Types of Asynchronous Diagnosability and the *Reveals*-Relation in Occurrence Nets

Stefan Haar, *Member, IEEE*
INRIA Saclay - LSV, Ecole Normale Supérieure de Cachan
61, ave. du Président Wilson, 94235 CACHAN Cedex, France

***Abstract*—We consider asynchronous diagnosis in (safe) Petri net models of distributed systems, using the partial order semantics of occurrence net unfoldings. Both the observability and diagnosability properties will appear in two different forms, depending on the semantics chosen: *strong* observability and diagnosability are the classical notions from the state machine model and correspond to interleaving semantics in Petri nets. By contrast, the *weak* form is linked to characteristics of nonsequential processes, and requires an asynchronous *progress* assumption on those processes. We give algebraic characterizations for both types, and give verification methods. The study of weak diagnosability leads us to the analysis of a relation in occurrence nets, first presented in [15]: given the occurrence of some event a that *reveals* b, the occurrence of b is inevitable. Then b may already have occurred, be concurrent to, or even in the future of a. We show that the *reveals*-relation can be effectively computed recursively - for each pair, a suitable finite prefix of bounded depth is sufficient -, and show its use in asynchronous diagnosis. Based on this relation, a decomposition of the Petri net unfolding into *facets* is defined, yielding an abstraction technique that preserves and reflects maximal partially ordered runs.**

***Index Terms*—Petri Nets; Fault detection; Discrete event systems**

I. Introduction

In highly distributed networked systems, events occur in an asynchronous way; moreover, the supervisor needs to receive alarms from sensors that are generally at a non-negligible distance. Due to asynchronicity between the system and its supervision, alarms collected at different distant sensors can not be meaningfully given a temporal precedence. In particular, it is appropriate to abandon the usual *interleaving* semantics which describes system behaviour by sequences of events. In fact, we will follow the approach of [8], [9] in which

- the system is modeled as a (safe) Petri net, thus taking into account the local and asynchronous nature of states and transitions, and
- the semantics on which diagnosis operates is that of *partially ordered executions* as obtained through the partial order *unfolding* of Petri nets.

Petri nets (see e.g. [18], [26], [29]) and their partial order unfoldings [6], [20], [25] have been increasingly used in recent years for both fault diagnosis [8], [9], [14] and control (see e.g. [13]) of asynchronous discrete event systems. The advantage of partial order semantics lies in the space reduction for representing nonsequential processes that have a high degree of parallelism. In unfoldings, sets of concurrent events are not ordered, which means they have to be represented only once (by one partial order) rather than by giving all their interleavings whose number is exponential in the size of the concurrent set. See also [9] and the discussion in [7].

The purpose of the present article is to investigate *diagnosability* for Petri net models under the partial order perspective. Not surprisingly, the work of Sampath et al.'s [30] classical characterization of diagnosability in languages of words obtained as *sequential* runs of *automata* will carry over - partly- to the asynchronous setting, where the languages are formed by *nonsequential* runs of *Petri nets*. However, important differences will become apparent between the situations in interleaving semantics on the one hand and in partial order semantics on the other. Our analysis will lead us to distinguish *weak* and *strong* versions of both observability and diagnosability. In short, strongly diagnosable systems allow fault diagnosis under any policy of execution, even those in which some subprocesses may move on quickly while others halt; for weak diagnosability, diagnosis needs only be successful in executions in which all parts progress in a balanced way.

We will also consider methods for verification of weak diagnosability; this requires to account for phenomena that are intrinsic to *concurrency* in system behaviour. It motivates a deeper analysis of the structure of occurrence nets, leading to the *reveals* relation $\triangleright$ which we first pointed out (under the name of *covering* relation) in [15]. It connects pairs (a, b) of events such that a *reveals* b in the sense that whenever a occurs, b must have occurred or will eventually occur as well (while a may not be necessary for occurrence of b). We will define the relation $\triangleright$, prove its key properties, and show that it can be effectively computed off-line on a bounded prefix of the model unfolding.

Once the $\triangleright$-relation is known, it can be used, e.g., to detect and identify invisible fault events: the observation of a allows to deduce that any b revealed by a either has already occurred, or will inevitably eventually occur (possibly in the future of a, or in parallel). This fact allows to generalize both (a posteriori) diagnosis and prediction.

A further application of the *reveals* relation is in a possible reduction of the size of occurrence net representations by suitable abstractions. *Facets* are subnets of the unfolding in

Work supported French national telecommunication research network (RNRT) project SWAN, decision no. 03 S 481, and by the INRIA Sabbatical program while the author was with School of Information Technology and Engineering, University of Ottawa, Canada. Email: Stefan.Haar@inria.fr or haar@lsv.ens-cachan.fr

which *any* two events reveal one another. As a consequence, if *any* event in a facet ψ occurs, eventually *all* other events of ψ have to occur. Facets enjoy some nice structural properties; their study opens the way to a new topic of *qualitative diagnosability* which is the subject of future work.

The paper is organized as follows: Section II gives basic definitions; Section III recalls the asynchronous diagnosis methodology from [8], [9], [14], and defines weak and strong diagnosability concepts. The characterizations for the two properties are given in Section IV. Section V studies effective verification of diagnosability. The *reveals* relation is introduced and studied in Section VI; Section VII presents and analyzes abstractions into facets and associated diagnosability issues, and Section VIII concludes.

II. Definitions

Nets and homomorphisms: A *net* is a triple $N = (\mathcal{P}, T, F)$, where $\mathcal{P}$ and T are disjoint sets of *places* and *transitions*, respectively, and $F \subset (\mathcal{P} \times T) \cup (T \times \mathcal{P})$ is the *flow relation*.[1] In figures, places are represented by circles, rectangular boxes represent transitions, and arrows represent F. For node $x \in P \cup T$, call ${}^\bullet x \triangleq \{x' \mid F(x', x)\}$ the *preset*, and $x^\bullet \triangleq \{x' \mid F(x, x')\}$ the *postset* of x. Let $<$ be the transitive closure of F and $\leqslant$ the reflexive closure of $<$; further, let $\lceil x \rceil \triangleq \{x' \mid x' < x\}$ be the *prime configuration* or *cone* of x, and $\lfloor x \rfloor \triangleq \lceil x \rceil \backslash \{x\}$ the *pre-cone* of x.

A *net homomorphism* from N to N' is a map $\pi : P \cup T \longmapsto P' \cup T'$ such that *(i)* $\pi(P) \subseteq P'$, $\pi(T) \subseteq T'$, and *(ii)* $\pi_{|{}^\bullet t} : {}^\bullet t \to {}^\bullet \pi(t)$ and $\pi_{|t^\bullet} : t^\bullet \to \pi(t)^\bullet$ induce bijections, for every $t \in T$.

Homomorphisms between nets allow to formalize branching processes, see below.

Definition 1 *Two nodes x, x' of a net N are* in conflict, *written $x \# x'$, if there exist $t, t' \in T$ such that* (i) $t \neq t'$, (ii) ${}^\bullet t \cap {}^\bullet t' \neq \emptyset$, *and* (iii) $t \leqslant x$ *and* $t' \leqslant x'$. *A node x is said to be in* self-conflict *iff $x \# x$. An* occurrence net *(ON) is a net $ON = (B, E, F, \mathbf{c}_0)$, with the elements of B called* conditions *and those of E* events, *satisfying the additional properties :*

1) *no self-conflict:* $\forall x \in B \cup E : \neg(x \# x)$;
2) *$\leqslant$ is a partial order:* $\forall x \in B \cup E : \neg(x < x)$;
3) *Finite cones:* $\forall x \in B \cup E : |\lceil x \rceil| < \infty$;
4) *no backward branching:* $\forall b \in B : |{}^\bullet b| \leq 1$.
5) *the set $\mathbf{c}_0 \triangleq \min(ON)$ of $\leqslant$-minimal nodes of ON is contained in B.*

A prefix *of ON is any downward closed subset $\mathcal{R} \subseteq B \cup E$, i.e. such that for every $x \in \mathcal{R}$, $\lceil x \rceil \subseteq \mathcal{R}$; by abuse of notation, we will identify a prefix $\mathcal{R}$ with the* subnet *of ON spanned by the* set *$\mathcal{R}$. Prefix $\mathcal{C}$ is a* configuration *iff it is conflict-free, i.e. $x \in \mathcal{C}$ and $x \# y$ imply $y \notin \mathcal{C}$. Denote as $\mathbf{Con}(ON)$ the set of ON's configurations. Call any $\subseteq$-maximal element of $\mathbf{Con}(ON)$ a **run** of ON; the set of runs is denoted as $\Omega(ON)$ or simply Ω if no confusion can occur.*

The right hand side of Figure 1 shows an occurrence net. The leftmost branch, with events labeled β, γ, β is an example of a configuration.

[1] only *ordinary* nets are considered here, i.e. with arc weights 0 or 1.

Without loss of generality and for convenience, we have added property 5) in Definition 1; it is not required, e.g., in [5], . Note further that, as a consequence of property 3) in Def. 1, $B \cup E$ is well-ordered by $\leqslant$, i.e. there exist no infinite strictly decreasing sequences. Occurrence nets are useful to represent executions of Petri nets, see below: essential dynamical properties are visible via the topological structure of the acyclic graph. Nodes x and x' are *concurrent*, written $x \ \mathbf{co}\ x'$, if neither $x \leqslant x'$, nor $x' \leqslant x$, nor $x \# x'$ hold. A *co-set* is a set $\mathcal{X} \subseteq b$ of pairwise concurrent conditions. A maximal co-set $\mathcal{X}$ w.r.t. set inclusion is called a *cut*, and generically denoted by the symbol $\mathbf{c}$; in particular, $\mathbf{c}_0$ is a cut, called the *initial* cut of ON. - We note for future reference that occurrence nets are a special case of *event structures* [27]:

Definition 2 *A tuple $\mathcal{E} = (E, <, \#)$ is a **prime event structure** or **PES** iff:*

1) *$(E, <)$ is a countable, partially ordered set,*
2) *$\lceil e \rceil$ is finite for all $e \in E$,*
3) *$\# \subseteq E \times E$ is symmetric and irreflexive, and for all $x, y, z \in E$, $x \# y$ and $y < z$ together imply $x \# z$.*

Petri Nets: Let $N = (P, T, F)$ be a finite net. A *marking* of net N is a set[2] $M \subseteq P$. A *Petri net* (PN) is a pair $\mathcal{N} = (N, M_0)$, where $M_0 \subseteq P$ is an *initial* marking. $t \in T$ is *enabled* at M, written $M \xrightarrow{t}$, iff ${}^\bullet t \subseteq M$. If $M \xrightarrow{t}$, then t can *fire*, leading to $M' = (M \backslash {}^\bullet t) \cup t^\bullet$; write in that case $M \xrightarrow{t} M'$. The set $\mathbf{R}(M_0)$ contains the markings of $\mathcal{N}$ *reachable* through $\longrightarrow$. A Petri net $\mathcal{N} = (N, M_0)$ is *safe* if for all $M \in \mathbf{R}(M_0)$ and $t \in T$, $M \xrightarrow{t}$ implies $(t^\bullet \cap M) \subseteq {}^\bullet t$. Only safe nets are considered in this article. If $p \in M$, we will draw a black *token* in the circle representing p.

Example: In Figure 1, the left hand side shows a safe Petri net whose initial marking is $M_0 = \{1, 4, 7\}$. In M_0, the enabled transitions are α, β and η. The net represents a simple model of fault propagation between two components: component 1 consists of transitions α, β ,and γ, and places 1 and 2; component 2 of transitions δ, η, ζ, with places $4, 5, 6$. The places 2 and 7 serve as an interface linking both components. Initially, both components are in an *ok* state reflected by M_0. Then, if η fires, component 2 will remain permanently in a faulty state (reflected by place 6), regardless of the actions in component 1. On the side of component 1, occurrence fault β has no outside effect; the effect of β on component 1 can be repaired by occurrence of γ. Fault α, on the other hand, marks place 3 and thus enables the induced fault δ on the side of component 2, thus exhibiting propagation of a fault; in this model, that fault can be repaired on either component, through γ and ζ, respectively.

Branching Processes and Unfoldings: A *branching process* of the safe Petri net $\mathcal{N} = (P, T, F, M_0)$ is given by a pair $\Pi = (ON, \pi)$, where $ON = (B, E, G, \mathbf{c}_0)$, and π is a homomorphism from ON to N, such that:

1) π is injective on $\mathbf{c}_0$, and $\pi(\mathbf{c}_0) = M_0$;

[2] we will only consider safe nets here, the general definition allowing for *multiset* markings is therefore not necessary.

