Diagnosability of Discrete-Event Systems

Meera Sampath, Raja Sengupta, Stéphane Lafortune, *Member, IEEE*,
Kasim Sinnamohideen, *Member, IEEE*, and Demosthenis Teneketzis, *Member, IEEE*

Abstract—**Fault detection and isolation is a crucial and challenging task in the automatic control of large complex systems. We propose a discrete-event system (DES) approach to the problem of failure diagnosis. We introduce two related notions of diagnosability of DES's in the framework of formal languages and compare diagnosability with the related notions of observability and invertibility. We present a systematic procedure for detection and isolation of failure events using diagnosers and provide necessary and sufficient conditions for a language to be diagnosable. The diagnoser performs diagnostics using on-line observations of the system behavior; it is also used to state and verify off-line the necessary and sufficient conditions for diagnosability. These conditions are stated on the diagnoser or variations thereof. The approach to failure diagnosis presented in this paper is applicable to systems that fall naturally in the class of DES's; moreover, for the purpose of diagnosis, most continuous variable dynamic systems can be viewed as DES's at a higher level of abstraction. In a companion paper [20], we provide a methodology for building DES models for the purpose of failure diagnosis and present applications of the theory developed in this paper.**

I. Introduction

In this paper, we study the diagnosability of discrete-event systems. The property of diagnosability is introduced in the context of the failure diagnosis problem. Failure detection and isolation is an important task in the automatic control of large complex systems. The increasingly stringent requirements on performance and reliability of complex man-made systems have necessitated the development of sophisticated and systematic methods for the timely and accurate diagnosis of system failures. The problem of failure diagnosis has received considerable attention in the literature, and a wide variety of schemes have been proposed. These include: i) quantitative methods based on mathematical models (see [1], [5], [7], [23], and references therein), ii) expert systems and other AI-based methods (see [6], [18], and references therein), and iii) discrete-event systems (DES's) methods (see [2], [8], [10], [11], [21], and [22]). The quantitative methods employ analytical models of the physical process, which allow for comparison of sensor measurements with their predicted values. The AI-based methods incorporate the knowledge of human experts and reasoning mechanisms into the diagnostic system. Methods combining these two approaches have also been proposed [5], [7].

We propose in this paper and in the companion paper [20] a DES approach to the problem of failure diagnosis that expands on the work in [21]. This approach is applicable to systems that fall naturally in the class of DES's; moreover, for the purpose of diagnosis, continuous variable dynamic systems can often be viewed as DES's at a higher level of abstraction. The states of the discrete-event model reflect the normal and the failed status of the system components while the failure events form part of the event set. The problem is to detect the occurrence of these events. The major advantage of this approach is that it does not require detailed in-depth modeling of the system to be diagnosed and hence is ideally suited for diagnosis of systems which are difficult to model. Typical examples include large and/or complex systems like heating, ventilation, and air conditioning (HVAC) units, power plants, and semiconductor manufacturing equipment. In [20], we discuss in detail discrete-event modeling of systems for failure diagnosis and illustrate our approach with several examples. Comparisons are also made between our approach and alternative approaches to failure diagnosis.

The focus of this paper is to develop the underlying theory for our approach. The system behavior is modeled as a regular language and is represented by a finite state machine (FSM). We propose two related notions of diagnosability in the framework of formal languages. Roughly speaking, a language is said to be diagnosable if it is possible to detect (with finite delay) occurrences of certain distinguished unobservable events, namely the failure events. We present a systematic procedure for detection and isolation of failure events using diagnosers. The diagnoser is an FSM built from the FSM model of the system. This machine performs diagnostics when it observes on-line the behavior of the system; states of the diagnoser carry failure information and occurrences of failures can be detected (with a finite delay) by inspecting these states. We provide necessary and sufficient conditions for a language to be diagnosable. These conditions are stated on the diagnoser or variations thereof. Thus, the diagnoser serves two purposes: i) on-line detection and isolation of failures and ii) off-line verification of the diagnosability properties of the system.

In Section II, we introduce the notion of diagnosability of DES's. We first present the system model and introduce the necessary notation. Next, we formally define the notions of diagnosability and I-diagnosability and illustrate these definitions by means of simple examples. This is followed by a comparison with related work in the DES literature, namely, other approaches to diagnosability, and the problems

Manuscript received April 15, 1994; revised December 15, 1994. Recommended by Associate Editor, R. Nikoukhah. This research was supported in part by NSF Grants ECS-9057967, ECS-9312134, and NCR-9204419 with additional support from DEC and GE.

M. Sampath, R. Sengupta, S. Lafortune, and D. Teneketzis are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA.

K. Sinnamohideen is with Johnson Controls, Inc., Milwaukee, WI 53201 USA.

IEEE Log Number 9413531.

of observability and invertibility. In Section III, we present the construction procedure of the diagnoser and illustrate this procedure with an example. Necessary and sufficient conditions for diagnosability and I-diagnosability are presented in Section IV. In Section V we discuss on-line diagnosis of failures in diagnosable systems using the diagnoser introduced in Section III. Finally, in Section VI we provide a brief summary of the main results of this paper. A summary of some of the results in this paper can be found in [19].

II. The Notion of Diagnosability

A. The System Model

The system to be diagnosed is modeled as an FSM or generator

$$G = (X, \Sigma, \delta, x_0) \tag{1}$$

where X is the state space, Σ is the set of events, δ is the partial transition function, and x_0 is the initial state of the system. The model G accounts for the normal and failed behavior of the system. The behavior of the system is described by the prefix-closed language [17] $L(G)$ generated by G. Henceforth, we shall denote $L(G)$ by L. L is a subset of Σ^*, where Σ^* denotes the Kleene closure of the set Σ [9].

Some of the events in Σ are observable, i.e., their occurrence can be observed, while the rest are unobservable. Thus the event set Σ is partitioned as $\Sigma = \Sigma_o \dot{\cup} \Sigma_{uo}$ where Σ_o represents the set of observable events and Σ_{uo} represents the set of unobservable events. The observable events in the system may be one of the following: commands issued by the controller, sensor readings immediately after the execution of the above commands, and changes of sensor readings. The unobservable events may be failure events or other events that cause changes in the system state not recorded by sensors (see [20]).

Let $\Sigma_f \subseteq \Sigma$ denote the set of failure events which are to be diagnosed. We assume, without loss of generality, that $\Sigma_f \subseteq \Sigma_{uo}$, since an observable failure event can be trivially diagnosed. Our objective is to identify the occurrence, if any, of the failure events, given that in the traces generated by the system, only the events in Σ_o are observed. In this regard, we partition the set of failure events into disjoint sets corresponding to different failure types

$$\Sigma_f = \Sigma_{f1} \cup \cdots \cup \Sigma_{fm}. \tag{2}$$

Let Π_f denote this partition. The partition Π_f is motivated by the following considerations:

1) Inadequate instrumentation may render it impossible to diagnose uniquely every possible fault.
2) We may not be required to identify uniquely the occurrence of every failure event. We may simply be interested in knowing if one of a set of failure events has happened as, for example, when the effect of the set of failures on the system is the same.

Hereafter, when we write that "a failure of type F_i has occurred," we will mean that some event from the set Σ_{fi} has occurred.

We make the following assumptions on the system under investigation:

A1) The language L generated by G is live. This means that there is a transition defined at each state x in X, i.e., the system cannot reach a point at which no event is possible.

A2) There does not exist in G any cycle of unobservable events, i.e.,

$$\exists\, n_o \in \mathbb{N} \quad \text{such that} \quad \forall\, ust \in L,\ s \in \Sigma_{uo}^* \Rightarrow ||s|| \leq n_o$$

where $||s||$ is the length of trace s.

The liveness assumption on L is made for the sake of simplicity. With slight modifications, all of the main results of this paper hold true when the liveness assumption is relaxed. Assumption A2) ensures that observations occur with some regularity. Since detection of failures is based on observable transitions of the system, we require that G does not generate arbitrarily long sequences of unobservable events.

In [20], we discuss in detail discrete-event modeling of systems for failure diagnosis. Suppose the system to be diagnosed consists of several distinct physical components and a set of sensors. We first build FSM models of the individual components. These models account for both the normal and the faulty behavior of the components. Consider, for example, a simple HVAC system consisting of a pump, valve, and controller. Fig. 1 depicts the component models for this system. Starting from the component models and sensor maps, we then generate a composite model which captures the interactions between the components and also incorporates in it the sensor maps. This composite model is the system G on which we perform diagnostics.

We conclude this section on the system model with some notation and the construction of the generator G' that will be used later.

1) Notation: The empty trace is denoted by ϵ. Let $\overline{s}$ denote the prefix-closure of any trace $s \in \Sigma^*$. We denote by L/s the postlanguage of L after s, i.e.,

$$L/s = \{t \in \Sigma^* | \, st \in L\}. \tag{3}$$

We define the projection $P: \Sigma^* \to \Sigma_o^*$ in the usual manner [17]

$$\begin{aligned} P(\epsilon) &= \epsilon \\ P(\sigma) &= \sigma \quad \text{if} \quad \sigma \in \Sigma_o \\ P(\sigma) &= \epsilon \quad \text{if} \quad \sigma \in \Sigma_{uo} \\ P(s\sigma) &= P(s)P(\sigma)\ s \in \Sigma^*, \quad \sigma \in \Sigma. \end{aligned} \tag{4}$$

Thus, P simply "erases" the unobservable events in a trace. The inverse projection operator P_L^{-1} is defined as

$$P_L^{-1}(y) = \{s \in L: P(s) = y\}. \tag{5}$$

Let s_f denote the final event of trace s. We define

$$\Psi(\Sigma_{fi}) = \{s\sigma_f \in L: \sigma_f \in \Sigma_{fi}\} \tag{6}$$

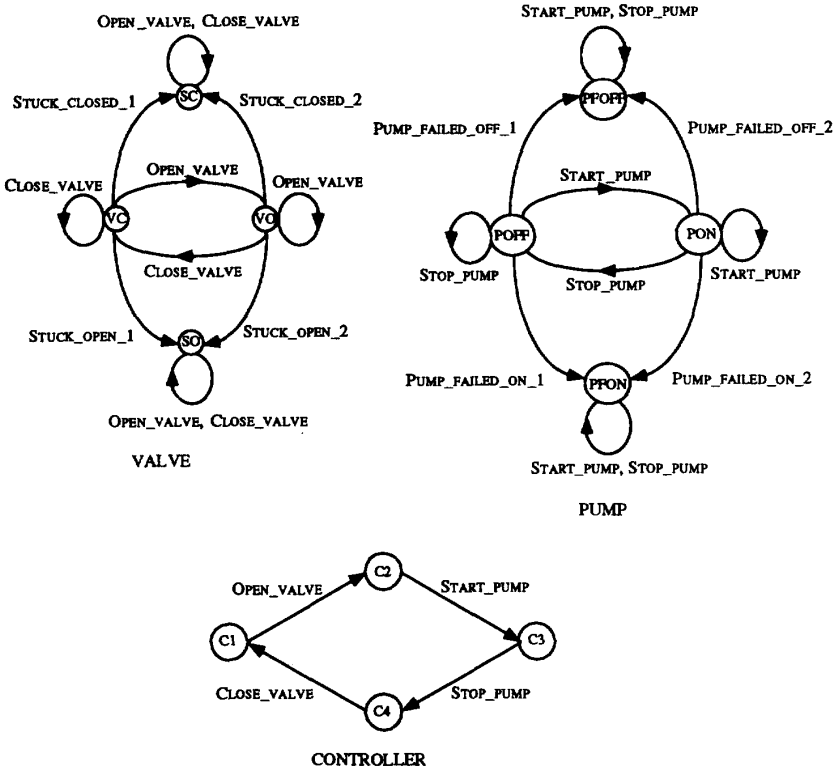


Fig. 1. Component models for a simple HVAC system.

i.e., $\Psi(\Sigma_{fi})$ denotes the set of all traces of L that end in a failure event belonging to the class Σ_{fi}. Consider $\sigma \in \Sigma$ and $s \in \Sigma^*$. We use the notation $\sigma \in s$ to denote the fact that σ is an event in the trace s. With slight abuse of notation, we write $\Sigma_{fi} \in s$ to denote the fact that $\sigma_f \in s$ for some $\sigma_f \in \Sigma_{fi}$, or, formally, $\overline{s} \cap \Psi(\Sigma_{fi}) \neq \emptyset$.

We define

$$X_o = \{x_0\} \cup \{x \in X : x \text{ has an observable event into it}\}. \tag{7}$$

Let $L(G, x)$ denote the set of all traces that originate from state x of G. We define

$$L_o(G, x) = \{s \in L(G, x) : s = u\sigma,\ u \in \Sigma_{uo}^*,\ \sigma \in \Sigma_o\} \tag{8}$$

and

$$L_\sigma(G, x) = \{s \in L_o(G, x) : s_f = \sigma\}. \tag{9}$$

$L_o(G, x)$ denotes the set of all traces that originate from state x and end at the first observable event. $L_\sigma(G, x)$ denotes those traces in $L_o(G, x)$ that end with the particular observable event σ.

2) The Generator G': In the following sections, we will need to use a specially constructed generator G' of the language

$$P(L) = \{t : t = P(s) \quad \text{for some} \quad s \in L\}. \tag{10}$$

G' will in general be nondeterministic, and it is constructed as follows

$$G' = (X_o, \Sigma_o, \delta_{G'}, x_0) \tag{11}$$

where X_o, Σ_o, and x_0 are as defined previously. The transition relation of G' is given by $\delta_{G'} \subseteq (X_o \times \Sigma \times X_o)$ and is defined as follows

$$(x, \sigma, x') \in \delta_{G'} \quad \text{if} \quad \delta(x, s) = x' \quad \text{for some} \quad s \in L_\sigma(G, x). \tag{12}$$

It is straightforward to verify that $L(G') = P(L)$. Figs. 4–6 illustrate the construction of G' from G for three different systems.

B. Approaches to Defining Diagnosability

We are now ready to define the notion of diagnosability. Roughly speaking, a language L is diagnosable if it is possible to detect with a finite delay occurrences of failures of any type using the record of observed events. We now present two definitions of diagnosability, with the first definition more stringent than the second. We shall henceforth refer to the first notion as diagnosability and to the second one as I-diagnosability.

1) Diagnosability: Formally, we define diagnosability as follows.

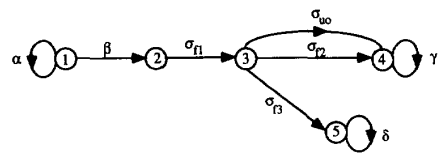


Fig. 2. Example of a system with multiple failures.

Definition 1: A prefix-closed and live language L is said to be diagnosable with respect to the projection P and with respect to the partition Π_f on Σ_f if the following holds

$$(\forall i \in \Pi_f)(\exists n_i \in \mathbb{N})[\forall s \in \Psi(\Sigma_{fi})](\forall t \in L/s) \\ [||t|| \geq n_i \Rightarrow D]$$

where the diagnosability condition D is

$$\omega \in P_L^{-1}[P(st)] \Rightarrow \Sigma_{fi} \in \omega.$$

The above definition of diagnosability means the following. Let s be any trace generated by the system that ends in a failure event from the set Σ_{fi}, and let t be any sufficiently long continuation of s. Condition D then requires that every trace belonging to the language that produces the same record of observable events as the trace st should contain in it a failure event from the set Σ_{fi}. This implies that along every continuation t of s one can detect the occurrence of a failure of the type F_i with a finite delay, specifically in at most n_i transitions of the system after s. Alternately speaking, diagnosability requires that every failure event leads to observations distinct enough to enable unique identification of the failure type with a finite delay.

