2021 61st IEEE Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Fault Diagnosis of Discrete-Event Systems under Non-Deterministic Observations with Output Fairness

Weijie Dong, Shang Gao, Xiang Yin and Shaoyuan Li

Abstract— **In this paper, we revisit the fault diagnosis problem of discrete-event systems (DES) under non-deterministic observations. Non-deterministic observation is a general observation model that includes the case of intermittent loss of observations. In this setting, upon the occurrence of an event, the sensor reading may be non-deterministic such that a set of output symbols are all possible. Existing works on fault diagnosis under non-deterministic observations require to consider all possible observation realizations. However, this approach includes the case where some possible outputs are permanently disabled. In this work, we introduce the concept of *output fairness* by requiring that, for any output symbols, if it has infinite chances to be generated, then it will indeed be generated infinite number of times. We use an assume-guarantee type of linear temporal logic formulas to formally describe this assumption. A new notion called output-fair diagnosability (OF-diagnosability) is proposed. An effective approach is provided for the verification of OF-diagnosability. We show that the proposed notion of OF-diagnosability is weaker than the standard definition of diagnosability under non-deterministic observations, and it better captures the physical scenario of observation non-determinism or intermittent loss of observations.**

I. Introduction

Engineering cyber-physical systems (CPS), such as manufacturing systems, transportation systems and power systems, are generally very complex due to their intricate operation logic and hybrid dynamics. For such large-scale safety-critical systems, failures during their operations are very common as millions of components/modules are working in parallel. Thefore, fault diagnosis and detection is crucial but challenging task in order to monitor the operation conditions and to maintain safety for CPSs. In this work, we investigate the fault diagnosis problem in the framework of discrete-event systems (DES) [5]–[7], [9]–[12], [15]–[19].

In the modeling of DES, the occurrences of events are essentially observed by the corresponding sensors. In practice, however, due to measurement noises, sensor failures or malicious attacks, the sensor readings can be *unreliable* or *non-deterministic*. Such non-deterministic observation issue was initially addressed in the context of *intermittent loss of observations* [3], [4], [8], where it is assumed that some observable events are unreliable in the sense that their occurrences may be observed or be lost non-deterministically. In [13], Takai and Ushio investigated the issue of non-deterministic observation using Mealy automata. Specifically, the observation of the system is modeled by a *state-dependent non-deterministic output function*. The model is quite general that captures the issue of intermittent loss of observations. Furthermore, it allows the observation symbols to be different from the original event set.

Our work is motivated by the aforementioned works on fault diagnosis under intermittent loss of observations [4] or, in a broad sense, non-deterministic observations [13]. In particular, we note that the existing models for non-deterministic observations essentially assume that *all possible realizations under the non-deterministic output mapping are feasible*. In the context of fault diagnosis, therefore, it needs to consider whether or not fault can be detected under all possible observation realizations. We argue in this work that this setting is somehow too strict for the purpose of diagnosis since it may exclude the possibility for the occurrence of some possible sensor reading, which is "unfair".

It is worth noting that, in [2], the authors investigated diagnosability for *fair systems*. However, the notion of fairness is different from our setting. Specifically, [2] assumes that the dynamic of the system is fair in the sense that each transition can be executed for infinite number of times whenever it is enabled infinitely. However, the observation mapping considered therein is still static modeled as a natural projection. In contrast, we impose fairness constraint on the observation mapping rather than the internal behavior of the system the system's dynamic. Therefore, the problem setting in our work is quite different from that of [2].

In this paper, we revisit the fault diagnosis problem under non-deterministic observations. Motivated by the above discussion, however, to better capture the physical scenario in which *all possible observations are always available*, we further require that each possible observation (or output symbol) is *fair* in the sense that if it has infinite chances to be observed then it will indeed be observed infinite number of times. To formally describe this fairness requirement, we use an assume-guarantee type linear temporal logic (LTL) formulas. Then, we propose a new condition called *output-fair diagnosability* (OF-diagnosability) that provides the necessary and sufficient condition for the existence of a diagnoser that works correctly under the fair-observation setting. Also, we utilize the structure properties of fairness requirement to provide an more effective procedure for checking this new condition. Our work bridges the gap between the existing mathematical definition of diagnosability under intermittent loss of observations and the physical setting, where those

This work was supported by the National Natural Science Foundation of China (62061136004, 61803259, 61833012) and the National Key Research and Development Program of China (2018AAA0101700).

W. Dong S. Gao, X. Yin and S. Li are with Department of Automation and Key Laboratory of System Control and Information Processing, Shanghai Jiao Tong University, Shanghai 200240, China. E-mail: {wjd_dollar,gaoshang,yinxiang,syli}@sjtu.edu.cn.

unreliable sensors will not fail permanently.

II. Preliminary

A. System Model

Let Σ be a finite set of events. A finite (infinite) string over Σ is a finite (infinite) sequence of events in Σ of form $s = \sigma_1\sigma_2\ldots\sigma_n(\ldots)$, where $\sigma_i \in \Sigma$, for $i = 1, \cdots, n$. A finite (infinite) language is a set of finite (infinite) strings. We denote by Σ^* and Σ^ω, respectively, the set of all finite and the set of all infinite strings over Σ. Note that the empty string, denoted by ε, is included in Σ^*. Given a finite language $L \subseteq \Sigma^*$, the prefix-closure of L is defined by $\mathsf{Pre}(L) = \{w \in \Sigma^* : \exists t \in \Sigma^* \text{ s.t. } wt \in L\}$. Similarly, for an infinite language $L \subseteq \Sigma^\omega$, its prefix-closure is the set of all its finite prefixes, i.e., $\mathsf{Pre}(L) = \{w \in \Sigma^* : \exists t \in \Sigma^\omega \text{ s.t. } wt \in L\}$. For any string $s \in \Sigma^* \cup \Sigma^\omega$, we write $\mathsf{Pre}(\{s\})$ as $\mathsf{Pre}(s)$ for the sake of simplicity. Given an infinite string $s \in \Sigma^\omega$, $\mathsf{Inf}(s)$ denotes the set of events that appear infinitely number of times in the string.