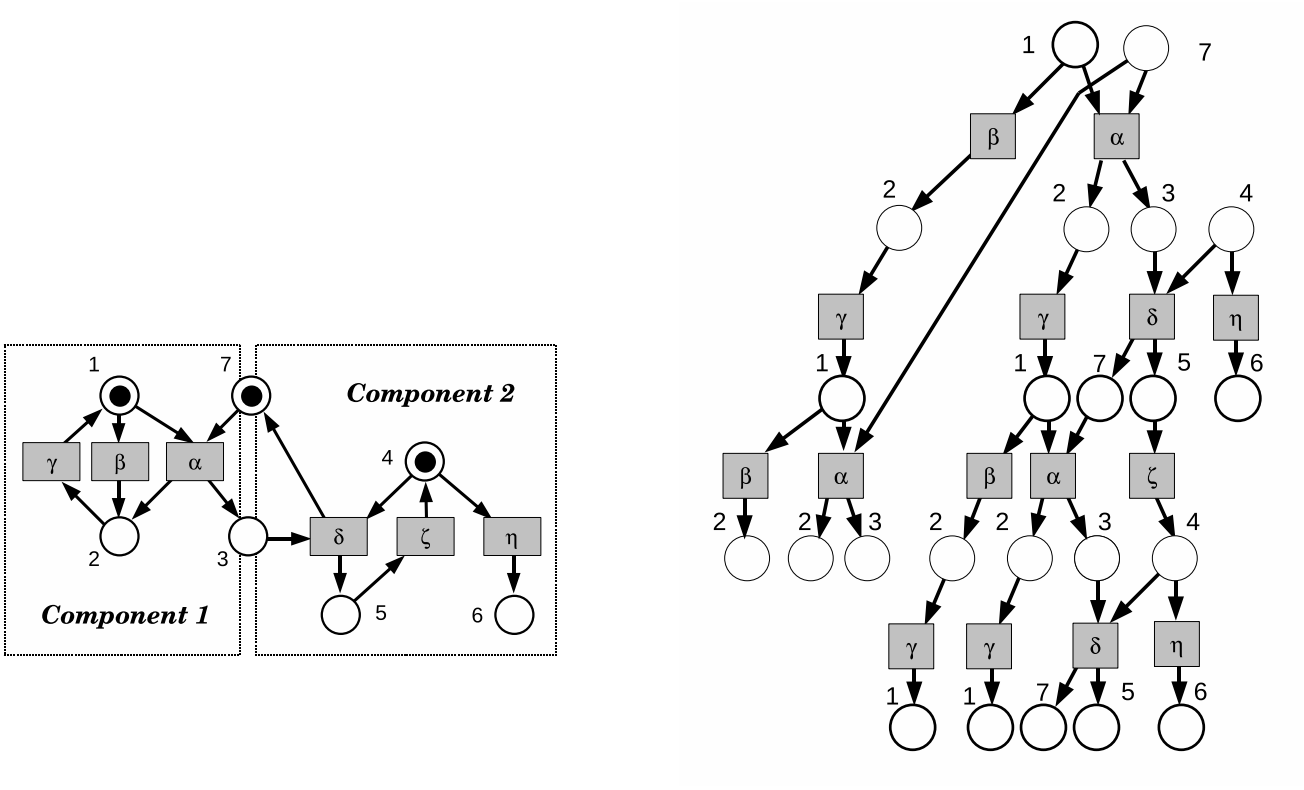


Fig. 1. A Petri net (left) and a prefix of its unfolding (right)

2) for all $e, e' \in E$, ${}^\bullet e = {}^\bullet e'$ and $\pi(e) = \pi(e')$ imply $e = e'$.

For Π_1, Π_2 two branching processes, Π_2 is a *prefix* of Π_1, written $\Pi_2 \sqsubseteq \Pi_1$, if there exists an injective homomorphism ψ from ON_2 into a prefix of ON_1, such that ψ induces a bijection between the initial cuts $\mathbf{c}_0^1$ and $\mathbf{c}_0^2$, and the composition $\pi_1 \circ \psi$ coincides with π_2.

By theorem 23 of [5], there exists a unique (up to an isomorphism) $\sqsubseteq$-maximal branching process, called the *unfolding* of $\mathcal{N}$ and denoted $\mathcal{U}(\mathcal{N})$; by abuse of notation, we will also use $\mathcal{U}(\mathcal{N})$ for the occurrence net obtained by the unfolding.

The unfolding of $\mathcal{N}$ can be computed using the following canonical algorithm by Esparza et al [6]. For any branching process $\Pi = (ON_\Pi, \pi_\Pi)$ of $\mathcal{N} = (P, T, F)$, with $ON_\Pi = (B_\Pi, E_\Pi, G_\Pi)$, denote as $pex(\Pi) \subseteq T \times \mathfrak{P}(B)$ the set of *possible extensions* of Π, i.e. of the pairs $(t, \mathcal{X})$ such that

- $\mathcal{X}$ is a co-set of ON_Π,
- ${}^\bullet t = \pi_\Pi(\mathcal{X})$,
- E_Π contains no event e such that $\pi_\Pi(e) = t$ and ${}^\bullet e = \mathcal{X}$.

Now, let $\mathbf{c}_0 \triangleq M_0 \times \{\emptyset\}$ and initialize $\Pi = (\mathbf{c}_0, \emptyset, \emptyset)$; recursively, for given $\Pi = (ON_\Pi, \pi_\Pi)$ with $ON_\Pi = (B_\Pi, E_\Pi, G_\Pi)$, compute $pex(ON_\Pi)$ and replace:

$$
\begin{array}{lll}
E_\Pi & by & E_\Pi \cup pex(ON_\Pi), \\
B_\Pi & by & B_\Pi \cup \{(P, e) \mid e \in pex(ON_\Pi), p \in \pi_\Pi(e)^\bullet\}, \ \textit{and} \\
G_\Pi & by & G_\Pi \cup \{(b, (t, \mathcal{X})) \mid (t, \mathcal{X}) \in pex(ON_\Pi), b \in \mathcal{X}\} \\
 & & \quad \cup \left\{(e, (P, e)) \mid e \in pex(ON_\Pi), p \in \pi_\Pi(e)^\bullet\right\}.
\end{array}
$$

We note the following technical properties:

Lemma 1 ***([5], [29])*** *If $\mathcal{U} = (B, E, G, \mathbf{c}_0, \pi)$ is the unfolding of safe Petri net $\mathcal{N} = (P, T, F, M_0)$, then:*

1) *If $\mathbf{c} \subseteq B$ is a cut then so is $\mathbf{c}' \triangleq (\mathbf{c} \backslash {}^\bullet e) \cup e^\bullet$ for every e such that ${}^\bullet e \subseteq \mathbf{c}$;*
2) *for $x, y \in B$, x **co** y implies $\pi(x) \neq \pi(y)$.*
3) *π maps all cuts of ON into $\mathcal{N}$-markings in $\mathbf{R}(M_0)$, and every marking in $\mathbf{R}(M_0)$ is the π-image of a cut of ON.*

Every *finite* configuration $\mathcal{C}$ terminates at a cut, which we denote $\mathbf{c}_\mathcal{C}$. The mapping $\mathcal{C} \mapsto \mathbf{c}_\mathcal{C}$ is bijective; for each cut $\mathbf{c}$, the union of the cones of all conditions in $\mathbf{c}$ yield the unique configuration $\mathcal{C}$ such that $\mathbf{c} = \mathcal{C}_\mathbf{c}$. Moreover, one has the following two correspondences:

• If $\mathcal{C}$ is a configuration of $\mathcal{U}_\mathcal{N}$ with $\mathcal{N} = (N, M_0)$, then every occurrence sequence σ obtained as a linear order extension, i.e. an *interleaving*, of the partial order $\leqslant_\mathcal{C}$ yields a firable transition sequence of $\mathcal{N}$. Conversely, every firable transition sequence of $\mathcal{N}$ corresponds to a linear order extension of some configuration of $\mathcal{U}_\mathcal{N}$. To sum up: the nonsequential executions of $\mathcal{N}$ are in one-to-one correspondence with the configurations of $\mathcal{U}(\mathcal{N})$. We will therefore speak of $\mathcal{N}$*'s configurations* and write $\mathbf{Con}(\mathcal{N}) \triangleq \mathbf{Con}(\mathcal{U}_\mathcal{N})$ and $\Omega(\mathcal{N}) \triangleq \Omega(\mathcal{U}_\mathcal{N})$.

• For every reachable marking $M \subseteq P$ of $\mathcal{N}$, there exists at least one cut $\mathbf{c}$ of $\mathcal{U}(\mathcal{N})$ such that $\pi(\mathbf{c}) = M$, and for the unique configuration $\mathcal{C}$ such that $\mathbf{c}_\mathcal{C} = \mathbf{c}$, execution of $\mathcal{C}$ takes M_0 to M; write $M_0 \xrightarrow{\mathcal{C}} M$ for this. Conversely, every finite

configuration $\mathcal{C}$ corresponds to a unique reachable marking $M(\mathcal{C})$ given by $M(\mathcal{C}) \triangleq \pi(\mathbf{c}_{\mathcal{C}})$. We call configurations such that $M(\mathcal{C}) = M(\mathcal{C}')$ *marking equivalent*, and denote this by $\mathcal{C} \equiv_M \mathcal{C}'$.

III. Asynchronous Diagnosability

Let us start with a reminder on Diagnosability for interleaved sequences, and recall the formal definition of Sampath et al. [30] for diagnosis in interleaved models (see also Lin [24]): let $\mathcal{L} \subseteq \Sigma^*$ be a prefix-closed language (the behavior of the system to be diagnosed) over the event alphabet Σ, denote $O \subseteq \Sigma$ the set of *observable* and $UO \triangleq \Sigma \backslash O$ that of *unobservable* events[3]. Denote $P : \Sigma^* \to O^*$ the projection to observable words, that is, the homomorphism that erases all unobservable events and leaves observable ones unchanged; moreover, let $\phi \in UO$ be a *fault*[4]. Then $\mathcal{L}$ is *diagnosable* iff there exists $n \in \mathbb{N}$ such that, for any word $w = w'\phi$ in $\mathcal{L}$, any $v \in \Sigma^*$ s. th. $wv \in \mathcal{L}$ and $|v| \geqslant n$ satisfies

$$x \in P^{-1}\left[P\left(wv\right)\right] \quad \Rightarrow \quad |x|_\phi \geqslant 1. \tag{1}$$

Here, $|u|$ denotes total length of word u, and $|u|_\phi$ the number of ϕ-occurrences in u. Condition (1) means that every behavior x that produces the same sequence of observable events as wv does, contains at least one fault event: all extensions of w of at least length n will make the fault apparent. A polynomial time algorithm for testing diagnosability is given by Kumar et al. [19]; see also Yoo and Lafortune [32].

Asynchronous Diagnosis: Moving to the non-sequential framework, we shall be using analogous terminology and symbols.

Definition 3 *Let $\mathcal{N} = (P, T, F, M_0)$ be a Petri net with unfolding $\mathcal{U} = (B, E, G, \pi)$, and Σ an alarm alphabet containing the empty symbol ε; further, let $\lambda : T \to \Sigma$, for Σ some non-empty alphabet, be a* labeling function *associating alarms to system transitions. Call* silent *or* unobservable *transitions the elements of $UO \triangleq \lambda^{-1}(\varepsilon)$, and let $O \triangleq T \backslash UO$ be the set of* observable *transitions, and $\phi \in UO$ the* fault *to be diagnosed.*

Here, $\mathcal{N} = (\mathcal{P}, T, F, M_0)$ is the underlying "true" system, with the places in P representing the local states. This framework allows for *silence* (i.e. labeling by ε) and *ambiguity* (the same label for distinct events). We assume that $\phi \in UO$; the diagnosis problem considered here concerns *silent* faults, whose associated "alarm" is ε. Set $E_\phi \triangleq \pi^{-1}(\{\phi\})$, $E_O \triangleq \pi^{-1}(O)$, and $E_{UO} \triangleq E \backslash E_O$. The approach carries over to *sets* of faults without deep changes, yet we will focus on the case with one fault event to keep notations simpler. We will illustrate below the effect of different labeling functions on the same net; that is, for $\mathcal{N}$ fixed, we will ask what constraints λ must satisfy to achieve observability and diagnosability. Requiring that e.g. transition α of the net on the left hand side of figure 1 be observable, means in practice that an active sensor needs to be put on the corresponding plant part, allowing to record some alarm $\lambda(\alpha)$ on each occurrence of α. Conversely, if we determine that visibility of α is not necessary, then such a sensor need not be deployed (or, if it is already in place, we need not record its alarms).

