FAILURE DIAGNOSIS OF STOCHASTIC AUTOMATA

DAVID THORSLEY AND DEMOSTHENIS TENEKETZIS

Abstract. Diagnosability of finite-state machines as introduced by Sampath et al. in [7] requires certainty that a failure has occurred in order to diagnose it. However, in many situations, it may be practical to diagnose failures based on them having a high likelihood of occurring, even if the failure has not occurred with certainty.

This paper extends the diagnosis methodology developed for logical finite state machines to stochastic automata. By incorporating probabilistic information into the model, we are able to define the weaker diagnosability conditions of A-diagnosability and AA-diagnosability. Off-line conditions for A-diagnosability and AA-diagnosability are determined through the construction of a stochastic diagnoser; we also show how the stochastic diagnoser can be used for online diagnosis of failure events.

1. Introduction

The purpose of this report is to study the diagnosability of stochastic discrete-event systems (DES) or stochastic automata. Discrete-event systems are systems whose evolution is guided by the occurrence of physical events that are separated by regular or irregular intervals of time. Stochastic automata are a more precise formulation of the general DES model, in which a probabilistic structure is appended to the model to estimate the likelihood of specific events occurring. An introduction to the theory of stochastic automata can be found in the textbook of Paz [6].

The failure diagnosis problem for DES is to detect the occurrence of specific pre-defined failure events that may not be directly observed by the sensors available to the system. Roughly speaking, a system is considered to be diagnosable if any instance of a failure event can eventually be detected from the observations made on the system.

The notion of diagnosability of DES is related to the notion of observability. Özeren and Willsky [5] formulate a definition of observability for nondeterministic systems. According to [5], a system is said to observable if, for any state in the system, any string of events that can occur from that state takes the system into a situation where there is only one possible state of the system consistent with the event observations. They introduce an observer DES which is used to produce state estimates of the system after every observable event.

Sampath et al. [7] consider the diagnosis problem using a similar methodology to Özeren and Willsky. In the diagnosis problem, the objective is to determine if specific events have occurred at some point in the evolution of the system, given a string of event observations. The conditions on diagnosability are less stringent than those for observability. It is not necessary to uniquely identify the state of the system; the state may be uncertain provided that if a failure has indeed occurred, all
possible behaviors consistent with the event observations indicate that this failure has occurred. The "diagnoser" machine developed in [7] can be interpreted as an extended version of Özveren and Willsky's observer.

The approach to the diagnosability problem for stochastic automata considered in this report is similar to the approach considered for logical automata in [7]. Comparisons with the results of that paper and the results presented here will be made throughout the text of this report.

Lunze and Schröder [4] approach the notions of observability and diagnosability in the context of stochastic automata. However, they model failures as permanent conditions of the system as opposed to events that may or may not occur. The concept of diagnosability considered by Lunze and Schröder considers a set of faults to be undiagnosable if faults are conditionally independent of the states and outputs of the system. Under this definition, a system is considered to be stochastically diagnosable if every output of the system provides information that can help diagnose the failure without delay.

The main differences between this report and the cited literature are the following. In this report, we use the stochastic properties of the model to determine the probable long-range behavior of the system. Logical discrete-event models can not distinguish between traces or states that are highly probable and those that are less probable. Therefore the notions that a state can be observed or a failure can be diagnosed after a finite delay are "all-or-nothing" propositions: one possible system behavior, however improbable, that does not allow the failure to be diagnosed is sufficient to consider a system to be undiagnosable. In this report, we present definitions of diagnosability that allow such improbable system behaviors to be disregarded.

While the probabilistic information is used in [4] to derive solutions to the state observation and diagnosis problems, it is not used to predict the future behavior of the system. Under their definition of diagnosability, a system is not stochastically diagnosable if there exists a possible system behavior that does not provide any information about which fault is the true fault. Because they require only the existence of such behavior, they do not check to see if this behavior is likely in the long-term evolution of the diagnoser. Lunze and Schröder also do not construct a diagnoser finite state machine to assist in the diagnosis problem, an approach used both in [7] and in this report.

The results presented in this report involve a class of stochastic discrete-event systems that are not diagnosable in the sense of [7], but display long-term properties that approximate that type of diagnosability. Roughly speaking, a failure event is considered to be diagnosable if the probability of correctly diagnosing any instance of that failure can be made arbitrarily close to 1 within a finite delay.

Because the model under consideration here allows us to formulate the probability distributions of various states and failures, we consider two different definitions of what it means for a failure to be "diagnosed." In the first situation, a failure is not said to be diagnosed until all possible system behaviors consistent with the observations of the system contain at least one instance the failure event. In the second situation, we merely require that of all the consistent system behaviors, the subset that contains the failure event has a probability above a pre-determined threshold.
This report is organized as follows. Section 2 introduces the stochastic automaton model under consideration and introduces the concepts and notation required to discuss diagnosability. Section 3 introduces new definitions of diagnosability motivated by the probabilistic nature of the automaton. Section 4 describes the construction of a stochastic diagnoser used to state conditions that ensure diagnosability. These conditions are presented in Section 5. Outstanding open problems and extensions of this report are discussed in Section 6.

2. The Model

The type of system to be diagnosed is a stochastic automaton. It is a quadruple

\[
G = (\Sigma, \mathcal{X}, p, x_0)
\]

where
- \( \Sigma \) is a finite set of events
- \( \mathcal{X} \) is a finite state space
- \( p(x', e | x) \) is a state transition probability defined for \( \forall x, x' \in \mathcal{X}, e \in \Sigma \)
- \( x_0 \in \mathcal{X} \) is the initial state

The total behavior of the system is summarized through the prefix-closed language \( \mathcal{L}(G) \), which for simplicity will be denoted as \( L \). \( L \) is a subset of \( \Sigma^* \), the Kleene-closure of the event set \( \Sigma \).

The system is observed through the events transitioning the system from one state to another. The event set \( \Sigma \) is partitioned as \( \Sigma = \Sigma_o \cup \Sigma_{uo} \), where \( \Sigma_o \) represents the set of observable events and \( \Sigma_{uo} \) represents the set of unobservable events. Observable events are events the occurrence of which is detected by any of the sensors available to the system; unobservable events are those events that cause state changes that the available sensors cannot detect.

When a string of events occurs in a system, the sequence of observable events is indicated by the projection of the string, which is defined as \( \text{Proj} : \Sigma^* \rightarrow \Sigma_o^* \)

\[
\text{Proj}(\epsilon) = \epsilon
\]

\[
\text{Proj}(\sigma) = \begin{cases} 
\sigma & \text{if } \sigma \in \Sigma_o \\
\epsilon & \text{if } \sigma \in \Sigma_{uo}
\end{cases}
\]

\[
\text{Proj}(s \sigma) = \text{Proj}(s) \text{Proj}(\sigma) \text{ for } s \in \Sigma^*, \sigma \in \Sigma
\]

where \( \epsilon \) denotes the empty trace. The inverse projection of a string of observable events \( s_o \) with respect to a language \( L \) is given by:

\[
\text{Proj}^{-1}_L(s_o) = \{ s \in L : \text{Proj}(s) = s_o \}
\]

We define a set of failure events \( \Sigma_f \subseteq \Sigma \). The objective of the diagnosis problem under consideration is to determine the likelihood of the occurrence of these failure events when only the events in \( \Sigma_o \) are observed. We can assume, without loss of generality, that \( \Sigma_f \subseteq \Sigma_{uo} \) because it is a trivial problem to determine when an observable failure has occurred.

To facilitate the diagnosis problem, the set of failure events is partitioned into a set of failure types:

\[
\Sigma_f = \Sigma_{f_1} \cup \cdots \cup \Sigma_{f_m}
\]

If a failure event \( \sigma_f \in \Sigma_f \) occurs, we will say that “a failure of type \( F_i \) has occurred.”
The probability transition function $p(x', e \mid x)$ defines the probability that a certain event $e$ will occur and transition the state of the machine from a given state $x$ to the specified state $x'$. For example, $p(z, a \mid y) = 0.7$ states that, if the system is in state $y$, with probability .7 the event $a$ will occur and transition the state of the system to $z$. We will assume, for the sake of simplicity, that $p(x', e \mid x) > 0$ for at most one $x' \in \mathcal{X}$.

Under this assumption, the probability transition function can be used to define the partial transition function, which is defined as $\delta : \mathcal{X} \times \Sigma \rightarrow \mathcal{X}$ where

\begin{equation}
\delta(x, e) = y \Rightarrow p(y, e \mid x) > 0
\end{equation}

If for some $x \in \mathcal{X}$ and $e \in \Sigma$, there does not exist $y \in \mathcal{X}$ such that $p(y, e \mid x) > 0$, then $\delta(x, e)$ is undefined.