In this paper, we consider DES modeled by a deterministic finite-state automaton (DFA) $G = (Q, \Sigma, f, q_0)$, where Q is the finite set of states; Σ is the finite set of events; $f : Q \times \Sigma \to Q$ is the partial transition function; and $q_0 \in Q$ is the initial state. The transition function f can be extended to $f : Q \times \Sigma^* \to Q$ recursively in the usual manner. The finite language generated by G is defined by $\mathcal{L}(G) = \{s \in \Sigma^* : f(q_0, s)!\}$, and the infinite language generated by G is defined by $\mathcal{L}^\omega(G) = \{s \in \Sigma^\omega : f(q_0, s)!\}$. We assume that system G is live, i.e., $\forall q \in Q, \exists \sigma : f(q, \sigma)!$.

B. Non-Deterministic Observations

In the partial observation setting, it is assumed that the occurrence of each event cannot be observed directly or perfectly. A commonly adopted simple approach is to partition the event set Σ into observable events Σ_o and unobservable events Σ_{uo}. Then natural projection $P : \Sigma^* \to \Sigma_o^*$ can be used to capture the issue of partial observability.

In many real-world scenarios, however, the sensor readings may be non-deterministic due to observation noises, sensor failures or malicious attacks. Furthermore, the sensor reading of the event occurrence can be *state-dependent*. Here, we adopt state-dependent non-deterministic observation model proposed by [13], [14], which captures both state-dependency and non-determinism of observations.

Formally, we assume that Δ is a new set of all possible observations or *output symbols*. Then a state-dependent non-deterministic output function is

$$\mathcal{O} : Q \times \Sigma \to 2^{\Delta_\varepsilon},$$

where $\Delta_\varepsilon = \Delta \cup \{\varepsilon\}$. The output function means that if event $\sigma \in \Sigma$ occurs at state $q \in Q$, then the system is possible to observe any symbol in $\mathcal{O}(q, \sigma)$ non-deterministically.

Remark 1: The above non-deterministic observation model is quite general in the sense it subsumes many observation models in the literature. For example, the standard natural-projection-based observation can be formulated by setting $\mathcal{O}(q, \sigma) = \{\sigma\}$ for all $\sigma \in \Sigma_o$ and $\mathcal{O}(q, \sigma) = \{\varepsilon\}$ for all $\sigma \in \Sigma_{uo}$. Also, it captures the so-called *intermittent loss of observations* [4]. In this setting, the event set is usually partitioned as $\Sigma = \Sigma_r \dot\cup \Sigma_{ur} \dot\cup \Sigma_{uo}$, where Σ_r is the set of reliable events whose occurrences can always be observed directly, Σ_{ur} is the set of unreliable events whose occurrences may be observed but can also be lost and Σ_{uo} is the set of unobservable events whose occurrences can never be observed. This setting can be captured by considering a non-deterministic output function with $\Delta = \Sigma$ and for any $q \in Q$ and $\sigma \in \Sigma$, we have

$$\mathcal{O}(q, \sigma) = \begin{cases} \{\sigma\} & \text{if } \sigma \in \Sigma_r \\ \{\sigma, \varepsilon\} & \text{if } \sigma \in \Sigma_{ur} \\ \{\varepsilon\} & \text{if } \sigma \in \Sigma_{uo} \end{cases}$$

Note that, for the general case we consider here, the output symbols Δ can be different from the original event set Σ.

Based on the output function $\mathcal{O} : Q \times \Sigma \to 2^{\Delta_\varepsilon}$, we can define a non-deterministic observation mapping $M : \mathcal{L}(G) \to 2^{\Delta^*}$, where Δ^* is the set of finite strings over Δ and we have $\varepsilon \in \Delta^*$, recursively as:

- $M(\varepsilon) = \{\varepsilon\}$;
- for any $s \in \Sigma^*$ and $\sigma \in \Sigma$, we have

$$M(s\sigma) = \{\alpha\beta \in \Delta^* : \alpha \in M(s) \wedge \beta \in \mathcal{O}(f(q_0, s), \sigma)\}$$

Intuitively, $M(s) \in \Delta^*$ is the set of all possible observations (or output strings) upon the occurrence of internal string $s \in \Sigma^*$. The observation mapping is also extended to $M : 2^{\Sigma^*} \to 2^{\Delta^*}$ by: for any language $L \subseteq \mathcal{L}(G)$, $M(L) = \cup_{s \in L} M(s)$.

Since the observation is non-deterministic, for any specific internal string $s \in \mathcal{L}(G)$, it may have different *output realizations*. To "embed" the actual observation occurred into the internal execution of the system, we define

$$\Sigma_e = Q \times \Sigma \times \Delta_\varepsilon$$

as the set of *extended events*. Then an extended string is a finite or infinite sequence of extended events. We say a finite extended string

$$s = (q_0, \sigma_0, \delta_0)(q_1, \sigma_1, \delta_1)\ldots(q_n, \sigma_n, \delta_n) \in \Sigma_e^*$$

is generated by G if $f(q_i, \sigma_i) = q_{i+1}, \forall i < n$ and $\delta_i \in \mathcal{O}(q_i, \sigma_i), \forall i \leq n$. We denote by $\mathcal{L}_e(G)$ the set of all finite extended strings generated by G. The infinite extended strings are defined analogously and the set of all infinite extended strings generated by G is denoted by $\mathcal{L}_e^\omega(G)$.

For any extended string $s = (q_0, \sigma_0, \delta_0)(q_1, \sigma_1, \delta_1)\cdots$, we define $\Theta_Q(s) = q_0 q_1 \cdots, \Theta_\Sigma(s) = \sigma_0\sigma_1\cdots$ and $\Theta_\Delta(s) = \delta_0\delta_1\cdots$ as its corresponding state sequence, (internal) event string and output string, respectively. Clearly, for any $s \in \mathcal{L}_e(G)$, we have $\Theta_\Delta(s) \in M(\Theta_\Sigma(s))$ because $\Theta_\Delta(s)$ is a specific observation realization of internal string $\Theta_\Sigma(s)$.

