Reformulation for the Diagnosis of Discrete-Event Systems

Alban Grastien
NICTA* and the Australian National University, Canberra, Australia

Gianluca Torta
Università di Torino, Torino, Italy

Abstract

Diagnosis is traditionally defined on a space of hypotheses (typically, all the combinations of zero or more possible faults). In the present paper, we argue that a suitable reformulation of this hypothesis space can lead to more efficient computation of diagnoses, most notably by exploiting opportunities for various forms of model abstraction. The paper focuses on the diagnosis of Discrete Event Systems (DES), although the main ideas apply to diagnosis in general.

An important contribution of the paper is the study of several formal properties related to the correctness and precision of the diagnoses obtained through reformulation.

1 Introduction

Diagnosis is the problem of detecting abnormal behaviour of a system and, after detection, to determine the location and/or the type of system faults that caused the abnormal behaviour. A *diagnosis hypothesis* indicates which fault(s) occurred in the system, and the *diagnosis* is the set of alternative hypotheses that explain (i.e., are compatible) with the observed system behaviour. In this paper, we focus on Model-Based Diagnosis (MBD) of Discrete-Event Systems (DESs, see (Cassandras and Lafortune 1999)), where the diagnosis is computed by comparing a complete DES model of the system behaviour with a (partial) observation of the actual system behaviour (Sampath et al. 1995).

Since the size of the search space for diagnosis is usually exponential in the number of different faults, many recent works in diagnosis of DESs have tried to tackle this complexity issue, e.g. (Benveniste et al. 2005; Pencolé, Kamenetsky, and Schumann 2006). A possible approach already explored in MBD of static system models (e.g., (Sachenbacher and Struss 2005; Torta and Torasso 2008)) is to abstract the model in order to simplify the diagnosis process.

*NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

Model abstraction works well when there is a mismatch between the level of detail of the system model and the level of detail of the *hypothesis space*, i.e., the set of diagnosis hypotheses. For example, think of a model of a car which describes exactly how each component works, while the only diagnostic hypotheses we need to consider are whether the engine starts or not; in such a case, it is easy to imagine that the model can be significantly simplified (i.e., abstracted) without any loss in the possibility of discriminating among the two hypotheses.

In many real cases however, the hypothesis space is defined in such a way that only little abstraction can be applied to the model without incurring severe loss of precision. This problem stems from the fact that, usually, the diagnosis hypotheses are expressed in terms of detailed statements about the global system status: for each possible fault in the whole system, a diagnosis hypothesis needs to specify whether such a fault occurred or not. Moreover, all of the faults that occurred within the (possibly extended) time interval during which the system has been observed must be accounted for in the diagnosis. Considering again the diagnosis of a car, for each component we could be interested in knowing whether a fault has occurred to it during the last week; in such a case, it is difficult to perform a drastic abstraction of the model without losing any precision in the discrimination among different hypotheses.

In this article, we study a novel approach to reduce the complexity of DES diagnosis, based on a reformulation of the hypothesis space. Our approach consists in the following main steps:

1. the hypothesis space is formulated differently, i.e., we define a new hypothesis space,
2. the diagnosis is computed for this new hypothesis space,
3. (optionally) the diagnosis is *mapped back* to the original formulation of the hypothesis space.

In other words, we propose to reformulate the *language* that will be used to answer each diagnostic problem; since such an answer (i.e., the diagnosis) is expressed as a set of alternative hypotheses belonging to the hypothesis space, we need to reformulate the hypothesis

space itself. In the example above, we may transform the detailed hypothesis space which assigns a precise fault mode to each component of the car, to the simple hypothesis space that just contains the two hypotheses: the car starts, the car doesn't start.

The main benefit of this process is that a suitably defined new hypothesis space may allow powerful model abstractions, as pointed out above. It is important to note that, in our proposal, model abstraction is performed after the problem simplification introduced by the reformulation of the hypothesis space; this represents a somewhat reversed view w.r.t. most previous works on abstraction (with the notable exception of (Sachenbacher and Struss 2005)), which start with the abstraction of the system model and consider the change of the hypothesis space for diagnosis only as an effect of the model abstraction.

In this paper, we will focus on the first and last steps, i.e., on the operations related to the mapping from one hypothesis space to another one.

In general, when the reformulated diagnosis is mapped back to the original hypothesis space we do not obtain the same result as if we directly computed the diagnosis using the original model; in other words, the reformulation of the hypothesis space may in general lead to loss of diagnostic precision w.r.t. to the original hypothesis space. For example, a typical implementation of the scheme above is to diagnose every possible system failure separately instead of trying to solve the problem globally; in this way, the original diagnostic problem is mapped to a linear number of simpler diagnosis problems (Pencolé, Kamenetsky, and Schumann 2006) and, following our approach, a specific model abstraction can be applied to each of them. However, this process may result in the loss of dependencies among faults, e.g., we may end up knowing that each one of the faults f_1, f_2 possibly occurred, without knowing whether their occurrences are mutually exclusive.

One of the main contributions of this article is to study some properties of the system model and/or the applied reformulations which guarantee that an algorithm based on the reformulated hypothesis space leads to the same diagnosis as a classic MBD algorithm applied to the original hypothesis space.

However, we believe that separating the reformulation and abstraction processes is beneficial even when the reformulation causes loss of precision. Indeed, in many cases it is better to have a diagnosis expressed in a coarser language (such as in the car example) than being unable to compute diagnoses because of the prohibitive computational costs. Moreover, if the diagnosis is computed in several steps (as in the hierarchical diagnosis procedure), it is perfectly acceptable to get an imprecise intermediate result, which is then used to focus more precise reasoning in subsequent steps.

