Probabilistic Analysis of Predictability in Discrete Event Systems

Farid Nouioua, Philippe Dague, Lina Ye

Probabilistic Analysis of Predictability in Discrete Event Systems

Farid Nouioua[1] and **Philippe Dague**[2] and **Lina Ye**[3]
[1]LSIS, Aix-Marseille Univ., France
e-mail: farid.nouioua@lsis.org
[2]LRI IASI, Univ. Paris-Sud 11., France
e-mail: philippe.dague@lri.fr
[3] INRIA Grenoble - Rhône-Alpes.,France
e-mail: Lina.Ye@inria.fr

Abstract

Predictability is a key property allowing one to expect in advance the occurrence of a fault in a system based on its observed events. Existing works give a binary answer to the question of knowing whether a system is predictable or not. In this paper, we consider discrete event systems where probabilities of the transitions are available. We show how to take advantage of this information to perform a Markov chain-based analysis and extract a variety of probability values that give a finer appreciation of the degree of predictability. This analysis is particularly important in case of non predictable systems. We consider a "light" analysis that focuses only on predictability as well as a "deep" analysis that handles in a uniform framework both predictability and diagnosability.

1 Introduction

Nowadays, real-life systems are more and more complex and increasingly need to be highly autonomous. Faults diagnosability is a key property to increase the autonomy of such systems. This property has been extensively studied in the last years. The seminal work in [Sampath *et al.*, 1995] provided an algorithm to verify diagnosability in discrete event systems (DES) represented by finite automata, based on the so-called diagnoser. As the number of states in a diagnoser is exponential in the number of states in the system, polynomial algorithms have been proposed that are based on the twin plant approach [Jiang *et al.*, 2001; Yoo and Lafortune, 2002]. Further developments have been done around diagnosability including: diagnosability of single faults as well as patterns in distributed systems [Schumann and Pencolé, 2007; Ye and Dague, 2012], the use of model-checking techniques [Cimatti *et al.*, 2003], algebraic languages [Console *et al.*, 2000] or satisfiability formalism [Rintanen and Grastien, 2007] and the verification of diagnosability in probabilistic DES [Nouioua and Dague, 2008; Thorsley and Teneketzis, 2005].

Diagnosability ensures the ability to detect faults after their occurrences. However, since it is not always easy to recover the system after the faults occurred, a stronger property has to be considered: the ability of the system to predict the faults before their occurrences. The prediction of a fault may be very useful in practice. Indeed, when the fault is predicted, appropriate measures may be taken to avoid its negative effects. In [Genc and Lafortune, 2009] the diagnoser and the twin plant approaches have been adapted to verify predictability. The work in [Jeron *et al.*, 2008] concerns the predictability of patterns and that in [Cassez and Grastien, 2013] deals with timed DES. The predictability of distributed DES has been studied in [Ye *et al.*, 2013]. In the previous works, the decision about predictability tells simply either the system is predictable or not. However, the "degree" of non predictability is not the same from a system to another. Indeed, if a first system contains only a low proportion of traces where the fault cannot be predicted while a second one contains a much greater proportion of such traces, it would be plausible to associate a measure of non predictability that is more important in the latter system than in the former one. Moreover, this kind of measure may be beneficial in practice. For instance, it may be better in contexts where the consequences of a fault are not very critical, to tolerate a system with a sufficiently low degree of non-predictability than to add the missing sensors which can be very expensive. More generally, the more the application is critical and the consequences of their faults are dangerous, the higher is the threshold representing the minimal degree of probability that may be tolerated for predictability.

This paper extends the approach proposed in [Nouioua and Dague, 2008] for the analysis of diagnosability in probabilistic DES to deal with predictability. Two approaches are proposed. The first one based on a so called light estimator, is devoted to analyze only predictability while the second one uses a deep estimator as a uniform framework that allows one to analyze jointly predictability and diagnosability. In both cases, the idea consists in extracting an appropriate Markov chain explaining the dynamics of the system. Then, the results of the asymptotic behavior of this chain determines the probabilities of different classes of traces. These probabilities give a synthetic quantitative appreciation of the degree of non predictability and/or non diagnosability.

The outline of the paper is as follows. The probabilistic model is presented in section 2. Section 3 recalls the diagnoser-based approach to verify predictability and diagnosability in classical DES. Section 4 is devoted to the presentation of the light and the deep estimators. Section 5 shows how to use these estimators to perform a Markov-chain probabilistic analysis of predictability and/or diagnosability in probabilistic DES. Finally, in section 6. we conclude and give some perspectives of future work.

2 Probabilistic Discrete Event Model

The model used in this paper is that of a probabilistic discrete event system (PDES) which consists of a classical DES enriched by probability values on its transitions.

Definition 1. A probabilistic discrete event system $(PDES)$ is modeled by the structure $\Gamma = (X, E, \theta, x_0)$ where $X = \{x_0, ..., x_{n-1}\}$ is a finite set of states ($|X| = n$), $E = \{e_0, ..., e_{m-1}\}$ is a finite set of events ($|E| = m$), x_0 is the initial state and $\theta : X \times E \times X \longrightarrow [0..1]$ is a probabilistic transition function: $\theta(x, e, x') = \alpha$ $(0 \leq \alpha \leq 1)$ is the probability that the event e occurs in x and causes the transition of the system from state x to state x'.

To a PDES $\Gamma = (X, E, \theta, x_0)$ we associate a classical DES $G = (X, E, \delta, x_0)$ where the transition function $\delta \subseteq X \times E \times X$ of G is defined by: $(x, e, x') \in \delta \Leftrightarrow \theta(x, e, x') > 0$. Thus, the DES G is obtained from Γ by removing the probability values. δ can be generalized as usual to words of E^* (Kleene closure of E). For $s \in E^*$ with $s = a_1 \dots a_k$, $(x_0, s, x) \in \delta$ iff there is sequence of states $x_1, \dots, x_k$ sucht that $x_k = x$ and $(x_{i-1}, a_i, x_i) \in \delta$ for $1 \leq i \leq k$.

We denote by $L \subseteq E^*$ the language generated by G. L is prefix closed. $E = E_o \cup E_{uo}$ where E_o (resp. E_{uo}) contains the observable (resp. unobservable) events. $E_f \subseteq E_{uo}$ is a subset of unobservable faulty events. Moreover, faults are partitioned into disjoint sets corresponding to the different fault types: $E_f = E_{f_1} \cup \dots \cup E_{f_p}$. In the sequel, we will focus, without loss of generality, on one fault type as in [Yoo and Lafortune, 2002; Schumann and Pencolé, 2007]. Indeed, since the system is predictable if and only if it is predictable for each fault type then, to check the predictability of a system, it suffices to check its predictability for each fault type by considering all the other faults as non observable. For the sake of simplicity, we will denote by f each occurrence of the considered fault type. We suppose also that L is live, i.e., there is no cycle in G with only unobservable events, and that we represent in the model all the possible transitions of the system in each state. Thus, for each $x \in X$: $\sum_{y \in X} \sum_{e \in E} \theta(x, e, y) = 1$.