The case of multiple failures from the same set of the partition deserves special attention. When more than one failure of the same type, say, F_i, occurs along a trace s of L, the above definition of diagnosability does not require that each of these occurrences be detected. It suffices to be able to conclude, within finitely many events after the occurrence of the first failure, that along s, a failure from the set Σ_{fi} happened. In later sections we shall see how this feature distinguishes the case of possible multiple failures from the case of no multiple failures from any set of the partition.

We illustrate by a simple example the above notion of diagnosability. Consider the system represented in Fig. 2. Here, α, β, γ, and δ are observable events, σ_{uo} is an unobservable event while σ_{f1}, σ_{f2}, and σ_{f3} represent failure events. Let the initial state x_0 of the system be state 1. If one chooses the partition $\Sigma_{f1} = \{\sigma_{f1}, \sigma_{f2}\}$ and $\Sigma_{f2} = \{\sigma_{f3}\}$, i.e., it is not required to distinguish between failures σ_{f1} and σ_{f2}, then the above system is diagnosable with $n_1 = 2$ and $n_2 = 1$. On the other hand, if the partition is $\Sigma_{f1} = \{\sigma_{f1}\}$, $\Sigma_{f2} = \{\sigma_{f2}\}$, and $\Sigma_{f3} = \{\sigma_{f3}\}$, then the system is not diagnosable since it is not possible to deduce the occurrence of failure σ_{f2}.

2) I-Diagnosability: The preceding definition of diagnosability requires condition D to hold for all traces of L containing a failure event. We now propose a relaxed definition of diagnosability (termed I-diagnosability) that requires the diagnosability condition D to hold not for all traces containing a failure event, but only for those in which the failure event is followed by certain indicator observable events associated with every failure type. This modification is motivated by the following physical consideration. Consider, for example, an HVAC system with a controller unit. In normal mode of operation, the controller responds by issuing the command "open valve" whenever it senses a heating load on the system. Likewise, it issues the command "close valve" when the load is removed. Assume that when the controller fails, it does not sense the presence of any load on the system and hence does not issue any commands to the valve. Suppose that during operation, the controller does fail, and suppose further that it is possible for the system to execute an arbitrarily long sequence of events, which does not involve any of the valve commands. Under such conditions, it is obvious that one cannot diagnose any failure of the valve. Such a system is considered not diagnosable according to the previous definition. In the case of the modified definition, we associate as indicator events, "open valve" and "close valve," respectively, with the valve failure events, "stuck-closed" and "stuck-open," and require the system to execute the "open valve" event or the "close valve" event before deciding on its diagnosability. The system is considered diagnosable if after the execution of the corresponding indicator events it is possible to detect valve failures, while it is termed not diagnosable if even after the indicator event is executed the corresponding valve failure remains undetectable. To summarize, I-diagnosability requires detection of failures only after the occurrence of an indicator event corresponding to the failure.

We first associate to every failure event in Σ_f one or more observable indicator events. Let $\Sigma_I \subseteq \Sigma_o$ denote the set of indicator events, and let $I_f: \Sigma_f \rightarrow 2^{\Sigma_I}$ denote the indicator map. Next we choose a partition Π_f on Σ_f such that

$$\bigcup_{i \in \Pi_f} \Sigma_{fi} = \Sigma_f$$

as before, with the additional constraint that for each $i = 1, \cdots, m$

$$\sigma_{f1}, \sigma_{f2} \in \Sigma_{fi} \Rightarrow I_f(\sigma_{f1}) = I_f(\sigma_{f2})$$

and define

$$I(\Sigma_{fi}) = I_f(\sigma_f) \quad \text{for any} \quad \sigma_f \in \Sigma_{fi}. \tag{13}$$

We now have a set of observable indicator events $I(\Sigma_{fi})$ associated with each failure type F_i. (See [20] for more details on the choice of indicator events for physical systems.)

We now propose the following definition of I-diagnosability.

Definition 2: A prefix-closed and live language L is said to be I-diagnosable with respect to the projection P, the partition Π_f on Σ_f, and the indicator map I if the following holds

$$(\forall i \in \Pi_f)(\exists n_i \in \mathbb{N})(\forall s \in \Psi(\Sigma_{fi}) \\ (\forall t_1 t_2 \in L/s: st_1 \in \Psi[I(\Sigma_{fi})]) \quad [||t_2|| \geq n_i \Rightarrow D]$$

where the diagnosability condition D is

$$\omega \in P_L^{-1}[P(st_1 t_2)] \Rightarrow \Sigma_{fi} \in \omega.$$

Note that $\Psi[I(\Sigma_{fi})]$ denotes the set of all traces of L that end in an observable event from the set $I(\Sigma_{fi})$. Therefore, in

the case of the I-diagnosability, we require that occurrences of failure events of the type F_i should be detected in at most n_i transitions of the system after the occurrence of an indicator event from the set $I(\Sigma_{fi})$.

Consider the system represented in Fig. 2. Suppose that the indicator events are chosen as follows: $I(\Sigma_{f1}) = \{\gamma\}$, $I(\Sigma_{f2}) = \{\delta\}$, and $I(\Sigma_{f3}) = \{\delta\}$. Let the desired partition be $\Sigma_{f1} = \{\sigma_{f1}\}$, $\Sigma_{f2} = \{\sigma_{f2}\}$, and $\Sigma_{f3} = \{\sigma_{f3}\}$. This system is I-diagnosable with $n_1 = 0$ and $n_3 = 0$. It is to be noted that although it is not possible to deduce the occurrence of failure σ_{f2}, the indicator event corresponding to σ_{f2}, i.e., δ does not follow this failure event and hence the diagnosability condition is not violated.

C. Comparison with Related Work

Partial observation problems in DES's have been investigated by several researchers. While the problem of diagnosability itself has not been studied in detail, the related notions of observability, observability with delay, and invertibility have been the subject of several papers, among them [3], [4], and [12]–[16]. Though closely related to these other problems, diagnosability is a distinctly different notion for the following reasons: partitioning of the failure events, need to identify every failure type with a finite delay, possibility of multiple failures, possible presence of unobservable events other than the failure events, no requirement of diagnosis or detection during normal system operation, and absence of "locking-on" phenomenon (explained below). In this section, we first discuss other approaches to diagnosability that have been proposed in recent DES literature [2], [10]. Afterward, we discuss briefly the differences between diagnosability and the other notions mentioned above.

1) Other Approaches to Diagnosability: Lin, in [10] (also see [11]), proposes a state-based approach to diagnosability. He assumes partial state information available via an output function. He addresses separately the problems of off-line and on-line diagnosis. In off-line diagnosis, the system to be diagnosed is not in normal operation and can be thought of as being in a "test-bed." The diagnostic procedure involves issuing a sequence of test commands, observing the resulting outputs, and drawing inferences on the set of possible states the system could be in. The off-line diagnosis problem can be considered equivalent to the problem of "verification." In on-line diagnosis, the system is assumed to be in normal operation. The goal of diagnostics, as before, is to issue a sequence of commands and identify uniquely, up to a partition, the state of the system. Unlike the case of off-line diagnosis, however, one now has to account for the possible occurrences of other uncontrollable events during the diagnostic process. The author gives an algorithm for computing a diagnostic control or a sequence of test commands for diagnosing system failures. This algorithm is guaranteed to converge if the system is indeed on-line diagnosable.

In [2], Bavishi and Chong study extensions of the above work. In particular, they consider testability of DES's (which is equivalent to the off-line diagnosability problem studied in [10]) and present algorithms i) for determining the optimal set of sensors which would ensure testability of a given system and ii) given a fixed set of sensors, for determining the infimal partition of the state space, with respect to which the system is testable.

2) Related Notions in DES's:

Language Observability: Lin and Wonham study in [12] the supervisory control problem with partial event observations. They introduce a language-based definition of observability and state conditions for the existence of a solution to the supervisory control problem in terms of observability and controllability of languages. The control problem addressed there does not require explicit determination of the occurrences of unobservable events or identification of the system state. Thus, the notion of observability introduced there is different from the problem of diagnosability.

Observability of State Machines: In his paper on observability of DES's [16], Ramadge explicitly addresses the problem of state identification for discrete-event systems. In his framework, the system is modeled by a nondeterministic automaton with full event observability and partial state observability via an output map defined on the states (as in a Moore automaton). The problem is to reconstruct exactly the state of the system after the occurrence of every event. The motivation for the observability problem addressed there is an observer-state feedback approach to controller synthesis. The work in [16] is set in a different framework and is incomparable with the diagnosability problem studied here.

Özveren and Willsky adopt in [13] a slightly different notion of observability from that of Ramadge. They assume a partial event observation model with no direct state observations. A system is termed observable if, using a record of observable events, it is possible to determine the current state exactly at intermittent (but not necessarily fixed) points in time, separated by a bounded number of events. An observer is a DES which produces estimates of the state of the system after the occurrence of every observable event. In [13], the authors also address the problem of observability with delay. A system is said to be observable with delay, if, at intermittent points in time, it is possible to have perfect knowledge not of the current state of the system but of the state some finite number of transitions into the past. In our framework, diagnosability is posed as an event detection problem. When viewed as a problem of state identification, diagnosability is a stronger notion than observability with delay since the former requires that every failure state should be identifiable uniquely (up to a partition). In contrast, in [13], there is no notion of a particular state or set of states being observable. A system is observable (or observable with delay) as long as there exists at least one state which is uniquely identifiable at intermittent points in time. On the other hand, diagnosability only requires that the failure states be identifiable with finite delay; there are no similar requirements on the normal states. Thus, a system could execute arbitrarily long sequences of events, while in normal (failure-less) operation, with no single state being uniquely determinable even with delay. Further, a system could fail to be observable (with or without delay) if in the post-failure operation, there exists no state that is uniquely identifiable. This system could still be diagnosable,

however, since we require unique identification not of every failure state but only of every set of the partition. See Appendix A of this paper for examples illustrating differences between diagnosability and observability with delay.

In [3], Caines *et al.* study the state estimation problem for partially observed automata. The system is modeled as input-state-output automaton with partial state information available via an output function. A state output automaton is taken to be a special case of the above automaton where the input set is a singleton. They address the problems of initial state observability and current state observability using two different kinds of observers: classical dynamical observers and logic-based dynamical observers. The classical dynamic observer is a finite state automaton which takes for its input the observed system behavior, namely, the sequence of input–output pairs, and generates a sequence of state estimates (either of the initial state or of the current state). The logic-based observer, on the other hand, is a logic-based dynamical system built in the framework of predicate calculus. This observer generates a sequence of logic propositions which describe the properties of the system. An interesting feature of these logic-based observers is their adaptability to changes in the system model. Observability as studied in [3] and observability as discussed in [16] and [13] differ in the following important aspect. In [3], the authors assume that once the current state of the system is determined, then it is known for all future time, i.e., once the observer estimate converges to the true state of the system, it will thenceforth stay locked on and will always provide the correct system state as its output, for all observed input–output behavior.

Invertibility: In [14], Özveren and Willsky introduce the concept of invertibility which is closely related to the problem of diagnosability. A language is said to be invertible if, at any time, using knowledge of the observed event sequence up to that time, we can reconstruct the full event sequence (corresponding to this observed sequence) up to a finite, bounded number of events in the past. Invertibility is a stronger notion than diagnosability. For a system to be diagnosable, we do not require reconstruction of entire event sequences; we are interested in identifying the occurrence of specific failure events only. Further, when the failure events are partitioned into sets, one is interested only in identifying if one of a set of events has happened. Also, as mentioned before, in the case of multiple failures from the same set of the partition, diagnosability does not require detection of every single occurrence of these failures; it is enough to be able to conclude that a failure event from that set has occurred at least once. Hence, a system that is diagnosable could be noninvertible. We present in Appendix A an example of a noninvertible system which is diagnosable.

The problem of eventual invertibility of timed DES's modeled by generalized semi-Markov schemes is addressed by Park and Chong in [15]. In this modeling framework, the timed behavior of a system is described by an automaton in conjunction with a set of event lifetimes. Partial state as well as partial event information is assumed available. In addition, all transition firing times are assumed to be observable. The problem of eventual invertibility is to determine from observations of events, states, and transition epochs, the corresponding event lifetimes up to a finite time in the past. The authors establish in [15] the equivalence between the problem of extracting event lifetimes and that of constructing the event trajectory from observations of the system behavior.

This concludes the comparison of our notion of diagnosability with other related notions that have appeared in the literature.

III. The Diagnoser

We now introduce the diagnoser which is an FSM built from the system model G. This machine is used to perform diagnostics when it observes on-line the behavior of G. The diagnoser is also used to state necessary and sufficient conditions for diagnosability. While the "basic" diagnoser presented in this section is adequate for the purpose of diagnosis, additional modifications as presented in Section IV are necessary to test for diagnosability. In this section we present the construction procedure of the diagnoser. On-line diagnosis of failures in diagnosable systems using this diagnoser is discussed in Section V.

Construction: We define the set of failure labels $\Delta_f = \{F_1, F_2, \cdots F_m\}$ where $|\Pi_f| = m$ and the complete set of possible labels

$$\Delta = \{N\} \cup 2^{\{\Delta_f \cup \{A\}\}}. \tag{14}$$

Here N is to be interpreted as meaning "normal," A as meaning "ambiguous" (to be explained shortly), and F_i, $i \in \{1, \cdots, m\}$ as meaning that a failure of the type F_i has occurred. Recall from Section II-A-1) the definition of X_o and define

$$Q_o = 2^{X_o \times \Delta}. \tag{15}$$

The diagnoser for G is the FSM

$$G_d = (Q_d, \Sigma_o, \delta_d, q_0) \tag{16}$$

where Q_d, Σ_o, δ_d, and q_0 have the usual interpretation. The initial state of the diagnoser q_0 is defined to be $\{(x_0, \{N\})\}$. The transition function δ_d of the diagnoser is constructed as explained below. The state space Q_d is the resulting subset of Q_o composed of the states of the diagnoser that are reachable from q_0 under δ_d. Since the state space Q_d of the diagnoser is a subset of Q_o, a state q_d of G_d is of the form

$$q_d = \{(x_1, \ell_1), \cdots, (x_n, \ell_n)\}$$

where $x_i \in X_o$ and $\ell_i \in \Delta$, i.e., ℓ_i is of the form $\ell_i = \{N\}$, $\ell_i = \{A\}$, $\ell_i = \{F_{i_1} F_{i_2}, \cdots, F_{i_k}\}$, or $\ell_i = \{A, F_{i_1} F_{i_2}, \cdots, F_{i_k}\}$ where in the last two cases $\{i_1, i_2, \cdots, i_k\} \subseteq \{1, 2, \cdots, m\}$.

An observer for G (see [13]) gives estimates of the current state of the system after the occurrence of every observable event. The diagnoser G_d can be thought of as an extended observer where we append to every state estimate a label of the form mentioned above. The labels attached to the state estimates carry failure information and failures are diagnosed by checking these labels. We assume the system G is normal to start with, hence we define $q_0 = \{(x_o, \{N\})\}$.

Before defining the transition function δ_d of the diagnoser, we define the following three functions: the label propagation function LP, the range function R, and the label correction function LC.

Definition 3: The label propagation function $LP: X_o \times \Delta \times \Sigma^* \rightarrow \Delta$.

