Chiachío, Manuel and Chiachío, Juan and Prescott, Darren and Andrews, John (2016) An information theoretic approach for knowledge representation using Petri nets. In: Future Technologies Conference 2016, 6-7 Dec 2016, San Francisco, USA.

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/39038/1/An%20information%20theoretic%20approach%20for%20knowledge%20representation%20using%20Petri%20nets.pdf

Copyright and reuse:
The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:
The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk
An information theoretic approach for knowledge representation using Petri nets

Manuel Chiachío, Juan Chiachío, Darren Prescott and John Andrews
Resilience Engineering Research Group
Faculty of Engineering
University of Nottingham
Nottingham, NG7 2RD, UK
Email:{manuel.chiachio, juan.chiachioruano, darren.prescott, john.andrews}@nottingham.ac.uk

Abstract—A new hybrid approach for Petri nets (PNs) is proposed in this paper by combining the PNs principles with the foundations of information theory for knowledge representation. The resulting PNs have been named Plausible Petri nets (PPNs) mainly because they can handle the evolution of a discrete event system together with uncertain (plausible) information about the system using states of information. This paper overviews the main concepts of classical PNs and presents a method to allow uncertain information exchange about a state variable within the system dynamics. The resulting methodology is exemplified using an idealized expert system, which illustrates some of the challenges faced in real-world applications of PPNs.

Keywords—Petri nets; Expert Systems; Knowledge representation

I. INTRODUCTION

A. Classical Petri Nets

Petri Nets (PNs) are bipartite directed graphs (digraph) used for modeling the dynamics of systems, which are known after the celebrated thesis dissertation Kommunikation mit Automaten by Carl Petri in 1962 [1]. Two types of nodes are represented in a PN: transitions and places, which are temporarily visited by tokens, the abstract moving units of a PN that can adopt different meanings depending on the application (also different meanings within a specific PN). The distribution of tokens over the PN at a specific time of execution is referred to as marking.

From a mathematical perspective, a PN can be defined as an ordered 6-tuple $\Psi$ as follows [2]:

$$\Psi = (P, T, E, M_0, D, W)$$

where $P \in \mathbb{N}^n_p$ and $T \in \mathbb{N}^n_t$ are vectors to denote the set of $n_p$ places and $n_t$ transitions of the PN respectively, $M_0 \in \mathbb{N}^{n_p}$ is the initial marking, and $D \in \mathbb{R}^{n_t}$ is the non-negative vector of switching delays of the transitions (0 by default). The connections between transitions and places are expressed through the set of edges $E \subseteq (P \times T) \cup (T \times P)$, so that $E \subseteq \mathbb{N}^{n_p} \times \mathbb{N}^{n_t}$ contains ordered pairs of nodes. Each edge (also referred to as arc) has assigned a weight, a non-negative integer value (1 by default) defined in the set of weights $W$.

There are three main laws that govern a PN, that can be enunciated as follows:

1) Transitions always consume from all the input arcs at the same time;

2) Transitions always produce from all out-coming arcs the same amount;

3) Transitions cannot consume tokens from an empty place.

The dynamics of classical PNs can be described through a state equation defined as follows:

$$M_{k+1} = M_k + A^T u_k$$

where $u_k$ is the firing vector at state $k$, an $n_t$-dimensional vector of Boolean values whose elements are obtained according to the firing rule, as will be explained further below. $A$ is an $n_t \times n_p$ matrix typically referred to as the incidence matrix, whose elements are the result of subtracting the forward ($A^+$) and backward ($A^-$) incidence matrices respectively, as follows:

$$A = A^+ - A^-$$

In the last equation, the element $a_{ij}^+$ is the weight of the arc from transition $t_i$ to output place $p_j$, whereas $a_{ij}^-$ is the weight of the arc to transition $t_i$ from input place $p_j$, where $i = 1, \ldots, n_t$, $j = 1, \ldots, n_p$. If transition $t_i$ is activated at state $k$, then $u_{i,k} \in u_k$ is modified according to the firing rule, which can be expressed as follows:

$$u_{i,k} = \begin{cases} 1, & \text{if } M(j) \geq a_{ij}^- \forall p_j \in \mathbb{P}_{t_i} \\ 0, & \text{otherwise} \end{cases}$$

where $M(j) \in \mathbb{N}$ is the marking for place $p_j$, and $\mathbb{P}_{t_i}$ denotes the set of places that belong to the pre-set of transition $t_i$, $i = 1, \ldots, n_t$, as will be next described.