Since the asynchronous semantics of $\mathcal{N}$ is given by the set of nonsequential processes, i.e. the *configurations* of its partial order unfolding $\mathcal{U}_\mathcal{N}$, these take over the role that is played by the word-language for automata in the above. A *configuration language (CL)* is a set of finite partially ordered configurations such that $\mathcal{C} \in \mathbb{L}$ and $\mathcal{C}' \sqsubseteq \mathcal{C}$ imply $\mathcal{C}' \in \mathbb{L}$. For a given safe Petri net $\mathcal{N} = (P, T, F, M_0)$, let the configuration language of $\mathcal{N}$ be

$$\mathbb{L}(\mathcal{N}) \quad \triangleq \quad \{\mathcal{C} \cap E \mid \mathcal{C} \in \mathbf{Con}(\mathcal{N})\};$$

that is, the language of $\mathcal{N}$ consists of its configurations, considered as sets of *events*.

Height and Progress: As Fig. 2 shows, concurrent systems may exhibit non-sequential processes whose local parts do not progress at the same pace. Suppose the fault to be diagnosed is γ. On some interleaved behaviors, γ may go undetected: if the net performs δ and an infinite number of cycles involving α and β, no decision on γ will be available. However, it is clear that if η never occurs, γ eventually occurs with certainty unless the right hand part of the net remains idle forever. In most applications, the assumption that "something will eventually happen", is realistic for every process involved. In particular, if a transition is enabled, it will eventually either fire or become disabled by another transition. Here, γ is not in any way influenced by α and β since its only conflict is with η. As opposed to the interleaved case, we therefore consider two different notions of diagnosability:

- the restrictive one of **strong diagnosability** which requires faults to be detected by *all* infinite executions;
- and **weak diagnosability** which requires that all faults be detectable at least on those executions which progress in a balanced way on all local components.

The examples will show that the two notions do not coincide. To formalize things, we have to dwell on the notion of *height*, which is the measure for progress of the system in logical time. Measuring the progress in a concurrent processes can be done by counting events, like for sequences; this leads to a notion of *length*, see [4]. This length is to be contrasted with *height*, in which the causal relations between events are taken into account: the height of a prefix, e.g. a configuration, is the length of its longest causal chain; call this the *upper height*. A more sophisticated height function measures, so to speak, the advancement of the slowest parts of the process. This concept - which we will call *lower height* - is based on the "measuring scale" formed by the prefixes $\mathcal{R}_n$, see below, which are formed by all nodes whose *upper height* is at most n. These prefixes $\mathcal{R}_n$ grow uniformly "on all ends" as n grows.

Let us formalize things now. We first define the *upper height* of a prefix $\mathcal{R}$ to be the number of events in the longest $<$-chain in $\mathcal{R}$. That is, we set recursively

$$\overline{\mathcal{H}}(\mathbf{c}_0) \quad \triangleq \quad 0 \tag{2}$$

$$\overline{\mathcal{H}}(\lceil e \rceil) \quad \triangleq \quad 1 + \overline{\mathcal{H}}(\lfloor e \rfloor), \tag{3}$$

[3] see Kumar and Shayman [21] on observability and co-observability.

[4] for simplicity, we assume there is only one *fault type* in the sense of [30]; the developments given below extend to the general case.

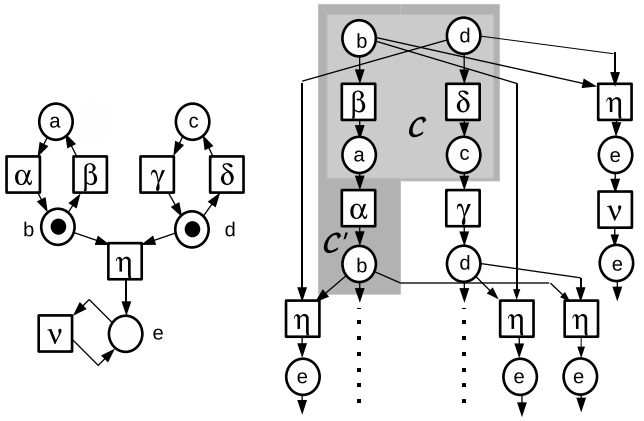


Fig. 2. Left: a Petri Net ; right: a prefix of its unfolding with two of its configurations

and for arbitrary prefixes $\mathcal{R}$,

$$\overline{\mathcal{H}}(\mathcal{R}) \quad \triangleq \quad \sup\{\overline{\mathcal{H}}(\lceil e \rceil) \mid e \in E \cap \mathcal{R}\} \tag{4}$$

(note that with $\sup(\emptyset) \triangleq 0$, this is consistent with $\overline{\mathcal{H}}(\mathbf{c}_0) = 0$). Let $n \in \mathbb{N}_0$, and define $\mathcal{R}_n$ to denote the maximal prefix whose height does not exceed n; that is, $\mathcal{R}$ is the set of all nodes of height not exceeding n:

$$\begin{array}{rcl} E_n & \triangleq & \{e \in E \mid \overline{\mathcal{H}}(\lceil e \rceil) \leqslant n\} \\ \mathcal{R}_n & \triangleq & \cup_{e \in E_n} (\lceil e \rceil \cup e^\bullet)\,; \end{array}$$

call $\mathcal{R}_n$ $\mathcal{N}$'s *nth prime prefix.*

Upper and lower height

Specializing on configurations, we have *two* different height functions: First, applying (4) directly to configurations - seen as special prefixes - yields the **upper height** $\overline{\mathcal{H}}(\mathcal{C}) \in \mathbb{N} \cup +\infty$ for $\mathcal{C} \in \mathbf{Con}(\mathcal{N})$. Second, define the **lower height** $\underline{\mathcal{H}}(\mathcal{C})$ of $\mathcal{C}$ as the index of the greatest $\mathcal{R}_n$ in which the trace $\mathcal{C} \cap \mathcal{R}_n$ of $\mathcal{C}$ is a maximal configuration of the occurrence net $\mathcal{R}_n$. Formally, we define, setting $\sup(\emptyset) \triangleq 0$:

$$\underline{\mathcal{H}}(\mathcal{C}) \quad \triangleq \quad \sup\{n \in \mathbb{N} \mid \exists\, \omega \in \Omega :\ \omega \cap \mathcal{R}_n = \mathcal{C} \cap \mathcal{R}_n\}\,.$$

Since the $\mathcal{R}_n$ consist only of executions of perfectly balanced, uniform progress, the lower height of a configuration $\mathcal{C}$ can be seen as the *time until progress imbalance sets in* on $\mathcal{C}$. This height function can be used to define a metric that is standard in partial order semantics, see e.g. [4], [22], [23]. The resulting topology on runs can be used to describe system properties; application of those topological tools to diagnosability is the subject of ongoing work whose discussion would lead us too far afield here.

Of course, $\underline{\mathcal{H}}(\mathcal{C}) \leqslant \overline{\mathcal{H}}(\mathcal{C})$, with equality iff either

- $\underline{\mathcal{H}}(\mathcal{C}) = +\infty$, or
- all runs $\omega, \omega' \in \Omega$ that extend $\mathcal{C}$, i.e. $\mathcal{C} \sqsubseteq \omega$ and $\mathcal{C} \sqsubseteq \omega'$, agree on $\mathcal{R}_n$ with $n = \overline{\mathcal{H}}(\mathcal{C})$, i.e.

$$\omega \cap \mathcal{R}_{\overline{\mathcal{H}}(\mathcal{C})} = \mathcal{C} = \omega' \cap \mathcal{R}_{\overline{\mathcal{H}}(\mathcal{C})}.$$

Progress

Call the finite configurations that satisfy $\underline{\mathcal{H}}(\mathcal{C}) = \overline{\mathcal{H}}(\mathcal{C})$ **progressive**. By extension, call an infinite configuration $\mathcal{C}$ progressive iff all its finite truncations $(\mathcal{C} \cap \mathcal{R}_n)_{n \in \mathbb{N}}$ are progressive. A non-progressive configuration $\mathcal{C}$ may allow an extension by events whose height is inferior to $\overline{\mathcal{H}}(\mathcal{C})$. By contrast, progressive configurations cannot be extended without increasing the lower height.

The term of 'progressive' configurations is justified by the fact that their local processes all progress in a *fair* way, none of them lagging behind indefinitely[5].

Example: In Fig. 2, consider the configurations $\mathcal{C}$ (light gray) and $\mathcal{C}'$, with $\mathcal{C} \sqsubseteq \mathcal{C}'$; we have $\overline{\mathcal{H}}(\mathcal{C}) = \underline{\mathcal{H}}(\mathcal{C}') = 1$, but $\overline{\mathcal{H}}(\mathcal{C}') = 2$ and $\underline{\mathcal{H}}(\mathcal{C}') = 1$. Clearly, $\mathcal{C}$ is progressive and $\mathcal{C}'$ is not; however, $\mathcal{C}'$ can be extended into progressive configurations, e.g. $\mathcal{C}' \cup \{\gamma\}$. Denote as $\mathbb{L}_{\mathbf{prog}}$ the set of *progressive* configurations.

Faulty configurations : For $\mathcal{C} \in \mathbb{L}$ let $\mathcal{C}_O$ be the labeled partial order induced by $\mathcal{C}$ on $\mathcal{C} \cap E_O$. Write $\mathcal{C} \sim_O \mathcal{C}'$ iff $\mathcal{C}_O$ and $\mathcal{C}'_O$ are isomorphic. Let $\equiv_\phi$ be the equivalence on $\mathbb{L}$ given by

$$\mathcal{C} \equiv_\phi \mathcal{C}' \quad iff \quad (\mathcal{C} \cap E_\phi = \emptyset \iff \mathcal{C}' \cap E_\phi = \emptyset)\,;$$

that is, two configurations are ϕ-equivalent if either both contain a fault, or neither of them does.

Live and dead configurations: In analogy with the *liveness* requirement in [30], let us say that a configuration $\mathcal{C}$ is **dead** iff it has no infinite extension, i.e. iff $\mathcal{C} \sqsubseteq \mathcal{C}'$ implies that $\overline{\mathcal{H}}(\mathcal{C})' < \infty$. This finishes our preparations.

Definition 4 *Let height measure $H : \mathbb{L} \to [0, \infty)$ be either $H \equiv \underline{\mathcal{H}}(\bullet)$ or $H \equiv \overline{\mathcal{H}}(\bullet)$. A CL $\mathbb{L}$ is H-***diagnosable** *w.r.t. O and ϕ iff there exists $n \in \mathbb{N}$ such that for all configurations $\mathcal{C} \in \mathbb{L}$ that have the form $\mathcal{C} = \lceil e_\phi \rceil$ with $e_\phi \in E_\phi$, every $\mathcal{C} \in \mathbb{L}$ such that*

(a) $\mathcal{C}_\phi \sqsubseteq \mathcal{C}$,
(b) $\mathcal{C}$ is not dead, and
(c) $H(\mathcal{C}) \geqslant H(\mathcal{C}_\phi) + n$, satisfies:

$$\forall \mathcal{C}' \in \mathbb{L} :\ \mathcal{C}' \sim_O \mathcal{C} \quad \Rightarrow \quad E_\phi \cap \mathcal{C}' \neq \emptyset. \tag{5}$$

Now, we lift diagnosability from languages to nets:

Definition 5 *Let $\mathcal{N} = (P, T, F, M_0)$ a safe Petri net, $\mathcal{U}_\mathcal{N} = (B, E, G, \mathbf{c}_0)$ its unfolding, and $\mathbb{L}$ and $\mathbb{L}_{\mathbf{prog}}$ as above. Further, let $E_O \triangleq \pi^{-1}(O) \subseteq E$ be the set of observable events, $\phi \notin E_O$ a and $E_\phi \triangleq \pi^{-1}(\{\phi\})$. Then $\mathcal{N}$ is said to*

1) *be* observable, *or* satisfy *$\mathcal{OBS}$ (for O), iff $\forall\, \mathcal{C}, \mathcal{C}' \in \mathbb{L}$,*

$$\left.\begin{array}{rc} & \mathcal{C} \sqsubseteq \mathcal{C}' \\ \wedge & \mathcal{C} \neq \mathcal{C}' \\ \wedge & M_\mathcal{C} = M_{\mathcal{C}'} \end{array}\right\} \quad \Rightarrow \quad (\mathcal{C} \not\sim_O \mathcal{C}') \tag{6}$$

2) *be* observable, *or* satisfy *$\mathcal{WOBS}$ (for O), iff (6) holds in $\mathbb{L}_{\mathbf{prog}}$.*

[5]In fact, progressive executions for safe nets are necessarily *fair* in the sense that any transition which is enabled an infinite number of times must also fire an infinite number of times. The converse is not true; fair executions do not necessarily lead to progressive configurations.