Example 1: Let us consider system G_1 in Figure 1, where we have $\Sigma = \{a, b, c, \sigma_f, u\}$ and $\Delta = \{a, b, c\}$. The output function $\mathcal{O} : Q \times \Sigma \to 2^{\Delta_\varepsilon}$ is specified by the label of each transition, where the LHS of "\" denotes the internal event and the RHS of "\" denotes the set of all possible output symbols. For example, $b\backslash\{b, \varepsilon\}$ from state 3 to state 4 means

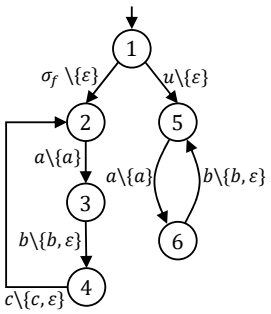


Fig. 1: System G_1.

that system moves to state 4 from state 3 by firing event b, i.e., $f(3,b) = 4$, and we can non-deterministically observe any output in set $\mathcal{O}(3,b) = \{b, \varepsilon\}$. In this example, we either observe b itself or observe nothing, which corresponds to the case of intermittent loss of observations. Then for finite string $\sigma_f ab \in \mathcal{L}(G_1)$, the set of all possible output is $M(\sigma_f ab) = \{ab, a\}$. These two different outputs lead to the following two extended strings $(1,\sigma_f,\varepsilon)(2,a,a)(3,b,b)$ and $(1,\sigma_f,\varepsilon)(2,a,a)(3,b,\varepsilon)$, respectively.

C. Fault Diagnosis

In this work, we assume that the system is subject to faults whose occurrences need to be diagnosed within a finite number of steps. Specifically, we assume that $\Sigma_F \subset \Sigma$ is the set of fault events. For the sake of simplicity, we do not distinguish among different fault types. We say a string $s \in \Sigma^* \cup \Sigma^\omega$ is faulty if some event in Σ_F appears in s and we write $\Sigma_F \in s$ with a slight abuse of notation; otherwise, it is normal. We define $\mathcal{L}_F(G)$ and $\mathcal{L}_F^\omega(G)$ as the sets of all finite and infinite faulty strings generated by G, respectively. Similarly, we define $\Sigma_{e,F} = Q \times \Sigma_F \times \Delta^\varepsilon$ as the set of extended fault events and denote by $\mathcal{L}_{e,F}(G) = \{s \in \mathcal{L}_e(G) : \Sigma_{e,F} \in s\}$ the set of all finite extended faulty strings generated by G; the extended infinite faulty language $\mathcal{L}_{e,F}^\omega(G)$ is also defined analogously. Finally, we define $\Psi_e(G)$ as the set of all finite extended faulty strings in which fault events occur *for the first time*, i.e.,

$$\Psi_e(G) = \{s \in \mathcal{L}_{e,F}(G) : \forall t \in \mathsf{Pre}(s) \setminus \{s\}, \Sigma_{e,F} \notin t\}.$$

To capture whether or not the occurrences of fault events can be detected within a finite number of steps, the notion of diagnosability (under non-deterministic observations) has been proposed in the literature [13].

Definition 1: System G is said to be *diagnosable* w.r.t. output function $\mathcal{O}$ and fault events Σ_F if

$$(\forall s \in \Psi_e(G))(\exists n \in \mathbb{N})(\forall st \in \mathcal{L}_e(G))[|t| \geq n \Rightarrow \mathsf{diag}] \quad (1)$$

where the diagnostic condition diag is

$$(\forall \omega \in \mathcal{L}_e(G))[\Theta_\Delta(\omega) = \Theta_\Delta(st) \Rightarrow \Sigma_{e,F} \in \omega].$$

Intuitively, the above definition says that, for any faulty extended string in which fault events appear for the first time, there exists a finite detection bound such that, for any of its continuation longer than the detection bound, any other extended strings having the same observation must also contain fault events. Note that we consider extended strings rather than the internal strings in order to capture the issue of non-deterministic observations.

Example 2: Again, we consider system G_1 depicted in Figure 1 with $\Sigma = \{a,b,c,\sigma_f,u\}, \Delta = \{a,b,c\}$ and $\Sigma_F = \{\sigma_f\}$. Let us consider faulty extended string $(1,\sigma_f,\varepsilon) \in \Psi_e(G_1)$, which can be extended arbitrarily long as $s_F = (1,\sigma_f,\varepsilon)[(2,a,a)(3,b,b)(4,c,\varepsilon)]^n$. However, for any n, we can find a normal extended string $s_N = (1,u,\varepsilon)[(5,a,a)(6,b,b)]^n$ such that $\Theta_\Delta(s_F) = \Theta_\Delta(s_N) = (ab)^n$. Therefore, system G_1 is not diagnosable under $\mathcal{O}$.

III. Diagnosability with Always-Fairness Assumption

A. Motivating Example

As we discussed in Remark 1, the non-deterministic output function can be used to capture the issue of *intermittent loss of observations*. Here, however, we argue that the original definition of diagnosability in Definition 1 may be too strong for the case of intermittent loss of observations.

To see this, we still consider system G_1 shown in Figure 1, where the reliable event set is $\{a\}$, the unreliable events set is $\{b,c\}$ and the unobservable event set is $\{\sigma_f, u\}$. As we have discussed in Example 2, this system is not diagnosable because there are two infinite extended strings:

- one is faulty $s_F = (1,\sigma_f,\varepsilon)[(2,a,a)(3,b,b)(4,c,\varepsilon)]^\omega$;
- the other is non-faulty $s_N = (1,u,\varepsilon)[(5,a,a)(6,b,b)]^\omega$

and they have the same observation, i.e., $\Theta_\Delta(s_1) = \Theta_\Delta(s_2) = (ab)^\omega$. As a result, the occurrence of fault in string s_F can never be detected within a finite number of steps.

