Exploiting Agent Diagnosis for Plan Repair in MAS (Extended Abstract)

Roberto Micalizio
Università di Torino
corso Svizzera 185
10149 - Torino, Italy
micalizio@di.unito.it

Pietro Torasso
Università di Torino
corso Svizzera 185
10149 - Torino, Italy
torasso@di.unito.it

ABSTRACT

In this paper we formalize a local strategy for plan repair in a Multi Agent Plan, where agents exhibit a collaborative behavior to reach a common global goal in a partially observable environment. The recovery strategy is based on a plan failure analysis where two main results are inferred: the agent diagnosis (which explains the action failure in terms of faults in agent functionalities) and the set of missing goals (i.e. services that the faulty agent can no longer provide to the other agents). Relying on the agent diagnosis, a faulty agent activates a local re-planner to restore the healthy status in its functionalities and achieve the set of missing goals.

Categories and Subject Descriptors

I.2.11 [**Multiagent Systems**]; I.2.8 [**Problem Solving, Control Methods, and Search**]: *Plan execution, formation, and generation*

Keywords

Agent Reasoning, Multi-Agent Plan Recovery, Multi-Agent Plan Diagnosis

1. INTRODUCTION

A Multi-Agent Plan (MAP) is a system where actions are carried on concurrently by a team of cooperative agents, which typically are spatially distributed in the environment. The adoption of a MAP presents a number of benefits (e.g., flexibility, robustness, parallelism) in solving real complex tasks. However, the execution of MAP in real-world scenarios can be affected by unexpected plan threats, whose occurrence can cause the failure of one or more actions in the MAP, and consequently the interruption of the plan progress. In some cases, however, the agents can still complete their activities despite the occurrence of plan threats, but the given MAP needs to be revised i.e., the MAP must be repaired.

The problem of repairing a MAP has been recently addressed both in the scenario of *self-interested agents* (e.g., [6]) and in the scenario of *collaborative agents* (e.g., [2, 3]). In these approaches each agent has its own local plan and goals, and possibly interact by competing or cooperating with other agents. Plan repair typically consists in a dynamic reallocation of resources to recover from unexpected interaction flaws rising during the execution. Only in [3] a recovery strategy driven by a diagnostic process is discussed. Wielding action failures is not an easy task as the occurrence of faults is unpredictable and their effects may be non deterministic; moreover, the environment where the agents act is just partially observable, hence the status of the agents is not precisely known at each time instant. It follows that precompiled recovery solutions cannot easily be adopted as they should anticipate all the possible conditions in which a fault can occur and all its possible effects on the system.

Also in this paper we advocate a distributed approach, where the global MAP is decomposed into as many local plans as the agents in the team, and each of these local plans is executed by a specific agent in the team. However, rather than coordination flaws, in this paper plan threats are physical faults affecting the functionalities of the agents.

2. THE PROPOSED APPROACH

The solution we propose consists in establishing a distributed control for the MAP execution; more precisely, each agent supervises the actions it executes by performing, in a closed loop, three fundamental activities: *on-line monitoring*, *plan failure analysis,* and *plan repair.* The monitoring task estimates the status of the agent at each time instant and detects action failures. Whenever an action failure is detected the plan failure analysis is activated; such an analysis involves two main inferences: the *agent diagnosis*, which explains the action failure in terms of unhealthy agent functionalities; and the set of *missing goals*, namely, services that the agent can no longer provide given the failure.

Once the plan failure analysis has been completed a local plan repair strategy is attempted. In our scenario we consider two kinds of faults: *recoverable faults*, that can be autonomously repaired by the agent in trouble by means of a proper sequence of repairing actions, and *not recoverable faults* (no recovery actions exist). For this reason the repair strategy we propose consists of two steps. In the first step, the agent in trouble synthesizes on-the-fly a new local plan including the repairing actions required for restoring a healthy condition in its functionalities. This re-planning phase heavily relies on the results of the plan failure analysis; in fact, the agent diagnosis highlights the functionalities that are assumed to be unhealthy ; while the set of missing goals drives the re-planning process as, in principle, it is sufficient to find an alternative way for providing the system with these goals to reach the global goal of the MAP despite

Cite as: Exploiting Agent Diagnosis for Plan Repair in MAS, (Extended Abstract), Roberto Micalizio, Pietro Torasso, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1231–1232

the occurrence of the action failure.

The repairing plan, however, may not exist (e.g., the action failure is due to not recoverable faults). Therefore, the second step of recovery strategy tries to mitigate the effects of the failure on the global system by moving the agent in trouble into a *safe status*, where the agent does not represent a latent menace for other team members. When the action failure cannot be recovered by means of a local strategy, global (or team level) strategies need to be activated; for example a new global MAP could be inferred where resources and goals originally assigned to the faulty agent are redistributed to the other agents. Due to space reasons we will sketch just the first step of the repair strategy.

3. A LOCAL STRATEGY TO PLAN REPAIR

As formalized by Cox et al. in [1] the global MAP is the tuple $\langle A, E, CL, CC, NC \rangle$, where A is the set of action instances the agents in the team $\mathcal{T}$ have to execute; E is the set of precedence links $a \prec a'$ between actions in A; CL is a set of causal links of the form $l : a \xrightarrow{q} a'$; namely, the action a provides the action a' with the service q; finally, CC and NC are respectively *concurrency* and *non-concurrency* symmetric relations over action instances in A.

