Self-healablity = diagnosability + repairability

Marie-Odile Cordier (+), Yannick Pencolé (*), Louise Travé-Massuyès (*), Thierry Vidal (+)
(+) IRISA/INRIA/University of Rennes1 - Campus de Beaulieu, 35042 Rennes Cédex – FRANCE
email: *Marie-Odile.Cordier, Thierry.Vidal@irisa.fr*
(*) LAAS-CNRS, University of Toulouse - Avenue Colonel Roche, 31077 Toulouse Cédex – FRANCE
email: *Louise.Trave-Massuyes, Yannick.Pencole@laas.fr*

Abstract

Real-life complex systems are often required to have a high level of autonomy, even in faulty situations. The diagnosability analysis is the a priori study of the capability of a system to be self-aware about its current state by analysing the observations received by the sensors. The repairability analysis is the a priori study of the capability of a system to react to faults by applying repair actions. Even though they are strongly related, these properties are generally analysed independently. This paper builds upon the state of the art in both domains and proposes a joint property, called self-healability, which achieves a bridge between diagnosability and repairability. We first revisit and extend the classical definitions of diagnosability and repairability. Then, we give our definition of self-healability illustrated with several examples.

1 Introduction

Real-life complex systems are often required to have a high level of autonomy, even in faulty situations. They are expected to be self-aware of their current state and reactive to faults by applying repair actions that take them back to nominal operation. In other words, such systems must be self-healing (Ghosh *et al.* 2007). Designing self-healing systems requires to be able to evaluate the joint degree of self-awareness and reactiveness. In the scientific community, these two properties are better known as diagnosability, i.e. the capability of a system and its monitors to exhibit different observables for different anticipated faulty situations, and repairability, i.e. the ability of a system and its repair actions to cope with any unexpected situation. However these two properties are generally analysed independently. On the one hand, the diagnosis community has proposed a formal definition of diagnosability (Sampath *et al.* 1995; Cordier, Travé-Massuyès, & Pucel 2006) and developped several approaches to check it (Jiang *et al.* 2001; Yoo & Lafortune 2002; Cimatti, Pecheur, & Cavada 2003; Jéron *et al.* 2006; Travé-Massuyès, Escobet, & Olive 2006; Schumann & Pencolé 2007). On the other hand, repairability has received much less attention, and refers to the software systems community that is more interested in control-based approaches such as fault tolerance or dependability as well as to the planning community that uses repair plans often without strong guarantees on their applicability.

A first naive approach could be suggested: first identify the set of all possible basic faults, then analyse separately their diagnosability (i.e. can the observed traces of any fault always be matched back to that fault with certainty ?) and their repairability (i.e. does there exist, for any fault, at least one plan of actions capable of repairing it once it has occurred ?). If one gets positive answers to both questions, then the system is self-healing. But that would be too strong a requirement because, as regards self-healability, complete diagnosability may not be needed: consider for instance that observations can always provide the information that fault f_1 or fault f_2 has occurred, but the system is not always able to discriminate them. One would conclude the system is not diagnosable with respect to these faults. But, if one has a repair plan r_k that is fine for repairing both f_1 and f_2, then this is enough to ensure self-healability! This example calls for a more general and integrated definition of diagnosability and repairability to serve the purpose of self-healability. The basic idea that is developped throughout this paper is to identify a set of so-called *macrofaults* (e.g. 'f_1 or f_2') such that these macrofaults are both diagnosable and repairable.

What our work shows is that facing diagnosability and repairability together calls for extending the classical diagnosability concept. This extension defines diagnosability for macrofaults that may overlap in terms of their underlying sets of basic faults, unlike diagnosability that is classically defined for partitions of basic faults. Self-healability is then defined as achieving a bridge between diagnosability and repairability. The strong form of self-healability checks that there exists at least one set of macrofaults which can always be diagnosed with certainty and be repaired. It is then shown that weakened forms of self-healability can be defined. The first weakened property requires full repairability of the macrofaults but accepts that they are diagnosed with certainty only for a subset of the possible behaviours of the system. The second property accepts partial repairability of a subset of macrofaults but requires that they are diagnosed with certainty for all the possible behaviours of the system.

Section 2 starts with some related work that helps positioning our work within the overall area of dynamic systems supervision, then section 3 gives some preliminary definitions and hypotheses on the faults, the observables and the repair actions. Section 4 focuses on the strong form of self-healability. We first give our definition of diagnosability, which is an extention of the classical definition. We then define repairability and conclude by joining these two properties to define self-healability. Section 5 defines two weakened forms of self-healability, namely weak self-healability and partial self-healability. Section 6 concludes by giving some perspectives on the way self-healability can be extended to include temporal constraints. Another issue is to show the part that (weakened) self-healability can play at design time to help improving the system properties by augmenting either the observable events or the set of repair actions in a coordinated way.

2 Related Work

In (Ghosh *et al.* 2007) self-healing systems are presented as a rather new and rapidly increasing domain of interest in the software systems community. The definition keeps relatively general, and is stated as:

> *Self-healability is the property that enables a system to perceive that it is not operating correctly and, without human intervention, make the necessary adjustments to restore itself to normality.*

That definition is related to the even more general definition of *dependable* systems, that defines the system as globally trustworthy with respect to its ability to always deliver its service, or *fault-tolerant* systems, in which faults that occur do not affect the efficiency of the system. But there are many ways to guarantee such a general property. *Controllability* of dynamic systems (Ramadge & Wonham 1989) is a way to proactively modify the architecture of a system at design time to hinder any occurrence of a catastrophic fault. Another way is to achieve fault-tolerance by on-line mechanisms, to be used whenever a fault occurs. That might be addressed through so-called *passive approaches* that rely on robust feedback mechanisms accounting for possible deviations from the normal behaviour. See for instance (Tornil, Escobet, & Travé-Massuyès 2003)(Garcia, Bernussou, & Arzelier 1994) in the continuous processes domain. In the software domain, approaches like *homeostasis* rely on the same idea, arguing that the distinction between health and unhealth is sometimes tenuous and proposing to maintain a stable internal state for the system despite external variations (Shaw 2002).

