Formulating Event-Based Critical Observations in Diagnostic Problems

Cody James Christopher[1,2] and Alban Grastien[2,1]

[1]Artificial Intelligence Group, The Australian National University.
[2]Optimisation Research Group, NICTA*

Abstract

We claim that in scenarios involving a human operator with responsibility over systems being monitored by diagnoser, presenting said operator with a concise set of observations capturing the essence of a failure improves the operator's understanding of the diagnosis.

We take this in the context of Discrete Event Systems and demonstrate how the idea can be applied to systems utilising event-based observations, which can contain implicit information. We introduce the notion of an abstracted event stream, called a sub-observation, that makes the implicit information explicit for the operator and allows a diagnoser to arrive at the same diagnosis. We call the most abstract of these the critical observation. We provide relevant definitions, properties, and a procedure for computing the critical observation in a diagnosis problem.

1 Introduction

Diagnosis problems are concerned with the detection and identification of occurrences of specific events in a system, generally called faults or failures. These occurrences are difficult to detect as the fault events are typically not directly observable, however, they can be inferred from the system model (a description of the system behaviour) and the observations produced by the system.

Diagnosis is the first step in the fault recovery process. Once a fault has been detected and identified, the appropriate actions can be taken to mitigate its effects. The issue, however, is that this procedure acts as a black box; given a model and a sequence of observations, a diagnoser asserts a fault by claiming that there is no possible nominal execution of the system that would produce the observation sequence.

The present work is written under the assumption that a diagnosis procedure is fundamentally built for a human operator in charge of taking actions after a fault is identified. In this scenario, a black box approach does not allow for the presentation of the information relevant to the diagnosis. We assume that providing the operators with explanatory evidence is useful in convincing them of the validity of the diagnosis, in addition to providing information as to the causes of the fault.

Further, we assume that a more concise explanation is strictly preferred to more verbose explanation, and consequently that there is merit to isolating the "smallest" amount of supporting evidence, or what we call the *critical observations*. In cognitive psychology, the seminal paper on the topic of working memory in humans supports this view, giving the average working memory capacity as 7 ± 2 distinct pieces of information [1]. Providing only the observations critical to the diagnosis also has the additional benefit of ameliorating privacy concerns in systems where privacy is considered important.

We extend the results of Christopher et al. [2] to event-based observations. We first present preliminary theory and notation, before going on to show that event-based observations contain implicit information. We then introduce what we call *sub-observations* that can capture this implicit information and make it available for use in diagnosis procedures. We then provide formal definitions of *sufficiency* and *criticality* in addition to several important properties that allow for a terminating algorithm. We present an algorithm for computing the critical observation and discuss its complexity. A discussion of alternate ways of defining sub-observations precedes a brief discussion of related work and a conclusion.

2 Preliminaries and Notations

The present work takes place in the context and standard framework of discrete event systems (DES) [3]. We denote as Σ the set of events that can take place on the system. A system *run* is a finite sequence of events, $w = e_1 e_2 \dots e_k$, and the system is modeled as the prefix-closed language $\mathcal{L}_{\mathrm{M}} \subseteq \Sigma^\star$ that represents all possible runs.

The set of events is partitioned into *observable* events Σ_{o}—events that are recorded—and unobservable events Σ_{u}—those that are not. The observation o generated by run $w = e_1 e_2 \dots e_k$, hereafter called the *trace* of w, is the projection of w on the set of observable events (i.e., all unobservable events of the run are deleted):

$$o = P_{\Sigma_{\mathrm{o}}}(w) = \begin{cases} \varepsilon & \text{if } k = 0 \\ e_1 P_{\Sigma_{\mathrm{o}}}(e_2 \dots e_k) & \text{if } k > 0 \text{ and } e_1 \in \Sigma_{\mathrm{o}} \\ P_{\Sigma_{\mathrm{o}}}(e_2 \dots e_k) & \text{otherwise.} \end{cases}$$

The observed language of a trace o, denoted $\mathcal{L}_o$, is the set of finite sequences of events that could produce the observed sequence: $\mathcal{L}_o = P^{-1}_{\Sigma_{\mathrm{o}}}(o) = \{w \in \Sigma^\star \mid P_{\Sigma_{\mathrm{o}}}(w) = o\}$.

*NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

The set of unobservable events includes a subset of fault events, $\Sigma_\mathrm{f} \subseteq \Sigma_\mathrm{u}$. With slight abuse of notation we write $f \in w$ as short for $w \in \Sigma^\star f \Sigma^\star$ (or "f appears in w") and $F \cap w$ as short for $\{f \in F \mid f \in w\}$ (or "the subset of events from F that appear in w").