Example 1. Figure 1 shows an example of a PDES $\Gamma = (X, E, \theta, x_0)$ where: $X = \{x_0, x_1, x_2, x_3, x_4, x_5, x_6\}$, $E = E_o \cup E_{uo}$ with $E_o = \{a, b, c\}$ and $E_{uo} = \{f, u\}$, the set of fault events is $E_f = \{f\}$, the initial state is x_0 and the transition function is shown in figure 1[1]. Note that in absence of a transition from a state x to a state x' with an event e we have: $\theta(x, e, x') = 0$.

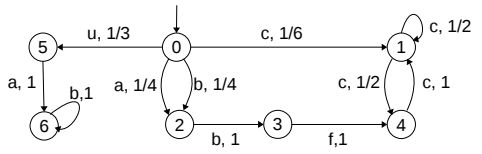

Figure 1: A Probabilistic DES

A word of L is also called trace. The empty trace is denoted by ϵ. The postlanguage of L after a trace s is: $L/s = \{t \in E^* | st \in L\}$. The set of prefixes of a word s is denoted by $\overline{s}$. $P : E^* \longrightarrow E_o^*$ is a projection function that erases from any trace its unobservable events: $P(\sigma) = \epsilon$ if $\sigma = \epsilon$ or $\sigma \in E_{uo}$, $P(\sigma) = \sigma$ if $\sigma \in E_o$ and $P(s\sigma) = P(s)P(\sigma)$ for $s \in E^*$ and $\sigma \in E$. P_L^{-1} is the inverse projection: for any $w \in E_o^*$, $P_L^{-1}(w) = \{s \in L | P(s) = w\}$. It provides, for a sequence of observable events w, all traces of L whose projection is w. We denote by s_f the final event of a trace s and by $\Psi(f)$ all traces ending in the fault event f: $\Psi(f) = \{s \in L | s_f = f\}$ and we define: $X_o = \{x_0\} \cup \{x \in X \mid \exists y \in X, \exists e \in E_o, (y, e, x) \in \delta\}$. Let $L(G, x)$ denote the set of traces originating from x, $L_o(G, x)$ denotes the set of traces originating from x and ending at the first observable event and $L_\sigma(G, x)$ the subset of $L_o(G, x)$ containing traces that end at the observable event σ: $L_o(G, x) = \{s \in L(G, x) \mid s = u\sigma, u \in E_{uo}^*, \sigma \in E_o\}$, $L_\sigma(G, x) = \{s \in L_o(G, x) \mid s_f = \sigma\}$.

3 The "Binary" Predictability

3.1 Basic definitions

Intuitively, a fault is predictable in a system iff, based on observed events, one can deduce each occurrence of this fault, before it actually occurs. It is diagnosable iff we can deduce without confusion after a finite delay of observations whether the fault occurred or not:

Definition 2. [Genc and Lafortune, 2009; Sampath *et al.*, 1995]. f is predictable iff: $(\exists n \in \mathbb{N})(\forall s \in \Psi(f))(\exists t \in \overline{s})[(f \notin t) \wedge P]$, where the predictability condition P is: $(\forall u \in L)(\forall v \in L/u)[(P(u) = P(t)) \wedge (f \notin u) \wedge (\|v\| \geq n) \Rightarrow (f \in v)]$.

f is diagnosable iff: $(\exists n \in \mathbb{N})[\forall s \in \Psi(f)](\forall t \in L/s)[\|t\| \geq n \Rightarrow D]$), where the diagnosability condition D is: $w \in P_L^{-1}[P(st)] \Rightarrow f \in w$.

It has been shown that predictability is stronger than diagnosability [Genc and Lafortune, 2009], i.e., a predictable fault is always diagnosable but the inverse is not true. Indeed, it is easy to check that in example 1., the fault f is diagnosable but it is not predictable.

3.2 Checking the "binary" predictability

We start by recalling how to check the "binary" predictability (where probabilities are not taken into account). We use for that the algorithm described in [Genc and Lafortune, 2009] and based on the notion of diagnoser introduced first in [Sampath *et al.*, 1995] to check diagnosability. It is worth mentioning that the twin plant approach [Jiang *et al.*, 2001] is better in terms of complexity. However, the structure of the diagnoser is much more adapted to the Markov chain-based analysis we perform in this work (see a more detailed discussion of this issue in the conclusion).

Before recalling the notion of diagnoser, let us first recall the notion of generator: The generator G' is defined by $G' = (X_o, E_o, \delta_{G'}, x_0)$ where X_o, E_o and x_0 have already been defined. $\delta_{G'}$ is such that: $(x, \sigma, x') \in \delta_{G'}$ *if* $(x, s, x') \in \delta$ *for some* $s \in L_\sigma(G, x)$. The corresponding probabilistic generator is defined by $\Gamma' = (X_o, E_o, \theta_{\Gamma'}, x_0)$ where X_o, E_o and x_0 are the same as in G' and the probabilistic transition function $\theta_{\Gamma'} : X_o \times E_o \times X_o \longrightarrow [0, 1]$ is defined by: $\theta_{\Gamma'}(x, \sigma, x') = \sum_{s \in L_\sigma(G,x)} \theta(x, s, x')$. We have that the

[1] To simplify the figures, we represent a state name x_i by its index i.

sum of the probabilities of all transitions issued from each state of Γ' equals 1:

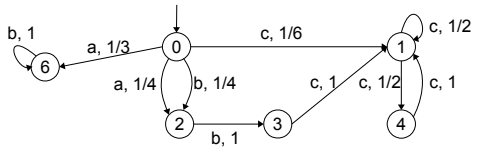

Figure 2: The probabilistic generator

Proposition 1. $\forall x \in X_0, \sum_{\sigma \in E_o} \sum_{x' \in X_0} \theta_{\Gamma'}(x, \sigma, x') = 1$.

Proof. Let us put $\lambda_x = \sum_{\sigma \in E_o} \sum_{x' \in X_0} \theta_{\Gamma'}(x, \sigma, x')$. By the definition of the generator we have : $\lambda_x = \sum_{\sigma \in E_o} \sum_{x' \in X_0} \sum_{s \in L_\sigma(G,x)} \theta(x, s, x')$. By the commutativity of addition, we have: $\lambda_x = \sum_{x' \in X_0} \sum_{\sigma \in E_o} \sum_{s \in L_\sigma(G,x)} \theta(x, s, x')$ which may be rewriten as: $\lambda_x = \sum_{x' \in X_0} \sum_{s \in L_o(G,x)} \theta(x, s, x')$.

