal is that when a robot is tasked to search for something and it finds (or loses) it, the robot tells the requestor of this status change—something obvious to humans but typically not considered explicitly in robot operations.

We have previously noted the need to design-in collaborative capabilities in robots at a more basic level than is typically done [28]. One area we are focusing on for the future is determining how to code robots in a manner that allows for a finer grained progress appraisal. It would be useful, not only to know if an action is completed or aborted, but also if the robot is struggling or delayed. We have been working on several examples of robotic behavior where we can provide this type of information.

A final challenge is dealing with the more subjective aspects of progress appraisal (see Section 3). We note, for example, the difference in difficulty between assessing progress on a more objective task that depends on closing a known distance between a robot and its target and a more subjective task that depends on aggregating a number of imperfectly known estimates, perhaps even reflecting different points-of view or conceptualizations of the meaning of the activity. This will be a daunting problem [24].

7 Conclusions

Complex operations involving mixed teams of humans, software agents, and robots, require strong tools to support coordination of the interdependencies among the activities of the components. Since, in the real world, activities usually do not play out exactly as expected, successful coordination often requires that adjustments in planned activities be made to accommodate disturbances in progress within the

³ The Commander has higher authority in this case; extant policy automatically detects the conflict in commands and de-conflicts according to chain of command.

coordinating activities. This, in turn, requires that agents have an understanding of their progress so that they can convey this to other participants. Historically, this kind of assessment of progress (with associated warnings of impending trouble) has been a critical limitation of automation, and this lack has contributed to some automation disasters. Our research group is confronting this specific kind of limitation. The work is just beginning, but we are optimistic that important progress can be made.

Acknowledgements

Our thanks to the following individuals for their substantial contributions: James Allen, Maggie Breedy, Marco Carvalho, Tom Eskridge, Lucian Galescu, Hyuckchul Jung, James Lott, Robin Murphy, Jerry Pratt, Anil Raj, Signe Redfield, Richard Smith, Niranjani Suri, Choh Man Teng, and Andrzej Uszok. The research was supported in part by grants and contracts from the U.S. Air Force Office of Scientific Research to Dartmouth College through subcontract (5-36195.5710), the Office of Naval Research (N00014-06-1-0775), and the U.S. Army Research Laboratory through the University of Central Florida under Cooperative Agreement Number W911NF-06-2-0041. William J. Clancey has been supported in part by funding from NASA's Constellation Program. The article does not necessarily reflect the views of any of these agencies. We also thank the fine reviewers and the editors of this volume.

References

1. Allen, J.F., Ferguson, G.: Human-machine collaborative planning. In: Proceedings of the NASA Planning and Scheduling Workshop, Houston, TX (2002)
2. Arkin, R.C.: Homeostatic control for a mobile robot: Dynamic replanning in hazardous environments. *Journal of Robotic Systems* 9(2), 197–214 (1992)
3. Biever, C.: Bots as peers: It's all a matter of teamwork. *San Francisco Chronicle*, San Francisco, CA, May 6 (2007) (accessed September 15, 2007), <http://sfgate.com/cgi-bin/article.cgi?f=/c/a/2007/05/06/ING9GPK9U71.DTL>
4. Biever, C.: If you're happy, the robot knows it. *The New Scientist* 193(2596), 30–31 (March 24, 2007) (accessed September 15 (2007), <http://technology.newscientist.com/article/mg19325966.500-if-youre-happy-the-robot-knows-it.html>
5. Bradshaw, J.M., Sierhuis, M., Acquisti, A., Feltovich, P., Hoffman, R., Jeffers, R., Prescott, D., Suri, N., Uszok, A., Van Hoof, R.: Adjustable autonomy and human-agent teamwork in practice: An interim report on space applications. In: Hexmoor, H., Falcone, R., Castelfranchi, C. (eds.) *Agent Autonomy*, pp. 243–280. Kluwer Academic Publishers, Dordrecht (2003)
6. Bradshaw, J.M., Acquisti, A., Allen, J., Breedy, M.R., Bunch, L., Chambers, N., Feltovich, P., Galescu, L., Goodrich, M.A., Jeffers, R., Johnson, M., Jung, H., Lott, J., Olsen Jr., D.R., Sierhuis, M., Suri, N., Taysom, W., Tonti, G., Uszok, A.: Teamwork-centered autonomy for extended human-agent interaction in space applications. In: *AAAI 2004 Spring Symposium*, 22–24 March, 2004. Stanford University, CA (2004)