Note that, the existence of above extended string s_F is due to the following physical scenario: the infinite internal string $\sigma_f(abc)^\omega$ is executed, i.e., the system loops forever in the cycle formed by states $2, 3$ and 4, and each time when event c is executed at state 4, the sensor reads ε due to observation loss. In other words, the sensor corresponds to transition $4 \xrightarrow{c}$ has to fail *permanently* in order to draw the conclusion that the system is not diagnosable. However, this source of non-diagnosability seems violate the setting of intermittent loss of observations. In the context of intermittent loss of observations or non-deterministic observations, it makes more sense to assume that each possible observation is *eventually possible*. Therefore, if event c at state 4 corresponds to a sensor that may be unreliable but will not fail permanently, then along the infinite loop of $\sigma_f(abc)^\omega$, once one has a single chance to observe output c for transition $4 \xrightarrow{c}$, it will conclude immediately that fault events have occurred.

B. Δ'-Fairness Assumption

The above discussion suggests that the exiting definition of diagnosability under non-deterministic observations is a bit strong than its underlying physical setting because it includes the situation where some output symbols are disabled permanently. In order to bridge this discrepancy between the definition of diagnosability and the physical meaning of non-deterministic observations, in this work, we put an additional *fairness* assumption on those sensors

that are subject to intermittent loss of observations, or non-deterministic observations in a broad sense. Specifically, we assume that $\Delta' \subseteq \Delta$ is a set of "fair" output symbols in the sense that if they have infinite opportunities to occur, then they will indeed occur infinite number of times.

To formalize this fairness requirement, we use the linear temporal logic (LTL) formulas. Formally, an LTL formula φ is constructed based on a set of atomic propositions $\mathcal{AP}$, Boolean operators and temporal operators as follows [1]:

$$\varphi ::= true \mid p \mid \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \bigcirc \varphi \mid \varphi_1 U \varphi_2,$$

where $p \in \mathcal{AP}$ is an atomic proposition; $\bigcirc$ and U denote "next" and "until", respectively. Other Boolean connectives can be induced by $\wedge$ and $\neg$, e.g., $\varphi_1 \vee \varphi_2 = \neg(\neg\varphi_1 \wedge \neg\varphi_2)$ and $\varphi_1 \rightarrow \varphi_2 = \neg\varphi_1 \vee \varphi_2$. Temporal operators $\Box$ "always" and $\Diamond$ "eventually" can be induced by until operator, i.e., $\Diamond\varphi = trueU\varphi$ and $\Box\varphi = \neg\Diamond\neg\varphi$. LTL formulas are evaluated over infinite sequences of atomic proposition sets (called words). For any infinite word $s \in (2^{\mathcal{AP}})^\omega$, we denote by $s \models \varphi$ if it satisfies LTL formula φ.

In our context of non-deterministic observations, we choose extended events as atomic propositions, i.e., $\mathcal{AP} = \Sigma_e$. For the sake of simplicity, each extended event (q, σ, δ) will also be written as σ_q^δ meaning that event σ occurs at state q and generates observation δ. For simplicity, we define

$$\sigma_q := \bigvee_{\delta \in \mathcal{O}(\sigma, q)} \sigma_q^\delta$$

as the proposition formula meaning that transition $q \xrightarrow{\sigma}$ occurs no matter what output it generates. Then we introduce the notion of Δ'-fairness as follows.

Definition 2: (Δ'-Fairness) Let $\Delta' \subseteq \Delta_\varepsilon$ be a set of output symbols. We say an infinite extended string $s \in \mathcal{L}_e^\omega(G)$ is Δ'-fair if

$$s \models \varphi_{fair} := \bigwedge_{q \in Q, \sigma \in \Sigma} \left(\Box\Diamond\sigma_q \rightarrow \bigwedge_{\delta \in \Delta' \cap \mathcal{O}(q,\sigma)} \Box\Diamond\sigma_q^\delta \right) \quad (2)$$

The above subset of outputs $\Delta' \subseteq \Delta_\varepsilon$ is referred to as the *fair outputs* meaning that these outputs will always have the opportunity to be measured. Specifically, this requirement is captured by formula φ_{fair} saying that, for any transition $q \xrightarrow{\sigma}$, if it is fired infinite number of times, then any of its fair outputs, i.e., $\sigma \in \Delta' \cap \mathcal{O}(q, \sigma)$, can actually be observed infinite number of times. This excludes the case where some fair outputs are disabled permanently. We define

$$\mathcal{L}_e^\varphi(G) = \{s \in \mathcal{L}_e^\omega(G) : s \models \varphi_{fair}\}$$

as the set of infinite extended strings generated by G satisfying φ_{fair}. We also define $\mathcal{L}_{e,F}^\varphi(G) = \mathcal{L}_e^\varphi(G) \cap \mathcal{L}_{e,F}^\omega(G)$ as the set of infinite faulty extended strings satisfying φ_{fair}.

C. OF-Diagnosability

Based on the above notion of Δ'-fairness on extended strings, now we modify the existing definition of diagnosability as shown in Definition 1 to the *output-fair diagnosability* (OF-diagnosability) defined as follows.

Definition 3: (OF-diagnosability) System G is said to be *output-fairly diagnosable* (OF-diagnosable) w.r.t. fault events Σ_F, output function $\mathcal{O}$ and fair outputs $\Delta' \subseteq \Delta_\varepsilon$ if

$$(\forall s \in \mathcal{L}_{e,F}^\varphi(G))(\exists t \in \mathsf{Pre}(s))[\textsf{fair-diag}] \quad (3)$$

where the fair-diagnostic condition fair-diag is

$$(\forall w \in \mathcal{L}_e(G))[\Theta_\Delta(w) = \Theta_\Delta(t) \Rightarrow \Sigma_{e,F} \in w].$$

Intuitively, OF-diagnosability revises the standard diagnosability by restricting our attention only to those infinite extended strings satisfying the fairness assumption and investigates whether or not faults in those "output-fair" strings can be detected.

