

Adding uncertainty: unobservable events and observers for finite state automata and PNs

From observability to privacy and security in discrete event systems

Prof. Gianmaria DE TOMMASI

Email: detommas@unina.it

December 2020

- 1 Discrete Event Systems (DES), Languages and Automata
- 2 Petri nets (PNs) and their twofold representation to model DES
- 3 MILP and ILP formulations: logical conditions, binary variables “do everything”, and variable connecting
- 4 **Adding uncertainty: unobservable events and observers for finite state automata and PNs**
- 5 Augmenting the observers: diagnosability of prefix-closed languages, diagnosers and the fault detection for finite state automata
- 6 Diagnosability and fault detection in PNs - Part I: graph-based approaches
- 7 Diagnosability and fault detection in PNs - Part II: algebraic approaches for bounded systems
- 8 Security issues in DES: non-interference and opacity
- 9 Non-interference and opacity enforcement
- 10 Open issues

- 1** The automata case
 - Source of nondeterminism in DES modelled as logic automata
 - Nondeterministic automata
 - Observer automata

- 2** The Petri nets case
 - Source of nondeterminism in DES modelled as Petri nets
 - Observer coverability graph
 - State estimation in labeled net systems

- The primary source of **nondeterminism** is the limitations of the sensors attached to the system
- This results in **unobservable events** that causes a change in the state that cannot be directly *measured*
- From the point of view of an *external observer*, the occurrence of an unobservable event is equivalent to the occurrence of the silent event ε

- Another way to model uncertainty about the system behaviour can be the **lack of knowledge about the initial state**
- Sometime it is assumed that the initial state of a DES is **one among a set of states**



- There can be also **uncertainty on the effects** due to the occurrence of an event. . .
- . . . or **uncertainty due to undistinguishable events**
- Both sources of uncertainty can be modelled as an event that, from a given state x , can cause transitions to more than one state
- In this case **the state transition function becomes nondeterministic**

$$f : X \times E \mapsto 2^X$$

- When unobservable events are used to model the uncertain system, we can assume that $E = E_o \cup E_{uo}$ with
 - E_o the set of **observable** events
 - E_{uo} the set of **unobservable** events
 - $E_o \cap E_{uo} = \emptyset$
- For an *external* observer the occurrence of and event $e \in E_{uo}$ is equivalent to the occurrence of ε
- The projection function can be used to *filter out* the unobservable events from the words generated by the system

Projection

$$Pr : E^* \mapsto E_o^*$$

$$\left\{ \begin{array}{ll} Pr(\varepsilon) := \varepsilon & \\ Pr(e) := e & \text{if } e \in E_o \\ Pr(e) := \varepsilon & \text{if } e \in E_{u0} \\ Pr(we) := Pr(w)Pr(e) & w \in E^*, e \in E \end{array} \right.$$

- Given a word $w \in E^*$ generated by the uncertain model, its protection $Pr(w) \in E_o^*$ represents what an *external* observer can measure

- **Nondeterministic automata** permit to take into account all the sources of uncertainty that have been introduced so far
- A nondeterministic automata (**NDA**) is defined a 6-ple

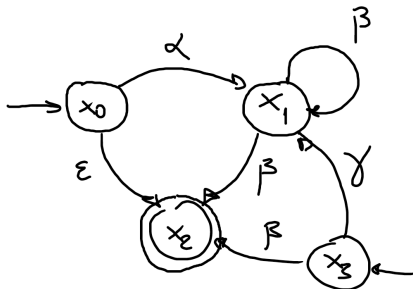
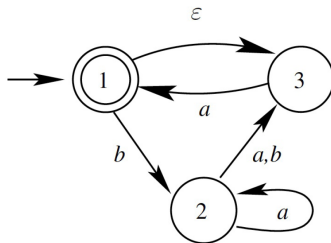
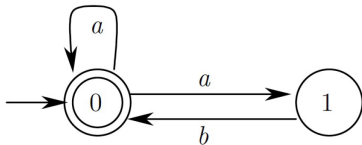
$$G_{nd} = (X, E \cup \{\varepsilon\}, f_{nd}, \Gamma, x_0, X_m)$$

- The silent event ε is included in the set of events that drive the systems dynamic
- The transition function is defined as

$$f_{nd} : X \times E \cup \{\varepsilon\} \mapsto 2^X$$

that is $f_{nd}(x, e) \subseteq X$, when defined (uncertainty on the *consequences* of a given event)