3) *be* **(strongly) diagnosable**, *or satisfy* $\mathbb{D}$ *w.r.t.* O *and* ϕ, *iff*
   a) $\mathcal{N}$ *satisfies* $\mathcal{OBS}$ *(for* O*), and*
   b) $\mathbb{L}$ *is* $\overline{\mathcal{H}}(\bullet)$*-diagnosable w.r.t.* O *and* ϕ.
4) *be* **weakly diagnosable**, *or satisfy* $\mathbb{W}$ *w.r.t.* O *and* ϕ, *iff*
   a) $\mathcal{N}$ *satisfies* ($\mathcal{WOBS}$), *and*
   b) $\mathbb{L}_{\mathbf{prog}}$ *is* $\underline{\mathcal{H}}(\bullet)$*-diagnosable w.r.t.* O *and* ϕ.

Some remarks are in order. First, a *strongly* diagnosable net $\mathcal{N}$ is also diagnosable in the sense of [12], [30] (see (1)), and vice versa. In fact, consider the interleavings of $\mathcal{N}$s runs. The existence of the constant bound n, such that the fault can be decided with certainty at most n actions after occurrence of the fault, corresponds to the fact that only a finite number of invisible transition firings can occur concurrently to any visible transition.

Secondly, note that while strong diagnosability trivially implies weak diagnosability, the converse is not true[6]: In Fig. 2, suppose β is the fault action $\phi \triangleq \beta$, $O = \{\alpha\}$, and for $m \in \mathbb{N}$, let $\mathcal{C}^{(m)}$ be the smallest configuration such that

(i) β never occurs on $\mathcal{C}^{(m)}$, and
(ii) δ occurs exactly m times on $\mathcal{C}^{(m)}$.

Then $\overline{\mathcal{H}}(\mathcal{C}^{(m)}) = 2m+1$, yet $\mathcal{C}^{(m)} \sim_O \mathcal{C}^{(1)}$ for all m, so the system is not strongly diagnosable. Note that the $\mathcal{C}^{(m)}$ are not progressive; all *progressive* configurations of height at least $2k+1$ contain at least k instances of $\alpha \in O$. It follows from the above that the system is weakly diagnosable.

IV. Characterization of Diagnosability

After these preparations, we are now ready to state and prove our characterizations of weak and strong diagnosability. As in the classical setting, diagnosability is *violated* iff the system is able to perform two indiscernible, non-fault-equivalent cycles. That is, there must be O-equivalent configurations $\mathcal{C}_1$ and $\mathcal{C}_2$ with, respectively, extensions $\mathcal{C}_1'$ and $\mathcal{C}_2'$ that are O-equivalent and marking-equivalent, but not ϕ-equivalent; then the system may repeat that cyclic behavior indefinitely, without a decision about occurrence of faults. In fact:

Theorem 1 *With labeling* $\lambda : T \to \Sigma$, *and* ϕ, O, UO, $\mathbb{L}$ *and* $\mathbb{L}_{\mathbf{prog}}$ *as above, a safe Petri net* $\mathcal{N} = (P, T, F, M_0)$ *that satisfies* $\mathcal{OBS}$ *is* **strongly diagnosable** *w.r.t.* O *and* ϕ *iff for all* $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_1', \mathcal{C}_2' \in \mathbb{L}$,

$$\left.\begin{array}{l} \mathcal{C}_1 \sim_O \mathcal{C}_2 \ \wedge\ \mathcal{C}_1' \sim_O \mathcal{C}_2' \\ \mathcal{C}_1 \neq \mathcal{C}_1' \\ \forall\ i \in \{1,2\} : \left\{ \begin{array}{rl} & M_{\mathcal{C}_i} = M_{\mathcal{C}_i'} \\ \wedge & \mathcal{C}_i \sqsubseteq \mathcal{C}_i' \end{array} \right. \end{array}\right\} \Rightarrow \mathcal{C}_1' \equiv_\phi \mathcal{C}_2'. \qquad (7)$$

A net $\mathcal{N}$ *that satisfies* $\mathcal{WOBS}$ *is* ***weakly diagnosable*** *w.r.t.* O *and* ϕ *iff the restriction of (7) to* $\mathbb{L}_{\mathbf{prog}}$ *holds.*

In other words, violations of diagnosability are characterized by the presence of configurations $\mathcal{C}_1$ and $\mathcal{C}_2$ such that *(i)* $\mathcal{C}_1$ and $\mathcal{C}_2$ are observationally equivalent, *(ii)* $\mathcal{C}_1$ and $\mathcal{C}_2$ have extensions $\mathcal{C}_1'$ and $\mathcal{C}_2'$ that are observationally equivalent to one another, and such that $\mathcal{C}_1'$ is a *proper* marking-equivalent extension of $\mathcal{C}_1$ and $\mathcal{C}_2'$ *any* marking-equivalent extension of $\mathcal{C}_2$; and *(iii)* $\mathcal{C}_1'$ is faulty and $\mathcal{C}_2'$ is not. Note, before we proceed to the proof, that (7) allows $\mathcal{C}_2 = \mathcal{C}_2'$ in the assumption. In preparation of the proof below, denote as $\mathcal{C}_1 \circ \mathcal{C}_2$ the *concatenation* configuration obtained from $\mathcal{C}_1$ in $\mathcal{N} = (N, M_0)$ and $\mathcal{C}_2$ in $\mathcal{N}_{\mathcal{C}_1}(N, M(\mathcal{C}_1))$ appended after $\mathcal{C}_1$. Define further $\mathcal{C}^1 \triangleq \mathcal{C}$ and $\mathcal{C}^{k+1} \triangleq \mathcal{C}^k \circ \mathcal{C}$.

Proof: We show the strong diagnosability case; the result for weak diagnosability is obtained in the same way with $\mathbb{L}$ replaced by $\mathbb{L}_{\mathbf{prog}}$. For the **"only if"** part, let $\mathcal{C}_i \sqsubseteq \mathcal{C}_i'$, $i \in \{1,2\}$, constitute a violation of (7), i.e.

1) $\mathcal{C}_2' \cap E_\phi \neq \emptyset$ and $\mathcal{C}_1' \cap E_\phi = \mathcal{C}_1 \cap E_\phi = \emptyset$;
2) defining configurations μ_i by $\mathcal{C}_i' = \mathcal{C}_i \circ \mu_i$ for $i \in \{1,2\}$, one has that μ_1 contains at least one event, and
3) $\mathcal{C}_1 \sim_O \mathcal{C}_2$, $\mathcal{C}_1' \sim_O \mathcal{C}_2'$ and $M_{\mathcal{C}_i} = M_{\mathcal{C}_i'}$.

From 2, it follows that a copy of μ_i can be appended to $\mathcal{C}_i'$ as well, and so forth; let $\mathcal{C}_i^k \triangleq \mathcal{C}_i \circ \mu_i^k$ be the configuration obtained after appending k copies of μ_i to $\mathcal{C}_i$. Observe that $\overline{\mathcal{H}}(\mathcal{C}_1^k) \geqslant \max(k, \overline{\mathcal{H}}(\mathcal{C}_1))$. Thus $\overline{\mathcal{H}}(\mathcal{C}_1^k) \to \infty$ as $k \to \infty$. Now, by assumption we have $\mathcal{C}_2^k \sim_O \mathcal{C}_2$; by construction, $\mu_2 \cap E_\phi \neq \emptyset$ and thus $\mathcal{C}_2^k \cap E_\phi \neq \emptyset$. Hence (5) is violated.

To show the **"if"** part, suppose (5) does *not* hold: for every $n \in \mathbb{N}$, there exists $\mathcal{C}(n) \in \mathbb{L}(\mathcal{N})$ such that

1) some $e \in E_\phi$ is $\leqslant$-maximal in $E \cap \mathcal{C}(n)$, and
2) there exist $\mathcal{C}_1(n), \mathcal{C}_2(n) \in \mathbb{L}(\mathcal{N})$ such that

$$\begin{array}{rcl} \mathcal{C}(n) \sqsubseteq \mathcal{C}_1(n) & \wedge & \overline{\mathcal{H}}(\mathcal{C}_1(n)) \geqslant \overline{\mathcal{H}}(\mathcal{C}(n)) + n \\ \wedge\ \mathcal{C}_2(n) \sim_O \mathcal{C}_1(n) & \wedge & \mathcal{C}_2(n) \cap E_\phi = \emptyset. \end{array}$$

Assume first that one can choose $\mathcal{C}_1'(n)$ with

$$\mathcal{C}(n) \ \sqsubseteq\ \mathcal{C}_1'(n) \sqsubseteq \mathcal{C}_1(n)$$

such that

$$\begin{array}{rcl} M_{\mathcal{C}_1(n)} & = & M_{\mathcal{C}_1'(n)}, \\ \mathcal{C}_1(n) \sim_O \mathcal{C}_1'(n) & \wedge & \mathcal{C}_1(n) \neq \mathcal{C}_1'(n); \end{array}$$

then we are done by setting $\mathcal{C}_2' \triangleq \mathcal{C}_2(n)$. Thus assume

$$\forall\ \mathcal{C}_1' : \left[\left\{ \begin{array}{l} \mathcal{C}(n) \sqsubseteq \mathcal{C}_1'(n) \sqsubseteq \mathcal{C}_1(n) \\ \mathcal{C}_1(n) \sim_O \mathcal{C}_1'(n) \\ M_{\mathcal{C}_1(n)} = M_{\mathcal{C}_1'(n)} \end{array} \right\} \Rightarrow \mathcal{C}_1(n) = \mathcal{C}_1'(n) \right]. \qquad (8)$$

For any $\mathcal{C}_1 \sqsubseteq \mathcal{C}_1(n)$, let $Simil(\mathcal{C}_1, n)$ be the set of configurations $\mathcal{C}_2 \sqsubseteq \mathcal{C}_2(n)$ such that $\mathcal{C}_2 \sim_O \mathcal{C}_1$. For any reachable marking M of $\mathcal{N}$, let $S_1(M, n)$ be the set of configurations $\mathcal{C}_1$ such that *(i)* $\mathcal{C}_1 \sqsubseteq \mathcal{C}_1(n)$ and *(ii)* $M = M(\mathcal{C}_1)$. Let K be the number of all reachable markings of $\mathcal{N}$. Then for all $n > K$, there is at least one marking M such that $|S_1(M, n)| \geqslant 2$; repeating the argument, one finds using (8) that for all $n > K^2$ there exists a marking M such that $|S_1(M, n)| > K$. With

$$Simil_2(M, n) \ \triangleq\ \left\{ \begin{array}{l} \mathcal{C}_2 \in \mathbb{L}(\mathcal{N}), \\ \mathcal{C}_2 \sqsubseteq \mathcal{C}_2(n) \end{array} \ \middle|\ \begin{array}{r} \exists \mathcal{C}_1 \in S_1(M, n) : \\ \mathcal{C}_1 \sim_O \mathcal{C}_2 \end{array} \right\},$$

we therefore have $|Simil_2(M, n)| > K$. Thus there exist $\mathcal{C}_2, \mathcal{C}_2' \in U_2(M, n)$ such that $\mathcal{C}_2 \neq \mathcal{C}_2'$ and $M_{\mathcal{C}_2} = M_{\mathcal{C}_2'}$. By definition of $Simil_2(M, n)$, $\mathcal{C}_1 \sim_O \mathcal{C}_2$ and $\mathcal{C}_1' \sim_O \mathcal{C}_2'$. Since, by construction, $\mathcal{C}_1 \sqsubseteq \mathcal{C}_1' \sqsubseteq \mathcal{C}_1(n)$ and $M_{\mathcal{C}_1} = M_{\mathcal{C}_1'}$, property

[6] similarly for weak and strong *observability*

(7) is violated, q.e.d. □

Note that the progressive and non-progressive cases do not require separate proofs: the difference is only in the set of configurations over which the different $\mathcal{C}$-variables in the proof may range. However, recall that the strong and weak diagnosability *properties* are not equivalent, as the examples above and below show.