Remark 2: In Definition 1, it is known that "$\forall s \in \Psi_e(G)$" and "$\exists n \in \mathbb{N}$" can be swapped, which means that if the system is diagnosable, then there exists a uniform detection bound after the occurrence of any fault events. However, in Definition 3, the length of detection prefix t can be arbitrarily long, depending on how the fairness assumption is satisfied in the specific infinite faulty string $s \in \mathcal{L}_{e,F}^\varphi(G)$, since the Δ'-fairness assumption only guarantees that all fair outputs *eventually* occur but the delay can be arbitrarily large.

We show that the proposed notion of a OF-diagnosability provides the necessary and sufficient condition for the existence of diagnoser that works "correctly" under the Δ'-fairness assumption. Formally, a diagnoser is a function

$$D : M(\mathcal{L}(G)) \rightarrow \{0, 1\}$$

that decides whether a fault has happened (by issuing "1") or not (by issuing "0") based on the output string. We say that a diagnoser works correctly under the Δ'-fairness assumption if it satisfies the following conditions:

C1) By assuming that each fair-output will actually be observed infinitely if they have infinite chances to be observed, the diagnoser will issue a fault alarm for any occurrence of fault events, i.e.,

$$(\forall s \in \mathcal{L}_{e,F}^\varphi(G))(\exists s' \in \mathsf{Pre}(s)))[D(\Theta_\Delta(s')) = 1].$$

C2) The diagnoser will not issue a false alarm if the execution is still normal, i.e.,

$$(\forall s \in \mathcal{L}_e(G) : \Sigma_{e,F} \notin s)[D(\Theta_\Delta(s)) = 0].$$

The following theorem says that there exists a diagnoser working "correctly" under the Δ'-fairness assumption if and only if the system is OF-diagnosable.

Theorem 1: There exists a diagnoser satisfying conditions C1 and C2 if and only if G is OF-diagnosable w.r.t. fault events Σ_F, output function $\mathcal{O}$ and fair outputs $\Delta' \subseteq \Delta_\varepsilon$.

We use the following examples to illustrate the concept of output fairness as well as the notion of OF-diagnosability.

Example 3: Again, let us consider system G_1 shown in Figure 1 with $\Sigma_F = \{\sigma_f\}$ and suppose that the fair outputs are $\Delta' = \{b, c\}$. Then by Definition 2, the Δ'-fairness assumption is

$$\varphi_{fair} = (\bigwedge_{i=3,6} \Box\Diamond b_i \rightarrow \Box\Diamond b_i^b) \wedge (\Box\Diamond c_4 \rightarrow \Box\Diamond c_4^c)$$

Note that for infinite faulty extended string

$$s_F = (1, \sigma_f, \varepsilon)[(2, a, a)(3, b, b)(4, c, \varepsilon)]^\omega,$$

we have $s_F \not\models \varphi_{fair}$ because $\square\lozenge c_4$ holds but $\square\lozenge c_4^c$ does not hold. Therefore, $s_F \notin \mathcal{L}^{\varphi}_{e,F}(G)$ is not considered in the analysis of OF-diagnosability. Clearly, for any faulty string in $\mathcal{L}^{\varphi}_{e,F}(G)$, c_4^c must occur, which means output c will be observed and upon the occurrence of which the fault can be detected. Therefore, this system is actually OF-diagnosable.

IV. Verification of OF-diagnosability

A. Augmented System

To verify OF-diagnosability, our first step is to augment both the state-space and the event-space of G such that

- the information of whether or not a fault event has occurred is encoded in the augmented state-space; and
- the information of where the event occurs and which specific output is observed are encoded in the augmented event-space.

Formally, given system $G = (Q, q_0, \Sigma, f)$, fault events Σ_F and output function $\mathcal{O}$, we define the *augmented system* as a new DFA

$$\tilde{G} = (\tilde{Q}, \tilde{q}_0, \Sigma_e, \tilde{f}), \tag{4}$$

where

- $\tilde{Q} \subseteq Q \times \{F, N\}$ is the set of augmented states;
- $\tilde{q}_0 = (q_0, N)$ is the initial augmented state;
- Σ_e is the set of augmented events, which are just extended events;
- $\tilde{f} : \tilde{Q} \times \Sigma_e \to \tilde{Q}$ is the transition function defined by: for any $\tilde{q} = (q, l) \in \tilde{Q}$ and $\tilde{\sigma} = (q, \sigma, \delta) \in \Sigma_e$, we have $\tilde{f}(\tilde{q}, \tilde{\sigma})!$ if $f(q, \sigma)!$ and $\delta \in \mathcal{O}(q, \sigma)$. Furthermore, when $\tilde{f}(\tilde{q}, \tilde{\sigma})!$, we have

$$\tilde{f}(\tilde{q}, \tilde{\sigma}) = \begin{cases} (f(q, \sigma), N) & \text{if } l = N \wedge \tilde{\sigma} \notin \Sigma_{e,F} \\ (f(q, \sigma), F) & \text{otherwise} \end{cases}.$$

The above constructed augmented system $\tilde{G}$ has the following properties:

- First, the augmented system $\tilde{G}$ generates extended strings. Essentially, it still tracks the original dynamic of the system by putting both the output realization and the current state information together with the internal event. Therefore, we have

$$\mathcal{L}(\tilde{G}) = \mathcal{L}_e(G) \text{ and } \mathcal{L}^{\omega}(\tilde{G}) = \mathcal{L}^{\omega}_e(G).$$

- Second, each augmented state $(q, l) \in Q \times \{F, N\} = \tilde{Q}$ has two components. The first component q is the actual state in the original system G and the second component $l \in \{N, F\}$ is a label denoting whether fault events have occurred. By the construction, the label will change from N to F only when an extended fault event occurs and once the label becomes F, it will be F forever. We denote by $\tilde{Q}_N = \{(q, l) \in \tilde{Q} : l = N\}$ the set of normal augmented states and the set of faulty states $\tilde{Q}_F$ is defined analogously.

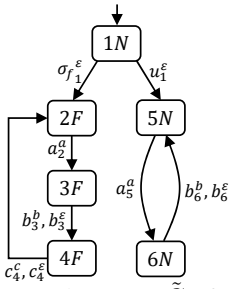


Fig. 2: Augmented system $\tilde{G}_1$ for system G_1.