In all of these cases, thanks to reformulation the user has a means to explicitly control the transformation of the hypothesis space, and thus to limit the loss of precision due to the transformation(s) of the problem.

After introducing the basic concepts on which our work is based (section 2), we precisely define reformulation (section 3) and study some of its properties (section 4). Then, we analyze some relevant examples of the possible applications of reformulation (section 5) and conclude the paper with a discussion. The proofs of the theorems are omitted here due to lack of space, but may be found in (Grastien and Torta 2010).

2 Preliminaries

In this section, we review the classical framework of the MBD of DESs, slightly rephrasing it to better fit the concept of reformulation introduced in the next section.

2.1 Diagnosis

We consider a system denoted as $\mathcal{S}$. Because of misconception, misuse or inavoidable failures, the system may exhibit a number of *faults* denoted by the set E_f. The system is monitored by *sensors* which produce an observation θ. *Diagnosis* is the problem of using the observation θ to determine whether the system $\mathcal{S}$ exhibited faults, and in this case to identify which fault(s) did occur.

More formally, we call *diagnosis hypothesis* (or simply *hypothesis*), denoted as $h : E_f \to \mathbf{B}$, a function that associates a Boolean with each fault. The semantics of hypothesis h is that fault $f \in E_f$ occurred iff $h(f) = \top$. The *space of diagnosis hypotheses* H is defined as the set $\mathbf{B}^{E_f} = \{(h : E_f \to \mathbf{B})\}$ of all the possible functions from E_f to $\mathbf{B}$. The *diagnosis problem* is then defined as the tuple $\langle \mathcal{S}, \theta, H \rangle$. A *diagnosis* Δ is formally defined as a subset of hypotheses: $\Delta \subseteq H$. Note that the diagnosis is defined with respect to a space of diagnosis hypotheses H.

The definition of $H = \mathbf{B}^{E_f}$ provided above focuses on the *set* of faults that occurred in the system, and is widely adopted in the literature on DES diagnosis. While such a definition will provide a basis for deriving specific results on reformulation, it is worth pointing out that most discussions made in the paper would be unaffected by the adoption of alternative definitions of DES diagnosis hypotheses found in the literature (most notably the one whereby a diagnosis hypothesis is a *sequence* of faults, i.e. $H = E_f\star$ where $\star$ is the usual Kleene closure).

2.2 Model-Based Diagnosis of DES

Let E be a set of labels. A *language* $\mathcal{L}$ on set E is a set of *words* $\sigma \in \mathcal{L}$ defined as sequences of labels: $\mathcal{L} \subseteq E\star$.

We consider that the system can be accurately modeled by a finite DES. In practice, the behaviour of the system is represented by a model M (automaton, Petri net, etc.) that defines a language $\mathcal{L}_M$ on the set of system *events* $E = E_u \cup E_o \cup E_f$, where E_u is the set of *unobservable events*, E_o the set of *observable events*, and E_f the set of faults. A specific behaviour of the system is represented by a word $\sigma \in \mathcal{L}_\mathcal{M}$, and generates an observation $obs(\sigma)$ defined as the projection $Proj_{E_o}(\sigma)$

of σ on the set of observable events; unobservable events and faults are not observed.

The *semantics* of hypothesis $h \in H$ is defined as the set of behaviours $sem(h) \subseteq E\star$ that "agree" with hypothesis h; if $\sigma \in sem(h)$, we say that σ *belongs* to h. In hypothesis space $H = \mathbf{B}^{E_f}$, the definition:

$$sem(h) = \{\sigma \in E\star \mid \forall f \in E_f,\ f \in \sigma \leftrightarrow h(f) = \top\}$$

captures the intended meaning of each hypothesis h.

Given a model and an observation, the possible behaviours of the system are the behaviours σ of the model that generate the observation. However, we are not so much interested in these behaviours, as to the hypotheses they belong to.

Definition 1 *A* model-based diagnosis problem *(or MBD problem) is a tuple $P = \langle M, \theta, H\rangle$ where M is a DES model, $\theta \in E_o\star$ is an observation, and H is a space of diagnosis hypotheses.*

The model-based diagnosis *(or MBD) Δ_P of problem $P = \langle M, \theta, H\rangle$ is defined by:*

$$\Delta_P = \{h \in H \mid \exists \sigma \in \mathcal{L}_M : \sigma \in sem(h) \wedge obs(\sigma) = \theta\}.$$

In the context of this paper, an MBD problem is a particular case of a diagnosis problem where the system is modeled by a DES. The meaning of MBD $\Delta_P = \{h_1, \ldots, h_k\}$ is that each one of the hypotheses $h_1, \ldots, h_k$ is *possible* according to observation θ and model M.

An important point about Δ_P is the following. Consider a number of sets $H_1, \ldots, H_k$ which cover H (i.e., $H = \bigcup_{i\in\{1,\ldots,k\}} H_i$), and compute the diagnosis $\Delta_{P,i}$ for each problem $P_i = \langle M, \theta, H_i\rangle$; then, it is easy to see that $\Delta_P = \bigcup_{i\in\{1,\ldots,k\}} \Delta_{P,i}$. In other words, Δ_P can be computed by considering each subset of hypotheses H_i separately, and then unioning the results. This makes it possible to apply specific model abstractions for each sub-problem P_i; we will come back to the relevance of this possibility when we describe some applications of reformulation in section 5.