This relationship demonstrates that the stochastic model is a more specific model than the logical finite state machine model discussed in [7]. In the logical automaton model, the partial transition function is defined as part of the specification of the system, but here it is derived from the state transition probabilities.

The partial transition function can be extended to strings of events as follows:

\begin{align*}
\delta(x, e) &= x \\
\delta(x, se) &= \delta(\delta(x, s), e)
\end{align*}

Figure 1 provides a pictorial representation of a stochastic automaton. The set of states is $\mathcal{X} = \{0, 1, 2\}$, and the initial state $x_0 = 0$ is denoted by an unconnected transition. The set of events is $\Sigma = \{a, b, c, \sigma_f\}$, which is partitioned into $\Sigma_o = \{a, b, c\}$ and $\Sigma_w = \sigma_f = \{\sigma_f\}$. A transition arc is drawn between two states if the probability of that transition occurring is greater than zero.

In order to facilitate the solution to the diagnosis problem, we make two assumptions about the stochastic automaton $G$:

(A1) The language $L$ generated by $G$ is live. That is to say, for every state in $\mathcal{X}$, the probability of a transition occurring from that state is one, or, equivalently, for $\forall x \in \mathcal{X}$

\begin{equation}
\sum_{x' \in \mathcal{X}} \sum_{e \in \Sigma} p(x', e \mid x) = 1
\end{equation}
(A2) The generator $G$ does not have any cycle of unobservable events, or equivalently
\[
\exists n_0 \in \mathbb{N} \text{ such that } \forall s \in L, s \in \Sigma^* \Rightarrow ||s|| \leq n_0
\]
Together these two assumptions force observable events to occur with regularity. Assumption (A1) requires that transitions will continue to occur regardless of the state of the system, and (A2) requires that after at most a finite delay, one of these transitions will be an observable event.

The liveness assumption (A1) also forces all states in $G$ to satisfy the Markov property, allowing the use of techniques of Markov chain analysis in subsequent sections of this report.

2.1. Discrete Event Notation. The symbol $\bar{s}$ will be used to denote the prefix-closure of a string $s \in \Sigma^*$. The postlanguage $L/s$ is the set of possible continuations of a string $s$, i.e.
\[
L/s = \{ t \in \Sigma^* | st \in L \}
\]
When defining diagnosability, it will be important to consider traces that end in a failure event of a specific type. Define
\[
\Psi_f = \{ s \sigma_f \in L : \sigma_f \in \Sigma_f \}
\]
\[
X_\sigma = \{ x_0 \} \cup \{ x \in X : p(x, e | y) > 0 \text{ for some } e \in \Sigma \land y \in X \}
\]
Let $L(G, x)$ represent the set of all traces that originate from the state $x$ in the state space of $G$. Let the final event in any trace be denoted by $s_f$. Define
\[
L_o(G, x) = \{ s \in L(G, x) : s = u \sigma, u \in \Sigma^*, \sigma \in \Sigma_o \}
\]
\[
L_o(G, x) = \{ s \in L_o(G, x) : s_f = \sigma \}
\]
Convention: In the examples of stochastic automata in this paper, observable events will be marked using lowercase Roman letters, while unobservable events that are not failure events will be denoted by lowercase Greek letters. Failures will be denoted as $\sigma_f$.

2.2. Probabilistic Notation. From our assumption that $p(x', e | x) > 0$ for only one $x'$ for each pair $(x, e)$, we can write $p(x', e | x) = p(e | x)$. Therefore, the probability of an event $e \in \Sigma$ being the next event given the system is in state $x$ is given by:
\[
\text{Prob}(e | x) = p(e | x)
\]
If we wish to find the probability of a particular string being the true future system behavior given the system is state $x$, we can calculate this recursively as:
\[
\text{Prob}(es | x) = p(e | x)p(s | \delta(x, e))
\]
Because our tests for diagnosability can only be based on observable events, it will be important to determine the probability that $e_o \in \Sigma_o$ will be the next observable event given the system is in state $x$. This can be calculated as:
\[
\text{Prob}(e_o | x) = \sum_{s \in L_o(x)} \text{Prob}(s | x)
\]
If our state observation is incomplete, we will need to determine the probability being in a state $x$, given that we have observed the string $s_o$. This probability is:

$$Prob(x \mid s_o) = \frac{\sum_{s \in Proj^{-1}_L(s_o)} Prob(s \mid x_0)}{\sum_{s \in Proj^{-1}_L(s_o)} Prob(s \mid x_0)}$$

By combining equations (15) and (16), we can determine the probability of the next observable event being $e_o$, given that the string of observed events to date is $s_o$.

$$Prob(e_o \mid s_o) = \sum_{x \in \delta(x_0, s) \text{ for some } s \in Proj^{-1}_L(s_o)} Prob(e_o \mid x)Prob(x \mid s_o)$$

Lastly, if we are in a state $x$, the probability that after the next event $e_0$, the system has been transitioned to the state $x'$ is given by:

$$Prob(x', e_0 \mid x) = \sum_{s \in L_{e_0}(x) \delta(x, s) = x'} Prob(s \mid x)$$

3. APPROACHES TO DEFINING DIAGNOSABILITY

The objective of the diagnosis problem is to detect the occurrence of an unobservable failure in the system, based on the information available from the record of observed events.

**Definition 1.** (Diagnosability) (as defined in [7]) A live, prefix-closed language $L$ is $F_i$-diagnosable with respect to a projection $Proj$ if

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

where the diagnosability condition function $D : \Sigma^* \rightarrow \{0, 1\}$ is given by

$$D(st) = \begin{cases} 1 & \text{if } \omega \in Proj^{-1}_L[Proj(st)] \Rightarrow \Sigma_f \in \omega \\ 0 & \text{otherwise} \end{cases}$$
Figure 2 shows an example of a system that is $F$-diagnosable. If the event $\sigma_f$ occurs, the next observable event will be either $a$ or $b$. If $b$ is observed, then we know that the only possible system behavior consists with our observation of $b$ is $\sigma_f b$, and the failure will be diagnosed. On the other hand, if $a$ is observed, it will necessarily be followed by $c$. The only behavior consistent with the string of observations $ac$ is $\sigma_f ac$, so once again the failure will be diagnosed. Regardless of whether the first observed event is $a$ or $b$, in this example we will not have to wait for more than two events after the failure to determine that the failure has indeed occurred.

This definition of diagnosability was developed for logical automata models and the necessary and sufficient conditions for this type of diagnosability are stated in [7]. This definition therefore makes no use of the probabilistic information that the stochastic model under consideration in this report contains. We now present weaker definitions of diagnosability that take into account the stochastic structure of the model.

Consider the system in Figure 3, and suppose that the behavior of the system is the trace $s = \sigma_f$. Clearly, $s \in \Psi(\Sigma_f)$. The postlanguage of $s$ is given by $L/s = a^*b^*$, meaning that it consists of an arbitrary number of $a$'s followed by an arbitrary number of $b$'s.

Let $n \in \mathbb{N}$. Let $t \in L/s$ such that $||t|| = n$. Then $t$ is of the form $a^{n-k}b^k$, $0 \leq k \leq n$.

Suppose $k > 0$. Then $Proj_L^{-1}[Proj(st)] = \sigma_f a^{n-k}b^k$. The failure is therefore diagnosed, as every string in the projection contains the failure event $F$.

Now suppose $k = 0$, that is to say, $t = a^n$. Then $Proj_L^{-1}[Proj(st)] = \{\sigma_f a^n, \mu a^n\}$. The diagnosability condition in Definition 1 is not satisfied, as $\sigma_f \not\in \mu a^n$. Since the string $a^n$ is part of the postlanguage $L/s$ for an arbitrarily large $n$, there is potentially an infinite delay before the failure can be diagnosed. Therefore the system is not diagnosable.

The reason this system is not diagnosable is because of the existence of the trace $a^n \in L/s$, a trace along which the failure cannot be diagnosed. However, because we have appended probabilities to the system model, we can now consider
the probability of the trace $a^n$ being the actual behavior of the system. At each moment, the probability of $a$ occurring is 0.9, so $\text{Prob}(a^n) = (0.9)^n$. As $n$ becomes larger and larger, the probability of the trace that prevents diagnosis from occurring becomes smaller and smaller.

Although in this system we can never guarantee that the failure will be diagnosed after a finite delay, the probability of a string of events that allows diagnosis becomes greater as we observe more events after the failure event.

This observation is the motivation for this weaker definition of diagnosability, which is created by making condition (19) less stringent.