Example 4: Still, we consider system G_1 shown in Figure 1 with the same setting in Example 3. Its augmented system $\tilde{G}_1$ is depicted in Figure 2, where $(1, \sigma_f, \varepsilon)$ is the extended fault event and after occurrence of which all states is augmented with label F, e.g., states $2F$ and $3F$. Furthermore, the transitions of the augmented system $\tilde{G}_1$ are defined according to the actual transitions and the underlying observations of original system G_1. For example, because $f(3, b) = 4$ and $\mathcal{O}(3, b) = \{b, \varepsilon\}$, we have two new transitions in $\tilde{G}_1$: $\tilde{f}(3F, b_3^b) = 4F$ and $\tilde{f}(3F, b_3^{\varepsilon}) = 4F$.

B. Verification Structure

Let $r \in 2^{\tilde{Q}}$ be a set of states representing the current state estimation of the augmented system. By observing a new output symbol $\delta \in \Delta$, the estimate is updated by the following observable reach

$$\mathsf{Next}(r, \delta) = \left\{ \tilde{q}' \in \tilde{Q} : \begin{array}{c} \exists \tilde{q} \in r, \tilde{\sigma} \in \Sigma_e \text{ s.t} \\ \Theta_\Delta(\tilde{\sigma}) = \delta \wedge \tilde{q}' = \tilde{f}(\tilde{q}, \tilde{\sigma}) \end{array} \right\}.$$

Also, the system can execute silently without generating an output symbol. This is captured by the unobservable reach defined by:

$$\mathsf{UR}(r) = \left\{ \tilde{q}' \in \tilde{Q} : \begin{array}{c} \exists \tilde{q} \in r, s \in \Sigma_e^* \text{ s.t} \\ \Theta_\Delta(s) = \varepsilon \wedge \tilde{q}' = \tilde{f}(\tilde{q}, s) \end{array} \right\}.$$

Now, based on the augmented system $\tilde{G} = (\tilde{Q}, \tilde{q}_0, \tilde{\Sigma}, \tilde{f})$, we construct the *verification structure* as follows:

$$V = (Q_V, q_V^0, \Sigma_V, f_V) \tag{5}$$

where

- $Q_V \subseteq \tilde{Q} \times 2^{\tilde{Q}}$ is the set of states;
- $q_V^0 = (\tilde{q}_0, \mathsf{UR}(\{\tilde{q}_0\}))$ is the initial state;
- $\Sigma_V = \Sigma_e$ is the event set, which is just the set of extended events;
- $f_V : Q_V \times \Sigma_V \to Q_V$ is the transition function defined by: for any $(\tilde{q}, r) \in \tilde{Q} \times 2^{\tilde{Q}}$ and $\tilde{\sigma} \in \Sigma_V$, we have

$$f_V((\tilde{q}, r), \tilde{\sigma}) = \begin{cases} (\tilde{q}', r) & \text{if } \Theta_\Delta(\tilde{\sigma}) = \varepsilon \\ (\tilde{q}', r') & \text{if } \Theta_\Delta(\tilde{\sigma}) \in \Delta \end{cases}$$

where $\tilde{q}' = \tilde{f}(\tilde{q}, \tilde{\sigma})$ and $r' = \mathsf{UR}(\mathsf{Next}(r, \Theta_\Delta(\tilde{\sigma})))$.

Intuitively, each state $(\tilde{q}, r)$ in the verification structure V has two components. The first component $\tilde{q} \in \tilde{Q}$ tracks the current (augmented) state in $\tilde{G}$. Therefore, the transition of the first part is consistent with $\tilde{f}$. As a result, we also have

$$\mathcal{L}(V) = \mathcal{L}(\tilde{G}) = \mathcal{L}_e(G)$$

and the same for the generated infinite language. On the other hand, the second component $r \in 2^{\tilde{Q}}$ captures the current state estimation of system $\tilde{G}$. Therefore, this component is updated only when a new output symbol is generated, i.e., $\Theta_\Delta(\tilde{\sigma}) \in \Delta$. Furthermore, if it is updated, then it first updates the estimate by the observable reach to compute the set of state that can be reached immediately following the output. Also, we need to include all states that can be reached silently by using the unobservable reach. Essentially, we can image V as the synchronization of the augmented system $\tilde{G}$ and its current-state estimate under the non-deterministic observation setting.

For any state $(\tilde{q}, r) \in Q_V$ in V, we say the state is

- *certain* if $\tilde{q} \in \tilde{Q}_F$ and $r \subseteq \tilde{Q}_F$;
- *uncertain* if $\tilde{q} \in \tilde{Q}_F$ and $r \cap \tilde{Q}_N \neq \emptyset$.

We denote by Q_V^{cer} and Q_V^{unc} the set of certain states and uncertain states in V, respectively. Then for any string $s \in \mathcal{L}(V)$, based on the definition of uncertain and certain states, we know that (i) for any faulty extended string $s \in \mathcal{L}(V) = \mathcal{L}_e(G)$, there exists a normal extended string $s' \in \mathcal{L}_e(G)$ such that $\Theta_\Delta(s) = \Theta_\Delta(s')$ if and only if $f_V(q_V^0, s) \in Q_V^{unc}$; and (ii) for any faulty extended string $s \in \mathcal{L}(V) = \mathcal{L}_e(G)$, if $f_V(q_V^0, s) \in Q_V^{cer}$, then for any of its continuation $st \in \mathcal{L}(V)$, we have $f_V(q_V^0, st) \in Q_V^{cer}$.

C. Checking OF-Diagnosability

Now we investigate the verification of OF-diagnosability. According to Definition 2, a system is not OF-diagnosable if and only if there exists an infinite extended faulty string satisfying the Δ'-fairness assumption but all states in V reached along the string are uncertain. Then, since the systems is finite, only loops can generate infinite strings. This leads to the definition of *fairly uncertain loop* (FU-loop).