Given $x \in X_o$, $\ell \in \Delta$, and $s \in L_o(G, x)$, LP propagates the label ℓ over s, starting from x and following the dynamics of G, i.e., according to $L(G, x)$. It is defined as follows

$$LP(x, \ell, s) = \begin{cases} \{N\} & \text{if } \ell = \{N\} \wedge \forall i[\Sigma_{fi} \notin s] \\ \{A\} & \text{if } \ell = \{A\} \wedge \forall i[\Sigma_{fi} \notin s] \\ \{F_i : F_i \in \ell \vee \Sigma_{fi} \in s\} & \text{otherwise.} \end{cases}$$

Definition 4: The range function $R: Q_o \times \Sigma_o \rightarrow Q_o$ is defined as follows

$$R(q, \sigma) = \bigcup_{(x,\ell)\in q} \bigcup_{s \in L_\sigma(G,x)} \{(\delta(x, s), LP(x, \ell, s))\}.$$

Definition 5: The label correction function $LC: Q_o \rightarrow Q_o$ is defined as follows

$$LC(q) = \{(x, \ell) \in q : x \text{ appears only once in all the pairs in } q\} \cup \{(x, \{A\} \cup \ell_{i1} \cap \cdots \cap \ell_{ik}) \text{ whenever } \exists \text{ two or more pairs } (x, \ell_{i1}), \cdots, (x, \ell_{ik}) \text{ in } q\}.$$

The use of the label correction function LC and the label A is explained as follows. The label acquired by any state x along a trace s indicates the occurrence or otherwise of a failure when the system moves along trace s and transitions into state x. Suppose that there exist two pairs (x, ℓ), (x, ℓ') in $R(q, \sigma)$ for some state q of the diagnoser. Then this implies that the state x could have resulted from a failure event of a particular type, say F_i, or not. Under this condition, we attach the label A to x to denote the fact that there is an ambiguity. In other words, the A label is to be interpreted as meaning "either F_i or not F_i" for $i \in \{1, \cdots, m\}$. It is to be noted here that we do not distinguish between cases "F_i or F_j," "F_j or F_k," "N or F_i," and so on. In all of these situations, we simply use the label A. While this may lead to loss of information necessary for determining diagnosability of a language, it is adequate for the purpose of diagnosis to treat alike all cases mentioned above. We will explain this in more detail in Sections IV and V.

The transition function of the diagnoser $\delta_d: Q_o \times \Sigma_o \rightarrow Q_o$ is now defined as

$$q_2 = \delta_d(q_1, \sigma) \Leftrightarrow q_2 = LC[R(q_1, \sigma)] \tag{17}$$

with $\sigma \in e_d(q_1)$ where

$$e_d(q_1) = \bigcup_{(x,\ell)\in q_1} \{P(s) : s \in L_o(G, x)\}. \tag{18}$$

In words, $e_d(q_1)$ is the active event set of G_d at q_1, i.e., the set of all possible transitions of the diagnoser at the state q_1.

To summarize, the diagnoser G_d is constructed as follows. Let the current state of the diagnoser (i.e., the set of estimates of the current state of G with their corresponding labels) be q_1, and let the next observed event be σ. The new state of the diagnoser q_2 is computed following a three-step process:

1) For every state estimate x in q_1, compute the reach due to σ, given by $S(x, \sigma) = \{\delta(x, s\sigma) \text{ where } s \in \Sigma_{uo}^*\}$.
2) Let $x' \in S(x, \sigma)$ with $\delta(x, s\sigma) = x'$. Propagate the label ℓ associated with x to the label ℓ' associated with x' according to the following rules:
   a) If $\ell = \{N\}$ and s contains no failure events, then the label ℓ' is also $\{N\}$.
   b) If $\ell = \{A\}$ and s contains no failure events, then the label ℓ' is also $\{A\}$.
   c) If $\ell = \{A, F_i\}$ and s contains no failure events, then the label ℓ' is $\{F_i\}$.
   d) If $\ell = \{N\}$ or $\{A\}$ and s contains failure events from Σ_{fi}, Σ_{fj}, then $\ell' = \{F_i, F_j\}$.
   e) If $\ell = \{F_i, F_j\}$ or $\{A, F_i, F_j\}$ and s contains failure events from Σ_{fk}, then $\ell' = \{F_i, F_j, F_k\}$.
3) Let q_2 be the set of all (x', ℓ') pairs computed following Steps 1) and 2) above, for each (x, ℓ) in q_1. Replace by (x', A, F_i, F_j) all (x', ℓ'), $(x', \ell'') \in q_2$ such that F_i and F_j are components of both ℓ' and ℓ''. That is, if the same state estimate x' appears more than once in q_2 with different labels, we associate with x' all common components of these labels, and in addition, we attach to x the ambiguous label A.

Note that in cases c), d), and e) above, we do not propagate the A label from one state to the next. While this leads to a reduction in the state space of the diagnoser, it leads to no loss of information necessary for determining the diagnosability properties of a language or for implementing diagnostics. The reasons for this will become evident in the subsequent sections.

We now give a simple example illustrating the construction of the diagnoser. Fig. 3 illustrates a system G and its diagnoser G_d. Here α, β, γ, δ, and σ are observable events while σ_{uo}, σ_{f1}, σ_{f2}, and $\sigma_{f2'}$ are unobservable. $\Sigma_{f1} = \{\sigma_{f1}\}$ and $\Sigma_{f2} = \{\sigma_{f2}, \sigma_{f2'}\}$. In all illustrations that follow, we represent (x, ℓ) pairs simply as $x\ell$ for clarity. Also, the initial state x_0 of G is chosen to be state 1.

Remark: In the above construction procedure we have assumed knowledge of the initial state of the system, since the diagnoser is assumed to run in parallel with the system from the start of operation. It is to be noted that the above procedure remains valid, however, even in the case of unknown initial state.

IV. Necessary and Sufficient Conditions for Diagnosability

In this section, we present necessary and sufficient conditions for a language L to be diagnosable, followed by similar conditions for L to be I-diagnosable. These conditions are stated on the diagnoser G_d or variations thereof. To test for these conditions, we use, in addition to the diagnoser, the machine G' introduced in Section II-A-2).

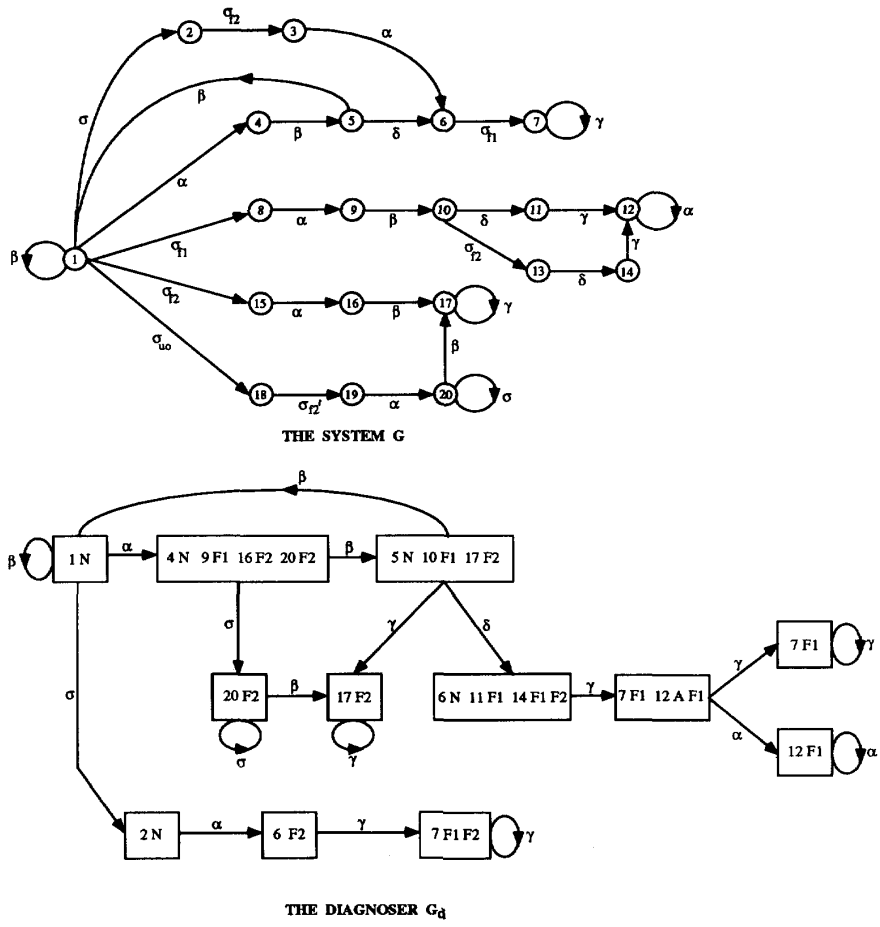


Fig. 3. Example illustrating construction of the diagnoser G_d.

A. Conditions for Diagnosability

We investigate separately the case of no multiple failures of the same type and that of possible multiple failures of the same type. The former corresponds to the situation where along every trace s of L, no more than one failure from the same set of the partition can occur; the latter corresponds to the situation where it is possible to have multiple failures from the same set of the partition occurring along any trace s. The reason for the separate investigation of these two cases will become apparent as we proceed.

1) The Case of No Multiple Failures:

Properties and Definitions of the Diagnoser: We now state a few properties of the diagnoser that follow from its construction. These properties and the definitions that follow will be used subsequently to state and prove the conditions for diagnosability.

P1) By construction, any $x_i \in X_o$ appears in at most one pair (x_i, ℓ_i) in any state of Q_d.

P2) Let $q \in Q_d$. Then

$$(x_1, \ell_1), (x_2, \ell_2) \in q \Leftrightarrow \quad \exists\, s_1, s_2 \in L$$

such that

$$s_{1f}, s_{2f} \in \Sigma_o, \delta(x_0, s_1) = x_1, \delta(x_0, s_2) = x_2$$

and

$$P(s_1) = P(s_2).$$

P3) Let $q_1, q_2 \in Q_d$ and $s \in \Sigma^*$ such that $(x_1, \ell_1) \in q_1$, $(x_2, \ell_2) \in q_2$, $\delta(x_1, s) = x_2$, and $\delta_d[q_1, P(s)] = q_2$. Then

$$(F_i \notin \ell_2) \wedge (A \notin \ell_2) \Rightarrow F_i \notin \ell_1.$$

Property P3) above simply states that the failure labels F_i propagate from state to state, unless replaced by the A label as a consequence of the label correction function LC. Hence, along any trace s of L, if a state x receives an F_i label, every successor x' of x also carries the F_i label, unless ambiguity arises, in which case, x' receives the A label. Also note that if along a trace $s \in L$, a state x carries the label N, then so do all of its predecessors.

Definition 6:

1) A state $q \in Q_d$ is said to be F_i-certain if $\forall (x, \ell) \in q$, $F_i \in \ell$.
2) A state $q \in Q_d$ is said to be F_i-uncertain if $\exists (x, \ell)$, $(y, \ell') \in q$, such that $F_i \in \ell$ and $F_i \notin \ell'$.

3) A state $q \in Q_d$ is said to be ambiguous if $\exists (x, \ell) \in q$, such that $A \in \ell$.

Note that in the above definition of an F_i-uncertain state, $x \neq y$ by Property P1). Also note that if a state q is not F_i-uncertain, it does not necessarily imply that q is F_i-certain, since a state $q \in Q_d$ such that $\forall (x, \ell) \in q$, $F_i \notin \ell$ is neither F_i-certain nor F_i-uncertain. The following results are a direct consequence of the construction of the diagnoser.

Lemma 1:

i) Let $\delta_d(q_0, u) = q$. If q is F_i-certain, then $[\forall \omega \in P_L^{-1}(u)]$ $\Sigma_{fi} \in \omega$.

ii) If a state $q \in Q_d$ is F_i-uncertain, then $\exists s_1, s_2 \in L$ such that: $\Sigma_{fi} \in s_1$, $\Sigma_{fi} \notin s_2$, $P(s_1) = P(s_2)$, $\delta_d[q_0, P(s_1)] = q$, and $\delta(x_0, s_1) \neq \delta(x_0, s_2)$.

iii) If a state $q \in Q_d$ is ambiguous, then $\exists s_1, s_2 \in L$ and $\exists i \in \Pi_f$ such that: $\Sigma_{fi} \in s_1$, $\Sigma_{fi} \notin s_2$, $P(s_1) = P(s_2)$, $\delta_d[q_0, P(s_1)] = q$, and $\delta(x_0, s_1) = \delta(x_0, s_2)$.

From the definition of an F_i-certain state and the above lemma, it is obvious that if the current state of the diagnoser is F_i-certain, then we can conclude that a failure of the type F_i has occurred, regardless of what the current state of G is. This is precisely the type of diagnosis that is addressed in this paper. On the other hand, presence of an F_i-uncertain state in G_d corresponds to the situation where there are two traces s_1 and s_2 in L such that s_1 contains a failure event of type F_i while s_2 does not and in addition, the traces s_1 and s_2 produce the same record of observable events. Whenever the diagnoser hits an F_i-uncertain state, we conclude that a failure of the type F_i may have occurred but it is not possible to ascertain from the observed event sequence up to that point whether the failure has indeed occurred. Finally, the presence of an ambiguous state in G_d corresponds to the situation where there are two traces s_1 and s_2 in L such that the set of all possible continuations of s_1 in L is the same as that of s_2, s_1 contains a failure event of a particular type, say F_i, while s_2 does not, and in addition the traces s_1 and s_2 produce the same record of observable events. We shall henceforth refer to such traces as F_i -ambiguous traces.

Definition 7: A set of states $x_1, x_2, \cdots, x_n \in X$ is said to form a cycle in G if $\exists s \in L(G, x_1)$ such that $s = \sigma_1 \sigma_2 \cdots \sigma_n$ and $\delta(x_l, \sigma_l) = x_{(l+1) \bmod n}$, $l = 1, 2, \cdots, n$.

The following definition of an F_i-indeterminate cycle is based upon examination of cycles in G_d and G'.

Definition 8: A set of F_i-uncertain states $q_1, q_2, \cdots, q_n \in Q_d$ is said to form an F_i-indeterminate cycle if

1) $q_1, q_2, \cdots, q_n$ form a cycle in G_d with $\delta_d(q_l, \sigma_l) = q_{l+1}, l = 1, \cdots, n-1, \delta_d(q_n, \sigma_n) = q_1$, where $\sigma_l \in \Sigma_o$, $l = 1, \cdots, n$, and

2) $\exists (x_l^k, \ell_l^k), (y_l^r, \tilde{\ell}_l^r) \in q_l$, $l = 1, \cdots, n$, $k = 1, \cdots, m$, and $r = 1, \cdots, m'$ such that

   a) $F_i \in \ell_l^k$, $F_i \notin \tilde{\ell}_l^r$ for all l, k, and r;

   b) The sequences of states $\{x_l^k\}$, $l = 1, \cdots, n$, $k = 1, \cdots, m$ and $\{y_l^r\}$, $l = 1, \cdots, n$, $r = 1, \cdots, m'$ form cycles in G' with

$$(x_l^k, \sigma_l, x_{(l+1)}^k) \in \delta_{G'}, \quad l = 1, \cdots, n-1, \quad k = 1, \cdots, m,$$

$$(x_n^k, \sigma_n, x_1^{k+1}) \in \delta_{G'}, \quad k = 1, \cdots, m-1$$

and

$$(x_n^m, \sigma_n, x_1^1) \in \delta_{G'}$$

and

$$(y_l^r, \sigma_l, y_{(l+1)}^r) \in \delta_{G'}, \quad l = 1, \cdots, n-1, \quad r = 1, \cdots, m',$$

$$(y_n^r, \sigma_n, y_1^{r+1}) \in \delta_{G'}, \quad r = 1, \cdots, m'-1$$

and

$$(y_n^{m'}, \sigma_n, y_1^1) \in \delta_{G'}.$$

In other words, an F_i-indeterminate cycle in G_d is a cycle composed exclusively of F_i-uncertain states for which there exist:

1) A corresponding cycle (of observable events) in G' involving only states that carry F_i in their labels in the cycle in G_d (this is the sequence $\{x_l^k\}$) and

2) A corresponding cycle (of observable events) in G' involving only states that do not carry F_i in their labels in the cycle in G_d (this is the sequence $\{y_l^r\}$).

Observe that in the above definition, m and m' denote the number of times the cycle $q_1, q_2, \cdots, q_n$ in G_d is completed before the cycle in G' is completed, i.e., nm and nm' are the cycle lengths in G' for $\{x_l^k\}$ and $\{y_l^r\}$, respectively.

