Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning
Main Track

# Diagnosis of Active Systems with Abstract Observations and Compiled Knowledge

**Gianfranco Lamperti**[1] , **Marina Zanella**[1] , **Xiangfu Zhao**[2]

[1]Department of Information Engineering, University of Brescia, Via Branze 38, Brescia 25123, Italy
[2]School of Computer and Control Engineering, Yantai University, 30, Qingquan RD, Laishan District, Yantai 264005, China
{gianfranco.lamperti,marina.zanella}@unibs.it, xiangfuzhao@gmail.com

**Abstract**

An active system (AS) is a discrete-event system (DES) with asynchronous behavior, which is represented by a network of components that are modeled as communicating automata. When being operated, an AS performs a trajectory within its behavior space, while generating a sequence of observations, namely a *temporal observation*. The model of the AS and a temporal observation are the two key ingredients of the diagnosis task, which aims to find out possible faulty behavior via abductive reasoning. Among other knowledge, such reasoning requires knowing what is observable and what is not. This essential distinction constitutes the *observability* of the AS. In the literature, the observability of a DES boils down to qualifying each state transition either as observable or unobservable, which contrasts with the way humans observe reality, typically by mapping a collection of observations to a single, abstract perception. Moreover, the occurrence of single state transitions is not necessarily what we can observe or what we want to observe for diagnosis purposes. This paper presents an extended notion of observability, where each observation is associated with a behavioral scenario rather than a single state transition, where a scenario is defined as a regular language on state transitions. To speed up the online diagnosis engine, specific diagnosis-oriented knowledge is compiled offline. Eventually, the diagnosis technique based on abstract observability is extended to cope with temporal uncertainty.

## 1 Introduction

Automated diagnosis of physical systems is still a topic of considerable research in Artificial Intelligence. A popular approach is *model-based diagnosis* (Hamscher, Console, and de Kleer 1992), which exploits the model of a system in order to find the causes of its abnormal behavior, based on some observations. Model-based diagnosis can be either *consistency-based* (Reiter 1987), initially conceived for static systems (like combinational circuits), or *abduction-based* (McIlraith 1998), like in this paper. Diagnosing a dynamical system (Struss 1997) may be facilitated by modeling it as a discrete-event system (DES). Typically, a DES (Cassandras and Lafortune 2008) can be either a Petri net (Jiroveanu, Boel, and Bordbar 2008; Cabasino, Giua, and Seatzu 2010; Basile 2014; Cong et al. 2018) or a net of communicating automata, one automaton for each component (Baroni et al. 1999; Debouk, Lafortune, and Teneketzis 2000; Pencolé and Cordier 2005; Grastien, Cordier, and Largouët 2005; Kan John and Grastien 2008; Kwong and Yonge-Mallo 2011; Grastien, Haslum, and Thiébaux 2012; Lamperti, Zanella, and Zhao 2018b), like in this paper.

Following the seminal work by Sampath et al. (1995; 1996), each state transition is qualified as either *normal* or *faulty*, even if, in principle, the component model may incorporate the normal behavior only, as proposed by Pencolé et al. (2018). For a DES, the input of the diagnosis task is a sequence of observations generated when the DES is being operated, called a *temporal observation*. The output is a set of *candidates*, where each candidate is a set of *faults*, each fault being associated with a (faulty) state transition. For several years, both notions of *abnormality* (defining what is normal and what is faulty) and *observability* (defining what is observable and what is not) have been tightly coupled to the description of the model of the DES. These notions started being separated from the DES modeling by Lamperti and Zanella (2006). Abnormality in DESs was further generalized to a *pattern* that can represent specific combinations of faults (Jéron et al. 2006; Lamperti and Zanella 2011; Lamperti and Zhao 2014).

The generalization of abnormality somewhat spurred the generalization of the notion of DES observability presented in this paper. After all, the simplistic notion of observability provided in the literature, where observations are associated with state transitions, contrasts with the way humans observe reality, typically by mapping a collection of observations to a single, abstract perception. In the new perspective adopted in this paper, each observation is associated with a behavioral scenario rather than a single state transition, where a scenario is defined as a regular language on state transitions. In a sense, observations become *abstract*, as they represent fragments of the DES behavior rather than single state transitions.

Generalized observability allows for the modeling of several real-world scenarios that are all considered by the observer. The observer can figure out the occurrence of (possibly complex and overlapping) evolutions of the DES within each single scenario as well as across the scenarios, the same way as a human being can perceive several phenomena at the same time. Abstract observability, besides resembling human perception, supports the representation of observations when sensors are adopted. Let us assume that some state transitions of a component can be detected by sensors if we

consider the component in isolation. If we assemble several components in a composite system, maybe we cannot place all the above sensors, while we can place a sensor that detects a chain of transitions involving several components. Moreover, the individual state transitions that are meaningful from a behavioral point of view are not necessarily the units that we want to observe for diagnosis purposes.

This paper focuses on the diagnosis of a class of asynchronous DESs, called *active systems* (ASs), already presented in the literature (Lamperti, Zanella, and Zhao 2018b), when these are endowed with an *abstract observability*, which is a new notion. The proposal is to process the given AS model and the abstract observation scenarios offline, that is, once and for all, so as to obtain compiled knowledge, including the *watchers* and the *diagnosis reference manual* of the AS, to be exploited online for the efficient generation of the candidates. The knowledge compilation stage is independent of the specific temporal observation, whereas the online stage, consisting in a call to a diagnosis engine, takes as input the given temporal observation.

Eventually, the technique proposed in this paper to diagnose ASs endowed with abstract observability is extended to cope with *uncertain observations*, which are characterized by a *partial* temporal order of the observed events instead of the usual *total* order. In fact, in a real world context, the total order of the observations is possibly unknown for several reasons, such as the distribution of the communication channels that convey the observations from the AS to the observer and the synchronization errors of the clocks relevant to these channels. Uncertain observations are not new (Lamperti and Zanella 2002), however here they are considered in the novel *abstract* perspective.

## 2 System Characterization

An AS is a network of components connected by links, where each component is endowed with input/output pins. A link connects an output pin of a component with an input pin of another component. Each component is modeled as a communicating automaton (Brand and Zafiropulo 1983), where a transition is triggered by an event either occurring in the external world or being ready at an input pin. The occurrence of a transition consumes the triggering event and possibly generates new events on some output pins, thereby providing triggering events to other components. A transition can be triggered only if all the links, in which the events are generated, are empty. This results in a reaction of the AS, where a series of component transitions move the AS from its initial state to a final state, where all events are consumed.

**Example 1** (Active System)**.** Outlined on the left of Fig. 1 is an AS, called $\mathcal{Z}$, which includes a transducer $z$, a breaker $b$, and a link from $z$ to $b$. The communicating automata modeling $z$ and $b$ are outlined above and below $\mathcal{Z}$, respectively, both of them including two states. The transducer is designed to detect a low-voltage external event (possibly indicating a short circuit), and to command the breaker to open by emitting an *op* event (transition $z_1$). However, the transducer may misbehave by not commanding the breaker to open ($z_3$). When the short circuit is vanished, the transducer commands the breaker to close again by emitting a $cl$ event ($z_2$). When in state 0 (closed), the breaker is designed to open when an $op$ event is ready at its input pin, in other words, when the event is within the link (transition $b_1$). Conversely, when a $cl$ event is ready, the breaker closes ($b_2$). When the breaker is required to close while being closed (state 0), it consumes the event $cl$ without changing its state ($b_4$). Still, like the transducer, the breaker may exhibit abnormal behavior by not opening when required ($b_3$).

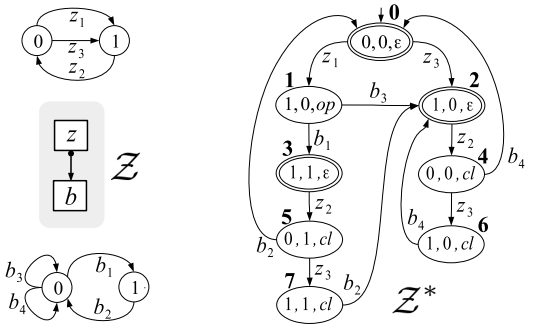


Figure 1: Active system $\mathcal{Z}$ (left), where transitions $z_3$ and $b_3$ are faulty, and relevant behavior space $Bsp(\mathcal{Z})$ (right).

The behavior of an AS is constrained by its topology and the models of its components. These constraints confine the behavior of the AS within a deterministic finite-automaton (DFA), called *behavior space*.

**Definition 1** (**Behavior Space).** *Let $\mathcal{X}$ be an AS. The* behavior space *of $\mathcal{X}$ is a DFA*

$$Bsp(\mathcal{X}) = (\Sigma, X, \tau, x_0, X_\mathrm{f}) \tag{1}$$

*where $\Sigma$ is the alphabet, comprising the set of component transitions, $X$ is the set of states $(S, E)$, where $S$ is a tuple of component states and $E$ is a tuple of (possibly empty) events that are ready at the input pins of components, $x_0 = (S_0, E_0)$ is the initial state, where all events in $E_0$ are empty, $X_\mathrm{f} \subseteq X$ is the set of final states $(S_\mathrm{f}, E_\mathrm{f})$ such that all events in $E_\mathrm{f}$ are empty, $\tau : X \times \Sigma \mapsto X$ is the transition function, where $\tau(x, t) = x'$ iff $t$ is triggerable at state $x$ and $x'$ is the state reached by the consumption of the input event and the generation of the output events relevant to $t$.*

**Definition 2** (**Trajectory).** *A sequence $T = [t_1, \ldots, t_q]$ of component transitions in the language of a behavior space $Bsp(\mathcal{X})$ is a* trajectory *of $\mathcal{X}$. A prefix of $T$ is a* semi-trajectory *of $\mathcal{X}$. Let $\mathbb{T}$ be a set of component transitions in $\mathcal{X}$. The* restriction *of $T$ on $\mathbb{T}$ is $T_{[\mathbb{T}]} = [t \mid t \in T, t \in \mathbb{T}]$.*[1]

**Example 2** (Behavior Space)**.** Shown on the right of Fig. 1 is the behavior space $Bsp(\mathcal{Z})$ of $\mathcal{Z}$ (cf. Example 1), with each state being identified by a triple $(\bar{z}, \bar{b}, \bar{e})$, where $\bar{z}$ is a state of the transducer $z$, $\bar{b}$ is a state of the breaker $b$, and $\bar{e}$ is an event ready in the link ($\varepsilon$ denotes the empty event, in other words, an empty link). States are renamed $0 \cdots 7$,

[1] Based on Definition 2, a trajectory $T$ is a string in the regular language of $Bsp(\mathcal{X})$, namely $T \in Bsp(\mathcal{X})$, which therefore ends in a final (accepting) state.

where 0 is the initial state, while the final states are 0, 2, and 3. A trajectory of $\mathcal{Z}$ is $T = [z_3, z_2, z_3, b_4, z_2, z_3, b_4]$, ending in state 2. Note how $Bsp(\mathcal{Z})$ involves abnormal transitions also, namely $z_3$ and $b_3$.

Although irrelevant to our simple example, an AS is assumed to be described as a network of communicating automata because the specification of the (whole) behavior space is in general impractical for real systems owing to the exponential explosion of the system states.

