Diagnosis of Single and Multi-Agent Plans

Cees Witteveen
Faculty EEMCS
Delft University of Technology
P.O.Box 5031, NL-2600 GA Delft
witt@ewi.tudelft.nl

Nico Roos
Dept. of Computer Science
Universiteit Maastricht
P.O.Box 616, NL-6200 MD Maastricht
roos@cs.unimaas.nl

Roman van der Krogt
Faculty EEMCS
Delft University of Technology
P.O.Box 5031, NL-2600 GA Delft
r.p.j.vanderkrogt@ewi.tudelft.nl

Mathijs de Weerdt
Faculty EEMCS
Delft University of Technology
P.O.Box 5031, NL-2600 GA Delft
m.m.deweerdt@ewi.tudelft.nl

ABSTRACT

We discuss the application of Model Based Diagnosis in agent-based planning. We model a plan as a system to be diagnosed and assume that agents can monitor the execution of the plan by making partial observations during plan execution. These observations are used by the agents to explain plan deviations (errors) by qualifying some action instances as behaving abnormally. We prefer those qualifications that are maximum informative, i.e. explain as much as possible. Since in a plan several instances of the same action might occur, an error occurring in one instance might be used to predict the occurrence of the same error in an action instance to be executed later on. To account for such correlations, we introduce causal rules to generate diagnoses from action instances qualified as abnormally and we introduce *Pareto minimal causal diagnoses* as the right extension of classical minimal diagnoses.

Next, we consider the multi-agent perspective where each agent is responsible for a part of the total plan, we show how plan-diagnoses of these partial plans are related to diagnoses of the total plan and how global diagnoses can be obtained in a distributed way.

Categories and Subject Descriptors

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

General Terms

Reliability, Theory

Keywords

Model-Based Diagnosis, Planning

1. INTRODUCTION

Model-Based Diagnosis (MBD) [4, 5, 13] is a well-known technique to infer abnormalities of internal components of a given system S from its input-output behavior. To this end a model of S is given, where the set of components C, the possible behaviors of each of the components $c \in C$ and their relations have been specified. Usually, for each component c several *health modes* are distinguished: a *normal* mode and several *fault* modes. For each health mode of c a particular behavior of c is given. Once the health mode of each component $c \in C$ is specified the behavior of the total system is defined and the output of the system S can be inferred from its input unambiguously. The diagnostic engine is triggered whenever, under the assumption that all components are functioning normally, there is a discrepancy between the output as predicted from the input observations, and the actually observed output. The result of MBD then is a suitable assignment of health modes to the components, called a *diagnosis*, such that the actually observed output is *consistent* with this health mode qualification or can be *explained* by this qualification. Usually, in a diagnosis one requires the number of components qualified as abnormally to be minimized.

Our contribution in this paper is an adaptation and extension of MBD to both single agent and multi-agent *planning systems*. First of all, to enable plan diagnosis, we will introduce an *object-oriented* description of actions in contrast to the traditional *state based approach* [7]. In this approach action instances in a plan (also called *steps* cf. [12]) can be viewed as components whose outputs influence the states of objects we are interested in. This object-oriented view of actions is closely related to the resource-oriented view presented in [17], which offers important advantages in multiagent planning.

The object oriented view of plans enables us to apply the MBD-approach to plan execution of single and multiagent plans. By viewing action instances as the components and inputs and outputs of a component as objects of which the status is influenced by an action, the plan itself can be viewed as the structural description of a system. This description offers us a very natural model for describing and analyzing plan execution. The relations between plan steps (action instances) are specified by viewing them as instances of the more abstract notion of an *action scheme* (also called a *plan operator* cf. [12]). At the level of an action scheme we define the generic behavior of the action.

Secondly, we will introduce the notion of *plan diagnosis* in a single agent system, by showing how the object oriented description of plans enables us to apply the standard MBD approach to plans using (partial) *observations* of *plan states*. Observations plus assumptions about the (mal)functioning of action instances can be used to make predictions about future plan states. Distinguishing between normal and abnormal execution of actions in a plan, we introduce inclusion-minimal sets of actions qualified as abnormal to explain deviations between expected plan states and observed plan states. Such minimal qualifications directly correspond to minimal diagnoses. Within such minimal plan diagnoses we distinguish maximum informative qualifications that can be used to enhance the explanatory power of diagnoses.

Remark Within the MBD-approach one usually distinguishes the classical system description approach and the Discrete Event Systems (DES) [1] approach. Instead of a system consisting of a set of interrelated components, DES uses a more abstract description based on a finite state machine representation of the components and is especially suited for modeling dynamic systems. Since plan execution is a dynamic process, plan diagnosis by modeling the plan as a DES may seem to be an obvious choice. The problem, however, is that the components distinguished in a DES are meant to be reactive finite automata, where state transitions take place as the consequence of external inputs or messages received from neighboring links. Modeling the action instances as our primary components would imply that either the automata representing the actions would become trivial or we would be forced to model the combined states of all objects as a single state of the finite state machine, and the actions themselves as events causing state transitions. For future extensions of our approach we see possibilities to adapt the DES approach, especially if the fault modes of the components are further differentiated. ■

Thirdly, we introduce the concept of a *causal diagnosis*. The idea of establishing a minimal diagnosis in MBD is governed by the principle of *minimal change*: explain the abnormalities in the behavior observed by changing the qualification from normal to abnormal for as few system components as necessary. Using this principle is intuitively acceptable if the components qualified as abnormal are failing *independently*. However, if there exist *dependencies* between such components, the choice for minimal diagnoses cannot be justified. As we will argue, the existence of dependencies between failing actions in a plan is often the rule instead of an exception. Therefore, we will refine the concept of a plan diagnosis by introducing the concept of a *causal diagnosis*. We relate the anomalous execution of actions to anomalous execution of other actions in the form of causal rules. These rules enable us to replace a set of dependent failing actions (e.g. a plan diagnosis) by a set of unrelated *causes* of the original diagnosis. This independent and usually smaller set of causes constitutes a causal diagnosis, consisting of a set of failing actions. Such a causal diagnosis always generates a cover of a minimal diagnosis. More importantly, such causal diagnoses can also be used to predict failings of actions that have to be executed in the plan and thereby also can be used to assess the consequences of such failures for goal realizability.

Finally, we concentrate on multi-agent plan diagnosis. Here, the agents together are assumed to execute a common plan that is partitioned over the agents. Each agent is responsible for the execution of its sub-plan and has to respect the dependencies with sub-plans of other agents. We show that global diagnoses can be obtained in a distributive way by establishing partial diagnoses of the sub-plans.