7. Bradshaw, J.M., Feltovich, P., Jung, H., Kulkarni, S., Taysom, W., Uszok, A.: Dimensions of adjustable autonomy and mixed-initiative interaction. In: Nickles, M., Rovatsos, M., Weiss, G. (eds.) *AUTONOMY 2003. LNCS (LNAI)*, vol. 2969. Springer, Heidelberg (2004)
8. Bradshaw, J.M., Feltovich, P.J., Jung, H., Kulkarni, S., Allen, J., Bunch, L., Chambers, N., Galescu, L., Jeffers, R., Johnson, M., Sierhuis, M., Taysom, W., Uszok, A., Van Hoof, R.: Policy-based coordination in joint human-agent activity. In: *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, The Hague, Netherlands, October 10-13 (2004)
9. Bradshaw, J.M., Jung, H., Kulkarni, S., Johnson, M., Feltovich, P., Allen, J., Bunch, L., Chambers, N., Galescu, L., Jeffers, R., Suri, N., Taysom, W., Uszok, A.: Toward trustworthy adjustable autonomy in KAOs. In: Falcone, R. (ed.) *Trusting Agents for Trustworthy Electronic Societies*. Springer, Berlin (2005)
10. Bradshaw, J.M., Feltovich, P., Johnson, M., Bunch, L., Breedy, M.R., Jung, H., Lott, J., Uszok, A.: Coordination in Human-Agent Teamwork. Invited Paper and Presentation. In: *AAAI Fall Symposium*, November 8-10 (2007)
11. Bruemmer, D.J., Marble, J.L., Dudenhoeffer, D.D.: Mutual initiative in human-machine teams. In: *Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants*, vol. 7, pp. 22–30 (2002)
12. Burke, J.J., Murphy, R.R., Coovert, M., Riddle, D.L.: Moonlight in Miami: A field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human-Computer Interaction* 19(1-2), 85–116 (2004)
13. Christofferson, K., Woods, D.D.: How to make automated systems team players. In: Salas, E. (ed.) *Advances in Human Performance and Cognitive Engineering Research*, vol. 2. JAI Press, Elsevier (2002)
14. Clancey, W.J.: *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge University Press, Cambridge (1997)
15. Clancey, W.J.: *Conceptual Coordination: How the Mind Orders Experience in Time*. Lawrence Erlbaum, Hillsdale (1999)
16. Clancey, W.J.: Simulating activities: Relating motives, deliberation, and attentive coordination. *Cognitive Systems Review* (2002); Special issue on Situated and Embodied Cognition
17. Clancey, W.J.: Automating Capcom: Pragmatic operations and technology research for human exploration of Mars. In: Cockell, C. (ed.) *Martian Expedition Planning*. AAS Science and Technology Series, vol. 107, pp. 411–430 (2004)
18. Clancey, W.J.: Roles for agent assistants in field science: Understanding personal projects and collaboration. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews* 32(2) (2004)
19. Clancey, W.J., Sierhuis, M., Damer, B., Brodsky, B.: Cognitive modeling of social behavior. In: Sun, R. (ed.) *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*. Cambridge University Press, New York City (2005)
20. Clark, H.H.: *Using Language*. Cambridge University Press, Cambridge (1996)
21. Cohen, P.R., Levesque, H.J.: *Teamwork*. SRI International, Menlo Park (1991)
22. Davis, L.E., Rough, J., Ceccine, G., Gareban-Schaefer, A., Zeman, L.L.: *Hurricane Katrina: Lessons for Army Planning and Operations*. Rand Corporation, Santa Monica (2007)
23. Feltovich, P., Bradshaw, J.M., Jeffers, R., Suri, N., Uszok, A.: Social order and adaptability in animal and human cultures as an analogue for agent communities: Toward a policy-based approach. In: Omacini, A., Petta, P., Pitt, J. (eds.) *Engineering Societies for the Agents World IV*. LNCS, vol. 3071, pp. 21–48. Springer, Berlin (2004)