For a path $x_1, \ldots, x_{n+1}$ labelled by the trace $e_1, \ldots e_n (n \geq 1)$ i.e., $\delta(x_i, e_i) = x_{i+1}$ for $1 \leq i \leq n$, we put $events(t) = e_1, \ldots, e_n$. The probability of the path t is $p(t) = \prod_{i=1}^{n} \theta(x_i, e_i, x_{i+1})$. for $x \in X$, we denote by T_x the set containing every path t starting from x and such that $events(t) = u\sigma$ with $u \in E_{uo}^*$ and $\sigma \in E_o$, i.e. paths that are labelled by a trace formed by an arbitrary number of unobservable events followed by one observable event. It is clear that $L_o(G, x) = \{events(t) | t \in T_x\}$. Thus, the previous expression for λ_x may be reformulated by : $\lambda_x = \sum_{t \in T_x} p(t)$.

Now, let us denote by $T_{x,y}^e$ the subset of T_x containing the paths that start by the transition (x, y) labelled by the event e, then $\lambda_x = \sum_{y \in X} \sum_{e \in E} \sum_{t \in T_{x,y}^e} p(t)$. We have to prove that for any $x, y \in X$ and $e \in E$, we have $\sum_{t \in T_{x,y}^e} p(t) = \theta(x, e, y)$. Indeed, by proving this equality, we obtain: $\lambda_x = \sum_{y \in X} \sum_{e \in E} \theta(x, e, y) = 1$ (by the definition of the model).

Let us put $\beta_{x,y}^e = \sum_{t \in T_{x,y}^e} p(t)$ and prove that $\beta_{x,y}^e = \theta(x, e, y)$. We will use an induction on the lenght of the longer path in $T_{x,y}^e$. This maximal lenght is denoted $MaxT_{x,y}^e$. Since every path in $T_{x,y}^e$ contains at least the transition (x, y) labelled by e, we have $MaxT_{x,y}^e \geq 1$.

- for $MaxT_{x,y}^e = 1$, $T_{x,y}^e$ contains one path constituted from the unique transition (x, y) labelled by e. It is obvious in this case that $\beta_{x,y}^e = \theta(x, e, y)$.
- Suppose that for any $x, y \in X$ and $e \in E$, if $T_{x,y}^e \leq n$ then $\beta_{x,y}^e = \theta(x, e, y)$. Let us take $x, y \in X$ and $e \in E$ such that $MaxT_{x,y}^e = n + 1$.
  $\beta_{x,y}^e = \sum_{t \in T_{x,y}^e} p(t) = \sum_{t' \in T_y} (\theta(x, e, y) \times p(t')) = \theta(x, e, y) \times \sum_{t' \in T_y} p(t') = \theta(x, e, y) \times \sum_{z \in X} \sum_{e' \in E} \sum_{t'' \in T_{y,z}^{e'}} p(t'')$. Clearly $MaxT_{y,z}^{e'} \leq n$, then by the induction hypothesis, we obtain: $\sum_{t'' \in T_{y,z}^{e'}} p(t'') = \theta(y, e', z)$. From the previous equality, we have: $\beta_{x,y}^e = \theta(x, e, y) \times \sum_{z \in X} \sum_{e' \in E} \theta(y, e', z) = \theta(x, e, y)$.

□

Now let us recall the notion of diagnoser:

Definition 3. A diagnoser is a deterministic automaton which is defined by $G_d = (Q_d, E_o, \delta_d, q_0)$ where:

- $Q_d \subseteq 2^{X_o \times \{N,F\}}$. A state q_d of Q_d is of the form: $q_d = \{(x_1, l_1), \ldots, (x_k, l_k)\}$ where $x_i \in X_o$ and $l_i \in \{N, F\}$.
- $q_0 = \{(x_0, N)\}$ is the initial state of the diagnoser G_d.
- E_o is the set of the observable events.
- $\delta_d : Q_d \times E_o \longrightarrow Q_d$ is the transition function of the diagnoser defined by: $\delta_d(q, \sigma) = \bigcup_{(x,l) \in q} \bigcup_{s \in L_\sigma(G,x)} \bigcup_{(x,s,x') \in \delta} \{(x', LP(x, l, s))\}$ where $LP : X_o \times \{N, F\} \times E^* \longrightarrow \{N, F\}$ is a label propagation function defined by: if $l = N$ and $f \notin s$ then $LP(x, l, s) = N$ else $LP(x, l, s) = F$.

Figure 3-(a) depicts the diagnoser of the system presented in example 1.

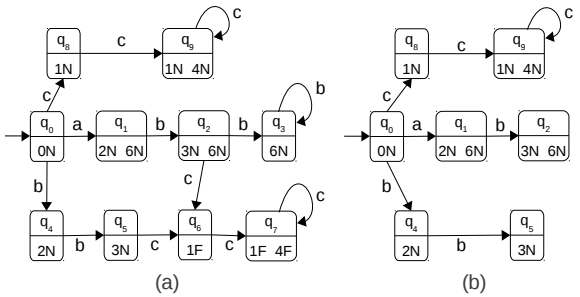

Figure 3: (a) The diagnoser G_d (b) The simplified diagnoser G_d'

A state q of G_d is f-uncertain if $\exists (x, l), (x', l') \in q$ such that $l = N$ and $l' = F$. It is f-certain (resp. normal) if $\forall (x, l) \in q, l = F$ (resp. $\forall (x, l) \in q, l = N$). A set of f-uncertain states forms an f-indeterminate cycle if this set forms a cycle in G_d to which correspond (with the same observed trace) in the generator G' a cycle reached from x_0 involving the fault and a cycle reached from x_0 without involving the fault.

We denote by Q^N the set of normal states of G_d. Let $\mathcal{C}$ be the set of normal states having an immediate successor that is not normal (i.e., f-uncertain or f-certain). We call these states, the critical states: $\mathcal{C} = \{q \in Q^N | \exists q' = \delta_d(q, o)$ such that $o \in E_o$ and $q' \notin Q^N\}$. We put $\mathcal{C}_{OK} = \{q \in \mathcal{C}|$ all the accessible cycles from q contain only f-certain states $\}$ and $\mathcal{C}_{KO} = \mathcal{C} \setminus \mathcal{C}_{OK}$. We have:

- f is predictable iff all the accessible cycles from each state of $\mathcal{C}$ are formed exclusively by f-certain states (i,e. $\mathcal{C}_{KO} = \emptyset$).
- f is diagnosable if and only if its diagnoser G_d contains no f-indeterminate cycle.

Example 1 (cont). From figure 3-(a), it is clear that f is diagnosable because the diagnoser does not contain any f-uncertain state. Moreover, we have: $\mathcal{C} = \{\{(3, N)\}, \{(3, N), (6, N)\}\}$. While the only accessible cycle from $\{3, N\}$ is constituted from f-certain states, we may reach from $\{(3, N), (6, N)\}$ either a cycle of normal states or a cycle of f-certain states. The fault f is then not predictable.