A set $\delta \subseteq \Sigma_\mathrm{f}$ of faults is *consistent* with the model $\mathcal{L}_\mathrm{M}$ and the trace o if there exists a run $w \in \mathcal{L}_\mathrm{M}$ that would produce this trace ($P_{\Sigma_\mathrm{o}}(w) = o$) and that exhibits exactly these faults ($w \cap \Sigma_\mathrm{f} = \delta$). The diagnosis of trace o, denoted $\Delta(o)$, is the collection of all consistent sets of faults:

$$\Delta(o) = \left\{ \delta \subseteq \Sigma_\mathrm{f} \;\middle|\; \begin{array}{l} \exists w \in \mathcal{L}_\mathrm{M}. \\ P_{\Sigma_\mathrm{o}}(w) = o \;\wedge\; \delta = w \cap \Sigma_\mathrm{f} \end{array} \right\} \quad (1)$$

Hereafter we use the hat notation ($\hat{}$) to indicate that the given symbol represents what actually occurred. Given a run $\hat{w}$, $\hat{\delta} = \hat{w} \cap \Sigma_\mathrm{f}$ is the set of faults that occurred during the run; then the following result is trivial: $\hat{w} \in \mathcal{L}_\mathrm{M} \Rightarrow \hat{\delta} \in \Delta(P_{\Sigma_\mathrm{o}}(\hat{w}))$. (The premise, completeness of the model, is assumed.)

We find it more convenient to define the diagnosis in terms of emptiness of languages. Let $\mathcal{L}_\delta$ be the language that represents all sequences that contain exactly δ:

$$\mathcal{L}_\delta = \{w \in \Sigma^\star \mid w \cap \Sigma_\mathrm{f} = \delta\} = \bigcap_{f \in \delta} \Sigma^\star f \Sigma^\star \cap \bigcap_{f \in \Sigma_\mathrm{f} \setminus \delta} (\Sigma \setminus \{f\})^\star$$

That is, $\mathcal{L}_\delta$ represents the set of all runs containing all of the faults of δ, intersected with all possible runs where the faults not in δ never occur—the result is a set of all runs where the only faults that occur are those in δ. With $\mathcal{L}_\delta$ defined, we can equivalently express the diagnosis as an emptiness of languages problem:

$$\delta \in \Delta(o) \iff \mathcal{L}_\mathrm{M} \cap \mathcal{L}_o \cap \mathcal{L}_\delta \neq \emptyset. \quad (2)$$

3 Sub-Observations

We first discuss event-based observations, and in particular that event-based observations contain implicit information that must be taken into consideration when performing diagnosis. We then introduce the notion of *sub-observations*, providing formal definitions and an explanatory example. Once this has been established, a procedure is given for diagnosing with sub-observations.

3.1 Event-Based Diagnosis and Implicit Information

Event-based diagnosis, contrasted with state-based diagnosis, comes with a subtlety; specifically, there is a type of implicit information encoded in the trace. Take for example the repeated observation of a window being closed without there ever being an observation of the window opening; in this case, the fact that we never observed an open event is distinctly relevant to a diagnosis procedure.

To further illustrate this, we provide a simple abstract example in the form of a DES: Take $\Sigma = \{a, b, c, d, e, f_1, f_2\}$, with $\Sigma_\mathrm{o} = \{a, b, c, d, e\}$, $\Sigma_\mathrm{u} = \Sigma_\mathrm{f} = \{f_1, f_2\}$. We provide the system model in the form of a NFA in Figure 1 and consider some example traces over it:

$o_1 = ababab c$. The model specifies that f_2 must have occurred in strings containing a followed by c. In this case, the intervening sequence is long ($babab$), and could be much longer. The important information, however, is that a was at some point followed by c. Reporting in some abstract sense that a was followed by c is enough to convince an operator of the correctness of the diagnosis.

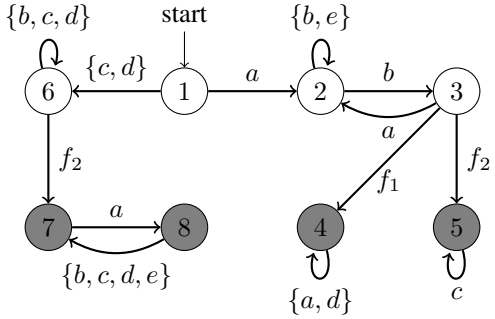

Figure 1: Example DES

$o_2 = ababaa$. The model specifies that f_1 must have occurred for there to be two a events that are not separated by another observable event. More specifically, the *lack* of an intervening event is the crucial piece of information that determines the fault. In this case, reporting in some abstract sense that multiple a occurred consecutively is enough to indicate the fault convincingly.

3.2 Framework

We first present a general framework for sub-observation, which is then further specified for our particular choice of implementation.

General Definition

Definition 1 *We define a framework for sub-observations as a tuple: $\langle \mathbb{O}, \preceq, sub \rangle$:*

1. *A* sub-observation*, θ, is an abstraction over a trace that represented an intentional relaxation (or weakening) of the concrete knowledge contained in the trace.*
2. *$\mathbb{O}$ is the space of possible sub-observations.*
3. *The symbol $\preceq$ is a binary relation and partial order over $\mathbb{O}$ and relates two sub-observations θ, θ' such that $\theta' \preceq \theta$ iff θ' is a more abstracted form of θ.*
4. *sub is an injective function, mapping traces to maximal (w.r.t. $\preceq$) sub-observations $\theta \in \mathbb{O}$:*

$$sub : \Sigma_\mathrm{o}^* \to \mathbb{O}$$

A sub-observation θ implicitly represents the set of traces for which it is a more abstract form of:

$$\psi(\theta) = \{o \in \Sigma_\mathrm{o}^* \mid \theta \preceq sub(o)\}$$

Therefore, $\theta' \preceq \theta \Rightarrow \psi(\theta') \supseteq \psi(\theta)$.