**Definition 2.** (A-diagnosability) A live, prefix-closed language $L$ is asymptotically $F_i$-diagnosable (or $F_i$-A-diagnosable) with respect to a projection $\text{Proj}$ and a set of transition probabilities $p$ if

$$\text{(21)} \quad (\forall \varepsilon > 0)(\exists n_i \in \mathbb{N})(\forall s \in \Psi(\Sigma_{f_i}))
\{\text{Prob}(t : D(st) = 0 \land t \in L/s \land ||t|| \geq n_i) < \varepsilon\}$$

where the diagnosability condition function $D$ is as in (20):

$$\text{(22)} \quad D(st) = \begin{cases} 1 & \text{if } \omega \in \text{Proj}_L^{-1}[\text{Proj}(st)] \Rightarrow \Sigma_{f_i} \in \omega \\ 0 & \text{otherwise} \end{cases}$$

The system in Figure 3 is $F_i$-A-diagnosable. The only string in the postlanguage $L/s$ that does not allow diagnosis of the failure $F_i$ is $a^n$, a string whose probability of occurring approaches zero as $n$ becomes arbitrarily large. This indicates that, after a failure occurs, we can let the probability of diagnosing the failure after a finite delay become arbitrarily high by selecting a value $n$ such that $\text{Prob}(a^n)$ is sufficiently small.

However, this is not the only way in which the diagnosability conditions in Definition 1 can be weakened using a probabilistic model. Consider the system shown in Figure 4. Let $s = a\sigma_f$. The postlanguage $L/s$ is given by $(b\alpha + b\alpha \sigma_f)^\ast$. Let $t \in L/s$ and let $m \in \mathbb{N}$ be an integer such that $t$ contains $2m$ observable events. Then $\text{Proj}(st) = a(ba)^m$, and $\text{Proj}_L^{-1}[\text{Proj}(st)] = (ab + a\sigma_f b)^m a$.

The string $(ab)^m a$ is part of the set $\text{Proj}_L^{-1}[\text{Proj}(st)]$, regardless of the length of the continuation $t$. Similarly, if $t$ has $2m - 1$ observable events, then $(ab)^m \in \text{Proj}_L^{-1}[\text{Proj}(st)]$. 

---

**Figure 4.** A system that is AA-diagnosable but not A-diagnosable.
Because $\sigma_f \not\in (ab)^m$ and $\sigma_f \not\in (ab)^n$, the diagnosability condition function $D(st)$ is equal to zero for all continuations $t$ of the string $s$. Therefore this system is neither diagnosable nor A-diagnosable, as we can never say for any continuation $t$ that all possible true system behaviors consistent with the observed behavior contain the failure event $\sigma_f$.

The problematic trace is the trace $(ab)^m$. However, just as when the condition for A-diagnosability was developed, we can consider the probability of the trace $(ab)^m$ which does not contain the failure event $\sigma_f$. The probability of this trace is $(1)(0.9)(1)(0.9) \cdots (1)(0.9) = (0.9)^m$.

The probability of the trace that does not contain a failure approaches zero as the number of events observed becomes large. Therefore, although we cannot assure that a correct diagnosis is made, we can force the probability of the failure event being included in the actual behavior to be arbitrarily close to one by waiting for a sufficiently long, yet finite, amount of observations.

This observation allows us introduce a third notion of diagnosability that is weaker than both the original definition and A-diagnosability:

**Definition 3.** (AA-Diagnosability) A live, prefix-closed language $L$ is asymptotically approximately $F_i$-diagnosable (or $F_i$-AA-diagnosable) with respect to a projection $Proj$ and a transition probability function $p$ if

\[(\forall \varepsilon > 0 \land \forall \alpha < 1) (\exists n_i \in \mathbb{N}) [\forall s \in \Psi(\Sigma_{f_i})] \quad \{ \text{Prob}(t : D_\alpha(st) = 0 \land t \in L/s \land ||t|| \geq n_i) < \varepsilon \}\]

where the diagnosability condition function $D_\alpha$ is:

\[(24) \quad D_\alpha(st) = \begin{cases} 1 & \text{if } \text{Prob}(\omega : \omega \in \text{Proj}_{L}^{-1}[\text{Proj}(st)] \land \Sigma_{f_i} \in \omega) > \alpha \\ 0 & \text{otherwise} \end{cases} \]

The system in Figure 4 is $F_i$-AA-diagnosable, because as $||t||$ becomes large, the probability that the system behavior that does not contain a failure approaches zero.

AA-diagnosability takes advantage of the probabilistic structure of the system model to provide a weaker diagnosability condition. In strict diagnosability (Definition 1), we require that all possible continuations after a failure lead, after a finite delay, into situations where all possible true system behaviors contain the failure event. To interpret AA-diagnosability, we replace the word “all” with the phrase “with arbitrarily high probability.” Hence, in AA-diagnosability, we require that, with arbitrarily high probability, possible continuations of the system after a failure lead, after a finite delay, into situations where we can say, with arbitrarily high probability, that the possible true system behavior contains the failure event.

Because they refer to all $\epsilon > 0$ or all $\alpha < 1$, the definitions of A-diagnosability and AA-diagnosability are based on the limiting behavior of the system as the length of the continuation following a failure grows without bound. Therefore, in order to determine if a stochastic automaton has these properties, we will need to either let the system run for an infinitely long time or derive off-line conditions that are necessary and sufficient to confirm A-diagnosability and AA-diagnosability.

However, if we wish to determine if the diagnosability condition $D_\alpha(st) = 1$ is satisfied for a specific $\alpha$, we can determine this by observing the behavior of the system online. Given an observed string $s_\alpha$, we can calculate the probability that a failure of type $F_i$ has occurred using a machine called the stochastic diagnoser.
If the probability of failure is greater than \( \alpha \), we can declare that a failure has occurred. The stochastic diagnoser can also be used as a tool to help accomplish the primary goal of this report, which is to determine necessary and sufficient conditions to ensure A-diagnosability and AA-diagnosability.

4. The Stochastic Diagnoser

The stochastic diagnoser is a machine constructed from the stochastic automaton \( G \). The diagnoser serves two purposes: one is to perform online diagnostics of the system \( G \), while the other is to allow us to formulate necessary and/or sufficient conditions for A-diagnosability and AA-diagnosability.

4.1. Construction of the Stochastic Diagnoser. In order to construct the stochastic diagnoser, we need to first define a set of failure labels \( \Delta_f = \{F_1, F_2, \cdots, F_m\} \) where \( m \) is the number of different failure types in the system. The set of possible failure labels is defined as:

\[
\Delta = \{N\} \cup 2^{\{\Delta_f\}}
\]

The \( \{N\} \) label should be interpreted as representing the “normal” behavior of the system, while a label of the form \( \{F_i, F_j\} \) should be interpreted to mean that “at least one failure of type \( i \) and at least one failure of type \( j \) have occurred.”

The set of observable states of the system is defined as:

\[
x \in X_o \Rightarrow (x = x_0) \vee (\exists s \in L : \delta(x_0, s) = x \wedge s_f \in \Sigma_o)
\]

The set of possible diagnoser states is defined as follows:

\[
Q_o = 2^{X_o \times \Delta}
\]

Each diagnoser state consists of a subset of the observable states of the system with failure labels attached.

The stochastic diagnoser for a finite-state machine \( G \) is the machine:

\[
G_d = (Q_d, \Sigma_o, \delta_d, q_0, \Phi, \phi_0)
\]

where

- \( Q_d \) is the set of discrete states
- \( \Sigma_o \) is the set of observable events
- \( \delta_d \) is the transition function of the diagnoser (to be defined below)
- \( q_0 \) is the initial discrete state, defined as \( \{(x_0, \{N\})\} \)
- \( \Phi \) is the set of probability transition matrices (to be defined below)
- \( \phi_0 \) is the initial probability mass function on the initial state \( q_0 \)

The stochastic diagnoser can be thought of consisting of two interconnected models. The first model is a finite-state discrete event model, which estimates the current state and current failure information of the system. The second is an infinite-state probabilistic model, which gives the probability for each component of the state estimate based on the observed events.

The set of discrete states, \( Q_d \), is the subset of \( Q_o \) that is reachable from \( q_0 \) under \( \delta_d \). A state \( q_d \in Q_d \) is a set of the form

\[
q_d = \{(x_1, l_1), \ldots, (x_n, l_n)\}
\]

where \( x_i \in X_o \) and \( l_i \in \Delta \). The elements of each state of the diagnoser are called components. The set of components of the diagnoser is defined as the set of all
triples \((q, x, l)\) such that \(q \in Q_d\), \(x \in X_o\), \(l \in \Delta\), and \((x, l) \in q\). The number of components in a diagnoser state \(q_d\) will be denoted by \(\|q_d\|\).