4 The Light and Deep Estimators

The diagnoser gives us a general information about the state of the system after the observation of a sequence of observable events. For example, let us suppose that we have, in the diagnoser, a path from the initial state $q_0 = \{(x_0, N)\}$ to some f-uncertain state $q = \{(x_1, N), (x_2, F)\}$ and that this path is labeled by the trace w of observable events. If we observe the trace w, we can deduce that the system is either in state x_1 and no fault occurred or in state x_2 with the occurrence of a fault. In a probabilistic framework, the probabilities to be in state x_1 or in state x_2 are not necessarily the same. However, the probability to observe w, independently from the target state, represents the probability to be in a f-uncertain state.

The estimator is a PDES which makes explicit this information: a state of the estimator is composed of a state name from the original system, a fault label (N or F) and new attributes that indicate further information wrt occurrence of the fault, ambiguity of the corresponding diagnoser state, etc. Thus, the transitions of the estimators correspond to a refinement of those of the diagnoser. Indeed, if there is a transition in the diagnoser from a state q_1 to a state q_2 labelled by an observable σ then, for each sub-state in q_2, we have at least one sub-state in q_1 which transits to it by σ, of course, with some probability. The estimator makes explicit these internal transitions and their probabilities.

4.1 Light estimator

The light estimator (L-estimator) is constructed from a part of the diagnoser (henceforth called the simplified diagnoser) which ignores any state that is accessible necessarily through a critical state.

Definition 4. Let $G_d = (Q_d, E_o, \delta_d, q_0)$ be a diagnoser. The simplified diagnoser associated to G_d is $G'_d = (Q'_d, E_o, \delta'_d, q_0)$ where $Q'_d = Q_d \setminus B$ where $B = \{q \in Q_d : q \notin \mathcal{C}$ and q cannot be accessible from q_0 without passing by a state in $\mathcal{C}\}$ and δ'_d is the restriction of δ_d to Q'_d.

G'_d is obtained from G_d by: (1) removing all the transitions originating from any state $q \in \mathcal{C}$ then, (2) keeping only the part of the diagnoser that is accessible from the initial state q_0. Note that a state which is reachable both by passing by a state in $\mathcal{C}$ and without passing by such a state is kept, but only the paths from q_0 to that state that do not contain any state from $\mathcal{C}$ are kept. Note also that critical states no longer have outgoing transitions. Figure 3-(b) depicts the simplified diagnoser of the system of example 1.

The light estimator represents explicitly the transitions between the sub-states of the simplified diagnoser as well as their probability values. It proposes also an adequate treatment of the sub-states of the critical states:

Definition 5. The L-estimator is defined by $H = (T, E'_o, \psi, t_0)$ where:

- $E'_o = E_o \cup \{\alpha\}$ is the set of observable events, such that $\alpha \notin E_o$ is a new event standing for any observable event. This event is added to ensure a coherent definition of the L-estimator and will not play any role in the following development.
- Let $q_0, \ldots, q_m$ be the states of the simplified diagnoser G_d such that $q_0 = \{(x_0, N)\}$. The set T of the states of H is included in $X_o \times \{N, F\} \times \{0, \ldots, m\}$.
- $\psi : T \times E'_o \times T \longrightarrow [0..1]$ is the probabilistic transition function of H. Let $t = (x, l, i)$ be a state of T such that $q_i \notin \mathcal{C}$ and $t' = (x', l', i')$ be a state of T and $\sigma \in E_o$. The transition probability $\psi(t, \sigma, t')$ is different from 0 if only if there is a possible transition from t to t'. From the construction of the diagnoser G_d, this corresponds to the case where there is at least some trace $s \in L_\sigma(G, x)$ such that $l' = LP(x, l, s)$ and $(x, s, x') \in \delta$. Let S be the set of all such traces: $S = \{s \in L_\sigma(G, x) | l' = LP(x, l, s)$ and $(x, s, x') \in \delta\}$. The transition probability $\psi(t, \sigma, t')$ is then the sum of the probabilities of transitions from x to x' by the different traces of S: $\psi(t, \sigma, t') = \sum_{s \in S} \theta(x, s, x')$. For each state $t = (x, l, i)$ of T such that $q_i \in \mathcal{C}$, we put $\psi(t, \alpha, t) = 1$.
- $t_0 = (x_0, N, 0)$ is the initial state.

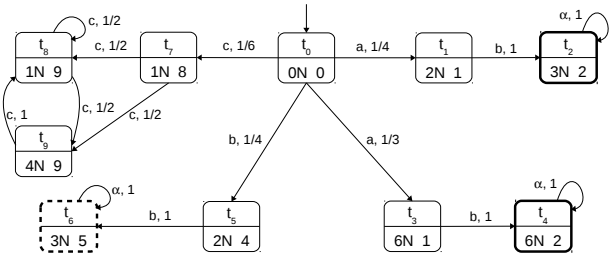

Figure 4: The light estimator

Intuitively, a state $t = (x, l, i)$ of the light estimator contains the relevant information about predictability. A trace w of observable events leading from t_0 to $t = (x, l, i)$ in the L-estimator leads in the diagnoser from q_0 to q_i such that $(x, l) \in q_i$. Moreover, in case $q_i \in \mathcal{C}_{KO}$ (resp. $q_i \in \mathcal{C}_{OK}$), the fault may or may not occur (resp. will necessarily occur) before observing the next observable event after w and in any observed trace having w as prefix, the fault cannot be predicted (resp. can be predicted) before its occurrence. Thus, as soon as a state $t = (x, l, i)$ such that $q_i \in \mathcal{C}$ is reached, the decision about predictability can be taken independently from the possible subsequent continuations. This explains the addition of the loop on each such state with the probability 1. Figure 4 shows the estimator of the system given in example 1. We have that the sum of the probabilities of all transitions issued from each state of H is 1:

Proposition 2. $\forall t \in T, \sum_{\sigma \in E'_o} \sum_{t' \in T} \psi(t, \sigma, t') = 1$.

Proof. Let $t = (x, l, i)$ be a state of the L-estimator. Two cases are possible :

- For $q_i \in \mathcal{C}$, only one transition from t to t is defined and labelled by α such that $\psi(t, \alpha, t) = 1$. The proposition holds for such states.
- For $q_i \notin \mathcal{C}$, by the construction of the L-estimator, every transition from $t = (x, l, i)$ to a state $t' = (x', l', i')$ in the estimator labelled by an event σ corresponds to one and only one transition in the generator G' from state x to state x' labelled by

the same observable σ. Moreover $\psi(t,\sigma,t') = \theta_{\Gamma'}(x,\sigma,x')$. Thus $\sum_{\sigma\in E'_o}\sum_{t'\in T}\psi(t,\sigma,t') = \sum_{\sigma\in E_o}\sum_{x'\in X_0}\theta_{\Gamma'}(x,\sigma,x') = 1$.