On the contrary, self-healing systems are clearly *recovery*-oriented: a clear difference is made between healthy and unhealthy situations, and the system is monitored to

- detect when the system moves from a normal mode to a faulty mode,
- diagnose the situation,
- choose and execute an appropriate recovery strategy.

That principle can be seen as an *active approach* in the fault-tolerant systems community. It may however result, as can be seen in (Saboori & Zad 2005), in integrating recovery modes in the global model of the system, making then tenious the boundary between passive and active approaches.

Our work is clearly positioned towards self-healability, adopting a *diagnose/repair*-based view. The model specifying the (normal and faulty) behaviours of the system is clearly separated from the repair strategies that are available for bringing back the system from a faulty mode to a normal mode. In the planning community, *plan revision*, *plan adaptation* and *replanning techniques* are variations of the same general idea: in case of a disruption in the current execution of a plan, the nominal plan is halted and some alternative sequence of actions is executed, the latter being either precomputed off line or computed on line. Our work is somehow related to the work in (Washington, Golden, & Bresina 1999), in which so-called *alternate plans* have been computed off line and uploaded on a space system, and the monitoring module tracks deviations and switch to one of such alternate plans whenever needed, to put the system back in a state in which the nominal plan can be restarted. But such planning approaches show two shortcomings :

- full observability is assumed, turning diagnosis rather straightforward, and
- there is no formal analysis of the 'self-healing' property of the system, i.e. there is no guarantee that there always exists a plan execution that successfully repairs the system whatever fault occurs.

Both issues are addressed in this paper.

Removing the assumption of full observability requires a thorough analysis of *diagnosability*. Diagnosability covers a set of properties that have been studied for many years in different fields. In the context of continuous systems, diagnosability is stated in terms of detectability and isolability (Chen & Patton 1994; Basseville 2001). In the DES context, the first definitions have been proposed in (Sampath *et al.* 1995) and several extensions exist for intermittent faults (Contant, Lafortune, & Teneketsis 2004) and for more generic patterns (Jéron *et al.* 2006). A unified diagnosability definition for continuous systems and discrete-event systems is proposed in (Cordier, Travé-Massuyès, & Pucel 2006). Checking diagnosability is computationally complex and several algorithms have been proposed (Yoo & Lafortune 2002; Jiang *et al.* 2001; Cimatti, Pecheur, & Cavada 2003; Schumann & Pencolé 2007). One of the main purpose of this analysis is to provide an assistance for the design of the system, like sensor placements (Travé-Massuyès, Escobet, & Milne 2001) or communication protocol specifications (Pencolé 2005).

On the other hand, assuring, in any situation and at any time, successfull repair, calls for the definition of a formal *repairability* property, which to our knowledge has never been thoroughly defined.

Ultimately, our work aims at combining diagnosability and repairability to reach a self-healability property, consistent with (Ghosh *et al.* 2007), that can be ensured at design time.

3 Preliminaries

In this paper, we adopt the generic viewpoint defined in (Cordier, Travé-Massuyès, & Pucel 2006). A system may be a discrete-event system or a continuous system. Depending on its inputs, the system may have several *behaviours* and produce several outputs. We first define preliminary notions that are used throughout this paper.

3.1 Faults

The set of possible faults $\mathcal{F}$ of the system is composed of nf *basic faults* $\mathcal{F} = \{f_1, \ldots, f_{nf}\}$. This paper relies on the following single fault assumption.

Assumption 1 *If the behaviour of the system is faulty, only one fault is present.*

In the following and without loss of generality, f_i denotes either the type of the fault or its presence in the system. Moreover, in order to be generic, we extend $\mathcal{F}$ with the symbol $norm$ which stands for the absence of fault. In the following, $norm$ is considered just like any other f_i of $\mathcal{F}$.

Definition 1 (Macrofault) *A* macrofault F_j *is a set of faults* $F_j \subseteq \mathcal{F}, F_j \neq \emptyset$*. If* F_j *is present in the system, it means that one and only one of the basic faults* $f_i \in F_j$ *is present in the system.*

For instance, the macrofault $\{f_1, f_2\}$ represents the fact that either f_1 or f_2 is present, the macrofault $\{f_2, norm\}$ means that either the fault f_2 is present or there is no fault.

A macrofault may be a singleton ($F_j = \{f_i\}$), i.e. basic faults may be seen as special cases of macrofaults. A set of macrofaults is then denoted by $E(\mathcal{F})$ (with $E(\mathcal{F}) \subseteq 2^{\mathcal{F}}$).

Definition 2 (Covering set) *A set of macrofaults* $E(\mathcal{F})$ *is a* covering set *of* $\mathcal{F}$ *if and only if* $\forall f_i \in \mathcal{F}, \exists F_j \in E(\mathcal{F})$ *such that* $f_i \in F_j$.

3.2 Observations

Depending on the type of system, observations consist of sequences of observable events or sets of measured values for observable variables. These observations are the observable part of the system's behaviour. This paper focuses on defining the generic notion of self-healability which does not require the distinction between these different types of observations. Thus, for the sake of generality, the observability of the system is represented by the set OBS of all the possible observations of the system. $\sigma \in OBS$ is called an *observable* of the system.

In the case of discrete event systems (DES), OBS is usually specified through the set of observable events $\mathcal{O} = \{o_1, ..., o_{no}\}$. Complementing $\mathcal{O}$ with the set of unobservable events $\mathcal{U} = \{u_1, ..., u_{nu}\}$ determines the whole set of events of the system $\mathcal{E} = \mathcal{O} \cup \mathcal{U}$. Faults are specific unobservable events $\mathcal{F} \subseteq \mathcal{U}$.

