Verification of Detectability in Probabilistic Finite Automata

Christoforos Keroglou\textsuperscript{a}, Christoforos N. Hadjicostis\textsuperscript{b}

\textsuperscript{a}Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor

\textsuperscript{b}Department of Electrical and Computer Engineering, University of Cyprus

Abstract

In this paper we analyze state estimation in stochastic discrete event systems (SDES) that can be modeled as probabilistic finite automata (PFAs). For a given PFA, we obtain the necessary and sufficient conditions that guarantee exact state estimation, at least asymptotically, with increasing certainty as more information is acquired from observing the behavior of the given PFA, by defining the notion of AA-detectability, and providing necessary and sufficient conditions that can be used to verify it. The characterization and analysis of AA-detectability is transformed to a problem of classification between two (or more) PFAs, which capture the recurrent behaviour of an underlying Markov process that is obtained by ignoring output behavior and focusing on state transitions in the given PFA. Our approach combines techniques used in classification between two (or more) PFAs with state estimation methods used in logical discrete event systems (DES). We prove that the proposed verification of AA-detectability is of polynomial complexity with respect to the size of the state space of the given PFA.

Keywords: Discrete event systems, Finite automata, Markov models, State estimation

1 Introduction and Motivation

The state estimation problem is key in many control engineering applications involving complex systems. Early instances of state estimation problems in nondeterministic finite automata appear in [13] and [14], both of which formulate the observability problem by requiring (eventually) perfect knowledge of the current state in a partially observable system. State estimation techniques are needed in many applications of discrete event systems (DES). For example, opacity [15,16,8] requires that a given set of states remain opaque by an intruder who has access to the generated sequence of observations, regardless of the underlying activity in the system; opacity has direct implications on issues of privacy and security in DES. Another related application is fault diagnosis [17] which requires discrimination, within a finite time interval following the occurrence of a fault, between the set of states that are possible under normal behaviour and the set of states that are possible under faulty behaviour, for every possible trace that can be executed in the system; disambiguation between these two sets of states hinges on state estimation techniques. Issues of diagnosability and opacity have also been studied in stochastic settings [23,5]. A relevant problem in probabilistic finite automata (PFAs) is the classification between two given models, typically hidden Markov models (HMMs) or PFAs [1,9,11]. Classification is closely related to diagnosability if we can treat these models as having separable sets of states.

State estimation aims at accurate characterization of the possible current states following a possibly long observation sequence generated by the underlying DES. In deterministic settings, a key concept is the notion of detectability [20]. In particular, the notion of strong detectability holds if all observation sequences lead to an accurate estimate of the current state after a finite number of observations. However, in stochastic DES (SDES), we can relax the strong notion of accurate state estimation described above by requiring that accurate state estimation is only achieved asymptotically. The availability of stochastic information allows us to compute the probability of any trace of events and determine not only if a state is a possible candidate as a current state, but also how probable it is. In other words, the available probabilistic information can be used by the estimator to determine the posterior likelihood of a certain state, conditioned on a particular sequence of observations and, via the verification process, to determine the likelihood of problematic observation sequences.

The authors of [21] introduced notions of detectability
in PFAs (stochastic detectability). The approach in [21] takes the viewpoint that the problematic behaviour generates sequences of observations that do not allow us to estimate the exact state with increasing certainty; it analyzes all possible observation sequences (infinite sequences) and declares the system not stochastically detectable, if there is at least one such problematic sequence (however improbable, as long as it is feasible). Our previous work in [10] introduced the notion of A-detectability which requires that, with probability that increases as the length of observation sequences increases, long enough observation sequences lead to perfect state estimation. In [10] we provided necessary and sufficient conditions for A-detectability, along with a proof that A-detectability is a PSPACE-hard problem.

Our main contributions in this paper are twofold. First, we introduce the notion of AA-detectability, which requires increasing certainty to a single state estimate for a subset of such sequences (namely, sequences that can be generated repeatedly by the system with non vanishing probability). Second, we propose a method that can be used to verify AA-detectability with polynomial complexity. The notion of AA-detectability that we introduce and analyze in this paper allows problematic sequences to lead to conditional probabilities that are nonzero for more than one state, but requires a single state estimate for the problematic behaviour. Thus, our method can be used to verify AA-detectability with polynomial complexity.

The remaining sections of this paper are organized as follows. Section 2 introduces the system model and reviews some useful results from Markov theory. Section 3 reviews previous notions of stochastic detectability and introduces AA-detectability. We illustrate important differences between AA-detectability and previous notions, and also present an example. Section 4 provides the verification algorithm for AA-detectability, which is summarized in Algorithm 1. Finally, Section 5 concludes the paper.