- The initial state may be itself a set of states, that is $x_0 \subseteq X$



Extending the transition function to the nondeterministic case

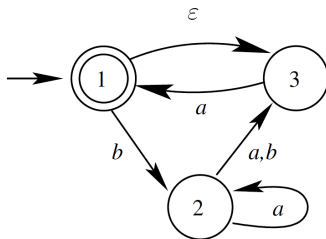
- For (logic) deterministic automata it is $f(x, \varepsilon) = x$ (in the deterministic case ε is used as *empty string*, rather than silent event)
- ε -reach of a state x

$\varepsilon R(x) = \{\text{all the states that can be reached from } x \text{ following a silent transition}\}$

- **By definition it is** $x \in \varepsilon R(x)$
- If $B \in X$, then

$$\varepsilon R(B) = \bigcup_{x \in B} \varepsilon R(x)$$

- It is then possible to extend f_{nd} as
 - $f_{nd}^{ext}(x, \varepsilon) := \varepsilon R(x)$
 - $f_{nd}^{ext}(x, we) := \varepsilon R(\{z \in X \mid z \in f_{nd}(y, e) \text{ for some } y \in f_{nd}^{ext}(x, w)\})$
with $w \in E^*$ and $e \in E$
- In general it is $f_{nd}(x, e) \subseteq f_{nd}^{ext}(x, e)$ with $e \in E \cup \{\varepsilon\}$



- $f_{nd}(1, \varepsilon) = \{3\}$; $f_{nd}^{ext}(1, \varepsilon) = \{1, 3\}$
- $f_{nd}(3, a) = \{1\}$; $f_{nd}^{ext}(3, a) = \{1, 3\}$
- $f_{nd}(3, a) = \{1\}$; $f_{nd}^{ext}(3, a) = \{1, 3\}$
- $f_{nd}(2, a) = f_{nd}^{ext}(2, a) = \{2, 3\}$
- $f_{nd}(2, b) = f_{nd}^{ext}(2, b) = \{3\}$
- $f_{nd}(1, bba) = \{1\}$; $f_{nd}^{ext}(1, bba) = \{1, 3\}$

Given the notion of **extended transition function** f_{nd}^{ext} , it is possible to define the languages generated and marked by a NDA

Language generated by $G_{nd} - \mathcal{L}(G_{nd})$

$$\mathcal{L}(G_{nd}) = \{w \in E^* \mid \exists x \in x_0 \text{ s.t. } f_{nd}^{ext}(x, w) \text{ is defined}\}$$

Language marked by $G_{nd} - \mathcal{L}_m(G_{nd})$

$$\mathcal{L}_m(G_{nd}) = \{w \in \mathcal{L}(G_{nd}) \mid \exists x \in x_0 \text{ s.t. } f_{nd}^{ext}(x, w) \cap X_m \neq \emptyset\}$$

Logic nondeterminism vs stochastic nondeterminism

- Here we are dealing with nondeterminism in the context of **logic automata**
- Nondeterminism can be associated also to the timing of event occurrences
- The inclusion of this further source of nondeterminism calls for the use of **stochastic models**...
 - Stochastic automata
 - Generalized Semi-Markov Process
 - Markov chains
 - ...
- ...which are out of the scope of these lectures :)

- The **observer** is a **deterministic** automaton that is equivalent to a given NDA
 - **Equivalence in terms of languages**
- If the NDA has finite state space, then also the observer will be a FSM
- The observer allows us to estimate the state of a NDA
- First results for fault detection have been obtained by extending the concept of observer
 - **Be patient and wait for Lecture #5 by Prof. Basile**



M. Sampath et al.

Diagnosability of Discrete-Event Systems

IEEE Transactions on Automatic Control, 1995

Let $G_{nd} = (X, E \cup \{\varepsilon\}, f_{nd}, x_0, X_m)$ be a NDA. Its observer is the deterministic automaton