The remainder of this paper is organized as follows: In the next section, we discuss some approaches related to plan-based diagnosis. Section 3 introduces the preliminaries of plan-based diagnosis, while Section 4 formalizes plan-based diagnosis. Section 5 extends the formalization to multi-agent plans. Section 6 concludes the paper.

2. RELATED RESEARCH

In this section we briefly discuss some related approaches to plan diagnosis. Like we use MBD as a starting point to plan diagnosis, Birnbaum et al. [2] apply MBD to *planning agents* relating health states of agents to *outcomes* of their planning activities. They do not take into account faults that can be attributed to actions occurring in a plan as a separate source of errors. Instead of focusing upon the relationship between agent properties and outcomes of plan executions, we take a more detailed approach, distinguishing two separate sources of errors (actions and properties of the executing agents) and focusing upon the detection of anomalies during the plan execution. This enables us to predict the outcomes of a plan beforehand instead of using them only as observations.

De Jonge et al. [6] propose another approach that directly applies model-based diagnosis to plan execution. Their paper focuses on agents each having an individual plan, and where conflicts between these plans may arise (e.g. if they require the same resource). Diagnosis is applied to determine those factors that are accountable for *future* conflicts. The authors, however, do not take into account dependencies between health modes of actions and do not consider agents that collaborate to execute a common plan.

Kalech and Kaminka [10, 11] apply *social diagnosis* in order to find the cause of an anomalous plan execution. They consider hierarchical plans consisting of so-called *behaviors*. Such plans do not prescribe a (partial) execution order on a set of actions. Instead, based on its observations and beliefs, each agent chooses the appropriate behavior to be executed. Each behavior in turn may consist of primitive actions to be executed, or of a set of other behaviors to choose from. Social diagnosis then addresses the issue of determining what went wrong in the joint execution of such a plan by identifying the disagreeing agents and the causes for their selection of incompatible behaviors (e.g., belief disagreement, communication errors). This approach might complement our approach when conflicts not only arise as the consequence of faulty actions, but also as the consequence of different selections of sub-plans in a joint plan.

Lesser et al. [3, 9] also apply diagnosis to (multi-agent) plans. Their research concentrates on the use of a *causal model* that can help an agent to refine its initial diagnosis of a failing *component* (called a *task*) of a plan. As a consequence of using such a causal model, the agent would be able to generate a new, situation-specific plan that is better suited to pursue its goal. While their approach in its ultimate intentions (establishing anomalies in order to find a suitable plan repair) comes close to our approach, their approach to diagnosis concentrates on specifying the exact causes of the failing of one single *component* (task) of a plan. Diagnosis is based on observations of a component without taking into account the consequences of failures of such a component w.r.t. the remaining plan. In our approach, instead, we are interested in applying MBD-inspired methods to *detect* plan failures. Such failures are based on observations during plan execution and may concern individual components of the plan, but also agent properties. Furthermore, we do not only concentrate on failing components themselves, but also on the consequences of these failures for the future execution of plan elements.

3. PRELIMINARIES

We start with considering plan-based diagnosis as a simple extension of the model-based diagnosis approach where the model is not a description of an underlying system but a *plan* of an agent.

States We take an *object-* or *resource-based* view on the world, assuming that for the planning problem at hand, the world can be simply described by a set $Obj = \{o_1, o_2, \ldots, o_n\}$ of objects, their respective *value domains* S_i and and their (current) values $s_i \in S_i$.[1] A *state of the world* σ then is an element of $S_1 \times S_2 \times \ldots \times S_n$. It will not always be possible to give a complete state description. Therefore, we introduce a *partial state* as an element $\pi \in S_{i_1} \times S_{i_2} \times \ldots \times S_{i_k}$, where $1 \leq i_1 < \ldots < i_k \leq n$. We use $O(\pi)$ to denote the set of objects $\{o_{i_1}, o_{i_2}, \ldots, o_{i_k}\} \subseteq Obj$ specified in such a (partial) state π. The value s_j of object $o_j \in O(\pi)$ in π will be denoted by $\pi(j)$. The value of an object $o_j \in Obj$ not occurring in a partial state π is said to be unknown (or unpredictable) in π, denoted by $\perp$. Partial states can be ordered with respect to their information content: π is said to be contained in π', denoted by $\pi \sqsubseteq \pi'$, iff $O(\pi) \subseteq O(\pi')$ and $\pi'(j) = \pi(j)$ for every $o_j \in O(\pi)$. We say that two partial states π, π' are *equivalent* modulo a set of objects O, denoted by $\pi =_O \pi'$, if for every $o_j \in O$, $\pi(j) = \pi'(j)$. Finally, we define the partial state π restricted to a given set O, denoted by $\pi \upharpoonright O$, as the state $\pi' \sqsubseteq \pi$ such that $O(\pi') = O \cap O(\pi)$.

Goals An (elementary) goal g of an agent specifies a set of states an agent wants to bring about using a plan. Here, we specify each such a goal g as a constraint, that is a relation over some product $S_{i_1} \times \ldots \times S_{i_k}$ of domains. We say that a goal g is satisfied by a partial state π, denoted by $\pi \models g$, if the relation g contains some tuple (partial state) $(v_{i_1}, v_{i_2}, \ldots v_{i_k})$ such that $(v_{i_1}, v_{i_2}, \ldots v_{i_k}) \sqsubseteq \pi$. We assume each agent to have a set G of such elementary goals $g \in G$. We use $\pi \models G$ to denote that all goals in G hold in π, i.e. for all $g \in G$, $\pi \models g$.

Action schemes An *action scheme* or *plan operator* α is represented as a function that replaces the values of a subset $O_\alpha \subseteq Obj$ by other values, dependent upon the values of another set $O'_\alpha \supseteq O_\alpha$ of objects. Hence, every action scheme α can be modeled as a (partial) function $f_\alpha : S_{i_1} \times \ldots \times S_{i_k} \to S_{j_1} \times \ldots \times S_{j_l}$, where $1 \leq i_1 < \ldots < i_k \leq n$ and $\{j_1, \ldots, j_l\} \subseteq \{i_1, \ldots, i_k\}$. The objects whose value domains occur in $dom(f_\alpha)$ will be denoted by $dom_O(\alpha) = \{o_{i_1}, \ldots, o_{i_k}\} = O'_\alpha$ and, likewise $ran_O(\alpha) = \{o_{j_1}, \ldots, o_{j_l}\} = O_\alpha$. Note that $ran_O(\alpha) \subseteq dom_O(\alpha)$. This functional specification f_α constitutes the *normal* behavior of the action, denoted by f_α^{nor}. The correct execution of an action may fail either because of an inherent malfunctioning or because of a malfunctioning of an agent responsible for executing the action, or because of unknown external circumstances. In all these cases we would like to model the effects of executing such failed actions. Therefore, we introduce a set of *health modes* M_α for each action scheme α. M_α contains at least the normal mode *nor*, the mode *ab* indicating the most general abnormal behavior, and possibly several other specific fault modes. The most general abnormal behavior of action scheme α is specified by the function f_α^{ab}, where $f_\alpha^{ab}(s_{i_1}, s_{i_2}, \ldots, s_{i_k}) = (\perp, \perp, \ldots, \perp)$.[2] To keep the discussion simple, in the sequel we distinguish only the health modes *nor* and *ab*. Specific action instances of the action scheme α will be denoted by small roman letters a_i. If $type(a_i) = \alpha$, such an instance a_i is said to be of type α. If the context permits we will use "actions" and "instances of actions" interchangeably.