2.3 Quality of Diagnosis

Let $\overline{\sigma} \in \mathcal{L}_M$ be the representation in our model of the actual (real) behaviour of the system. The *perfect diagnosis* $\overline{\Delta}$ is defined as the set of diagnosis hypotheses matched by this behaviour:

$$\overline{\Delta} = \{h \in H \mid \overline{\sigma} \in sem(h)\}.$$

Clearly, if the hypotheses in H are mutually-exclusive, the perfect diagnosis $\overline{\Delta}$ contains at most one element; plus, if the set of hypotheses *covers* the set of behaviours (for all $\sigma \in \mathcal{L}_M, \exists h \in H : \sigma \in sem(h)$), the perfect diagnosis contains at least one element; note, however, that our definition of diagnosis hypothesis is general enough for $\overline{\Delta}$ to contain zero, one or several elements.

Ideally, the diagnosis procedure should return the perfect diagnosis. In practice, this may be impossible because the observability of the system is partial and the sensors do not provide precise enough an observation to diagnose perfectly. Moreover, the model itself may be imprecise.

Diagnoses can be evaluated and compared thanks to two criteria: *d-correctness* defines the property that hypotheses $h \in \overline{\Delta}$ are indeed included in the diagnosis; *d-precision* defines the property that hypotheses $h \notin \overline{\Delta}$ are indeed excluded from the diagnosis. In this paper we take the view that bad d-correctness (or low coverage) is more serious than bad d-precision (or high false coverage) (Krysander and Nyberg 2008), and therefore focus our interest on d-correct diagnoses.

Theorem 1 *Given a problem $P = \langle M, \theta, H\rangle$, the MBD Δ_P is the most d-precise diagnosis which is certainly d-correct given the available model M and observation θ.*

Provided that, among the d-correct diagnoses, Δ_P is the most d-precise diagnosis which can be computed given an MBD problem P, we will say that a diagnosis Δ is d-correct (resp. d-precise) w.r.t. P if $\Delta \supseteq \Delta_P$ (resp. $\Delta \subseteq \Delta_P$).

3 Reformulation

The framework developed in the previous section defines a diagnosis as a set of hypotheses, each of which is possible according to the model and the observations. In particular, each hypothesis h in the hypothesis space $\mathbf{B}^{E_f}$ refers to the (non) occurrence of each faulty event in E_f over the entire period of observation. Therefore, knowing whether h is possible or not requires to reason globally over the whole system and for the whole time period during which observations have been collected.

A powerful technique for alleviating such a complexity is model abstraction, namely the simplification of the model by forgetting *irrelevant* details. However, it is usually difficult to apply such a technique to the MBD reasoning task, because abstracting the model often has undesired effects on the computed diagnosis; in particular, spurious hypotheses may easily appear because the abstraction forgot some relevant details of the model.

In this paper we want to argue that it is beneficial to reformulate the hypothesis space before abstraction. To this end, we introduce the notion of reformulation of the hypothesis space H to a new space H', which is expected to allow more efficient model abstraction.

This idea is depicted in Fig. 1:

1. the diagnosis problem $P = \langle M, \theta, H\rangle$ is reformulated to a problem $P^\rho = \langle M, \theta, H'\rangle$ with the same model M and a new hypothesis space H';
2. the hypothesis space H' of problem P^ρ may allow for the abstraction of model M to a model M', yielding a problem $P^\rho_A = \langle M', \theta', H'\rangle$ with the same diagnosis Δ_{P^ρ} as P^ρ;
3. the diagnosis Δ_{P^ρ} is computed in space H' (i.e. $\Delta_{P^\rho} \subseteq H'$) by solving problem P^ρ_A;
4. the diagnosis Δ_{P^ρ} is mapped back to a diagnosis Δ^ρ_P in the original space H (i.e. $\Delta^\rho_P \subseteq H$).

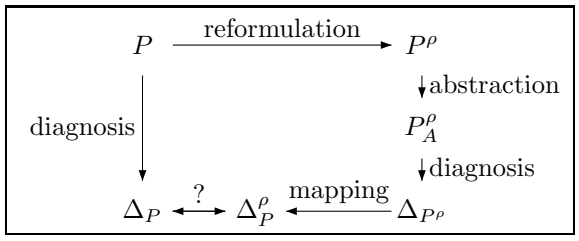

Figure 1: Principle of diagnosis through reformulation

Note that, in the second step, we assume that the model abstraction from M to M' fully preserves the solution of problem P^ρ in hypothesis space H'. A discussion on model abstractions that preserve the correctness and precision of diagnosis can be found in (Grastien and Torta 2011).

Despite this assumption about the abstraction step, an important question we shall discuss in the next section is whether Δ^ρ_P matches the original diagnosis Δ_P (represented by the question mark in the figure), i.e. whether the reformulation itself preserves the correctness and precision of diagnosis.

We are now ready to formally define reformulations.

Definition 2 *Given a hypothesis space H, a* reformulation *is a pair $\rho = \langle g, H' \rangle$, where H' is a set of hypotheses and g is a function that associates with each hypothesis $h \in H$ a set $\{\Delta'_1, \ldots, \Delta'_l\}$ of sets of hypotheses $\Delta'_i = \{h'_{i1}, \ldots, h'_{ik_i}\}$ with $h'_{ij} \in H'$.*

The intended meaning of the reformulation $g(h) = \{\Delta'_1, \ldots, \Delta'_l\}$ of a hypothesis h is that whenever h is possible (i.e., it belongs to diagnosis Δ_P), then all of the hypotheses in Δ'_i are possible, for at least one $i \in \{1, \ldots, l\}$. In section 4 we will formally define the semantics of $g(h)$.