The execution of the MAP is distributed: each agent i in $\mathcal{T}$ executes the sequence of actions $[a_0^i, a_1^i, \ldots, a_\infty^i]$ within a specific local plan P^i, which reaches the local goal G^i.

In the following we will briefly recall the main steps in local a control loop performed by each agent in $\mathcal{T}$.

Monitoring. The monitoring task has been deeply discussed in [5]; here it is sufficient to say that, given the partial observability of the environment, the agent i cannot precisely determine its own status at each time instant t, rather the agent i can just estimate a belief state $\mathcal{B}^i(t)$; i.e, a set of alternative agent states which are all consistent with the observations received so far by the agent. Given the belief state $\mathcal{B}^i(t)$, the agent i can also determine the outcome of the last action it has executed. Intuitively, an action is *successfully completed* when its expected nominal effects hold in every state included in the belief state $\mathcal{B}(t)^i$; the action is *failed* otherwise.

Plan Failure Analysis. *Agent diagnosis.* Intuitively, given the agent belief state $\mathcal{B}^i(t)$, the agent diagnosis D^i can be inferred by projecting every state in $\mathcal{B}^i(t)$ over the status variables *healthVar* referring to the health status of the agent functionalities. Observe that, since $\mathcal{B}^i(t)$ is in general ambiguous, the agent diagnosis may result to be ambiguous too; namely, the agent diagnosis D^i is a set of alternative explanations, such that each explanation $exp \in D^i$ is a complete assignment of values to the status variables in *healthVar*.

Missing Goals. The failure of an action a_l^i prevents the execution of the actions in the plan segment $[a_{l+1}^i, \ldots, a_\infty^i]$. Therefore, all the services provided by these actions will never be provided to the system. The set of these services is denoted as the set of *missing goals*; singling out these services is essential as, in principle, it would be sufficient to find an alternative way to provide them in order to reach the global goal of the MAP despite the action failure; formally:

DEFINITION 1. *The set $\mathcal{M}G(i)$ of missing goals is the set of services q such that, for each action $a_k^i \in [a_l^i, a_\infty^i]$ the following conditions hold*

i. q is an expected nominal effect of action a_k^i and
ii. it holds either
- q is an atom of the local goal G^i
- $\exists$ a causal link $l \in CL$ such that $l : a_k^i \xrightarrow{q} a_h^j$; i.e., the action a_k^i provides the action a_h^j (assigned to another agent j) with the service q.

Plan repair. The plan repair step substitutes the plan segment $[a_l^i, \ldots, a_\infty^i]$ with a repairing plan P^{*i}, inferred on-the-fly, whose goal G^{*i} is defined as follows:

DEFINITION 2. $G^{*i} = \mathcal{M}G(i) \cup healthy(D^i)$

Where $healthy(D^i)$ is the healthy status of the agent that must be restored by the recovery plan; namely, let $\{f_1, ..., f_n\}$ be the set of functionalities assumed to be faulty in D^i, the healthy formula results as $healthy(D^i) = \neg ab(f_1) \wedge \ldots \wedge \neg ab(f_n)$. To reach the (new) local goal G^{*i} the agent exploits a conformant planner (see [4]); in fact, the repairing plan P^{*i} exists *iff* it is executable no matter the actual health status of the agent is (note that the agent diagnosis D^i is, in general, a set of alternative explanations).

4. CONCLUSIONS.

In this paper we have sketched a local strategy for plan repair through a closed loop of control. Two main reasons support such a local strategy. First, even though global repair strategies could infer more efficient recovery plans, these solutions tend to be expensive, especially because the system is just partially observable: a global replanner has to deal with ambiguous global states. The local strategy we propose, instead, has to deal with just the ambiguity of the agent in trouble, and the solution is computationally cheaper. Second, the strategy is driven by the local plan failure analysis, that is, it is not necessary to anticipate all the possible circumstances in which a fault can occur, the repair is performed taking into account the (assumed) current status of the agent.

Preliminary results we have collected in a simulated environment of service robots are very promising. For example, considering the execution of 30 MAPs, each of which was perturbed by the injection of one fault, the team of robots reaches (on average) 80% of the assigned sub-goals when the local repair strategy is switched on; whereas just the 20% of the sub-goals are reached when the local repair is off and no other repair process is activated. These results encourage further steps in the field of MAP repair, including the integration of local and global strategies.

5. REFERENCES

[1] J. S. Cox, E. H. Durfee, and T. Bartold. A distributed framework for solving the multiagent plan coordination problem. In *Proc. AAMAS05*, pages 821–827, 2005.

[2] K. Decker and J. Li. Coordinating mutually exclusive resources using gpgp. *JAAMAS*, 3(2):113–157, 2000.

[3] B. Horling and V. Benyo, B. Lesser. Using self-diagnosis to adapt organizational structures. In *Proc. ICAA'01*, pages 529–536, 2001.

[4] R. Micalizio and P. Torasso. Recovery from plan failures in partially observable environments. In *Research and Development in Intelligent Systems XXVII*, pages 321–334, 2007.

[5] R. Micalizio and P. Torasso. Monitoring the execution of a multi-agent plan:dealing with partial observability. In *Proc. of ECAI'08*, pages 408–412, 2008.

[6] R. van der Krogt and de Weerdt. Plan repair as an extension of planning. In *Proc. of ICAPS'05*, pages 284–259, 2005.