$$Obs(G_{nd}) = (X_{obs}, E, f_{obs}, x_{0,obs}, X_{m,obs})$$

where

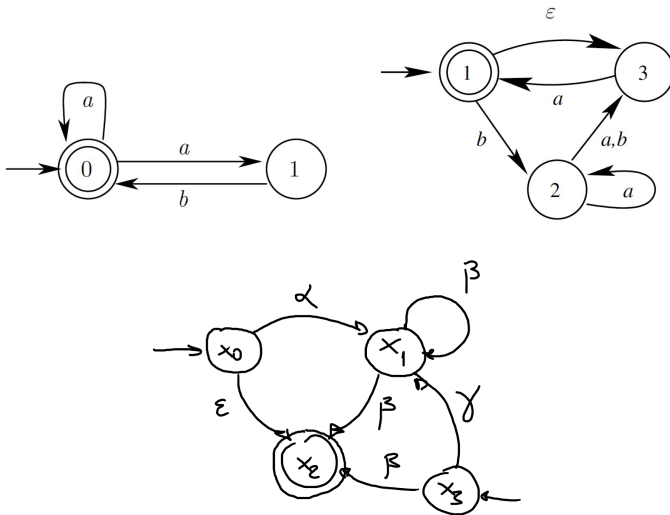
- $x_{0,obs} := \varepsilon R(x_0)$
- For each $B \in X_{obs}$ and $e \in E$, the transition function of the observer is defined as

$$f_{obs}(B, e) := \varepsilon R(\{x \in X \mid \exists x_e \in B \text{ s.t. } x \in f(x_e, e)\})$$

therefore the state $f_{obs}(B, e)$ is included in X_{obs}

- $X_{m,obs} := \{B \in X_{obs} \mid B \cap X_m \neq \emptyset\}$

Let's try to build the observer for these NDA!



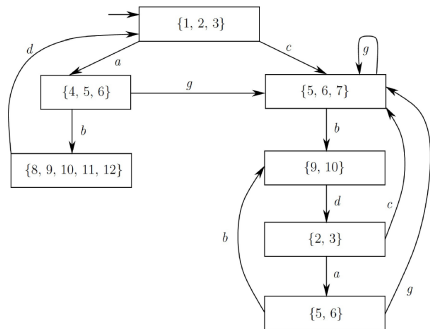
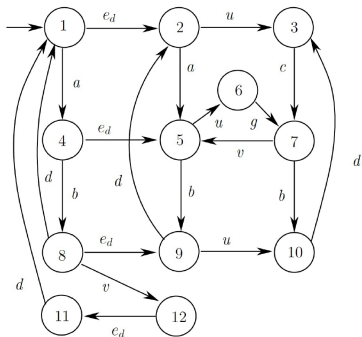
- Given a NDA G_{nd} with finite state space, its observer $Obs(G_{nd})$
 - has finite state space as well
 - is **equivalent** to G_{nd}
- Indeed, by definition it is
 - $\mathcal{L}(G_{nd}) = \mathcal{L}(Obs(G_{nd}))$
 - $\mathcal{L}_m(G_{nd}) = \mathcal{L}_m(Obs(G_{nd}))$
- **Finite-state NDA speaks regular languages as deterministic FSM**

Observer of deterministic automata with unobservable events



- The observer can be used to estimate the state of a **partially observed deterministic** automaton, i.e. a deterministic automaton with $E = E_o \cup E_{uo}$
- It is sufficient to *replace* the unobservable events with the silent transition \rightarrow a NDA is derived from the deterministic automaton with unobservable events

Set of unobservable events $E_{uo} = e_d, u, v$



- 1 **Unknown initial marking** (state) – it applies also to unlabeled PNs
- 2 ε -free (λ -free) PNs \rightarrow **partial knowledge of the system dynamic** or **undistinguishable events**
- 3 Unlabeled PNs with unobservable transitions \rightarrow the unobservable transitions are mapped on the silent event $\varepsilon \rightarrow$ **lack of sensors**
- 4 Arbitrarily labeled PNs \rightarrow both **2** and **3**

- For bounded PNs with *relatively* small reachability set $R(N, \mathbf{m}_0)$, state estimation can be achieved by building the observer of the **nondeterministic reachability graph**
- However, specific approaches have been developed for PNs

The **observer coverability graph (OCG)** can be built to estimate the marking m of an unlabeled PNs under the following. . .