Note that we give two degrees of freedom in the definition of a reformulation $\rho = \langle g, H' \rangle$: first of all, it is possible to choose the target set of hypotheses H' of the reformulation and their semantics, i.e., for each $h' \in H'$ the set of behaviours $sem(h') \subseteq E\star$ that agree with h'. Moreover, it is possible to choose the way hypotheses in H are mapped to hypotheses in H'. As we shall see below, our choices may be constrained if we want our reformulation to be correct and precise; however, this still gives us a lot of freedom in defining reformulations. Such a freedom can be exploited in order to choose a reformulation that makes model abstraction easier.

Definition 3 *Given a reformulation $\rho = \langle g, H' \rangle$ and an MBD problem $P = \langle M, \theta, H \rangle$ we define:*

1. *Reformulated problem: the MBD problem $P^\rho = \langle M, \theta, H' \rangle$.*
2. *Reformulated diagnosis: the MBD Δ_{P^ρ} computed starting from problem P^ρ (or, equivalently, from its abstraction P^ρ_A)*
3. *Diagnosis through reformulation: the diagnosis $\Delta^\rho_P = g^{-1}(\Delta_{P^\rho})$ obtained by mapping Δ_{P^ρ} back to the hypothesis space H, i.e.: $\Delta^\rho_P = \{h \in H \mid g(h) = \{\Delta'_1, \ldots, \Delta'_l\} \wedge \exists i \in \{1, \ldots, l\} : \Delta'_i \subseteq \Delta_{P^\rho}\}$.*

4 Quality of Reformulation

As noted in the previous section, the diagnosis Δ^ρ_P computed through reformulation may be different from the model-based diagnosis Δ_P. In this section, we study this issue.

As a first step, since the diagnosis is defined with respect to the semantics $sem(h)$ of hypotheses h, i.e., the set of behaviours that belong to h, it is useful to provide a definition of the semantics of $g(h)$.

Definition 4 *Let $g(h) = \{\Delta'_1, \ldots, \Delta'_l\}$, $\Delta'_i = \{h'_{i1}, \ldots, h'_{ik_i}\}$; the* semantics *of $g(h)$ is defined as:*

$$sem(g(h)) = \{\sigma \in \mathcal{L}_M \mid \exists i \in \{1, \ldots, l\} : \forall j \in \{1, \ldots, k_i\},\ \sigma \in sem(h'_{ij})\}.$$

This definition reflects the fact that $g(h)$ represents the disjunction of sets Δ'_i, and that each set Δ'_i represents the conjunction of the hypotheses h'_{ij}. Therefore, a trace σ belongs to $sem(g(h))$ iff σ belongs to $sem(h'_{ij})$ for all h'_{ij} in some Δ'_i.

The semantics of h and $g(h)$ should clearly be related, in order for the reformulation to yield meaningful diagnoses. We shall see however, that requiring that $sem(h) = sem(g(h))$ is not sufficient to ensure $\Delta^\rho_P = \Delta_P$.

4.1 Correctness of Reformulation

Definition 5 *The reformulation $g(h)$ of a hypothesis $h \in H$ is* r-correct *iff $sem(h) \subseteq sem(g(h))$. A reformulation $\rho = \langle g, H' \rangle$ is r-correct iff for each $h \in H$, $g(h)$ is r-correct.*

The following theorem relates r-correctness with d-correctness.

Theorem 2 *Let $\rho = \langle g, H' \rangle$ be a reformulation of H and $P = \langle M, \theta, H \rangle$ be an MBD problem. If ρ is r-correct and Δ' is a d-correct diagnosis for $P^\rho = \langle M, \theta, H' \rangle$, then $\Delta = g^{-1}(\Delta')$ is a d-correct diagnosis for P.*

Note that the theorem (as the following ones) holds in particular for the diagnosis through reformulation $\Delta^\rho_P = g^{-1}(\Delta_{P^\rho})$.

Corollary 1 *If the reformulation is r-correct, the diagnosis through reformulation is d-correct.*

The r-correctness of reformulation ρ is not, in general, a necessary condition for the d-correctness of Δ^ρ_P. However, this definition of r-correctness can be easily checked by considering just the hypotheses in the spaces H, H' and their semantics, and provides an effective sufficient condition to guarantee d-correctness of diagnosis through reformulation.

4.2 Precision of Reformulation

Definition 6 *The reformulation $g(h)$ of a hypothesis $h \in H$ is* r-precise *iff $sem(h) \supseteq sem(g(h))$. A reformulation $\rho = \langle g, H' \rangle$ is r-precise iff for each $h \in H$, $g(h)$ is r-precise.*

In general, it is not possible to make a statement about the d-precision of diagnosis through reformulation analogous to the one made in Theorem 2 about its d-correctness. However, it is possible to identify some important special cases.

Disjunctions

Definition 7 *Given a reformulation $\rho = \langle g, H' \rangle$, we say that g* disjunctively decomposes *hypothesis $h \in H$ if $g(h) = \{\{h'_1\}, \ldots, \{h'_l\}\}$. We also say that $g(h)$ is a* disjunction.

A disjunction is a decomposition of a hypothesis h of space H into a set of (non-exclusive) alternatives.

A typical disjunction consists in enumerating possible expressions of hypothesis h. Assume for instance a network composed of a master component and a collection of ten slave components; assume further that at any given moment, the master component is connected to exactly one component and that this connection can be determined from the observation of the master component; assume finally that hypothesis h represents a fault on the master component. Diagnosing h is difficult because it requires considering all eleven components of the network. A possible disjunction of h is the set $g(h) = \{\{h_1\}, \ldots, \{h_{10}\}\}$ where h_i corresponds to a fault when the master component is connected with its ith slave. Diagnosing h_i requires to monitor the observations from the master component and the ith slave, which is fairly simple. In this example, a hard decision problem is reformulated in a reasonable number of simple decision problems.