In the case of continuous systems, OBS is usually specified as the set of all possible observation tuples for observable variables, i.e., $OBS = \{(o_1, o_2, ..., o_k)\}$ where k is the number of sensoring devices. Faults can be thought as unknown disturbances that affect the system's behaviour and hence the observations.

3.3 Repairs

We first define, in a simplified way, what a repair plan is. Planning contexts may of course require more elaborated definitions, but the global property definitions proposed in this article remain the same. For our purpose, we do not really need to know what a repair plan is semantically speaking, we merely need to have a set of such repair plans and to be able to match them to (basic) faults.

Definition 3 (Repair Plan) *A* repair plan r_k *is defined as a sequence of actions* $A_k = \{a_{k1}, ..., a_{kn}\}$*, each* a_{ki} *belonging to the set of all possible recovering basic actions (which are seen as events in DES). A repair plan is always assumed to be applicable (no preconditions). Each repair plan* r_k *has a goal* g_k *which is the nominal state of the system in which the repair plan is expected to bring the system back.*

The set of available repair plans is denoted $\mathcal{R} = \{r_1, ..., r_{nr}\}$ One can see that the definition of a repair plan is independent from the faults it might apply to, one then needs to establish that relation (just like signatures are sets of observations that are related to faults). That is done through the predicate $Repair$:

$Repair(r_k, f_i)$ means that the repair plan r_k, if executed in a state in which f_i is present, brings the system back to a state in which the goal g_k should be true and $norm$ present[1].

The $Repair$ predicate can also be applied to a macrofault, and there is an equivalence between repairing a macrofault and repairing each of the basic faults belonging to the macrofault, i.e.

Property 1 $Repair(r_k, F_j) \equiv \forall f_i \in F_j, Repair(r_k, f_i).$

4 Self-healability

Self-healability is intuitively defined by the following:

A system is self-healing if, and only if, after the occurence of any basic fault, a diagnosis is issued that automatically raises a repair plan fitted for the fault.

Behind this intuitive definition, two properties of the system are hidden: diagnosability and repairability. What is also hidden is *what* the system should be able to diagnose and *which* plan should then be applied. In this section, we first introduce these two notions to finally propose a first definition for self-healability.

4.1 Diagnosability

A first requirement for a diagnosable system is that every behaviour (faulty or not) must produce some observations. If no observation is available, the system is not diagnosable at all. Therefore, one requirement for self-healability is:

Property 2 *The system is not silent: every possible behaviour has an associated observable.*

In the case where the observations are the result of some measurements (in a continuous system for instance), this property is usually satisfied since the measures can be performed whatever the behaviour of the system really is. In the case of discrete-event models, this property implies some consequences on the system itself because the observations are usually events emitted by the system (in the literature about DES, this property is usually called the *liveness* of the observations (Sampath *et al.* 1995)).

Diagnosability analysis relies on the notion of *fault signatures* (Cordier, Travé-Massuyès, & Pucel 2006). Intuitively,

[1] We consider that there always exists a repair plan r_{norm} such that $Repair(r_{norm}, norm)$, r_{norm} being the empty plan with a set of actions $A_k = \emptyset$.

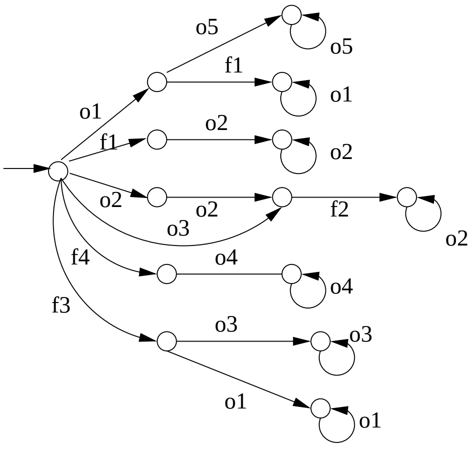

Figure 1: An example.

a fault signature is the association between a fault and a set of possible observations. The following definition defines the concept of elementary signature, noted *e-signature*. An e-signature corresponds to one observable in the presence of the fault and in a given context.

Definition 4 (Fault e-signature) *An* e-signature *σ of F_j is an observable such that there exists a behaviour of the system for which σ is observed when a fault $f_i \in F_j$ is present.*

Example:

Figure 1 presents the global model of a discrete-event system Σ. The set of basic faults is $\mathcal{F} = \{norm, f_1, f_2, f_3, f_4\}$ and the fault events are not observable. The other events (i.e. o_1, o_2, o_3, o_4 and o_5) are observable.

$o1o5^\infty$ is an e-signature of the normal behaviour; it is composed of one observation $o1$ followed by an infinite sequence of $o5$. $o2^\infty$ is an e-signature of the fault $\{f_1\}$ and of the fault $\{f_2\}$, therefore it is also an e-signature of $\{f_1, f_2\}$, but it can also be considered as an e-signature of, for instance, $\{f_1, f_2, f_3\}$.

The observation of an e-signature σ denotes that the fault F_j may be present in the system. The presence of F_j is not certain since it is possible that σ is also the e-signature of another fault $F_{j'}$.

Definition 5 (Characteristic e-signature) *An e-signature σ* characterizes *F_j if and only if in every behaviour that produces the e-signature σ, a fault $f_i \in F_j$ is present.*

If an e-signature σ *characterizes* the presence of F_j, it means that the observation of σ guarantees that the fault F_j is certainly present in the system. Note that an e-signature that characterizes F_j may characterise a fault $F_{j'}$ where $F_{j'} \cap F_j \neq \emptyset$, and any e-signature characterizes the macrofault $\mathcal{F}$.

In other words, when observing a characteristic σ, one is sure that one of the $f_i \in F_j$ has occurred, and that other $f_{i'}$ which are not in F_j have not occurred. But each $f_i \in F_j$ may have other e-signatures that will be characteristic of other macrofaults that also include f_i.