2 Notation and Background

Let $\Sigma$ be an alphabet (set of events) and denote by $\Sigma^*$ the set of all finite-length strings of elements of $\Sigma$ (sequences of events), including the empty string $\varepsilon$. The length of string $s$ is denoted by $|s|$ with $|\varepsilon| = 0$. A language $L \subseteq \Sigma^*$ is a subset of finite-length strings in $\Sigma^*$ [4] (i.e., sequences of events with the convention that the first event appears on the left). Given strings $s, t \in \Sigma^*$, the string $st$ denotes the concatenation of $s$ and $t$, i.e., the sequence of events captured by $s$ followed by the sequence of events captured by $t$. For a string $s$, $\pi$ denotes the prefix-closure of $s$, and is defined as $\bar{s} = \{t \in \Sigma^* | \exists t' \in \Sigma^* \{t' = s\}\}$.

In general, only a subset $\Sigma_{\text{obs}} (\Sigma_{\text{obs}} \subseteq \Sigma)$ of the events can be observed $^1$, so that $\Sigma$ is partitioned into the set of observable events $\Sigma_{\text{obs}}$ and the set of unobservable events $\Sigma_{\text{un}} = \Sigma - \Sigma_{\text{obs}}$. The natural projection $P_{\Sigma_{\text{obs}}} : \Sigma^* \rightarrow \Sigma_{\text{obs}}^*$ can be used to map any trace executed in the system to the sequence of observable events associated with it. This projection is defined recursively as $P_{\Sigma_{\text{obs}}} (\sigma s) = P_{\Sigma_{\text{obs}}} (\sigma) \Pi_{\Sigma_{\text{obs}}} (s)$, for $s \in \Sigma^*$, $\sigma \in \Sigma$, with $P_{\Sigma_{\text{obs}}} (\varepsilon) = \sigma$ if $\sigma \in \Sigma_{\text{obs}}$ and $P_{\Sigma_{\text{obs}}} (\varepsilon) = \varepsilon$ if $\sigma \in \Sigma_{\text{un}} \cup \{ \varepsilon \}$ (where $\varepsilon$ represents the empty trace $[4]$). In the sequel, the subscript $\Sigma_{\text{obs}}$ in $P_{\Sigma_{\text{obs}}}$ will be dropped when it is clear from context. We denote an observation sequence of length $n$ as $\omega = \omega_1 \omega_2 ... \omega_n$, where $\forall i, \omega_i \in \Sigma_{\text{obs}}$.

2.1 Probabilistic Finite Automata and Markov Chains

Definition 1 (Probabilistic Finite Automata (PFA)). A probabilistic finite automaton (PFA) is captured by $H = (X, \Sigma, p, \pi_0)$, where $X = \{x_1, x_2, \ldots, x_{|X|}\}$ is the set of states, $\Sigma$ is the set of events, $\pi_0$ is the initial-state probability distribution vector, and $p(i', \sigma | i)$ is the state transition probability defined $^2$ for $i, i' \in X$, and $\sigma \in \Sigma$, as the conditional probability that event $\sigma$ occurs and the system transitions to state $i'$ given that the system is in state $i$. Clearly, we have $\sum_{i' \in X} \sum_{\sigma \in \Sigma} p(i', \sigma | i) = 1, \forall i \in X$ and $\sum_{i \in X} \pi_0 (i) = 1$.

We can associate with a given PFA $H = (X, \Sigma, p, \pi_0)$ a unique NFA $^3$ $G = (X, \Sigma, \delta, \x_0)$ where the state transition function $\delta : X \times \Sigma \rightarrow 2^X$ is defined for $i \in X$, $\sigma \in \Sigma$ as $\delta(i, \sigma) = \{ i' | p(i', \sigma | i) > 0 \}$ and the set of possible initial states is defined as $\x_0 = \{ i | \pi_0 (i) > 0 \}$. The function $\delta$ can be extended from the domain $X \times \Sigma$ to the domain $X \times \Sigma^*$ in the routine recursive manner: $\delta(x, \sigma s) = \delta(\delta(x, \sigma), s)$, if $\delta(x, \sigma)$ is defined and for a subset of states $S \subseteq X$ we have $\delta(S, \sigma) = \cup_{x \in S} \delta(x, \sigma)$. In this way, the behavior of the PFA $H$ is mapped to the behavior of the associated NFA $G$, i.e., $L(H) = L(G)$.