□

4.2 Deep estimator

The deep estimator (D-estimator) is constructed from the whole diagnoser and thus is similar to that proposed in [Nouioua and Dague, 2008] for diagnosability analysis. However, in order to analyze also predictability, a D-estimator state contains an additional label which propagates relevant information about predictability.

Definition 6. The D-estimator is defined by: $\Delta = (Z, E_o, \varphi, z_0)$ where:

- E_o is the set of observable events.
- Let $q_0,\ldots,q_k$ be the states of the diagnoser G_d such that $q_0 = \{(x_0, N)\}$. The set Z of the states of Δ is included in $X_o \times \{N,F\} \times \{0,\ldots,k\} \times \{NA, A\} \times \{OK, KO\}$ where the label NA (resp. A) stands for non ambiguous (resp. ambiguous) and the label OK (resp. KO) stands for being predictable (resp. not predictable).
- The initial state of Δ is defined by:
$$z_0 = \begin{cases} (x_0, N, 0, NA, KO) & if \quad q_0 \in \mathcal{C}_{KO} \\ (x_0, N, 0, NA, OK) & otherwise \end{cases}$$
- $\varphi : Z \times E_o \times Z \longrightarrow [0..1]$ is the probabilistic transition function of Δ. Let $z = (x, l, i, Att_D, Att_P)$ and $z' = (x', l', i', Att'_D, Att'_P)$ be two states of Z and σ be an observable event. The transition probability $\varphi(z,\sigma,z')$ is different from 0 if only if there is a possible transition from z to z'. From the construction of the diagnoser G_d, this corresponds to the case where there is at least some trace $s \in L_\sigma(G,x)$ such that $l' = LP(x,l,s)$ and $(x,s,x') \in \delta$. Let S be the set of all such traces: $S = \{s \in L_\sigma(G,x) | l' = LP(x,l,s) \ and \ (x,s,x') \in \delta\}$. The transition probability $\varphi(z,\sigma,z')$ is then the sum of the probabilities of transitions from x to x' by the different traces of S: $\varphi(z,\sigma,z') = \sum_{s\in S}\theta(x,s,x')$. The label Att'_P is given by: $Att'_P = \begin{cases} KO & if \quad q_{i'} \in \mathcal{C}_{KO} \ (1^{st} \ occurrence \ of \ KO) \\ KO & if \quad Att_P = KO \ (propagation \ of \ KO) \\ OK & otherwise \end{cases}$
The label Att_D depends only on the corresponding diagnoser state. For a state $z = (x, l, i, Att_D, Att_P)$ of Z, we have: $Att_D = \begin{cases} A & if \quad q_i \ is \ an \ f-uncertain \ state \\ NA & otherwise \end{cases}$

Intuitively, a state $z = (x, l, i, Att_D, Att_P)$ of the deep estimator contains all the relevant information about both diagnosability and predictability if the system follows a trace whose projection is the observable trace leading from z_0 to z. Namely, the trace w leading from z_0 to z in the deep estimator leads in the diagnoser from q_0 to the state q_i such that $\{(x,l)\} \in q_i$. Moreover, if q_i is f-uncertain then $Att_D = A$, otherwise $Att_D = NA$ and if $q_i \in \mathcal{C}_{KO}$ or q_i is reached from q_0 by following a sequence of states containing a state $q \in \mathcal{C}_{KO}$ then $Att_P = KO$ otherwise $Att_P = OK$. Figure 5 shows the estimator of the system

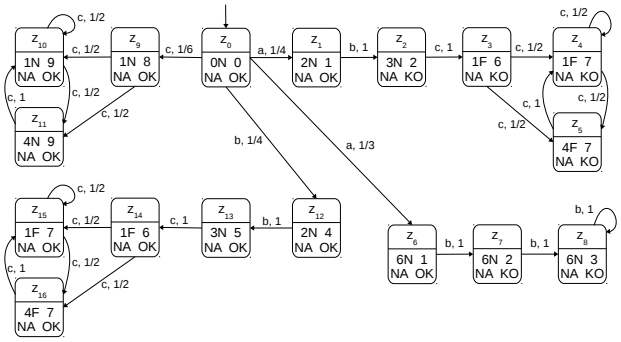

Figure 5: The deep estimator

given in example 1[2]. We have the following result :

Proposition 3. $\forall z \in Z, \sum_{\sigma\in E_o}\sum_{z'\in Z}\varphi(z,\sigma,z') = 1$.

Proof. The proof is similar to the second case in the proof of the previous proposition. □

5 Probabilistic Analysis

In this section, we show how to extract from (the light and the deep) estimator an homogeneous and discrete Markov chain and then to exploit the well known results about the asymptotic behaviors of such chains (for more details about that, see for example [Kemeny and Snell, 1983]) to obtain a finer appreciation of predictability. We believe that such a refinement may be very useful in practice to deal with non-predictable systems.

5.1 Markov chains associated with the estimators

To the light estimator $H = (T, E'_o, \psi, t_0)$ (resp. the deep estimator $\Delta = (Z, E_o, \varphi, z_0)$), we associate the homogeneous and discrete time Markov chain $\{M_i, i = 0, 1...|T|-1\}$ (resp. $\{K_i, i = 0, 1...|Z|-1\}$) where M_i (resp. K_i) is a random variable whose value is the state of the system after the observation of a set of events. T (resp. Z) is the state space of the Markov chain.

The transition matrix tr^L of the L-estimator is defined by: $\forall (t_1,t_2) \in T \times T, tr^L_{t_1,t_2} = \sum_{\sigma\in E'_o}\psi(t_1,\sigma,t_2)$. The transition matrix tr^D of the D-estimator is defined by: $\forall (z_1,z_2) \in Z \times Z, \quad tr^D_{z_1,z_2} = \sum_{\sigma\in E_o}\varphi(z_1,\sigma,z_2)$.

Example 1 (cont). Since from each couple of states (t,t') (resp. (z,z')) of the L-estimator (resp. the D-estimator) there is at most one transition from t to t' (resp. from z to z') with a probability different from 0, the graphical representation of the Markov chain M_i (resp. K_i) is obtained from figure 4 (resp. figure 5) by just removing the observable events. Here is the transition matrix tr^L of the L-estimator of our system:

[2]Note that any f-uncertain cycle in the diagnoser that is not f-indeterminate, i.e., does not correspond to a cycle in the model, disappears in the D-estimator.

$$
\begin{array}{c} \\ t_0 \\ t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \\ t_8 \\ t_9 \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} t_0 & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 & t_8 & t_9 \end{array} \\
\begin{pmatrix}
0 & \frac{1}{4} & 0 & \frac{1}{3} & 0 & \frac{1}{4} & 0 & \frac{1}{6} & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

A Markov chain is reducible if its representative graph contains more than one strongly connected component. This is the case for the Markov chains associated with the D-estimator and the L-estimator except in one particular case for the latter (see proposition 4. below).