In the following theorem (as in subsequent ones), we assume for simplicity that the reformulation ρ maps each hypothesis $h \in H$ to itself, except for the hypotheses whose mapping is explicitly mentioned in the theorem. More complex reformulations can be viewed just as successive applications of these basic reformulations.

Theorem 3 *Let $\rho = \langle g, H' \rangle$ be a reformulation of H s.t. $g(\overline{h})$ is a disjunction $\{\{h'_1\}, \ldots, \{h'_l\}\}$ for some $\overline{h} \in H$, and let $P = \langle M, \theta, H \rangle$ be an MBD problem. If $g(\overline{h})$ is r-precise and Δ' is a d-precise diagnosis for $P^\rho = \langle M, \theta, H' \rangle$, then $\Delta = g^{-1}(\Delta')$ is a d-precise diagnosis for P.*

Conjunctions

Definition 8 *Given a reformulation $\rho = \langle g, H' \rangle$, we say that g* conjunctively decomposes *hypothesis $h \in H$ if $g(h) = \{\Delta'\}$, with $\Delta' = \{h'_1, \ldots, h'_k\}$. We also say that $g(h)$ is a* conjunction.

A conjunction is a decomposition of a hypothesis h of space H into a set of sub-hypotheses.

Unfortunately, an r-precise reformulation which contains a conjunction, does not guarantee that a diagnosis through reformulation $\Delta^\rho_P = g^{-1}(\Delta_{P^\rho})$ is d-precise, as shown in the following example.

Example 1 *Consider the DES modeled by the automaton in Figure 2 where a, b and c are the only observable events and f_1 and f_2 are the two faulty events. The space of diagnosis hypotheses is defined by $H = \mathbf{B}^{\{f_1, f_2\}}$. A hypothesis will be written $h_S \in H$ such that $h_S(f) = \top$ iff $f \in S$; for instance, the hypothesis which states that no fault occurred is written $h_\emptyset$.*

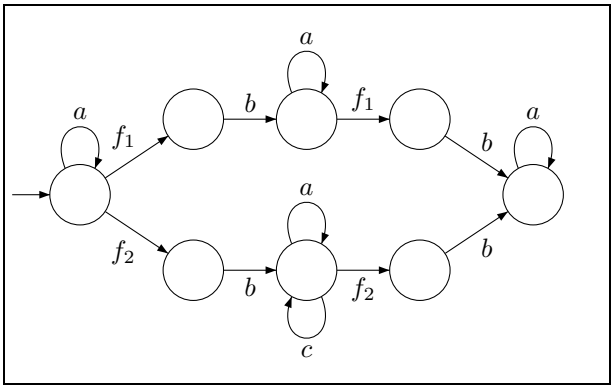

Figure 2: Illustration of loss of precision: if faults f_1 and f_2 are diagnosed separately, the correlation is lost.

Consider now observation $\theta = [b, a, a \ldots, a]$. MBD of observation θ is $\{h_{\{f_1\}}, h_{\{f_2\}}\}$ because seeing only one b and no c means one fault occurred.

Assume now the reformulation where each fault is diagnosed separately. The reformulated space is $H' = \{h_1, h_2, h'_1, h'_2\}$ where h_i corresponds to the traces where fault f_i occurs (and nothing is assumed about fault $f_j, j \neq i$), and h'_i corresponds to the traces where fault f_i does not occur. The reformulation of $h_{\{f_1, f_2\}}$ is $\{\{h_1, h_2\}\}$; it is clearly r-precise and r-correct.

The reformulated MBD diagnosis of θ is $\{h_1, h_2, h'_1, h'_2\}$, i.e., nothing can be decided about any fault. Therefore, the MBD through reformulation is $\{h_\emptyset, h_{\{f_1\}}, h_{\{f_2\}}, h_{\{f_1, f_2\}}\}$, and is therefore not d-precise.

We need to explore more closely what properties of the system guarantee that the diagnosis through reformulation is d-precise. We therefore introduce the following notation: the observations of a hypothesis h are the set of observations that can be emitted by some behaviour of h: $obs(h) = \{\theta \in Proj_{E_o}(\mathcal{L}_M) \mid \exists \sigma \in sem(h) : obs(\sigma) = \theta\}$. The observations of a set of hypotheses $\{h_1, \ldots, h_k\}$ are the set of observations that belong to the observations of each hypothesis h_i: $obs(\{h_1, \ldots, h_k\}) = \{\theta \in Proj_{E_o}(\mathcal{L}_M) \mid \forall i \in \{1, \ldots, k\}, \theta \in obs(h_i)\}$.

Theorem 4 *Let $\rho = \langle g, H' \rangle$ be a reformulation of H s.t. $g(\overline{h})$ is a conjunction $\{\{h'_1, \ldots, h'_k\}\}$ for some $\overline{h} \in$*

H, and let $P = \langle M, \theta, H\rangle$ be a diagnosis problem. If the observations of $\overline{h}$ cover the observations of $\{h'_1, \ldots, h'_k\}$, i.e. $obs(\overline{h}) \supseteq obs(\{h'_1, \ldots, h'_k\})$ and Δ' is a d-precise diagnosis for $P^\rho = \langle M, \theta, H'\rangle$, then $\Delta = g^{-1}(\Delta')$ is a d-precise diagnosis for P.