$^1$ Notice that, as discussed later, an NFA with $\Sigma = \Sigma_{\text{obs}} \cup \Sigma_{\text{un}}$ can be transformed to an NFA with $\Sigma = \Sigma_{\text{obs}}$ via an appropriate transformation.

$^2$ For simplicity, sometimes we refer to a state $x_i$ as state $i$.

$^3$ Later in this paper, we will use a deterministic finite automaton $D = (X, \Sigma, \delta, x_0)$, which is an NFA in which i) the initial state is not a set of states, but a single state (denoted by $x_0 \in X$), and ii) for all states $x_i$ and all events $\sigma$, if the transition is possible, then $\delta(x, \sigma) \subseteq X$ which means that the next state, if defined, is a single state.
We next quickly review useful results regarding Markov chains and PFAs.

1. A Markov chain is captured by $M = (X, A, \pi_0)$, where $X = \{x_1, x_2, ..., x_N\}$ is the set of states; $A$ is the state transition matrix defined so that its $i,j$th entry captures the probability of a transition from state $i$ to state $j$ given by $p_{ij} = \sum_{x} p(j, x | i)$; and $\pi_0$ is the initial state probability distribution vector.

2. A state $x_i \in X$ of a Markov chain is said to be periodic if the greatest common divisor $d$ of the set $\{n > 0 : \Pr(x[n] = x_i | x[0] = x_i) > 0\}$ is $d \geq 2$ (note that $x[t] = x_i$ denotes the event that the state of the Markov chain at step $t$ is $x_i$). If $d = 1$, state $x_i$ is said to be aperiodic. MC is aperiodic if all states $x_i \in X$ are aperiodic.

3. MC is irreducible if for all $j, i \in X$, there exists some $n \in \mathbb{N}$ such that $A^n(j, i) > 0$, where $A^n(j, i)$ captures the transition probability from state $i$ to state $j$ in exactly $n$ steps (given by the $(j, i)$th entry of matrix $A^n$) [18,4].

4. Let $\pi[t]$ be a $|X|$-dimensional vector whose $j$th entry denotes the probability of being in state $x_j$ after $t$ steps. We have $\pi[0] = \pi_0$ and $\pi[t] = A^t \pi_0$ for $t = 1, 2, ...$

5. If MC is irreducible and aperiodic, then $\lim_{t \to \infty} \pi[t]$ exists and is called the stationary distribution of the Markov chain denoted by $\pi = \{\pi_s(x_1), \pi_s(x_2), ..., \pi_s(x_N)\}$'s.

6. The recurrence time of state $x_i \in X$ is defined as $T_i = \inf\{n > 0 : x[n] = x_i\}$ [3,4]. Recurrence implies that a state is visited infinitely often. We write $x_i \in X_R$, (where $X_R \subseteq X$ is the set of recurrent states) if $\Pr(T_i < \infty | x[0] = x_i) = 1$ [3,4].

7. For $x_k, x_i \in X$ the corresponding element of the one-step transition probability matrix $\{12,7\}$ is given as

8. A PFA $H_0 = (X, \Sigma_{obs}, p_0, \pi_0)$ is obtained from a Markov chain by omitting the unobservable events ($\Sigma_{unobs} = \Sigma - \Sigma_{obs}$). The transition probability matrix $p_{ij} = \Pr(x[j+1] | x[j], \sigma)$ for $i, j \in X$ and $\sigma \in \Sigma_{unobs}$ is obtained by setting $\pi_0(i) = 1$, and calculating $p_{ij} = \sum_{\sigma \in \Sigma_{unobs}} \Pr(j, \sigma | i)$.

9. A closed strongly connected component (CSCC) is a PFA $H_i = (X_i, \Sigma_i, p_i, \pi_0)$ with $X_i \subseteq X$ such that $\forall x, x' \in X_i$ and $\forall \sigma \in \Sigma_i, p_i(x', \sigma | x) = p(x', \sigma | x)$, whenever the former is defined (zero otherwise) and $\pi_0$ is an $|X_i|$-dimensional probability vector that captures the initial probabilities associated with states in $X_i$. We say that $H_i$ is a strongly connected component or irreducible if its associated Markov chain is strongly

\[ \text{in the case of a periodic Markov chain with period } d, \text{ there exist } \pi_{s,r} = \lim_{t \to \infty} A^{d+r} \pi_0, \text{ } r \in \{0, ..., d-1\}. \text{ In that case, we use the average stationary distribution } \pi_s = \frac{1}{d} \sum_{r=0}^{d-1} \pi_{s,r}. \]
connected or irreducible. Note that \( H_i = (X_i, \Sigma, p_i, \pi_{0,i}) \) being a CSCC implies that \( \sum_{\sigma \in \Sigma} \sum_{x' \in X_i} P(x', \sigma | x) = 1 \) for \( x \in X_i \).