Assumptions

- 1 The initial marking m_0 of the system is **completely unknown** (the only considered source of nondeterminism)
- 2 The structure of the net $N = (P, T, \mathbf{Pre}, \mathbf{Post})$ is known
- 3 All the transition occurrences can be observed \rightarrow the system is unlabeled and $T_{uo} = \emptyset$

The OCG has been proposed in



A. Giua and C. Seatzu

Observability of Place/Transition Nets

IEEE Transactions on Automatic Control, 2002

If $\sigma = t^1 t^2 \dots$ is a sequence enabled under the (unknown) initial marking, i.e. $\mathbf{m}_0[\sigma]$, then the following algorithm can be used to estimate the marking

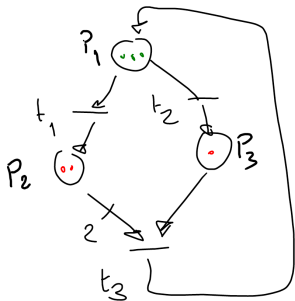
Marking estimation by using the observation of the transition occurrences

- 1 Let the initial estimate be $\mu_0 = \mathbf{0}$
- 2 Let $i = 1$
- 3 Wait until t^i fires
- 4 Set μ'_i equal to

$$\mu'_i(p) = \max\left(\mu_{i-1}(p), \mathbf{Pre}(p, t^i)\right), \quad \forall p \in P$$

- 5 Let $\mu_i = \mu'_i + \mathbf{C}(\cdot, t^i)$
- 6 Let $i = i + 1$
- 7 Goto step 3

Example of marking estimation



UNKNOWN INITIAL MARKING

$$m_0 = (3 \ 2 \ 1)^T$$

INITIAL ESTIMATION

$$M_0 = (0 \ 0 \ 0)^T$$

OBSERVED SEQUENCE

$$\sigma = t_3 t_2$$

$$M'_1 = (0 \ 2 \ 1)^T \Rightarrow M_1 = (1 \ 0 \ 0)^T$$

$$M'_2 = M_1 = (1 \ 0 \ 0)^T \Rightarrow M_2 = (0 \ 1 \ 0)^T$$

$$m_0 [\sigma > m \text{ with } m = (3 \ 1 \ 0)^T$$

Given a sequence $\sigma \in L(N, \mathbf{m}_0)$ with $\mathbf{m}_0[\sigma] \mathbf{m}$ and let μ be the marking estimate built by means of the proposed algorithm **after the occurrence of σ** , then the following two definitions can be given

p -complete sequence

Given $p \in P$, the sequence σ is said to be p -complete if $\mu(p) = \mathbf{m}(p)$

Marking complete sequence

The sequence σ is said to **marking complete** if it is p -complete for all $p \in P$

Marking observability

A system $\mathcal{S} = \langle N, \mathbf{m}_0 \rangle$ is said to be **Marking Observable (MO)** if there exists a marking complete sequence $\sigma \in L(N, \mathbf{m}_0)$

Strong marking observability

A system $\mathcal{S} = \langle N, \mathbf{m}_0 \rangle$ is said to be **Strongly Marking Observable (SMO)** in k steps, if

- $\forall \sigma \in L(N, \mathbf{m}_0)$ such that $|\sigma| \geq k$, σ is marking complete (*every sufficiently long sequence is marking complete*)
- $\forall \sigma \in L(N, \mathbf{m}_0)$ such that $|\sigma| < k$, either σ is marking complete or $\exists t \in T$ such that $\mathbf{m}_0[\sigma t)$ (*short and non marking complete sequences can be always extended to marking complete ones*)

Uniform marking observability

A system $\mathcal{S} = \langle N, \mathbf{m}_0 \rangle$ is said to be **Uniformly Marking Observable (uMO)** if $\forall \mathbf{m} \in R(N, \mathbf{m}_0)$ the system $\langle N, \mathbf{m} \rangle$ is MO

Uniform strong marking observability

A system $\mathcal{S} = \langle N, \mathbf{m}_0 \rangle$ is said to be **Uniformly Strongly Marking Observable (uSMO)** in k steps, if $\forall \mathbf{m} \in R(N, \mathbf{m}_0)$ the system $\langle N, \mathbf{m} \rangle$ is SMO in k steps

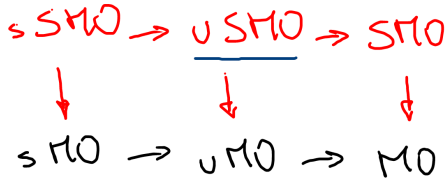
Structural marking observability

A net N is said to be **Structurally Marking Observable (sMO)** if $\langle N, \mathbf{m}_0 \rangle$ is MO $\forall \mathbf{m}_0 \in \mathbb{N}^m$