Plans A plan is a partial order $P = \langle A, < \rangle$ where A is a set of instances of actions. The partial order specifies a precedence relation between these instances: $a < a'$ implies that the (instance) a must finish before a' may start. We will denote the *transitive reduction* of $<$ by $\ll$, i.e., $\ll$ is the smallest subrelation of $<$ whose transitive closure equals $<$.

We assume that if in a plan P two action instances a and a' are $<$-independent, in principle they may be executed concurrently. This means that the dependency relation $<$ at least should capture all resource dependencies that would prohibit concurrent execution of actions. Therefore, we assume $<$ to satisfy the following *concurrency requirement*:

If $ran_O(a) \cap dom_O(a') \neq \varnothing$ then $a < a'$ or $a' < a$.[3]

That is, for concurrent instances, domains and ranges do not overlap. If the range of a overlaps with the domain of a', the action a may influence the effect of the action a' depending on the order in which a and a' are executed.

Figure 1 gives an illustration of a plan. Arrows denote the objects an action uses as inputs. In this plan, the dependency relation is specified as $a_1 \ll a_3$, $a_2 \ll a_4$, $a_4 \ll a_5$, $a_4 \ll a_6$ and $a_1 \ll a_5$. Note that the last dependency has to be included because a_5 changes the value of o_2 needed by a_1. The action a_1 shows that not every object occurring in the domain of an action needs to be affected by the action. The actions a_5 and a_6 illustrate that concurrent actions may have overlapping domains.

Qualifications In order to predict the result of an anomalous plan execution, we introduce the notion of a *qualified* plan. A qualified version P_Q of a plan $P = (A, <)$ is a tuple $P_Q = (A, <, Q)$, where $Q \subseteq A$ is the subset of actions

[1] In contrast to the conventional approach to state-based planning, cf. [7].

[2] This definition implies that the behavior of abnormal actions is essentially unpredictable.

[3] Note that since $ran_O(a) \subseteq dom_O(a)$, this requirement excludes overlapping ranges of concurrent actions, but domains of concurrent actions are allowed to overlap as long as the values of the object in the overlapping domains are not affected by the actions.

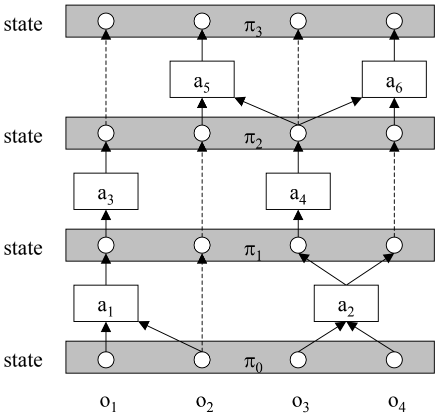

Figure 1: Plans & states

qualified as abnormal and therefore, $A - Q$ the subset of actions qualified as normal. Since a qualification Q corresponds with assigning the health mode ab to every action in Q and since $f_a^{ab}(s_{i_1}, s_{i_2}, \dots, s_{i_k}) = (\bot, \bot, \dots, \bot)$ for every action instance $a \in Q$, the results of anomalously executed actions are unpredictable.[4]

Note that a normal plan execution of plan P corresponds with the qualified version $P_\varnothing$.

Plan execution For simplicity, we will assume that every action takes a unit of time to execute. We are allowed to observe the execution of a plan P at discrete times $t = 0, 1, 2, \dots, k$ where k is the depth of the plan, i.e., the longest $<$-chain of actions in P. Let $depth_P(a)$ be the depth of action a in plan $P = (A, <)$ that is, $depth_P(a) = 0$ if $\{a' \mid a' \ll a\} = \varnothing$ and $depth_P(a) = 1 + max\{depth_P(a') \mid a' \ll a\}$, else. If the context is clear, we often will omit the subscript P. We assume that the plan starts to be executed at time $t = 0$ and that concurrency is fully exploited, i.e., all actions a with $depth(a) = 0$ are completed at time $t = 1$ and every action a with $depth(a) = k$ will be started at time k and will be completed at time $k + 1$. Note that thanks to the above specified *concurrency requirement* concurrent execution of actions having the same depth leads to a well-defined result.

Let P_t denote the set of actions a with $depth(a) = t$, let $P_{>t} = \bigcup_{t'>t} P_{t'}$, $P_{<t} = \bigcup_{t'<t} P_{t'}$ and $P_{[t,t']} = \bigcup_{k=t}^{t'} P_k$. We say that an action a is *enabled* in a state σ if $dom_O(a) \subseteq O(\sigma)$.

Execution of P on a given initial state σ_0 will induce a sequence of states $\sigma_0, \sigma_1, \dots, \sigma_k$, where σ_{t+1} is generated from σ_t by applying the subset of actions P_t enabled in σ_t.

This idea can be easily generalized to inducing a sequence of *partial* states using the action instances occurring in P given a (partial) state π at time $t \geq 0$, denoted by (π, t).

We say that $(\pi', t+1)$ is (directly) generated by execution of P_Q from (π, t), abbreviated by $(\pi, t) \rightarrow_{Q;P} (\pi', t+1)$, iff the following conditions hold:

1. $\pi' \upharpoonright ran_O(a) = f_a^{nor}(\pi \upharpoonright dom_O(a))$ for each $a \in P_t - Q$ with $dom_O(a) \subseteq O(\pi)$, that is, the consequences of all actions a enabled in π can be predicted and occur in π'.
2. $O(\pi') \cap ran_O(a) = \varnothing$ for each $a \in Q \cap P_t$, since the result of executing an abnormal action cannot be predicted (even if such an action is enabled in π);
3. $O(\pi') \cap ran_O(a) = \varnothing$ for each $a \in P_t$ with $dom_O(a) \not\subseteq O(\pi)$, that is, even if an action a is enabled in (the complete state) σ_t, if a is not enabled in $\pi \sqsubseteq \sigma_t$, the result is not predictable and therefore does not occur in π', since it is not possible to predict the consequences of actions that depend on values not defined in π.
4. $\pi'(i) = \pi(i)$ for each $o_i \notin ran_O(P_t)$, that is, the value of any object not occurring in the range of an action in P_t should remain unchanged. Here, $ran_O(P_t)$ denotes the union of the sets $ran_O(a)$ with $a \in P_t$.