Proposition 4. Under the assumption that there is at least one occurrence of a fault in the system: (1) The Markov chain $\{K_i\}$ associated with the D-estimator Δ is reducible. (2) If the L-estimator H contains at least two states, the Markov chain $\{M_i\}$ associated with the L-estimator H is reducible.

Proof. (1) Suppose that there is at least one fault occurrence in the system, the D-estimator must contain at least one state whose fault label is F and it is clear that from such a state we can never come back to the initial state z_0 whose fault label is N.

(2) Suppose that there is at least one fault occurrence in the system. If the L-estimator contains only one state, then this state must be the initial state having a self loop with a probability that equals 1. This case occurs when the initial state of the simplified diagnoser is a critical state ($q_0 \in \mathcal{C}$). If the L-estimator contains more than one state, then it must contain at least a state t coming from a diagnoser state $q \in \mathcal{C}$. Since the only transition originating from t is a loop on t itself, there is no path from t to t_0. □

5.2 The asymptotic behavior

From the study of the asymptotic behavior of the Markov chain associated to the estimator, we can compute relevant probability measures concerning classes of possible infinite observed traces of the system and estimate the average number of steps (observables) after which the system converges to a stage where it is or not predictable/diagnosable. The next two sections give the details of the probability values that one can extract from the L-estimator and the D-estimator. Here, we present briefly the steps to follow in order to study the asymptotic behavior of a reducible Markov chain $\{M_i\}$.

1. Classify the states of the chain $\{M_i\}$. A class is simply a strongly connected component in the representative graph of $\{M_i\}$; a persistent class is a class where states have no successor outside it; an absorbent class is a persistent class that contains only one state. A non persistent is called transitory. Let $\zeta = \{C_1, \ldots, C_h\}$ be the set of the persistent classes of $\{M_i\}$ and $\mu = \{\mu_1, \ldots, \mu_r\}$ be the set of states belonging to transitory classes.
2. Put the transition matrix in the canonical form in which persistent classes are put at the beginning and the states of each persistent class are put together. We obtain:

$$
tr = \begin{pmatrix}
Tr_1 & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & Tr_h & 0 \\
R_1 & \cdots & R_h & Q
\end{pmatrix}
$$

   Tr_i is the stochastic matrix containing the transition probabilities inside the persistent class C_i. $R = [R_1, \ldots, R_h]$ (resp. Q) contains the transition probabilities from transitory states to persistent states (resp. to transitory states).
3. Compute the fundamental matrix given by: $N = (I - Q)^{-1}$ (I is the unit matrix of size r) and the absorption matrix given by $B = N.R$. We have the following results: the probability to be in a transitory state after an infinite number of steps is 0; the average number of steps (observed events) before absorption in a persistent class starting from the transitory state i is the sum of the terms of the i^{th} row of N and the probability of absorption in the persistent state j when we start from state i is given by the term $B_{i,j}$ of B. The absorption probability of a persistent class is then the sum of the absorption probabilities of its states.

In our context, we suppose without loss of generality that the initial state t_0 (resp. z_0) is the first transitory state[3]. Thus, since our starting point is always the initial state t_0 (resp. z_0), we are interested in the next sections only on the first rows of N and B respectively.

5.3 Probability values from the L-estimator

Let $\Gamma = (X, E, \theta, x_0)$ a PDES and $H = (T, E'_o, \psi, t_0)$ be its L-estimator. Let M_i be the associated Markov chain and let N^L and B^L its fundamental and absorption matrices respectively. Let $\mathcal{T}_{ko}$, (resp. $\mathcal{T}_{ok}$) be the subset of persistent classes whose states correspond to critical states of the diagnoser where the fault is not predictable (resp. where the fault is predictable), i.e. states $t = (x, l, i)$ where $q_i \in \mathcal{C}_{KO}$ (resp. where $q_i \in \mathcal{C}_{OK}$). Let $\mathcal{T}_{nf}$ be the subset of all the other persistent classes i.e. where the fault does not occur. Then, we can define the relevant probabilities $\mathcal{P}_{ko}$, $\mathcal{P}_{ok}$ and $\mathcal{P}_{nf}$ as follows:

- $\mathcal{P}_{ko}$ is the probability to follow a trace where the fault cannot be predicted, $\mathcal{P}_{ko} = \sum_{c \in \mathcal{T}_{ko}} \sum_{t \in c} (B^L)_{0,t}$[4].
- $\mathcal{P}_{ok}$ is the probability to follow a trace where the fault surely occurs and is predicted, $\mathcal{P}_{ok} = \sum_{c \in \mathcal{T}_{ok}} \sum_{t \in c} (B^L)_{0,t}$.
- p_{nf} is the probability to follow a trace where the fault surely does not occur. $\mathcal{P}_{nf} = \sum_{c \in \mathcal{T}_{nf}} \sum_{t \in c} (B^L)_{0,t} = 1 - (\mathcal{P}_{ko} + \mathcal{P}_{ok})$.

[3] z_0 is always transitory. See the explanation given in the previous footnotes. There is only one (trivial) case where t_0 is not transitory. In this case t_0 is the only state of the Markov chain.

[4] Note that a fault may be predictable in a trace but not predictable in the system. Predictability in the system is achieved when the fault is predictable in all the traces containing it. In our example the fault is predictable (resp. not predictable) in any observed trace starting by b (resp. by a).

Table 1: Relevant probability values from the L-estimator

Pr	$\mathcal{P}_{ko}$	$\mathcal{P}_{ok}$	$\mathcal{P}_{NF}$
Val	$7/12$	$1/4$	$1/6$

- In addition to these probability values, we can obtain the average number of steps before absorption starting from state t_0: $\overline{Nb}^L_{Abs} = \sum_{j=0}^{r-1}(N^L)_{0,j}$. This number represents the average number of events that must be observed to either predict that the fault will occur or that it will not occur or to decide that the fault is not predictable.

Example 1 (cont). In our example (see the transition matrix in section 4.3.1. and the representative graph in figure 4), we have four persistent classes: $C_1 = \{t_2\}$, $C_2 = \{t_4\}$, $C_3 = \{t_6\}$ and $C_4 = \{t_8, t_9\}$ where $\mathcal{T}_{ko} = \{C_1, C_2\}$, $\mathcal{T}_{ok} = \{C_3\}$ and $\mathcal{T}_{nf} = \{C_4\}$. Each remaining state forms alone a transitory class. After putting the transition matrix in the canonical form we compute the matrices N^L and B^L. The first rows corresponding to t_0 of these matrices are:

$$(N^L)_0 = \begin{array}{c} t_0 \quad t_1 \quad t_3 \quad t_5 \quad t_7 \\ \left(\begin{array}{ccccc} 1 & \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & \frac{1}{6} \end{array}\right) \end{array} \quad \textit{and}$$

$$(B^L)_0 = \begin{array}{c} t_2 \quad t_4 \quad t_6 \quad t_8 \quad t_9 \\ \left(\begin{array}{ccccc} \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & \frac{1}{12} & \frac{1}{12} \end{array}\right) \end{array}$$

From $(N^L)_0$ we have: $\overline{Nb}^L_{Abs} = 1 + 1/4 + 1/3 + 1/4 + 1/6 = 2\ steps$. Table 1 sums up the relevant probabilities for our example:

This means that, in average, after the observation of 2 events, we have a probability of $7/12$ (resp. $1/4$) to be in a trace where the fault may occur or not and it cannot be predicted (resp. the fault will occur and it is predicted) and a probability of $1/6$ to be in a trace where the fault will not occur.