An F_i-indeterminate cycle in G_d indicates the presence in L of two traces s_1 and s_2 of arbitrarily long length, such that they both have the same observable projection, and s_1 contains a failure event from the set Σ_{fi} while s_2 does not. The notion of an F_i-indeterminate cycle is the most crucial element in the development of necessary and sufficient conditions for diagnosability. We now present examples to better illustrate this notion.

Figs. 4 and 5 depict two different systems and their corresponding diagnosers. The diagnoser in Fig. 4 has a cycle of F_1-uncertain states, with the corresponding event sequence $\beta\gamma\delta$. Corresponding to this cycle in the diagnoser, there are two cycles in the state machine G': the first involves states 3–5 which appear with an F_1 label in the cycle in the diagnoser and the second involves states 7, 11, and 12 which carry a N label in the cycle in the diagnoser. Thus the cycle in G_d is a F_1-indeterminate cycle with $m = m' = 1$, $x_1^1 = 3$, $x_2^1 = 4$, $x_3^1 = 5$, and $y_1^1 = 7$, $y_2^1 = 11$, $y_3^1 = 12$. The diagnoser in Fig. 5 also has a cycle of F_1-uncertain states. In fact, on closer inspection, one sees that the diagnosers of the systems in Figs. 4 and 5 are identical. This time the cycle is not F_1-indeterminate, however, as there is no corresponding cycle in G' involving states that carry the F_1 label in the cycle in G_d, namely states 3–5, 9, and 10.

In the above examples, the cycle in G_d corresponds directly to a cycle in G', in the sense that the loop in G' is completed with just one completion of the loop in the diagnoser G_d, i.e., $m = m' = 1$. We now give an example of a system where more than one traversal of the loop in G_d is required to complete the loop in G'. In Fig. 6, the set $\{x_l^k\}$ in Definition 8

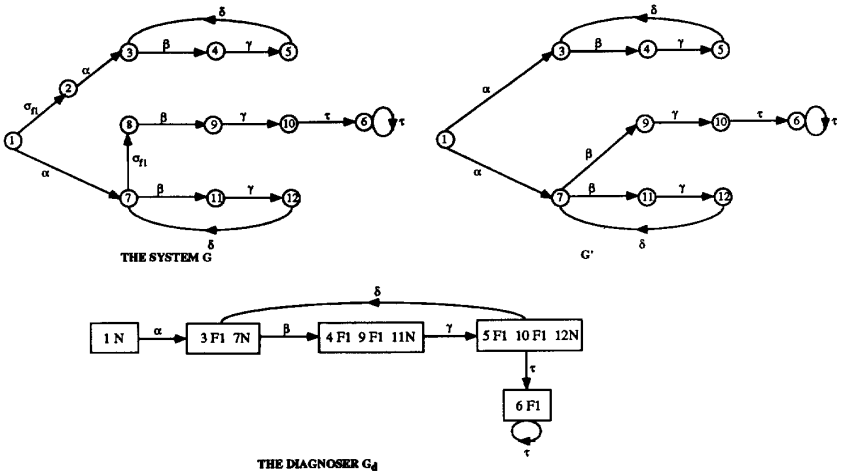


Fig. 4. Example of a system with an F_1-indeterminate cycle in its diagnoser G_d.

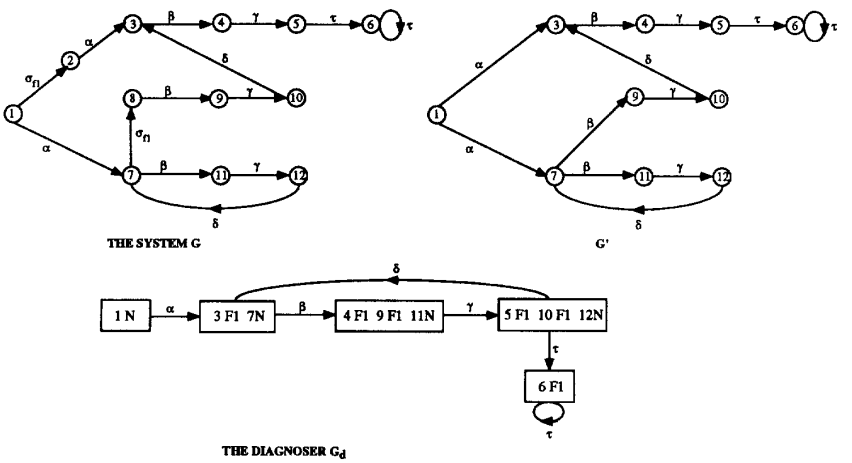


Fig. 5. Example of a system with a cycle of F_1-uncertain states in its diagnoser G_d.

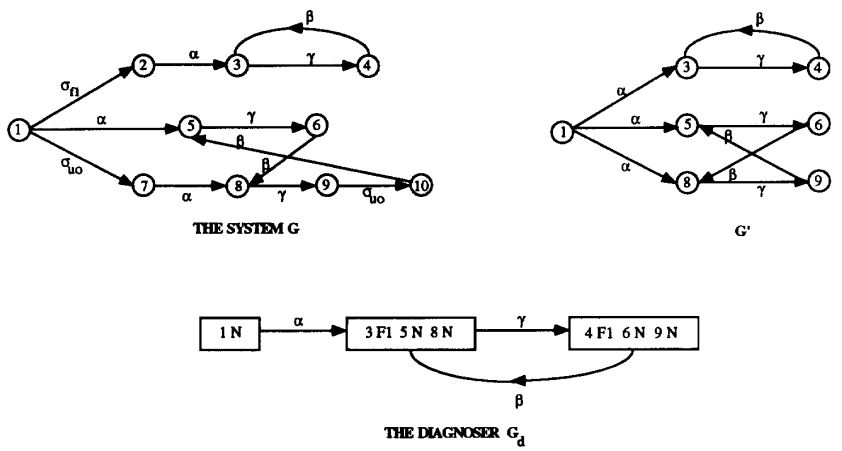


Fig. 6. Another example of a system with an F_1-indeterminate cycle in its diagnoser.

is $\{3, 4\}$ while the set $\{y_l^r\}$ is $\{5, 6, 8, 9\}$ (or, $\{8, 9, 5, 6\}$). Here $m = 1$ and $m' = 2$.

We are now ready to state the necessary and sufficient conditions for diagnosability in the case of no multiple failures.

Necessary and Sufficient Conditions:

Theorem 1: A language L without multiple failures of the same type is diagnosable if and only if its diagnoser G_d satisfies the following two conditions:

C1) There are no F_i-indeterminate cycles in G_d, for all failure types F_i.

C2) No state $q \in Q_d$ is ambiguous.

Proof:

Necessity: We first prove that if L is diagnosable, then it satisfies Condition C1). By contradiction, assume there exist states $q_1, q_2, \cdots, q_n \in Q_d$ such that they form an F_i-indeterminate cycle and let $\delta_d(q_i, \sigma_i) = q_{(i+1) \bmod n}$. Let $(x_l^k, \ell_l^k), (y_l^r, \tilde{\ell}_l^r) \in q_l$, $l = 1, \cdots, n$, $k = 1, \cdots, m$, and $r = 1, \cdots, m'$ form corresponding cycles in G' with $F_i \in \ell_l^k$, $F_i \notin \tilde{\ell}_l^r$. Then we have

$$\begin{aligned} \delta(x_l^k, s_l^k \sigma_l) &= x_{(l+1)}^k, \quad l = 1, \ldots, n-1, \\ & \qquad k = 1, \cdots, m, \\ \delta(x_n^k, s_n^k \sigma_n) &= x_1^{k+1}, \quad k = 1, \cdots, m-1, \\ \delta(x_n^m, s_n^m \sigma_n) &= x_1^1 \end{aligned}$$

and

$$\begin{aligned} \delta(y_l^r, \tilde{s}_l^r \sigma_l) &= y_{(l+1)}^r, \quad l = 1, \cdots, n-1, \\ & \qquad r = 1, \cdots, m', \\ \delta(y_n^r, \tilde{s}_n^r \sigma_n) &= y_1^{r+1}, \quad r = 1, \cdots, m'-1 \end{aligned}$$

and

$$\delta(y_n^{m'}, \tilde{s}_n^{m'} \sigma_n) = y_1^1$$

where

$$m, m' \in \mathbb{N}, \; s_l^k \in L(G, x_l^k), \; \tilde{s}_l^r \in L(G, y_l^r)$$

and

$$s_l^k, \tilde{s}_l^r \in \Sigma_{uo}^*.$$

Since $(x_1^1, \ell_1^1), (y_1^1, \tilde{\ell}_1^1) \in q_1$, $\exists\, s_0, \tilde{s}_0 \in L$ such that $\delta(x_0, s_0) = x_1^1$, $\delta(y_0, \tilde{s}_0) = y_1^1$ and $P(s_0) = P(\tilde{s}_0)$ from Property P2). Further, since $F_i \in \ell_1^1$, then $\Sigma_{fi} \in s_0$ and since $F_i \notin \tilde{\ell}_l^r$, we have that $\Sigma_{fi} \notin \tilde{s}_0$ and $\Sigma_{fi} \notin \tilde{s}_l^r$ for all l, r.

Consider the two traces

$$\begin{aligned} \omega = s_0 (& s_1^1 \sigma_1 s_2^1 \sigma_2 \cdots s_n^1 \sigma_n s_1^2 \sigma_1 s_2^2 \sigma_2 \cdots \\ & s_n^2 \sigma_n \cdots s_1^m \sigma_1 s_2^m \sigma_2 \cdots s_n^m \sigma_n)^{km'} \\ \tilde{\omega} = \tilde{s}_0 (& \tilde{s}_1^1 \sigma_1 \tilde{s}_2^1 \sigma_2 \cdots \tilde{s}_n^1 \sigma_n \tilde{s}_1^2 \sigma_1 \tilde{s}_2^2 \sigma_2 \cdots \\ & \tilde{s}_n^2 \sigma_n \cdots \tilde{s}_1^{m'} \sigma_1 \tilde{s}_2^{m'} \sigma_2 \cdots \tilde{s}_n^{m'} \sigma_n)^{km} \end{aligned}$$

for arbitrarily large k. We have that $\omega, \tilde{\omega} \in L$, $P(\omega) = P(\tilde{\omega}) = P(s_0)(\sigma_1 \sigma_2 \cdots \sigma_n)^{kmm'}$, and $\Sigma_{fi} \in \omega$ while $\Sigma_{fi} \notin \tilde{\omega}$. Let $s \in \bar{s}_0$ be such that $s \in \Psi(\Sigma_{fi})$, and let $t \in L/s$ be such that $\omega = st$. By choosing k to be arbitrarily large, we can get $||t|| > n$ for any given $n \in \mathbb{N}$. Thus, we have $\tilde{\omega} \in P_L^{-1}[P(st)]$ and $\Sigma_{fi} \notin \tilde{\omega}$. Therefore, the chosen s violates the definition of diagnosability for F_i. Hence L is not diagnosable.

We now prove that if L is diagnosable, then it satisfies Condition C2). By contradiction, let $q \in Q_d$ be ambiguous. Then for some $i \in \Pi_f$, $\exists\, F_i$-ambiguous traces $s_1, s_2 \in L$ satisfying Lemma 1-iii). Let $\delta(x_0, s_1) = x$. Since $\delta(x_0, s_1) = \delta(x_0, s_2)$, $t \in L/s_1$ iff $t \in L/s_2$. Since, by assumption, multiple failures from the same set of the partition do not occur, and since $\Sigma_{fi} \in s_1$, $\Sigma_{fi} \notin t \; \forall t \in L/s_1$. Hence, $s_2 t \in P_L^{-1}[P(s_1 t)]$ and $\Sigma_{fi} \notin s_2 t \; \forall t \in L/s_1$. Choosing $s = s_1$ and $\omega = s_2 t$, we see that Definition 1 is violated and L is not diagnosable.

Sufficiency: Assume that the diagnoser G_d for L satisfies Conditions C1) and C2). Pick any $s \in L$ and any F_i such that $s \in \Psi(\Sigma_{fi})$ and let $\delta(x_0, s) = x$. Pick any $t_1 \in L_o(G, x)$. From Assumption A2) of a finite bound n_o on the length any sequence of unobservable events in L, $||t_1|| \leq n_o$. Let $\delta(x_0, st_1) = x_1$ and correspondingly in G_d, let $\delta_d[q_0, P(st_1)] = q_1$. Since $\Sigma_{fi} \in st_1$, and since we assume that there are no ambiguous states in G_d, we have $(x_1, \ell_1) \in q_1$ with $F_i \in \ell_1$.

We now have two distinct cases to consider: I) q_1 is F_i-certain and II) q_1 is F_i-uncertain.

Case I: Suppose q_1 is F_i-certain. Then, by Lemma 1-i)

$$(\forall\, \omega \in P_L^{-1}[P(st_1)]) \quad \Sigma_{fi} \in \omega.$$

Hence L is diagnosable for F_i with $n_i = n_o$. Since this is true for any F_i, L is diagnosable.

Case II: Suppose q_1 is F_i-uncertain. Consider any $(z, \ell) \in q_1$ such that $F_i \in \ell$. We shall then refer to z as an "x-state" of q_1. Likewise, if $(z', \ell') \in q_1$ such that $F_i \notin \ell'$, then we shall denote z' as a "y-state" of q_1. We have assumed that there are no F_i-indeterminate cycles in G_d. Recalling the definition of an F_i-indeterminate cycle, this assumption means that one of the following is true: i) there are no cycles of F_i-uncertain states in G_d, or ii) there exists one or more cycles of F_i-uncertain states $q_1, q_2, \cdots, q_n$ in G_d but corresponding to any such cycle in G_d, there do not exist two sequences $\{x_l^k\}$ and $\{y_l^r\}$, $l = 1, \cdots, n$, and $k, r \in \mathbb{N}$ such that both of these form cycles in G', where the sequence $\{x_l^k\}$ is composed of "x-states" of q_l, and the sequence $\{y_l^r\}$ is composed of "y-states" of q_l, $l = 1, \cdots, n$.

Case i): Suppose that there are no cycles of F_i-uncertain states in G_d. Then this implies that every F_i-uncertain state should lead to an F_i-certain state in a bounded number of transitions by Condition C2) and by Property P3) of label propagation.

Case ii): Suppose that there exists a cycle of F_i-uncertain states $q_1, q_2, \cdots, q_n$ in G_d as in ii) above. We now show that whenever a failure happens, i.e., when the true state of the system is an "x-state," it is not possible to loop for arbitrarily long in this cycle in G_d and thereby never detect the failure.

Pick any "y-state" $y_l \in q_l$, and let the corresponding label be $\tilde{\ell}_l$. Since $F_i \notin \tilde{\ell}_l$, the pair $(y_l, \tilde{\ell}_l) \in q_l$ could only have resulted from a pair $(y_{l-1}, \tilde{\ell}_{l-1}) \in q_{l-1}$ such that $F_i \notin \tilde{\ell}_{l-1}$ and not from any $(x_{l-1}, \ell_{l-1}) \in q_{l-1}$ where $F_i \in \ell_{l-1}$, because of Condition C2) and Property P3). That is, the "y-state" y_l cannot be a successor of any "x-state" x_{l-1} along the corresponding trace in G'. Thus, by backward induction, we can always build a cycle of states in G' involving some or all of the "y-states" of q_l, $l = 1, \cdots, n$. These "y-states" then constitute the sequence $\{y_l^r\}$. But since the cycle of F_i-uncertain states $q_1, q_2, \cdots, q_n$ is not F_i-indeterminate, there cannot be a corresponding cycle in G' involving the "x-states" of q_l, i.e., there cannot exist a sequence $\{x_l^k\}$ of "x-states"

that form a cycle in G'. Hence, if we pick any "x-state" x_l in any state q_l in the cycle in G_d, then a sufficiently long trace $p \in L(G, x_l)$ [guaranteed by the liveness Assumption A1)] will leave the cycle of F_i-uncertain states. Specifically, let l^x be the number of "x-states" in any q_l, $l = 1, \cdots, n$. Then, we can stay in the cycle formed by states $q_1, q_2, \cdots, q_n$ for as long as $\sum_{l=1}^{n} l^x$ transitions of G_d, before leaving it. Since this is true for any cycle of F_i-uncertain states in G_d, we can conclude that we will eventually hit an F_i-certain state from q_l.