For arbitrary values of $t \leq t'$ we say that (π', t') is (directly or indirectly) generated by execution of P_Q from (π, t), denoted by $(\pi, t) \rightarrow^*_{Q;P} (\pi', t')$, iff the following conditions hold: *(i)* if $t = t'$ then $\pi' = \pi$; *(ii)* if $t' = t + 1$ then $(\pi, t) \rightarrow_{Q;P} (\pi', t')$; *(iii)* if $t' > t + 1$ then there must exists some state $(\pi'', t' - 1)$ such that $(\pi, t) \rightarrow^*_{Q;P} (\pi'', t' - 1)$ and $(\pi'', t' - 1) \rightarrow_{Q;P} (\pi', t')$.

Note that $(\pi, t) \rightarrow^*_{\varnothing;P} (\pi', t')$ denotes the a normal execution of a normal plan $P_\varnothing$. Such a normal plan execution will also be denoted by $(\pi, t) \rightarrow^*_P (\pi', t')$.

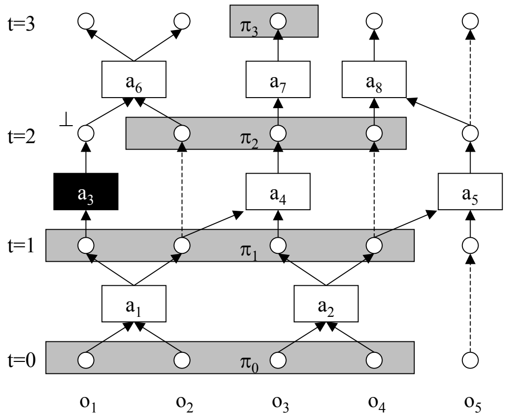

Figure 2: Execution of abnormal actions

Figure 2 gives an illustration of an execution of a plan with abnormal actions. Suppose action a_3 is qualified as abnormal and generates a result that is unpredictable ($\bot$). Given the qualification $Q = \{a_3\}$ and the partially observed state π_0 at time point $t = 0$, we predict the partial states of the time points $t > 0$; i.e. $(\pi_0, t_0) \rightarrow^*_{Q;P} (\pi_i, t_i)$ with $i \in \{1, 2, 3\}$. Since the value of o_1 and of o_5 cannot be predicted at time $t = 2$, the result of action a_6 and of action a_8 cannot be predicted and π_3 contains only the value of o_3.

4. PLAN DIAGNOSIS

Suppose that we have a (partial) observation $obs(t) = (\pi, t)$ of the state of the world at time t and an observation $obs(t') = (\pi', t')$ at time $t' > t \geq 0$ during the execution of the plan P. We would like to use these observations to infer the health states of the actions occurring in P. Assuming a normal execution of P, we can (partially) predict the state of the world at a time point t' given the observation $obs(t)$: if all actions behave normally, we predict a partial state $\pi'_\varnothing$

[4] Note that in our context "undefined" is considered to be equivalent to "unpredictable".

at time t' such that $obs(t) \rightarrow^*_P (\pi'_\varnothing, t')$. If the normality assumption holds, the values of the objects that occur in both the predicted state and the observed state at time t' should match, i.e, we should have $\pi' =_{O(\pi') \cap O(\pi'_\varnothing)} \pi'_\varnothing$.[5] If not, the execution of some actions must have gone wrong. To determine which action(s) may have failed, we may apply the following straight-forward extension of the diagnosis concept in MBD to plan diagnosis [5]:

Definition 1. *Let $\langle P, obs(t), obs(t')\rangle$ with $P = (A, <)$ be a plan with observations $obs(t) = (\pi, t)$ and $obs(t') = (\pi', t')$, where $t < t' \leq depth(P)$. Let $obs(t) \rightarrow^*_{Q;P} (\pi'_Q, t')$ be a derivation assuming a qualification Q. Then Q is said to be a* plan diagnosis *of $\langle P, obs(t), obs(t')\rangle$ iff $\pi' =_{O(\pi') \cap O(\pi'_Q)} \pi'_Q$.*

So in a plan diagnosis Q the observed partial state (π') at time t' and the predicted state assuming the qualification Q (π'_Q) at time t' agree upon the values of all objects defined in both states.

Note that for all objects in $O(\pi') \cap O(\pi'_Q)$, the qualification Q provides an *explanation* for the observation π' made at time point t'. Hence, for these objects the qualification provides an *abductive diagnosis* [4]. For all observed objects in $O(\pi') - O(\pi'_Q)$, no value can be predicted given the qualification Q. Hence, by declaring them to be unpredictable, possible conflicts with respect to these objects if a normal execution of all actions is assumed, are resolved. This corresponds with the idea of a *consistency-based diagnosis* [13].

Minimum diagnosis and maximum informative diagnosis If Q is a plan diagnosis of $\langle P, obs(t), obs(t')\rangle$, then every superset $Q' \supseteq Q$ is also a plan diagnosis: since $\pi'_{Q'} \sqsubseteq \pi'_Q$, we have $O(\pi') \cap O(\pi'_{Q'}) \subseteq O(\pi') \cap O(\pi'_Q)$ and therefore $\pi' =_{O(\pi') \cap O(\pi'_Q)} \pi'_Q$ implies $\pi' =_{O(\pi') \cap O(\pi'_{Q'})} \pi'_{Q'}$. Clearly, then, the smaller a diagnosis, the more informative it will be, i.e., the more values it will predict that are also actually observed. Therefore, like in MBD, we will concentrate on *minimum* diagnoses. But there is a caveat: a minimum diagnosis i.e., a cardinality minimal diagnosis, is not necessarily a most informative diagnosis. Therefore, we also define the notion of a *maximum informative diagnosis*:

Definition 2. *Given plan observations $\langle P, (\pi, t), (\pi', t)\rangle$, a qualification Q is said to be a*

- minimum plan diagnosis *if for every plan diagnosis Q' it holds that $|Q| \leq |Q'|$ and*
- maximum informative plan-diagnosis *iff for all plan diagnoses Q^*, it holds that $|O(\pi') \cap O(\pi'_Q)| \geq |O(\pi') \cap O(\pi'_{Q^*})|$.*

Note that for every maximum informative diagnosis Q we have $O(\pi') \cap O(\pi'_Q) \subseteq O(\pi') \cap O(\pi'_\varnothing)$, where $obs(t) \rightarrow^*_{\varnothing;P} (\pi'_\varnothing, t')$ is the partial state derivation assuming a *normal plan* execution.