Because of Theorem 4, we say that h is (conjunctively) *decomposable* in $\{h'_1, \ldots, h'_k\}$ if the observations of h cover the observations of $\{h'_1, \ldots, h'_k\}$.

In principle, decomposability of h into $\{h'_1, \ldots, h'_k\}$ can be tested with a procedure such as the one illustrated in algorithm 1.

Algorithm 1 Testing decomposability.

input: model M, hypothesis h, set of hypotheses $\{h'_1, \ldots, h'_k\}$
$\mathcal{L} := Proj_{E_o}(\mathcal{L}_M) \setminus obs(h)$
for $i = 1 \ldots k$ **do**
  $\mathcal{L} := \mathcal{L} \cap Proj_{E_o}(sem(h'_i))$
end for
return $\mathcal{L} \stackrel{?}{=} \emptyset$

The algorithm starts with the set $\mathcal{L}$ of observations θ that are not observations of h, and incrementally discards from such a set the observations that cannot be simultanesously explained with the hypotheses $\{h'_1, \ldots, h'_i\}$, for increasing values of i up to k. If it ends up with an empty set, it means that $obs(h)$ covers $obs(\{h'_1, \ldots, h'_k\})$, i.e. that h is decomposable in $\{h'_1, \ldots, h'_k\}$.

In practice, an interesting way of ensuring decomposability is through the well-known *diagnosability* property. Let us first recast diagnosability in our framework.

Definition 9 *A hypothesis h is diagnosable on a model M if $\forall \sigma \in sem(h)$, $\forall \sigma' \in \mathcal{L}_M$, $obs(\sigma) = obs(\sigma') \Rightarrow \sigma' \in sem(h)$.*

Note that the original definition of diagnosability (Sampath et al. 1995) included a delay between the time instant when the hypothesis becomes true and the time instant when the fault can be diagnosed with certainty. This does not quite fit with our definition of hypothesis where the semantics $sem(h)$ of a hypothesis h is not necessarily stable (or "extension-closed") (Jéron et al. 2006).

Theorem 5 *Let M be a model and let $\rho = \langle g, H'\rangle$ be a reformulation of H s.t. $g(\overline{h})$ is a conjunction $\{\{h'_1, h'_2\}\}$ for some $\overline{h} \in H$. If ρ is r-precise and h'_1 is diagnosable on model M, then $\overline{h}$ is decomposable in $\{h'_1, h'_2\}$.*

This result can be easily extended for a decomposition in more than two elements.

Combining theorem 4 and theorem 5, it follows that, if hypothesis h is reformulated in a conjunction $\{h'_1, \ldots, h'_k\}$ s.t. all the (possibly but one) hypotheses h'_i are diagnosable, then the diagnosis through reformulation is still precise. A similar result can be derived from (Cordier, Travé-Massuyès, and Pucel 2006), where it is demonstrated that diagnosability of sets of faults is equivalent to diagnosability of every fault. Note that the condition on diagnosability we have discussed is sufficient but not necessary.

Aggregations

Definition 10 *Given a reformulation $\rho = \langle g, H'\rangle$, we say that g* aggregates *hypotheses $h_1, \ldots, h_m \in H$ if $g(h_1) = \ldots = g(h_m) = \{\{h'\}\}$. We also say that $\{\{h'\}\}$ is an* aggregation *of $h_1, \ldots, h_m$. An aggregation is said to be* overall precise *if $sem(h') \subseteq sem(h_1) \cup \ldots \cup sem(h_m)$.*

An *aggregation* is a composition of two or more hypotheses of space H into a single hypothesis in another space H'. Note that the notion of overall precision is weaker than that of precision, i.e. a precise aggregation is also overall precise but not viceversa.

Before providing a sufficient condition for d-precise diagnosis through reformulation in the presence of aggregations, we introduce the notion of *indistinguishability* between hypotheses.

Definition 11 *Two hypotheses h_1, h_2 are said to be indistinguishable w.r.t. a model M if $\forall \sigma_1 \in sem(h_1)$, $\exists \sigma_2 \in sem(h_2)$: $obs(\sigma_1) = obs(\sigma_2)$, and viceversa.*

It is easy to see that if h_1, h_2 are indistinguishable, then $obs(h_1) = obs(h_2)$. Moreover, if $\{\{h'\}\} = g(h_1) = g(h_2)$ is an overall precise aggregation of such indistinguishable hypotheses, also $obs(h')$ is equal to $obs(h_i), i \in \{1, 2\}$.

Theorem 6 *Let $\rho = \langle g, H'\rangle$ be a reformulation of H s.t. $\{\{h'\}\}, h' \in H'$ is an overall precise aggregation of $h_1, h_2 \in H$, and $P = \langle M, \theta, H\rangle$ be a diagnostic problem. If h_1, h_2 are indistinguishable w.r.t. M and Δ' is a d-precise diagnosis for $P^\rho = \langle M, \theta, H'\rangle$, then $\Delta = g^{-1}(\Delta')$ is a d-precise diagnosis for P.*

It is easy to extend this result to the aggregation of m indistinguishable hypotheses $h_1, \ldots, h_m$.

5 Examples of Reformulations

In this section we consider some specific reformulations that can be linked to previous works in the literature. Such works have shown that the reformulations they (implicitly) use are both practical and useful; by analyzing them within our framework we would like, on the one hand to demonstrate the applicability of our framework and, on the other hand, to hint at the benefits that specific applications of reformulation can get from a general framework.

5.1 Spatial decomposition

Very large networks – such as the Internet or electricity distribution networks – encompass thousands of interconnected components. A priori, the behaviour of any component and any sensor in the network may provide relevant information for the diagnosis task, so that applying model abstraction to alleviate complexity is all

but trivial. However, one of the important features of such networks is their distributive aspect; there are no "central" components in the network, which means that any defect from some component can usually be confined to a relatively small part of the network.