Structural strong marking observability

A net N is said to be **Structurally Strongly Marking Observable (sSMO)** in k steps, if $\langle N, \mathbf{m}_0 \rangle$ is SMO in k steps $\forall \mathbf{m}_0 \in \mathbb{N}^m$ (with k dependent on \mathbf{m}_0)

Relationship between observabilities in PNs



$USMO$ guarantees that it is possible to estimate the current marking even if we start to observe the system not since the "beginning", i.e., since any $m \in R(N, m_0)$

- The **Observer coverability graph (OCG)** permits to represent both the set of reachable markings of a net system (also unbounded), and an upper bound for the estimation error computed in accordance with the proposed algorithm
- Similarly to the coverability graph, the construction of the OCG is based the observer coverability **tree**

Observer coverability tree (taken from Giua & Seatzu, *IEEE TAC 2002*)

Algorithm 21 (Observer Coverability Tree)

1. Let $u_0 = M_0$. Label the initial node (M_0/u_0) as the root and tag it "new".
2. If "new" nodes exist, select a new node (M/u) and:
 - 2.1. If (M/u) is identical to a node labeled "old" then tag (M/u) "old" and go to step 2.
 - 2.2. If no transitions are enabled at M , tag (M/u) "dead" and go to step 2.
 - 2.3. For each transition t enabled at M do the following:
 - 2.3.1. $\forall p \in P$, if $M(p) = \omega$ then let $\tilde{M}(p) = M(p)$ and $\tilde{u}(p) = u(p)$, else let $\tilde{M}(p) = M(p) + C(p,t)$ and $\tilde{u}(p) = \min\{u(p), M(p) - Pre(p,t)\}$;
 - 2.3.2. on the path from the root to (M/u) if there exists a marking $\tilde{M} \leq \tilde{M}$ and $\tilde{M} \neq \tilde{M}$, i.e., \tilde{M} is covered by \tilde{M} , then let $\tilde{M}(p) = \omega$ for each p such that $\tilde{M}(p) > \tilde{M}(p)$;
 - 2.3.3. introduce (\tilde{M}/\tilde{u}) as a node, draw an arc with label t from (M/u) to (\tilde{M}/\tilde{u}) , and tag (\tilde{M}/\tilde{u}) "new".
 - 2.4 Tag (M/u) "old" and go to step 2. ■

The u vector
represents an
upper bound
for the estimation
error $e = m - \mu$

- 1 A net system $\mathcal{S} = \langle N, \mathbf{m}_0 \rangle$ is **MO** if there exists a node in its OCG such that $\mathbf{u} = \mathbf{0}$
- 2 A net system $\mathcal{S} = \langle N, \mathbf{m}_0 \rangle$ is **SMO** in k steps *iff* $\mathbf{u} = \mathbf{0}$ for each node (\mathbf{m}, \mathbf{u}) in the OCG such that
 - the node belongs to a cycle
 - the node is dead
- 3 Conditions to check uMO, uSMO, sMO and sSMO require additional *tools* (see Giua & Seatzu, *IEEE TAC*, 2002, for more details)

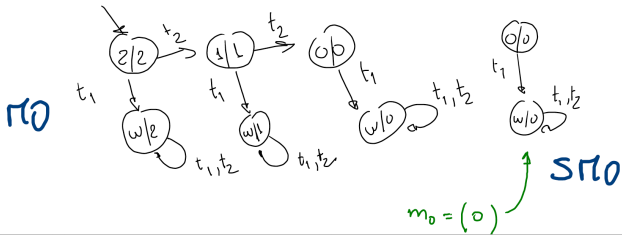
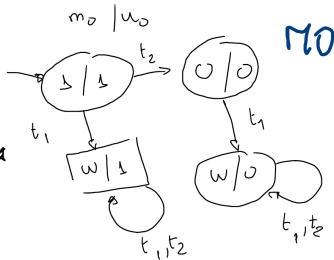
OCG – Examples