In order to construct the probability transition matrices \(\Phi\), we will need to impose an order on the set of components in each state \(q \in Q_d\). This order can be chosen arbitrarily. By convention, the \(i\)th component of a state \(q\) will be denoted by \(c_{q,i}\).

Each state of the diagnoser consists of the set of components that are possible true states consistent with the observed system behavior. If the component \((x, l)\) is an element of the state \(q\), it means that for every sequence of observed events that transitions the diagnoser to the state \(q\), there exists at least one string \(s\) in the inverse projection of that sequence such that \(s\) transitions the stochastic model to the state \(x\) and failures of all types included in the label \(l\) are included in \(s\). The properties of components of the diagnoser will be essential to providing conditions for diagnosability in Section 5.

In order to define \(\delta_d\), the transition function of the diagnoser, we must first define how the labels change from one state to another. Define the label propagation function \(LP : X_o \times \Delta \times \Sigma_o^{*} \rightarrow \Delta\) as

\[
LP(x, l, s) = \begin{cases} 
\{N\} & \text{if } l = \{N\} \land \forall i[\Sigma_f_i \not\subset s] \\
\{F_i : F_i \in l \lor \Sigma_f_i \subset s\} & \text{otherwise}
\end{cases}
\]

Using the label propagation function, we can define the transition function of the diagnoser as:

\[
\delta_d(q, \sigma) = \bigcup_{(x,l) \in \mathcal{G}} \bigcup_{s \in L_x(q, x)} \{(\delta(x, s), LP(x, l, s))\}
\]

The function \(LP\) shows that a label \(F_i\) is added whenever the true behavior of the system contains an event \(\sigma_f \in \Sigma_f_i\). Once this label is appended, it cannot be removed regardless of whether or not an event in \(\Sigma_f_i\) occurs or not in the system behavior following the label.

From the transition functions of the original stochastic automaton and the stochastic diagnoser, we can define the component transition function \(\delta_{comp}\), which determines the component of the diagnoser that is the true state, given that the true behavior of the system is \(s \in \Sigma^{*}\) and \(\sigma_f \in \Sigma_o\). Given \(q \in Q_d\), \((x, l) \in q\), and \(\sigma\),

\[
\delta_{comp}(q, x, l, s) = (\delta_d(q, Proj(s), \delta(x, s), LP(x, l, s))
\]

The quadruple \((Q_d, \Sigma_o, \delta_d, q_0)\) that has been defined above is equivalent to the diagnoser presented in [7], with the modification that the “ambiguous” label has been removed from the set of possible labels. This quadruple is used to provide estimates of the state and information on the possible failure events. This is the “discrete-event” part of the stochastic diagnoser, which is used to determine the state of the diagnoser. In order to derive the probabilities of each component in each state, we now append a probabilistic structure to make the diagnoser “stochastic.”

In order to calculate these, we need to define a set of probability transition matrices as \(\Phi : Q_d \times \Sigma_o \rightarrow \mathcal{M}[0,1]\)

\[
\Phi_{ij}(q, \sigma_o) = \sum_{s \in L_x(q, x)} Prob(s) \\
=[Proj(c_{q_{ij}}(q, \sigma_o, s), \Sigma_o | q_{ij})
\]

where the range \(\mathcal{M}[0,1]\) represents the set of finite-dimensional matrices whose values are contained in the interval \([0,1]\). The size of the matrix outputted by \(\Phi(q, \sigma)\) is
\[ \|q\| \times \|\delta_d(q, \sigma_o)\| \]. So, for example, if an event takes the diagnoser from a state with \( m \) components to a state with \( n \) components, the size of the matrix associated with that event will be \( m \times n \). The initial probability vector of the system corresponds to the probability mass function of the initial state. Since the only component of the initial state is \( (x_0, \{N\}) \), we define \( \phi_0 = [1] \).

The “total state” of the diagnoser is a pair \((q, \phi)\), a combination of the discrete state of the diagnoser and the probability vector obtained from the trace required to reach that state.

Given the total state of a diagnoser \((q, \phi)\), where \( q \) is an order set \( \{c_1, c_2, \ldots, c_n\} \) and \( \phi \) is a vector \( [\phi_1, \phi_2, \ldots, \phi_n] \), we conclude that \( \text{Prob}(c_i) = \phi_i, i = 1, \ldots, n \). The following theorem makes clear the procedure of calculating the probability vector \( \phi \) from the set of matrices \( \Phi \):

**Theorem 1.** The state probability vector \( \phi(s_o e_o) \) can be calculated recursively as follows:

\[
\phi(\epsilon) = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{34}
\]

\[
\phi(s_o e_o) = \frac{\phi(s_o) \Phi(q_i e_o)}{\text{Prob}(e_o \mid s_o)} \tag{35}
\]

**Proof.** If the observed trace is the empty trace, then, by its construction, the stochastic diagnoser is in the initial state and \( \phi(\epsilon) = \phi_0 = [1] \). Now suppose there exists an observable event \( e_o \) that transitions the system from a state \( q_i \) where \( \|q_i\| = n \) to a state \( q_j \) where \( \|q_j\| = m \). The recursive equation for the probability vector is
given by the following derivation:

\[
\begin{align*}
\phi(s_o e_o) & = [\text{Prob}(c_{j,1} | s_o e_o) \cdots \text{Prob}(c_{j,m} | s_o e_o)] \\
& = \frac{1}{\text{Prob}(e_o | s_o)} [\text{Prob}(c_{j,1} | s_o e_o) \text{Prob}(e_o | s_o) \cdots \\
& \quad \text{Prob}(c_{j,m} | s_o e_o) \text{Prob}(e_o | s_o)] \\
& = \frac{1}{\text{Prob}(e_o | s_o)} [\text{Prob}(c_{j,1}, e_o | s_o) \cdots \text{Prob}(c_{j,m}, e_o | s_o)] \\
& = \frac{1}{\text{Prob}(e_o | s_o)} \left[ \sum_{k=1}^{n} \text{Prob}(c_{j,1}, e_o | c_{i,k}, s_o) \text{Prob}(c_{i,k} | s_o) \cdots \\
& \quad \cdots \sum_{k=1}^{n} \text{Prob}(c_{j,m}, e_o | c_{i,k}, s_o) \text{Prob}(c_{i,k} | s_o) \right] \\
& = \frac{1}{\text{Prob}(e_o | s_o)} [\text{Prob}(c_{i,1} | s_o) \cdots \text{Prob}(c_{i,n} | s_o)] \cdots \\
& \quad \left[ \text{Prob}(c_{j,1}, e_o | c_{i,1}, s_o) \cdots \text{Prob}(c_{j,m}, e_o | c_{i,1}, s_o) \right] \\
& \quad \cdots \left[ \text{Prob}(c_{j,1}, e_o | c_{i,n}, s_o) \cdots \text{Prob}(c_{j,m}, e_o | c_{i,n}, s_o) \right] \\
& \quad \phi(s_o e_o) = \frac{\phi(s_o) \Phi(q_i, e_o)}{\text{Prob}(e_o | s_o)}
\end{align*}
\]

The dependence on the observed string \( s_o \) in Equation (41) does not affect the values of the terms of the matrix, because given the current component of the diagnoser state, the transition probability to the next component is independent of the path used to reach that state.

This result gives us a method to perform online diagnosis by calculating the probability vector from the matrices in the stochastic diagnoser. Suppose the observed behavior of the system is \( s_o = e_1 e_2 \ldots e_n \), and the sequence of states observed is \( (q_1, q_2, \ldots, q_n) \). Then the unnormalized probability vector is given by \( \phi_{un}(s_o) = \phi \Phi(q_1, e_1) \Phi(q_2, e_2) \cdots \Phi(q_n, e_n) \). To find the normalized probability, we need to divide this vector by \( \text{Prob}(e_o | s_o) \), which is simply the sum of the terms of \( \phi_{un}(s_o) \).

To perform online diagnosis of failures using the stochastic diagnoser, select a threshold \( \beta \) such that \( 0 < \beta < 1 \). Suppose we observe online the string of events \( s_o \), and let \( s \in \text{Proj}^{-1}_L(s_o) \). We say that a failure has been diagnosed online if \( D_{\beta}(s) = 1 \), or equivalently if \( \text{Prob}(F | s_o) > \beta \).