Example : $o3o2^\infty$ is a characteristic e-signature of the fault $\{f_2\}$ and of the fault $\{f_1, f_2\}$. $o2^\infty$ is a characteristic e-signature of the fault $\{f_1, f_2\}$.

Definition 6 (Characteristic signature) *A* characteristic signature *of F_j is a set of e-signatures that characterise F_j. Such a set is denoted $Sig(F_j)$.*

Example : $\{o1^\infty, o2^\infty\}$, $\{o2^\infty, o3o2^\infty\}$,$\{o1^\infty, o3^\infty\}$ are the respective sets of all the e-signatures of $\{f_1\}$,$\{f_2\}$, $\{f_3\}$ but are not characteristic signatures of $\{f_1\}$,$\{f_2\}$, $\{f_3\}$. $\{o4^\infty\}$ is the set of all the e-signatures of $\{f_4\}$ and is a characteristic signature of $\{f_4\}$, noted $Sig(\{f_4\})$. $\{o1^\infty, o3^\infty\}$ is a characteristic signature of $\{f_1, f_3\}$, as well as $\{o3^\infty\}$ and $\{o1^\infty\}$.

Now, we are ready to propose a definition for diagnosability that is suitable for self-healability. Classically, diagnosability is formally defined on a partition of the basic faults of $\mathcal{F}$ (Cordier, Travé-Massuyès, & Pucel 2006). In this paper, the definition of diagnosability is extended to a covering set of faults $E(\mathcal{F}) = \{F_1, \ldots, F_m\}$.

Definition 7 (Diagnosability of a set of macrofaults) *The covering set $E(\mathcal{F})$ is* diagnosable *if and only if there exists m sets of characteristic signatures $Sig(F_1), \ldots, Sig(F_m)$ such that:*

1. $\bigcup_{i=1}^{m} Sig(F_i) = OBS$;
2. $\forall i, j, i \neq j, Sig(F_i) \cap Sig(F_j) = \emptyset$.

In other words, the set $PSig = \{Sig(F_1), \ldots, Sig(F_m)\}$ is a *partition* of the e-signatures, and $E(\mathcal{F})$ is *diagnosable* if and only if a partition $PSig$ exists such that the observation of $\sigma \in Sig(F_j)$ guarantees that F_j is present.

Example: The set of macrofaults $E_0(\mathcal{F}) = \{\{norm, f_1, f_2, f_3, f_4\}\}$ is diagnosable, the partition of e-signatures is $PSig_0 = \{\{o1o5^\infty, o_1^\infty, o_2^\infty, o_3o_2^\infty, o_3^\infty, o_4^\infty\}\}$. The set of macrofaults $E_1(\mathcal{F}) = \{\{norm\}, \{f_1, f_2, f_3\}, \{f_4\}\}$ is diagnosable, the partition of e-signatures is $PSig_1 = \{\{o1o5^\infty\}, \{o_1^\infty, o_2^\infty, o_3o_2^\infty, o_3^\infty\}, \{o_4^\infty\}\}$. The set $E_2(\mathcal{F}) = \{\{norm\}, \{f_1, f_2\}, \{f_1, f_3\}, \{f_4\}\}$ is also diagnosable and the corresponding partition is $PSig_2 = \{\{o1o5^\infty\}, \{o_2^\infty, o_3o_2^\infty\}, \{o_1^\infty, o_3^\infty\}, \{o_4^\infty\}\}$. The set $E_3(\mathcal{F}) = \{\{norm\}, \{f_1\}, \{f_2\}, \{f_3\}, \{f_4\}\}$ is not diagnosable because there are some cases in which f_1 and f_2 are not discriminated (the same for f_1 and f_3) . $E_4(\mathcal{F}) = \{\{norm\}, \{f_1\}, \{f_2, f_3\}, \{f_4\}\}$ is not diagnosable. $E_5(\mathcal{F}) = \{\{norm\}, \{f_1, f_3\}, \{f_2\}\{f_4\}\}$ is not diagnosable.

This diagnosability definition is generic and covers different concepts. For instance, if there is a macrofault in $E(\mathcal{F})$ which contains $norm$ and at least a basic fault f_i, and $E(\mathcal{F})$ is diagnosable, it means the system is not *fault-detectable* with respect to that set of macrofaults: some faulty behaviour, whatever the fault is, may not be detected. On the contrary, if $\{norm\} \in E(\mathcal{F})$ and $E(\mathcal{F})$ is diagnosable and none of the other macrofaults contain $norm$ then all faulty behaviours are detectable.

In the specific case where $E(\mathcal{F})$ is a partition of the basic faults with $\{norm\} \in E(\mathcal{F})$, the diagnosability of $E(\mathcal{F})$

corresponds to the classical definition over a partition of basic faults (see (Cordier, Travé-Massuyès, & Pucel 2006)).

We use the predicate $Diagnosable(E(\mathcal{F}))$ to tell that $E(\mathcal{F})$ is diagnosable.

Then the definition of the diagnosability of the system, through the simple predicate $Diagnosable$, comes directly:

Definition 8 (Diagnosability) $Diagnosable \equiv \exists E(\mathcal{F})$ *such that* $Diagnosable(E(\mathcal{F}))$

Let us remark that a system is always diagnosable as it is always true that $Diagnosable(\{\mathcal{F}\})$.

4.2 Repairability

Repairability is actually straightforward and comes directly from the basic definition of sets of macrofaults and through the predicate $Repair$. First, a macrofault F_j is repairable if and only if there exists a repair plan that repairs it:

Definition 9 (Repairability of a macrofault) $Repairable(F_j) \equiv \exists r_k$ *such that* $Repair(r_k, F_j)$

Then, considering the basic faults belonging to a macrofault, we have the implication:

Property 3 $Repairable(\{f_i, f_j\}) \models Repairable(f_i)$ *and* $Repairable(f_j)$

The converse is of course not always true (only when a common plan repairs both f_i and f_j).