## 3 Diagnosis Setting

In order to perform the diagnosis task, the specification of an AS needs to be extended with information indicating which behavior is normal and which is abnormal. In our approach, abnormality is associated with faulty transitions.[2]

**Definition 3** (**Abnormality**). *Let* $\mathbf{T}$ *be the domain of component transitions of an AS* $\mathcal{X}$*, and let* $\mathbf{F}$ *be a domain of symbols called* faults*. The* abnormality *of* $\mathcal{X}$ *is a set of associations between component transitions and faults, namely* $Abn(\mathcal{X}) \subseteq \mathbf{T} \times \mathbf{F}$*. If* $(t, f) \in Abn(\mathcal{X})$*, then* $t$ *is* faulty*, else* $t$ *is* normal*.*

Based on the description of the abnormality of an AS, a diagnosis can be associated with each trajectory.

**Definition 4** (**Diagnosis**). *Let* $T = [t_1, \ldots, t_q]$ *be a trajectory of an AS* $\mathcal{X}$*. The* diagnosis $\delta$ *of* $T$ *is the set of faults associated with the faulty transitions in* $T$*, namely*

$$\delta(T) = \{f \mid t \in T, (t, f) \in Abn(\mathcal{X})\}. \tag{2}$$

Since a diagnosis is a set (rather than a multiset or a sequence), possible repetitions of the same fault are ignored. Ignoring the repetition of the same fault (set-oriented approach to diagnosis) prevents a diagnosis from embedding temporal information on faults (Bertoglio et al. 2020a).

**Example 3** (Abnormality)**.** For the AS $\mathcal{Z}$ introduced in Example 1, we define $Abn(\mathcal{Z}) = \{(z_3, \mathbf{z}), (b_3, \mathbf{b})\}$. In general, however, several faults may be relevant to the same component, as several transitions may be faulty for the same component. Let $T = [z_3, z_2, z_3, b_4, z_2, z_3, b_4]$ (cf. Example 2). We have $\delta(T) = \{\mathbf{z}\}$. Still, a diagnosis may involve several faults, as for $T' = [z_1, b_3, z_2, z_3, b_4]$, where $\delta(T') = \{\mathbf{b}, \mathbf{z}\}$. In particular, a diagnosis may be empty, for instance, for $T'' = [z_1, b_1, z_2, b_2]$, we have $\delta(T'') = \emptyset$.

To complete the information on the AS, for diagnosis purposes we need to specify the mode in which the behavior of the AS is observable. To this end, each observation is associated with a regular language that is defined by a regular expression on a set of component transitions.

**Definition 5** (**Observability**). *Let* $\mathbf{T}$ *be the domain of component transitions of an AS* $\mathcal{X}$*, let* $\mathbf{L}$ *be a set of regular languages on subsets of* $\mathbf{T}$*, and let* $\mathbf{O}$ *be a domain of symbols called* observations*. The* observability *of* $\mathcal{X}$ *is a relation*

$$Obs(\mathcal{X}) \subseteq 2^{\mathbf{T}} \times \mathbf{L} \times \mathbf{O}. \tag{3}$$

*where each observation in* $\mathbf{O}$ *may appear only once.*

[2]The distinction between faults and faulty transitions is grounded on the fact that, depending on which sort of information the diagnosis is expected to incorporate, in $Abn(\mathcal{X})$ different faulty transitions may be associated with the same fault.

Each element in $Obs(\mathcal{X})$ is a triple $(\mathbb{T}, \mathcal{L}, o)$, where $\mathbb{T}$ is a set of component transitions, $\mathcal{L}$ is a regular language on $\mathbb{T}$ defined by a regular expression, and $o$ is an observation. Each triple represents a behavioral scenario that can be perceived by the observer.

**Example 4** (Observability)**.** For the AS $\mathcal{Z}$ in Example 1, we define $Obs(\mathcal{Z}) = \{(\mathbb{T}, \mathcal{L}_z, z), (\mathbb{T}, \mathcal{L}_b, b), (\mathbb{T}, \mathcal{L}_a, a)\}$, where $\mathbb{T} = \{z_1, z_2, z_3, b_1, b_2, b_3, b_4\}$, $\mathcal{L}_z = (z_1 \mid z_2)$, $\mathcal{L}_b = (b_1 \mid b_2 \mid b_4)$, and $\mathcal{L}_a = (z_2\, b_4 \mid b_4\, z_2)$.[3] As such, each normal transition is observable via the same observation ($z$ for the sensor and $b$ for the breaker), while $a$ is emitted when the transitions $z_2$ and $b_4$ occur sequentially (in either order). We say that $a$ is an *abstract observation*, since it is emitted in correspondence of a specific combination of transitions.

Given a triple $(\mathbb{T}, \mathcal{L}, o) \in Obs(\mathcal{X})$ and a trajectory $T$ of $\mathcal{X}$, the observation $o$ occurs when the projection of $T$ on $\mathbb{T}$ includes a subsequence that is a string in $\mathcal{L}$. Since several observations may occur at the same time, in theory, $T$ would manifest itself as a sequence of sets of observations. However, we assume that observations in the same set are perceived as sequences, where the temporal ordering of each sequence is unpredictable.[4] In other words, a trajectory $T$ of $\mathcal{X}$ is perceived by the observer as a temporal sequence of observations, called a *temporal observation* of $\mathcal{X}$.

**Definition 6** (**Observation Space**)**.** *The* space *of a set* $O$ *of observations is the set of sequences of observations*[5]

$$O^* = \{\mathcal{O} \mid \mathcal{O} = [\, o \mid o \in O\,]\}. \tag{4}$$

*Let* $\mathbb{O} = [O_1, \ldots, O_n]$ *be a sequence of sets of observations. The* space *of* $\mathbb{O}$ *is the set of sequences of observations*

$$\mathbb{O}^* = \left\{ \mathcal{O} \mid \mathcal{O} = \bigsqcup_{i=1}^{n} [o \mid o \in O_i^*] \right\}. \tag{5}$$

*where '*$\sqcup$*' denotes the concatenation of sequences.*

**Example 5** (Observation Space)**.** For $\mathbb{O} = [\{a, b\}, \emptyset, \{c, d\}]$, $\mathbb{O}^* = \{[a, b, c, d], [b, a, c, d], [a, b, d, c], [b, a, d, c]\}$.

**Definition 7** (**Temporal Observation**)**.** *Let* $T = [t_1, \ldots, t_q]$ *be a trajectory in* $Bsp(\mathcal{X})$*. The* signature *of* $T$ *is the sequence of sets of observations*

$$\begin{aligned} Sgn(T) = [\, O_i \mid i \in [1\,..\,q], O_i = \{\, o \mid j \in [1\,..\,i], \\ T' = [t_j, \ldots, t_i], (\mathbb{T}, \mathcal{L}, o) \in Obs(\mathcal{X}), T'_{[\mathbb{T}]} \in \mathcal{L}\,\}\,]. \end{aligned} \tag{6}$$

*The space of* $Sgn(T)$ *is denoted* $Sgn^*(T)$*. A sequence* $\mathcal{O} \in Sgn^*(T)$ *is a* temporal observation *of* $\mathcal{X}$*, where* $T$ *is said to* conform *with* $\mathcal{O}$*.*

[3]A regular expression is defined inductively over an alphabet $\Sigma$ as follows. The empty symbol $\varepsilon$ and every $a \in \Sigma$ is a regular expression. If $x$ and $y$ are regular expressions, then the followings are regular expressions: $(x)$ (parentheses may be used), $x \mid y$ (alternative), $xy$ (concatenation), $x?$ (optionality), $x^*$ (repetition zero or more times), and $x^+$ (repetition one or more times).

[4]Formally, observations in the same set occurs simultaneously, but, since in reality simultaneity is difficult to detect, we make the practical assumption that they are perceived one at a time, without any constraint on their reciprocal ordering.

[5]Intuitively, $O^*$ comprises all the 'permutations' of $O$.

As such, each $O_i$ in the signature $Sgn(T)$ is the set of observations that are generated at the occurrence of the $i$-th component transition in $T$.

**Example 6** (Temporal Observation)**.** Let $T = [z_3, z_2, z_3, b_4, z_2, z_3, b_4]$ be a trajectory of $\mathcal{Z}$ (cf. Example 2). We have the signature $Sgn(T) = [\emptyset, \{z\}, \emptyset, \{b\}, \{a, z\}, \{b\}]$ and the space $Sgn^*(T) = \{[z, b, a, z, b], [z, b, z, a, b]\}$, where both sequences in $Sgn^*(T)$ are temporal observations of $\mathcal{Z}$.

**Definition 8** (**Candidate Set**)**.** *Let $\mathcal{O}$ be a temporal observation of $\mathcal{X}$. The* candidate set *of $\mathcal{O}$ is defined as*

$$\Delta(\mathcal{O}) = \{\, \delta(T) \mid T \in Bsp(\mathcal{X}), \mathcal{O} \in Sgn^*(T)\}\,. \qquad (7)$$

Intuitively, if $T$ is a trajectory in $Bsp(\mathcal{X})$ and $\mathcal{O}$ is a sequence in $Sgn^*(T)$, then $\mathcal{O}$ may have been generated by $T$ and, hence, $T$ may be the actual trajectory of $\mathcal{X}$. Hence, based on $\mathcal{O}$, $\delta(T)$ is a possible diagnosis of $\mathcal{X}$. A *diagnosis problem* amounts to determining the (whole) candidate set of a temporal observation of an AS being operated online.

**Example 7** (Candidate Set)**.** Let $\mathcal{O} = [z, b, z, a, b]$ be a temporal observation of $\mathcal{Z}$ (cf. Example 6). Based on the behavior space $Bsp(\mathcal{Z})$ in Fig. 1 and $Obs(\mathcal{Z})$ in Example 4, the only trajectory $T \in Bsp(\mathcal{Z})$ such that $\mathcal{O} \in Sgn^*(T)$ is $T = [z_3, z_2, z_3, b_4, z_2, z_3, b_4]$. Based on Example 3, $\delta(T) = \{\mathbf{z}\}$; hence, $\Delta(\mathcal{O}) = \{\{\mathbf{z}\}\}$: the candidate set is a singleton. In general, several trajectories may fulfill eqn. (7) and, thus, several candidates may be included in $\Delta(\mathcal{O})$.

## 4 Compiled Knowledge

In order to speed up the online diagnosis engine, it is convenient to compile specific knowledge offline based on the properties of the AS, including its observability and abnormality. In particular, the notion of observability of an AS (Definition 5) requires the diagnosis engine to match trajectories of $\mathcal{X}$ with regular languages specified by regular expressions. Based on eqn. (7), a candidate in $\Delta(\mathcal{O})$ is the diagnosis of a trajectory $T$ such that $\mathcal{O} \in Sgn^*(T)$. Based on Definition 7, $\mathcal{O} \in Sgn^*(T)$ means that we need to understand when observations occur based on the sequence of component transitions in $T$. Specifically, for each $(\mathbb{T}, \mathcal{L}, o) \in Obs(\mathcal{X})$, at any point of a prefix $T_i$ of $T$, namely $T_i = [t_1, \ldots, t_i]$, we need to check if the projection on $\mathbb{T}$ of a suffix of $T_i$ is a string in $\mathcal{L}$. If so, the observation $o$ should be in a proper position in $\mathcal{O}$ (otherwise $T$ does not conform with $\mathcal{O}$). The critical point is therefore to keep tracking possible strings in $\mathcal{L}$ based on sequences of component transitions in $T$. Since $\mathcal{L}$ is regular, it can be recognized by a finite automaton. However, a classical recognizer of the language is not sufficient for this task, as strings of the same language may overlap in $T$. To cope with possibly overlapping strings in the languages associated with observations, the notion of a *watcher* is introduced.