Examples:

1) Fig. 1: Let us ask under which choices of observable set $O \subseteq T$ the net $\mathcal{N}$ satisfies $\mathcal{OBS}$, and if so, whether $\mathcal{N}$ is then diagnosable for that O and a given fault ϕ. First, we claim that $\mathcal{OBS}$ (and even $\mathcal{WOBS}$) is *equivalent* with

$$\gamma \in O \quad \vee \quad [\beta \in O \wedge (|\{\alpha, \delta, \zeta\} \cap O \geqslant 1))]\,. \qquad (9)$$

In fact, every $\mathcal{C} \in max_{ON}$ contains γ-labeled events, so the implications $(\gamma \in O) \Rightarrow \mathcal{OBS} \Rightarrow \mathcal{WOBS}$ are immediate. On the other hand, suppose $\gamma \notin O$; then we deduce from the configuration $\mathcal{C}$ on shaded background in the figure that $\beta \in O$ (otherwise $\mathcal{C}$ and two of its prefixes yield witnesses of non-diagnosability). Inspecting the other non-dead configurations of $max^*_{\mathcal{R}}$ in a similar way, we see that $\alpha \notin O$ entails $(\delta \in O) \vee (\zeta \in O)$; we deduce that (9) is necessary for (both weak and strong) observability, and thus for (both weak and strong) diagnosability. Now, let us check sufficiency, i.e. whether $\gamma \in O$ makes $\mathcal{N}$ *diagnosable*. For this, let us consider the cases $\phi = \eta$ and $\phi = \beta$. Since we have to respect $\phi \notin O$, (9) is reduced , in the case $\phi = \beta$, to

$$[\gamma \in O]\,. \qquad (10)$$

One has in the case $\phi = \beta$ that the conjunction of

(i) $\phi^{-1}(\beta) \cap \mathcal{C} \neq \emptyset$ and
(ii) $\overline{\mathcal{H}}(\mathcal{C}') > \overline{\mathcal{H}}(\mathcal{C}) + 1$ or $\underline{\mathcal{H}}(\mathcal{C}') > \underline{\mathcal{H}}(\mathcal{C}) + 1$

implies $\phi^{-1}(\gamma) \cap \mathcal{C} \neq \emptyset$.

For the case $\phi = \eta$, consider the set max_η of configurations from max_{ON} that contain an η-event. Inspection of Fig. 1 shows that for every $\mathcal{C}_\eta \in max_\eta$, any extension $\mathcal{C}'_\eta$ of $\mathcal{C}_\eta$ that satisfies

- either $\overline{\mathcal{H}}(\mathcal{C}') > \overline{\mathcal{H}}(\mathcal{C}) + 1$ or
- $\underline{\mathcal{H}}(\mathcal{C}') > \underline{\mathcal{H}}(\mathcal{C}) + 2$,

contains a γ-instance. Summing up, $\gamma \in O$ is necessary and sufficient for $\mathcal{OBS}$, $\mathcal{WOBS}$, $\mathbb{D}$, and $\mathbb{W}$.

2) Fig. 2: Since both cycles in this net can perform an arbitrary number of rounds independently of one another, strong observability and strong diagnosability clearly require that at least one transition out of $\{\alpha, \beta\}$ *and* at least one transition out of $\{\gamma, \delta\}$ be in O. For weak observability, it is both necessary and sufficient that one transition out of $\{\alpha, \beta\}$ *or* one transition out of $\{\gamma, \delta\}$ be in O. Supposing, as in the above discussion, that γ is the fault transition, having $\delta \in O$ is clearly sufficient to diagnose γ. In fact, it is not necessary to have α, β, or η in O for detecting γ's occurrence since the number of occurrences of δ gives sufficient information under the progress assumption : if δ occurs only once, then γ has not occurred ; in all other cases, δ occurs more than once, and γ must have occured. Note that this net, with $O = \{\delta\}$ or $O = \{\delta, \eta\}$, is weakly observable: it is not possible to reproduce, by executing progressive configurations, any reachable marking without firing δ; any additional firing of $\{\beta\}$ *alone* is non-progressive unless it is balanced by an additional firing of δ.

V. Verification of Diagnosability

A detailed analysis of checking *strong* diagnosability is given by Cabasino et al. [11], [12], using Net invariants. We will focus our attention on *weak* diagnosability here.

Finite Complete Prefixes: The runs of $\mathcal{U}(\mathcal{N})$ represent all maximal nonsequential executions. That is, any firing *sequence* of $\mathcal{N}$ is obtained as the linear order extension of (some prefix of) some run $\omega \in \Omega(\mathcal{U}(\mathcal{N}))$. Even if unfoldings are infinite in general, any safe Petri net admits finite *complete* prefixes that contain at least one copy of every reachable marking; this is what allows using branching processes in Model Checking [6], [25]. Methods for obtaining and optimizing such *complete prefixes* have received considerable attention in the literature, see e.g. [20].The definition and size of such prefixes varies with the intended purpose. We use here the following definition, similar to that in [13]:

Definition 6 *The order 1 unfolding, denoted $\mathcal{U}_1(\mathcal{N})$, is a finite prefix of the unfolding obtained by stopping the construction when we reach a* ***cut-off*** *event e, i.e., an event such that:*

- *EITHER $M(\lceil e \rceil) = M_0$;*
- *OR there exists another event e' such that* (i) $\lceil e' \rceil \subseteq \lceil e \rceil$, *and* (ii) $M(\lceil e \rceil) = M(\lceil e' \rceil)$.

In the following we call e' the mirror transition *of e in $\tilde{\mathcal{N}}_1(M_0)$. Once we have constructed $\mathcal{U}_1 \triangleq \mathcal{U}_1(\mathcal{N})$, assume we continue the unfolding until we reach an event e such that there exist another event e' with the following properties:*

- *if e' belongs to $\mathcal{U}_1$, it is a cut-off event of $\mathcal{U}_1$;*
- *$\lceil e' \rceil \subseteq \lceil e \rceil$;*
- *$M(\lceil e \rceil) = M(\lceil e' \rceil)$.*

The resulting net, denoted $\mathcal{U}_2(\mathcal{N})$, is called 2nd order unfolding*; by iterating the above, one obtains a nested family $(\mathcal{U}_n(\mathcal{N}))_{n \in \mathbb{N}}$ of* n-th order unfoldings.

Note that the initial definition from [25] used as cutoff criterion the *cardinality*, i.e. $|\lceil e' \rceil| < |\lceil e \rceil|$, which would lead to a shorter prefix in general yet not guarantee completeness w.r.t. computing the *reveals* relation below.

Lemma 2 *Let $\mathcal{R}$ be any prefix of the unfolding $\mathcal{U}_{ON}$. If there exist* **witnesses of non-diagnosability** *in $\mathcal{R}$, configurations $\mathcal{C}_i, \mathcal{C}'_i$ for $i \in \{1, 2\}$ such that the left hand side of (7) holds, but $\kappa'_1 \not\equiv_\phi \kappa'_2$, then κ'_1, κ_2 can be chosen maximal for ON.*

Proof: By assumption, there exist *(i)* $\mathcal{C}''_i$ such that $\mathcal{C}_i \sqsubseteq \mathcal{C}''_i \sqsubseteq \mathcal{C}'_i$ (thus $\mathcal{C}''_i \sim_O \mathcal{C}'_i$ and $\mathcal{C}''_1 \sim_O \mathcal{C}''_2$) and *(ii)* a maximal configuration $\mathcal{C}'''_i$ of ON such that $M_{\mathcal{C}''_i} = M_{\mathcal{C}'''_i}$, i.e. $\mathcal{C}''_1, \mathcal{C}''_2, \mathcal{C}'''_1, \mathcal{C}'''_2$ are witnesses of non-diagnosability. □

We have:

Theorem 2 *For a given net $N = (P, T, F)$, there exists a finite number $Z = Z(\mathcal{N})$ such that for any 1-safe marking*

$M_0 \subseteq P$ *of* N*, the* Z*-th prefix* $\mathcal{R}_Z$ *of the unfolding of* $\mathcal{N} = (N, M_0)$ *is sufficient to verify (strong or weak) diagnosability: if there exist any* $\mathcal{C}_1, \mathcal{C}'_1, \mathcal{C}_2, \mathcal{C}'_2$ *such that (7) is violated, one can choose them with this property such that* $\max(\overline{\mathcal{H}}(\mathcal{C}'_1), \overline{\mathcal{H}}(\mathcal{C}'_2)) \leqslant Z$.

Proof: Let $\mathcal{A} \subseteq \mathbf{Con}(\mathcal{N})$ be a $\sim_O$-equivalence class. We say that $\mathcal{A}$ is *reducible* iff for all $\mathcal{C} \in \mathcal{A}$, there exist $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ such that *(i)* $\overline{\mathcal{H}}(\mathcal{C}_2) > 0$, *(ii)* $\mathcal{C} = \mathcal{C}_1 \circ \mathcal{C}_2 \circ \mathcal{C}_3$, and *(iii)* for all $i \in \mathbb{N}$, $\mathcal{C}_1 \circ \mathcal{C}_2^i \circ \mathcal{C}_3 \subseteq \mathbb{L}(\mathcal{N})$; call $\mathcal{A}$ *irreducible* otherwise. Since $K < \infty$, the result follows from Lemma 2 with the pigeonhole principle once the following claim is proved: *the number* J *of irreducible* $\sim_O$*-classes of* $\mathcal{N}$ *is bounded above by* 2^K. For this, note that the height $\overline{\mathcal{H}}(\mathcal{C})$ of any configuration $\mathcal{C} \in \mathbb{L}(\mathcal{N})$ that does not contain two comparable markings, i.e. such that $\mathcal{C}_1 \sqsubseteq \mathcal{C}_2 \sqsubseteq \mathcal{C}$ and $M_{\mathcal{C}_1} = M_{\mathcal{C}_2}$ imply $\mathcal{C}_1 = \mathcal{C}_2$, is bounded above by K. Thus all $\mathcal{A}$ such that

$$\overline{\mathcal{H}}(\mathcal{A}) \quad \triangleq \quad \inf\{\overline{\mathcal{H}}(\mathcal{C}) \mid \mathcal{C} \in \mathcal{A}\}$$

Finally, the number of configurations of height K or less is bounded above by 2^K, so we are done. □

The upper bounds on the size of the complete prefix are far from sharp; $\mathcal{R}^*$ can be chosen moderate if there is a high degree of parallelism in $\mathcal{N}$ and no excessive branching. The efficiency of diagnosability checking thus requires a careful choice of prefixes; see [6], [20].

One obtains thus the following algorithm for checking weak diagnosability of $\mathcal{N} = (P, T, F, M_0)$:

(A) Compute a complete prefix $\mathcal{R}^*$ as above, and its set $max_{ON} \triangleq \Omega(\mathcal{R}^*)$ of maximal configurations.

(B) For any pair $\mathcal{C}'_1, \mathcal{C}'_2$ of maximal configurations such that $\mathcal{C}_1 \sim_O \mathcal{C}_2$, check whether there exist $\mathcal{C}_i \sqsubseteq \mathcal{C}'_i$ such that $\mathcal{C}_1 \sim_O \mathcal{C}'_1$, $M_{\mathcal{C}_1} = M_{\mathcal{C}'_1}$ and $\mathcal{C}_1 \sqsubseteq \mathcal{C}'_1$.