In this context, we envision an important use of reformulation related to the decomposition of global hypotheses into sets of local hypotheses. Consider a problem where $H = \mathbf{B}^{E_f}$, i.e. each hypothesis contains information about all possible faults. Now, consider $S \subset 2^{E_f}$, a collection of subsets of E_f that covers E_f (i.e. $\left(\bigcup_{E \in S} E\right) = E_f$). Define the hypothesis space H' as $H' = \left(\bigcup_{E \in S} H'_E\right)$, where $H'_E = \mathbf{B}^E$. Each hypothesis h'_E belonging to a subspace H'_E of H' contains information about all faults in subset E; we let $sem(h'_E) = \{\sigma \mid \forall f \in E,\ f \in \sigma \leftrightarrow h'_E(f) = \top\}$. Define ρ between H and H' s.t. for all h, $g(h)$ is the conjunction $\{\{\Delta'\}\}$ of the set of hypotheses $\Delta' \subseteq H'$ consistent with h.

This reformulation yields two benefits. First, the number of hypotheses is shrinked from $2^{|E_f|}$ to $\Sigma_{E \in S} 2^{|E|}$ which can be very beneficial, especially if for each E, $|E| \ll |E_f|$; in the best case, i.e., if the size of each $|E|$ is one, the number of hypotheses is reduced to $2 \times |E_f|$. Second, following our discussion in section 2, it is possible to solve the reformulated problem $P^\rho = \langle M, \theta, H' \rangle$ by solving sub-problems $P^\rho_E = \langle M, \theta, H'_E \rangle$. The main benefit of this is that specific abstractions can be applied to model M for each problem P^ρ_E, and such model abstractions can be extremely powerful if set E contains only a small part of set E_f.

Let us now consider if (and when) diagnosis through this kind of reformulation is d-correct and d-precise. D-correctness is guaranteed by theorem 2, since it is easy to see that the definition of ρ given above is r-correct. D-precision can be guaranteed by theorem 4, provided that each $h \in H$ s.t. $g(h) = \{\{\Delta'\}\}$ is decomposable in Δ'. Since decomposability is related to diagnosability by theorem 5, a possible technique to ensure the required decomposability of each $h \in H$ is to decide the sensors that will be placed on the system (Brandán Briones, Lazovik, and Dague 2008).

A previous work that proposes a technique close to this kind of reformulation is (Pencolé, Kamenetsky, and Schumann 2006). In the paper, the authors refuse the global picture of diagnosis and propose to test each fault separately with a *specialised diagnosers*. In terms of reformulation, the set of faults E_f is split into subsets $E_i = \{f_i\}$, each one containing just one fault f_i. Consequently, the reformulated space H' is defined as $\left(\bigcup_{f_i \in E_f} H'_{E_i}\right)$, and a specialized diagnoser is built for diagnosing each fault separately.

The approach in (Pencolé, Kamenetsky, and Schumann 2006) highly reduces the complexity, making it linear in the number of possible faults. However, as the authors acknowledge, the *information about fault correlations is lost by the specialised diagnosers*; we illustrated this fact in Example 1 where the fault correlation corresponds to *d*-precision. Such a loss of precision may be studied (and possibly avoided) by applying the framework developed in this paper. In particular, reasoning in terms of reformulation may suggest that the set of faults E_f should be better split in (small) subsets that are not necessary singletons; moreover, it may help in the choice of additional sensors that ensure precision.

5.2 Temporal decomposition

When the temporal window on which the diagnosis is defined is large, it may be more convenient to split it in smaller windows that can be diagnosed separately. Consider on the one hand a hypothesis $h \in H$ which states that a specific fault f occurred in the time window W during which the whole observation θ was collected; it is possible to define two hypotheses h'_1 and h'_2 each of which states that f occurred during a subwindows W_1 and W_2 respectively.

The reformulation $g(h)$ could be a disjunction $\{\{h'_1\}, \{h'_2\}\}$, as it suffices that one h_i is true for h to be true. Each hypothesis h'_i is concerned with the occurrence of event f during window W_i, but some other observations received in W might be required for a precise diagnosis; however, it should be possible to ignore most observation fragments in the subwindows associated with $h'_j \neq h'_i$. This has already been done in the chronicle-based approach (Cordier and Dousson 2000), where only contextual observation fragments are used; it also relates to *finite trackability* in (Grastien and Anbulagan 2009), where a decision about the behaviour at some time can be made within a limited time window.

5.3 Aggregated faults

Consider that, in order to build a hierarchical model of the system, we want to aggregate a set of components $c_1, \ldots, c_k$ into a subsystem Γ. For the sake of simplicity, let us assume that the model of each component c_i has its own fault event f_i which represents the failure of c_i, and let us ignore multiple fault hypotheses. If hypotheses $h_i \in H$ represent the occurrence of single faults f_i, we may want to map all the hypotheses h_i to the same hypothesis $h'_s \in H'$ which represents the failure of the subsystem s. If we let $sem(h'_s) = \bigcup_{i=1,\ldots,k} sem(h_i)$, such a reformulation is a correct and overall precise aggregation (section 4.2).