**Definition 9** (**Watcher**)**.** *Let $\mathcal{X}$ be an AS, let* $\mathbf{T}$ *be the set of component transitions in $\mathcal{X}$, and let $(\mathbb{T}, \mathcal{L}, o) \in Obs(\mathcal{X})$. Let $\mathcal{R}_o = (\mathbb{T}, R, \tau_r, r_0, R_f)$ be a finite automaton recognizing $\mathcal{L}$. Let $\mathcal{R}_o^\varepsilon$ be the nondeterministic finite automaton (NFA) obtained from $\mathcal{R}_o$ by inserting an $\varepsilon$-transition from each non-initial state to the initial state $r_0$. The* watcher $\mathcal{W}_o$ *of $o$ is a DFA obtained by the determinization of $\mathcal{R}_o^\varepsilon$.*

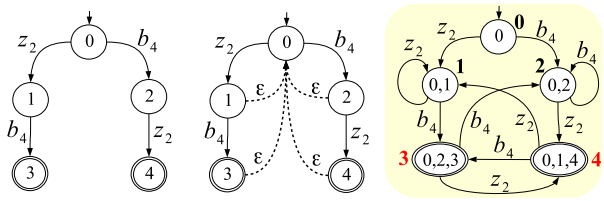


Figure 2: From left to right: $\mathcal{R}_a$, $\mathcal{R}_a^\varepsilon$, and watcher $\mathcal{W}_a$.

**Example 8** (Watcher)**.** With reference to $Obs(\mathcal{Z})$ (cf. Example 4), consider the language $\mathcal{L}_a = (z_2\, b_4 \,|\, b_4\, z_2)$, which is associated with the abstract observation $a$. Shown in Fig. 2 are the recognizer $\mathcal{R}_a$, the NFA $\mathcal{R}_a^\varepsilon$, and the watcher $\mathcal{W}_a$.[6] Note how the $\varepsilon$-transitions in $\mathcal{R}_a^\varepsilon$ allow for a continuous matching of (possibly overlapping) strings, which is in general not possible using a simple recognizer. To clarify, assume the following trajectory in $Bsp(\mathcal{Z})$:

$$T = [\, z_3, z_2, z_3, \overbrace{b_4}^{T'}, \underbrace{z_2, b_4}_{T''}, z_1, b_1\,]. \qquad (8)$$

$T$ includes two overlapping subtrajectories in $\mathcal{L}_a$, namely $T' = [\, b_4, z_2\,]$ and $T'' = [\, z_2, b_4\,]$, where the last transition $z_2$ of $T'$ is the first transition of $T''$. Hence, the observation $a$ is emitted twice in $T$, namely at the last transition of $T'$ and $T''$, respectively. Assume further to trace the emission of $a$ based on the recognizer $\mathcal{R}_a$. When the final state 4 is reached, $a$ is emitted. At this point, since no transition exits the final state 4, the recognizer starts again from the initial state 0 in order to keep matching.[7] It first changes state to 2 in correspondence of $b_4$, and with $z_1$ (mismatch) it returns to 0. The result is that, owing to the overlapping of the subtrajectories $T'$ and $T''$, the second emission of $a$ goes undetected. By contrast, consider matching $T$ based on the watcher $\mathcal{W}_a$. After the detection of $a$ at the final state **4**, the next transition $b_4$ moves to **3**, the other final state, thereby also detecting the emission of the second occurrence of $a$.

Given an AS $\mathcal{X}$, an essential knowledge structure to be generated offline is the *teleological space* of $\mathcal{X}$, denoted $Tsp(\mathcal{X})$. As the name suggests, a teleological space is conceived for a specific purpose, namely diagnosis, thereby involving information on observations and faults, based on the observability $Obs(\mathcal{X})$ and the abnormality $Abn(\mathcal{X})$, respectively. Roughly, $Tsp(\mathcal{X})$ is a DFA whose language equals the language of the behavior space $Bsp(\mathcal{X})$. Each state $x^\diamond$ of $Tsp(\mathcal{X})$ incorporates, beyond a state of $Bsp(\mathcal{X})$, a diagnosis $\bar{\delta}$ and a tuple $w$ of states of the watchers of the (abstract) observations. Notably, each state $w_i \in w$ tracks the recognition of the observation $o_i$ based on the watcher $\mathcal{W}_{o_i}$.

[6] Based on the classical SUBSET CONSTRUCTION determinization algorithm (Hopcroft, Motwani, and Ullman 2006), each state of the DFA is identified by a subset of the states of the NFA.

[7] A recognizer simply checks whether a given string is within a regular language: when the transition function is no longer applicable, it fails to match the rest of the string. What is relevant, however, is not that the recognizer would restart from the initial state, but that it cannot recognize overlapping strings.

A transition entering a state $(x, \delta, w)$ is marked with a pair $(t, O)$, where $t$ is a component transition, as in $Bsp(\mathcal{X})$, and $O$ is the (possibly empty) set of observations associated with the final states in $w$. Notably, a trajectory in $Tsp(\mathcal{X})$ ending in a state $(\bar{x}, \bar{\delta}, \bar{w})$ is a sequence of pairs $(t, O)$, where the projection on $t$ equals a trajectory in $T \in Bsp(\mathcal{X})$ such that $\delta(T) = \bar{\delta}$ (cf. Proposition 1).

**Definition 10 (Teleological Space).** *Let $\mathcal{X}$ be an AS, where $Bsp(\mathcal{X}) = (\Sigma, X, \tau, x_0, X_\mathrm{f})$, let $\mathbf{T}$ be the domain of component transitions in $\mathcal{X}$, let $Obs(\mathcal{X}) = \{(\mathbb{T}_1, \mathcal{L}_1, o_1), \ldots, (\mathbb{T}_k, \mathcal{L}_k, o_k)\}$, let $\mathbf{O}$ be the domain of observations, let $\mathbf{F}$ be the domain of faults, let $\mathcal{W}_i = (\mathbb{T}_i, W_i, \tau_i, w_{0i}, W_{\mathrm{f}i})$ be the watcher of $o_i$, $i \in [1\,..\,k]$, and let $W = (W_1 \times \cdots \times W_k)$. The* teleological space *of $\mathcal{X}$ is a DFA*

$$Tsp(\mathcal{X}) = (\Sigma^\diamond, X^\diamond, \tau^\diamond, x_0^\diamond, X_\mathrm{f}^\diamond) \tag{9}$$

*where $\Sigma^\diamond \subseteq \mathbf{T} \times 2^{\mathbf{O}}$ is the alphabet, $X^\diamond \subseteq X \times 2^{\mathbf{F}} \times W$ is the set of states, $x_0^\diamond = (x_0, \emptyset, w_0)$ is the initial state, with $w_0 = (w_{01}, \ldots, w_{0k})$, $X_\mathrm{f}^\diamond \subseteq X^\diamond$ is the set of final states, with $(x, \delta, w) \in X_\mathrm{f}^\diamond$ iff $x \in X_\mathrm{f}$, and $\tau^\diamond : X^\diamond \times \Sigma^\diamond \mapsto X^\diamond$ is the transition function, such that $\tau^\diamond((x, \delta, w), (t, O)) = (x', \delta', w')$, $w = (w_1, \ldots, w_k)$, $w' = (w_1', \ldots, w_k')$, iff $(x', \delta', w')$ is connected with a final state, $\tau(x, t) = x'$, and:*

$$\delta' = \begin{cases} \delta \cup \{f\} & \text{if } (t, f) \in Abn(\mathcal{X}) \\ \delta & \text{otherwise} \end{cases} \tag{10}$$

*$\forall i \in [1\,..\,k]$, the new state $w_i'$ of the watcher $\mathcal{W}_i$ is*

$$w_i' = \begin{cases} \bar{w}_i & \text{if } t \in \mathbb{T}_i \text{ and } \bar{w}_i = \tau_i(w_i, t) \\ w_{0i} & \text{if } t \in \mathbb{T}_i \text{ and } \tau_i(w_i, t) \text{ is undefined} \\ w_i & \text{if } t \notin \mathbb{T}_i \end{cases} \tag{11}$$

$$O = \{o_i' \mid i \in [1\,..\,k], \tau_i(w_i, t) = w_i', w_i' \in W_{\mathrm{f}i}\}\,. \tag{12}$$

*Similarly to the behavior space, a string in the regular language of $Tsp(\mathcal{X})$ is a* trajectory *in $Tsp(\mathcal{X})$.*

**Example 9** (Teleological Space)**.** Shown in Fig. 3 is $Tsp(\mathcal{Z})$, which comprises 32 states, renamed $0\,..\,31$, where 0 is the initial state and the double circles denote the final states. Each state $z^\diamond$ is identified by a triple: a state in $Bsp(\mathcal{Z})$, a diagnosis $\delta$, and a state $w_a$ of the watcher $\mathcal{W}_a$ (cf. Fig. 2). Note that, according to Definition 10 and Example 4, the third field of $z^\diamond$ should be a tuple $(w_z, w_b, w_a)$, where $w_z \in \mathcal{W}_z$ and $w_b \in \mathcal{W}_b$. The reason for $w_z$ and $w_b$ being missing is that both languages $\mathcal{L}_z$ and $\mathcal{L}_b$ include strings that are composed of one transition only, for instance, $\mathcal{L}_z = \{[z_1], [z_2]\}$. Thus, the observation can be detected directly based on the component transition only. In other words, when the observation $o$ is not abstract (that is, when $o$ is associated with single component transitions), the watcher $\mathcal{W}_o$ becomes unnecessary, as it is for $\mathcal{W}_z$ and $\mathcal{W}_b$. Each transition in $Tsp(\mathcal{Z})$ is marked with a pair $(t, O)$, where $t$ is a component transition and $O$ is the (possibly empty) set of observations emitted at the occurrence of $t$. For instance, we have $\langle 0, (z_1, \{z\}), 1\rangle$, where $\{z\}$ is represented as $z$ in Fig. 3, since, according to $Obs(\mathcal{Z})$, the transition $z_1$ of the transducer generates the observation $z$. Likewise, we have

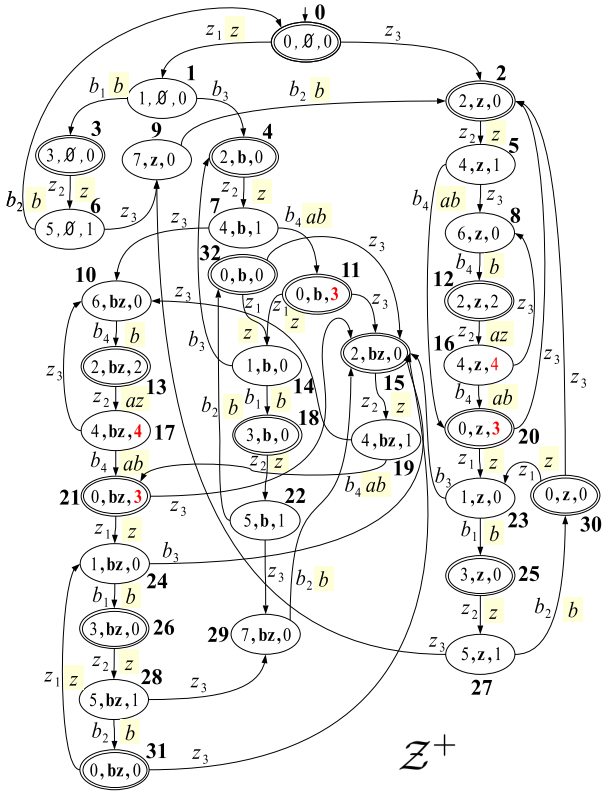


Figure 3: Teleological space $Tsp(\mathcal{Z})$ (cf. $Bsp(\mathcal{Z})$ in Fig. 1).