Formally, given the verification structure V, we define a *run* in V as a finite sequence

$$\pi = q_V^1 \xrightarrow{\sigma_V^1} q_V^2 \xrightarrow{\sigma_V^2} \cdots \xrightarrow{\sigma_V^{n-1}} q_V^n$$

where $q_V^1, \cdots, q_V^n \in Q_V, \sigma_V^1, \cdots, \sigma_V^n \in \Sigma_V$ and $q_V^{i+1} = f_V(q_V^i, \sigma_V^i), i = 1, \ldots, n-1$. A run of the above form is called a *loop* if $q_V^1 = q_V^n$.

Then given a loop $\pi = q_V^1 \xrightarrow{\sigma_V^1} q_V^2 \cdots \xrightarrow{\sigma_V^{n-1}} q_V^n$ in the verification structure, where $\sigma_V^i = (q^i, \sigma^i, \delta^i)$, we say π is

- *fair* if for each transition that occurs in the loop, any of its fair output symbols must occur in the loop associating with the same transition, i.e.,

$$(\forall i \in \{1, \cdots, n\})(\forall \delta \in \mathcal{O}(q^i, \sigma^i) \cap \Delta')$$
$$(\exists j \in \{1, \cdots, n\})[(q^j, \sigma^j) = (q^i, \sigma^i) \wedge \delta^j = \delta]$$

- *uncertainty* if all states in the loop are uncertain, i.e.,

$$\forall i \in \{1, \cdots, n\} : q_V^i \in Q_V^{unc}$$

- *reachable* if there exists a finite string $s \in \mathcal{L}(V)$ such that $f_V(q_V^0, s) = q_V^1$.

Clearly, by properties of the verification structure, we know that an uncertain loop will only be reached by faulty strings. Furthermore, if some state in a loop is uncertain, then all states in it are uncertain.

The following main result shows that, to check OF-diagnosaiblity, it suffices to check the existence of a reachable fair and uncertain loop (FU-loop) in the verification structure V.

Theorem 2: A system G is not OF-diagnosable w.r.t. fault events Σ_F, output function $\mathcal{O}$ and fair outputs $\Delta' \subseteq \Delta$, if and only if, there exists a reachable FU-loop in the verification structure V.

The condition in Theorem 2 can be checked as follows. First, in verification structure, we find all strongly connected components (SCCs) that are reachable by fault events. Clearly, each of above SCCs only consists of either certain states or uncertain states and we only need to consider those uncertain SCCs. Then we need to check if any of the uncertain SCCs contains a fair loop. To this end, for each extended event (q, σ, δ), we check if it is fair in the sense that any $\delta \in \Delta' \cap \mathcal{O}(q, \sigma)$ also appears in the same SCC. If not, we need to remove such an extended event, and then recompute the SCCs and repeat the above removal procedure. When no such "unfair" extended event can be removed, the SCC remained (if exists) will contain a fair-loop; otherwise, no fair-lopp can be found. The above procedure is in polynomial-time in the size of the verification structure since computing all SCCs can be done in linear time and the above procedure repeats at most $|Q|^2|\Sigma||\Delta|$ times. However, the size of the verification structure is exponential in the size of the original system. Therefore, the overall complexity for checking OF-diagnosability is exponential.

Example 5: Still, let us consider system G_1 shown Figure 1 with $\Sigma_F = \{\sigma_f\}$ and suppose that the fair outputs are $\Delta' = \{b, c\}$. As we have discussed in Example 3, this system is OF-diagnosability. Here, we analyze this more formally using Theorem 2. Based on the augmented system $\tilde{G}_1$, we construct the verification structure V_1; part of it is shown in Figure 3, where we focus on the only reachable uncertain loop, as highlighted with red lines in the Figure 3, and omit other parts without loss of generality for the purpose of verification. To form a loop, we have to fire extended event c_4^ϵ infinite number of times at state $4F$. However, the fair output of transition $4 \xrightarrow{c}$ is $c \in \Delta'$, i.e., $\mathcal{O}(4, c) \cap \Delta' = \{c\}$ and extended c_4^c is not included. Therefore, we conclude the V_1 cannot form a FU-loop, which means that system G_1 is OF-diagnosable according to Theorem 2.

In the previous running example, the non-deterministic observation is either to see the occurrence of the internal event or to lose the observation. This actually corresponds to the special case of intermittent loss of observations. As we mentioned, our framework captures the general case of state-dependency and allows the output symbols to be different from the event set.

Example 6: Let us consider system G_2 shown in Figure 4(a) with $\Sigma = \{a, b, c, \sigma_f, u\}$, $\Delta = \{o_1, o_2, o_3\}$ and $\Sigma_F = \{\sigma_f\}$. We assume that the fair outputs are $\Delta' = \{o_2, o_3\}$. For example, for transition $4 \xrightarrow{c}$, we may observe any output $\delta \in \mathcal{O}(4, c) = \{o_2, o_3, \varepsilon\}$ non-deterministically.

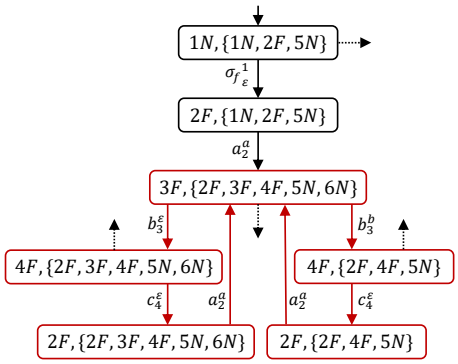
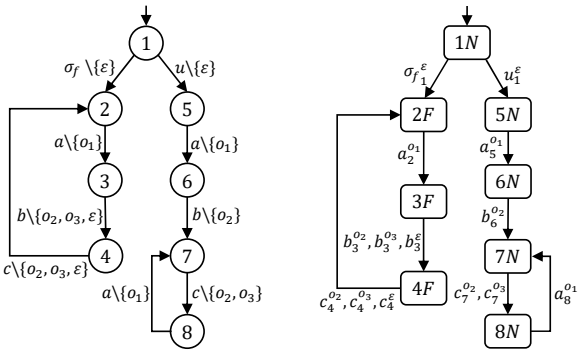


Fig. 3: Part of verification structure V_1 corresponding to system G_1. The uncertain loop is highlighted by red lines.