5.4 Probability values from the D-estimator

Let $\Gamma = (X, E, \theta, x_0)$ a PDES and $\Delta = (Z, E_o, \varphi, z_0)$ be its D-estimator. Let K_i be the associated Markov chain and let N^D and B^D its fundamental and absorption matrices respectively. Note that in a persistent class, all states share the same value for the attribute Att_D. The same applies to the attributes Att_P and l. We define the following subsets of persistent classes:

- ξ_A (resp. ξ_{NA}) is the subset of persistent classes containing only ambiguous states (resp. non ambiguous states), i.e. states $z = (x, l, i, Att_D, Att_P)$ where $Att_D = A$ (resp. $Att_D = NA$)
- ξ_F (resp. ξ_N) is the subset of persistent classes containing only faulty states (resp. normal states), i.e., states $z = (x, l, i, Att_D, Att_P)$ where $l = F$ (resp. $l = N$)
- ξ_{ko} (resp. ξ_{ok}) is the subset of persistent classes containing only states of the form: $z = (x, l, i, Att_D, Att_P)$ where $Att_P = KO$ (resp. $Att_P = OK$)
- $\xi_{F\wedge A}$ is the subset of persistent classes containing only states that are faulty and ambiguous: $\xi_{F\wedge A} = \xi_F \cap \xi_A$.
- $\xi_{F\wedge ko}$ is the subset of persistent classes containing only states of the form: $z = (x, l, i, Att_D, Att_P)$ such that $l = F$ and $Att_P = KO$: $\xi_{F\wedge ko} = \xi_F \cap \xi_{ko}$.
- $\xi_{NA\wedge ko}$ is the subset of persistent classes containing only states of the form: $z = (x, l, i, Att_D, Att_P)$ such that $Att_D = NA$ and $Att_P = KO$: $\xi_{NA\wedge ko} = \xi_{NA} \cap \xi_{ko}$.

We define different probability values that give a good evaluation of the situation wrt both diagnosability and predictability:

- $\mathcal{P}_A$ (resp. $\mathcal{P}_{NA}$) is the probability to follow a trace where the fault is not diagnosable (resp. is diagnosable). $\mathcal{P}_A = \sum_{c\in\xi_A}\sum_{z\in c}(B^D)_{0,z}$ and $\mathcal{P}_{NA} = 1 - \mathcal{P}_A$.
- $\mathcal{P}_F$ (resp. $\mathcal{P}_N$) is the probability to follow a trace where the fault occurs (resp. does not occur). $\mathcal{P}_F = \sum_{c\in\xi_F}\sum_{z\in c}(B^D)_{0,z}$ and $\mathcal{P}_N = 1 - \mathcal{P}_F$.
- $\mathcal{P}_{ko}$ (resp. $\mathcal{P}_{ok}$) is the probability to follow a trace where the fault is not predictable (resp. is predictable). $\mathcal{P}_{ko} = \sum_{c\in\xi_{ko}}\sum_{z\in c}(B^D)_{0,z}$ and $\mathcal{P}_{ok} = 1 - \mathcal{P}_{ko}$.
- $\mathcal{P}_{F/A}$ is the probability to follow a trace where the fault occurs known that it is not diagnosable. Using Bayes formula we obtain: $\mathcal{P}_{F/A} = \frac{\sum_{c\in\xi_{F\wedge A}}\sum_{z\in c}(B^D)_{0,z}}{\sum_{c\in\xi_A}\sum_{z\in c}(B^D)_{0,z}}$.
- $\mathcal{P}_{F/ko}$ is the probability to follow a trace where the fault occurs known that it is not predictable. Using Bayes formula we obtain: $\mathcal{P}_{F/ko} = \frac{\sum_{c\in\xi_{F\wedge ko}}\sum_{z\in c}(B^D)_{0,z}}{\sum_{c\in\xi_{ko}}\sum_{z\in c}(B^D)_{0,z}}$.
- $\mathcal{P}_{ko/NA}$ is the probability to follow a trace where the fault is not predictable known that it is diagnosable. Using Bayes formula, we obtain the following formula: $\mathcal{P}_{ko/NA} = \frac{\sum_{c\in\xi_{NA\wedge ko}}\sum_{z\in c}(B^D)_{0,z}}{\sum_{c\in\xi_{NA}}\sum_{z\in c}(B^D)_{0,z}}$.
- The average number of steps before absorption starting from state z_0 is: $\overline{Nb}^D_{Abs} = \sum_{j=0}^{r-1}(N^D)_{0,j}$. Note that this value is relevant only for diagnosability, since for predictability the decision is made as soon as a critical state is reached. The good value of this parameter for predictability is $\overline{Nb}^L_{Abs}$ given by the L-estimator.

The fact that predictability is stronger than diagnosability is captured here by the fact that the set of persistent classes with the label OK (resp. KO) is a subset (resp. superset) of that of persistent classes with the label NA (resp. A). Thus, the probability that a fault is (resp. is not) predictable is smaller (resp. greater) than the probability that this fault is (resp. is not) diagnosable.

Proposition 5. We have: $\mathcal{P}_{ok} \leq \mathcal{P}_{NA}$ and equivalently $\mathcal{P}_{ko} \geq \mathcal{P}_A$.

Proof. From the construction of the D-estimator, it is easy to see that if a state has the label A then necessarily it has the label KO because it belongs necessarily to a diagnoser state accessible from $\mathcal{C}_{KO}$. □

Table 2: Relevant probability values for example 1.