From theorem 2, we know that diagnosis through this kind of reformulation is d-correct. Moreover, theorem 6 tells us that, if hypotheses h_i are mutually indistinguishable, it is also d-precise. It is important to note that we may not be interested in the precision of diagnosis (i.e. we may be willing to apply reformulation even if hypotheses h_i are not mutually indistinguishable). One reason is because we may not have a practical interest in distinguishing the exact fault that occurred in subsystem s (see also the discussion below); in this case, the reformulation framework gives us the formal means to express our desired level of granularity of diagnosis,

which should be taken into account for performing useful model abstractions (Sachenbacher and Struss 2005). Another reason why we may accept imprecise diagnosis is because, after computing the diagnosis Δ_P^ρ through reformulation, we may want to refine it with further reasoning in the original space H, as it is typically done in hierarchical diagnosis. Also in this case the reformulation framework can be helpful, because the explicit definition of ρ gives us important information on how (computationally) easy will be the refinement step.

In (Cordier et al. 2007), the authors define *macro-faults* which correspond to sets of behaviours for which a similar recovery procedure can be taken. These macro-faults cover but, as in the present work, do not form a partition of the set of behaviours, as a single behaviour may be recovered from by different procedures. The macro-faults correspond to aggregations of hypotheses, and do not need to be mapped back. The authors put emphasis on precision at the level of macro faults: the diagnoser should return a macro-fault only if the system behaviour actually matches the macro-fault. Interestingly, the correctness property, as defined in the present paper, is not considered by the authors; if a behaviour belongs to several macro-faults, the diagnoser needs to return only one of these macro-faults.

In (Perrot and Travé-Massuyès 2007), the authors address the abstraction of static systems; their abstractions are defined as the aggregation of states (complete assignments of health variables), which constitute their diagnostic hypothesis space. Therefore, the abstractions studied in (Perrot and Travé-Massuyès 2007) can be viewed as reformulations (more specifically, aggregations). In such a context, the notions of *Concrete Solution Increasing* and *Concrete Solution Decreasing* abstractions (taken from the literature on abstraction and mentioned by the authors) closely match our notions of r-precision and r-correctness of reformulations.

6 Conclusion

In this paper, we have presented a framework for reformulating diagnosis hypotheses in the context of DES diagnosis.

As discussed in the examples, reformulations can open the way to performing powerful model abstractions, thus improving diagnosis efficiency; moreover, the possibility of explicitly mapping an hypothesis space into another one allows an improved control of the relevant diagnostic information that is (is not) lost with the transformation, and of the (computational) cost of retrieving such information if desired.

We have paid a particular attention to reformulations that fully preserve the correctness and precision of diagnosis, identifying a number of sufficient conditions which guarantee that a diagnosis procedure based on reformulation exhibits such properties. In future work, we would like to explore additional conditions related to correctness and precision, focusing on efficient ways of testing their truth for given diagnostic problems; we would like to include in our study also cases when it is not convenient (or desired) to completely preserve the precision of diagnosis across the reformulation.

Another future direction of the present research will be to study more deeply the relation between reformulations and the opportunities for model abstraction.

References

Benveniste, A.; Haar, St.; Fabre, É.; and Jard, Cl. 2005. Distributed monitoring of concurrent and asynchronous systems. *Journal of Discrete Event Dynamical Systems (JDEDS)* 15(1):33–84.

Brandán Briones, L.; Lazovik, A.; and Dague, Ph. 2008. Optimal observability for diagnosability. In *Proc. DX-08*, 31–38.

Cassandras, Ch., and Lafortune, St. 1999. *Introduction to Discrete Event Systems*. Kluwer Academic Publ.

Cordier, M.-O., and Dousson, Ch. 2000. Alarm driven monitoring based on chronicles. In *Proc. SafeProcess-00*, 286–291.

Cordier, M.-O.; Pencolé, Y.; Travé-Massuyès, L.; and Vidal, Th. 2007. Self-healablity = diagnosability + repairability. In *Proc. DX-07*, 251–258.

Cordier, M.-O.; Travé-Massuyès, L.; and Pucel, X. 2006. Comparing diagnosability in continuous and discrete-event systems. In *Proc. DX-06*, 55–60.

Grastien, A., and Anbulagan. 2009. Incremental diagnosis of DES with a non-exhaustive diagnosis engine. In *Proc. DX-09*, 345–352.

Grastien, Al., and Torta, Gi. 2010. Reformulation for the diagnosis of discrete-event systems. In *Proc. DX-10*, 63–70.

Grastien, Al., and Torta, Gi. 2011. A theory of abstraction for diagnosis of discrete-event systems. In *Proc. SARA-11*.

Jéron, Th.; Marchand, H.; Pinchinat, S.; and Cordier, M.-O. 2006. Supervision patterns in discrete-event systems diagnosis. In *Proc. DX-06*, 117–124.

Krysander, M., and Nyberg, M. 2008. Statistical properties and design criterions for fault isolation in noisy systems. In *Proc. DX-08*, 101–108.

Pencolé, Y.; Kamenetsky, D.; and Schumann, A. 2006. Towards low-cost fault diagnosis in large component-based systems. In *Proc. SafeProcess-06*.

Perrot, F., and Travé-Massuyès, L. 2007. Choosing abstractions for hierarchical diagnosis. In *Proc. DX-07*, 354–360.

Sachenbacher, M., and Struss, P. 2005. Task-dependent qualitative domain abstraction. *Artificial Intelligence (AIJ)* 162(1–2):121–143.

Sampath, M.; Sengupta, R.; Lafortune, St.; Sinnamohideen, K.; and Teneketzis, D. 1995. Diagnosability of discrete-event systems. *IEEE Transactions on Automatic Control (TAC)* 40(9):1555–1575.

Torta, G., and Torasso, P. 2008. A symbolic approach for component abstraction in model-based diagnosis. In *Proc. DX-08*, 355–362.