Doctoral Consortium Abstract in *Proceedings of the 16th National Conference on Artificial Intelligence (AAAI-99).* July, 1999.

Execution Monitoring and Diagnosis in Multi-Agent Environments

Gal A. Kaminka

Computer Science Department and Information Sciences Institute
University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292
galk@isi.edu

Agents in complex, dynamic, multi-agent environments face uncertainty in the execution of their tasks, as their sensors, plans, and actions may fail unexpectedly, e.g., the weather may render a robot's camera useless, its grip too slippery, etc. The explosive number of states in such environments prohibits any resource-bounded designer from predicting all failures at design time. This situation is exacerbated in multi-agent settings, where interactions between agents increase the complexity. For instance, it is difficult to predict an opponent's behavior.

Agents in such environments must therefore rely on run-time execution monitoring and diagnosis to detect a failure, diagnose it, and recover. Previous approaches have focused on supplying the agent with *goal-attentive* knowledge of the ideal behavior expected of the agent with respect to its goals. These approaches encounter key pitfalls and fail to exploit key opportunities in multi-agent settings: (a) only a subset of the sensors (those that measure achievement of goals) are used, despite other agents' sensed behavior that can be used to indirectly sense the environment or complete the agent's knowledge; (b) there is no monitoring of social relationships that must be maintained between the agents regardless of achievement of the goal (e.g., teamwork); and (c) there is no recognition of *failures in others*, though these change the ideal behavior expected of an agent (for instance, assisting a failing teammate).

To address these problems, we investigate a novel complementary paradigm for multi-agent monitoring and diagnosis. *Socially-Attentive Monitoring* (SAM) focuses on monitoring the social relationships between the agents as they are executing their tasks, and uses models of multiple agents and their relationships in monitoring and diagnosis. We hypothesize that failures to maintain relationships would be indicative of failures in behavior, and diagnosis of relationships can be used to complement goal-attentive methods. In particular, SAM addresses the weaknesses listed above: (a) it allows inference of missing knowledge and sensor readings through other agents' sensed behavior; (b) it directly monitors social relationships, with no attention to the goals; and (c) it allows recognition of failures in others (even if they are not using SAM!).

SAM uses the STEAM teamwork model, and a role-similarity relationship model to monitor agents. It relies on plan-recognition to infer agents' reactive-plan hierarchies from their observed actions. These hierarchies are compared in a top-down fashion to find relationship violations, e.g., cases where two agents selected different plans despite their being on the same team. These trigger diagnosis that uses the relationship models to facilitate recovery. For example, in teamwork, a commitment to joint selection of plans further mandates mutual belief in preconditions. Thus a difference in selected plans may be explained by a difference in preconditions, and can lead to recovery using negotiations.

We empirically and analytically investigate SAM in two dynamic, complex, multi-agent domains: the ModSAF battlefield simulation, where SAM is employed by helicopter pilot agents; and the RoboCup soccer simulation where SAM is used by a coach agent to monitor teams' behavior. We show that SAM can capture failures that are otherwise undetectable, and that distributed monitoring is better and simpler (sound and complete detection, no representation of ambiguity) than a centralized scheme (complete and unsound, requiring representation of ambiguity). Key contributions and novelties include: (i) a general framework for socially-attentive monitoring, and a deployed implementation for monitoring teamwork; (ii) rigorously proven guarantees on the applicability and results of practical socially-attentive monitoring of teamwork under conditions of uncertainty; (iii) procedures for diagnosis based on a teamwork relationship model. Future work includes the use of additional relationship models and formalization of social diagnosis capabilities.

An example. Three helicopters (using SAM) were to fly to a specified land-mark, and switch to a scouting plan in which two of them land. One pilot did not see the land mark, while its team-mate detected it and landed. Using SAM, the agents involved detected that the team is no longer in agreement on the plan being executed. Diagnosis lead them to realize that the landmark was seen only by some pilots, and the failure was recovered from.

References

Kaminka, G. A., and Tambe, M. 1999. I'm OK, You're OK, We're OK: Experiments in Distributed and Centralized Relational Execution Monitoring. In *Proc. of the 3rd International Conf. on Autonomous Agents (Agents*-99).

Copyright © 1999 American Association for Artificial Intelligence (www.aaai.org). All rights reserved.