Pr	$\mathcal{P}_A$	$\mathcal{P}_F$	$\mathcal{P}_{ko}$	$\mathcal{P}_{F/A}$	$\mathcal{P}_{F/ko}$	$\mathcal{P}_{ko/NA}$
Val	0	$1/2$	$7/12$	undef	$3/7$	$7/12$

Example 1 (cont). Let us now come back to our example. We have three persistent classes: $C_1 = \{z_4, z_5\}$, $C_2 = \{z_8\}$, $C_3 = \{z_{15}, z_{16}\}$ and $C_4 = \{z_{10}, z_{11}\}$. Each remaining state forms alone a transitory class. The first rows (corresponding to z_0) of the matrices N^D and B^D are:

$$(N^D)_0 = \begin{array}{cccccccccc} z_0 & z_1 & z_2 & z_3 & z_6 & z_7 & z_9 & z_{12} & z_{13} & z_{14} \\ \big(1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \big); \end{array}$$

$$(B^D)_0 = \begin{array}{ccccccc} z_4 & z_5 & z_8 & z_{15} & z_{16} & z_{10} & z_{11} \\ \big(\frac{1}{8} & \frac{1}{8} & \frac{1}{3} & \frac{1}{8} & \frac{1}{8} & \frac{1}{12} & \frac{1}{12} \big) \end{array}$$

We have: $\overline{Nb}^D_{Abs} = 3.33$ *steps* and the different probability values obtained from the D-estimator are given in table 2. Since the system is diagnosable we have $\mathcal{P}_A = 0$ (so $\mathcal{P}_{NA} = 1$) and $\mathcal{P}_{F/A}$ is not defined. Of course, for a non diagnosable system $\mathcal{P}_{F/A}$ is defined. The probability that the system executes a faulty trace is $\mathcal{P}_F = 1/2$ and the probability to follow a trace where the fault is not predictable is $7/12$. However, being in a trace where the fault is not predictable, the probability that this trace is faulty is $\mathcal{P}_{F/ko} = 3/7$. Thus, it is more probable that the system follows a non predictable trace but inside the non predictable traces, it is more probable that the fault will not occur. Because the system is diagnosable we have $\mathcal{P}_{ko/NA} = \mathcal{P}_{ko} = 7/12$. For a non diagnosable system, $\mathcal{P}_{ko/NA} \neq \mathcal{P}_{ko}$ and $\mathcal{P}_{ko/NA}$ is the probability that a diagnosable trace is not predictable.

6 Conclusion

This paper investigated the use of information about probabilities of transitions in a DES to refine the decision about fault predictability. In particular, the proposed approach allows one to quantify the degree of non-predictability and accordingly to deal in a more flexible way with non predictable systems. Moreover, a uniform framework to analyze both predictability and diagnosability is proposed and several relevant probability values are extracted to better appreciate these notions. These relevant probability values may be used differently according to the criticality of the application.

Despite the interest of this analysis, a limit of the present work is that it is based on the diagnoser approach which is costly because of its exponential complexity. Of course, it is quite natural to try to use twin plants instead of diagnosers. However, it turns out that the structure of the diagnoser is much more adequate to that of the twin plant to achieve our objective. Indeed, the analysis we wanted to do aims at separating different classes of traces (three classes of traces for predictability: the class of traces where we can decide that the fault will appear, the class of traces where we can decide that the fault will not appear and the class of traces where we cannot take a decision before the occurrence of the fault. More classes are considered in the joint analysis of predictability and diagnosability). It turns out that the diagnoser gives directly this separation: the class of each continuation of a sub-path arriving at a critical state is determined by the nature of this state. However, this is not the case for the twin plant: a sub-path arriving at a critical state of the twin plant may be the prefix of a path where the fault is predictable. Accordingly, the Markov chain, which is the basis of our probabilistic analysis, follows directly from the L-estimator and the D-estimator constructed easily from the diagnoser. It is not so easy to extract such a Markov chain from the twin plant.

Not being easy does not mean necessarily to be impossible. The polynomial complexity of the twin plant approach allows one to apply it on a larger class of complex systems and this is a strong motivation for us to continue the efforts in order to adapt it to the kind of probabilistic analysis performed in the present work. This issue is left for a future work. The present work has opened the way for several other perspectives including the generalization of the probabilistic-based approach to pattern predictability, to the case of distributed discrete event systems and to other discrete event models such as Petri nets.

References

[Cassez and Grastien, 2013] F. Cassez and A. Grastien. Predictability of event occurrences in timed systems. In *International Workshop on Formal Modeling and Analysis of Timed Systems*, 2013.

[Cimatti *et al.*, 2003] A. Cimatti, C. Pecheur, and R. Cavada. Formal verification of diagnosability via symbolic model checking. In *International Joint Conference on Artificial Intelligence*, pages 363–369, 2003.

[Console *et al.*, 2000] L. Console, C. Picardi, and M. Ribaudo. Diagnosis and diagnosability analysis using pepa. In *European Conference on Artificial Intelligence*, pages 131–136, 2000.

[Genc and Lafortune, 2009] S. Genc and S. Lafortune. Predictability of event occurrences in partially-observed discrete-event systems. *Automatica*, 45(2):301–311, 2009.

[Jeron *et al.*, 2008] T. Jeron, H. Marchand, S. Genc, and S. Lafortune. Predictability of sequence patterns in discrete event systems. In *IFAC World Congress*, pages 537–543, 2008.

[Jiang *et al.*, 2001] S. Jiang, Z. Huang, V. Chandra, and R. Kumar. A polynomial algorithm for testing diagnosability of discrete event systems. *IEEE Trans. On Aut. Cont.*, 46(8):1318–1321, 2001.

[Kemeny and Snell, 1983] J.G. Kemeny and J.L. Snell. *Finite Markov Chains*. Springer-Verlag, 1983.

[Nouioua and Dague, 2008] F. Nouioua and P. Dague. A probabilistic analysis of diagnosability in discrete event systems. In *European Conference on Artificial Intelligent*, pages 224–228, 2008.

[Rintanen and Grastien, 2007] J. Rintanen and A. Grastien. Diagnosability testing with satisfiability algorithms. In *International Joint Conference on Artificial Intelligence*, pages 532–537, 2007.

[Sampath *et al.*, 1995] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. Diagnosability of discrete-event systems. *IEEE Trans. On Aut. Cont.*, 40(9):1555–1575, 1995.

[Schumann and Pencolé, 2007] A. Schumann and Y. Pencolé. Scalable diagnosability checking of event-driven systems. In *International Joint Conference on Artificial Intelligence*, pages 575–580, 2007.

[Thorsley and Teneketzis, 2005] D. Thorsley and D. Teneketzis. Diagnosability of stochastic discrete-event systems. *IEEE Trans. on Aut. Cont.*, 50(4):476–492, 2005.

[Ye and Dague, 2012] L. Ye and P. Dague. A general algorithm for pattern diagnosability of distributed discrete event systems. In *IEEE International Conference on Tools with Artificial Intelligence*, pages 130–137, 2012.

[Ye *et al.*, 2013] L. Ye, P. Dague, and F. Nouioua. Predictability analysis of distributed discrete event systems. In *IEEE Conference on Decision and Control*, pages 5009–5015, 2013.

[Yoo and Lafortune, 2002] T. Yoo and S. Lafortune. Polynomial-time verification of diagnosability of partially-observed discrete-event systems. *IEEE Trans. On Aut. Cont.*, 47(9):1491–1495, 2002.