Therefore, for both situations i) and ii), we conclude that $\forall t_2 \in L(G, x_1)$ of sufficiently long length, $\delta_d[q_1, P(t_2)] = \delta_d[q_0, P(st_1t_2)] = q_2$ is F_i-certain. Let $t = t_1t_2$. We conclude that $\exists n_i \in \mathbb{N}$ such that $\forall t \in L/s$

$$||t|| \geq n_i \Rightarrow (\forall \omega \in P_L^{-1}[P(st)]) \quad \Sigma_{fi} \in \omega.$$

Hence L is diagnosable. Further, we can obtain a bound on n_i, $\forall i \in \Pi_f$ as follows. First, recall that $||t_1|| \leq n_o$. Next, define

$$C_i = \sum_{q \in Q_d : q \text{ is } F_i\text{-uncertain}} \#x\text{-states in } q. \tag{19}$$

Finally, recall that at most n_o unobservable events can occur between any two observable events in L. Hence we have that

$$n_i \leq C_i \times n_o + n_o. \tag{20}$$

Q.E.D.

It follows from the above proof that Conditions C1) and C2), together with the liveness assumption on L, imply that if L is indeed diagnosable, then every F_i-uncertain state leads to an F_i-certain state in a bounded number of transitions of the diagnoser. We now have the following important corollary.

Corollary 1: Consider a prefix-closed and live language L. Let Σ_{fi}, $i = 1, 2, \cdots, m$ denote disjoint sets of failure events in Σ. Assume that multiple failures of the same type do not occur in the traces in L. If L is diagnosable with delay n_i corresponding to failure type F_i, then the diagnoser G_d transitions into an F_i-certain state in at most $n_i + n_o$ events of L following the occurrence of a failure event of type F_i.

Proof: Let L be diagnosable with delay n_i corresponding to failure type F_i. From the proof of sufficiency of Conditions C1) and C2) of Theorem 1, it is obvious that every trace of L containing a failure event of type F_i leads to an F_i-certain state of the diagnoser in a bounded number of transitions. We now show that this happens in at most $n_i + n_o$ transitions of the system following the failure event. Consider any $s \in \Psi(\Sigma_{fi})$, and consider any $t \in L/s$ such that $||t|| \geq n_i$. Since L is diagnosable with delay n_i, we have that $(\forall \omega \in P_L^{-1}[P(st)])$, $\Sigma_{fi} \in \omega$. First suppose that $t_f \in \Sigma_o$. It follows then from the construction of the diagnoser G_d that $\delta_d[q_0, P(st)]$ is F_i-certain. Next suppose that $t_f \notin \Sigma_o$. Since the state of the diagnoser G_d corresponding to the trace st is defined only after the occurrence of the first observable event following st, and since the length of any sequence of unobservable events in L is bounded by n_o, we have that $\forall v \in L: (v = stu\sigma_o)(u \in \Sigma_{uo}^*)(\sigma_o \in \Sigma_o)$, $||u\sigma_o|| \leq n_o$ and $\delta_d[q_0, P(v)]$ is F_i-certain. **Q.E.D.**

Recall the systems represented in Figs. 4 and 5. In both of these systems multiple failures of the same type do not occur along any trace. The diagnoser corresponding to the system in Fig. 5 does not have any F_i-indeterminate cycle or ambiguous states, and hence this system is diagnosable. The bound on the delay n_1 for this system is calculated from (20) to be six; inspection of the system reveals that the actual value of $n_1 = 6$. Inspection of the diagnoser shows that the detection delay for this system is also six. (Here, $n_1 + n_o = 7$.) The system represented in Fig. 4 is not diagnosable since the diagnoser G_d for this system contains an F_i-indeterminate cycle as explained earlier. Fig. 3 represents another system that is not diagnosable. This again is an example of a system in which multiple failures are not possible; inspection of the diagnoser for this system reveals the presence of an ambiguous state.

Remark: One could interpret Conditions C1) and C2) of Theorem 1 as generalizations to the case of diagnosability of Özveren and Willsky's conditions for invertibility stated in [14].

2) The Case of Multiple Failures: We now consider the case of possible multiple failures from the same set of the partition. First, recall that when more than one failure event of the same failure type occurs along any trace of the system, our definition of diagnosability does not require that all of these events be detected. We only require that it be possible to conclude with finite delay (after the first occurrence of a failure) that a failure event of that type happened. This is what distinguishes the case of multiple failures from the case of no multiple failures and leads to the following consequences on the diagnosability of a language.

In the case of no multiple failures discussed in the last section, we saw that a necessary condition for L to be diagnosable is that no state of Q_d is ambiguous. In other words, L should contain no two F_i-ambiguous traces $\forall i \in \Pi_f$. Such a requirement is not necessary when we allow for the possibility of multiple failures. Recall from Lemma 1-iii) that any two F_i-ambiguous traces s_1 and s_2 produce the same record of observable events and, in addition, share the same future behavior. Thus, no future observations can help identify which of the two traces was actually executed by the system. If every trace in the post-language of these ambiguous traces contains failure events of the same type that caused the ambiguity, namely, failures from the set Σ_{fi} occurring in a bounded number of transitions of the system following the first occurrence of the failure, and if it were possible to detect with finite delay the occurrence of these failures, the language L would still satisfy our condition of diagnosability. Hence, presence of two ambiguous traces does not necessarily imply that L is not diagnosable. To determine in the case of multiple failures if L is indeed diagnosable, one needs to record what failure types caused the ambiguity and test if these failure types reappear. For these reasons, the "basic" diagnoser introduced in Section III is not adequate for checking diagnosability of a language in which multiple failures of the same type are possible. In this regard, we now introduce some modifications to the diagnoser G_d of Section III.

First define the new set of possible labels

$$\Delta^{mf} = \{N\} \cup 2^{\Delta_f} \tag{21}$$

(as opposed to $\{N\} \cup 2^{\{\Delta_f \cup \{A\}\}}$ in the previous case). The modified diagnoser for G is the FSM

$$G_d^{mf} = (Q_d^{mf}, \Sigma_o, \delta_d^{mf}, q_0). \tag{22}$$

Here $q_0 = \{(x_0, \{N\})\}$ as before, and the label propagation function LP^{mf}, the range function R, the transition function δ_d^{mf}, and the state space Q_d^{mf} of G_d^{mf} are defined as follows.

Definition 9: The label propagation function $LP^{mf}: X_o \times \Delta^{mf} \times \Sigma^* \to \Delta^{mf}$.

Given $x \in X_o$, $\ell \in \Delta^{mf}$, and $s \in L_o(G, x)$, LP^{mf} propagates the label ℓ over s, starting from x and following the dynamics of G, i.e., according to $L(G, x)$. It is defined by

$$LP^{mf}(x, \ell, s) = \begin{cases} \{N\} & \text{if } \ell = \{N\} \wedge \forall i[\Sigma_{fi} \notin s] \\ \{F_i: F_i \in \ell \vee \Sigma_{fi} \in s\} & \text{otherwise.} \end{cases}$$

The range function $R: Q_o \times \Sigma_o \to Q_o$ is the same as before but with the new LP^{mf}

$$R(q, \sigma) = \bigcup_{(x,\ell) \in q} \bigcup_{s \in L_o(G,x)} \{(\delta(x, s), LP^{mf}(x, \ell, s))\}.$$

The label correction function LC, which assigns the A label, is now dropped.

The transition function $\delta_d^{mf}: Q_o \times \Sigma_o \to Q_o$ is now defined as

$$q_2 = \delta_d^{mf}(q_1, \sigma) \Leftrightarrow q_2 = R(q_1, \sigma) \tag{23}$$

with $\sigma \in e_d(q_1)$ defined as before. The state space Q_d^{mf} is the resulting subset of Q_o composed of the states of the diagnoser that are reachable from q_0 under the transition function δ_d^{mf}.

Properties and Definitions of G_d^{mf}: We now restate some of the properties of the diagnoser and the definitions of Section IV-A-1) taking into account the modifications discussed above. Note that we have now dropped the label correction function LC. Therefore, Property P1) no longer holds, i.e., there may exist $q \in Q_d$ such that $(x, \ell), (x, \ell') \in q$ with $\ell \neq \ell'$. Property P2) remains true for the case of multiple failures. Property P3) is restated as follows.

P3)-MF): Let $q_1, q_2 \in Q_d$ and $s \in \Sigma^*$ such that $(x_1, \ell_1) \in q_1$, $(x_2, \ell_2) \in q_2$, $\delta(x_1, s) = x_2$, and $\delta_d[q_1, P(s)] = q_2$. Then

$$F_i \notin \ell_2 \Rightarrow F_i \notin \ell_1.$$

Property P3)-MF) simply states that the failure labels F_i propagate from state to state and if along any trace s of L a state x receives an F_i label, then every successor x' of x also carries the F_i label.

The definition of an F_i-certain state holds as before, and the definition of an ambiguous state is now irrelevant. We add to the definitions of an F_i-uncertain state and an F_i-indeterminate cycle, respectively, the qualifiers "x not necessarily distinct from y" and "x_l^k not necessarily distinct from y_l^r." Therefore, states $q \in Q_d$ that were ambiguous in the case of no multiple failures are now F_i-uncertain states.

Lemma 1 is restated as follows.

Lemma 2:

i) Let $\delta_d(q_0, u) = q$. If q is F_i-certain, then $(\forall \omega \in P_L^{-1}(u))\ \Sigma_{fi} \in \omega$.

ii) If a state $q \in Q_d$ is F_i-uncertain, then this implies that $\exists s_1, s_2 \in L$ such that $\Sigma_{fi} \in s_1$, $\Sigma_{fi} \notin s_2$, $P(s_1) = P(s_2)$, and $\delta_d[q_0, P(s_1)] = q$.

The proof of the above lemma is obvious by the construction of the diagnoser G_d^{mf}. Note that Lemma 1-ii) and 1-iii) of Section IV-A-1) have been restated together as Lemma 2-ii) since ambiguous states have now become F_i-uncertain states.

Fig. 7 illustrates construction of the diagnoser G_d^{mf} for the case of multiple failures. In this system α, β, γ, and δ are observable events while σ_{uo} is unobservable. The only failure event in the system is σ_{f1} and hence the partition is given by $\Sigma_{f1} = \{\sigma_{f1}\}$.

Necessary and Sufficient Conditions:

Theorem 2: A language L is diagnosable if and only if its diagnoser G_d^{mf} satisfies the following condition:

C-MF): There are no F_i-indeterminate cycles in G_d, for all failure types F_i.

Proof: The proof of the necessity of the above condition is identical to the proof of the necessity of Condition C1) of Theorem 1 since the latter proof does not require that the x_l^ks and the y_l^rs be distinct. The proof of the sufficiency of Condition C-MF) is essentially the same as the proof of the sufficiency of Conditions C1) and C2) of Theorem 1. The only difference is that the absence of ambiguous states is true by assumption in the case of Theorem 1 whereas it is true by construction in the present case. Hence, reasoning along lines identical to the proof of Theorem 1, we conclude that the condition of no F_i-indeterminate cycles in G_d^{mf}, for all failure types F_i, is necessary and sufficient for L to be diagnosable in the case of multiple failures. Further, reasoning as before, we have the following bound on the delay n_i, $\forall i \in \Pi_f$

$$n_i \leq C_i^{mf} \times n_o + n_o \tag{24}$$

where

$$C_i^{mf} = \sum_{q \in Q_d^{mf}:\, q \text{ is } F_i\text{-uncertain}} \#x\text{-states in } q. \tag{25}$$

This, however, is a very conservative bound. In Section V we shall provide a better bound on the delay n_i. **Q.E.D.**

As before, note that Condition C-MF), together with the liveness assumption on L, implies that if L is diagnosable, then every F_i-uncertain state of the diagnoser G_d^{mf} leads to an F_i-certain state in a bounded number of transitions of G_d^{mf}. Hence we have the following corollary, whose proof is analogous to that of Corollary 1.

Corollary 2: Consider a prefix-closed and live language L. Let Σ_{fi}, $i = 1, 2, \cdots, m$ denote disjoint sets of failure events in Σ. If L is diagnosable with delay n_i corresponding to failure type F_i, then the diagnoser G_d^{mf} transitions into an F_i-certain state in at most $n_i + n_o$ events of L following the occurrence of a failure event of type F_i.

Fig. 7 represents a system where multiple failures of the same type are possible. This system is diagnosable since it

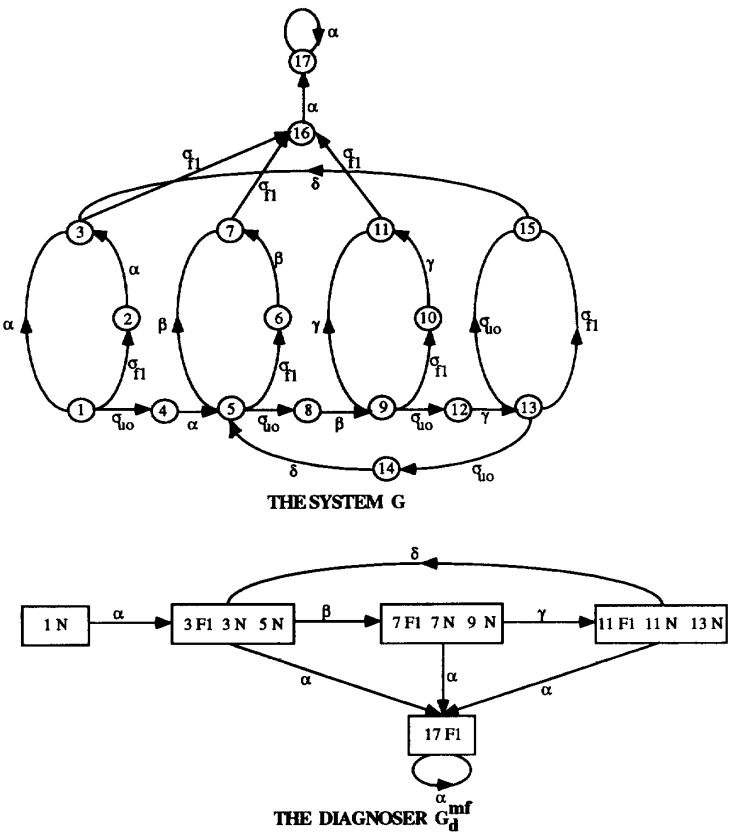


Fig. 7. Example illustrating construction of the diagnoser G_d^{mf} for the case of multiple failures.

is easily verified that the cycle of F_i-uncertain states in the corresponding diagnoser G_d^{mf} is not F_i-indeterminate.

B. Conditions for I-Diagnosability

We now study necessary and sufficient conditions for a language to be I-diagnosable. Recall from Section II-B-2) that in the case of I-diagnosability we are interested in detecting failure events only after the occurrence of the corresponding indicator events, i.e., we require the diagnosability condition D to hold only for those traces in which an indicator event follows a failure event. Based on this requirement, we introduce the following modifications to the basic diagnoser G_d.