The language of a sub-observation, denoted $\mathcal{L}_\theta$, represents the set of all possible runs θ could represent. However, these runs are already captured by $\mathcal{L}_o$, and so $\mathcal{L}_\theta$ can be expressed as the union of the languages of the traces it is a more abstract form of:

$$\mathcal{L}_\theta = \bigcup_{o \in \psi(\theta)} \mathcal{L}_o \quad (3)$$

Specific Definition

For the purposes of our specific definition of sub-observations, it is necessary to distinguish between what we call *hard* and *soft* events. A *hard event* is a singleton observable event, $x \in \Sigma_\mathrm{o}$, and represents the firm occurrence of an

event in the system. A *soft event* is a subset of observable events, $y \subseteq \Sigma_o$, that any number (including zero) of which may have occurred along with any number of unobservable events.

We now explicitly characterize our construction of sub-observations based on the general framework presented in Definition 1:

Definition 2 *A sub-observation, θ, is a strict time-ordered alternating sequence of soft and hard events, commencing and ending with a soft event: $\theta = y_0 x_1 y_1 \dots x_n y_n$. We denote $\mathbb{O}(o)$ the space of sub-observations for a given trace o. $\theta \in \mathbb{O}$ has length $|\theta| = n$. For readability, sub-observations may occasionally be written as a comma separated list. The language of θ can then also be expressed:*

$$\mathcal{L}_\theta = (y_0 \cup \Sigma_u)^* x_1 (y_1 \cup \Sigma_u)^* \dots x_n (y_n \cup \Sigma_u)^*$$

By way of example, take the sub-observation $\theta = (\{b,d\}, a, \emptyset, c, \{a\})$ – in this case, we say the singleton events $x_1 = a$ and $x_2 = c$ are *hard* and occurred in the specified order. The first soft event, $y_0 = \{b,d\}$, represents the possibility of any number of b or d events in any order having occurred before the first hard event – similarly, $y_1 = \emptyset$ indicates that no events occurred between the hard events x_1 and x_2, and $y_2 = \{a\}$ that any number of a events could have occurred after the final hard event. There are multiple traces $\hat{o}$ that this could represent, ac being the simplest, but traces such as $ddacaa$ or bac, or indeed up to infinite (or bounded length depending) other possibilities.

Definition 3 *The function sub generates a sub-observation in $\mathbb{O}$ from a given trace by inserting empty soft events at the head of the trace, and after every hard event:*

$$\begin{aligned} &\text{For } o = e_1 \dots e_n \\ &sub(o) = \emptyset x_1 \emptyset \dots x_n \emptyset \in \mathbb{O} \\ &\text{Where } \forall i : x_i = e_i \end{aligned}$$

Definition 4 *The relation $\preceq$ over $\mathbb{O}$ is defined such that $\theta' \preceq \theta$ if and only if there exists a mapping function f:*

$$\begin{aligned} &\text{Given } |\theta'| = n, |\theta| = m \\ &f : \{0, \dots, n+1\} \to \{0, \dots, m+1\} \text{ such that} \\ &f(i) < f(i+1),\ f(0) = 0,\ f(n+1) = m+1 \\ &x'_i = x_{f(i)} \\ &y'_i \supseteq \bigcup_{f(i) \le j \le f(i+1)-1} y_j \cup \bigcup_{f(i) < j < f(i+1)} x_j \end{aligned}$$

The relation $\preceq$ is provably a partial order.

In words: $\theta' \preceq \theta$ if there exists some f that maps the hard events in θ' to an equivalent sequence in θ, retaining the time-ordering of both, and each y'_i in θ' captures the union of all intervening events – y_j (inclusive) and x_j (exclusive), for j ranging between $f(i)$ and $f(i+1) - 1$.
For example take $\theta = (\{ac\}, b, \{cd\}, a, \{c\}, d, \{c\}, a, \emptyset)$ and $\theta' = (\{abcd\}, a, \{bcd\}, a, \emptyset)$. The hard events in θ' are matched to x_2 and x_4 in θ, and each y'_i "swallows" the other information. Specifically, $f(1) = 2, f(2) = 4$, satisfies the constraints for $\theta' \preceq \theta$. This is illustrated in Figure 2.

To summarize, a sub-observation, in a practical sense, can be thought of as a relaxation of the information presented in the original trace. By including soft events in the sub-observation, we are allowing for the "hiding" (abstraction) of events such that an operator can be presented with only the most relevant information.

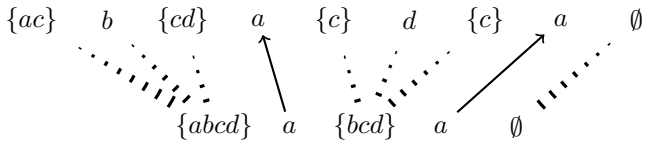

Figure 2: An example map satisfying $\preceq$

3.3 Diagnosis of Sub-Observations

We now formalize the usage of sub-observations in a diagnosis procedure by extending the procedure introduced for event-based diagnosis presented in §2. This involves checking the consistency of a set of possible faults.