The classification problem between Markov chains, which is needed in our AA-detectability analysis, characterizes the ability of an observer to distinguish, among two available systems, the correct system, that generated the specific observation sequence. A pair of systems is classifiable if any of the available systems can be classified correctly, with arbitrarily high probability, as the observation sequence increases in length. Classification between Markov chains is a special case of the classification problem among PFAs, which is described in detail in [9]. Theorem 1 describes the necessary and sufficient conditions for this problem.

**Theorem 1 ([12], Part III, Chapter 12)** Consider two irreducible Markov chains denoted by \( MC_1 = (X_1, A_1, \pi_{0,1}) \) and \( MC_2 = (X_2, A_2, \pi_{0,2}) \) with the same set of transitions, (and thus the same transition diagrams). We can classify them iff their one-step transition probability matrices \( P^{(1)} \) and \( P^{(2)} \) are different.

### 3 Detectability in PFAs

Now we revisit two previously used stochastic detectability notions, and we present an example to demonstrate that AA-detectability considers more general classes of systems than either of these two notions.

**Definition 3 (Strong Stochastic Detectability [21].)** A PFA \( H = (X, \Sigma, p, \pi_0) \) is strongly (stochastic) detectable with respect to a set of observable events \( \Sigma_{\text{obs}} \subseteq \Sigma \) if from equally likely initial states (i.e., \( \pi_0 = \frac{1}{|\Sigma|} \)) all infinite sequences are convergent. This means

\[
(\forall 0 < \alpha < 1) (\exists N \in \mathbb{N})(\forall n \geq N) \{ (s \in \Sigma^* : \Pr(s) > 0 \land |s| = n) \rightarrow (\rho(P(s)) \geq \alpha) \}.
\]

**Definition 4 (A-Detectability [10].)** A PFA \( H = (X, \Sigma, p, \pi_0) \) is \( A \)-detectable with respect to a set of observable events \( \Sigma_{\text{obs}} \subseteq \Sigma \) if

\[
(\forall \epsilon > 0) (\exists N \in \mathbb{N})(\forall n \geq N) \{ \Pr\{ s \in \Sigma^* : |s| = n, |R(X_0, P(s))| > 1 \} < \epsilon \},
\]

where the projection \( P \) is taken with respect to the set of observable events and \( R(X_0, P(s)) \) is taken with respect to the NFA \( G \) associated with PFA \( H \).

Unlike Strong (Stochastic) Detectability in Definition 3, A-Detectability in Definition 4 takes into consideration the probability of occurrence of all possible observation sequences \( \omega \in P(L(H)) \), with \( |R(X_0, \omega)| > 1 \). This probability is required to be under a specific threshold \( \epsilon > 0 \) (which can be as small as we want). In other words, A-detectability takes into account the probability of a given set of (problematic) observation sequences. The verification of A-detectability was shown in [10] to be a PSPACE-hard problem.

**Definition 5 (AA-detectability).** A PFA \( H = (X, \Sigma, p, \pi_0) \) is AA-detectable with respect to a set of observable events \( \Sigma_{\text{obs}} \subseteq \Sigma \) if

\[
(\forall \epsilon > 0) (\forall \alpha > 0) (\exists N \in \mathbb{N})(\forall n \geq N) \{ \Pr\{ s \in \Sigma^* : |s| = n \land \rho(P(s)) \geq \alpha \} \geq 1 - \epsilon \}.
\]

**Example 1** The following example is used to clarify the notion of AA-detectability. Consider the PFA \( H = (X, \Sigma, p, \pi_0) \) depicted in Fig. 1 with \( \pi_0 = [1,0,0,0,0,0] \), where \( X = \{1,2,3,4,5,6\}, \Sigma = \{ \alpha, \beta \}, p \) is as defined by the transitions in the figure, and \( \Sigma_{\text{obs}} = \Sigma \). According to Definition 4 the system is not AA-detectable because \( \forall s \in L(H) : |R(X_0, P(s))| > 1 \). Note also that the infinite sequence \( s = \alpha \beta^n \), where \( n \) is arbitrarily large (\( n \geq 1 \)) is not convergent: for any \( n \), we have \( \omega = P(s) = P(\alpha \beta^n) = \alpha \beta^n \) and \( \rho(P(\omega)) = \rho_P(\omega)(3) = \frac{\pi_0(3)}{\pi_0(3) + \pi_0(5)} = \frac{1/3}{1/3 + 2/15} < 1 \), because \( \pi_0(3) = \frac{2(0.5)^n}{3} \) and \( \pi_0(5) = \frac{2(0.5)^n}{15} \). Although there exists an infinite sequence which is not convergent, the system in this example is, in fact, AA-detectable, because AA-detectability is related to the overall probability of non-convergent sequences. We will analyze this example in detail in Section 4.