**Example 1.** Consider the diagnoser in Figure 5, and suppose the observed behavior of the system is \( s_o = aabcaa \). Then the probability vector is given by:

\[
\phi_{un}(s_o) = [1] \begin{bmatrix} .7 & .05 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} .7 & .05 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} .2 & .05 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} .7 & .05 \\ 0 & .7 \end{bmatrix} \begin{bmatrix} .7 & .05 \\ 0 & .5 \end{bmatrix} = \begin{bmatrix} 0.04802 & 0.026705 & 0.00915 \\ 0.5725 & 0.3184 & 0.1091 \end{bmatrix}
\]

\[
\phi(s_o) = [0.5725 \ 0.3184 \ 0.1091]
\]
The discrete state of the diagnoser after observing $s_o$ is $\{(0, N), (0, F), (1, F)\}$. Therefore the total state of the stochastic diagnoser is

$$ (q, \phi) = \{(0, N), (0, F), (1, F)\}, \begin{bmatrix} 0.5725 & 0.3184 & 0.1091 \end{bmatrix} $$

From the total state, we can conclude that:

$$ \Pr((0, N) \mid s_o) = 0.5725 $$
$$ \Pr((0, F) \mid s_o) = 0.3184 $$
$$ \Pr((1, F) \mid s_o) = 0.1091 $$

From this we can conclude that

$$ \Pr(F \mid s_o) = \Pr((0, F) \mid s_o) + \Pr((1, F) \mid s_o) = 0.4275 $$

If we had set a failure threshold of, say, $\beta = 0.4$, we would now declare that a failure has occurred and take appropriate steps to repair the system. If we had set a higher threshold ($\beta > 0.43$), we would continue to observe the system until such time as a string $s_o$ occurs where $\Pr(F \mid s_o) > \beta$.

4.2. The Stochastic Diagnoser as Markov Chain. Despite the fact that the stochastic diagnoser has matrices associated with each transition instead of simple probabilities, at a fundamental level the diagnoser can be described as a Markov Chain. This will allow us to use techniques of Markov chain analysis to derive conditions for diagnosability in Section 5.

From any stochastic diagnoser, we can construct a Markov Chain where the components of each state of the diagnoser can be thought of as the states of the Markov Chain. To do this, first define $\Omega : Q_d \times Q_d \to \mathcal{M}_{[0,1]}$ as

$$ \Omega(q_i, q_j) = \sum_{e_o \in \Sigma_o, \delta(q_i, e_o) = q_j} \Phi(q_i, e_o) $$

$\Omega(q_i, q_j)$ is simply the sum of all the matrices associated with the transitions from $q_i$ to $q_j$. If there are no transitions from $q_i$ to $q_j$, $\Omega(q_i, q_j)$ is a matrix of zeros of size $\|q_i\| \times \|q_j\|$.

We can construct a Markov transition matrix from $\Omega(q_i, q_j)$ as follows.

**Theorem 2.** Let $G_d$ be a stochastic diagnoser. Then the matrix $\Pi(G_d)$ is a Markov transition matrix, where $\Pi(G_d)$ is defined as

$$ \Pi(G_d) = \begin{bmatrix} \Omega(q_1, q_1) & \cdots & \Omega(q_1, q_n) \\ \vdots & \ddots & \vdots \\ \Omega(q_n, q_1) & \cdots & \Omega(q_n, q_n) \end{bmatrix} $$

**Proof.** Let $q_a \in Q_d$ and let $c_{q_a,i} \in q_a$ be the $i$th component of $q_a$. By construction

$$ \Phi_{ij}(q_a, e_o) = \Pr(b(c_{q_a,j}, e_o \mid c_{q_a,i}) $$

$$ \Omega_{ij}(q_a, q_b) = \sum_{e_o \in \Sigma_o} \Pr(b(c_{q_a,j}, e_o \mid c_{q_a,i})) $$

The sum of the $i$th row of $\Omega(q_a, q_b)$ is therefore given by:

$$ \sum_{j=1}^{\|q_b\|} \Omega_{ij}(q_a, q_b) = \sum_{j=1}^{\|q_b\|} \sum_{e_o \in \Sigma_o} \Pr(b(c_{q_b,j}, e_o \mid c_{q_a,i})) $$

$$ \sum_{j=1}^{\|q_b\|} \Omega_{ij}(q_a, q_b) = \sum_{j=1}^{\|q_b\|} \sum_{e_o \in \Sigma_o} \Pr(b(c_{q_b,j}, e_o \mid c_{q_a,i})) $$
When the matrix $\Pi(G_d)$ is constructed, the $k$th row of $\Pi(G_d)$ is constructed from the $i$th rows of $\Omega_{q_i, q}$, where $q$ is an arbitrary stochastic diagnoser state. So the sum of the $k$th row of $\Pi(G_d)$ is given by:

\begin{align}
\sum_{q \in Q_d} \sum_{j=1}^n \Omega_{ij}(q_o, q) &= \sum_{q \in Q_d} \sum_{e_o \in \Sigma_o} Prob(c_{q,j}, e_o | c_{q_o,i}) \\
&= \sum_{e_o \in \Sigma_o} Prob(e_o | c_{q_o,i}) \\
&= 1
\end{align}

Therefore the sum of each row of $\Pi(G_d)$ is 1. The number of rows in $\Pi(G_d)$ is equal to the sum of the number of rows in $\Omega(q_i, q)$ for any $q \in Q_d$, i.e., $\sum_{i=1}^n ||q_i||$. Similarly, the number of columns in $\Pi(G_d)$ is equal to the sum of the number of columns in $\Omega(q, q_i)$ for any $q \in Q_d$, i.e. $\sum_{i=1}^n ||q_i||$. Because it has an equal number of rows and columns, $\Pi(G_d)$ is a square matrix. Therefore, $\Pi(G_d)$ is a Markov transition matrix.

Example 2. To illustrate the results of Theorem 2, consider again the stochastic diagnoser shown in Figure 5. Denote this diagnoser by $G_d$. $G_d$ can be described by a Markov chain with matrix of transition probabilities $\Pi(G_d)$ that is given below.

\[
\begin{bmatrix}
(q_1, 0, N) & 0.7 & 0.05 & 0.2 & 0.05 & 0 & 0 & 0 & 0 \\
(q_2, 0, N) & 0.7 & 0.05 & 0.2 & 0.05 & 0 & 0 & 0 & 0 \\
(q_2, 1, F) & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\
(q_3, 2, N) & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
(q_3, 2, F) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
(q_4, 0, N) & 0 & 0 & 0 & 0.2 & 0.05 & 0.7 & 0 & 0.05 & 0 & 0 \\
(q_4, 0, F) & 0 & 0 & 0 & 0 & 0.25 & 0 & 0.7 & 0.05 & 0 & 0 \\
(q_4, 1, F) & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 \\
(q_5, 0, N) & 0 & 0 & 0 & 0 & 0 & 0.7 & 0.05 & 0 & 0 & 0 \\
(q_5, 0, F) & 0 & 0 & 0 & 0 & 0 & 0.7 & 0.05 & 0 & 0 & 0 \\
\end{bmatrix}
\]

The components to the left of the matrix indicate which component is associated with each row of the matrix. The horizontal and vertical lines in the matrix delineate the boundaries between the different matrices $\Omega(q_i, q_j)$. By inspecting $\Pi(G_d)$, the matrices associated with each transition of $G_d$ are clearly identifiable.

4.3. Relevant Results from Markov Chain Theory. Because the components of the diagnoser can be thought of as states of a Markov Chain, we will be able to use the theory of Markov Chains to derive conditions for A-diagnosisability and AA-diagnosisability. The following subsection is a review of the results in Markov Chain theory that will be essential for this report; readers seeking a more thorough review of the subject should consult [1] or [3].

Suppose that $x$ and $y$ are two states of a Markov Chain. The notation $\rho_{xy}$ indicates the probability that if the Markov Chain is in state $x$, it will, at some point in the future, visit state $y$.

If, for a state $x$, $\rho_{xx} = 1$, that state is called recurrent. Otherwise, if $\rho_{xx} < 1$, then $x$ is a transient state. If a state is transient, then at some point in the evolution of the Markov chain the system will leave that state and never return; on the other hand, if a state is recurrent, if the Markov Chain visits the state once, the chain
will return to that state infinitely often. If the Markov Chain is finite-state, there
must be at least one recurrent state in the chain.

If $\rho_{xy} > 0$, it is said that $y$ is reachable from $x$; this is denoted by $x \rightarrow y$. If
$x$ is a recurrent state and $x \rightarrow y$, then $y$ is also a recurrent state and $y \rightarrow x$. If
there is a set of recurrent states $\{x_1, x_2, \ldots, x_n\}$ such that $x_i \rightarrow x_j$ and $x_j \rightarrow x_i$ for
$\forall i, j \in \{1, \ldots, n\}$, then that set is called a recurrence class. Paz [6] indicates that
determining whether or not a state is recurrent is a decidable problem, while Xie
[9] provides an efficient algorithm for finding recurrent classes.