We therefore provide the construction of the diagnosis of θ, $\Delta(\theta)$, the set of faults consistent with a given sub-observation:

Definition 5 *The diagnoses of a sub-observation θ is the union of the diagnoses of the traces for which θ is the more abstract form of, represented by $\psi(\theta)$ as given in Definition 1:*

$$\Delta(\theta) = \bigcup_{o \in \psi(\theta)} \Delta(o)$$

From Definition 5 we note that, given $\hat{\delta} \in \Delta(\hat{o})$, that if $\theta \preceq sub(\hat{o})$ then $\hat{\delta} \in \Delta(\theta)$. That is, the actual diagnosis $\hat{\delta}$ of the actual trace $\hat{o}$, will by definition be in $\Delta(\theta)$ if θ is an abstraction of $\hat{o}$.

First, we observe the following lemma:

Lemma 3.1 *The possible traces permitted by the language of a more abstracted sub-observation strictly contains all the permitted traces of all its ascendants:*

$$\theta' \preceq \theta \Longrightarrow \mathcal{L}_\theta \subseteq \mathcal{L}'_\theta$$

Proof This is a direct consequence of Equation 3

Equation 2 provided a formulation of the diagnosis as a question of emptiness in the intersection of languages – that is, is there some run that is simultaneously possible according to the system model, the observations, and the faults that occurred during the run. This can similarly be extended to a similar question for sub-observations. As $\mathcal{L}_\theta$ is defined in Definition 2, then $\Delta(\theta)$ can be equivalently extended:

$$\Delta(\theta) \equiv \{\delta \mid \mathcal{L}_M \cap \mathcal{L}_\delta \cap \mathcal{L}_\theta \neq \emptyset\} \tag{4}$$

Definition 5 and Equation 4 provide a formal definition and a characterization of the diagnosis of a sub-observation, but do not specify how to implement the procedure, in particular given that $\psi(\theta)$ may be infinitely large.

The scientific literature is rich in works dealing with abstract traces. These approaches were developed to handle situations where observations can be lost [4]; sensors can fail [5; 4]; the order between observations may be only partially known [6; 4; 7; 8]; the observability can vary [9]; etc.

It is possible to interpret Equation 4 quite literally—build three finite-state machines representing all three languages $\mathcal{L}_M$, $\mathcal{L}_\delta$, and $\mathcal{L}_\theta$, synchronize them, and verify emptiness. Similarly, this emptiness verification can be reduced to a planning problem [10; 11] or a model-checking one [12].

When the model is represented by a finite-state machine, the specific definition of sub-observations makes it possible to solve the problem by tracking the belief state (the set of states that the system could be in) after each soft and hard

event in the sub-observation. Assuming the system state incorporates the diagnosis information, then the diagnosis can be inferred from the belief state at the end of the sub-observation. This procedure can be used on-line [13] or pre-processed in a fashion akin to the diagnoser [14].

4 Critical Observations

The primary objective of this work is to compute a minimal sub-observation that preserves the assertiveness of the diagnosis: a critical observation. We first give a formal definition of this notion, followed by a discussion of some relevant properties and a procedure for computing the critical observation.

4.1 Definition of a Critical Observation

We say a sub-observation is sufficiently precise if it allows us to infer a given diagnosis:

Definition 6 *Given a diagnosis D, a sub-observation θ is* sufficient *to prove D if $\Delta(\theta) = D$. Given a trace $\hat{o}$, a sub-observation $\theta \preceq sub(\hat{o})$ is* sufficient *for $\hat{o}$ if $\Delta(\theta) = \Delta(\hat{o})$.*

A corollary of Definition 5 gives us $\Delta(\theta) \supseteq \Delta(\hat{o})$. As previously noted, abstracting away details to produce a sub-observation sacrifices some information about the system behavior—sufficiency, then, is the property that this information loss did not affect the diagnosis by making feasible other potential diagnoses:

$$\Delta(\theta) \setminus \Delta(\hat{o}) = \emptyset \tag{5}$$

Our goal, then, is to return a sub-observation that is sufficient for the actual trace, $\hat{o}$. By Definitions 1, and 5, we see that the naïve sub-observation, $sub(\hat{o})$, satisfies the criteria to be sufficient for $\hat{o}$, and means that at least one solution can be found:

$$\Delta(sub(\hat{o})) = \bigcup_{o \in \psi(sub(\hat{o}))} \Delta(o) = \Delta(\hat{o})$$

Given two sub-observations θ and θ', a human operator will, from our initial assumptions, better understand and assimilate a diagnosis with θ' if θ' is more abstract than θ. We therefore search for a "most abstract", or *critical*, sub-observation, defined as follows:

Definition 7 *Given a trace $\hat{o}$, a sub-observation $\theta \preceq sub(\hat{o})$ is* critical *for $\hat{o}$ if it is sufficient for $\hat{o}$ and there is no strict sub-observation of θ that is also sufficient:*

$$\forall \theta' \in \mathbb{O}.\ (\theta' \preceq \theta) \wedge (\Delta(\theta') = \Delta(\hat{o})) \Rightarrow (\theta' = \theta)\,. \tag{6}$$

A critical sub-observation (more simply called a critical observation) is therefore a sufficient sub-observation that cannot be abstracted more without damaging (complicating) the precision of the diagnosis.

As $\preceq$ is only a partial order, it is possible that there could be several critical sub-observations. For instance, using the example in Figure 1, both $\theta_1 = \Sigma_o c \Sigma_o a \Sigma_o$ (the system emitted a c and later an a) and $\theta_2 = (\Sigma_o \setminus \{a\}) d \Sigma_o a \Sigma_o$ (the system emitted anything bar an a, then a d and later an a) are critical observations for the trace $\hat{o} = cda$.