Example To illustrate the difference between minimum plan diagnosis en maximum informative diagnosis, consider again the plan execution depicted in Figure 2. Given $obs(0)$ and $obs(3)$ and a deviation in the value of o_2 at time $t = 3$, there are three possible minimum diagnoses: $D_1 = \{a_1\}$, $D_2 = \{a_3\}$ and $D_3 = \{a_6\}$. D_2 and D_3 are also maximum-informative diagnoses. ■

Note that in general a maximum informative diagnosis need not be a minimum diagnosis.

Minimal Causal Diagnosis Maximum informative diagnoses have to be preferred among the minimum diagnoses. Even maximum informed diagnoses, however, are not always the best ones we can obtain. The reason is that in a plan, instances of the same action may occur at several places thereby inducing *causal dependencies* between abnormalities. For example, suppose that we have a plan for carrying luggage from a depot to a number of waiting planes. Such a plan might contain several instances of a drive action pertaining to the same carrier. Suppose that an instance a_i of some drive action (type) a behaves abnormally because of malfunctioning of the carrier. Then it is reasonable to assume that other instances a_j of the same drive action that occur in the plan *after* a_i can be predicted to behave abnormally, too. This implies that instead of taking a qualification Q consisting of all instances of these actions, it would suffice to consider only the earliest occurrence of such an instance in Q as a *cause* of the malfunctioning of the remaining instances in Q.

In general, to capture such *causal relations* between instances of actions, we specify a set of *causal rules* Φ. Each such a rule is of the form $\alpha_1, \alpha_2, \ldots, \alpha_k \rightarrow \alpha_{k+1}$, where $\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}$ are action schemes, expressing that whenever a qualification Q contains action instances a_i occurring in $P_{\leq t}$ such that $type(a_i) = \alpha_i$ for $i = 1, \ldots, k$, then the action instances $a_{k+1} \in P_{>t}$ such that $type(a_{k+1}) = \alpha_{k+1}$, will be qualified as abnormal, too. The set Φ is said to specify a *causal theory* for the set of instances A. The set $inst_P(\Phi)$ is used to denote all instantiations of action schemes $(a_{i_1}, a_{i_2}, \ldots, a_{i_k}) \rightarrow a_{i_{k+1}} \in \Phi$ with respect to a plan P such that for some $t \geq 0$ there holds that $\{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\} \subseteq P_{\leq t}$ and $a_{i_{k+1}} \in P_{>t}$.

To define the set of causal consequences of qualifications, we consider the set of instances $Inst_P(\Phi)$ as a simple propositional Horn theory over the set A of instances of actions acting as atomic propositions. The set of causal consequences $C_\Phi(Q)$ of a qualification Q using Φ then equals the set of atomic propositional consequences of $Inst_P(\Phi) \cup Q$:

$$C_\Phi(Q) = Cn_A(Inst_P(\Phi) \cup Q).$$

Example Let $\mathcal{A}$ be the set of actions schemes of action instances in A where abnormal behavior is preserved, that is, if some instance of $\alpha \in \mathcal{A}$ is detected as behaving abnormally in a plan, then every future instance a of type α will also behave abnormally. The driving action we mentioned above is such a type of action. Now we can define a simple causal theory Φ as $\Phi = \{\alpha \rightarrow \alpha \mid \alpha \in \mathcal{A}\}$. As a result, whenever one instance of an action in $\mathcal{A}$ is qualified as abnormal, all subsequent instances of such an action will be qualified as abnormal, too. ■

It is easy to see that the operator C_Φ satisfies inclusion, monotony and idempotency. Using this operator, we can easily define a set of *causes* of a minimum plan diagnosis Q as a minimal set Q' that generates (a superset) of Q with the help of the causal rules in Φ:

[5] Since we do not assume to have full control over the plan observations we cannot assume $O(\pi') = O(\pi'_\varnothing)$, that is, observations of objects might only partially overlap.

Definition 3. *Let Q be a plan diagnosis of some plan P with causal theory Φ and let $obs(t)$ and $obs(t')$ be two observations with $t < t'$. Then a causal diagnosis (associated with Q) is a (subset) minimal set $Q_{min} \subseteq P_{[t,t']}$ such that $Q \subseteq [C_\Phi(Q_{min})]_{\leq t'}$, where $[C_\Phi(Q_{min})]_{\leq t'} = C_\Phi(Q_{min}) \cap P_{\leq t'}$.*

Such a minimal set of causes Q_{min} will always exist: take an arbitrary plan diagnosis Q. Since C_Φ satisfies inclusion, we have $Q \subseteq C_\Phi(Q)$ and since $Q \subseteq P_{\leq t'}$, it follows that $Q \subseteq [C_\Phi(Q)]_{\leq t'}$. If Q is a minimal solution satisfying this last inclusion relation, it is a causal diagnosis, else, there exists a subset Q_{min} of Q satisfying the inclusion.

In general, a set of causes of Q will be smaller than Q itself. On the other hand, since $Q \subseteq [C_\Phi(Q_{min})]_{\leq t'}$, this latter set, generated by the causes Q_{min}, is also a plan-diagnosis, but might be less informative than Q. So in trying to minimize the set of causes, we could loose explanatory power. We will keep a balance between minimality and explanatory power by defining *Pareto minimal causal* diagnoses as follows:

Definition 4. *Let $P = (A, <)$ be a plan, Φ a causal theory and let $obs(t) = (\pi, t)$ and $obs(t') = (\pi', t')$ with $t < t'$ be two observations. Let $obs(t) \rightarrow^*_{\varnothing;P} (\pi'_\varnothing, t')$ and let $obs(t) \rightarrow^*_{Q;P} (\pi'_Q, t')$ be the plan execution assuming a set Q of abnormal actions. Then Q is said to be a* Pareto minimum causal diagnosis *of $\langle P = (A, <), \Phi, obs(t), obs(t') \rangle$ iff there is no causal diagnosis Q' that dominates Q. Here, a diagnosis Q' is said to dominate Q iff* (i) *$|Q'| < |Q|$ and $[C_\Phi(Q')]_{\leq t'}$ is at least as informative as $[C_\Phi(Q)]_{\leq t'}$ or* (ii) *$|Q'| = |Q|$ and $[C_\Phi(Q')]_{\leq t'}$ is a more informative diagnosis than $[C_\Phi(Q)]_{\leq t'}$.*