Suppose a chain starts in state $x$. Then the number of times that the evolution
of the chain takes it to state $y$ will be denoted by $N_x(y)$. The probability that after
$n$ transitions, the state of the Markov chain will have transitioned from $x$ to $y$ is
denoted by $P^n(x, y)$. We can rewrite this probability in discrete event notation as:

$$ P^n(x, y) = \text{Prob}(t : t \in L(G, x) \land ||t|| = n \land \delta(x, t) = y) $$

As the number of transitions in a Markov Chain grows large, the probability of
being in a transient state approaches zero. As this idea is central to the develop-
ment of conditions for A-diagnosability and AA-diagnosability, it will be expressed
formally in the following lemma.

**Lemma 1.** Let $\mathcal{X}$ be the finite state space of a Markov Chain, and let $\mathcal{T} \subset \mathcal{X}$ be
the set of transient states of the chain.

Let $x \in \mathcal{X}$ be an arbitrary state of the Markov Chain, and let $t$ be an arbitrary
sequence of state transitions beginning at $x$. Then $\forall \epsilon > 0$, $\exists n \in \mathbb{N}$ such that

$$ \text{Prob}(t : t \in L(G, x) \land ||t|| \geq n \land \delta(x, t) \in \mathcal{T}) < \epsilon $$

**Proof.** Let $y \in \mathcal{T}$. Then

$$ \text{Prob}(N_x(y) = m) = \rho_{xy} \rho_{yy}^{m-1} (1 - \rho_{yy}) $$

$$ E(N_x(y)) = \sum_{m=1}^{\infty} m \text{Prob}(N_x(y) = m) $$

$$ = \sum_{m=1}^{\infty} m \rho_{xy} \rho_{yy}^{m-1} (1 - \rho_{yy}) $$

$$ = \frac{\rho_{xy}}{1 - \rho_{yy}} < \infty $$

as the fact that $y$ is a transient state implies that $1 - \rho_{yy} > 0$.

The expected number of visits to state $y$ is the sum of the expected number of
visits to state $y$ after any specific number of transitions are made, so

$$ \sum_{n=0}^{\infty} P^n(x, y) = E(N_x(y)) < \infty $$

In order for the sum of an infinite series to be finite, the terms of that infinite
series must approach zero. Therefore $\forall \epsilon > 0$, $\exists n \in \mathbb{N}$ such that

$$ ||t|| \geq n \Rightarrow P||t|| (x, y) < \frac{\epsilon}{l} $$

This relationship can be rewritten in discrete event notation as:

$$ \text{Prob}(t : t \in L(G, x) \land ||t|| \geq n \land \delta(x, t) = y) < \frac{\epsilon}{l} $$
where $l$ is the number of transient states of the Markov chain. Since $y$ is an arbitrary element of $\mathcal{T}$, we can conclude

$$\text{(60)}$$

$$\text{Prob}(t : t \in L(G, x) \land \|t\| \geq n \land \delta(x, t) = y) = \sum_{y \in \mathcal{T}} \text{Prob}(t : t \in L(G, x) \land \delta(x, t) = y) < \sum_{y \in \mathcal{T}} \frac{1}{l} \epsilon < \epsilon$$

$$\text{(61)}$$

\hfill \Box

5. Conditions for Diagnosability

In this section we present necessary and sufficient conditions for a stochastic automaton to be $A$-diagnosable, and sufficient conditions for a stochastic automaton to be $AA$-diagnosable. This conditions will be expressed in terms of the stochastic diagnoser introduced in the previous section and will take advantage of the Markov properties of the diagnoser shown by Theorem 2. Before determining conditions for $A$-diagnosability and $AA$-diagnosability, we present some additional properties of the stochastic diagnoser.

5.1. Properties of the Stochastic Diagnoser. These properties of the stochastic diagnoser can be deduced from the properties of the label propagation function and the Markovian structure of the relationship between the components.

Property 1. All components that are reachable from a component with the label $F_i$ also bear the label $F_i$.

Proof. If $F_i \in l_1$, then $F_i \in LP(x, l_1, s)$ for $\forall x \in X, s \in L_0(G, x)$. Essentially, once a failure label is appended, it cannot be removed. \hfill \Box

Property 2. If a diagnoser state is $F_i$-certain, then all traces reaching the state contain some event $\sigma$ such that $\sigma \in \Sigma_{F_i}$.

A state $q$ of the diagnoser is said to be $F_i$-certain if for all $(x, l) \in q$, either $F_i \in l$ or $F_i \not\in l$. From this definition, this property is shown in Lemma 1-i of [7].

Property 3. All components in the same recurrence class have the same failure label.

Proof. Suppose $c_1$ and $c_2$ are components in the same recurrence class. Then $c_2$ is reachable from $c_1$ and vice versa. From Property 1, if a label were appended to $c_2$ by a trace reaching from $c_1$ to $c_2$, it cannot be removed again by any trace reaching from $c_2$ to $c_1$. Therefore, $c_1$ and $c_2$ must carry the same label. \hfill \Box

Property 4. All components reachable from a recurrent component bearing the label $F_i$ in an $F_i$-uncertain state are elements of $F_i$-uncertain states.

Proof. Let $c_r$ denote a recurrent component bearing the label $F_i$ in an $F_i$-uncertain state.

Consider a state $q_f \in Q_d$ where all components have labels that include $F_i$. By Property 1, all components reachable from any component in this state also bear the label $F_i$. The diagnoser transition function $\delta_d$ shows that any state reachable from $q_f$ contains only those components that are reachable from the components
of \( q_f \). From Property 1, the only components reachable from the components of \( q_f \) carry the label \( F_i \); therefore, all states reachable from \( q_f \) must be certain that \( F_i \) has occurred.

Since \( c_r \) is in an \( F_i \)-uncertain state, it cannot be reached from the state \( q_f \). Therefore, no component of \( q_f \) can be in the same recurrence class as \( c_r \), which implies that no component of \( q_f \) is reachable from \( c_r \).

Furthermore, from Property 1, no component that is reachable from \( c_r \) can not carry a label \( F_i \), so no state that is certain that \( F_i \) did not occur can be reachable from \( c_r \). Therefore the only states that can be reached from \( c_r \) are \( F_i \)-uncertain states. 

\[ \square \]

5.2. \textbf{Necessary and Sufficient Conditions for A-diagnosability.} Using the above properties of the stochastic diagnoser, we can state conditions for a language \( L \) to be \( A \)-diagnosable or \( PA \)-diagnosable in terms of the properties of the diagnoser.

\textbf{Theorem 3.} A language \( L \) generated by a stochastic automaton \( G \) is asymptotically \( F_i \)-diagnosable if, and only if, every state in its diagnoser \( G_d \) containing a recurrent component with the label \( F_i \) is \( F_i \)-certain.

\textbf{Proof.} \textit{Necessity:} Necessity will be shown by contradiction. Suppose there exists a state \( q \in Q_d \) that such that \( q \) is not \( F_i \)-certain and \( q \) contains a recurrent component \( q_f = (q, x, l_f) \) such that \( F_i \in l_f \). We will then show that

\begin{equation}
(\exists \epsilon > 0)(\exists s \in \Psi(\Sigma_{f_i}))(\forall n \in \mathbb{N})
\{ \text{Prob}(t : D(st) = 0 \land t \in L/s \land \|t\| \geq n) > \epsilon \}
\end{equation}

By construction, every component of every state of the diagnoser is accessible from the initial state. Therefore, there exists a trace \( st \), where \( s \in \Psi(\Sigma_{f_i}) \) and \( t \in L/s \) such that \( \delta_{\text{comp}}(c_{or}, st) = c_f \) and \( \text{Prob}(t) > 0 \). Because \( c_f \) is in an \( F_i \)-uncertain state, \( D(st) = 0 \).

Let \( n \in \mathbb{N} \). Let \( u \in L/st \) such that \( \|tu\| \geq n \). By Property 4, for \( \forall u \in L/st \), the string \( u \) transits the diagnoser to an \( F_i \)-uncertain state. Therefore

\begin{equation}
\text{Prob}(u : D(stu) = 0 \land \|tu\| \geq n \land tu \in L/s) = 1
\end{equation}

Choose \( \epsilon > 0 \) such that

\begin{equation}
0 < \epsilon < \text{Prob}(t)
\end{equation}

Then, because of Equations 63 and 64, and the fact that \( \text{Prob}(tu) = \text{Prob}(t) \text{Prob}(u) \), it follows that for \( \forall n \in \mathbb{N} \)

\begin{equation}
\text{Prob}(tu : D(stu) = 0 \land \|tu\| \geq n \land tu \in L/s) = \text{Prob}(t) \text{Prob}(u : D(stu) = 0 \land \|tu\| \geq n \land tu \in L/s) > \epsilon
\end{equation}

Therefore, if there is a recurrent component carrying the label \( F_i \) in an \( F_i \)-uncertain state, the stochastic automaton is not \( A \)-diagnosable.