$\langle 12, (z_2, \{a, z\}), 16\rangle$, where the set $\{a, z\}$ is represented as the string $az$ in Fig. 3, since the transition $z_2$ of the transducer generates the observation $z$ and the state 4 is final in $\mathcal{W}_a$. To simplify the figure, $O$ is missing when it is empty, as in $\langle 0, (z_3, \emptyset), 2\rangle$. Also, according to eqn. (10), the set of faults $\delta$ is extended when the transition $t$ is faulty. For instance, based on $Abn(\mathcal{Z})$, in the transition $\langle 11, (z_3, \emptyset), 15\rangle$, where $11 = (0, \{\mathbf{b}\}, 3)$, we have $\delta = \{\mathbf{b}\}$ being extended by $\mathbf{z}$, yielding $\delta' = \{\mathbf{b}, \mathbf{z}\}$ in the state 15.

The projection of the language of the teleological space on the component transitions equals the language of the behavior space. Moreover, the diagnosis marking the final state of a trajectory in the teleological space equals the diagnosis of the corresponding trajectory in the behavior space. In other words, the teleological space extends each trajectory in the behavior space with the relevant diagnosis (Proposition 1).

**Proposition 1.** *Let $Tsp(\mathcal{X}) = (\Sigma^\diamond, X^\diamond, \tau^\diamond, x_0^\diamond, X_\mathrm{f}^\diamond)$.*

$$\{[t \mid (t, O) \in T^\diamond, T^\diamond \in Tsp(\mathcal{X})]\} = \{T \mid T \in Bsp(\mathcal{X})\}\,. \tag{13}$$

*Also, if $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ is a trajectory in $Tsp(\mathcal{X})$, ending in $(x, \bar{\delta}, w)$, then $[O_1, \ldots, O_q]$ equals the signature of $T = [t_1, \ldots, t_q] \in Bsp(\mathcal{X})$, where $\delta(T) = \bar{\delta}$.*

**Proof** (*sketch*). If $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ is a trajectory in $Tsp(\mathcal{X})$, then, based on the definition of the transition function $\tau^\diamond$ in Definition 10, $\tau(x, t) = x'$, in other words, $T \in Bsp(\mathcal{X})$, where $T = [t_1, \ldots, t_q]$. If $T = [t_1, \ldots, t_q] \in Bsp(\mathcal{X})$, then, based on the definition

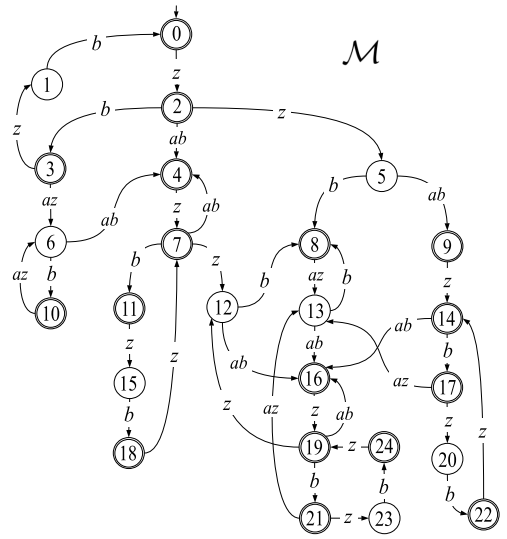


Figure 4: Diagnosis reference manual $\mathcal{M}$ of $\mathcal{Z}$ (cf. Table 1).

| State $\mu$ | States of $Tsp(\mathcal{Z})$ within $\mu$ | $\Delta(\mu)$ |
|---|---|---|
| **0** | $\{\mathbf{0}, \mathbf{2}\}$ | $\{\emptyset, \{\mathbf{z}\}\}$ |
| 1 | $\{6, 9\}$ | $\emptyset$ |
| **2** | $\{1, \mathbf{4}, 5, 8\}$ | $\{\{\mathbf{b}\}\}$ |
| **3** | $\{\mathbf{3}, \mathbf{12}\}$ | $\{\emptyset, \{\mathbf{z}\}\}$ |
| **4** | $\{\mathbf{2}, \mathbf{20}\}$ | $\{\{\mathbf{z}\}\}$ |
| 5 | $\{7, 10\}$ | $\emptyset$ |
| 6 | $\{8, 16\}$ | $\emptyset$ |
| **7** | $\{5, 8, \mathbf{15}, 23\}$ | $\{\{\mathbf{b}, \mathbf{z}\}\}$ |
| **8** | $\{\mathbf{13}\}$ | $\{\{\mathbf{b}, \mathbf{z}\}\}$ |
| **9** | $\{\mathbf{11}, \mathbf{15}\}$ | $\{\{\mathbf{b}\}, \{\mathbf{b}, \mathbf{z}\}\}$ |
| **10** | $\{\mathbf{12}\}$ | $\{\{\mathbf{z}\}\}$ |
| **11** | $\{\mathbf{12}, \mathbf{25}\}$ | $\{\{\mathbf{z}\}\}$ |
| 12 | $\{10, 19\}$ | $\emptyset$ |
| 13 | $\{10, 17\}$ | $\emptyset$ |
| **14** | $\{\mathbf{4}, 10, 14, 19\}$ | $\{\{\mathbf{b}\}\}$ |
| 15 | $\{9, 27\}$ | $\emptyset$ |
| **16** | $\{\mathbf{15}, \mathbf{21}\}$ | $\{\{\mathbf{b}, \mathbf{z}\}\}$ |
| **17** | $\{\mathbf{13}, \mathbf{18}\}$ | $\{\{\mathbf{b}\}, \{\mathbf{b}, \mathbf{z}\}\}$ |
| **18** | $\{\mathbf{2}, \mathbf{30}\}$ | $\{\{\mathbf{z}\}\}$ |
| **19** | $\{10, \mathbf{15}, 19, 24\}$ | $\{\{\mathbf{b}, \mathbf{z}\}\}$ |
| 20 | $\{22, 29\}$ | $\emptyset$ |
| **21** | $\{\mathbf{13}, \mathbf{26}\}$ | $\{\{\mathbf{b}, \mathbf{z}\}\}$ |
| **22** | $\{\mathbf{15}, \mathbf{32}\}$ | $\{\{\mathbf{b}\}, \{\mathbf{b}, \mathbf{z}\}\}$ |
| 23 | $\{28, 29\}$ | $\emptyset$ |
| **24** | $\{\mathbf{15}, \mathbf{31}\}$ | $\{\{\mathbf{b}, \mathbf{z}\}\}$ |

Table 1: Details of the states in the diagnosis reference manual $\mathcal{M}$ outlined in Fig. 4, where bold numbers denote final states.

of $\tau^\diamond$, $\tau(x,t) = x'$ (without any constraints from the definition of $\delta'$ and $w'$). Hence, $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)] \in Tsp(\mathcal{X})$. This proves eqn. (13). Also, let $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ be a trajectory in $Tsp(\mathcal{X})$ ending in $(x, \bar{\delta}, w)$. Based on eqn. (10), $\delta'$ is extended by collecting the faults associated with the faulty transitions in $T = [t_1, \ldots, t_q]$. Hence, based on eqn. (2), $\bar{\delta} = \delta(T)$. □

**Example 10** (Proposition 1)**.** With reference to $Bsp(\mathcal{Z})$ displayed in Fig. 1, consider the trajectory $T = [z_3, z_2, z_3, b_4, z_2, z_3, b_4]$ in Example 2. There is a trajectory $T^\diamond = [(z_3, \emptyset), (z_2, \{z\}), (z_3, \emptyset), (b_4, \{b\}), (z_2, \{a, z\}), (z_3, \emptyset), (b_4, \{b\})]$ in $Tsp(\mathcal{Z})$, ending in the state $12 = (2, \{\mathbf{z}\}, 2)$, whose projection on component transitions equals $T$. Besides, $\{\mathbf{z}\} = \delta(T)$ (cf. Example 3).

The teleological space $Tsp(\mathcal{X})$ is only an intermediate representation of the actual knowledge structure aimed at supporting the diagnosis engine being operated online. The ultimate goal is to construct a DFA whose language is the set of signatures (cf. Definition 7) of the trajectories of $\mathcal{X}$, with empty sets being missing. This is a sort of diagnostic manual, against which any given temporal observation $\mathcal{O}$ can be easily matched online by the diagnosis engine in order to determine the set of candidates. Specifically, each state is marked with a set of diagnoses, in such a way that the final states reached by matching $\mathcal{O}$ allow for the exact computation of the candidate set $\Delta(\mathcal{O})$.

**Definition 11 (Diagnosis Reference Manual).** *Let $Tsp(\mathcal{X}) = (\Sigma^\diamond, X^\diamond, \tau^\diamond, x_0^\diamond, X_\mathrm{f}^\diamond)$ be a teleological space. The* nondeterministic reference manual *of $\mathcal{X}$ is an NFA $\mathcal{M}^\varepsilon$ obtained from $Tsp(\mathcal{X})$ by replacing the pair $(t, O)$ marking each transition in $Tsp(\mathcal{X})$ with a new symbol $\sigma$, namely*

$$\sigma = \begin{cases} O & \text{if } O \neq \emptyset \\ \varepsilon & \text{otherwise.} \end{cases} \tag{14}$$

*The* diagnosis reference manual *of $\mathcal{X}$ is a DFA $\mathcal{M}$ obtained by determinization of $\mathcal{M}^\varepsilon$. A string in the language of $\mathcal{M}$ is a* trajectory *in $\mathcal{M}$. The* diagnosis set *of a state $\mu$ of $\mathcal{M}$ is a set of diagnoses*

$$\Delta(\mu) = \{\delta \mid x^\diamond \in (\mu \cap X_\mathrm{f}^\diamond), x^\diamond = (x, \delta, w)\}\,. \tag{15}$$

Based on eqn. (15), $\Delta(\mu) \neq \emptyset$ iff $\mu$ includes a final state.

**Example 11** (Diagnosis Reference Manual)**.** With reference to $Tsp(\mathcal{Z})$ displayed in Fig. 3, the nondeterministic reference manual of $\mathcal{Z}$, namely $\mathcal{M}^\varepsilon$, is obtained by substituting the pairs $(t, O)$ marking the transitions of $Tsp(\mathcal{Z})$ in accordance with eqn. (14), where transitions are, consequently, possibly marked with the empty symbol $\varepsilon$ (this is why $\mathcal{M}^\varepsilon$ is nondeterministic). The reference manual $\mathcal{M}$ of $\mathcal{Z}$, obtained by determinization of $\mathcal{M}^\varepsilon$, is outlined in Fig. 4, with state details being listed in Table 1, where bold numbers denote final states. In fact, according to the SUBSET CONSTRUCTION determinization algorithm, each state $\mu$ in $\mathcal{M}$ is identified by a subset of the $Tsp(\mathcal{Z})$ states in $\mathcal{M}^\varepsilon$. For instance, the state $\mathbf{3}$ in $\mathcal{M}$ is identified by the set $\{\mathbf{3}, \mathbf{12}\}$, where both states $\mathbf{3} = (3, \emptyset, 0)$ and $\mathbf{12} = (2, \{\mathbf{z}\}, 2)$ are final in $Tsp(\mathcal{Z})$. Hence, based on eqn. (15), we have $\Delta(\mathbf{3}) = \{\emptyset, \{\mathbf{z}\}\}$. Note that $\Delta(5) = \emptyset$ since the state $5 = \{7, 10\}$ does not include any final state in $Tsp(\mathcal{Z})$.

Notably, the language of the diagnosis reference manual is the set of the signatures of the trajectories in the behavior space, with the additional information of the relevant diagnoses being stored in the final states (Proposition 2).