We define, as before, the set of failure labels $\Delta_f = \{F_1, F_2, \cdots F_m\}$ where $|\Pi_f| = m$. In addition, we now define a set of indicator labels $\Delta_i = \{I_1, I_2, \cdots I_m\}$. We interpret $\{I_{i1}, \cdots, I_{ik}\}$ as meaning that indicator events of the type I_{i1} through I_{ik} have occurred. The complete set of possible labels is now defined as

$$\Delta^I = \{N\} \cup 2^{\Delta_f \cup \Delta_i} \tag{26}$$

with the restriction that

$$(\forall \ell \in \Delta^I) \quad I_i \in \ell \Rightarrow F_i \in \ell$$

(explained in the subsequent paragraphs).

The modified diagnoser G_d^I is the FSM

$$G_d^I = (Q_d^I, \Sigma_o, \delta_d^I, q_0) \tag{27}$$

with the initial state $q_0 = \{(x_0, \{N\})\}$ as in Section III. The label propagation function LP^I, the range function R, the label correction function LC^I, the transition function δ_d^I, and the state space Q_d^I of G_d^I are defined as follows.

Definition 10: The label propagation function $LP^I: X_o \times \Delta^I \times \Sigma^* \to \Delta^I$.

Given $x \in X_o$, $\ell \in \Delta^I$, and $s \in L_o(G, x)$, LP^I propagates the label ℓ over s, starting from x and following the dynamics of G, i.e., according to $L(G, x)$. It is defined by

$$LP^I(x, \ell, s) = \begin{cases} \{N\} & \text{if } \ell = \{N\} \wedge \forall i[\Sigma_{fi} \notin s] \\ \{F_i: F_i \in \ell \vee \Sigma_{fi} \in s\} & \\ \quad \bigcup \{I_i: I_i \in \ell \vee [I(\Sigma_{fi}) & \\ \quad \in s \wedge (F_i \in \ell \vee \Sigma_{fi} \in s)]\} & \text{otherwise.} \end{cases}$$

Fig. 8 illustrates propagation of labels according to LP^I as defined above. Here σ_{f1} refers to a failure event of type F_1 and σ_{I1} refers to an indicator event of type I_1. The range function $R: Q_o \times \Sigma_o \to Q_o$ is defined to be

$$R(q, \sigma) = \bigcup_{(x,\ell) \in q} \bigcup_{s \in L_\sigma(G,x)} \{(\delta(x, s), LP^I(x, \ell, s))\}.$$

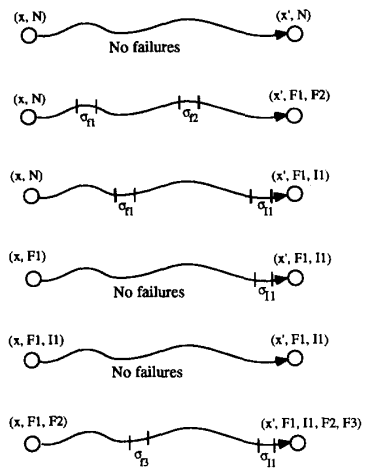


Fig. 8. Figure illustrating propagation of labels along traces of L.

Definition 11: The label correction function $LC^I: Q_o \rightarrow Q_o$ is defined as follows

$$LC^I(q) = q - \{(x, \ell) \in q : (x, \ell') \in q \wedge \forall i[F_i \in \ell \Leftrightarrow F_i \in \ell'] \wedge [\ell \subset \ell']\}.$$

The use of the label correction function LC^I is explained as follows. Suppose there exist two pairs (x, ℓ) and (x, ℓ') as described above, in $R(q, \sigma)$ for some state q of G_d^I. This implies the presence in L of two traces s_1 and s_2 such that they have identical projections and lead to the same state x, and s_1 contains an indicator event of type I_i following a failure event of type F_i while s_2 does not. Since for I-diagnosability, we are concerned only with traces in which the indicator event follows the failure, we can drop the pair (x, ℓ) which does not contain the I_i label with no loss of generality.

The transition function $\delta_d^I: Q_o \times \Sigma_o \rightarrow Q_o$ is defined as

$$q_2 = \delta_d^I(q_1, \sigma) \Leftrightarrow q_2 = LC^I[R(q_1, \sigma)] \tag{28}$$

with $\sigma \in e_d(q_1)$ defined as before. The state space Q_d^I is the resulting subset of Q_o composed of the states of the diagnoser that are reachable from q_0 under the transition function δ_d^I. A state q_d of G_d^I is now of the form

$$q_d = \{(x_1, \ell_1), \cdots, (x_n, \ell_n)\}$$

where $x_i \in X_o$ and $\ell_i \in \Delta^I$, i.e., ℓ_i is of the form $\ell_i = \{N\}$ or $\ell_i = \{F_{i_1}, F_{i_2}, \cdots, F_{i_k}, I_{j_1} I_{j_2}, \cdots, I_{j_l}\}$ where $\{i_1, i_2, \cdots, i_k\} \subseteq \{1, 2, \cdots, m\}$ and $\{j_1, j_2, \cdots, j_l\} \subseteq \{i_1, i_2, \cdots, i_k\}$.

We make the following observations on the modified diagnoser G_d^I:

1) In addition to failure information, the labels now carry information on occurrences of indicator events following the failure events.
2) We append the I_i label to any ℓ only if an indicator event from $I(\Sigma_{fi})$ follows a failure event from Σ_{fi}. The set of I_i labels is always a subset of the set of F_i labels in any (x, ℓ) pair $\in q \in Q_d$.
3) The I_i labels propagate from state to state just like the F_i labels.
4) We do not use the A label here. As mentioned earlier, we are now concerned only with traces where the failure event is followed by an appropriate indicator event. Therefore, there could be present in L two F_i-ambiguous traces for some $i \in \Pi_f$ and yet L could be diagnosable if no trace in the post-language of these traces contains an indicator event from the set $I(\Sigma_{fi})$. Hence, to check for I-diagnosability, we need to remember which failure types caused the ambiguity, even in the case of no multiple failures. Therefore, we do not need to distinguish between the case of possible multiple failures and the case of no multiple failures in this section.

Fig. 9 illustrates the construction of the diagnoser G_d^I. Here, $\sigma_i, i \in \{1, \cdots, 4\}$, and σ_{I1} are observable events while σ_{uo} is unobservable. The indicator event corresponding to the failure event σ_{f1} is $I(\sigma_{f1}) = \{\sigma_{I1}\}$ and the partition is $\Sigma_{f1} = \{\sigma_{f1}\}$.

Properties and Definitions of G_d^I: Since we do not use the A label, properties P1) through P3) of the diagnoser correspond to those discussed in Section IV-A-2). Likewise, the remarks on the definition of an F_i-certain state, an F_i-uncertain state, and an F_i-indeterminate cycle, and Lemma 2 cited in Section IV-A-2) remain valid. We now introduce the notions of (F_i, I_i)-uncertain states and (F_i, I_i)-indeterminate cycles.

Definition 12: A state $q \in Q_d^I$ is said to be (F_i, I_i)-uncertain if $\exists (x, \ell), (y, \ell') \in q$, such that $\{F_i, I_i\} \subseteq \ell$ and $F_i \notin \ell'$.

Lemma 3: If a state $q \in Q_d$ is (F_i, I_i)-uncertain, then this implies that $\exists s_1 = p_1 t_1 \in L$ and $s_2 \in L$ such that: $p_1 \in \Psi(\Sigma_{fi})$, $I(\Sigma_{fi}) \in t_1$, $\Sigma_{fi} \notin s_2$, $P(s_1) = P(s_2)$, and $\delta_d^I[q_0, P(s_1)] = q$.

The above lemma simply states that presence of an (F_i, I_i)-uncertain state in G_d^I corresponds to the situation where there are two traces s_1 and s_2 in L such that s_1 contains a failure event of type F_i followed by an indicator event corresponding to this failure type while s_2 does not contain a failure event of type F_i. In addition, the traces s_1 and s_2 produce the same record of observable events. Proof of this lemma follows directly from the construction of G_d^I.

Definition 13: A set of (F_i, I_i)-uncertain states $q_1, q_2, \cdots, q_n \in Q_d^I$ is said to form an (F_i, I_i)-indeterminate cycle if:

1) $q_1, q_2, \cdots, q_n$ form a cycle in G_d^I with $\delta_d^I(q_l, \sigma_l) = q_{l+1}$, $l = 1, \cdots, n-1$ and $\delta_d^I(q_n, \sigma_n) = q_1$ where $\sigma_l \in \Sigma_o$, $l = 1, \cdots, n$, and
2) $\exists (x_l^k, \ell_l^k), (y_l^r, \tilde{\ell}_l^r) \in q_l$, $l = 1, \cdots, n$, $k = 1, \cdots, m$, and $r = 1, \cdots, m'$ (x not necessarily distinct from y), such that
   a) $\{F_i, I_i\} \subseteq \ell_l^k$, $F_i \notin \tilde{\ell}_l^r$ for all l, k, and r;
   b) The sequences of states $\{x_l^k\}$, $l = 1, \cdots, n$, $k = 1, \cdots, m$, and $\{y_l^r\}$, $l = 1, \cdots, n$, $r = 1, \cdots, m'$ form cycles in G' with

$$(x_l^k, \sigma_l, x_{(l+1)}^k) \in \delta_{G'}, \quad l = 1, \ldots, n-1, \quad k = 1, \cdots, m,$$

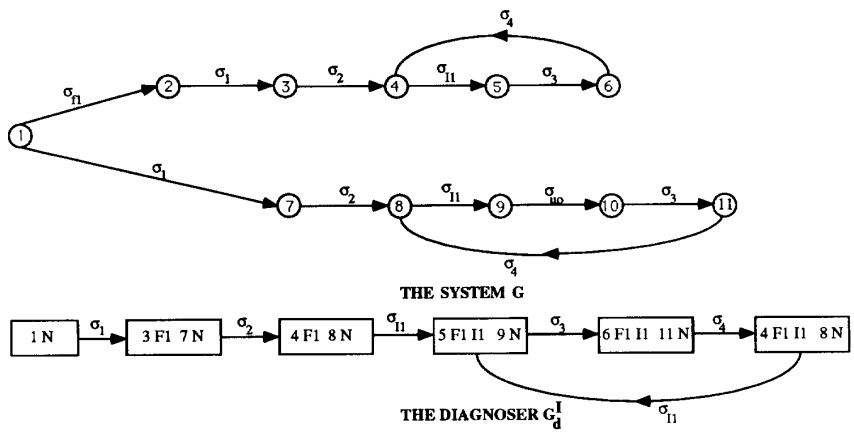


Fig. 9. Example illustrating construction of the diagnoser G_d^I.

$$(x_n^k,\ \sigma_n,\ x_1^{k+1}) \in \delta_{G'}, \quad k = 1, \ldots, m-1$$

and

$$(x_n^m,\ \sigma_n,\ x_1^1) \in \delta_{G'}$$

and

$$(y_l^r,\ \sigma_l,\ y_{(l+1)}^r) \in \delta_{G'}, \quad l = 1, \cdots, n-1, \quad r = 1, \cdots, m',$$
$$(y_n^r,\ \sigma_n,\ y_1^{r+1}) \in \delta_{G'}, \quad r = 1, \cdots, m'-1$$

and

$$(y_n^{m'},\ \sigma_n,\ y_1^1) \in \delta_{G'}.$$

An (F_i, I_i)-indeterminate cycle in G_d^I indicates the presence in L of two traces s_1 and s_2 of arbitrarily long length, such that they both have the same observable projection, and s_1 contains a failure event from the set Σ_{fi} followed by an indicator event from the set $I(\Sigma_{fi})$ while s_2 does not contain any event from the set Σ_{fi}.

Consider the system shown in Fig. 9. Inspection of the diagnoser G_d^I for this system reveals the presence of an (F_i, I_i)-indeterminate cycle. Here the set $\{x_l^k\}$ of Definition 13 is $\{5, 6, 4\}$ (these states carry the label $\{F_1, I_1\}$ in the diagnoser G_d^I), the set $\{y_l^r\}$ is $\{9, 11, 8\}$ (these states carry the label $\{N\}$), and $m = m' = 1$.

Necessary and Sufficient Conditions:

Theorem 3: A language L is I-diagnosable if and only if the diagnoser G_d^I satisfies the following condition:

C-I) There are no (F_i, I_i)-indeterminate cycles in G_d^I, for all failure types F_i.

Proof: This proof is very similar to the proof of Theorem 1 with the exceptions that we now consider the I labels and (F_i, I_i)-indeterminate cycles and that there are no ambiguous states. For the sake of clarity, it is presented in its entirety.

Necessity: We prove necessity by contradiction. Assume there exist states $q_1, q_2, \cdots, q_n \in Q_d^I$ such that they form an (F_i, I_i)-indeterminate cycle and let $\delta_d^I(q_i, \sigma_i) = q_{(i+1) \bmod n}$. Let (x_l^k, ℓ_l^k), $(y_l^r, \tilde{\ell}_l^r) \in q_l$, $l = 1, \cdots, n$, $k = 1, \cdots, m$, and $r = 1, \cdots, m'$ form corresponding cycles in G' with $\{F_i, I_i\} \subseteq \ell_l^k$, $F_i \notin \tilde{\ell}_l^r$. Then we have

$$\delta(x_l^k,\ s_l^k\sigma_l) = x_{(l+1)}^k, \quad l = 1, \cdots, n-1, \quad k = 1, \cdots, m$$
$$\delta(x_n^k,\ s_n^k\sigma_n) = x_1^{k+1}, \quad k = 1, \cdots, m-1$$

and

$$\delta(x_n^m,\ s_n^m\sigma_n) = x_1^1$$

and

$$\delta(y_l^r,\ \tilde{s}_l^r\sigma_l) = y_{(l+1)}^r, \quad l = 1, \cdots, n-1, \quad r = 1, \cdots, m'$$
$$\delta(y_n^r,\ \tilde{s}_n^r\sigma_n) = y_1^{r+1}, \ r = 1, \cdots, m'-1$$

and

$$\delta(y_n^{m'},\ \tilde{s}_n^{m'}\sigma_n) = y_1^1$$

where

$$m, m' \in \mathbb{N},\ s_l^k \in L(G, x_l^k),\ \tilde{s}_l^r \in L(G, y_l^r)$$

and

$$s_l^k,\ \tilde{s}_l^r \in \Sigma_{uo}^*.$$

Since $(x_1^1, \ell_1^1), (y_1^1, \tilde{\ell}_1^1) \in q_1, \exists\, s_0, \tilde{s}_0 \in L$ such that $\delta(x_0, s_0) = x_1^1$, $\delta(y_0, \tilde{s}_0) = y_1^1$ and $P(s_0) = P(\tilde{s}_0)$ from Property P2). Further, since $\{F_i, I_i\} \subseteq \ell_1^1$, then $\Sigma_{fi} \in s_0$, and $\exists\, st_1 \in \overline{s_0}$ such that $s \in \Psi(\Sigma_{fi})$ and $st_1 \in \Psi[I(\Sigma_{fi})]$, i.e., the trace st_1 contains a failure event of the type F_i and ends in an indicator event associated with the failure type F_i. Also, since $F_i \notin \tilde{\ell}_l^r$, we have that $\Sigma_{fi} \notin \tilde{s}_0$ and $\Sigma_{fi} \notin \tilde{s}_l^r$, for all l, r.

Consider the two traces

$$\omega = s_0(s_1^1\sigma_1 s_2^1\sigma_2 \cdots s_n^1\sigma_n s_1^2\sigma_1 s_2^2\sigma_2 \cdots s_n^2\sigma_n \cdots s_1^m\sigma_1 s_2^m\sigma_2 \cdots s_n^m\sigma_n)^{km'}$$

and

$$\tilde{\omega} = \tilde{s}_0(\tilde{s}_1^1\sigma_1\tilde{s}_2^1\sigma_2 \cdots \tilde{s}_n^1\sigma_n\tilde{s}_1^2\sigma_1\tilde{s}_2^2\sigma_2 \cdots \tilde{s}_n^2\sigma_n \cdots \tilde{s}_1^{m'}\sigma_1\tilde{s}_2^{m'}\sigma_2 \cdots \tilde{s}_n^{m'}\sigma_n)^{km}$$

for arbitrarily large k. We have that $\omega, \tilde{\omega} \in L$, and $P(\omega) = P(\tilde{\omega}) = P(s_0)(\sigma_1\sigma_2 \cdots \sigma_n)^{kmm'}$. Let $t_2 \in L/st_1$ be such that $\omega = st_1t_2$. By choosing k to be arbitrarily large, we can

get $||t_2|| > n$ for any given $n \in \mathbb{N}$. Thus, $\tilde{\omega} \in P_L^{-1}[P(st)]$ and $\Sigma_{fi} \notin \tilde{\omega}$. Therefore, the chosen s violates the definition of I-diagnosability for F_i. Hence L is not I-diagnosable.