![Fig. 1. PFA used in Example 1.](image-url)

**Remark 1** It is clear from the definitions of A-detectability and AA-detectability, that A-detectability implies AA-detectability, because \( \{ |R(X_0, P(s))| = 1 \} \rightarrow \{ \rho(P(s)) = 1 \} \rightarrow \{ \forall \alpha : \rho(P(s)) \geq \alpha \} \). On the other hand, there are cases (like Example 1 above) which clarify that AA-detectability does not imply A-detectability. It is worth pointing out that A-detectability does not take into account conditional probabilities in

\[
\sum_{\sigma \in \Sigma} \sum_{x' \in X} P(x', \sigma | x) = 1
\]
defining problematic sequences; AA-detectability generalizes both A-detectability (by taking into account conditional probabilities) and stochastic detectability (by considering the probability of non-convergent sequences). Also note that one can compare the notions of AA-diagnosability and AA-detectability in the most general context of differences and similarities between the problems of detectability and diagnosability as discussed in [10]. It is worth pointing out that some recent results indicate that AA-diagnosability can be polynomially verified [2].

4 Polynomial Verification of AA-Detectability

We first argue that AA-detectability for a PFA $H$ can be verified by focusing on continuations $t$ of any string $s$ that reaches the recurrent states of $H$. To see this, we first notice that if we concentrate on strings $s$ of length $|s| = N'$, we can make the combined probability of such strings that keep us in a non-recurrent state smaller than any (arbitrarily small) $\epsilon'$ by choosing $N'$ to be large enough. Suppose we also verify that AA-detectability holds for continuations $t$ of such strings $s$ in the recurrent states of $H$, i.e., for any $\epsilon$ and any $\alpha$, we can find an $N$ such that for strings $t$ of length $|t| \geq N$, the requirement in Definition 5 holds; then, the requirement also holds in the overall NFA for $\epsilon + \epsilon'$ and $\alpha$ for strings $st$ of length $|st| \geq N + N'$ (in the worst case, all the strings that we ignored by focusing on recurrent behavior are problematic); since $\epsilon'$ and $\epsilon$ can be arbitrarily small, the claim follows.

Theorem 2 We argue that a given PFA $H = (X, \Sigma, p, \pi_0)$ with $m$ CSCCs, $H_1, H_2, ..., H_m$, $m \geq 1$, is AA-detectable iff both of the following hold: (i) each $H_i$, $i = 1, 2, ..., m$, has a corresponding NFA that is actually a DFA (i.e., it has a deterministic transition function) which is a direct consequence of Theorem 3 (in the appendix), and (ii) for all CSCCs $H_{i_1}, H_{i_2}, ..., H_{i_k}$, $k \in \{1, 2, ...\}$ (with repetitions allowed) that are reachable under strings $s_1, s_2, ..., s_k$, with $P(s_1) = P(s_2) = ... = P(s_k)$, the probability of error when trying to classify between these different PFAs tends asymptotically to zero.

The remaining paper serves as proof of Theorem 2, which essentially is the description and justification of the verification process outlined as Algorithm 1 below.

1) We identify all CSCCs of PFA $H$. The CSCCs depend only on the graph structure of the state transition diagram of the given PFA. In case there exists at least one CSCC associated with a finite automaton with non-deterministic transitions, then the system is immediately declared as not AA-detectable (see appendix, Theorem 3).

2) We find all pairs of CSCCs that are reachable, possibly under different strings that have identical projection (i.e., $(H_i, H_j)$).

3) We use a polynomial complexity construction called Detector (Definition 6), which captures only pairs of state estimates. We can use the Detector, because:

a) according to Theorem 3 it is necessary for all CSCCs $H_i, i = 1, ..., m$, of PFA $H$, to have associated NFAs, $D_i, i = 1, ..., m$ that are DFAs:

b) we can distinguish all CSCCs by distinguishing all possible pairs of them.