**Proposition 2.** *Let $Tsp(\mathcal{X})$ be the teleological space of $\mathcal{X}$ and let $\mathcal{M}$ be the diagnosis reference manual of $\mathcal{X}$. Let $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ be a trajectory in $Tsp(\mathcal{X})$ and let $T = [t_1, \ldots, t_q]$. There is $[O \mid (t, O) \in T^\diamond, O \neq \emptyset] \in \mathcal{M}$ ending in a final state $\mu$ such that $\delta(T) \in \Delta(\mu)$. Conversely, let $\bar{\mathbb{O}} = [\bar{O}_1, \ldots, \bar{O}_m] \in \mathcal{M}$ ending in a final state $\mu$. There is a trajectory $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)] \in Tsp(\mathcal{X})$ such that $\bar{\mathbb{O}} = [O \mid (t, O) \in T^\diamond, O \neq \emptyset]$ and $\delta(T) \in \Delta(\mu)$, where $T = [t_1, \ldots, t_q]$.*

**Proof** (*sketch*). Assume $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ is a trajectory in $Tsp(\mathcal{X})$, with $T = [t_1, \ldots, t_q]$. According to the SUBSET CONSTRUCTION determinization algorithm generating $\mathcal{M}$ based on $\mathcal{M}^\varepsilon$ in Definition 11, the string $\bar{\mathbb{O}} = [O \mid (t, O) \in T^\diamond, O \neq \emptyset]$ is in the language of $\mathcal{M}$, ending in a final state $\mu$, such that the final state $(x, \bar{\delta}, w)$ of $T^\diamond$ is included in $\mu$. Based on Proposition 1, $\delta(T) = \bar{\delta}$. Hence, based on eqn. (15), $\delta(T) \in \Delta(\mu)$. Now, assume $\bar{\mathbb{O}} = [\bar{O}_1, \ldots, \bar{O}_m]$ is a trajectory in $\mathcal{M}$ ending in a final state $\mu$. Based on the determinization of $\mathcal{M}^e$ in Definition 11, there is a trajectory $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ in $Tsp(\mathcal{X})$, ending in a final state $(x, \bar{\delta}, w)$, such that $\bar{\mathbb{O}} = [O \mid (t, O) \in T^\diamond, O \neq \emptyset]$. Based on Proposition 1, $\delta(T) = \bar{\delta}$, where $T = [t_1, \ldots, t_q]$. □

**Example 12** (Proposition 2)**.** With reference to $Tsp(\mathcal{Z})$ in Fig. 3 and the corresponding diagnosis reference manual $\mathcal{M}$ outlined in Fig. 4 and detailed in Table 1, let $T^\diamond = [(z_3, \emptyset), (z_2, \{z\}), (b_4, \{a, b\}), (z_1, \{z\}), (b_3, \emptyset)]$ be a trajectory in $Tsp(\mathcal{Z})$ ending in state $15 = (2, \{\mathbf{b}, \mathbf{z}\}, 0)$, where $T = [z_3, z_2, b_4, z_1, b_3]$, and $\delta(T) = \{\mathbf{b}, \mathbf{z}\}$. As claimed in Proposition 2, there is a trajectory $[O \mid (t, O) \in T^\diamond, O \neq \emptyset]$ in $\mathcal{M}$, namely $[\{z\}, \{a, b\}, \{z\}]$, ending in the state $\mu = \mathbf{7}$, where $\Delta(\mathbf{7}) = \{\{\mathbf{b}, \mathbf{z}\}\}$, such that $\delta(T) \in \Delta(\mathbf{7})$. Conversely, let $\bar{\mathbb{O}} = [\{z\}, \{b\}]$ be a trajectory in $\mathcal{M}$, ending in the final state $\mathbf{3}$, where $\Delta(\mathbf{3}) = \{\emptyset, \{\mathbf{z}\}\}$. There is a trajectory $T^\diamond = [(z_3, \emptyset), (z_2, \{z\}), (z_3, \emptyset), (b_4, \{b\})]$ in $Tsp(\mathcal{Z})$ where $\bar{\mathbb{O}} = [O \mid (t, O) \in T^\diamond, O \neq \emptyset]$, $T = [z_3, z_2, z_3, b_4]$, $\delta(T) = \{\mathbf{z}\}$, and $\delta(T) \in \Delta(\mathbf{3})$.

## 5 Diagnosis Engine

The process of knowledge compilation presented in Section 4 is performed offline; as such, it is independent of the particular diagnosis problem associated with a temporal observation $\mathcal{O}$. Solving a diagnosis problem, that is, determining the candidate set of $\mathcal{O}$, is a task that is performed online by a *diagnosis engine*, which is independent of the specific AS. The diagnosis engine takes as input a temporal observation $\mathcal{O}$ of $\mathcal{X}$, along with the diagnosis reference manual $\mathcal{M}$ of $\mathcal{X}$, and generates the candidate set $\Delta(\mathcal{O})$. In so doing, the diagnosis engine does not perform any model-based reasoning since all it needs is incorporated in $\mathcal{M}$, the ultimate result of knowledge compilation. Roughly, the diagnosis engine performs the simple task of matching $\mathcal{O}$ against $\mathcal{M}$ and, once determined the accepting (final) states of $\mathcal{M}$, generates the diagnosis set $\Delta(\mathcal{O})$ by collecting in one basket the diagnoses marking these states. To this end, it generates online a data structure called *abduction* of $\mathcal{O}$, which tracks $\mathcal{O}$ on $\mathcal{M}$.

**Definition 12 (Abduction).** *Let $\mathcal{M} = (\Sigma, M, \tau, \mu_0, M_\mathrm{f})$ be the diagnosis reference manual of $\mathcal{X}$, and let $\mathcal{O} = [o_1, \ldots, o_n]$ be a temporal observation of $\mathcal{X}$. The* abduction *of $\mathcal{O}$ is a DFA*

$$Abd(\mathcal{O}) = (\Sigma, A, \tau_\mathrm{a}, a_0, A_\mathrm{f}) \tag{16}$$

*where $A \subseteq M \times [0\,..\,n]$ is the set of states, $a_0 = (\mu_0, 0)$ is the initial state, $A_\mathrm{f} = \{a \mid a \in A, a = (\mu, n), \mu \in M_\mathrm{f}\}$ is the set of final states, and $\tau_\mathrm{a} : A \times \Sigma \mapsto A$ is the transition function, with $\tau_\mathrm{a}((\mu, \Im), O) = (\mu', \Im')$ iff $\tau(\mu, O) = \mu'$, $\Im' = \Im + |O|$, $\Im' \leq n$, and $O = \{o_{\Im+1}, \ldots, o_{\Im+|O|}\}$, where $|O|$ denotes the cardinality of the set $O$. A string in the language of $Abd(\mathcal{O})$ is a* trajectory *in $Abd(\mathcal{O})$.*

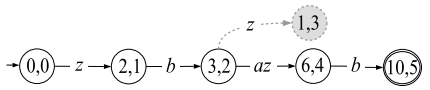


Figure 5: Generation of $Abd(\mathcal{O})$ for $\mathcal{Z}$, where $\mathcal{O} = [z, b, z, a, b]$.

**Example 13** (Abduction)**.** Let $\mathcal{O} = [z, b, z, a, b]$ be a temporal observation of $\mathcal{Z}$ (cf. Example 7). Based on the diagnosis reference manual $\mathcal{M}$ of $\mathcal{Z}$ (cf. Fig. 4 and Table 1), the generation of the abduction $Abd(\mathcal{O})$ is displayed in Fig. 5. In accordance with Definition 12, each state is composed of a pair $(\mu, \Im)$, where $\mu$ is a state of $\mathcal{M}$ and $\Im$ is an integer number in the range $[0\,..\,5]$, with 5 being the number of observations in $\mathcal{O}$. Note how the diagnosis engine may generate some *spurious* states which are not connected with any final state, like the state $(1, 3)$. These states, along with relevant transitions, are spurious. The initial state is $(0, 0)$, where $\mu = 0$ is the initial state of $\mathcal{M}$ and $\Im = 0$ indicates that no observations have yet been matched. Since the first observation of $\mathcal{O}$ is $z$, based on the transition function of $\mathcal{M}$, the state reached by matching $z$ is 2, thereby leading to the creation of the abduction transition $\langle (0, 0), \{z\}, (2, 1)\rangle$. Since the next (second) observation in $\mathcal{O}$ is $b$, the only transition in $\mathcal{M}$ that conforms with $b$ is $\langle 2, \{b\}, 3\rangle$, which leads to the creation of the transition $\langle (2, 1), \{b\}, (3, 2)\rangle$ in the abduction. Since the next (third) observation is $z$, two transitions in $\mathcal{M}$ are applicable, namely $\langle 3, \{z\}, 1\rangle$ and $\langle 3, \{a, z\}, 6\rangle$, where the former consumes one observation, whereas the latter consumes two observations, namely $a$ and $z$. Consequently, two transitions are created in the abduction, namely $\langle (3, 2), \{z\}, (1, 3)\rangle$ and $\langle (3, 2), \{a, z\}, (6, 4)\rangle$, where in state $(6, 4)$ the index $\Im$ is increased by two units (the cardinality of $\{a, z\}$). In the state $(1, 3)$, the next (fourth) observation is $a$. However, the only transition exiting the state 1 in $\mathcal{M}$ is marked with $\{b\}$, a mismatch. This is why $(1, 3)$ is marked as spurious, along with its entering transition. Instead, in state $(6, 4)$, the next (fifth) observation is $b$, which is matched by the transition $\langle 6, \{b\}, 10\rangle$ in $\mathcal{M}$, thereby allowing for the creation of the transition $\langle (6, 4), \{b\}, (10, 5)\rangle$ in $Abd(\mathcal{O})$. Note how $(10, 5)$ is final because 10 is final in $\mathcal{M}$ and 5 is the number of observations in $\mathcal{O}$. Eventually, no transition can exit $(10, 5)$ in the abduction, as no further observation can be matched in $\mathcal{O}$. Hence, the construction of $Abd(\mathcal{O})$ terminates, with $(10, 5)$ being the only final state.

The language of an abduction $Abd(\mathcal{O})$ is the set of the trajectories $\bar{\mathbb{O}}$ in the diagnosis reference manual such that $\mathcal{O}$ is 'compatible' with $\bar{\mathbb{O}}$, namely $\mathcal{O} \in \bar{\mathbb{O}}^*$ (Proposition 3).

**Proposition 3.** *Let $\mathcal{M}$ be the diagnosis reference manual of $\mathcal{X}$, let $\mathcal{O} = [o_1, \ldots, o_n]$ be a temporal observation of $\mathcal{X}$, and let $\bar{\mathbb{O}}$ be a trajectory in $\mathcal{M}$, ending in a final state $\mu$, such that $\mathcal{O} \in \bar{\mathbb{O}}^*$.*[8] *We have that $\bar{\mathbb{O}}$ is a trajectory in $Abd(\mathcal{O})$ ending in a final state $(\mu, n)$. Also, let $\bar{\mathbb{O}}$ be a trajectory in $Abd(\mathcal{O})$ ending in a final state $(\mu, n)$. We have that $\bar{\mathbb{O}}$ is a trajectory in $\mathcal{M}$ ending in state $\mu$, where $\mathcal{O} \in \bar{\mathbb{O}}^*$.*

[8] For the definition of $\bar{\mathbb{O}}^*$, see eqn. (5) in Definition 6.

