Towards Flexible Teamwork
Extended Abstract

Milind Tambe
Information Sciences Institute and Computer Science Department
University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292
tambe@isi.edu
http://www.isi.edu/soar/tambe

1 Introduction

teamwork: *cooperative effort by the members of a team to achieve a common goal.* – American Heritage Dictionary

Teamwork is becoming increasingly critical in many multi-agent environments, such as, virtual training(Tambe *et al.* 1995; Rao *et al.* 1993), interactive education, internet-based information integration(Williamson, Sycara, & Decker 1996), RoboCup robotic and synthetic soccer(Kitano *et al.* 1997), interactive entertainment(Hayes-Roth, Brownston, & Gen 1995; Reilly 1996), and potential multi-robotic space missions. Teamwork in these complex, dynamic domains is more than a simple union of simultaneous coordinated activity. An illustrative example provided by Cohen and Levesque(Cohen & Levesque 1991) — worth repeating, given that the difference between simple coordination and teamwork is often unacknowledged in the literature — focuses on the distinction between ordinary traffic and driving in a convoy. Ordinary traffic is simultaneous and coordinated by traffic signs, but it is not considered teamwork. Driving in a convoy, however, is an example of teamwork. The difference in the two situations is that while teamwork does involve coordination, in addition, it at least involves a common team goal and cooperation among team members.

This short note focuses on the development of a general model of teamwork to enable a team to act coherently, overcoming the uncertainties of complex, dynamic environments. In particular, in these environments, team members often encounter differing, incomplete and possibly inconsistent views of the world and (mental) state of other agents. To act coherently, team members must flexibly communicate to avoid miscoordination. Furthermore, such environments can often cause particular team members to unexpectedly fail in fulfilling responsibilities, or to discover unexpected opportunities. Teams must thus be capable of monitoring performance, and flexibly reorganizing and reallocating resources to meet any contingencies. Unfortunately, implemented multi-agent systems often fail to provide the necessary flexibility in coordination and communication for coherent teamwork in such domains(Jennings 1994; 1995). In particular, in these systems, agents are supplied only with preplanned, domain-specific coordination. When faced with the full brunt of uncertainties of complex, dynamic domains, the inflexibility of such preplanned coordination leads to drastic failures — it is simply difficult to anticipate and preplan for all possible contingencies. Furthermore, in scaling up to increasingly complex teamwork situations, these coordination failures continually recur. In addition, since coordination plans are domain specific, they cannot be reused in other domains. Instead, coordination has to be redesigned for each new domain.

The central hypothesis in our work is that providing agents with an explicit, general model of teamwork enables them to address such difficulties. Such a model enables agents to autonomously reason about coordination and communication, providing them the requisite flexibility in teamwork. Such general models also allow reuse of teamwork capabilities across domains. Not only does such reuse save implementation effort, but it also ensures consistency in teamwork across applications(Rich & Sidner 1997). Fortunately, recent theories of teamwork have begun to provide the required models for flexible reasoning about teamwork, e.g., *joint intentions*(Cohen & Levesque 1991; Levesque, Cohen, & Nunes 1990), *SharedPlan*(Grosz 1996) and joint responsibility(Jennings 1995), are some of the prominent ones among these. However, most research efforts have failed to exploit such teamwork theories in building practical applications(Jennings 1994; 1995).

In our work(Tambe 1997b; 1997a; 1996), we have developed an *implemented* general model of teamwork, called STEAM (simply, a Shell for **TEAM**work).[1] At its core, STEAM is based on the *joint intentions* theory(Levesque, Cohen, & Nunes 1990; Cohen & Levesque 1991); but it also parallels and in some cases borrows from the *SharedPlans* theory(Grosz 1996; Grosz & Kraus 1996). Thus, while STEAM uses joint intentions as the basic building block of teamwork, as in the SharedPlan theory, team members build up a complex hierarchical structure of joint intentions, individual intentions and beliefs about others' intentions. In STEAM, communication is driven by commitments embodied in the joint intentions theory — team members may communicate to attain mutual belief while building and

[1]STEAM code (with documentation/traces) is available at **www.isi.edu/soar/tambe/steam/steam.html**.

disbanding joint intentions. Thus, joint intentions provide STEAM a principled framework for reasoning about communication, providing significant flexibility. STEAM also facilitates monitoring of team performance by exploiting explicit representation of team goals and plans. If individuals responsible for particular subtasks fail in fulfilling their responsibilities, or if new tasks are discovered without an appropriate assignment of team members to fulfill them, team reorganization can occur. Such reorganization, as well as recovery from failures in general, is also driven by the team's joint intentions.

STEAM's operationalization in complex, real-world domains has been key in its development to address important teamwork issues discussed above. Two of the domains, Attack and Transport, are based on a real-world simulation environment for training, where pilot agent teams must participate large-scale synthetic exercises with hundreds of other synthetic agents(Tambe *et al.* 1995). In a third domain, RoboCup, our player agent team is now under development for participation in the forthcoming series of (simulated) soccer tournaments, beginning at IJCAI-97. Focusing on these complex domains has led STEAM to address some practical issues, not addressed in teamwork theories. One key illustration is in STEAM's detailed attention to communication overheads and risks, which can be significant. STEAM integrates decision theoretic communication selectivity — agents deliberate upon communication necessities vis-a-vis incoherency in teamwork. This decision theoretic framework thus enables improved flexibility in communication in response to unexpected changes in environmental conditions.