Sufficiency: Assume that the diagnoser G_d^I for L satisfies Condition C-I). Pick any $s \in L$ and any F_i such that $s \in \Psi(\Sigma_{fi})$ and any $t_1 \in L/s$ such that $t_{1f} \in I(\Sigma_{fi})$. Let $\delta(x_0, st_1) = x_1$, and correspondingly in G_d^I, let $\delta_d^I[q_0, P(st_1)] = q_1$. Since $\Sigma_{fi} \in st_1$ and $t_{1f} \in I(\Sigma_{fi})$, we have $(x_1, \ell_1) \in q_1$ with $\{F_i, I_i\} \subseteq \ell_1$.

We now have two distinct cases to consider: I) q_1 is F_i-certain, and II) q_1 is (F_i, I_i)-uncertain.

Case I: Suppose q_1 is F_i-certain. Then, by Lemma 2-i)

$$(\forall \omega \in P_L^{-1}[P(st_1)]) \quad \Sigma_{fi} \in \omega.$$

Hence L is I-diagnosable for F_i with $n_i = 0$. Since this is true for any F_i, L is I-diagnosable.

Case II: Suppose q_1 is (F_i, I_i)-uncertain. We have assumed that there are no (F_i, I_i)-indeterminate cycles in G_d^I. Recalling the definition of an (F_i, I_i)-indeterminate cycle, this assumption means that one of the following is true: i) there are no cycles of (F_i, I_i)-uncertain states in G_d^I, and ii) there exists one or more cycles of (F_i, I_i)-uncertain states in G_d^I but corresponding to any of these cycles in G_d^I, there do not exist two sequences $\{x_l^k\}$ and $\{y_l^r\}$, $l = 1, \cdots, n$ and $k, r \in \mathbb{N}$ such that both of these form cycles in G', where the sequence $\{x_l^k\}$ is composed of states that appear with an $\{F_i, I_i\}$ label in the cycle in G_d^I while the sequence $\{y_l^r\}$ is composed of states that do not appear with an F_i label.

Reasoning along lines similar to the proof of sufficiency of Conditions C1) and C2) of Theorem 1, we conclude as before that $\forall t_2 \in L(G, x_1)$ of sufficiently long length, $\delta_d^I[q_1, P(t_2)] = \delta_d^I[q_0, P(st_1t_2)] = q_2$ is F_i-certain. Note that q_2 cannot be F_i-uncertain. This is because q_1 is (F_i, I_i)-uncertain, and no (F_i, I_i)-uncertain state of G_d^I can lead to an F_i-uncertain state since the I_i labels propagate from state to state. Hence we conclude that $\exists\, n_i \in \mathbb{N}$ such that $\forall t_2 \in L/st_1$

$$||t_2|| \geq n_i \Rightarrow (\forall \omega \in P_L^{-1}[P(st_1t_2)]) \quad \Sigma_{fi} \in \omega.$$

Hence L is I-diagnosable. Further

$$n_i \leq C_i^I \times n_o \tag{29}$$

where

$$C_i^I = \sum_{q \in Q_d^I : q \text{ is } (F_i, I_i)\text{-uncertain}} \#x\text{-states in } q. \tag{30}$$

We note here that this bound on the delay n_i is conservative; in Section V we provide a better bound. **Q.E.D.**

Again, note that Condition C-I) and the liveness assumption on L together imply that if L is indeed I-diagnosable, every (F_i, I_i)-uncertain state leads to an F_i-certain state in a bounded number of transitions of the diagnoser. Also note that in the case of I-diagnosability, we are not concerned about F_i-uncertain states and F_i-indeterminate cycles which are not also (F_i, I_i)-uncertain and (F_i, I_i)-indeterminate, respectively.

Corollary 3: Consider a prefix-closed and live language L. Let Σ_{fi}, $i = 1, 2, \cdots, m$ denote disjoint sets of failure events in Σ, and let $I(\Sigma_{fi})$ denote the corresponding sets of indicator events. If L is I-diagnosable with delay n_i corresponding to failure type F_i, then the diagnoser G_d^I transitions into an F_i-certain state in at most $n_i + n_o$ events of L after the occurrence of an indicator event of type I_i following a failure event of type F_i.

Proof: Let L be I-diagnosable with delay n_i corresponding to failure type F_i. Proof of the sufficiency of Condition C-I) of Theorem 3 reveals that if L is I-diagnosable, then every trace of L containing a failure event of type F_i, followed by an indicator event of type I_i, leads to an F_i-certain state of G_d^I in a bounded number of transitions. We now show that this happens in at most $n_i + n_o$ transitions of the system following the indicator event. Consider any $s \in \Psi(\Sigma_{fi})$, and consider any $t_1t_2 \in L/s$ such that $st_1 \in I[\Psi(\Sigma_{f1})]$ and $||t_2|| \geq n_i$. Since L is diagnosable with delay n_i, we have that $(\forall \omega \in P_L^{-1}[P(st_1t_2)])\, \Sigma_{fi} \in \omega$. First suppose that $t_{2_f} \in \Sigma_o$. It follows then from the construction of the diagnoser G_d^I that $\delta_d^I[q_0, P(st_1t_2)]$ is F_i-certain. Next suppose that $t_{2_f} \notin \Sigma_o$. Since the state of the diagnoser G_d^I corresponding to the trace st_1t_2 is defined only after the occurrence of the first observable event following st_1t_2 and since the length of any sequence of unobservable events in L is bounded by n_o, we have that $\forall v \in L: (v = st_1t_2u\sigma_o)(u \in \Sigma_{uo}^*)(\sigma_o \in \Sigma_o)$, $||u\sigma_o|| \leq n_o$ and $\delta_d^I[q_0, P(v)]$ is F_i-certain. **Q.E.D.**

Fig. 9 provides an example of a system that is not I-diagnosable since the corresponding diagnoser G_d^I contains an (F_i, I_i)-indeterminate cycle.

This concludes the discussion on necessary and sufficient conditions for diagnosability and I-diagnosability. Note that checking for diagnosability and I-diagnosability amounts to cycle detection in the diagnosers and in G' and any of the standard cycle detection algorithms (which are of polynomial complexity) may be used.

V. On-Line Diagnosis of Diagnosable Systems

We show in this section that the basic diagnoser G_d introduced in Section III is adequate for diagnosing failures in diagnosable and I-diagnosable systems, with or without multiple failures. In other words, once it is established that L is diagnosable or I-diagnosable, we can restrict attention to G_d (as opposed to G_d^{mf} and G_d^I) for performing diagnostics; occurrences of failures in the system can be detected with a finite delay by inspecting the states of this diagnoser. This result is important from an implementation viewpoint, as G_d will in general have far fewer states than its counterparts G_d^{mf} and G_d^I.

Theorem 4: Consider a prefix-closed and live language L. Let Σ_{fi}, $i = 1, 2, \cdots, m$ denote disjoint sets of failure events in Σ. If L is diagnosable (respectively, I-diagnosable) with delay n_i corresponding to failure type F_i, then the diagnoser G_d detects occurrences of failure events of the type F_i in at most $n_i + n_o$ events of L after the occurrence of the failure events (respectively, after the occurrence of indicator events of type I_i following the failure events).

Proof:

Case I—Diagnosability: We first consider the case where L is such that multiple failures of the same type do not occur along any trace. Let L be diagnosable with delay n_i corresponding to failure type F_i. First recall that we conclude, by inspection of the states of the diagnoser, that a failure of type F_i has occurred when the diagnoser hits an F_i-certain state. It then follows directly from Corollary 1 that the diagnoser G_d detects occurrences of failures of type F_i with a delay of at most $n_i + n_o$ events.

Consider next the case where L is such that multiple failures of the same type are possible. Suppose that we again construct the diagnoser G_d. First, note that the only difference between G_d^{mf} [discussed in Section IV-A-2)] and G_d is in the treatment of ambiguous states. We have that every F_i-uncertain state of G_d^{mf} corresponds uniquely either to an ambiguous state or to an F_i-uncertain state of G_d, and every F_i-certain state of G_d^{mf} corresponds to a unique F_i-certain state of G_d. To be more specific, an F_i-uncertain state $q \in Q_d^{mf}$ such that $q = \{(x, \ell), (y, \ell')\}$ with $x \neq y$ will also be a state of G_d; an F_i-uncertain state $q \in Q_d^{mf}$ such that $q = \{(x, \ell), (x, \ell')\}$ will correspond to the state $q' = \{(x, \{A\} \cup \ell \cap \ell')\}$ of G_d; finally, two states $q_1, q_2 \in Q_d^{mf}$ such that $q_1 = \{(x, \ell), (x, \ell'), (y_1, \ell_1), \cdots, (y_k, \ell_k)\}$ and $q_2 = \{(x, \tilde{\ell}), (x, \tilde{\ell}'), (y_1, \ell_1), \cdots, (y_k, \ell_k)\}$ where q_1 is F_i-uncertain (due to ℓ and ℓ'), q_2 is F_j-uncertain (due to $\tilde{\ell}$ and $\tilde{\ell}'$), and $\ell \cap \ell' = \tilde{\ell} \cap \tilde{\ell}'$ will both correspond to the same ambiguous state $q_3 = \{(x, \{A\} \cup \ell \cap \ell'), (y_1, \ell_1), \cdots, (y_k, \ell_k)\} \in Q_d$. Note here that $L(G_d^{mf}, q) = L(G_d, q')$ and $L(G_d^{mf}, q_1) = L(G_d^{mf}, q_2) = L(G_d, q_3)$. Hence, if one considers a mapping of the states of G_d^{mf} onto the states of G_d, this map preserves the transition structure of G_d^{mf} in the sense of i) preserving the language generated by G_d^{mf} and ii) preserving the essential information for implementing diagnostics because whenever G_d^{mf} hits an F_i-certain state, so would G_d. It follows from Corollary 2 and the above reasoning that if L is diagnosable with delay n_i corresponding to failure type F_i, then the diagnoser G_d^{mf}, and consequently, the diagnoser G_d hits an F_i-certain state in at most $n_i + n_o$ events following the failure event. Thus, G_d detects occurrences of failures of the type F_i with a delay of at most $n_i + n_o$ events.

Case II—I-Diagnosability: Let L be I-diagnosable with delay n_i corresponding to failure type F_i. Suppose that we again construct the basic diagnoser G_d for L. As in the case of G_d^{mf}, every F_i-uncertain state of G_d^I corresponds uniquely to either an ambiguous state or an F_i-uncertain state of G_d, and every F_i-certain state of G_d^I corresponds to a unique F_i-certain state of G_d. Every (F_i, I_i)-uncertain state of G_d^I corresponds uniquely either to an ambiguous state or to an F_i-uncertain state of G_d. For example, any two states $q_1, q_2 \in Q_d^I$ of the form $q_1 = \{(x, \{F_i, I_i\}), (y_1, \ell_1), \cdots, (y_k, \ell_k)\}$ and $q_2 = \{(x, \{F_i\}), (y_1, \ell_1), \cdots, (y_k, \ell_k)\}$ correspond to the same state $q_2 = \{(x, \{F_i\}), (y_1, \ell_1), \cdots, (y_k, \ell_k)\}$ in G_d. Note again that $L(G_d^I, q_1) = L(G_d^I, q_2) = L(G_d, q_2)$. Hence, as before, if one considers a mapping of the states of G_d^I onto the states of G_d, this map preserves the transition structure of G_d^I in the sense of i) preserving the language generated by G_d^I and ii) preserving the essential information for implementing diagnostics because whenever G_d^I hits an F_i-certain state, so would G_d. From Corollary 3, we conclude that every trace of L containing a failure event of type F_i, followed by an indicator event of type I_i, leads to an F_i-certain state of G_d^I and consequently, to an F_i-certain state of G_d in at most $n_i + n_o$ events after the occurrence of the indicator event of the corresponding type. **Q.E.D.**

Based on the above theorem, we now improve upon the bounds on the delay n_i provided in Sections IV-A-2) and IV-B (cf., proofs of Theorems 2 and 3) for diagnosability in the case of multiple failures and for I-diagnosability, respectively. Recall from the proof of sufficiency of condition C-MF) of Theorem 2 that a bound on the delay n_i is given by $C_i^{mf} \times n_o + n_o$ where $C_i^{mf} = \sum_{q \in Q_d^{mf}: q \text{ is } F_i\text{-uncertain}} \#x$-states in q. We now provide a better bound on n_i which is given by $n_i \leq C_i \times n_o + n_o$ as in the case of no multiple failures. Note that this bound depends only on the states of the basic diagnoser G_d and not on the states of G_d^{mf}. The improved bound can be obtained as follows. First, recall that to obtain a bound on n_i for the case of multiple failures, we count the number of F_i-uncertain states in G_d^{mf} (that it is possible to visit before hitting an F_i-certain state). Next, recall from the proof of Theorem 4 that there exist in G_d^{mf} states of the form q_1 and q_2 as described there which have the property that $L(G_d^{mf}, q_1) = L(G_d^{mf}, q_2)$. It is obvious then that is not necessary to count more than once "duplicate" states like q_1 and q_2 because any trace passing through q_1 cannot pass through q_2, and vice-versa. Further, note that since both q_1 and q_2 correspond to the same state q_3 in G_d, these duplicate states get accounted for only once when we compute the bound in G_d.

Likewise, in the case of I-diagnosability, we can obtain a bound on the delay n_i that is better than the one presented in Section IV-C, namely, $n_i \leq C_i^I \times n_o$ where $C_i^I = \sum_{q \in Q_d^I: q \text{ is } F_i\text{-uncertain}} \#x$-states in q. The new bound depends only on the states of the diagnoser G_d and is given by $n_i \leq C_i \times n_o$. Note, as in the case of multiple failures discussed above, that "duplicated" states in G_d^I, of the form q_1 and q_2 described in the proof of Theorem 4, get accounted for twice when one counts the number of (F_i, I_i)-uncertain states that might be traversed before hitting an F_i-certain state in G_d^I, whereas these get accounted for (and correctly) only once in G_d. Hence we have the improved bound stated above.

We conclude, therefore, that in all cases, the bound on the detection delay n_i can be given as follows

$$n_i \leq C_i \times n_o + n_o \tag{31}$$

where $C_i = \sum_{q \in Q_d: q \text{ is } F_i\text{-uncertain}} \#x$-states in q.