4.2 Computing the Critical Observation

We now outline a procedure for computing a critical observation for a given problem. We rely on two fundamental properties: the finiteness of the set of sub-observations of interest, and the monotonicity of sufficiency.

Lemma 4.1 (Finiteness) *Given a trace $\hat{o}$, the set $\mathbb{O}(\hat{o})$ of sub-observations of $\hat{o}$ ($\{\theta \in \mathbb{O} \mid \theta \preceq sub(\hat{o})\}$) is finite.*

Proof This can be demonstrated by the fact that, by definition of $\preceq$, the length of a sub-observation of $\hat{o}$ must be equal to or smaller than that of $\hat{o}$. This can only decrease until $|\theta| = 1$, at which point the set is exhausted.

Lemma 4.2 (Monotonicity) *Given a trace $\hat{o}$ and two sub-observations θ_1, θ_2 such that $\theta_1 \preceq \theta_2 \preceq sub(\hat{o})$, if θ_1 is sufficient for $\hat{o}$, then so is θ_2.*

Proof This is a straightforward consequence of the fact that $\psi(\theta_1) \supseteq \psi(\theta_2)$.

Monotonicity guarantees that there is no unreachable "island" of sufficient sub-observations.

Finiteness provides us three decisive properties: One—that there always exists at least one critical observation (infinite domains can prevent the existence of minimal elements; e.g., there is no minimal real number strictly greater than 0), Two—that for any sufficient sub-observation θ, there exists a critical observation that is a sub-observation of θ (possibly θ itself), Three—the *depth* of a critical observation (the maximal number k of different sub-observations θ_i such that $\theta \preceq \theta_1 \preceq \cdots \preceq \theta_k \preceq sub(\hat{o})$) is finite.

As a consequence of these properties, as soon as a sufficient sub-observation θ is found the search for a critical observation can be limited to the set of sub-observations of θ (we call this a *greedy* approach). Another consequence of the above is that we can define a search algorithm that can find a sufficient, strict sub-observation of a given sub-observation (or return that no such sub-observation exists), that is guaranteed to terminate.

Finally, monotonicity together with finiteness, provides a practical characterization of criticality: a sufficient sub-observation θ is critical if and only if none of its children (defined next) are sufficient.

Definition 8 *A* child *of sub-observation θ is a strict sub-observation θ' of θ such that no sub-observation sits "between" θ' and θ.*

$$\theta' \in children(\theta) \iff (\theta' \prec \theta) \wedge (\nexists \theta'' \in \mathbb{O}.\ \theta' \prec \theta'' \prec \theta)\,.$$

If, on the other hand, we find that one child of θ is sufficient, then, according to the greedy approach described previously, we can iteratively check criticality of this child.

The set of children for our definition of sub-observation is readily computable. We can prove that the children of a sub-observation are exactly the sub-observations obtained by applying one of two operations which we will now define: the event-softening operation and the collapse operation.

Definition 9 *Given a sub-observation $\theta = y_0 x_1 \ldots x_k y_k$, the* event-softening *operation $\theta' = es(\theta, i, e)$ adds event e to the ith soft event of the sub-observation: $es(\theta, i, e) = y'_0 x'_1 \ldots x'_k y'_k$ (defined if $e \notin y_i$) such that*

- $\forall j \in \{1, \ldots, k\}.\ x'_j = x_j$,
- $\forall j \in \{0, \ldots, k\} \setminus \{i\}.\ y'_j = y_j$, *and*
- $y'_i = y_i \cup \{e\}$.

Procedure FINDCRITICALOBSERVATION
input: trace $\hat{o}$; *output*: critical observation
diag := $\Delta(\hat{o})$
$\theta := sub(\hat{o})$
candidates := *children*(θ)
while *candidates* $\neq \emptyset$ **do**
  $\theta' :=$ pop(*candidates*)
  if $\Delta(\theta') =$ *diag* **then**
    $\theta := \theta'$
    candidates := *children*(θ)
  end if
end while
return θ

Figure 3: Finding a critical observation

Definition 10 *Given a sub-observation $\theta = y_0 x_1 \dots x_k y_k$, the* collapse *operation $\theta' = coll(\theta, i)$ "forgets" the concrete occurrence of a hard event x_i. This operation requires the soft events before and after x_i to be equal and to allow for x_i: $coll(\theta, i) = y'_0 x'_1 y'_1 \dots x'_{k-1} y'_{k-1}$ (defined if $x_i \in y_i$ and $y_{i-1} = y_i$) such that*

- $\forall j \in \{1, \dots, i-1\}.\ x'_j = x_j$ *and* $y'_{j-1} = y_{j-1}$ *and*
- $\forall j \in \{i+1, \dots, k\}.\ x'_{j-1} = x_j$ *and* $y'_{j-1} = y_j$.
- $y'_{i-1} = y_{i-1} = y_i$

Lemma 4.3 *The children of a sub-observation θ are exactly all the sub-observations that can be obtained by applying either event-softening or collapse to θ.*
(See appendix for proof).

An algorithm for finding a critical observation is given in Figure 3. Starting from $\theta = sub(\hat{o})$, the algorithm verifies whether any child of θ is sufficient. If this is the case, then θ is replaced with this child and the verification continues iteratively.