4) We construct the associated PFAs, without the unobservable events (for simplicity we keep the same symbols $H_i$ and $H_j$), with underlying DFAs $D_i$ and $D_j$. We construct the Detector $G_{d_{ij}}$ for DFA $D_{ij}$ (see Definition 7), which is the union of $D_i$ and $D_j$. Then we construct the PFAs $H_{d_{ij},i}$ and $H_{d_{ij},j}$ that capture the common behaviour of $H_i$ and $H_j$ as their underlying DFA is the Detector $G_{d_{ij}}$. PFA $H_{d_{ij},ij}$ is assigned probabilities from $H_i$ and $H_{d_{ij},j}$ is assigned probabilities from $H_j$. We can classify PFAs $H_i$ and $H_j$ iff we can classify $H_{d_{ij},i}$ and $H_{d_{ij},j}$. The classification of these PFAs is equivalent to the classification of their underlying Markov chains, since the two PFAs are completely observable. Their underlying Markov chains are irreducible and share the same transition diagram, and their classification can be performed according to Theorem 1.

5) Finally, we decide that $H$ is AA-detectable iff we can classify all pairs of PFAs $H_i$ and $H_j$.

Remark 2 Steps 1, 2, and 3, are required to identify the pairs of CSCCs and equivalently the PFAs that we need to compare, to verify AA-detectability. We provide the justification of the use of a polynomial complexity construction which we call the Detector (we can transform the problem of classification among $m$ PFAs to a problem of classification among all possible pairs of these $m$ PFAs by extending the techniques in [9]). In Step 4 we perform classification among pairs of PFAs that were identified in previous steps. The general problem of classification is explored in many works, such as [22], where the problem is referred to as monitorability, and in [11], where the problem is referred to as distinguishability between two Hidden Markov chains. Theorem 3 helps us in transforming the problem of classification among two PFAs to the well studied problem of classification between two Markov chains.

In our case, Theorem 3 implies that the identified PFAs should be associated with deterministic finite automata. Therefore, our classification problem is a special case, of the general problem. This result provides the system designer with useful information about the general charac-

---

7 Polynomial graph algorithms for the identification of the CSCCs of a given graph can be found in [5], [6].

8 In fact, we could use the Detector even without Theorem 3. We are interested in being able to classify different CSCCs, without knowing how many of them are simultaneously reachable. This allows us to define these CSCCs as PFAs with only one possible initial state. Theorem 3 plays an important role later, when we need to assign probabilities, using the Detector.
Algorithm 1. Verification of AA-DETECTABILITY

Input: PFA $H = (X, \Sigma, p, \pi_0)$ with corresponding NFA $D = (X, \Sigma_{obs}, \delta, X_0)$

Output: AA-detectable/not AA-detectable

1. Label = 1
2. if $\exists s \in L(H), x_i \in X_i, x_j \in X_j$ : $x_i, x_j \in \delta_d(X_{0d}, p(s))$ then
3. Construct $D_{ij} = (X_{ij}, \Sigma_{obs}, \delta_{ij}, X_{0ij})$
4. Construct the Detector $D_{d,ij} = (X_{d,ij}, \Sigma_{obs}, \delta_{d,ij}, X_{0d,ij})$ of DFA $D_{ij}$
5. Construct PFA $H_{d,ij} \equiv (X_{d,ij}, \Sigma_{obs}, \pi_{0d,ij})$, with NFA $G_{d,ij}$ and Markov chain $M_{d,ij} = (X_{d,ij}, A_{d,ij}, \pi_{0d,ij})$ with one-step transition probability matrix $P^1$
6. Construct PFA $H_{d,ij} = (X_{d,ij}, \Sigma_{obs}, \pi_{0d,ij})$, with NFA $G_{d,ij}$ and Markov chain $M_{d,ij} = (X_{d,ij}, A_{d,ij}, \pi_{0d,ij})$, with one-step transition probability matrix $P^2$
7. if $P^1 = P^2$ then $H = (X, \Sigma_{obs}, \delta, X_0)$, with associated Markov chains that are irreducible (note that the PDAs are constructed from the initial PFAs by removing the unobservable events). Suppose that the finite automata associated with these PDAs are DFA’s $D_i = (X_i, \Sigma_{obs}, \delta_i, X_{0i})$ and $D_j = (X_j, \Sigma_{obs}, \delta_j, X_{0j})$. (1) We construct $D_{ij} = (X_{ij}, \Sigma_{obs}, \delta_{ij}, X_{0ij})$, with $X_{ij} = X_i \cup X_j$ (assume, without loss of generality, that $X_i \cap X_j = \emptyset$), $\delta_{ij}(x, \sigma) = x'$ iff $\{x, x' \in X_i \wedge \delta_i(x, \sigma) = x'\} \cup \{x, x' \in X_j \wedge \delta_j(x, \sigma) = x'\} \wedge X_{0ij} = X_{0i} \cup X_{0j}$
8. (2) We construct the detector (Definition 6) $G_{d,ij} = (X_{d,ij}, \Sigma_{obs}, \delta_{d,ij}, X_{0d,ij})$ of DFA $D_{ij}$
9. (3) We construct the PDAs $H_{d,ij} = (X_{d,ij}, \Sigma_{obs}, \pi_{0d,ij})$ and $H_{d,ij} = (X_{d,ij}, \Sigma_{obs}, \pi_{0d,ij})$, by assigning probabilities over the detector DFA $G_{d,ij}$. We have $p_{d,ij} = (x_k, x') \rightarrow \sigma(x_k, x') = p_{ij}(x', \sigma(x_k))$, where $x_k, x' \in D_i$ and $x_i, x' \in D_j$ (we can also have $x_k = \emptyset$ as a value, or the same state for the two components, which in the definition of the detector is a nondeterministic finite automaton, where $X_0d = R(X_0, e)$ is the set of all possible initial states for NFA $G$).