24. Feltovich, P., Bradshaw, J.M., Clancey, W.J., Johnson, M.: Toward an Ontology of Regulation: Socially-based Support for Coordination in Human and Machine Joint Activity. In: O'Hare, G.M.P., Ricci, A., O'Grady, M.J., Dikenelli, O. (eds.) *ESAW 2007. LNCS (LNAI)*, vol. 4457, pp. 175–192. Springer, Heidelberg (2007)
25. Fong, T.W., Thorpe, C., Baur, C.: Robot as partner: Vehicle teleoperation with collaborative control. In: *Workshop on Multi-Robot Systems*, Naval Research Laboratory, Washington, DC (March 2002)
26. Fong, T.W., Nourbaksh, I., Ambrose, R., Simmon, R., Scholtz, J.: The peer-to-peer human-robot interaction project. In: *AIAA Space 2005* (September 2005)
27. Goodrich, M.A., Olsen Jr., D.R., Crandall, J.W., Palmer, T.J.: Experiments in adjustable autonomy. In: *Proceedings of the IJCAI_01 Workshop on Autonomy, Delegation, and Control: Interacting with Autonomous Agents*, Seattle, WA (August 2001)
28. Johnson, M., Bradshaw, J.M., Feltovich, P., Jeffers, R., Uszok, A.: A semantically-rich policy-based approach to robot control. In: *Proceedings of the International Conference on Informatics in Control, Automation, and Robotics*, Lisbon, Portugal (2006)
29. Johnson, M., Feltovich, P.J., Bradshaw, J.M., Bunch, L.: Human-robot coordination through dynamic regulation. In: *Proceedings of the International Conference on Robotics and Automation (ICRA)*, Pasadena, CA (in press, 2008)
30. Johnson, M., Intlekofer Jr., K., Jung, H., Bradshaw, J.M., Allen, J., Suri, N., Carvalho, M.: Coordinated operations in mixed teams of humans and robots. In: Marik, V., Bradshaw, J.M., Meyer, J. (eds.) *Proceedings of the First IEEE Conference on Distributed Human-Machine Systems (DHMS 2008)*, Athens, Greece (in press, 2008)
31. Kelly, G.A.: *The Psychology of Personal Constructs*. Two vols. Norton, New York (1955)
32. Kelly, G.A.: *A Theory of Personality*. W. W. Norton & Co, New York City (1963)
33. King, I., Mason, W.: *Peace at Any Price: How the World Failed Kosovo*. Cornell University Press, Ithaca (2006)
34. Klein, G., Feltovich, P.J., Bradshaw, J.M., Woods, D.D.: Common ground and coordination in joint activity. In: Rouse, W.B., Boff, K.R. (eds.) *Organizational Simulation*, pp. 139–184. John Wiley, New York City (2004)
35. Klein, G., Woods, D.D., Bradshaw, J.M., Hoffman, R., Feltovich, P.: Ten challenges for making automation a team player in joint human-agent activity. *IEEE Intelligent Systems* 19(6), 91–95 (2004)
36. Leont'ev, A.N.: The problem of activity in psychology. In: Wertsch, J.V. (ed.) *The Concept of Activity in Soviet Psychology*. M. E. Sharpe, Armonk (1979)
37. Newton, D.: Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology* 28, 28–38 (1973)
38. Norman, D.A.: Categorization of action slips. *Psychological Review* 88, 1–15 (1981)
39. Norman, D.A.: The 'problem' with automation: Inappropriate feedback and interaction, not 'over-automation'. In: Broadbent, D.E., Reason, J., Baddeley, A. (eds.) *Human Factors in Hazardous Situations*, pp. 137–145. Clarendon Press, Oxford (1990)
40. Patterson, E.S., Watts-Perotti, J., Woods, D.D.: Voice loops as coordination aids in Space Shuttle Mission Control. *Computer Supported Cooperative Work* 8, 353–371 (1999)
41. Sarter, N., Woods, D.D., Billings, C.E.: Automation surprises. In: Salvendy, G. (ed.) *Handbook of Human factors/Ergonomics*, 2nd edn. John Wiley, New York (1997)
42. Sierhuis, M., Bradshaw, J.M., Acquisti, A., Van Hoof, R., Jeffers, R., Uszok, A.: Human-agent teamwork and adjustable autonomy in practice. In: *Proceedings of the Seventh International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS)*, Nara, Japan, May 19-23 (2003)

43. Smith, W.J.: The biological bases of social attunement. *Journal of Contemporary Legal Issues* 6, 361–371 (1995)
44. Tambe, M., Shen, W., Mataric, M., Pynadath, D.V., Goldberg, D., Modi, P.J., Qiu, Z., Salemi, B.: Teamwork in cyberspace: Using TEAMCORE to make agents team-ready. In: *Proceedings of the AAAI Spring Symposium on Agents in Cyberspace*, Menlo Park, CA (1999)
45. Uszok, A., Bradshaw, J.M., Johnson, M., Jeffers, R., Tate, A., Dalton, J., Aitken, S.: KAoS policy management for semantic web services. *IEEE Intelligent Systems* 19(4), 32–41 (2004)
46. Winograd, T.: A language/action perspective on the design of cooperative work. *Human-Computer Interaction* 3(1), 3–30 (1987-1988)