\textit{Sufficiency:} Let \( C \) be the set of components of a stochastic diagnoser, and let \( \mathcal{T}_c \in C \) be the set of transient components. Suppose that every state \( q \in Q_d \) that contains a recurrent component \( (q, x, l_f) \) such that \( \Sigma_{f_i} \in l_f \) is \( F_i \)-certain.

Let \( s \in \Psi(\Sigma_{f_i}) \). By Lemma 1, there exists \( n \in \mathbb{N} \) such that \( \forall c = (q, x, l) \in \mathcal{C} \)

\begin{equation}
\text{Prob}(t : t \in L(G, x) \land \|t\| \geq n \land \delta_{\text{comp}}(c,t) \in \mathcal{T}_c) < \epsilon
\end{equation}
Since $\delta(x_0, s)$ is a component of the diagnoser of the system reached by $s$, this implies that:

\[(67) \quad \text{Prob}(t : t \in L/s \land \|t\| \geq n \land \delta_{\text{comp}}(c, t) \in T_{\text{c}}) < \epsilon\]

Therefore, if at least $n$ events have occurred since the failure event, with probability greater than $1 - \epsilon$, we will reach a state that contains at least one recurrent component.

However, because $s \in \Psi(\Sigma_{F_i})$, $F_i \in LP(x_{\omega}, N, s)$. From Property 1, any label reachable after the string $s$ must contain $F_i$.

If the true behavior of the system reaches a recurrent component with label $F_i$, then, by assumption, the state is $F_i$-certain, Therefore $D(st) = 1$.

Since the probability of reaching an $F_i$-certain state is at least $1 - \epsilon$,

\[(68) \quad \text{Prob}(t : t \in L/s \land \|t\| \geq n \land D(st) = 1) > 1 - \epsilon\]

\[(69) \quad \text{Prob}(t : t \in L/s \land \|t\| \geq n \land D(st) = 0) < \epsilon\]

Therefore if every state containing a recurrent component bearing the label $F_i$ is $F_i$-certain, the system is $A$-diagnosable. $\square$

**Example 3.** Figure 6 shows the stochastic diagnoser of the stochastic automaton in Figure 3, which was shown in Section 3 to be $A$-diagnosable, but not diagnosable according to Definition 1.

From the conditions for diagnosability according to Definition 1 given in [7], this stochastic automaton is not diagnosable in that sense because there is a trace $a^n$ that takes the diagnoser into a cycle of $F$-uncertain states, and this cycle of $F$-uncertain states corresponds to two separate cycles in the original system model.

However, the Markov matrix associated with this stochastic diagnoser is:

\[
\begin{bmatrix}
(q_1, 0, N) & 0 & .5 & .45 & .05 \\
(q_2, 1, N) & 0 & 1 & 0 & 0 \\
(q_2, 2, F) & 0 & 0 & .9 & .1 \\
(q_3, 3, F) & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

From this matrix we can determine that the recurrent components are $(q_2, 1, N)$ and $(q_3, 3, F)$. To test for $A$-diagnosability, we need only consider the recurrent component carrying the label $F$. This component is the only component in the state $q_3$; therefore it is part of an $F$-certain state. Therefore the stochastic automaton in Figure 3 is $A$-diagnosable.

This condition indicates that in order to test for $A$-diagnosability, we only need to be able to determine the recurrent components of the stochastic diagnoser. This is equivalent to determining the recurrent states of a Markov Chain. Therefore in order to test for $A$-diagnosability, we need only know which events in our model of the system have a non-zero probability of occurring and we do not need to know the specific values of $p(x', e \mid x)$. This allows us to confirm if our system is $A$-diagnosable even if we have not modeled the transition probabilities exactly.

In the diagnosability conditions discussed in [7], the diagnoser itself was insufficient to confirm whether or not a system was diagnosable; in general, it was also necessary to consult a generator based on the logical automaton. However, the stochastic diagnoser contains sufficient information to test the necessary and sufficient conditions for $A$-diagnosability, without reference to a generator or to the
original stochastic automaton. The reason for this is that the set of matrices $\Phi$ contains information that is lost in the construction of the logical diagnoser; namely, whether certain components in one diagnoser state are reachable for certain components in other diagnoser states. This information allows the calculation of recurrent components to be made by consulting only the stochastic diagnoser.

5.3. A sufficient condition for AA-diagnosability.

**Theorem 4.** A language $L$ generated by a stochastic automaton $G$ is asymptotically approximately $F_1$-diagnosable if for every state in the diagnoser $G_d$ constructed from $G$, the set of recurrent components is $F_1$-certain.

**Proof.** Suppose that in each state of the diagnoser, the set of recurrent components is $F_1$-certain. Let $C$ be the set consisting of every component in every state of the stochastic diagnoser, and let $T_c \subseteq C$ and $R_c \subseteq C$ be the sets of transient and recurrent components of the diagnoser, respectively.

Let $t_t$ denote the set of traces such that for some $(q, x, l) \in C$

\[
t \in t_t \Rightarrow (t \in L(G, x)) \land (||t|| \geq n) \land (\delta_{\text{comp}}(q, x, l, t) \in T_c)
\]

Also, let $t_r$ denote the set of traces such that

\[
t \in t_r \Rightarrow (t \in L(G, x)) \land (||t|| \geq n) \land (\delta_{\text{comp}}(q, x, l, t) \in R_c)
\]

From Theorem 2, the components of the stochastic diagnoser can be treated as states of a Markov Chain. Therefore we can apply Lemma 1 to say that $\forall \epsilon > 0, \alpha < 1, \exists n \in \mathbb{N}$ such that $||t|| \geq n$ implies that for $\forall x \in X$

\[
\text{Prob}(t : t \in t_t \mid x) < \epsilon(1 - \alpha)
\]

\[
\text{Prob}(st : t \in t_t) < \epsilon(1 - \alpha)
\]

for any $s \in L$. Suppose we observe the a trace of events $s_0 t_0 \in \Sigma^*_o$. We can then condition the probability in Equation 73 on $s_0 t_0$, yielding:

\[
\sum_{s_0 t_0} \text{Prob}(st : t \in t_t \mid s_0 t_0) \text{Prob}(s_0 t_0) < \epsilon(1 - \alpha)
\]

Because every term in this summation is positive, we can consider only the subset of possible traces where $\text{Prob}(st : t \in t_t \mid s_0 t_0) \geq 1 - \alpha$, allowing the
following derivation:

\[
(75) \sum_{s \subseteq t, \mathbb{P}(s | t) \geq 1 - \alpha} \mathbb{P}(s_v t_v) \mathbb{P}(s_v t_o) < \epsilon (1 - \alpha)
\]

\[
(76) \sum_{s \subseteq t, \mathbb{P}(s | t) \geq 1 - \alpha} \frac{\mathbb{P}(s_v t_v)}{1 - \alpha} \mathbb{P}(s_v t_o) < \epsilon
\]

\[
(77) \sum_{s \subseteq t, \mathbb{P}(s | t) \geq 1 - \alpha} \mathbb{P}(s_v t_o) < \epsilon
\]

\[
(78) \sum_{s \subseteq t, \mathbb{P}(s | t) < \alpha} \mathbb{P}(s_v t_o) < \epsilon
\]

\[
(79) \sum_{s \subseteq t, \mathbb{P}(s | t) \geq 1 - \alpha} \mathbb{P}(s_v t_o) \geq 1 - \epsilon
\]

Therefore if \( s \in L \) and \( t \in L/s \) such that \( |t| \geq n \), with probability greater than \( 1 - \epsilon \), the total state of the stochastic diagnoser indicates the probability of being in a recurrent component is no less than \( \alpha \).

We have assumed that the set of recurrent components of each state of the stochastic diagnoser is \( F_i \)-certain. Therefore, with probability \( 1 - \epsilon \), the string \( st \) takes the stochastic diagnoser to a state where either \( \mathbb{P}(F_i) \geq \alpha \) or \( \mathbb{P}(F_i) < 1 - \alpha \).

Now suppose \( s \in \Psi(\Sigma_f) \). Then because \( F_i \in LP(x_0, N, s) \), any component reachable by \( st \) bears the label \( F_i \). Therefore the string \( st \), with probability greater than \( 1 - \epsilon \), will transition the stochastic diagnoser to a recurrent component with the label \( F_i \), which implies that \( \mathbb{P}(F_i) \geq \alpha \), or equivalently, \( D(\alpha)(st) = 1 \).