The repairability of a set of macrofaults is then defined as the repairability of all the macrofaults in the set:

Definition 10 (Repairability of a set of macrofaults) $Repairable(E(\mathcal{F})) \equiv \forall F_j \in E(\mathcal{F}) Repairable(F_j)$

The repairability of a system (that we simply note through the predicate $Repairable$ with no parameters) means that there exists a set of macrofaults which covers all the basic faults and that is repairable.

Definition 11 (Repairability of the system) $Repairable \equiv \exists E(\mathcal{F})$ *such that* $Repairable(E(\mathcal{F}))$ *and* $E(\mathcal{F})$ *is a covering set*

Example : If the only repair plan is r, with $Repair(r, \{f_1, f_3\})$, we get Repairable($\{f_1$,$f_3\}$), and also Repairable($\{f_1\}$) and Repairable($\{f_3\}$). However, the system is clearly not repairable as the faults f_2 and f_4 are not repairable.

4.3 Self-healability

Our first definition for self-healibility directly derives from the definitions of diagnosability and repairability.

Definition 12 (Self-healing covering set) *A covering set* $E(\mathcal{F})$ *is* self-healing *iff it is diagnosable and repairable, i.e.* $SelfHealing(E(\mathcal{F})) \equiv Diagnosable(E(\mathcal{F}))$ *and* $Repairable(E(\mathcal{F}))$

Given a self-healing covering set $E(\mathcal{F})$, the observation of a signature $\sigma \in OBS$ is always associated to a macrofault $F_\sigma \in E(\mathcal{F})$ because $E(\mathcal{F})$ is diagnosable. Moreover, since $E(\mathcal{F})$ is repairable, it means that there exists a repair plan for F_σ (a repair plan that repairs any basic fault of F_σ), therefore F_σ can be repaired. Finally, $E(\mathcal{F})$ is a covering set so every basic fault is covered by at least one macrofault, so every basic fault can be healed.

Definition 13 (Self-healing system) *A system is* self-healing *iff there exists a self-healing covering set* $E(\mathcal{F})$*, i.e.* $SelfHealing \equiv \exists E(\mathcal{F})$ *such that* $E(\mathcal{F})$ *is a covering set and* $SelfHealing(E(\mathcal{F}))$.

Example :

- Case 0 : Repairable($norm$) and Repairable($\{f_1$,$f_3\}$) and Repairable($\{f_1$,$f_2\}$) and Repairable(f_4)

  The set $E_2(\mathcal{F}) = \{\{norm\}, \{f_1, f_2\}, \{f_1, f_3\}, \{f_4\}\}$ is diagnosable and repairable. The system is then self-healing. It would also have been the case if we had repair plans such that Repairable($\{f_1$,f_2,$f_3\}$) and Repairable(f_4) because $E_1(\mathcal{F}) = \{\{norm\}, \{f_1, f_2, f_3\}, \{f_4\}\}$ is diagnosable.
- Case 1 : Repairable($norm$) and Repairable($\{f_1$,$f_3\}$) and Repairable(f_2) and Repairable(f_4)

  No set of macrofaults is self-healing as there does not exist a repair plan for $\{f_1, f_2\}$. The system is thus not self-healing.

As said previously, in a self-healing system every basic fault can be repaired (even in cases where it cannot be fully distinguished from other faults). This property is the strong form of self-healability. In the next section, we propose weakened forms of self-healability that can also be of interest, for instance during the design step.

5 Weakened forms of self-healability

The strong form of self-healability seen above is exactly what is needed when checking a system as it indicates that each basic fault can always be repaired, either because it can always be diagnosed and repaired, or because a macrofault containing it can always be diagnosed and repaired. In other words, it means that, for any occurrence of a fault, the observations and repair plans are sufficient to ensure a sufficiently precise diagnosis enabling to trigger a repair plan.

However, weakened forms of self-healability are interesting. For instance, when designing a system, it can be useful to guide the designer to improve the current self-healability. It can also be useful to identify a subpart of basic faults which are self-healing even if the whole system is not.

In the following, we define two weakened forms of self-healability. The first one, called *weak self-healability*, consists in assuring that each fault can be diagnosed and repaired at least in some contexts. The second one, called *partial self-healability*, consists in assuring that there exists a subset of faults that are always self-healing.

5.1 Weak Self-Healability

A basic fault can manifest itself according to a set of e-signatures. Weak self-healability is a way to decide if, at least for a subset of its e-signatures, each basic fault is repairable.

Weak Self-Healability relies on weak diagnosability and repairability. Let us first define what we call weak diagnosability.

Weak Diagnosability A covering set of macrofaults $E(\mathcal{F})$ is weakly diagnosable iff for each macrofault of $E(\mathcal{F})$, there exists at least an e-signature which characterises it. However, it may also exist e-signatures which do not characterize any macrofault of $E(\mathcal{F})$.

Definition 14 (Weak Diagnosability of a set of macrofaults) *The set of n macrofaults $E(\mathcal{F})$ is* weakly diagnosable *iff it is a covering set and there exists a partition of e-signatures $PSig = \{Sig_1, \ldots, Sig_r\}$ such that, $r \geq n$, for all $i \in \{1, \ldots, r\}, Sig_i \neq \emptyset$ and for all $i \in \{1, \ldots, n\}$, the signature Sig_i characterizes the macrofault F_i.*

Another way of expressing weak diagnosability is to say that the covering set of macrofaults $E(\mathcal{F})$ is *weakly diagnosable* iff there exists a set of macrofaults $E'(\mathcal{F})$, such that $E'(\mathcal{F})$ is diagnosable and $E(\mathcal{F}) \subseteq E'(\mathcal{F})$.

We will note weak diagnosability of a set of macrofaults $WeaklyDiagnosable(E(\mathcal{F}))$ and consequently weak diagnosability of the whole system can be defined as

Definition 15 (Weak Diagnosability) $WeaklyDiagnosable \equiv \exists E(\mathcal{F})$ *such that* $WeaklyDiagnosable(E(\mathcal{F}))$

As seen before, a system is always diagnosable and thus it is also always weakly diagnosable.