Theorem 4.4 *Algorithm* FINDCRITICALOBSERVATION *always terminates and returns a critical observation.*

This theorem is a direct consequence of the properties derived from the finiteness of $\mathbb{O}(sub(\hat{o}))$ and the monotonicity of the property, as described before.

4.3 Complexity

We now discuss the difficulty of finding a critical observation, defined in term of the number of $\Delta(\cdot)$ calls. Let $n = |\hat{o}|$ be the length of $\hat{o}$ (the number of observed events) and let $m = |\Sigma_o|$ be the number of observable events.

The maximal depth, D, of a sub-observation, namely that of $\theta_0 = \{\Sigma_o\}$, is provably $D = (n+1)\,m+n$: It is reached by softening m times each of the $n+1$ soft events followed by collapsing the n hard events. Furthermore each sub-observation (of length $k \leq n$), can be shown to have a bounded number of children, C, as given by Definition 8: At worst each soft event can be softened in any one of m ways ($(k+1)m$), and a potentially up to k hard events can be collapsed, giving $C = (k+1)\,m + k$, which we see is the same as D.

Consequently, the maximum number of $\Delta(\cdot)$ calls of Algorithm FINDCRITICALOBSERVATION is bounded by $D \times C$, and therefore in $O(n^2 m^2)$. It can even be shown that, for some traces, a naïve implementation may indeed call the diagnoser a number of times in $\Theta(n^2 m^2)$ with different sub-observations every time (see appendix for proof).

Fortunately it is possible to reduce this number drastically with heuristics. Indeed it is generally possible to prove that some children of θ' are not sufficient simply because some children of the parent of θ' were proven not sufficient, thus pruning the search tree significantly.

Consider for instance the sub-observation $\theta = \emptyset a \emptyset b \emptyset a \emptyset a \emptyset$ in the example of Figure 1 (with diagnosis: fault f_1). The softening by b of the soft event $y_3 = \emptyset$ between $x_3 = x_4 = a$ leads to a sub-observation ($\emptyset a \emptyset b \emptyset a \{b\} a \emptyset$) that is not sufficient, as the nominal diagnosis N becomes possible. Consider now the sub-observation $\theta' = \Sigma_o a \emptyset a \Sigma_o$ of θ. We can deduce automatically that the softening of $y'_1 = \emptyset$ by b in θ' leads to a non sufficient sub-observation, simply because the mapping function f of Definition 4 associates y'_1 with y_3.

It is therefore possible to "carry over" to the children of any sub-observation the information regarding which softening and collapse operations complicate the diagnosis and reduce precision. By doing so, the number of necessary calls provably drops to $\Theta(nm)$.

5 Other Definitions of Sub-Observations

In this article we presented one definition of sub-observation that, by no means, is the only viable one. We briefly discuss a few possible variants and then present the necessary elements that the reader would need to consider to use another definition.

In many circumstances the order between certain observed facts is irrelevant. In the example of Figure 1, the occurrence of both c and a, in any order, is symptomatic of fault f_2. Reminiscent of chronicles [15], a sub-observation could be a directed graph of hard events where a directed path between two hard events expresses a temporal precedence. This bears a similarity to temporal uncertainty in observations as described by Zanella and Lamperti [4].

The hard events are currently defined as a single specific observable event; one could alter the definition to allow for it could be replaced by a set of events. Indeed in the example of Figure 1, a fault can be diagnosed when observing either c or d before e. A reason for not distinguishing c from d in this specific scenario is that these events could represent the same message emitted by different components, or different messages emitted by the same component: the exact emitter of the message or the exact content may be irrelevant to diagnose the fault. This is similar to logical uncertainty in observations [4].

One more elaborate abstraction could be to use first-order representations. For instance, a fault may be identified by demonstrating that some user who was to be explicitly refused access to some data was actually given access to that data; the identity of the actual user may be irrelevant.

Defining New Sub-Observations

To apply the theory presented in this paper to a different definition of sub-observations, one needs to define the sub-observation space as given in Definition 1, i.e., the set of sub-observations $\mathbb{O}$, the partial order relation $\preceq$, and an inductive sub function that associates each observation with an equivalent maximal sub-observation in $\mathbb{O}$. This also needs to be additionally equipped with a procedure to compute $\Delta(\theta)$. Algorithm FINDCRITICALOBSERVATION

is guaranteed to return a critical observation if the sub-observation space is finite and if the *children* function exists and is specified.

We demonstrate a scenario where these conditions may not be satisfied: Assume that the set of observable events is infinite with each observable event associated with a rational number (modeling some continuous property, e.g., temperature). A natural abstraction would replace each event by a closed interval where the value associated with the event lies (the wider the interval, the most abstract the observation). There could, however, be no maximal interval in a situation where the relevant information about the observation is that the temperature measure is strictly positive. Furthermore, there is no notion of child in this particular sub-observation space since $\mathbb{Q}$ is a dense set. Special attention must therefore be taken when defining new types of sub-observations.

6 Related Work

Finding critical observations is a different issue from optimizing sensor placement [16] and dynamic observers [9]. These two problems aim at reducing the cost of monitoring a system (by reducing the number of sensors or switching them off). This reduction, however, needs to be conservative because the decision is made before any observations are available. Critical observations, on the other hand, can be computed after all observations are available.