Therefore if \( s \in \Psi(\Sigma_f) \), \( t \in L/s \) and \( |t| \geq n \), we can rewrite Equation 79 as

\[
(80) \sum_{s \subseteq t, D(\alpha)(st) = 1} \mathbb{P}(s_v t_o) \geq 1 - \epsilon
\]

\[
(81) \sum_{s \subseteq t, D(\alpha)(st) = 0} \mathbb{P}(s_v t_o) < \epsilon
\]

Therefore, for \( \forall s \in \Psi(\Sigma_f) \), if the true continuation is at least \( n \) events long, the probability of the set of strings of observable events that take the stochastic diagnoser to a total state where diagnosis can not be made has a probability of less than \( \epsilon \). Therefore,

\[
(82) \mathbb{P}(t : t \in L/s \land |t| \geq n \land D(\alpha)(st) = 0) < \epsilon
\]

that is to say, the system is AA-diagnosable. We have started from the assumption that the set of recurrent components of each state of the stochastic diagnoser if \( F_i \)-certain and shown that, if this is the case, the system is AA-diagnosable. \( \square \)

**Example 4.** The stochastic diagnoser of the system in Figure 4 is shown in Figure 7. It was previously shown the this stochastic automaton is AA-diagnosable but
not A-diagnosable. The Markov matrix associated with its stochastic diagnoser is:

\[
\begin{pmatrix}
(q_1,0,N) & 0 & 1 & 0 & 0 & 0 & 0 \\
(q_2,1,N) & 0 & 0 & 9 & 1 & 0 & 0 \\
(q_3,0,N) & 0 & 0 & 0 & 0 & 1 & 0 \\
(q_3,0,F) & 0 & 0 & 0 & 0 & 1 & 0 \\
(q_4,1,N) & 0 & 0 & 9 & 1 & 0 & 0 \\
(q_4,1,F) & 0 & 0 & 0 & 1 & 0 & 0 \\
\end{pmatrix}
\]

The recurrent components of the stochastic diagnoser are \((q_3,0,F)\) and \((q_4,1,F)\). Because these components appear in states that are not \(F\)-certain, the system is not A-diagnosable. However, because they are the only recurrent components in those states, the set of recurrent components in every state is \(F\)-certain. Therefore, the system is AA-diagnosable. In this system, we cannot ever be certain that a failure event has occurred, but the probability that a failure has occurred will approach 1 as the system evolves.

Although this condition developed in Theorem 4 is sufficient for AA-diagnosability, it is not necessary. The next example will show a system that is AA-diagnosable but does not meet the condition of Theorem 4.

**Example 5.** Consider the system in Figure 8. After the failure occurs, we can use the observations after the failure to determine if the system is in state 1 or state 2.

The AA-diagnosability of the system in Figure 8 can be determined using hypothesis testing techniques [8]. Because states 1 and 2 are separate recurrence classes, we can treat each state as a hypothesis for the true state of the system.

Let \(H_i\) denote the hypothesis that the system is in state \(i, i = 1, 2\). If we observe \(n\) events, a certain fraction will be \(a\) and the rest will be \(b\). Let \(\hat{a}_n\) denote the fraction of the \(n\) observed events that are \(a\).

To determine which hypothesis is correct, we consider the likelihood function \(L(H_1 \mid \hat{a}_n)\):

\[
L(H_1 \mid \hat{a}_n) = \frac{\text{Prob}(H_1 \mid \hat{a}_n)}{\text{Prob}(H_2 \mid \hat{a}_n)}
\]

\[
= \frac{\text{Prob}(\hat{a}_n \mid H_1)\text{Prob}(H_1)}{\text{Prob}(\hat{a}_n \mid H_2)\text{Prob}(H_2)}
\]

\[
= \frac{\binom{n}{a_n} \binom{.7}{.3} n^{-a_n} (3)^{n-a_n}}{\binom{n}{a_n} \binom{.5}{.5} n^{-a_n} (5)^{n-a_n}}
\]

\[
= (1.4)^{a_n} n(0.6)^{n-a_n}
\]
Taking the logarithm of the likelihood function gives:

(84) \[
\log L(H_1 \mid \hat{a}_n) = \hat{a}_n n \log(1.4) + (n - \hat{a}_n n) \log(0.6) \\
= n(\log(0.6) + \hat{a}_n(1.4 - \log(0.6)))
\]

As \( n \) grows large, the log-likelihood of \( H_1 \) grows large as well, provided the term \( \log(0.6) + \hat{a}_n(1.4 - \log(0.6)) \) is greater than zero, which is the case when \( \hat{a}_n > .61 \).

If the failure has occurred and the state of the system is state 2, we can determine from the law of large numbers that

\[
(\forall \varepsilon > 0)(\exists n_1 \in \mathbb{N}) \text{ such that } n \geq n_1 \Rightarrow \text{Prob}(|\hat{a}_n - 0.7| > 0.9) < \varepsilon
\]

Therefore \( \text{Prob}(\hat{a}_n < .61) < \varepsilon \). Let 0 < \( \alpha < 1 \). Now choose \( n_2 \in \mathbb{N} \) such that

\[
(n \geq n_2) \land (\hat{a}_n > .61) \Rightarrow \frac{\log L(H_1 \mid \hat{a}_n) \log \frac{\alpha}{1 - \alpha}}
\]

\[
\Rightarrow \text{Prob}(H_1 \mid \hat{a}_n) > \alpha
\]

Let \( n = \max(n_1, n_2) \). If a string \( st \) occurs such that \(|st| = n \) and \( \hat{a}_n(st) > .61 \), then \( D_{\alpha}(st) = 1 \). Also, the probability that \( \hat{a}_n > .61 \) is, by the law of large numbers, greater than \( 1 - \varepsilon \). Therefore, this system is AA-diagnosable, despite the fact that its stochastic diagnoser contains a state whose recurrent labels are \( F_1 \)-uncertain.

6. Future Work

This report has extended the methods for diagnosability of logical discrete-event systems developed by Sampath et al. [7] to stochastic automata. The properties of A-diagnosability and AA-diagnosability introduced in this report are less stringent types of diagnosability that take into account the probabilistic information in the stochastic automaton model.
The diagnoser finite-state machine introduced in [7] has been extended to a stochastic diagnoser by introducing the set of transition matrices $\Phi$, which allows for online updates of the probability mass function of the states of the system. Using the stochastic diagnoser, we have been able to derive conditions for $A$-diagnosability and $AA$-diagnosability.

However, there are several outstanding issues related to the stochastic diagnosis problem that are worthy of further study, including the following issues not discussed in this report.

6.1. **Necessary and Sufficient Conditions for AA-diagnosability.** Roughly speaking, in order for a system not to be $F_i$-$AA$-diagnosable, there must be a recurrence class bearing the label $F_i$ and recurrence class not bearing the label $F_i$ that are in some way *indistinguishable* from each other. Two recurrence classes can be distinguished from each other if the appear in different states of the stochastic diagnoser, a fact that is captured by the condition of Theorem 4.

However, as Example 5 shows, two recurrence classes can be distinguished from each other even if they consist of components of the same states of the stochastic diagnoser. Provided that for at least one state in the recurrence class, there is at least one event where the probability of that event occurring depends on which recurrence class the system is actually in, we will be able to distinguish between the two recurrence classes.

A necessary and sufficient condition for AA-diagnosability could likely be developed through a formalized notion of this idea of indistinguishability.

6.2. **Unobservable Reach.** The *unobservable reach* $U_x$ of a state $x$ is defined as the set of all states such that

$$y \in U_x \Rightarrow \exists s \in L(G, x) : s \in \Sigma^*_\omega \land \delta(x, s) = y$$

The diagnoser presented in the textbook by Cassandras and Lafortune [2] includes components in the unobservable reach of each component included in the diagnoser.

In the stochastic diagnoser, when we say that $Prob((x, l) \ | \ s_\omega) = \beta$, the proper interpretation of this statement is that “the probability the system in state $x$ at the instant the final event in $s_\omega$ is observed is $\beta$” or “the probability the system is in state $x$ or its unobservable reach after observing $s_\omega$ is $\beta$.”

Including components in the unobservable reach is less practical for the stochastic diagnoser because, without the existence of some notion of time, the probability distribution of a state $x$ and the states in the unobservable reach of $x$ cannot be determined. In order to provide a meaningful probabilistic interpretation to states in the unobservable reach, we will need to introduce either a discrete or continuous time clock to the model.

6.3. **Timeliness and Convergence Issues.** The definitions of $A$-diagnosability and $AA$-diagnosability imply that as a limiting behavior, systems that satisfy these properties approach the traditional sense of diagnosability. However, it may take a very long time for the probability of diagnosis to come close to 1; no condition has been placed on the rate at which the probability of diagnosis approaches 1. In practical situations, the time required before diagnosis becomes possible may be too long.
References