VI. The *Reveals* Relation

In the above discussion, we use implicitly reasonings of the form 'if x occurs, then y has already occured, or will occur eventually', in the sense that any infinite run that contains x also contains y. Under progress assumption (see above), this means that y *is inevitable given* x. In the context of the occurrence net in Fig. 3, it is obvious that, for any run ω,

$$k \in \omega \quad \Rightarrow \quad e \in \omega \Rightarrow b \in \omega; \tag{11}$$

in fact, (11) reflects the inheritance of $\#$ under $<$. But one also obtains the following facts in Fig. 3:

$$a \in \omega \quad \Longleftrightarrow \quad \neg(b \in \omega) \Longleftrightarrow c \in \omega \tag{12}$$

$$e \in \omega \quad \Longleftrightarrow \quad f \in \omega; \tag{13}$$

the reader is invited to check that (12) and (13) follow from the maximality of runs. Now, the inheritance of conflict along causality relations is not sufficient to derive (12) and (13); so one might suspect that, to obtain (12 and 13) from the relational structure, one would have to explore the entire set of configurations. However, we will show here that it suffices to consider an auxiliary relation, computable from the $\#$ relation in a finite bounded prefix $\mathcal{R}$ of the unfolding. To start, let us define:

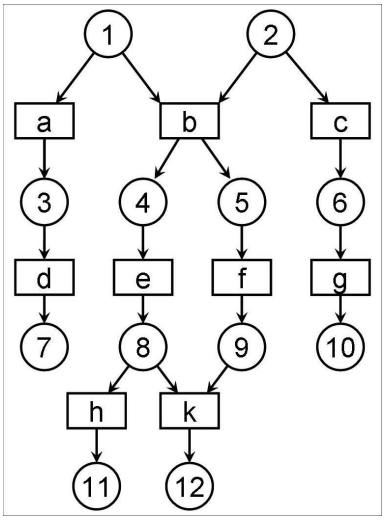


Fig. 3. On the relation $\triangleright$

Definition 7 *For a node* $x \in (B \cup E)$*, the* ***conflict set*** *of* x *is defined as* $\#[x] \triangleq \{x' \mid x \# x'\}$*. The* root conflict set *is given by* $\#_\mu[x] \triangleq \{y \mid x \# y \ \wedge \ \forall\, z : z < y \Rightarrow \neg(z \# x)\}$*; the symbol* $\#_\mu[\bullet]$ *is borrowed from [1] where it denotes immediate conflict in event structures. Node* x reveals y*, written* $x \triangleright y$*, iff* $\#[x] \supseteq \#[y]$*. The* revealed range *of* x *is* $\triangleright[x] \triangleq \{y \mid x \triangleright y\}$.

One immediately checks that $\triangleright$ is reflexive and transitive.

Lemma 3 $x \triangleright y$ *holds iff for all runs* ω*,*

$$x \in \omega \quad \Rightarrow \quad y \in \omega \tag{14}$$

Proof: If $x \in \omega$ and $y \notin \omega$, there exists a node $z \in \#[y] \cap \omega$; in fact, otherwise $\omega \cup \lceil y \rceil$ would be a configuration, and ω could not be maximal. If $x \triangleright y$, then $z \in \#[x] \cap \omega$, which is impossible, so we must have $\neg(x \triangleright y)$. Conversely, suppose that (14) holds for every ω; then there exists z such that $z \# y$ and $\neg(z \# x)$. But then there exists a run ω_z such that $x, z \in \omega_z$, but by assumption $y \notin \omega_z$, hence (14) is violated for ω_z. □

Relation $\triangleright$ is asymmetric: in fact, in Fig. 5 (left) we have $h \triangleright f$ but $\neg(f \triangleright h)$. On the other hand, $\triangleright$ is not a partial order: consider $e \triangleright f$ and $f \triangleright e$. However, the following holds:

Lemma 4 $x < y$ *implies that* $y \triangleright x$.

Proof: $x < y \Rightarrow \#[x] \subseteq \#[y]$ by $\#$-inheritance. □

As a consequence, we have:

Lemma 5 $\triangleright[x]$ *is a configuration.*

Proof: Since $\lceil x \rceil \subseteq \triangleright[x]$ by Lemma 3, we have $\mathbf{c}_0 \subseteq \triangleright[x]$; thus Lemma 4 implies the result. □

The following result shows that in order to decide whether $x \triangleright y$, it suffices to know $\#_\mu[x]$ and $\#_\mu[y]$:

Theorem 3 $\#_\mu[x]$ *generates* $\#[x]$ *through inheritance:*

$$\#[x] \quad = \quad \{z \mid \exists\ y \in \#_\mu[x] :\ y \leqslant z\}\,. \tag{15}$$

As a consequence, $x_1 \triangleright x_2$ *iff* $\#_\mu[x_1] \supseteq \#_\mu[x_2]$.

Proof: The inclusion$\#[x] \supseteq \{z \mid \exists\ y \in \#_\mu[x] :\ y \leqslant z\}$ being obvious, it remains to show

$$\#[x] \quad \subseteq \quad \{z \mid \exists\ y \in \#_\mu[x] :\ y \leqslant z\}\,. \tag{16}$$

Take any $y \in \#[x]\backslash\#_\mu[x]$. Since $x\#y$, there exist a condition b_1 and events x_1, y_1 such that *(i)* $x_1 \neq y_1$; *(ii)* $b_1 \in {}^\bullet x_1 \cap {}^\bullet y_1$; and *(iii)* $x_1 \leqslant x$ and $y_1 \leqslant y$. We will now inspect a family of nodes y_n, starting with $n = 1$: Let $n \geq 1$. If $y_n \in \#_\mu[x]$, we are done; otherwise there exist $b_{n+1} \in B$ and $x_{n+1}, y_{n+1} \in E$ such that *(a)* $x_{n+1} \neq y_{n+1}$; *(b)* $b_{n+1} \in {}^\bullet x_{n+1} \cap {}^\bullet y_{n+1}$; *(c)* $x_{n+1} \leqslant x$ and $y_{n+1} < y_n$. If we find recursively infinitely many such $y_1, y_2, \ldots$, this contradicts property 3) of Definition 1, since $y \geqslant y_1 > y_2 > \ldots$ There thus exists $n \in \mathbb{N}$ such that $y_n \in \#_\mu[x]$, proving (16). □

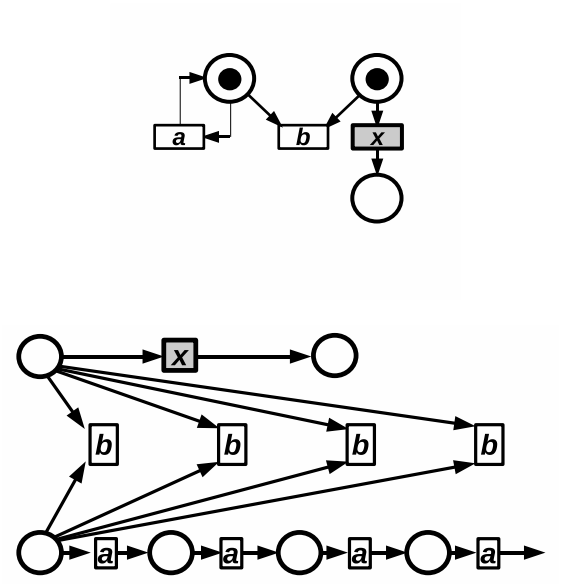


Fig. 4. Top: a safe Petri net; bottom: a prefix of its unfolding, exhibiting an infinite root conflict set.

So far we were able to reduce the computation of the *reveals* relation to comparison of root conflict sets. These sets can be infinite, as the example in Fig. VI shows: $\#_\mu[x]$ consists of all the b-labeled events in the central horizontal axis of the figure. However, the relation $\triangleright$ can be recursively computed, thanks to the following result:

Theorem 4 *Denote as* $round(x)$ *the smallest* n *such that* x *belongs to* $\mathcal{U}_n(\mathcal{N})$*, and as* K *the number of reachable markings of* $\mathcal{N} = (P, T, F, M_0)$*. Define, for nodes* x, y, z*, the "witness" predicate* $\mathbf{wit}(x, y, z)$ *by*

$$\mathbf{wit}(x, y, z) \quad \stackrel{\triangle}{\Longleftrightarrow} \quad [(z \ \#\ y) \ \wedge\ \neg(z \ \#\ x)]\,. \tag{17}$$

Then, with $m \triangleq \max(round(x), round(y))$*, for any two nodes* x, y *such that* $\neg(x \triangleright y)$*, there exists a* $\triangleright$-witness *in* $\mathcal{U}_{m+K-1}$*, i.e. a node* z *such that* $\mathbf{wit}(x, y, z)$ *holds.*

Proof: By the assumption $\neg(x \triangleright y)$, some node z satisfying $\mathbf{wit}x, y, z$ exists; it remains to show that z can be chosen in $\mathcal{U}_{n+K-1}$. If $x\#y$, we are done immediately, taking x as witness. Thus, suppose $\mathcal{C}_{xy} \triangleq \lceil x\rceil \cup \lceil y\rceil$ is a configuration; by the assumption, $\mathcal{C}_{xz}$ is also a configuration. Let M_{xy} be the marking generated by $\mathcal{C}_{xy}$. Choose $z \in E$ such that $\mathbf{wit}(x, y, z)$ holds, and such that $z' < z$ implies $\neg\mathbf{wit}(x, y, z')$. Then there exists $u \in E\backslash\{z\}$ such that *1)* ${}^\bullet u \cap {}^\bullet z \neq \emptyset$ and *2)* $u \leqslant y$. $\mathcal{C}^{zu} \triangleq \lceil {}^\bullet z\rceil \cup \lceil {}^\bullet u\rceil$ is a configuration by the choice of u and z. Let $M(zu)$ be the marking generated by $\mathcal{C}^{zu}$. Then by construction, for π the process label of the unfolding,

$$M(zu) \stackrel{\pi(z)}{\longrightarrow} \quad and \quad M(zu) \stackrel{\pi(u)}{\longrightarrow}, \tag{18}$$

$$and \qquad \mathcal{C}^{zu} \quad \sqsubseteq \quad \mathcal{C}_{xy}. \tag{19}$$

If $\max(round(u, z)) > n + K - 1$, the pigeonhole principle implies that there are two distinct configurations $\mathcal{C}_1, \mathcal{C}_2$ of $\mathcal{U}$ such that *(1)* $\lceil x\rceil \sqsubseteq \mathcal{C}_{xy} \sqsubseteq \mathcal{C}_1 \sqsubseteq \mathcal{C}_2 \sqsubseteq \mathcal{C}^{zu}$, and *(2)*$\mathcal{C}_1 \equiv_M \mathcal{C}_2$. We can then replace $\mathcal{C}_z$ by a different configuration $\mathcal{C}'$ that satisfies (18) and (19), obtained by 'removing' the section $\mathcal{C}_2\backslash\mathcal{C}_1$; in fact, $\mathcal{C}'$ shares $\mathcal{C}_1$ with $\mathcal{C}_z$ but follows the suffix of $\mathcal{C}_1$ isomorphic to the suffix of $\mathcal{C}_2$ in $\mathcal{C}_z$. Repeat this surgery until no such configurations $\mathcal{C}_1, \mathcal{C}_2$ can be found; the resulting configuration $\mathcal{C}'$ lies entirely within $\mathcal{U}_{n+K-1}$ by the definition of K. From (18), we also obtain the existence of an event e such that

$$\pi(e) = \pi(z) \quad \text{and} \quad \mathcal{C}' = \lceil e\rceil \cup \lceil u\rceil,$$

since u lies in $\mathcal{C}_{xy}$ and $M' \triangleq M(\mathcal{C}')$ satisfies $M' \stackrel{z}{\longrightarrow}$ and $M' \stackrel{u}{\longrightarrow}$. By construction, $e\#x$. We claim that

$${}^\bullet e \cap {}^\bullet u \quad = \quad {}^\bullet z \cap {}^\bullet u. \tag{20}$$

In fact, suppose there exists $b \in ({}^\bullet e\backslash{}^\bullet z) \cap {}^\bullet u$. By property *(ii)* of homomorphisms (Def. II), there must exist $b' \in {}^\bullet z \cap {}^\bullet u$ such that $\pi(b') = \pi(b)$. Then either *(i)* $b'\#b$, *(ii)* $b' < b$, *(iii)* $b < b'$ or *(iv)* b' **co** b. But *(i)* implies $b'\#b$; under *(ii)*, there must exist an event e_0 such that $b' < e_0 < b$, which also implies $b'\#b$; symmetrically, *(iii)* also leads to $b'\#b$; and *(iv)* contradicts Lemma 1.

Consider now the different possibilities for y; we have that:

- if $y\#u$ we are done;
- if $y < u$, then $y < x$, contradicting our assumption;
- If $u < y$, we obtain $z\#y$, another contradiction.