Example:

By definition, every diagnosable $E(\mathcal{F})$ is also weakly diagnosable, so $E_0(\mathcal{F})$, $E_1(\mathcal{F})$, $E_2(\mathcal{F})$ are weakly diagnosable. If we consider the set $E_3(\mathcal{F}) = \{\{norm\}, \{f_1\}, \{f_2\}, \{f_3\}, \{f_4\}\}$, it is not weakly diagnosable as $\{f_1\}$ has no characteristic e-signature. In the same way, $E_4(\mathcal{F})$ is not weakly diagnosable. However, $E_5(\mathcal{F})$ is weakly diagnosable with the corresponding partition of e-signatures $PSig_5 = \{\{o_1 o_5^\infty\}, \{o_1^\infty o_3^\infty\}, \{o_3 o_2^\infty\}, \{o_4^\infty\}, \{o_2^\infty\}\}\}$. The first four e-signatures characterize the first four macrofaults; the last one, o_2^∞, is a supplementary one which does not characterize any macrofault of the set $E_5(\mathcal{F})$.

If $DiagnosableSet(\mathcal{F})$ and $WeaklyDiagnosableSet(\mathcal{F})$ are the sets of sets of macrofaults $E(\mathcal{F})$ such that $E(\mathcal{F})$ is diagnosable and weakly diagnosable, respectively. In this example, we have $\{E_0(\mathcal{F}), E_1(\mathcal{F}), E_2(\mathcal{F}), E_5(\mathcal{F})\} \subseteq WeaklyDiagnosableSet(\mathcal{F})$ and by definition, one always gets: $DiagnosableSet(\mathcal{F}) \subseteq WeaklyDiagnosableSet(\mathcal{F})$.

Weak Self-healability The definition of weak self-healability follows:

Definition 16 (Weak self-healability of a set of macrofaults) *The covering set of macrofaults $E(\mathcal{F})$ is* weakly self-healing *iff it is weakly diagnosable and repairable.*

It can be shown from what was said before that the set of macrofaults $E(\mathcal{F})$ is *weakly self-healing* iff there exists a set of macrofaults $E'(\mathcal{F})$, such that $E'(\mathcal{F})$ is self-healing and $E(\mathcal{F}) \subseteq E'(\mathcal{F})$.

Definition 17 (Weak self-healability of the system) *A system is weakly self-healing iff there exists a covering set $E(\mathcal{F})$ that is weakly self-healing.*

Here again we can use a predicate notation: $WeaklySelfHealing \equiv \exists E(\mathcal{F})$ such that $WeaklySelfHealing(E(\mathcal{F}))$.

Hence the following propositions:

Proposition 1 *1. If a system is self-healing then it is weakly self-healing.*

2. If a system is weakly self-healing, it may not be self-healing.

Proof: The proof of 1 comes from the definitions. The proof of 2 comes from the case 1 of the following example.

Example :

- Case 0 : Repairable($norm$) and Repairable($\{f_1,f_3\}$) and Repairable($\{f_1,f_2\}$) and Repairable(f_4)

  The set $E_2(\mathcal{F}) = \{\{norm\}, \{f_1, f_2\}, \{f_1, f_3\}, \{f_4\}\}$ is diagnosable and repairable. The system is then self-healing and consequenly also weakly self-healing.

- Case 1 : Repairable($norm$) and Repairable($\{f_1,f_3\}$) and Repairable(f_2) and Repairable(f_4)

  No set of macrofaults is self-healing. $E_5(\mathcal{F}) = \{\{norm\}, \{f_1, f_3\}, \{f_2\}\{f_4\}\}$ is not self-healing as it is not diagnosable, but it is weakly self-healing as it is weakly diagnosable and repairable. The system is thus not self-healing but weakly self-healing.

- Case 2 : Repairable($norm$) and Repairable(f_1) and Repairable(f_2) and Repairable(f_3)and Repairable(f_4)

  No set of macrofaults is either self-healing or weakly self-healing as f_1 can never be diagnosed with certainty (even in some specific contexts). The system is thus not self-healing and not weakly self-healing.

5.2 Partial Self-Healability

Another weakened form of self-healability is to guarantee that a subset of macrofaults (or basic faults) are self-healing, even if not all of them. The idea is to guarantee that there exists a subset of basic faults which are self-healing, and thus both diagnosable and repairable.

We first need to introduce the following definition.

Definition 18 (No-overlap property) *The set of sets Y has the no-overlap property with the set of sets X iff $Y \subseteq X$ and $\forall y \in Y \ \forall x \in X \setminus Y \ x \cap y = \emptyset$.*

If $X = \{\{f1\}, \{f1, f2\}, \{f3\}\}$ then $Y = \{\{f1\}, \{f1, f2\}\}$ has the no-overlap property with X but $Y = \{\{f1\}\}$ has not.

Definition 19 (Partial Self-Healability of a set of macrofaults) *The (non covering) set of macrofaults $E(\mathcal{F})$ is* partially self-healing *iff there exists a set of macrofaults $E'(\mathcal{F})$, such that $E'(\mathcal{F})$ is diagnosable and $E(\mathcal{F})$* has the no-overlap property with *$E'(\mathcal{F})$ and $E(\mathcal{F})$ is repairable.*

The no-overlap property is important as it ensures that each macrofault of $E(\mathcal{F})$ can be diagnosed in any context.

Definition 20 (Partial self-healability of the system) *A system is partially self-healing iff there exists a set $E(\mathcal{F})$ that is partially self-healing.*

Here again we can use a predicate notation:
$PartiallySelfHealing \equiv \exists E(\mathcal{F})$ such that $PartiallySelfHealing(E(\mathcal{F}))$.