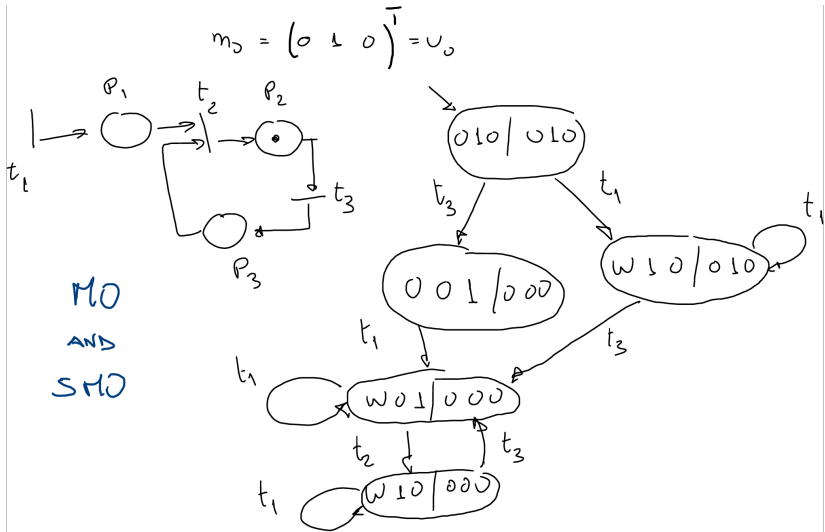


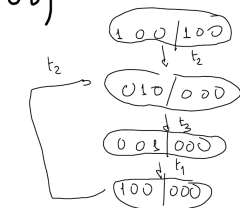
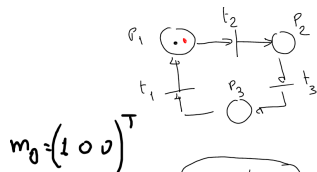
m/u
 $u_0 = m_0$

$m_0 = (1)$

$m_0 = (2)$

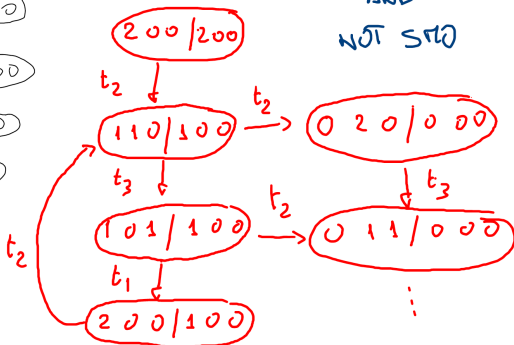






$$m_0 = (2\ 0\ 0)^T$$

MO
 AND
 $NOT\ SMO$



- In ε -free labeled systems, an **event can map on more than one transition**, to model undistinguishable events or uncertain dynamic
- The concept of **consistent markings** with a given observer word $w \in E^*$ can be used to build a state observer that does not require the construction of the reachability graph, and thus works for both bounded and **unbounded** systems

Assumptions

- 1 The structure of the net $N = (P, N, \mathbf{Pre}, \mathbf{Post})$ is known
- 2 The initial marking m_0 is known
- 3 The labeling function is ε -free and the events associated to transition firings can be observed



A. Giua, D. Corona, C. Seatzu

State estimation of λ -free labeled Petri nets with contact-free nondeterministic transitions

Discrete Event Dynamic Systems: Theory and Applications, 2005

Set of w -consistent markings $\mathcal{C}(w)$

Given an observed word w , the set of w -consistent markings $\mathcal{C}(w)$ is

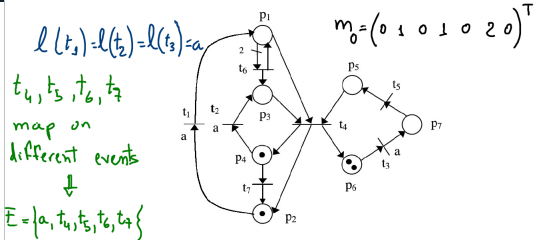
$$\mathcal{C}(w) = \{ \mathbf{m} \in \mathbb{N}^m \mid \exists \text{ a sequence } \sigma \in T^* \text{ such that } \mathbf{m}_0[\sigma] \text{ and } \ell(\sigma) = w \}$$

Algorithm to compute $\mathcal{C}(w)$

- 1 Let $w_0 = \varepsilon$ and $\mathcal{C}(w_0) = \mathbf{m}_0$
- 2 Let $i = 0$
- 3 Wait until a new event e is observed
- 4 Let $i = i + 1$
- 5 Let $w_i = w_{i-1}e$ and $\mathcal{C}(w_i) = \emptyset$
- 6 For all $\mathbf{m} \in \mathcal{C}(w_{i-1})$ do
 - For all t such that $\mathbf{m}[t]$ and $\ell(t) = e$
compute $\mathbf{m}' = \mathbf{m} + \mathbf{C}(\cdot, t)$ and let $\mathcal{C}(w_i) = \mathcal{C}(w_i) \cup \mathbf{m}'$
- 7 Goto step 3