**Proof** (*sketch*). Assume that $\bar{\mathbb{O}}$ is a trajectory in $\mathcal{M}$, ending in a final state $\mu$, where $\mathcal{O} \in \bar{\mathbb{O}}^*$. By induction on $\bar{\mathbb{O}}$, starting from the initial state $(\mu_0, 0)$ of $Abd(\mathcal{O})$ and based on Definition 6 and $\tau_a$ in Definition 12, the property $\mathcal{O} \in \bar{\mathbb{O}}$ assures the fulfillment of the key condition $O = \{O_{\Im+1}, \ldots, \mathcal{O}_{\Im+|O|}\}$ at each inductive step relevant to the transition $\langle(\mu, \Im), O, (\mu', \Im')\rangle$ in $\tau_{\mathrm{a}}$. Hence, $\bar{\mathbb{O}}$ is a trajectory in $Abd(\mathcal{O})$. Assume now that $\bar{\mathbb{O}}$ is a trajectory in $Abd(\mathcal{O})$, ending in a final state $(\mu, n)$. Based on Definition 12, by induction on $\bar{\mathbb{O}}$ and starting from the initial state $(\mu_0, 0)$, at each induction step, the condition $O = \{O_{\Im+1}, \ldots, \mathcal{O}_{\Im+|O|}\}$, for each transition $\langle(\mu, \Im), O, (\mu', \Im')\rangle$ in $\tau_a$, assures that $\mathcal{O} \in \bar{\mathbb{O}}^*$. Besides, the condition $\tau(\mu, O) = O'$ assures that $\bar{\mathbb{O}}$ is a trajectory in $\mathcal{M}$ also. □

**Example 14** (Proposition 3)**.** Consider the diagnosis reference manual $\mathcal{M}$ of $\mathcal{Z}$ (cf. Fig. 4 and Table 1) and $Abd(\mathcal{O})$ in Fig. 5, where $\mathcal{O} = [z, b, z, a, b]$ (cf. Example 13). Let $\mathbb{O} = [\{z\}, \{b\}, \{a, z\}, \{b\}]$ be a trajectory in $\mathcal{M}$ ending in the final state 10, where $\mathcal{O} \in \bar{\mathbb{O}}^*$ (cf. Definition 6). As claimed in Proposition 3, $\bar{\mathbb{O}}$ is a trajectory in $Abd(\mathcal{O})$ ending in a final state $(10, 5)$. The converse is also true.

A candidate set $\Delta(\mathcal{O})$ can be generated by collecting the diagnosis sets of the states of the diagnosis reference manual incorporated in the final states of $Abd(\mathcal{O})$ (Theorem 1).

**Theorem 1.** *Let $\mathcal{M}$ be the diagnosis reference manual of $\mathcal{X}$ and let $\mathcal{O} = [o_1, \ldots, o_n]$ be a temporal observation of $\mathcal{X}$. The candidate set of $\mathcal{O}$ can be computed based on the final states $A_{\mathrm{f}}$ of the abduction $Abd(\mathcal{O})$, specifically*

$$\Delta(\mathcal{O}) = \bigcup_{(\mu,n)\in A_{\mathrm{f}}} \Delta(\mu). \tag{17}$$

**Proof** (*sketch*). Let $\Delta_a$ denote $\bigcup_{(\mu,n)\in A_{\mathrm{f}}} \Delta(\mu)$ in eqn. (17). If $\bar{\delta} \in \Delta(\mathcal{O})$, then, based on eqn. (7), there is a trajectory $T \in Bsp(\mathcal{X})$, $T = [t_1, \ldots, t_q]$, such that $\mathcal{O} \in Sgn^*(T)$ and $\bar{\delta} = \delta(T)$. Based on Proposition 1, there is a trajectory $T^\diamond = [(t_1, O_1), \ldots, (t_q, O_q)]$ in $Tsp(\mathcal{X})$, ending in a final state $(x, \bar{\delta}, w)$, where $Sgn(T) = [O_1, \ldots, O_q]$ and $\mathcal{O} \in Sgn^*(T)$. Based on Proposition 2, there is a trajectory $\bar{\mathbb{O}} = [O \mid (t, O) \in T^\diamond, O \neq \emptyset]$ in the diagnosis reference manual $\mathcal{M}$ of $\mathcal{X}$, ending in a final state $\mu$, such that $(x, \bar{\delta}, w) \in \mu$ and $\mathcal{O} \in Sgn^*(T)$. Based on Proposition 3, there is a trajectory $\bar{\mathbb{O}}$ in $Abd(\mathcal{O})$, ending in a final state $(\mu, n)$. Hence, based on eqn. (15), $\bar{\delta} \in \Delta_a$.

If $\bar{\delta} \in \Delta_a$, then, based on eqn. (17), there is a trajectory $\bar{\mathbb{O}}$ in $Abd(\mathcal{O})$, ending in a final state $(\mu, n)$, where $\mathcal{O} \in \bar{\mathbb{O}}^*$ and $\bar{\delta} \in \Delta(\mu)$. Based on Proposition 3, there is a trajectory $\bar{\mathbb{O}}$ in $\mathcal{M}$, ending in $\mu$, where $\mathcal{O} \in \bar{\mathbb{O}}^*$ and $\bar{\delta} \in \Delta(\mu)$. Based on Proposition 2, there is a trajectory $T^\diamond = [(t_1, o_1), \ldots, (t_q, o_q)]$ in $Tsp(\mathcal{X})$, where $\bar{\mathbb{O}} = [O \mid (t, O) \in T^\diamond, O \neq \emptyset]$, $[O_1, \ldots, O_q] = Sgn(T)$, $T = [t_1, \ldots, t_q]$, and $\delta(T) = \bar{\delta}$. Based on Proposition 1, $T \in Bsp(\mathcal{X})$, $\mathcal{O} \in Sgn^*(T)$, and $\delta(T) = \bar{\delta}$. Hence, according to eqn. (7), $\bar{\delta} \in \Delta(\mathcal{O})$. □

**Example 15** (Theorem 1)**.** Let $\mathcal{O} = [z, b, z, a, b]$ be a temporal observation of $\mathcal{Z}$, where $Abd(\mathcal{O})$ is displayed in Fig. 5. Considering eqn. (17), the only final state in $Abd(\mathcal{O})$ is $(10, 5)$, where 10 is a final state in $\mathcal{M}$. Since $\Delta(10) = \{\{\mathbf{z}\}\}$ (cf. Table 1), according to Theorem 1, the candidate set is $\Delta(\mathcal{O}) = \{\{\mathbf{z}\}\}$, which is in fact the same singleton determined in Example 7 based on Definition 8. Instead, with $\mathcal{O} = [z, b]$, it is easy to find out that $Abd(\mathcal{O})$ includes the only final state $(3, 2)$ Hence, $\Delta(\mathcal{O}) = \Delta(3) = \{\emptyset, \{z\}\}$: either no faults occur or the transducer is faulty. Finally, with $\mathcal{O} = [z, z, b, a, z, b]$, the final states of $Abd(\mathcal{O})$ are $(8, 6)$ and $(17, 6)$. Hence, $\Delta(\mathcal{O}) = \Delta(8) \cup \Delta(17) = \{\{b\}, \{b, z\}\}$.

## 6 Coping with Temporal Uncertainty

In Section 3 we have assumed that the signature of a trajectory $T = [t_1, \ldots, t_q]$, namely $Sgn(T) = [O_1, \ldots, O_q]$, manifests itself to the external observer as a temporal observation $\mathcal{O} = [o_1, \ldots, o_n]$, where $n = \Sigma_{i=1}^{q} |O_i|$, namely $n$ is the sum of the cardinalities of the sets $O_i$. Specifically, each $O_i \in Sgn(T)$, $i \in [1\,..\,q]$, is mapped to a sequence $\bar{O}_i$, so that $\mathcal{O}$ is generated by the concatenation $\bigsqcup_{i=1}^{q} \bar{O}_i$ of such sequences. Hence, $\mathcal{O}$ is one of (in general) several temporal observations derivable from $Sgn(T)$, namely $\mathcal{O} \in Sgn^*(T)$. In diagnosis of ASs with simple (non abstract) observations, a temporal observation $\mathcal{O}$ is the projection of a trajectory $T$ on the observations associated with the observable transitions. Hence, for each $T$ there is just one temporal observation $\mathcal{O}$. However, owing to the distribution of the system and/or noise in the communication channels, $\mathcal{O}$ may become *uncertain* (Lamperti and Zanella 2002). A particular form of uncertainty is *temporal uncertainty*, where the total temporal ordering of observations in $\mathcal{O}$ is relaxed to partial ordering. The result is a graph where nodes represent observations and arcs represent partial temporal precedence between nodes. In this section, the diagnosis technique based on abstract observability is extended to cope with temporal uncertainty.

**Definition 13** (**Observation Graph**)**.** *Let $\mathcal{O} = [o_1, \ldots, o_n]$ be a temporal observation of $\mathcal{X}$. The* graph *of $\mathcal{O}$ is a directed acyclic graph (DAG) $\mathcal{G} = (\Omega, \mathcal{A})$, where $\Omega$ is the set of nodes and $\mathcal{A}$ the set of arcs, defined as follows. For each $o_i \in \mathcal{O}$, $i \in [1\,..\,n]$, there is a node $\omega_i \in \Omega$ that is marked with $o_i$. For each $o_i \in \mathcal{O}$, $i \in [1\,..\,(n-1)]$, there is an arc $(\omega_i, \omega_{i+1})$ in $\mathcal{A}$. If there is a path in $\mathcal{G}$ from a node $\omega$ to a node $\omega'$, then $\omega$* precedes *$\omega'$, denoted $\omega \prec \omega'$.*[9]

**Example 16** (Observation Graph)**.** Let $\mathcal{O} = [z, b, z, a, b]$ be a temporal observation of $\mathcal{Z}$. The graph $(\Omega, \mathcal{N})$ of $\mathcal{O}$ is shown on the left of Fig. 7, where $\Omega = \{1, \ldots, 5\}$.

**Definition 14** (**Uncertain Observation**)**.** *Let $\mathcal{G} = (\Omega, \mathcal{A})$ be the graph of a temporal observation $\mathcal{O} = [o_1, \ldots, o_n]$ of $\mathcal{X}$. Let $\mathcal{G}$ be transformed by applying a list of zero or more* temporal relaxations*, each of them being defined by three steps: (1) an arc $(\omega, \omega')$ is removed; (2) for each arc $(\omega', \omega_2)$, if $\omega \not\prec \omega_2$, then an arc $(\omega, \omega_2)$ is inserted; (3) for each arc $(\omega_1, \omega)$, if $\omega_1 \not\prec \omega'$, then an arc $(\omega_1, \omega')$ is inserted. The resulting graph $\mathcal{U}$ is an* uncertain observation *of $\mathcal{X}$ derived from $\mathcal{O}$. Let $Q = [\omega_1, \ldots, \omega_n]$ be a sequence*

[9] Since the arcs in $\mathcal{G}$ express total temporal ordering, $\omega \prec \omega'$ boils down to having $\omega$ before $\omega'$ in the linear graph. The definition of temporal precedence becomes significant when the graph involves partial temporal ordering (cf. Definition 14).

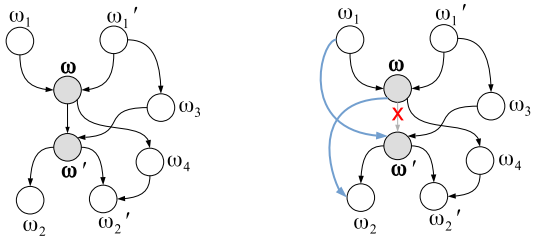

Figure 6: Temporal relaxation (cf. Definition 14).