Prediction of plan results Except for playing a role in establishing causal *explanations* of observations, (causal) diagnoses also can play a significant role in the *prediction* of future results (states) of the plan or even the attainability of the goals of the plan. First of all, we should realize that a diagnosis can be used to enhance observed state information as follows: Suppose that Q is a causal diagnosis of a plan P based on the observations $obs(t)$ and $obs(t')$ for some $t < t'$, let $obs(t) \rightarrow^*_{C_\Phi(Q);P} (\pi'_Q, t')$ and let $obs(t') = (\pi', t')$. Since $C_\Phi(Q)$ is a diagnosis, π' and π'_Q agree upon the values of all objects occurring in both states. Therefore we can combine the information contained in both partial states by merging them into a new partial state $\pi'_Q \sqcup \pi'$. Here, the merge $\pi^1 \sqcup \pi^2$ of two partial states π^1 and π^2 is simply defined as the partial state π where $\pi(j) = \pi^i(j)$ iff $\pi^i(j)$ is defined for $i = 1, 2$ and undefined otherwise. The partial state $\pi'_Q \sqcup \pi'$ can be seen as the partial state that can be obtained by direct observation at time t and indirectly by making use of previous observations and plan information.

In the same way, we can use this information and the causal consequences $C_\Phi(Q)$ to derive a prediction of the partial states derivable at times $t'' > t'$:

Definition 5. *Let Q is a causal diagnosis of a plan P based on the observations (π, t) and (π', t') where $t < t'$. Moreover, let $obs(t) \rightarrow^*_{C_\Phi(Q);P} (\pi'_Q, t')$ and let $obs(t') = (\pi', t')$.*

Then, for some time $t'' > t'$, (π'', t'') is the partial state predicted using Q and the observations if

$$(\pi'_Q \sqcup \pi', t') \rightarrow^*_{C_\Phi(Q);P} (\pi'', t'').$$

In particular, if $t'' = depth(P)$, i.e., the plan has been executed completely, we can predict the values of some objects that will result from executing P and we can check which goals $g \in G$ will still be achieved by the execution of the plan, based on our current knowledge. That is, we can check for which goals $g \in G$ it holds that $\tau \models g$. So causal diagnosis might also help in evaluating which goals will be affected by failing actions.

5. MULTI-AGENT DIAGNOSIS OF MULTI-AGENT PLANS

In the previous sections we discussed diagnosis of a single agent plan. In a multi-agent setting a group of agents is responsible for executing a common plan $P = (A, <)$ and each agent is responsible for a sub-plan P_i of P. Such a sub-plan P_i first of all contains a set A_i of actions only agent i is responsible for. We consider these sets A_i to constitute a partitioning of the set A. Furthermore, for each agent i we distinguish two sets of actions I_i and R_i belonging to other agents with which agent i has to synchronize the execution of actions.[6] The first set I_i is the set of actions that provide the input for the actions in A_i: $I_i = \{a \in A \mid a \ll a', a \notin A_i, a' \in A_i\}$. The set R_i is the set of actions that receive their input from actions performed by agent i: $R_i = \{a \in A \mid a' \ll a, a \notin A_i, a' \in A_i\}$.

If an action $a \in A_i$ of an agent is followed by an action $a' \in A_j$ of another agent ($a \ll a'$), we assume that the control over the objects $ran_O(a) \cap dom_O(a')$ is transferred from agent i to agent j. Note that this may result in transferring control over an object o to more than one agent if $a \ll a'$, $a \ll a''$, $o \in dom_O(a') - ran_O(a')$ and $o \in dom_O(a'') - ran_O(a'')$. However, the *concurrency requirement* guarantees that the agent executing an action that modifies the state of an object, will be the only agent having control over the object.

Each sub-plan P_i is generated by the set of actions A_i assigned to agent i together with the sets I_i and R_i and is simply defined as the plan $P_i = (A_i, I_i, R_i, <_i)$, where $<_i$ is the relation $<$ restricted to $((A_i \cup I_i) \times A_i)) \cup (A_i \times (A_i \cup R_i))$. Synchronization between two agents i and j is restricted to the outputs of actions $a \in R_i$ needed by actions $a' \in I_j$ with $a \ll a'$.

Figure 3 gives an illustration of a plan that is distributed over three agents. Initially, agent 1, 2 and 3 have control over respectively the objects o_1, o_2 and o_3. At $t = 2$ agent 2 transfers control over o_2 to both agent 1 and 3 and at $t = 3$ agent 1 transfers the control over o_2 completely to agent 3.

To discuss multi-agent diagnosis, we first have to concentrate on the derivation relation for partial plans. To derive the partial state $(\pi', t+1)$ of a sub-plan $P_i = (A_i, I_i, R_i, <_i)$ given a partial state (π, t) at time t, we must take into account the objects O_i^t under the control of the agent i at time t. The objects under the control of agent i are defined as: $O_i^t = \{o \in dom_O(a) \mid a \in A_i \cap P_{\geq t}, \forall a' \in P_{\geq t} : a' < a \Rightarrow o \notin ran_O(a')\} \cup \{o \in ran_O(a) \mid a \in A_i \cap P_{<t}, \forall a' : a \ll a' \Rightarrow o \notin dom_O(a')\}$.

Let π_i^a be the partial state of the objects $O(\pi_i^a) \subseteq ran_O(a) \cap O_i^t$ under the control of agent i at time t that are determined by the action $a \in I_i$ such that $a \ll a'$ for some $a' \in P_{i,t} = A_i \cap P_t$. We say that (π', t') is directly generated

[6] Here, we do not take into consideration how agents synchronize their actions; i.e., whether they synchronize their actions through communication or by observing a common environment.

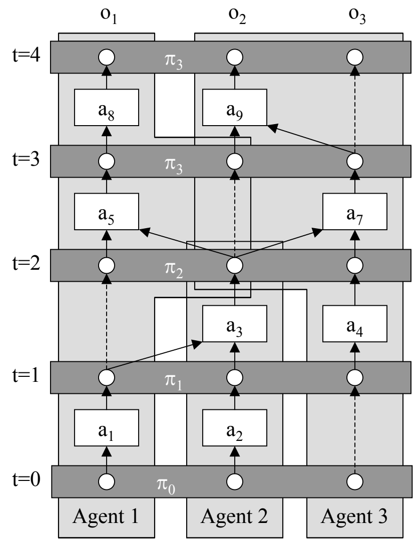

Figure 3: Multi-agent plan diagnosis

from (π, t) by the execution of sub-plan P_i using qualification Q_i, denoted by $(\pi, t); \{\pi_i^a \mid a \in I_i\} \rightarrow_{Q_i;P_i} (\pi', t+1)$, if there exists an information-minimal (w.r.t. $\sqsubseteq$) partial state π^* such that

- $(\pi \upharpoonright O_i^t) \sqsubseteq \pi^*$;
- for each $a \in I_i$ with $a \ll a'$ and $a' \in P_{i,t}$: $\pi_i^a \sqsubseteq \pi^*$;
- $(\pi^*, t) \rightarrow_{Q_i;P_i} (\pi', t+1)$.