Fig. 2. Detector for underlying system $G$ for PFA $H$ in Example 1. States $\{2, 4, 6\}$ and $\{3, 5\}$ form an automaton (inside the cycle) in which we assign probabilities (Definition 7).

Definition 7 (PDAs $H_{d,ij}$ and $H_{d,ij}$ with associated detector $G_{d,ij}$) Consider two PDAs $H_i = (X_i, \Sigma_{obs}, \pi_{0i})$ and $H_j = (X_j, \Sigma_{obs}, \pi_{0j})$ with associated Markov chains that are irreducible (note that the PDAs are constructed from the initial PDAs by removing the unobservable events). Suppose that the finite automata associated with these PDAs are DFA’s $D_i = (X_i, \Sigma_{obs}, \delta_i, X_{0i})$ and $D_j = (X_j, \Sigma_{obs}, \delta_j, X_{0j})$. (1) We construct $D_{ij} = (X_{ij}, \Sigma_{obs}, \delta_{ij}, X_{0ij})$, with $X_{ij} = X_i \cup X_j$ (assume, without loss of generality, that $X_i \cap X_j = \emptyset$), $\delta_{ij}(x, \sigma) = x'$ iff $\{x, x' \in X_i \wedge \delta_i(x, \sigma) = x'\} \cup \{x, x' \in X_j \wedge \delta_j(x, \sigma) = x'\} \wedge X_{0ij} = X_{0i} \cup X_{0j}$

Now we provide the relevant definitions and an example in order to complete the presentation of the verification algorithm.

Definition 6 (Detector [19].) Given an NFA $G = (X, \Sigma, \delta, X_0)$ under the natural projection map $P$ with respect to $\Sigma_{obs} \subseteq \Sigma$, the detector $G_d = (X_d, \Sigma_{obs}, \delta_d, X_{0d})$ is a nondeterministic finite automaton, where $X_0d = R(X_0, e)$ is the set of all possible initial states for NFA $G$.

(2) $X_d = X_p \cup X_u \cup \{X_0d\}$ is the finite set of states, with $X_p \equiv \{x_{d1}, x_{d2}, \ldots, x_{dn}\}$, where $D = |X \times X| - |X|$ with $x_{d1} = \{x_{11}, x_{12}, \ldots, x_{1m}\} \in X_p$, $x_{d1} \neq x_{m}, x_{1m}, x_{m} \in X$, and $X_0 = \{\{x_i\} | x_i \in X\}$.

(3) $\Sigma_{obs}$ is the finite set of observable events;

(4) $\delta_d : X_d \times \Sigma_{obs} \rightarrow X_d$ captures the state transitions and is defined as follows:

$$\delta_d(x_d, \sigma) = \begin{cases} \{x_{d1} \in X_p | x_{d1} \subseteq R(x_d, \sigma)\}, & \text{if } |R(x_d, \sigma)| > 1, \\ \{x_1\} \in X, & \text{if } R(x_d, \sigma) = \{x_1\}, \\ \text{undefined}, & \text{if } R(x_d, \sigma) = \emptyset. \end{cases}$$
Remark 3 The identification of all CSCCs and the construction of the detectors are of polynomial complexity with respect to the number of states of the PFA. The comparison for the one-step transition probabilities is also of polynomial complexity, because it is based on the computation of the steady-state probabilities for two Markov chains. Overall the verification procedure for AA-detectability is of polynomial complexity with respect to the number of states of the PFA: if we are given a PFA with $|X|$ states, then we have to check at most $|X|^2$ starting states in pairs of closed strongly connected components; each of these closed strongly connected components has size at most $|X|$, thus the product detectors have size at most $|X|^2$; all we have to do is to check whether the one-step transition probabilities for the two resulting Markov chains are different, which can be done with complexity $O(|X|^2)$ as it involves the computation of the steady state probabilities of a Markov chains of size $O(|X|^2)$ [7]. Thus, the overall complexity of our approach is $O(|X|^8)$.