Therefore y **co** u must hold. If we assume now that $e\#y$, there must exist an event $v \leqslant y$ such that ${}^\bullet e \cap {}^\bullet v \neq \emptyset$. By reasoning along the same lines as for (20) above, we obtain

$${}^\bullet e \cap {}^\bullet v \quad = \quad {}^\bullet z \cap {}^\bullet v; \tag{21}$$

as a consequence, $z\#y$, contradicting our assumptions. Therefore $\neg(e\#y)$, and we are done. □

Lifting $\triangleright$ to $\mathcal{N}$: Consider again Fig. 1. Every occurrence of δ is detected by a prior occurrence of α, and by a subsequent occurrence of ζ. That is, if δ is a fault event, then it suffices for $\mathcal{N}$ to be weakly δ-diagnosable if either δ or α are observable. This can be formalized as a *lifting* of $\triangleright$ to the level of $\mathcal{N}$:

Definition 8 *In* $\mathcal{N}$*, transition* $t_1 \in T$ reveals $t_2 \in T$*, written* $t_1 \triangleright_\mathcal{N} t_2$*, iff for all* $e_2 \in \pi^{-1}(t_2)$ *there exists* $e_1 \in \pi^{-1}(t_1)$ *such that* $e_1 \triangleright e_2$*, where* $\triangleright$ *is the* reveals *relation in* $\mathcal{U}(\mathcal{N})$.

We have the following obvious result:

Lemma 6 *Let O be as above, and $\phi \in T\backslash O$.*

1) *If there exists $t \in O$ such that $t \triangleright_{\mathcal{N}} \phi$, then $\mathcal{N}$ is weakly ϕ-diagnosable, and*
2) *if for all $t \in T\backslash O$, there exists $t' \in O$ such that $t' \triangleright_{\mathcal{N}} t$, then $\mathcal{N}$ is weakly observable.*

However, the converse of statement 1 in lemma **??** is not true. In fact, consider again Fig. 1. We obtain the following table for $\triangleright_{\mathcal{N}}$ ('+' at (x, y) means that $x \triangleright_{\mathcal{N}} y$, and '-' means $x \not\triangleright_{\mathcal{N}} y$):

$\triangleright_{\mathcal{N}}$	α	β	γ	δ	η	ζ
α	+	−	+	−	−	−
β	−	+	+	−	−	−
γ	−	−	+	−	−	−
δ	+	−	+	+	−	+
η	−	+	+	−	+	−
ζ	+	−	+	+	−	+

Now, let γ be the fault transition; then $\mathcal{N}$ is weakly diagnosable for *any* choice of $O \subseteq T\backslash\{\gamma\}$, even $O = \emptyset$, since all maximal runs are faulty.

Dwelling on the example a little further , we see that if α is the fault transition, then it suffice to have either $\delta \in O$ or $\eta \in O$ to obtain weak diagnosabiliy. Then $\mathcal{N}$ is clearly β-diagnosable, yet γ is not $\triangleright_{\mathcal{N}}$-revealed by either α or β. We see that $\triangleright_{\mathcal{N}}$ gives sufficient criteria for observability and diagnosability, and allows quick verification of both, if $\triangleright_{\mathcal{N}}$ has been precomputed offline; on the other hand, it has in general to be checked on a prefix of the unfolding (rather than $\mathcal{N}$) whether a particular occurrence of a transition t is revealed by some observable event.

VII. Facets and $\mathcal{Q}$-Diagnosability

An occurrence net ON can be decomposed into equivalence classes w.r.t. $\triangleright$, called *facets*; see Fig. 5.

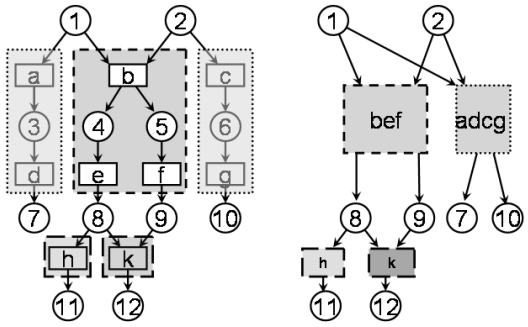


Fig. 5. Left: the example from Fig. 3 with facets highlighted; right: the occurrence net obtained from the example through facet abstraction

Definition 9 *A **facet** of ON is a strongly connected component of $\triangleright$, i.e. a maximal set $\psi \subseteq (E \cup B)$ such that for any $x, y \in \psi$, one has $x \triangleright y$ and $y \triangleright x$. Denote as $\psi(x)$ the unique facet that contains x.*

In Fig. 5, the facets are $\{a, d, c, g\}$, $\{b, e, f\}$, $\{h\}$, $\{k\}$; the right hand side shows the occurrence net obtained by abstracting every facet into a single event. Concerning the shape of facets, we obtain easily:

Lemma 7 *Facets are conflict-free.*

Proof: Suppose $e_1 \# e_2$; then $\#[e_1]\backslash\#[e_2] = \{e_1\}$ and $\#[e_2]\backslash\#[e_1] = \{e_2\}$, so neither $e_1 \triangleright e_2$ nor $e_2 \triangleright e_1$; e_1 and e_2 cannot belong to the same facet. □

More is true:

Lemma 8 *Facets are convex, i.e. $x, y \in \psi$ and $x < y < z$ together imply $z \in \psi$.*

Proof: By assumption, $\#[x] = \#[y]$; by Lemma 4, $\#[x] \subseteq \#[z] \subseteq \#[y]$, hence $\psi(x) = \psi(y) = \psi(z)$. □

Lemma 9 *For all $b \in B$, $b^\bullet \cap \psi(b) \neq \emptyset \Rightarrow |b^\bullet| = 1$.*

Proof: Suppose $|b^\bullet| \geqslant 1$. If $b^\bullet \cap \psi(b) \neq \emptyset$, then $\psi(b)$ contains a conflict pair, contradicting Lemma 8. So assume $e_1 \in b^\bullet \cap \psi(b)$ and $e_2 \in b^\bullet\backslash\psi(b)$. But then $e_2 \in \#[e_1]\backslash\#[]b)$, hence $\psi(b) \neq \psi(e_1)$, contradicting the assumption. □

A consequence of Lemma 9 is that maximal nodes in a facet are conditions.

Facets are Abstractions: We observe that the set of facets carries an induced event structure: let x_i be a node of ON, let $\psi_i \triangleq \psi(x_i)$, and set

$$\psi_1 \prec_\Psi \psi_2 \iff \left\{ \begin{array}{ll} & \psi_1 \neq \psi_2 \\ \exists & y_1 \in \psi_1, y_2 \in \psi_2 : \\ & y_1 < y_2 \end{array} \right. \quad (22)$$

$$\psi_1 \#_\Psi \psi_2 \iff [\exists\ y_1 \in \psi_1, y_2 \in \psi_2 :\ y_1 \# y_2] \quad (23)$$

In fact, relation $\prec_\Psi$ from (22) is a partial order by Lemma 8; $\#_\Psi$ is well-defined since $y_1 \# y_2$ implies $z_1 \# z_2$ for all z_1 from ψ_1 and z_2 from ψ_2, and since facets are conflict-free by Lemma 7.

One checks easily that $\psi_1 \# \psi_2 \prec_\Psi \psi_3$ implies $\psi_1 \# \psi_3$, and finds that $\mathcal{F} = (\Psi, \prec_\Psi, \#_\Psi)$ is an event structure in the sense of Definition 2. In fact, contracting every facet ψ into single events e_ψ whose output conditions are the maximal conditions of ψ, and whose input conditions are given by the pre-conditions of the minimal events in ψ, we obtain a reduced occurrence net $ON_{/_\Psi}$, see Fig. 5; below we will see that this abstraction operation preserves and respects runs. Let

$$\lceil \psi \rceil \triangleq \{\psi' \mid \psi' \prec_\Psi \psi\}.$$

By Lemma 8, the set union of all facets in $\lceil \psi \rceil$ spans a configuration of ON; we denote it by

$$\mathcal{C}(\psi). \quad (24)$$

Theorem 5 *$\omega_\Psi \subseteq \Psi$ is a run of $\mathcal{F} = (\Psi, \prec_\Psi, \#_\Psi)$ iff*

$$\omega_{\omega_\Psi} \triangleq \bigcup_{\psi \in \omega_\Psi} \psi \quad (25)$$

is a run of $\mathcal{E} = (E, \leqslant, \#)$.

Proof: First, assume ω_Ψ is a run of $\mathcal{F} = (\Psi, \prec_\Psi, \#_\Psi)$; then ω_{ω_Ψ} according to (25) is a configuration of $\mathcal{E} = (E, \leqslant, \#)$ by the above. Suppose ω_{ω_Ψ} is not maximal, and let $e \notin \omega_{\omega_\Psi}$ be such that $\lceil e \rceil \cup \omega_{\omega_\Psi}$ is a configuration. Then, by Lemma 3, the same is true for all $e' \in \psi(e)$, which contradicts maximality of ω_Ψ. Conversely, if ω_{ω_Ψ} is a run of $\mathcal{E} = (E, \leqslant, \#)$, assume there exists $\psi \notin \omega_\Psi$ such that $\lceil \psi \rceil \cup \omega_\Psi$ is a configuration in $\mathcal{F} = (\Psi, \prec_\Psi, \#_\Psi)$; then for any $e \in \psi$, we have that $\lceil e \rceil \cup \omega_{\omega_\Psi}$ is a configuration, contradicting the assumption. □

Theorem 6 *Let* $ON = (B, E, G, \mathbf{c}_0)$ *be an occurrence net, and* Ψ *its set of facets. Set* $\mathbf{c}_{0/_\Psi} \triangleq \mathbf{c}_0$ *and*

$$\begin{array}{rcl} E_{/_\Psi} & \triangleq & \Psi,\ B_{/_\Psi} \triangleq \mathbf{c}_0 \cup \{b \mid b^\bullet \cap \psi(b) = \emptyset\} \\ G_{/_\Psi} & \triangleq & \big\{(b,\psi) \in B_{/_\Psi} \times E_{/_\Psi} \mid b^\bullet \cap \psi(b) \neq \emptyset\big\} \\ & & \cup \big\{(\psi, e) \in B_{/_\Psi} \times E_{/_\Psi} \mid b^\bullet \cap \psi(b) = \emptyset\big\}; \end{array}$$

Then $ON_{/_\Psi} = (B_{/_\Psi}, E_{/_\Psi}, G_{/_\Psi}, \mathbf{c}_{0/_\Psi})$ *is an occurrence net.*

Proof: Note that the relation $(G_{/_\Psi})^2$ coincides with the immediate successor relation of $\mathcal{F}$. It therefore remains to show that

1) $ON_{/_\Psi}$ is a net, and
2) there is no backward branching;

then the induced relations on $E_{/_\Psi}$ can be easily seen to agree with those in $\mathcal{F}$, and we are done. For 1), disjointness and non-emptyness of $E_{/_\Psi}$ and $b_{/_\Psi}$ are immediate; by construction, $G_{/_\Psi} \subseteq (b_{/_\Psi} \times E_{/_\Psi}) \cup (E_{/_\Psi} \times b_{/_\Psi})$. To see 2), assume $G_{/_\Psi}$ contains two arcs $(e_{1/_\Psi}, b_{/_\Psi})$ and $(e_{2/_\Psi}, b_{/_\Psi})$ such that $e_{1/_\Psi} \neq e_{2/_\Psi}$. Then there must exist (in ON) $e_1' \in \psi(e_1)$ and $e_2' \in \psi(e_2)$ such that $b \in {e_1'}^\bullet \cap {e_2}^\bullet$, and moreover $e_1' \neq e_2'$ since facets are pairwise disjoint by construction; but then ON contains already a backward branching, a contradiction. □

$\mathcal{Q}$-Diagnosability: With the same setting and notations, define the *pro-cone* of a node $x \in E \cup B$ as

$$\lceil\lceil x \rceil\rceil \triangleq \mathcal{C}(\psi(x)); \tag{26}$$

the *closure* of a configuration $\mathcal{C}$ is defined as

$$\lceil\lceil \mathcal{C} \rceil\rceil \triangleq \bigcup_{x \in \mathcal{C}} \lceil\lceil x \rceil\rceil. \tag{27}$$

Configuration $\mathcal{C}$ is *closed* iff $\lceil\lceil \mathcal{C} \rceil\rceil = \mathcal{C}$. Note that $\lceil\lceil \mathcal{C} \rceil\rceil$ coincides with the configuration obtained by intersecting all runs that extend $\mathcal{C}$. One obtains closed configurations of ON as the configurations of the *facet* event structure $(\Psi, \prec_\Psi, \#_\Psi)$:

Lemma 10 *The configurations of* $ON_{/_\Psi}$ *correspond one-to-one to the closed configurations of* ON.