Consider again the trace $o = abaa$ in the example of Figure 1 whose critical observation is $\theta = \Sigma_o a \emptyset a \Sigma_o$. Consider the question of whether the first observable event of the trace is a c. The sub-observation θ does not provide this information since it is not necessary to infer the diagnosis. A dynamic observer however, has to check this information because it is necessary to dismiss fault f_2.

There has also been work on abstraction of event-based observations, as mentioned at the end of section 3. The subsumption ($\preceq$) between uncertain or partial observations has been studied by Lamperti et al. [17], although their motivation is different from ours: by identifying that the current uncertain observation θ is a refinement of a previous observation $\theta' \preceq \theta$, it is possible to reuse the diagnosis of θ' (that is, $\Delta(\theta) \subseteq \Delta(\theta')$).

7 Conclusion & Future Work

In this work we defined a notion of critical observations for the diagnosis of discrete event systems. A critical observation is a maximally abstracted observation that allows only the same diagnosis to be inferred as was from the complete observation. Critical observations are beneficial in that they contain the core proof that supports the diagnosis. An important assumption of this work is that more abstract observations are easier for a human operator to understand and act on; an important extension will be to minimize the amount of information from the model—and not only from the observations—necessary to infer the diagnosis.

We also want to be able to handle incremental and online diagnosis. Currently we assume that the critical observation is extracted once the diagnosis has been performed; however observations that are not critical for a given trace might become critical when more observations are produced by the system. We would like to identify as early as possible what abstraction of the currently received observations can be safely made without impairing the future diagnosis. Kurien and Nayak tried to address a similar problem [18] of removing intermediate (state-based) observations that do not provide additional information.

Critical observations are also good at reducing the amount of information disclosed about the system behaviour. In future work we want to explore this line of research and, in particular, examine the problem of finding sub-observations that satisfy a privacy criterion, for instance, one defined by opacity [19].

References

[1] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological review*, vol. 63, no. 2, p. 81, 1956.

[2] C. Christopher, M.-O. Cordier, and A. Grastien, "Critical observations in a diagnostic problem," in *IEEE Conference on Decision and Control*, 2014, pp. 382–387.

[3] C. Cassandras and S. Lafortune, *Introduction to discrete event systems*. Kluwer Academic Publishers, 1999.

[4] M. Zanella and G. Lamperti, *Diagnosis of active systems*. Kluwer Academic Publishers, 2003.

[5] L. Carvalho, M. Moreira, J. Basilio, and S. Lafortune, "Robust diagnosis of discrete-event systems against permanent loss of observations," *Automatica*, vol. 49, no. 1, pp. 223–231, 2013.

[6] R. Debouk, S. Lafortune, and D. Teneketzis, "Coordinated decentralized protocols for failure diagnosis of discrete event systems," *Journal of Discrete Event Dynamical Systems*, vol. 10, no. 1–2, pp. 33–86, 2000.

[7] Y. Pencolé and M.-O. Cordier, "A formal framework for the decentralised diagnosis of large scale discrete event systems and its application to telecommunication networks," *Artificial Intelligence (AIJ)*, vol. 164, no. 1–2, pp. 121–170, 2005.

[8] R. Su and W. Wonham, "Global and local consistencies in distributed fault diagnosis for discrete-event systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 12, pp. 1923–1935, 2005.

[9] F. Cassez and S. Tripakis, "Fault diagnosis with dynamic observers," in *International Workshop on Discrete Event Systems*, 2008, pp. 212–217.

[10] S. Sohrabi, J. Baier, and S. McIlraith, "Diagnosis as planning revisited," in *International Conference on the Principles of Knowledge Representation and Reasoning*, 2010, pp. 26–36.

[11] P. Haslum and A. Grastien, "Diagnosis as planning: two case studies," in *Scheduling and Planning Applications Workshop*, 2011, pp. 37–44.

[12] M.-O. Cordier and C. Largouët, "Using model-checking techniques for diagnosing discrete-event systems," in *International Workshop on Principles of Diagnosis*, 2001, pp. 39–46.

[13] A. Schumann, Y. Pencolé, and S. Thiébaux, "A spectrum of symbolic on-line diagnosis approaches," in *Conference on Artificial Intelligence*, 2007.

[14] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis, "Diagnosability of discrete-event systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 9, pp. 1555–1575, 1995.

[15] M.-O. Cordier and C. Dousson, "Alarm driven monitoring based on chronicles," in *IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, 2000, pp. 286–291.

[16] L. Brandán Briones, A. Lazovik, and P. Dague, "Optimal observability for diagnosability," in *International Workshop on Principles of Diagnosis*, 2008, pp. 31–38.

[17] G. Lamperti, F. Vivenzi, and M. Zanella, "On subsumption, coverage, and relaxation of temporal observations in reuse-based diagnosis of discrete-event systems: a unifying perspective," in *20th International Workshop on Principles of Diagnosis (DX-09)*, 2009, pp. 353–360.

[18] J. Kurien and P. Nayak, "Back to the future for consistency-based trajectory tracking," in *Conference on Artificial Intelligence*, 2000, pp. 370–377.

[19] F. Cassez, J. Dubreil, and H. Marchand, "Synthesis of opaque systems with static and dynamic masks," *Formal Methods in System Design*, vol. 40, no. 1, pp. 88–115, 2012.

8 Appendix

We provide proof sketches that will not be included in the final version of the paper.