5 Conclusions

In this paper we discussed the notion of AA-detectability in PFAs, which is inspired by analogous notions in detectability and stochastic diagnosability [23]. We applied methods closely related to those used in classification of PFAs to verify AA-detectability with polynomial complexity. A possible future research direction is the computation of bounds on the probability of error in state estimation problems in stochastic DES.

References

control coding, monitoring, diagnosis and control of large-scale discrete-event systems, and applications to network security, anomaly detection, and energy distribution systems.

A Proof of Theorem 3

Theorem 3 (A necessary condition for AA-detectability). Given a PFA $H = (X, \Sigma, p, \pi)$, its corresponding NFA $G = (X, \Sigma_{obs}, \delta, X_0)$, and its associated underlying Markov chain $M = (X, A, \pi_0)$, where $X = X_R \cup X_T$, with $X_R$ being the set of recurrent states and $X_T$ being the set of transient states, a necessary condition for $AA$-detectability is that there do not exist $x_j, x_k, x_i \in X_R$ and $s^k, s^j \in \Sigma_{nu}$ and $\sigma \in \Sigma$, such that $p(x_k, s^\sigma | x_j) > 0$ and $p(x_i, s^\sigma | x_j) > 0$.

PROOF. Let us suppose that the given PFA is AA-detectable, and let us assume that there exists at least one CSCC $H_1$. Consider all $s \in L(G)$, such that $s$ reaches a state $x \in X_R$ in the CSCC $H_1$, with some nonzero probability. For $0 < \alpha < 1$, let event $A(N, \alpha) = \{ s \in L(G) : |s| = n \geq N \land p(P(s)) > \alpha \}$, and let $A^C(N, \alpha)$ the complement of $A$. Furthermore, for $x_i, x_k$, let $A(N, \alpha, x_i) = \{ s \in L(G) : |s| = n \geq N \land p_{P(s)}(x_i) > \alpha \}$. The system being AA-detectable means

$$(\forall \epsilon > 0)(\forall n < \alpha < 1)(\exists N \in \mathbb{N})(\forall n \geq N) Pr(A(N, \alpha)) \geq 1 - \epsilon.$$
Pr(A^C(N', \alpha) \land A(N, \alpha)) \geq \epsilon' \text{ for some } \epsilon' \text{ to be specified.}

Note that

\[
\Pr(A^C(N', \alpha) \land A(N, \alpha)) = \Pr(A^C(N', \alpha) | A(N, \alpha)) \Pr(A(N, \alpha)) = \sum_{\forall x \in X_R} \Pr(A^C(N', \alpha) | A(N, \alpha, x)) \Pr(A(N, \alpha, x) | A(N, \alpha)) \Pr(A(N, \alpha)).
\]

Clearly, there exists \( x_i \in X_R \), for which \( \Pr(A(N, \alpha, x_i) | A(N, \alpha)) \geq \frac{1}{|X|}. \) For the chosen \( x_i \), following the line of thought of the previous paragraph, we can find an acyclic path \( t' = t(ij) \sigma' \), with length \( |t(ij)| + k + 1 \), which means that we can choose \( N' = N + |t(ij)| + k + 1 \), such that

\[
\Pr(A^C(N', \alpha) | A(N, \alpha, x_i)) \geq p_{\min}^{2|X|} (2|X|) \text{ is the maximum length of a possible } t'.
\]

Thus, \( \Pr(A^C(N', \alpha)) \geq p_{\min}^{2|X|} (1 - \epsilon). \) Due to the AA-detectability property, we need to have

\[
\frac{p_{\min}^{2|X|}}{|X|} (1 - \epsilon) \leq \Pr(A^C(N', \alpha)) < \epsilon,
\]

this double inequality holds if and only if \( \epsilon \geq \epsilon_0 \), where \( \epsilon_0 = \frac{p_{\min}^{2|X|}}{|X|} + 1 \).

In particular, for all \( \epsilon < \epsilon_0 \), the inequality does not hold, and therefore, for \( \epsilon < \epsilon_0 \), we cannot find \( N \in \mathbb{N} \), so that AA-detectability holds. Thus, we have reached a contradiction, and the proof is completed. \( \square \)