Operationalizing general models of teamwork, such as STEAM, necessitates key modifications in the underlying agent architectures. Agent architectures such as Soar(Newell 1990), RAP(Firby 1987), PRS(Rao *et al.* 1993), BB1(Hayes-Roth, Brownston, & Gen 1995), and IRMA(Pollack 1992) have so far focused on individual agent's flexible behaviors via mechanisms such as commitments and reactive plans. Such architectural mechanisms need to be modified for flexible teamwork. In particular, an explicit representation of mutual beliefs, reactive *team* plans and team goals is essential. Additional types of commitments, suitable for a team context, may need to be embodied in the architectures as well. Without such architectural moorings, agents are unable to exploit general models of teamwork, and reason about communication and coordination. This view concurs with Grosz(Grosz 1996), who states that "capabilities for teamwork cannot be patched on, but must be designed in from the start".

Our operationalization of STEAM is based on modifications to the Soar architecture(Newell 1990), plus a set of about 300 domain-independent Soar rules. Three different teams have been developed based on this operationalization of STEAM in the domains discussed above, including pilot agent teams for military simulations and player teams for RoboCup soccer. Indeed, the pilot agent teams have participated in several large-scale synthetic military exercises with hundreds of agents. Here, domain experts (expert human pilots) have set the tests for these pilot teams and issued favorable written and verbal performance evaluations.

STEAM is among just a very few implemented general models of teamwork. Other models include Jennings' *joint responsibility* framework in the GRATE* system(Jennings 1995) (based on Joint Intentions theory), and Rich and Sidner's COLLAGEN(Rich & Sidner 1997) (based on the SharedPlans theory), that both operate in complex domains. STEAM significantly differs from both these frameworks, via its focus on a different (and arguably wider) set of teamwork capabilities that arise in domains with teams of more than two-three agents, with more complex team organizational hierarchies, and with practical emphasis on communication costs. In particular, STEAM's key contributions include: (i) use of joint intentions as a building block for a team's joint mental attitude (Levesque, Cohen, & Nunes 1990; Cohen & Levesque 1991) — STEAM builds up a hierarchical structure of joint intentions and individual intentions, analogous to the partial SharedPlans (Grosz & Kraus 1996); (ii) integration of novel techniques for explicit establishment of joint intentions (Smith & Cohen 1996); (iii) principled communication based on commitments in joint intentions; (iv) use of explicit monitoring facilities as well as repair methods based on joint intentions; and (v) application of decision-theoretic techniques for communication selectivity and enhancements, within the context of the joint intentions framework.

Of course, STEAM is far from a complete model of teamwork, and several major issues remain open for future work. One key issue is investigating STEAM's interactions with learning. Failure detection and recovery is another a key topic for future work. Enriching STEAM's communication capabilities in a principled fashion is also a key topic for future work. Such enriched communication may form the basis of multi-agent collaborative negotiation.

References

Cohen, P. R., and Levesque, H. J. 1991. Teamwork. *Nous* 35.

Firby, J. 1987. An investigation into reactive planning in complex domains. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*.

Grosz, B., and Kraus, S. 1996. Collaborative plans for complex group actions. *Artificial Intelligence* 86:269–358.

Grosz, B. 1996. Collaborating systems. *AI magazine* 17(2).

Hayes-Roth, B.; Brownston, L.; and Gen, R. V. 1995. Multiagent collaobration in directed improvisation. In *Proceedings of the International Conference on Multi-Agent Systems (ICMAS-95)*.

Jennings, N. 1994. Commitments and conventions: the foundation of coordination in multi-agent systems. *The Knowledge Engineering Review* 8.

Jennings, N. 1995. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence* 75.

Kitano, H.; Tambe, M.; Stone, P.; Veloso, M.; Noda, I.; Osawa, E.; and Asada, M. 1997. The robocup synthetic agents' challenge. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.

Levesque, H. J.; Cohen, P. R.; and Nunes, J. 1990. On acting together. In *Proceedings of the National Conference on Artificial Intelligence*. Menlo Park, Calif.: AAAI press.

Newell, A. 1990. *Unified Theories of Cognition*. Cambridge, Mass.: Harvard Univ. Press.

Pollack, M. 1992. The uses of plans. *Artificial Intelligence* 57:43–68.

Rao, A. S.; Lucas, A.; Morley, D.; Selvestrel, M.; and Murray, G. 1993. Agent-oriented architecture for air-combat simulation. Technical Report Technical Note 42, The Australian Artificial Intelligence Institute.

Reilly, W. S. 1996. *Believable Emotional and Social Agents*. Ph.D. Dissertation, School of Computer Science, Carnegie Mellon University.

Rich, C., and Sidner, C. 1997. COLLAGEN: When agents collaborate with people. In *Proceedings of the International Conference on Autonomous Agents (Agents'97)*.

Smith, I., and Cohen, P. 1996. Towards semantics for an agent communication language based on speech acts. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*.

Tambe, M.; Johnson, W. L.; Jones, R.; Koss, F.; Laird, J. E.; Rosenbloom, P. S.; and Schwamb, K. 1995. Intelligent agents for interactive simulation environments. *AI Magazine* 16(1).

Tambe, M. 1996. Teamwork in real-world, dynamic environments. In *Proceedings of the International Conference on Multi-agent Systems (ICMAS)*.

Tambe, M. 1997a. Agent architectures for flexible, practical teamwork. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*.

Tambe, M. 1997b. Implementing agent teams in dynamic multi-agent environments. *Applied Artificial Intelligence*. (to appear).

Williamson, M.; Sycara, K.; and Decker, K. 1996. Executing decision-theoretic plans in multi-agent environments. In *Proceedings of the AAAI Fall Symposium on Plan Execution: Problems and Issues*.