The diagnosis of a sub-plan An agent i executing its sub-plan P_i may make partial observations at times t and t' with $t < t'$. This agent can predict the expected state of the world at time t' using the knowledge of its plan and information received from other agents about the expected effects of actions in I_i. If agent i notices a difference between the expected and the observed state of the world at time t', diagnosis is required.

Definition 6. *Let $P_i = (A_i, I_i, R_i, <_i)$ be the sub-plan of agent i, let $obs(t)$ and $obs(t')$ with $t < t'$ be two observations and let $(\pi, t); \{\pi^a \mid a \in I_i\} \rightarrow^*_{Q_i;P_i} (\pi'_{Q_i}, t')$ using qualification $Q_i \subseteq A_i$.*

Then Q_i is a sub-plan-diagnosis of $\langle P_i, obs(t), \{\pi^a \mid a \in I_i\}, obs(t')\rangle$ iff $\pi' =_{O(\pi') \cap O(\pi'_{Q_i})} \pi'_{Q_i}$.

Multi-agent plan-diagnosis Suppose that a qualification Q is a plan-diagnosis of the global plan $\langle P, obs(t), obs(t')\rangle$. We would like to know whether a group of agents executing this plan in a distributed way is able to establish the same diagnosis in a distributed way. In other words, if Q is a global diagnosis, is the qualification $Q_i = Q \cap A_i$ a sub-plan diagnosis of P_i?

The inputs of the sub-plan P_i are, of course, the observation (π, t) and for each action $a \in I_i$, the partial state π_i^a with $(\pi, t) \rightarrow^*_{Q;P} (\pi^*, depth_P(a) + 1)$. If $a \in A_j$, then, of course, $(\pi, t); \{\pi_j^a \mid a \in I_j\} \rightarrow^*_{Q_j;P_j} (\pi_i^a, depth_{P_j}(a) + 1)$.

Proposition 1. *Let Q be a plan diagnosis of $\langle P, obs(t), obs(t')\rangle$. Moreover, for each action $a \in I_i$, let π_i^a be the partial state describing the input provided by a.*

Then, $Q_i = Q \cap A_i$ is a plan-diagnosis of $\langle P_i, obs(t), \{\pi_i^a \mid a \in I_i\}, obs(t')\rangle$ and for each $a' \in I_j \cap A_i$, $(\pi, t); \{\pi_i^a \mid a \in I_i\} \rightarrow^*_{Q_i;P_i} (\pi', depth_P(a') + 1)$ and $\pi_j^{a'} = \pi' \upharpoonright dom_O(a')$.

The following proposition shows that a diagnosis of the whole plan can be obtained by combining plan-diagnoses of sub-plans.

Proposition 2. *Let the qualification Q_i be a plan-diagnosis of $\langle P_i, obs(t), \{\pi^a \mid a \in I_i\}, obs(t')\rangle$.*

Then the qualification $Q = \bigcup_i Q_i$ is a plan-diagnosis of $\langle P, obs(t), obs(t')\rangle$ iff for each agent i and for each $a' \in I_j \cap A_i$, $(\pi, t); \{\pi_i^a \mid a \in I_i\} \rightarrow^*_{Q_i;P_i} (\pi', depth_P(a') + 1)$ and $\pi_j^{a'} = \pi \upharpoonright dom_O(a')$.

Note that if an agent i qualifies an action $a \in A_i$ as being abnormal, i.e. $a \in Q_i$, then the information communicated to agent j for an action $a' \in I_j \cap A_i$ may contain unknown values for some of the objects in $dom_O(a')$. Also note that neither the combination of local minimum diagnoses needs to result in a global minimum diagnosis nor the combination of maximum-informative diagnoses needs to result in a global maximum-informative diagnosis.

Determining plan-diagnoses In a multi-agent system, the challenge is to determine a global diagnosis in a distributed way with agents using local knowledge only. As was already noted above, multi-agent plan-diagnosis can be seen as multi-agent diagnosis of a system with a spatial knowledge distribution over the agents [8]. This form of diagnosis raises a number of issues if agents try to establish their local diagnoses by exchanging information about predicted values of objects at specific time points [15, 16]. Roos et al. [16] therefore propose an indirect approach based on first determining dependency sets [13]. They use a focusing approach [14] that enables each agents to determine a set containing the likely broken components.

In plan diagnosis, we can use a similar approach. In plan diagnosis the use of dependency sets is even easier since circular dependencies between actions cannot occur.

Definition 7. *Let $P = (A, <)$ be a plan and let $obs(t) = (\pi, t)$ and $obs(t') = (\pi', t')$ with $t < t'$ be two (partial) observations. The dependency set of an object at time t^* is defined as the set*

$$Dep(o, t^*) = \{a' \in P_{[t,t^*]} \mid a \in P_{[t,t^*]}, o \in ran_O(a), a' \prec^* a\}$$

where $\prec^*$ is the reflexive and transitive closure of

$$\prec = \{(a, a') \mid a \ll a', ran_O(a) \cap dom_O(a') \neq \varnothing\}.$$

These dependency sets can be determined in a distributed way; see [16]. We will use $Dep_i(o, t^*)$ to denote the local part of the dependency set under the control of agent i.

If the observed value of an object $o' \in O(\pi')$ with $obs(t') = (\pi', t')$ does not correspond to the predicted value of o' assuming a normal plan execution, we say the dependency set $Dep(o', t')$ is a *conflict set.* Otherwise, $Dep(o', t')$ is a *confirmation set.* Clearly, there must have occurred an anomaly in at least one of the actions in a conflict set.

By using the conflict and confirmation sets, agents can either apply the focusing approach described in [16] or determine the local minimum diagnoses by determining the hitting sets the conflict sets [13]. Note that in some plans the combination of minimum diagnoses need not describe all minimum diagnoses of the whole system. In a system of physical components this is usually not a problem since additional measurements needed to eliminate diagnoses, will reveal the other minimum diagnoses. In plan-diagnosis we cannot make additional observations at past time points. Hence, agents have to exchange additional information to be able determine all minimum diagnoses.