The augmented system $\tilde{G}_2$ is depicted in Figure 4(b). For this system, the Δ'-fairness assumption is given by

$$\varphi_{fair} = \Big(\bigwedge_{i=4,7} (\Box\Diamond c_i \rightarrow \bigwedge_{\delta\in\{o_2,o_3\}} \Box\Diamond c_i^{\delta})\Big) \wedge (\Box\Diamond b_3 \rightarrow \bigwedge_{\delta\in\{o_2,o_3\}} \Box\Diamond b_3^{\delta}) \wedge (\Box\Diamond b_6 \rightarrow \Box\Diamond b_6^{o_2}).$$

Now, let us verify OF-diagnosability using Theorem 2. To this end, we construct the corresponding verification system V_2, part of which is shown in Figure 5. This loop is fair because for transition $3 \xrightarrow{b}$ all fair outputs in $\mathcal{O}(3,b)\cap\Delta' = \{o_2,o_3\}$ occur in the loop, and for transition $4 \xrightarrow{c}$ all fair outputs in $\mathcal{O}(4,c)\cap\Delta' = \{o_2,o_3\}$ occur in the loop. Furthermore, all states in the loop are uncertain. Thus, system G_2 is not OF-diagnosable.



(a) System G_2. (b) Augmented system $\tilde{G}_2$.

Fig. 4: Example of a non-OF-diagnosable system.

of diagnosability under non-deterministic observations and the physical setting of non-deterministic observations.

V. Conclusion

In this work, we revisited the problem of fault diagnosis of DES under non-deterministic observations. Compared with existing works, we introduced the notion of output fairness that excludes the case where some possible outputs are disabled permanently, which is formalized as Δ'-fairness assumption by LTL. We proposed a new notion called OF-diagnosability to capture the diagnostic requirement under Δ'-fairness assumption. Necessary and sufficient condition for testing OF-diagnosability was also provided based. Our work bridged the discrepancy between the existing definition

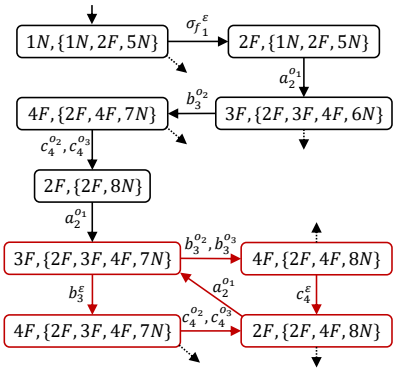


Fig. 5: Part of verification structure V_2 corresponding to system G_2. The uncertain loop is highlighted by red lines.

References

[1] C. Baier and J. Katoen. *Principles of Model Checking*. MIT press, 2008.

[2] P. Biswal and S. Biswas. A polynomial algorithm for diagnosability of fair discrete event systems. *Systems Science & Control Engineering*, 3(1):307–319, 2015.

[3] A. Boussif, M. Ghazel, and J. Basilio. Intermittent fault diagnosability of discrete event systems: an overview of automaton-based approaches. *Discrete Event Dynamic Systems*, 31(1):59–102, 2021.

[4] L. Carvalho, J. Basilio, and M. Moreira. Robust diagnosis of discrete event systems against intermittent loss of observations. *Automatica*, 48(9):2068–2078, 2012.

[5] L. Carvalho, M. Moreira, and J. Basilio. Comparative analysis of related notions of robust diagnosability of discrete-event systems. *Annual Reviews in Control*, 2021.

[6] Y. Hu, Z. Ma, Z. Li, and A. Giua. Diagnosability enforcement in labeled petri nets using supervisory control. *Automatica*, 131:109776, 2021.

[7] S. Lafortune, F. Lin, and C. Hadjicostis. On the history of diagnosability and opacity in discrete event systems. *Annual Reviews in Control*, 45:257–266, 2018.

[8] F. Lin. Control of networked discrete event systems: dealing with communication delays and losses. *SIAM Journal on Control and Optimization*, 52(2):1276–1298, 2014.

[9] Z. Ma, X. Yin, and Z. Li. Marking diagnosability verification in labeled petri nets. *Automatica*, 131:109713, 2021.

[10] V. Oliveira, F. Cabral, and M. Moreira. K-loss robust diagnosability of discrete-event systems. *IFAC-PapersOnLine*, 53(4):250–255, 2020.

[11] N. Ran, H. Su, A. Giua, and C. Seatzu. Codiagnosability analysis of bounded petri nets. *IEEE Trans. Automatic Control*, 63(4):1192–1199, 2018.

[12] S. Takai. A general framework for diagnosis of discrete event systems subject to sensor failures. *Automatica*, 129:109669, 2021.

[13] S. Takai and T. Ushio. Verification of codiagnosability for discrete event systems modeled by Mealy automata with nondeterministic output functions. *IEEE Trans. Aut. Control*, 57(3):798–804, 2012.

[14] T. Ushio and S. Takai. Nonblocking supervisory control of discrete event systems modeled by Mealy automata with nondeterministic output functions. *IEEE Trans. Aut. Control*, 61(3):799–804, 2015.

[15] G. Viana, M. Moreira, and J. Basilio. Codiagnosability analysis of discrete-event systems modeled by weighted automata. *IEEE Trans. Automatic Control*, 64(10):4361–4368, 2019.

[16] X. Yin, J. Chen, Z. Li, and S. Li. Robust fault diagnosis of stochastic discrete event systems. *IEEE Trans. Automatic Control*, 64(10):4237–4244, 2019.

[17] X. Yin and S. Lafortune. Codiagnosability and coobservability under dynamic observations: Transformation and verification. *Automatica*, 61:241–252, 2015.

[18] X. Yin and S. Lafortune. On the decidability and complexity of diagnosability for labeled Petri nets. *IEEE Trans. Automatic Control*, 62(11):5931–5938, 2017.

[19] J. Zaytoon and S. Lafortune. Overview of fault diagnosis methods for discrete event systems. *Annual Rev. Control*, 37(2):308–320, 2013.