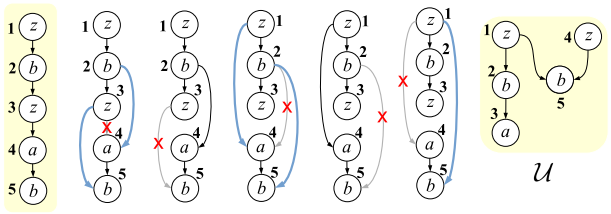

Figure 7: From a temporal observation $\mathcal{O} = [z, b, z, a, b]$ to an uncertain observation $\mathcal{U}$.

*including all the states in $\Omega$ where, for each $i, j \in [1\,..\,n]$, if $\omega_i \prec \omega_j$ in $\mathcal{U}$, then $i < j$, called a* topological sort *of $\Omega$ based on $\mathcal{A}$. A sequence $\mathcal{O} = [o'_1, \ldots, o'_n]$, where each $o'_i$ is the observation marking $\omega_i$ in $Q$, $i \in [1\,..\,n]$, is a* temporal observation embedded *in $\mathcal{U}$. The (finite) set of temporal observations embedded in $\mathcal{U}$ is denoted $\mathcal{U}^*$.*

Actions (1) and (2) in Definition 14 are meant to preserve the other temporal precedences after the removal of the arc $(\omega, \omega')$, so that only one precedence at a time is removed. To clarify, consider the observation graph displayed on the left of Fig. 6. The removal of the arc $(\omega, \omega')$ leads to the new graph shown on the right. Based on step (2) in Definition 14, for the arc $(\omega', \omega_2)$, since $\omega \not\prec \omega_2$, a new arc $(\omega, \omega_2)$ is inserted; by contrast, for $(\omega', \omega'_2)$, since $\omega \prec \omega'_2$, no new arc is inserted. Likewise, based on step (3), for the arc $(\omega_1, \omega)$, since $\omega_1 \not\prec \omega'$, a new arc $(\omega_1, \omega')$ is inserted; by contrast, for $(\omega'_1, \omega)$, since $\omega'_1 \prec \omega'$, no new arc is inserted.

**Example 17** (Uncertain Observation)**.** Consider the observation graph on the left of Fig. 7 (cf. Example 16). Five temporal relaxations are applied, which remove in cascade the arcs $(3, 4)$, $(3, 5)$, $(2, 4)$, $(2, 5)$, and $(1, 4)$, thereby leading to the uncertain observation $\mathcal{U}$ shown on the right of Fig. 7.

Among the temporal observations embedded in an uncertain observation $\mathcal{U}$ derived from a temporal observation $\mathcal{O}$ is $\mathcal{O}$ itself. Intuitively, $\mathcal{O}$ is still in $\mathcal{U}$, but in the company of other temporal observations (Proposition 4).

**Proposition 4.** *If $\mathcal{U}$ is an uncertain observation derived from a temporal observation $\mathcal{O}$, then $\mathcal{O} \in \mathcal{U}^*$.*

**Proof** (*sketch*). By induction on the sequence of temporal relaxations performed on $\mathcal{G}$. (*Basis*) $\mathcal{O} \in \mathcal{G}^*$, where $\mathcal{G}$ is the graph of $\mathcal{O}$. (*Induction*) Let $\mathcal{U} = (\Omega, \mathcal{A})$ be the DAG obtained by applying zero or more temporal relaxations on $\mathcal{G}$, where $\mathcal{O} \in \mathcal{U}^*$ is obtained from a topological sort $Q$ of $\Omega$ based on $\mathcal{A}$. Let $\mathcal{U}' = (\Omega, \mathcal{A}')$ be obtained by one relaxation on $\mathcal{U}$. Since a temporal relaxation only removes a temporal precedence, without adding new precedences, $Q$ is still a topological sort of $\Omega$ based on $\mathcal{A}'$. Hence, $\mathcal{O} \in \mathcal{U}'^*$. □

The notion of candidate set defined in Definition 8 can be extended to an uncertain observation.

**Definition 15** (**Candidate Set under Uncertainty**)**.** *Let $\mathcal{U}$ be an uncertain observation of $\mathcal{X}$ derived from a temporal observation $\mathcal{O}$. The* candidate set *of $\mathcal{U}$ is*

$$\Delta(\mathcal{U}) = \{\delta(T) \mid T \in Bsp(\mathcal{X}), Sgn^*(T) \cap \mathcal{U}^* \neq \emptyset\}\,. \quad (18)$$

Compared with Definition 8, eqn. (18) substitutes the condition $Sgn^*(T) \cap \mathcal{U}^* \neq \emptyset$ for $\mathcal{O} \in Sgn^*(T)$ in eqn. (7), as a consequence of the temporal relaxations moving $\mathcal{O}$ to $\mathcal{U}$. Still, based on Proposition 4, we have $\mathcal{O} \in \mathcal{U}^*$. Hence, if $\delta(T) \in \Delta(\mathcal{O})$ in eqn. (7), then $\delta(T) \in \Delta(\mathcal{U})$ in eqn. (18). The problem is now to compute $\Delta(\mathcal{U})$, which translates to the problem of extending the notion of abduction of a temporal observation $\mathcal{O}$ introduced in Definition 12, namely $Abd(\mathcal{O})$, to an uncertain observation, namely $Abd(\mathcal{U})$. To this end, we need to envisage a technique that allows for the indexing of $\mathcal{U}$. In fact, the index of $\mathcal{O} = [o_1, \ldots, o_n]$ is a natural number $\Im \in [0\,..\,n]$, the second field of an abduction state $(\mu, \Im)$. Consequently, in order to perform the indexing of $\mathcal{U}$, we also need to change the nature of the index $\Im$.

**Definition 16** (**Indexing**)**.** *Let $\mathcal{U} = (\Omega, \mathcal{A})$ be an uncertain observation, where $\mathbf{O}$ is the set of observations marking the nodes in $\Omega$. Let $\mathcal{N} = (\mathbf{O}, S, s_0, \tau, s_\mathrm{f})$ be an NFA where $S \subseteq 2^\Omega$ is the set of states, $s_0 = \emptyset$ is the initial state, $s_\mathrm{f} = \Omega$ is the (unique) final state, and $\tau : S \times \mathbf{O} \mapsto 2^S$ is the transition function, where $s' \in \tau(s, o')$ iff $s' = s \cup \{\omega'\}$, $\omega'$ is a node marked with $o'$, and $\forall(\omega, \omega') \in \mathcal{A}$, $\omega \in s$. The* indexing *of $\mathcal{U}$ is a DFA $Idx(\mathcal{U})$ obtained by determinization of $\mathcal{N}$.*

**Example 18** (Indexing)**.** Take the uncertain observation $\mathcal{U}$ in Fig. 7. Shown on the left of Fig. 8 is the NFA $\mathcal{N}$ involved in Definition 16, where the states are renamed $0\,..\,10$, with 10 being the final state. Next to it is $Idx(\mathcal{U})$, obtained by determinization of $\mathcal{N}$, where the states are renamed $\Im_0\,..\,\Im_9$, with $\Im_9$ incidentally being the unique final state.

The temporal observations embedded in an uncertain observation $\mathcal{U}$ equals the language of $Idx(\mathcal{U})$ (Proposition 5).

**Proposition 5.** *The language of $Idx(\mathcal{U})$ equals $\mathcal{U}^*$.*

**Proof** (*sketch*). Since $Idx(\mathcal{U})$ is generated by determinization of $\mathcal{N}$, the proof boils down to showing that the language of $\mathcal{N}$ equals $\mathcal{U}^*$. This can be proven by induction on the transition function of $\mathcal{N}$ starting from the initial state $s_0$. In fact, since each state of $\mathcal{N}$ is identified by a set of nodes of $\mathcal{U}$, the creation of a new transition $\langle s, o', s' \rangle$ is such that $s'$ is an extension of $s$ by a node $\omega'$ marked by $o'$ provided that, for each arc $(\omega, \omega')$ in $\mathcal{U}$, $\omega$ belongs to $s$. In other words, this condition allows for the generation of any temporal observation in $\mathcal{U}$ (soundness). Based on similar considerations, completeness is also true. □

The states of $Idx(\mathcal{U})$ are what we need for tracking $\mathcal{U}$ in the abductive reasoning aimed at computing $\Delta(\mathcal{U})$.

**Definition 17** (**Abduction under Uncertainty**)**.** *Let $\mathcal{M} = (\Sigma, M, \tau, \mu_0, M_\mathrm{f})$ be the diagnosis reference manual of $\mathcal{X}$, and let $\mathcal{U}$ be an uncertain observation of $\mathcal{X}$, where*

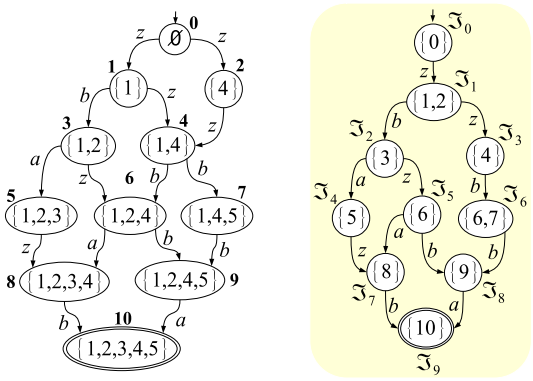

Figure 8: NFA $\mathcal{N}$ derived from $\mathcal{U}$ (left) and $Idx(\mathcal{U})$ (right).

$Idx(\mathcal{U}) = (\mathbf{O}, I, \tau_{\mathrm{i}}, \Im_0, I_{\mathrm{f}})$. *The* abduction *of $\mathcal{U}$ is a DFA $Abd(\mathcal{U}) = (\Sigma, A, \tau_{\mathrm{a}}, a_0, A_{\mathrm{f}})$, where $A \subseteq M \times I$ is the set of states, $a_0 = (\mu_0, \Im_0)$ is the initial state, $A_{\mathrm{f}} = \{a \mid a \in A, a = (\mu, \Im_{\mathrm{f}}), \mu \in M_{\mathrm{f}}, \Im_{\mathrm{f}} \in I_{\mathrm{f}}\}$ is the set of final states, and $\tau_{\mathrm{a}} : A \times \Sigma \mapsto A$ is the transition function, with $\tau_{\mathrm{a}}((\mu, \Im), O) = (\mu', \Im')$, $O = \{o_1, \ldots, o_k\}$, iff $\tau(\mu, O) = \mu'$ and there is a sequence $[o'_1, \ldots, o'_k]$, where $\{o'_1, \ldots, o'_k\} = O$, such that $\tau_{\mathrm{i}}(\Im, o'_1) = \Im_1$, $\tau_{\mathrm{i}}(\Im_1, o'_2) = \Im_2$, …, and $\tau_{\mathrm{i}}(\Im_{k-1}, o'_k) = \Im'$.*

**Example 19** (Abduction under Uncertainty)**.** Consider the reference manual $\mathcal{M}$ of $\mathcal{Z}$ outlined in Fig. 4 and the uncertain observation $\mathcal{U}$ displayed on the right of Fig. 7. Based on the indexing $Idx(\mathcal{U})$ in Fig. 8, shown in Fig. 9 is the generation of the abduction $Abd(\mathcal{U})$, where the dashed (gray) part is spurious. Compared with the abduction $Abd(\mathcal{O})$ in Fig. 5, where $\mathcal{O} = [z, b, z, a, b]$, which is consistent with the trajectory $[\{z\}, \{b\}, \{a, z\}, \{b\}]$ in $\mathcal{M}$, it is no surprise that $Abd(\mathcal{U})$ includes an extra trajectory $[\{z\}, \{a, b\}, \{z\}, \{b\}]$, which is consistent with $\mathcal{O}' = [z, b, a, z, b]$, $\mathcal{O}' \in Idx(\mathcal{U})$, $\mathcal{O}' \neq \mathcal{O}$, leading to the final state $(11, \Im_9)$. This is due to the temporal relaxations embedded in $\mathcal{U}$, which preserve $\mathcal{O}$ while adding new embedded temporal observations.