The confirmation sets enable the agents to determine all maximum informative diagnoses. The following proposition shows that by removing from each conflict set those actions that occur in at least one confirmation set, the maximum informative diagnoses can be determined.

PROPOSITION 3. *Let $\langle P, obs(t), obs(t')\rangle$ be a plan $P = (A, <)$ with observations $obs(t) = (\pi, t)$ and $obs(t') = (\pi', t')$, where $t < t' \leq depth(P)$. Let $T_1, ..., T_k$ be the conflict sets and let $N_1, ..., N_l$ be the confirmation sets determined using the observations. A qualification Q is a maximum informative* plan diagnosis *of $\langle P, obs(t), obs(t')\rangle$ iff Q is a minimum hitting set of $T_1 - N^*, ..., T_k - N^*$ with $N^* = \bigcup_{i=1}^{l} N_i$.*

Example To illustrate the protocol, consider the plan in Figure 3. Observations are made at $t = 0$ and $t' = 4$. At $t' = 4$ agents 1 and 3 observe expected values for o_1 and o_3 respectively while agent 3 observes an anomalous value for o_2. This implies that there is one global conflict set: $T = \{a_1, a_2, a_3, a_4, a_7, a_9\}$ for o_2, and two global confirmation sets: $N_1 = \{a_1, a_2, a_3, a_5, a_8\}$ for o_1 and $N_2 = \{a_1, a_2, a_3, a_4, a_7\}$ for o_3. By passing on identifications for the observed objects together with the status of the observations, the agents can determine their local conflict and confirmation sets in a distributed way [16]. So, agent 1 has one local conflict set: $T^1 = \{a_1\}$ and two local confirmation sets: $N_1^1 = \{a_1, a_5, a_8\}$ and $N_2^1 = \{a_1\}$, agent 2 has one local conflict set: $T^2 = \{a_2, a_3\}$ and two local confirmation sets: $N_1^2 = \{a_2, a_3\}$ and $N_2^2 = \{a_2, a_3\}$, and agent 3 has one local conflict set: $T^3 = \{a_4, a_7, a_9\}$ and one local confirmation set: $N_2^3 = \{a_4, a_7\}$. Given this information, the agents can determine their local minimum diagnoses, and their local maximum informative diagnoses. In this example agent 1 has the minimum local diagnosis $Q^1 = \{a_1\}$, agent 2 has the minimum local diagnoses $Q_1^2 = \{a_2\}$ and $Q_2^2 = \{a_2\}$, and agent 3 has the minimum local diagnoses $Q_1^3 = \{a_4\}$, $Q_2^3 = \{a_7\}$ and $Q_3^3 = \{a_9\}$. Agent 3 has the only maximum informative diagnosis, namely Q_3^3. ■

Determining the causal diagnoses is a straight forward problem once the diagnoses have been determined. However, if the actions in the antecedent of an instantiated rule belong to multiple agents, coordination between the agents is required. This a topic for further research.

6. CONCLUSION

We have presented a new object-oriented model for describing multi-agent plans. This model enables agents to apply techniques developed for multi-agent diagnosis to identify (*i*) minimum sets of anomalously executed actions and (*ii*) maximum informative (w.r.t. to predicting the observations) sets of anomalously executed actions. Due to the occurrence of several instances of the same action in a plan, anomalously executed actions might be correlated. Therefore, (*iii*) causal diagnoses have been introduced and we have extended the diagnostic theory enabling the prediction of future failure of actions. Finally, we have extended the plan-diagnosis to the multi-agent case. Issues for further research are handling dynamic changes that influence the applicability of causal rules, extending the diagnosis to the executing agents, and developing efficient protocols for distributed minimum, maximum informative and causal diagnosis.

7. REFERENCES

[1] P. Baroni, G. Lamperti, P. Poglianob, and M. Zanella. Diagnosis of large active systems. *Artificial Intelligence*, (110):135–183, 1999.

[2] L. Birnbaum, G. Collins, M. Freed, and B. Krulwich. Model-based diagnosis of planning failures. In *AAAI 90*, pages 318–323, 1990.

[3] N. Carver and V. Lesser. Domain monotonicity and the performance of local solutions strategies for CDPS-based distributed sensor interpretation and distributed diagnosis. *Autonomous Agents and Multi-Agent Systems*, 6(1):35–76, 2003.

[4] L. Console and P. Torasso. Hypothetical reasoning in causal models. *International Journal of Intelligence Systems*, 5:83–124, 1990.

[5] L. Console and P. Torasso. A spectrum of logical definitions of model-based diagnosis. *Computational Intelligence*, 7:133–141, 1991.

[6] F. de Jonge and N. Roos. Plan-execution health repair in a multi-agent system. In *PlanSIG 2004*, 2004.

[7] R. E. Fikes and N. Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 5:189–208, 1971.

[8] P. Frohlich, I. de Almeida Mora, W. Nejdl, and M. Schroeder. Diagnostic agents for distributed systems. In J.-J. C. Meyer and P.-Y. Schobbens, editors, *Formal Models of Agents. ESPRIT Project ModelAge Final Report Selected Papers. LNAI 1760*, pages 173–186. Springer-Verlag, 2000.

[9] B. Horling, B. Benyo, and V. Lesser. Using Self-Diagnosis to Adapt Organizational Structures. In *Proceedings of the 5th International Conference on Autonomous Agents*, pages 529–536. ACM Press, 2001.

[10] M. Kalech and G. A. Kaminka. On the design of social diagnosis algorithms for multi-agent teams. In *IJCAI-03*, pages 370–375, 2003.

[11] M. Kalech and G. A. Kaminka. Diagnosing a team of agents: Scaling-up. In *AAMAS 2004*, 2004.

[12] D. Nau, M. Ghallab, and P. Traverso. *Automated Planning: Theory & Practice*. Morgan Kaufmann Publishers Inc., 2004.

[13] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987.

[14] N. Roos. Efficient model-based diagnosis. *Intelligent System Engineering*, pages 107–118, 1993.

[15] N. Roos, A. ten Teije, A. Bos, and C. Witteveen. An analysis of multi-agent diagnosis. In *AAMAS 2002*, pages 986–987, 2002.

[16] N. Roos, A. ten Teije, and C. Witteveen. A protocol for multi-agent diagnosis with spatially distributed knowledge. In *AAMAS 2003*, pages 655–661, 2003.

[17] H. Tonino, A. Bos, M. de Weerdt, and C. Witteveen. Plan coordination by revision in collective agent based systems. *Artificial Intelligence*, 142:121–145, 2002.