- To compute the set of markings that are consistent with an observed word w with $|w| = k$, requires to compute the set of markings that are consistent with all prefixes of w
- Therefore, given a word w , each set $\mathcal{C}(\tilde{w})$ with $\tilde{w} \in \overline{\{w\}}$ must be explicitly enumerated
- However, note that the cardinality of the set of consistent markings may either increase or decrease as the length of the observed word increases

Example



$$C(E) = \{m_0\}$$

"a is observed" ($\Rightarrow w = a$)

$$C(a) = \left\{ (1 \ 0 \ 0 \ 1 \ 0 \ 2 \ 0)^T, (0 \ 1 \ 1 \ 0 \ 0 \ 2 \ 0)^T, (0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 1)^T \right\}$$

"a is observed again" ($\Rightarrow w = aa$)

$$C(aa) = \left\{ (1 \ 0 \ 1 \ 0 \ 0 \ 2 \ 0)^T, (0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1)^T, (0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 2)^T, (1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1)^T \right\}$$

"t₇ is observed" ($\Rightarrow w = aat_7$)

t₇ requires a token in p₄ to fire

$$C(aat_7) = \left\{ (0 \ 2 \ 0 \ 0 \ 0 \ 0 \ 2)^T, (1 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1)^T \right\}$$



A **linear algebraic characterization of $\mathcal{C}(w)$** can be given, with a fixed number of constraints, when the following additional assumption is made

- **Nondeterministic transitions** are *contact-free*, i.e., being t_i and t_j non deterministic, it is


$$\bullet t_i \cap \bullet t_j = \emptyset \quad \text{and} \quad \bullet t_i \cap t_i^{\bullet} = \emptyset$$

- The nondeterministic transition are those ones that share the event with other transitions
- **A linear characterization permits to avoid the enumeration of elements in $\mathcal{C}(w)$**
- **In some applications enumeration is not needed, while an algebraic characterization is sufficient (see also the diagnosability case in Lecture #7)**
- The details can be found in A. Giua, D. Corona, C. Seatzu, *Discrete Event Dynamic Systems: Theory and Applications*, 2005

State estimation for system with silent events

A **linear algebraic characterization of $\mathcal{C}(w)$** has been given also in the case of labeled system whose transitions map on the **silent event ε**

Assumptions

- 1 The structure of the net $N = (P, N, \mathbf{Pre}, \mathbf{Post})$ is known
 - 2 The initial marking m_0 is known
 - 3 The labels associated to the firing of transitions that do not map on ε can be observed, and a different label is associated to each of these transitions
 - 4 The subnet induced by the *silent* transitions is *acyclic*
 - 5 The subnet induced by the *silent* transitions is *backward conflict-free*, i.e., any two distinct silent transitions have no common output place
- Assumption 3 prevent to model undistinguishable events and some uncertainty on the dynamic
 - The details can be found in
 -  A. Giua, C. Seatzu, D. Corona
Marking Estimation of Petri Nets With Silent Transitions
IEEE Transactions on Automatic Control, 2007

- Diagnoser and Diagnosability – Chapter 2 (section 2.5.3) in



C. G. Cassandras and S. Lafortune
Introduction to Discrete Event Systems
Springer, 2008

- Diagnoser and Diagnosability – the seminal work



M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, D. Teneketzis
IEEE Transaction on Automatic Control
vol. 40, n. 9, pp. 1555-1575, 2005

- Observer coverability graph



A. Giua and C. Seatzu
Observability of Place/Transition Nets
IEEE Transactions on Automatic Control, 2002

- State estimation labeled systems



A. Giua, D. Corona, C. Seatzu
State estimation of λ -free labeled Petri nets with contact-free nondeterministic transitions
Discrete Event Dynamic Systems: Theory and Applications, 2005



A. Giua, C. Seatzu, D. Corona
Marking Estimation of Petri Nets With Silent Transitions
IEEE Transactions on Automatic Control, 2007

Adding uncertainty: unobservable events and observers for finite state automata and PNs

From observability to privacy and security in discrete event systems

Prof. Gianmaria DE TOMMASI
Email: detommas@unina.it

December 2020