Let us remark that, as we consider the normal behaviour as a basic fault, denoted $norm$, and if there exists a repair plan, for instance the empty one, associated to $norm$, then a detectable system is always partially self-healing.

Example :

- Case 0 : Repairable($norm$) and Repairable($\{f_1,f_3\}$) and Repairable($\{f_1,f_2\}$) and Repairable(f_4)

  The set $E_2(\mathcal{F}) = \{\{norm\}, \{f_1, f_2\}, \{f_1, f_3\}, \{f_4\}\}$ is diagnosable and repairable. The system is then self-healing and consequenly also weakly and partially self-healing.

- Case 1 : Repairable($norm$) and Repairable(f_1,f_3) and Repairable(f_2) and Repairable(f_4)

  $E_6(\mathcal{F}) = \{\{f_4\}\}$ is partially self-healing as it exists a covering set of macrofaults $E_2(\mathcal{F})$ such that $E_6(\mathcal{F}) \subset E_2(\mathcal{F})$, $E_6(\mathcal{F})$ and $E_2(\mathcal{F})$ has the no-overlap property, $E_2(\mathcal{F})$ is diagnosable, and $E_6(\mathcal{F}) = \{\{f_4\}\}$ is repairable. It shows that the basic fault f_4 can be repaired for any of its occurrences.

  The set $E_9(\mathcal{F}) = \{\{f_1, f_3\}\}$ is not partially self-healing. It exists a covering set of macrofaults $E_2(\mathcal{F})$ such that $E_9(\mathcal{F}) \subset E_2(\mathcal{F})$, $E_2(\mathcal{F})$ is diagnosable, and $E_9(\mathcal{F}) = \{\{f_1, f_3\}\}$ is repairable but $E_9(\mathcal{F})$ and $E_2(\mathcal{F})$ do not satisfy the no-overlap property; it means that there exists at least an e-signature, here o_2^{∞}, which is an e-signature of a macrofault of the considered set, here of $\{f_1, f_3\}$, but also of another macrofault not in the considered set, here, $\{f_1, f_2\}$. Even if $E_9(\mathcal{F})$ is repairable, it means that the basic fault f_1 cannot be repaired in any contexts. It would only be true if we had also a repair plan for $\{f_1, f_2\}$.

  The system is thus not self-healing but weakly self-healing and partially self-healing.

- Case 2 : Repairable($norm$) and Repairable(f_1) and Repairable(f_2) and Repairable(f_3) and Repairable(f_4)

  The system is not self-healing and not weakly self-healing. However, as in case 1, $E_6(\mathcal{F}) = \{\{f_4\}\}$ is partially self-healing and the system is thus partially self-healing.

- Case 3 : Repairable($norm$) and Repairable($\{f_1,f_2\}$) and Repairable($\{f_1,f_3\}$)

  Again, the system is not self-healing and not weakly self-healing. However, $E_7(\mathcal{F}) = \{\{f_1, f_2\}, \{f_1, f_3\}\}$ is partially self-healing as it exists a covering set of macrofaults $E_2(\mathcal{F})$ such that $E_7(\mathcal{F}) \subset E_2(\mathcal{F})$, $E_7(\mathcal{F})$ and $E_2(\mathcal{F})$ has the no-overlap property, $E_2(\mathcal{F})$ is diagnosable, and $E_7(\mathcal{F}) = \{\{f_1, f_2\}, \{f_1, f_3\}\}$ is repairable. The two macrofaults $\{f_1, f_2\}$ and$\{f_1, f_3\}$ can be diagnosed with certainty and are both repairable. It demonstrates that the basic faults f_1, f_2 and f_3 can thus be repaired for any of their occurrences. The system is partially self-healing.

  $E_8(\mathcal{F}) = \{\{f_1, f_2, f_3\}\}$ is not partially self-healing as there does not exist any repair plan for $\{f_1, f_2, f_3\}$. In case we had a repair plan for $\{f_1, f_2, f_3\}$, $E_7(\mathcal{F}) = \{\{f_1, f_2, f_3\}\}$ would be partially self-healing (wrt to $E_1(\mathcal{F})$).

To end up with that property, one could notice that partial self-healability could be related to a partial repairability property. The search for a partially self-healing set of macrofaults will indeed usually be driven by partial repairability: if one knows that some faults are not repairable, it is enough to check diagnosability for the remaining.

Partial repairability means that there exists a set of macrofaults (not necessarily covering all the basic faults) that is repairable.

Definition 21 (Partial repairability of the system) $PartiallyRepairable \equiv \exists E(\mathcal{F})$ *such that* $Repairable(E(\mathcal{F}))$

Property 4 $Repairable \models PartiallyRepairable$

In other words, partial self-healability is consistent (though it is not mandatory) with a system that is only partially repairable. On the other hand, a system that is self-healing or weakly self-healing must necessarily be (fully) repairable.

Property 5 $SelfHealing \models Repairable$

Property 6 $WeaklySelfHealing \models Repairable$

Property 7 $PartiallySelfHealing \models PartiallyRepairable$

To summarize, a self-healing set of macrofaults is such that each macrofault is diagnosable and repairable, which means also, as it is a covering set, that each basic fault is repairable (but not necessarily diagnosable).

A weakly self-healing set of macrofaults is such that, for each macrofault, there exist some contexts, generally not all, in which the macrofaults are diagnosable and repairable, which means also, as it is a covering set, that there are contexts, generally not all, in which each basic fault is repairable.

A partially self-healing set of macrofaults characterises the only subset of basic faults (the ones it covers) which are always repairable (even if they are not necessarily diagnosable).

Other weakened forms of self-healability could be defined and be interesting during the design phase of a system, as for instance coupling weak and partial self-healability, i.e. the sets of macrofaults which would not be but could become weakly self-healing by adding some repair plans.