We now provide an example that illustrates implementation of diagnostics for an I-diagnosable system using the diagnoser G_d. The system G, the diagnoser G_d^I, and the diagnoser G_d that is implemented are shown in Fig. 10. Here, the events $\alpha, \beta, \gamma, \delta, \sigma_{I1}, \sigma_{I2}$, and σ_{I3} are observable while σ_{uo} and the failure events $\sigma_{f1}, \sigma_{f2}, \sigma_{f3}$ are unobservable. The indicator events are chosen to be $I(\sigma_{f1}) = \{\sigma_{I1}\}$, $I(\sigma_{f2}) = \{\sigma_{I2}\}$, and $I(\sigma_{f3}) = \{\sigma_{I3}\}$; the partition is chosen to be $\Sigma_{f1} = \{\sigma_{f1}\}$,

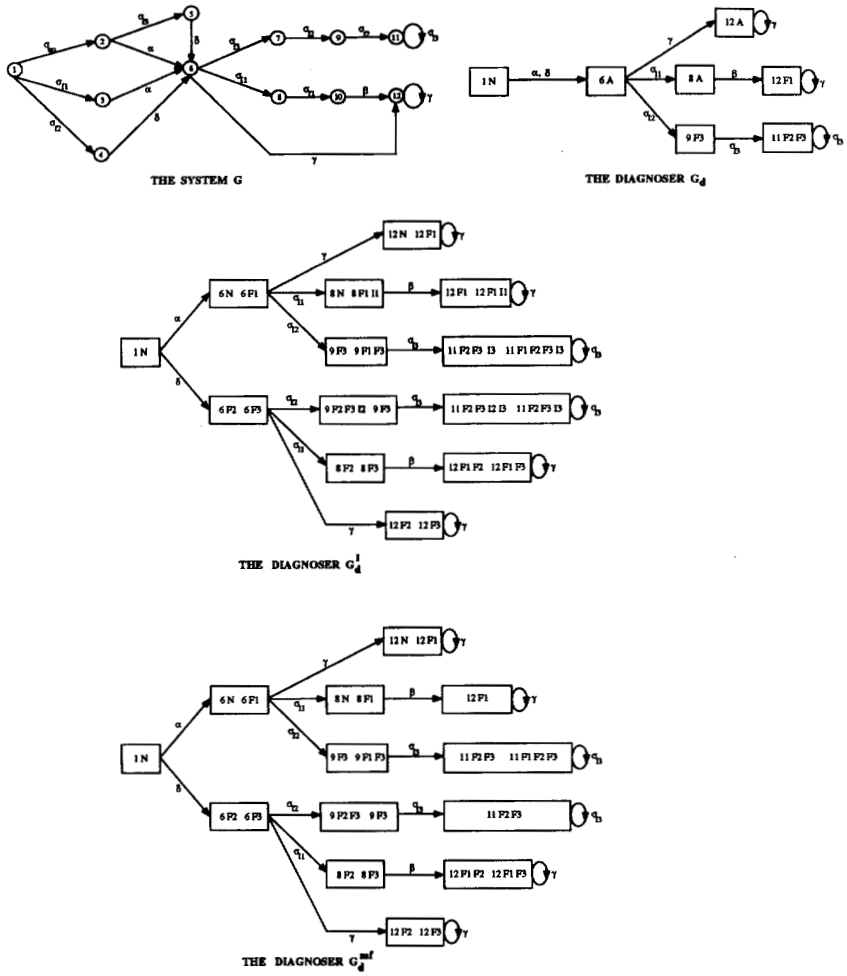


Fig. 10. Example illustrating implementation using diagnoser G_d.

$\Sigma_{f2} = \{\sigma_{f2}\}$, and $\Sigma_{f3} = \{\sigma_{f3}\}$. Inspection of G_d^I clearly shows that L is I-diagnosable. Knowing this fact, one is able to conclude that when G_d enters (and stays in) the state $\{(12, \{A\})\}$, no failures violating I-diagnosability have happened. Next, it is clear by inspecting the system G and the diagnoser G_d^I that when the trace $\alpha\sigma_{I2}\sigma_{I3}\sigma_{I3}^*$ is observed, the diagnoser enters into an F_1-indeterminate cycle and hence it is not possible to conclude whether a failure of type F_1 has happened or not. This, however, is not an (F_1, I_1)-indeterminate cycle since the corresponding trace in G, which contains the failure event σ_{f1}, does not contain the indicator event σ_{I1}. It is interesting to note that the corresponding state in the diagnoser G_d, $\{(11, \{F2, F3\})\}$, reveals nothing about the failure σ_{f1}.

Also shown in Fig. 10 is the diagnoser G_d^{mf}. Note that Condition C-MF) is violated in G_d^{mf}, and hence L is not diagnosable. Each of the states $\{(12, \{N\}), (12, \{F1\})\}$, $\{(12, \{F2\}), (12, \{F3\})\}$, $\{(12, \{F1, F2\}), (12, \{F1, F3\})\}$, and $\{(11, \{F2, F3\}), (11, \{F1, F2, F3\})\}$ forms an F_i-indeterminate cycle.

Finally, we make the observation that given an I-diagnosable language L, it is possible to have traces in L that satisfy the diagnosability condition D, but in which an indicator event of the appropriate type does not follow the failure event. Consider, for example, the trace $\sigma_{uo}\sigma_{f3}\delta\sigma_{f3}\sigma_{I2}\sigma_{f2}\sigma_{I3}$ in Fig. 10 and note that the corresponding state of G_d^I is F_2-certain.

VI. Conclusion

We have introduced the notions of diagnosability and I-diagnosability of systems in the framework of formal languages. We have compared this notion with the problems of observability, observability with delay, and invertibility, all of which fall in the general class of partial observation problems, and we have illustrated by means of examples that diagnosability is a distinctly different notion. We have provided a construction procedure for the diagnoser and presented necessary and sufficient conditions for diagnosability and I-diagnosability. These conditions can be verified using

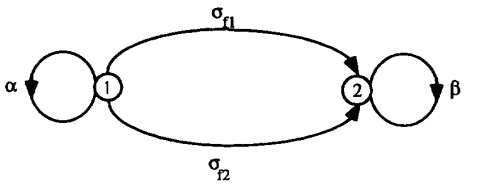


Fig. 11. Example of a nondiagnosable system that is observable with delay.

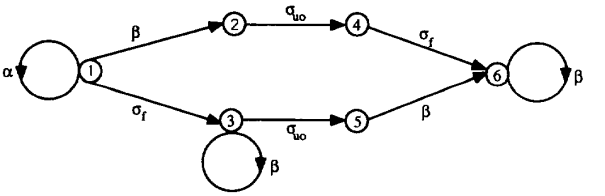


Fig. 12. Example of a diagnosable system that is not observable with delay.

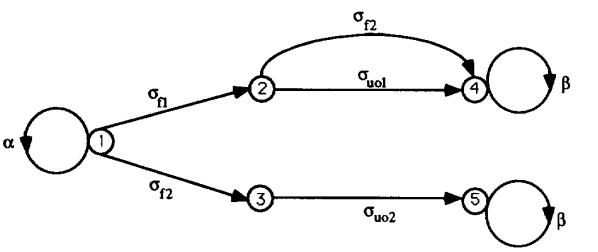


Fig. 13. Example of a noninvertible system that is diagnosable.

standard cycle detection algorithms on the diagnosers and the machine G'. We have shown that the "basic" diagnoser can be used to implement on-line diagnostics while suitably modified versions of this diagnoser can be used to check for diagnosability and I-diagnosability.

The theory presented in this paper is based on two assumptions on the system model. The first assumption, on the liveness of the system, can be relaxed and the definition of diagnosability can be extended to include terminating traces as well; the necessary and sufficient conditions for diagnosability can also be modified appropriately. While it is straightforward to do the above modifications, relaxing the liveness assumption tends to make the analysis cumbersome. The second assumption, on the absence of arbitrarily long traces of unobservable events in L, can also be relaxed if we require that the failures be detected within a bounded number of occurrences of observable events following the failure. Again, appropriate modifications of the theory presented in this paper are straightforward.

Finally, we point out that for the task of on-line diagnosis of diagnosable systems, it is not necessary to store the complete machine G_d whose state space may, in the worst-case, be exponential in the state space of G. It is sufficient to just remember its current state. Upon occurrence of an observable event, the new state of G_d could be built on-line from the current state of G_d and the relevant part of G, with polynomial complexity at each stage.

APPENDIX A

Diagnosability and Observability with Delay: Fig. 11 represents a system which is observable with delay but not diagnosable. Here α and β are observable events while σ_{f1} and σ_{f2} are unobservable failure events. This system is not diagnosable if the desired partition is $\Sigma_{f1} = \{\sigma_{f1}\}$ and $\Sigma_{f2} = \{\sigma_{f2}\}$.

Fig. 12 represents a system where the converse holds. In this figure α and β are observable while σ_{uo} is unobservable. The only failure event is σ_f. Here, a possible output sequence is β^*. When this sequence is observed, neither the current state nor the state any finite number of transitions in the past can be identified uniquely. On the other hand, it is possible to conclude the occurrence of a failure whenever the event sequence $\alpha^*\beta\beta\beta^*$ is observed. Hence, this is a diagnosable system which is not observable with delay.

Diagnosability and Invertibility: Fig. 13 depicts a noninvertible system which is diagnosable.

Here it is not possible to distinguish between the occurrence of traces $\sigma_{f1}\sigma_{uo1}\beta$, $\sigma_{f1}\sigma_{f2}\beta$, and $\sigma_{f2}\sigma_{uo2}\beta$. Hence the system is not invertible. If the required partition is $\Sigma_{f1} = \{\sigma_{f_1}, \sigma_{f_2}\}$, however, the system is diagnosable.

ACKNOWLEDGMENT

K. Sinnamohideen would like to acknowledge Johnson Controls Inc. for their support, especially Dr. S. Bomba, Vice-President for Technology, for his encouragement and the opportunities he provided. The authors also wish to acknowledge useful discussions with Prof. F. Lin.

REFERENCES

[1] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes—Theory and Application.* Englewood Cliffs, NJ: Prentice-Hall, 1993.
[2] S. Bavishi and E. Chong, "Automated fault diagnosis using a discrete event systems framework," in *Proc. 9th IEEE Int. Symp. Intelligent Contr.*, 1994, pp. 213–218.
[3] P. Caines, R. Greiner, and S. Wang, "Classical and logic based dynamic observers for finite automata," *IMA J. Math. Contr. Inform.*, vol. 8, pp. 45–80, 1991.
[4] R. Cieslak, D. Desclaux, A. Fawaz, and P. Varaiya, "Supervisory control of discrete-event processes with partial observations," *IEEE Trans. Automat. Contr.*, vol. 33 pp. 249–260, 1988.
[5] P. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge based redundancy—A survey and some new results," *Automatica*, vol. 26, pp. 459–474, 1990.
[6] W. Hamscher, L. Console, and J. Kleer, *Readings in Model Based Diagnosis.* San Mateo, CA: Morgan Kaufmann, 1992.
[7] D. Handelman and R. Stengel, "Combining expert systems and analytical redundancy concepts for fault tolerant flight control," *J. Guidance*, vol. 12, pp. 39–45, 1989.
[8] L. Holloway and S. Chand, "Time templates for discrete-event fault monitoring in manufacturing systems" in *Proc. 1994 Amer. Contr. Conf.*, 1994, pp. 701–706.
[9] J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages, and Computation.* Reading, MA: Addison Wesley, 1979.
[10] F. Lin, "Diagnosability of discrete-event systems and its applications," *J. DEDS*, vol. 4, no. 2, pp. 197–212, 1994.
[11] F. Lin, J. Markee, and B. Rado, "Design and test of mixed signal circuits: A discrete-event approach," in *Proc. 32nd Conf. Decis. Contr.*, 1993, pp. 246–251.
[12] F. Lin and W. M. Wonham, "On observability of discrete-event systems," *Inform. Sci.*, vol. 44, pp. 173–198, 1988.
[13] C. M. Özveren and A. S. Willsky, "Observability of discrete-event dynamic systems," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 797–806, 1990.

[14] ——, "Invertibility of discrete-event dynamic systems," *Math. Contr. Signals Syst.*, vol. 5, pp. 365–390, 1992.
[15] Y. Park and E. Chong, "On the eventual invertibility of discrete-event systems and its applications," in *Proc. 32nd Conf. Decis. Contr.*, 1993, pp. 680–685.
[16] P. J. Ramadge, "Observability of discrete-event systems," in *Proc. 25th Conf. Decis. Contr.*, 1986, pp. 1108–1112.
[17] P. J. Ramadge and W. M. Wonham, "The control of discrete-event systems," *Proc. IEEE*, vol. 77, no. 1, pp. 81–98, 1989.
[18] S. Rich and V. Venkatasubramanian, "Model-based reasoning in diagnostic expert systems for chemical process plants," *Comput. Chem. Eng.*, vol. 11, pp. 111–122, 1987.
[19] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis, "Diagnosability of discrete-event systems," in *Proc. 11th Int. Conf. Anal. Optim. Syst.*, Sophia-Antipolis, France (Lecture Notes in Control and Information Sciences), vol. 199. Berlin: Springer-Verlag, 1994, pp. 73–79.
[20] ——, "Failure diagnosis using discrete-event models," Dept. EECS, Univ. of Michigan, MI 48109, 1994, Tech. Rep. CGR 94-3. (Accepted for publication in *IEEE Trans. Contr. Syst. Tech.*.)
[21] K. Sinnamohideen, "Discrete-event based diagnostic supervisory control system," Presented at the in *Proc. AIChE Annual Meet.*, Los Angeles, CA, Nov. 17–22, 1991.
[22] N. Viswanadham and T. L. Johnson, "Fault detection and diagnosis of automated manufacturing systems," in *Proc. 27th Conf. Decis. Contr.*, Austin, TX, 1988, pp. 2301–2306.
[23] A. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, pp. 601–611, 1976.

Meera Sampath received the B.E. degree in electrical engineering from the College of Engineering, Guindy, Madras, India, and the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Kharagpur, India, in 1988 and 1990, respectively.

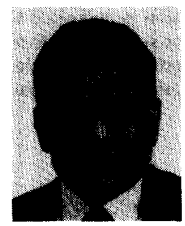
Since September 1991, she has been at the University of Ann Arbor, Michigan, where she is a doctoral candidate in electrical engineering. Her current research interests include failure diagnosis, discrete-event systems, and intelligent transportation systems.

Raja Sengupta received the Ph.D. degree in electrical engineering: systems from the University of Michigan, Ann Arbor, in February, 1995.

He is currently on the Research Staff at the California PATH program of the University of California, Berkeley. His areas of research include discrete-event systems, intelligent transportation systems, and discrete optimization.

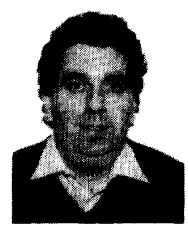
Stéphane Lafortune (S'78–M'86) received the B. Eng. degree from the École Polytechnique de Montreal in 1980, the M. Eng. degree from McGill University, Montreal, in 1982, and the Ph.D. degree from the University of California at Berkeley in 1986, all in electrical engineering.

Since 1986, he has been with the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, where he is an Associate Professor. His research interests are in discrete-event systems (modeling, on-line control, failure diagnosis, and applications) and in intelligent transportation systems.



Kasim Sinnamohideen (M'69) received the B.Sc. degree in physics form the University of Madras, India, in 1959 and the M.S. and Ph.D. degrees in mechanical engineering from Purdue University, West Lafayette, IN, in 1962 and 1968, respectively.

Since 1968 he has been with Johnson Controls Inc., Milwaukee, WI, where he has held various positions and participated in many areas of research. He has developed several diagnostics and configuration-type expert systems and knowledge-based real-time monitoring systems. He has worked in the areas of performance modeling of computer systems and polymer processing. He was a visiting scholar in the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, in 1992–1993.



Demosthenis Teneketzis (M'87) received the B.S. degree in electrical engineering from the University of Patras, Greece, in 1974 and the M.S., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, in 1976, 1977, and 1979, respectively.

From 1979 to 1980, he worked for Systems Control Inc., Palo Alto, CA, and from 1980 to 1984 he was with Alphatech Inc., Burlington, MA. Since September, 1984, he has been with the University of Michigan, Ann Arbor, where he is a Professor of Electrical Engineering and Computer Science. In winter and spring, 1992, he was a Visiting Professor at the Institute for Signal and Information Processing of the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. His current research interests include stochastic control, decentralized systems, queueing and communication networks, stochastic scheduling and resource allocation problems, and discrete-event systems.