We are now ready to give the definition of *$\mathcal{Q}$-diagnosability*:

Definition 10 *If* ON *satisfies* $\mathcal{WOBS}$ *w.r.t.* E_O*, then is* $\mathcal{Q}$-diagnosable *w.r.t.* ϕ *iff for all configurations* $\mathcal{C}, \mathcal{C}'$*,*

$$\left.\begin{array}{l} \lceil\lceil \mathcal{C} \rceil\rceil \sim_O \lceil\lceil \mathcal{C}' \rceil\rceil \\ \wedge\ \lceil\lceil \mathcal{C} \rceil\rceil \equiv_M \lceil\lceil \mathcal{C}' \rceil\rceil \end{array}\right\} \ \Rightarrow\ \lceil\lceil \mathcal{C} \rceil\rceil \equiv_\Phi \lceil\lceil \mathcal{C}' \rceil\rceil. \tag{28}$$

In words, ON is $\mathcal{Q}$-diagnosable iff for any two configurations $\mathcal{C}, \mathcal{C}'$ the following holds: if the *inevitable common parts* $\lceil\lceil \mathcal{C} \rceil\rceil$ or $\lceil\lceil \mathcal{C}' \rceil\rceil$ of all runs that extend $\mathcal{C}$ or $\mathcal{C}'$, respectively, produce the same observations and the same marking, they have to be also fault equivalent. Note that this definition is less restrictive than the one from [15] since it only applies to marking equivalent pairs. We observe that $\mathcal{Q}$-diagnosability includes both diagnosis of the past as 'prediction' of concurrent or future events. This notion of diagnosis is thus well adapted to asynchronous systems where the the precise interleaving of events is not available; concurrent events will occur and go unnoticed *unless* they change future branchings.

Verification of $\mathcal{Q}$-diagnosability for ON reduces - under some simplifying assumptions - to verification of weak diagnosability for $ON_{/_\Psi}$:

Theorem 7 *Assume that* ON *and* E_O *are such that for every facet* ψ *of* ON*,* $|\psi \cap E_O| \in \{0, 1\}$*, and that* $\psi \cap E_\phi = \psi \cap E_O = \emptyset$*. Define* $\lambda_{/_\Psi} : \Psi \to \mathfrak{A}$ *by setting*

$$\lambda_{/_\Psi}(\psi) \triangleq \begin{cases} \lambda(\pi(e)) & : \ \psi \cap E_O = \{e\} \\ \varepsilon & : \ \psi \cap E_O = \emptyset \end{cases}.$$

Further, let $\Psi_\phi \triangleq \{\psi \in \Psi(ON) \mid E_\phi \cap \psi \neq \emptyset\}$*. Then* ON *is* $\mathcal{Q}$*-diagnosable for* ϕ *iff* $ON_{/_\Psi}$ *is weakly diagnosable for* ϕ*.*

Proof: Suppose first that ON is $\mathcal{Q}$-diagnosable for E_O and $\Phi_{/_\Psi}$, and that $ON_{/_\Psi}$ is not weakly diagnosable. Then by Theorem 1, there exist configurations $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_1', \mathcal{C}_2'$ of $\mathbb{L}_{\mathbf{prog}}(ON_{/_\Psi})$ such that *(1)* $\mathcal{C}_1 \neq \mathcal{C}_1'$ and $\mathcal{C}_1 \sqsubseteq \mathcal{C}_1'$, $\mathcal{C}_2 \sqsubseteq \mathcal{C}_2'$, *(2)* $\mathcal{C}_1 \sim_O \mathcal{C}_2$ and $\mathcal{C}_1' \sim_O \mathcal{C}_2'$, *(3)* $\mathcal{C}_1 \equiv_M \mathcal{C}_1'$ and $\mathcal{C}_2 \equiv_M \mathcal{C}_2'$, but *(4)* $\mathcal{C}_1'$ contains ϕ while $\mathcal{C}_2'$ does not. But then the configurations $\mathcal{C}(\mathcal{C}_1), \mathcal{C}(\mathcal{C}_1'), \mathcal{C}(\mathcal{C}_2), \mathcal{C}(\mathcal{C}_2') \in \mathbf{Con}(ON_{/_\Psi})$ according to Lemma 10 constitute a counterexample to $\mathcal{Q}$-diagnosability of ON. The reverse implication follows from Lemma 10. □

Note that the assumption of only one observable event per facet is made here only to make the presentation simpler; in the general case, a more sophisticated labelling must be devised so that a generalization of Theorem 7 can hold, see [16].

Depending on the particular net under study, the facet net can be considerably smaller than the original unfolding; in some cases, it might be efficient to synthesize a generating Petri net from the quotient unfolding, and perform the diagnosis (or other analysis) on that net rather the original one. We think the tradeoff between this offline effort and the online complexity should be weighed carefully, as some nets will allow great reductions and speedup by quotienting, while for others there is no gain at all.

VIII. Conclusion

We have shown how the problem of diagnosability splits into several variants in the context of true concurrency in asynchronous systems. Characterizations of weak and strong diagnosability have been given. Investigating the relational structure of occurrence nets, for the purpose of finer analysis of observability and weak diagnosability, has lead us to the *reveals*-relation $\triangleright$ and the associated decomposition of occurrence nets into facets. We have seen that $\triangleright$ can

be effectively computed on sufficiently large prefixes, and that facets are adequate abstractions for preserving maximal nonsequential behaviour. The analysis of the nets obtained by facet abstraction, and their properties in terms of diagnosis, is an interesting new field. As noted above, knowledge of facets allows for *prediction* into the future. Obviously, the prognostic capacity of diagnosis using $ON_{/\Psi}$ depends directly on the size of ON's facets: the gain will thus be strongest in systems with a *high* degree of concurrency and a low to moderate degree of branching. It remains to optimize the exploration of the data structures of $\mathcal{U}(\mathcal{N})$ and Ψ for efficient verification of diagnosability. Computing the $\triangleright$-relation is polynomial in the size of $\mathcal{U}_2(\mathcal{N})$; on the other hand, the worst case size of $\mathcal{U}_2(\mathcal{N})$ is exponential in the size of P. However, many systems for which modeling with Petri nets is well suitable - namely highly distributed and asynchronous systems -, generally yield an order 2 unfolding of reasonable size.

Generally speaking, strong diagnosability is a notion inherited from sequential systems, while weak diagnosability and $\mathcal{Q}$-diagnosability are genuinely asynchonous properties with no sequential equivalent. The link between weak diagnosability and $\mathcal{Q}$-diagnosability is explicited by Theorem 7.

REFERENCES

[1] S. Abbes and A. Benveniste. True-concurrency probabilistic models: Branching cells and distributed probabilities for event structures. *Information and Computation* **204** (2) pp. 231–274, 2006.

[2] C.G. Cassandras and S. Lafortune. Introduction to Discrete Event Systems. Kluwer Academic Publishers, Boston etc, 1999.

[3] J. Desel and J. Esparza. Free Choice Petri Nets. *Cambridge Tracts in Theor. Comp. Sci.* vol. **40**, Cambridge Univ. Press, 1995.

[4] V. Diekert, G. Rozenberg (eds).*The Book of Traces* .World Sci., 1995

[5] J. Engelfriet. *Branching Processes of Petri Nets.* Acta Informatica **28**:575–591, 1991.

[6] J. Esparza, S. Römer, and W. Vogler. An improvement of McMillan's unfolding algorithm. *Form. Meth. in Syst. Des.* **20**(3):285-310, 2002.

[7] E. Fabre and A. Benveniste. Partial Order Techniques for Distributed Discrete Event Systems: why you can't avoid using them. Discrete Event Dynamic Systems, 2007 (17), 355-403.

[8] E. Fabre, A. Benveniste, C. Jard, and S. Haar. Diagnosis of Asynchronous Discrete Event Systems, a Net Unfolding Approach. *IEEE Trans. Aut. Control* **48**(5)714–727, May 2003.

[9] E. Fabre, A. Benveniste, C. Jard, and S. Haar. Distributed monitoring of concurrent and asynchronous systems. *Discrete Event Dynamic Systems* **15**(1):33–84, Mar. 2005

[10] A. Giua and C. Seatzu. Observability of Place/Transition Nets. *IEEE Trans. Aut. Control* **47**(9):1424–1437, 2002.

[11] M.P. Cabasino, A. Giua, S. Lafortune, C. Seatzu. Diagnosability analysis of unbounded Petri nets. CDC09: 48th IEEE Conf. on Decision and Control (Shanghai, China), December 2009.

[12] M.P. Cabasino, A. Giua, C. Seatzu. Diagnosability of bounded Petri nets. CDC09: 48th IEEE Conf. on Decision and Control (Shanghai, China), December 2009.

[13] A. Giua and C. Xie. Control of safe ordinary Petri nets using unfolding. *Discrete Event Dynamic Systems* **15**(4):349–373, Dec. 2005.

[14] S. Haar, A. Benveniste, E. Fabre, and C. Jard. Partial Order Diagnosability of Discrete Event Systems Using Petri Net Unfoldings. *Proc. of 42nd IEEE Conference on Decision and Control (CDC)*, 2003.

[15] S. Haar. Unfold and Cover: Qualitative Diagnosability for Petri Nets. Proc. 46th IEEE Conference on Decision and Control, 2007.

[16] S. Haar. Qualitative Diagnosability for Petri Nets revisited. Proc. 48th IEEE Conference on Decision and Control, 2009.

[17] S. Haar. Diagnosability and Branching Process Semantics. In: *Object Petri Nets, Processes, and Object Calculi*. Festschrift for R. Valk. Bericht (tech. report) **265**, pp.13–34, FB Informatik, University of Hamburg.

[18] L.E. Holloway, B.H. Krogh and A. Giua. A Survey of Petri Net Methods for Controlled Discrete event systems. *Discrete Event Dynamic Systems: Theory and Applications* **7**:151–190, 1997.

[19] S. Jiang, Z. Huang, V. Chandra and R. Kumar. A Polynomial Time Algorithm for Diagnosability of Discrete Event Systems.

[20] V. Khomenko, M. Koutny, and W. Vogler. Canonical Prefixes of Petri Net Unfoldings. *Acta Informatica* **40**:95–118, 2003.

[21] R. Kumar and M.A. Shayman. Formulae relating Controllability, Observability, and Co-Observability. *Automatica* **34** (2), 211–215, 1998.

[22] R. Kummetz and D. Kuske. The topology of Mazurkiewicz Traces. *Theor. Comp. Sci.* **305**:237–258, 2003.

[23] M.Z. Kwiatkowska. A Metric for Traces. *Inf. Proc. Lett.* **35**:129–135.

[24] F. Lin. Diagnosability of discrete event systems and its applications. *Discrete Event Dynamic Systems*, 4(1), 1994, pp. 197-212.

[25] K. McMillan. Using Unfoldings to avoid the state explosion problem in the verification of asynchronous circuits. *4th Workshop on Computer Aided Verification* 164–174, 1992.

[26] T. Murata. Petri Nets: Properties, Analysis and Applications. *Proc. of the IEEE* vol. **77** no 4, April 1989.

[27] M. Nielsen, G. Plotkin G. Winskel. Petri nets, event structures, and domains, Part I. *TCS* **13**:85–108, 1981.

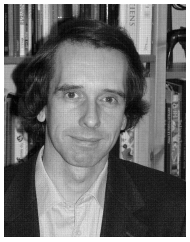
[28] J.L. Peterson. *Petri Net Theory and the Modeling of Systems.* Prentice-Hall, 1981.

[29] W. Reisig. *Petri nets.* Springer Verlag, 1985.

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

[31] G. Winskel. Event structures. *Advances in Petri nets,* LNCS **255**: 325–392, Springer Verlag, 1987.

[32] T. Yoo and S. Lafortune. Polynomial-Time Verification of Diagnosability of Partially-Observed Discrete-Event Systems. *IEEE Trans. Aut. Control* **47**(9):1491-1495 , 2002.



Stefan Haar studied mathematics at Hamburg University, Germany, and Johns Hopkins University, Baltimore, MD. He received the *Diplom* (M. Sc.) in mathematics - specializing in stochastic processes - and PhD degrees in computer science (Petri net theory) at Hamburg University. After post-doctoral positions in Berlin, Nancy and Paris, he has been a researcher with INRIA from 2001. Following a sabbatical stay at University of Ottawa, Canada, and with ALU Bell Labs Ottawa, he returned to France in 2008 and moved to INRIA Saclay, where he now directs the *MExICo* research team at *Ecole Normale de Cachan.* His research interests include Petri nets, event structures and graph grammars, conformance testing and probabilistic discrete event systemsfault diagnosis, quality of service contracts, and web service orchestration; he was an associate editor of IEEE Transactions of Automatic Control from 2005 through 2009.