6 Discussion

The main contribution of this paper is to propose a thorough and integrated definition of the *self-healability* of a dynamic system, through the analysis of what the system needs to recover from faults, when considering together its diagnosis and repair capabilities. Interestingly enough, the system does neither need to provide diagnosability for each basic fault nor to command a set of distinct specialized repair plans, one for each basic fault, but it needs to find a set of 'sufficient' diagnosable sets that can be matched to 'sufficient' repairs. As far as we know, it is the first time that such a definition is issued. Of course that is only a first step that calls for a lot of extensions.

First, in cases for which self-healability is met, there may exist many different sets of macrofaults establishing the property. An interesting concern is then to determine the 'best' (e.g. the one that provides the most information to the user, or the one that favours repairs that put back the system in preferred states, etc), and the search strategy (and hence the algorithms) to provide that optimal set of macrofaults.

Second, as we have shown, it is not always possible to find a set of macrofaults ensuring self-healability, and one may only get what we have called partial or weak self-healability. These weakened properties are useful at design time because it means that one needs to modify either the observables or the repair plans (i.e. increase sensor or actuator performance) that are available in the system. The interesting issue is then to find the 'best' set of macrofaults to isolate, i.e. here the one that minimizes the needs for extra diagnosis/repair capabilities. As we have shown, we could look for a set that is perfectly diagnosable but not repairable, or perfectly repairable but not always diagnosable, but such sets may be far from being optimal: hence the most interesting set of macrofaults to start with may very well be a set that is neither perfectly diagnosable nor repairable.

Other issues that we would like to deal with in the future are extensions of the current framework to address more elaborated (and more realistic) cases, mainly:

- Multiple faults: if f_2 occurs some time after f_1, maybe it is better to wait and apply a repair plan that will repair both than applying two successive plans...? Then we will need to be more precise in terms of the definition of repair plans and their chracterization in terms of their capabilities to repair only one or several faults, and to be able to repair a fault even though another one has occurred, etc.
- Temporal constraints: a very important issue is to consider the 'delay' that is needed to diagnose a fault, due to waiting for a sufficient number of observations; that delay may counteract some emergency delay for an adequate repair plan, i.e. the delay after which it will be too late to apply the plan. Such delays should be incorporated into our definitions to only match repair plans to macrofaults when such delay constraints are not violated.

References

Basseville, M. 2001. On fault detectability and isolability. *European Journal of Control* 7(8):625–637.

Chen, J., and Patton, R. 1994. A re-examination of fault detectability and isolability in linear dynamic systems. *Proceedings of the 2nd Safeprocess Symposium, Helsinki (Finland).*

Cimatti, A.; Pecheur, C.; and Cavada, R. 2003. Formal verification of diagnosability via symbolic model checking. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence IJCAI'03.*

Contant, O.; Lafortune, S.; and Teneketsis, D. 2004. Diagnosis of intermittent faults. *Journal of Discrete Event Dynamic Systems: Theory and Applications* 14:171–202.

Cordier, M.-O.; Travé-Massuyès, L.; and Pucel, X. 2006. Comparing diagnosability in continous and discrete-event systems. In González, C.; Escobet, T.; and Pulido, B., eds., *17th International Workshop on Principles of Diagnosis*, 55–60.

Garcia, G.; Bernussou, J.; and Arzelier, D. 1994. Robust stabilization of discrete-time linear systems with norm-bounded time-varying uncertainty. *Systems & Control Letters archive, Volume 22 , Issue 5* 327–339.

Ghosh, D.; Sharman, R.; Rao, H. R.; and Upadhyaya, S. 2007. Self-healing systems survey and synthesis. *Decision Support Systems* 42(4):2164–2185.

Jéron, T.; Marchand, H.; Pinchinat, S.; and Cordier, M.-O. 2006. Supervision patterns in discrete event systems diagnosis. In *Workshop on Discrete Event Systems, WODES'06.*

Jiang, S.; Huang, Z.; Chandra, V.; and Kumar, R. 2001. A polynomial time algorithm for diagnosability of discrete event systems. *IEEE Transactions on Automatic Control* 46(8):1318–1321.

Pencolé, Y. 2005. Assistance for the design of a diagnosable component-based system. In *17th International Conference on Tools with Artificial Intelligence (ICTAI 05)*, 549–556. IEE Computer Society.

Ramadge, P., and Wonham, W. 1989. The control of discrete event systems. *Proceedings of the IEEE* 81–98.

Saboori, A., and Zad, S. H. 2005. Fault recovery in discrete event systems. In *Proceedings of the ICSC congress on Computational Intelligence Methods and Applications (CIMA'05).*

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

Schumann, A., and Pencolé, Y. 2007. Scalable diagnosability checking of event-driven system. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI07)*, 575–580.

Shaw, M. 2002. Self-healing: Softening precision to avoid brittleness. In *Proceedings of the First Workshop on Self-Healing Systems WOSS '02, Charleston, South Carolina, USA, November 18-19.*, 111–113. ACM.

Tornil, S.; Escobet, T.; and Travé-Massuyès, L. 2003. Robust fault detection using interval models. In *12th European Control Conference ECC03, Cambridge, UK.*

Travé-Massuyès, L.; Escobet, T.; and Milne, R. 2001. Model-based diagnosability and sensor placement application to a frame 6 gas turbine subsystem. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI'01*, volume 1, 551–556.

Travé-Massuyès, L.; Escobet, T.; and Olive, X. 2006. Diagnosability analysis based on component supported analytical redundancy relations. *IEEE Transactions on Systems, Man and Cybernetics, Part A : Systems and Humans, Vol. 36, N6.*

Washington, R.; Golden, K.; and Bresina, J. 1999. Plan execution, monitoring, and adaptation for planetary rovers. In *Proceedings of the IJCAI'99 Workshop on 'Scheduling and Planning meet Real-time Monitoring in a Dynamic and Uncertain World'.*

Yoo, T., and Lafortune, S. 2002. Polynomial-time verification of diagnosability of partially-observed discrete-event systems. *IEEE Transactions of Automatic Control* 47(9):1491–1495.