Theorem 1, which allows for the generation of the candidate set $Cand(\mathcal{O})$ based on the abduction $Abd(\mathcal{O})$, can be naturally extended to uncertain observations (Theorem 2).

**Theorem 2.** *Let $\mathcal{M}$ be the diagnosis reference manual of $\mathcal{X}$ and let $\mathcal{U}$ be an uncertain observation of $\mathcal{X}$. The candidate set of $\mathcal{U}$ can be computed based on the final states $A_{\mathrm{f}}$ of the abduction $Abd(\mathcal{U})$, namely*

$$\Delta(\mathcal{U}) = \bigcup_{(\mu, \Im) \in A_{\mathrm{f}}} \Delta(\mu). \tag{19}$$

**Proof** (*sketch*). The proof is a variant of the proof of Theorem 1. Let $\Delta_a = \bigcup_{(\mu, \Im) \in A_{\mathrm{f}}} \Delta(\mu)$. Roughly, if $\bar{\delta} \in \Delta(\mathcal{U})$, then there is $T \in Bsp(\mathcal{X})$, where $Sgn^*(T) \cap \mathcal{U}^* \neq \emptyset$. Let $\mathcal{O} \in Sgn^*(T) \cap \mathcal{U}^*$ and let $(\mu, \Im_f)$ be the final state in $Abd(\mathcal{U})$ reached by matching $\mathcal{O}$. We have $\bar{\delta} \in \Delta(\mu)$. If $\bar{\delta} \in \Delta_a$, then there is a state $(\mu, \Im_f)$ in $Abd(\mathcal{U})$, reached by matching a temporal observation $\mathcal{O} \in \mathcal{U}^*$, such that there is $T \in Bsp(\mathcal{X})$ where $\mathcal{O} \in Sgn^*(T)$. Hence, $\bar{\delta} \in \Delta(\mathcal{U})$. □

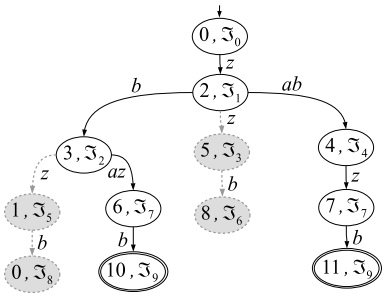

Figure 9: Abduction $Abd(\mathcal{U})$ for $\mathcal{Z}$, based on the reference manual outlined in Fig. 4 and the indexing $Idx(\mathcal{U})$ shown in Fig. 8 (right).

**Example 20** (Theorem 2)**.** Considering $Abd(\mathcal{U})$ in Fig. 9, with final states $(10, \Im_9)$ and $(11, \Im_9)$, we have $\Delta(11) = \Delta(10) = \{\{z\}\}$. Hence, $\Delta(\mathcal{U}) = \{\{z\}\}$. Compared with $Abd(\mathcal{O})$ in Fig. 5, incidentally, the extra trajectory $[\{z\}, \{a, b\}, \{z\}, \{b\}]$ in $Abd(\mathcal{U})$ yields no extra candidates.

## 7 Conclusion

Although being key to diagnosing DESs, observability has received little attention in the literature. This is why a notion of abstract observability has been proposed in the current paper, along with a diagnosis technique for a class of DESs to which this notion can be applied. Several application domains can be envisaged for the task of diagnosis with abstract observability, including networks of smart sensors and Internet of Things. Notably, online diagnosis with abstract observations benefit from compiled knowledge. To this end, we have adopted the most complete knowledge compilation approach, as total knowledge compilation brings to an efficient online matching of any given temporal observation against the diagnosis reference manual. However, total knowledge compilation requires heavy (possibly impractical) offline processing, as scalability is invariably an issue. This is why alternative online techniques that exploit lighter knowledge compilation are to be investigated in the future, including a minimal one, which is meant to generate the watchers only, and a partial one, which produces upfront a limited 'core reference manual', to be extended later, if needed, taking inspiration from Bertoglio et al. (2019; 2020b). Further research paths include defining *diagnosability* of DESs (Sampath et al. 1995; Jiang et al. 2001; Su, Zanella, and Grastien 2016) with abstract observations and with abstract uncertain observations, integrating abstract observability with abstract abnormality (Lamperti and Zanella 2011; Lamperti and Zhao 2014), and adapting abstract observability to complex ASs (Lamperti and Quarenghi 2016; Lamperti, Zanella, and Zhao 2018a) and deep DESs (Lamperti, Zanella, and Zhao 2020).

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (grant number 61972360).

## References

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

Basile, F. 2014. Overview of fault diagnosis methods based on Petri net models. In *Proceedings of the 2014 European Control Conference, ECC 2014*, 2636–2642.

Bertoglio, N.; Lamperti, G.; Zanella, M.; and Zhao, X. 2020a. Explanatory diagnosis of discrete-event systems with temporal information and smart knowledge-compilation. In Calvanese, D.; Erdem, E.; and Thielsher, M., eds., *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020)*. IJCAI Organization. 130–140.

Bertoglio, N.; Lamperti, G.; Zanella, M.; and Zhao, X. 2020b. Temporal-fault diagnosis for critical-decision making in discrete-event systems. In Cristani, M.; Toro, C.; Zanni-Merk, C.; Howlett, R.; and Jain, L., eds., *Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020*, volume 176 of *Procedia Computer Science*. Elsevier. 521–530.

Bertoglio, N.; Lamperti, G.; and Zanella, M. 2019. Intelligent diagnosis of discrete-event systems with preprocessing of critical scenarios. In Czarnowski, I.; Howlett, R.; and Jain, L., eds., *Intelligent Decision Technologies 2019*, volume 142 of *Smart Innovation, Systems and Technologies*. Springer, Singapore. 109–121.

Brand, D., and Zafiropulo, P. 1983. On communicating finite-state machines. *Journal of the ACM* 30(2):323–342.

Cabasino, M. P.; Giua, A.; and Seatzu, C. 2010. Fault detection for discrete event systems using Petri nets with unobservable transitions. *Automatica* 46:1531–1539.

Cassandras, C., and Lafortune, S. 2008. *Introduction to Discrete Event Systems*. New York: Springer, second edition.

Cong, X.; Fanti, M.; Mangini, A.; and Li, Z. 2018. Decentralized diagnosis by Petri nets and integer linear programming. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 48(10):1689–1700.

Debouk, R.; Lafortune, S.; and Teneketzis, D. 2000. Coordinated decentralized protocols for failure diagnosis of discrete-event systems. *Journal of Discrete Event Dynamic Systems: Theory and Applications* 10(1–2):33–86.

Grastien, A.; Cordier, M.; and Largouët, C. 2005. Incremental diagnosis of discrete-event systems. In *Nineteenth International Joint Conference on Artificial Intelligence (IJCAI 2005)*, 1564–1565.

Grastien, A.; Haslum, P.; and Thiébaux, S. 2012. Conflict-based diagnosis of discrete event systems: theory and practice. In *Thirteenth International Conference on Knowledge Representation and Reasoning (KR 2012)*, 489–499. Rome, Italy: Association for the Advancement of Artificial Intelligence.

Hamscher, W.; Console, L.; and de Kleer, J., eds. 1992. *Readings in Model-Based Diagnosis*. San Mateo, CA: Morgan Kaufmann.

Hopcroft, J.; Motwani, R.; and Ullman, J. 2006. *Introduction to Automata Theory, Languages, and Computation*. Reading, MA: Addison-Wesley, third edition.

Jéron, T.; Marchand, H.; Pinchinat, S.; and Cordier, M. 2006. Supervision patterns in discrete event systems diagnosis. In *Workshop on Discrete Event Systems (WODES 2006)*, 262–268. Ann Arbor, MI: IEEE Computer Society.

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

Jiroveanu, G.; Boel, R.; and Bordbar, B. 2008. On-line monitoring of large Petri net models under partial observation. *Journal of Discrete Event Dynamic Systems* 18:323–354.

Kan John, P., and Grastien, A. 2008. Local consistency and junction tree for diagnosis of discrete-event systems. In *Eighteenth European Conference on Artificial Intelligence (ECAI 2008)*, 209–213. Patras, Greece: IOS Press, Amsterdam.

Kwong, R., and Yonge-Mallo, D. 2011. Fault diagnosis in discrete-event systems: incomplete models and learning. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics* 41(1):118–130.

Lamperti, G., and Quarenghi, G. 2016. Intelligent monitoring of complex discrete-event systems. In Czarnowski, I.; Caballero, A.; Howlett, R.; and Jain, L., eds., *Intelligent Decision Technologies 2016*, volume 56 of *Smart Innovation, Systems and Technologies*. Springer International Publishing Switzerland. 215–229.

Lamperti, G., and Zanella, M. 2002. Diagnosis of discrete-event systems from uncertain temporal observations. *Artificial Intelligence* 137(1–2):91–163.

Lamperti, G., and Zanella, M. 2006. Flexible diagnosis of discrete-event systems by similarity-based reasoning techniques. *Artificial Intelligence* 170(3):232–297.

Lamperti, G., and Zanella, M. 2011. Context-sensitive diagnosis of discrete-event systems. In Walsh, T., ed., *Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011)*, volume 2, 969–975. Barcelona, Spain: AAAI Press.

Lamperti, G., and Zhao, X. 2014. Diagnosis of active systems by semantic patterns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 44(8):1028–1043.

Lamperti, G.; Zanella, M.; and Zhao, X. 2018a. Abductive diagnosis of complex active systems with compiled knowledge. In Thielscher, M.; Toni, F.; and Wolter, F., eds., *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference (KR 2018)*, 464–473. Tempe, Arizona: AAAI Press.

Lamperti, G.; Zanella, M.; and Zhao, X. 2018b. *Introduction to Diagnosis of Active Systems*. Springer, Cham.

Lamperti, G.; Zanella, M.; and Zhao, X. 2020. Diagnosis of deep discrete-event systems. *Journal of Artificial Intelligence Research* 69:1473–1532.

McIlraith, S. 1998. Explanatory diagnosis: conjecturing actions to explain observations. In *Sixth International Confer-*

*ence on Principles of Knowledge Representation and Reasoning (KR 1998)*, 167–177. Trento, I: Morgan Kaufmann, S. Francisco, CA.

Pencolé, Y., and Cordier, M. 2005. A formal framework for the decentralized diagnosis of large scale discrete event systems and its application to telecommunication networks. *Artificial Intelligence* 164(1–2):121–170.

Pencolé, Y.; Steinbauer, G.; Mühlbacher, C.; and Travé-Massuyès, L. 2018. Diagnosing discrete event systems using nominal models only. In Zanella, M.; Pill, I.; and Cimatti, A., eds., *28th International Workshop on Principles of Diagnosis (DX'17)*, volume 4, 169–183. Kalpa Publications in Computing.

Reiter, R. 1987. A theory of diagnosis from first principles. *Artificial Intelligence* 32(1):57–95.

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

Sampath, M.; Sengupta, R.; Lafortune, S.; Sinnamohideen, K.; and Teneketzis, D. 1996. Failure diagnosis using discrete-event models. *IEEE Transactions on Control Systems Technology* 4(2):105–124.

Struss, P. 1997. Fundamentals of model-based diagnosis of dynamic systems. In *Fifteenth International Joint Conference on Artificial Intelligence (IJCAI 1997)*, 480–485.

Su, X.; Zanella, M.; and Grastien, A. 2016. Diagnosability of discrete-event systems with uncertain observations. In *25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, 1265–1571.