Proof of Lemma 4.3

The proof is three-part:

a) proving that the event-softening operation produces only children;

b) proving that the collapse operation produces only children;

c) proving that there is no other child.

Event-Softenings It is easy to see that $\theta^2 \stackrel{def}{=} es(\theta^1, i, e) \prec \theta^1$.

Assume now that $\theta^2 \preceq \theta^3 \preceq \theta^1$ and let f_{23} and f_{31} be the two mapping functions—as presented in Definition 4—used to verify the two ordering relations.

By definition of $\preceq$, $|\theta^2| \leq |\theta^3| \leq |\theta^1|$. However since $|\theta^2| = |\theta^1|$ (by definition of event-softening), the size of all three sub-observations are equal and $f_{23} = f_{31}$ are the identity function.

As a consequence, $x_j^3 = x_j^2 = x_j^1$ for all j. Furthermore $y_j^2 \supseteq y_j^3 \supseteq y_j^1$ for all j. In particular, if $j \neq i$, since $y_j^2 = y_j^1$, then $y_j^3 = y_j^2 = y_j^1$. For i, $y_i^2 = y_i^1 \cup \{e\}$, meaning that either $y_i^3 = y_i^2$ or $y_i^3 = y_i^1$.

Therefore either $\theta^3 = \theta^2$ or $\theta^3 = \theta^1$.

Collapse Similarly, it is easy to see that $\theta^2 \stackrel{def}{=} coll(\theta^1, i) \prec \theta^1$.

Again assume that $\theta^2 \preceq \theta^3 \preceq \theta^1$ and let f_{23} and f_{31} be the functions defined as before.

The size of θ^3 now either equals that of θ^2 or θ^1; let $\ell \in \{1, 2\}$ denote the index such that $|\theta^3| = |\theta^\ell|$. Notice that either f_{23} or f_{31} is the identity function.

By definition of $\preceq$, we know that $x_j^3 = x_j^\ell$. Furthermore the set inclusions as well as the relations between y_j^2 and y_k^1 allow us to infer that $y_j^3 = y_j^\ell$ for all j.

Therefore $\theta^3 = \theta^\ell$.

No Other Children Assume now that θ' is a child of θ that cannot be obtained by event-softening or collapse. Let f be the mapping function used to verify the ordering relation.

By definition of the partial order $\preceq$, the size of θ' is smaller or equal to θ.

If $|\theta'| < |\theta|$ ("multiple collapse"), then let i be an index such that $f(i+1) > f(i)+1$ (such an index exists if the two sizes differ). If $y_{i+1} \backslash y_i \neq \emptyset$, then let $\theta'' = es(\theta, i, e)$ (where $e \in y_{i+1} \setminus y_i$) be the sub-observation obtained by softening y_i with e; then, $\theta' \prec \theta'' \prec \theta$. Similarly if $y_i \supseteq y_{i+1}$ with $\theta'' = es(\theta, i+1, e)$ (where $e \in y_i \setminus y_{i+1}$). Lastly the same applies if $y_i = y_{i+1}$ with $\theta'' = coll(\theta, i)$.

If θ and θ' have same size, then all x_i's equal the corresponding x_is, and all the y_i's are supersets of the corresponding y_is. Let i be an index such that $y_i' \neq y_i$ (if no such index exists, then $\theta' = \theta$). Let $\theta'' = es(\theta, i, e)$ where $e \in y_i' \setminus y_i$. Then $\theta' \prec \theta'' \prec \theta$.

Complexity of FINDCRITICALOBSERVATION

We show that the number of $\Delta(\cdot)$ calls in FINDCRITICALOBSERVATION could be in the order of $\frac{n^2 m^2}{4}$ where n is the length of the trace and m the number of observable events.

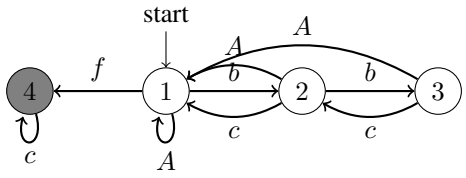

Figure 4: Example of a system: a fault is diagnosed if there are more cs than bs after the occurrence of the last a_i (A stands for $\{a_1, \ldots, a_{m-2}\}$).

We use the example of Figure 4 which involves faulty event f and observable events $\{a_1, \ldots, a_{m-2}, b, c\}$. Consider the trace of (odd) length n: $\hat{o} = \underbrace{a_1 \ldots a_1}_{n/2} \underbrace{bc \ldots bc}_{n/2} c$.

Clearly the trace reveals a faulty system since the number of cs exceeds the number of bs in this instance. The critical observation here is:

$$\Sigma_o a_1 \{c\} b \{c\} c \{c\} \ldots \{c\} b \{c\} c \{c\} c \Sigma_o,$$

i.e., all the second half of the trace needs to be kept.

We assume that FINDCRITICALOBSERVATION always tries to perform event-softening from the end of the sub-observation first, and only tries to collapse when no softening is possible. Neglecting the first steps where the c softenings are successful, the algorithm will need to make $U = \frac{n}{2} \times (m-1)$ calls to $\Delta(\cdot)$, unsuccessfully trying to softening the second half of the sub-observation. The number of successful softenings however is $S = \frac{n}{2} \times m$ (all the first half of the sub-observation), meaning that the number of $\Delta(\cdot)$ calls will be at least $U \times S = \frac{n^2 m(m-1)}{4}$ calls.