B. Tokens, probability, and information

Despite the efficiency that PNs have demonstrated for modelling expert system dynamics for a widespread range of applications, their pragmatism is not commensurate with the
In this work, a hybrid approach is proposed by combining traditional tokens with a new class of tokens that represent a degree of belief about a state variable at a certain mode which is specified by a place. The resulting framework has been named Plausible Petri Nets (PPN) since certain tokens encode a set of numerical values about a state variable \( x \) from a finite-dimensional set \( \mathcal{X} \subset \mathbb{R}^d \), along with a mapping over \( \mathcal{X} \) denoted by \( f : \mathcal{X} \rightarrow [0,1] \) that assigns each element \( x \) with a non-negative value interpreted as its relative plausibility, as will be discussed in the next section. In PPNs, the moving units are objects (in the sense of integer units as in classical PNs) and also states of information about \( x \). Moreover, the consumption and production of tokens in PPNs are not only associated with the concept of adding and subtracting integer values, but also with the operation of adding and subtracting information about the state variable \( x \).

The next section provides further insight into the concepts of plausibility, uncertainty, and states of information, since they are fully adopted throughout the text. Section III is devoted to providing the basis and main rules of the proposed PPNs. In Section IV, our framework is illustrated using an example of application. Finally, Section V gives concluding remarks.

### II. Basic operations in Plausible Petri Nets

Let \( A \subseteq \mathcal{X} \) be a subset representing a certain event or proposition over \( \mathcal{X} \) in such a way that there exist a probability measure \( P(A) \) through a density function \( f(x) \) that can be normalized such that \( \int_{\mathcal{X}} f(x)dx = 1 \). In this sense, \( P(A) \) is interpreted as the plausibility of a set of possible values \( x \) given a state of information about them provided by \( f(x) \) [10]–[12]. Let us now evoke the first principles of Boolean logic by recalling the concepts of the logic operators AND (\( \land \)) and OR (\( \lor \)) for the conjunction and disjunction of propositions, respectively. Then, the following logical relationships are compatible according to the De Morgan’s law [13], [14] for two probability measures \( P_a(A) \) and \( P_b(A) \), each one ascribed to the state of information \( f_a(x) \) and \( f_b(x) \), respectively:

\[
\begin{align*}
P_a(A) = 0 & \quad \text{OR} \quad P_b(A) = 0 \quad \Rightarrow \quad (P_a \land P_b)(A) = 0 \\
P_a(A) \neq 0 & \quad \text{OR} \quad P_b(A) \neq 0 \quad \Rightarrow \quad (P_a \lor P_b)(A) \neq 0
\end{align*}
\]

where \((P_a \land P_b)(A)\) and \((P_a \lor P_b)(A)\) can be expressed as [15]:

\[
\begin{align*}
(P_a \land P_b)(A) &= \int_{A} (f_a \land f_b)(x)dx \\
(P_a \lor P_b)(A) &= \int_{A} (f_a \lor f_b)(x)dx
\end{align*}
\]

In the last equation, the densities \((f_a \land f_b)(x)\) and \((f_a \lor f_b)(x)\) represent the conjunction and disjunction of states of information given by \( f_a(x) \), \( f_b(x) \), respectively [12], [16]. Besides, \((P_a \land P_b)(A)\) and \((P_a \lor P_b)(A)\) stand for the plausibility of the set of values \( x \in A \) in compliance with the information given by \((f_a \land f_b)(x)\) and \((f_a \lor f_b)(x)\), respectively.

Next, because commutativity is allowed in logic of propositions, i.e. \((P_a \lor P_b)(A) = (P_b \lor P_a)(A)\) and \((P_a \land P_b)(A) = (P_b \land P_a)(A)\), then a simple solution for the unnormalized densities \((f_a \land f_b)(x)\) and \((f_a \lor f_b)(x)\) are such that [17]:

\[
\begin{align*}
(f_a \lor f_b)(x) &= f_a(x) + f_b(x) \quad (8a) \\
(f_a \land f_b)(x) &= \frac{f_a(x)f_b(x)}{\mu(x)} \quad (8b)
\end{align*}
\]

where \( \mu(x) \) is the homogeneous density function [12], [18], a reference state of information that can be understood as a probability model for \( x \in \mathcal{X} \) in absence of any other information, which actually represents the state of total ignorance about \( \mathcal{X} \) [18], [19]. In case that \( \mathcal{X} \) is a linear space, it is demonstrated [17] that \( \mu(x) = \text{const.} \), i.e., \( \mu(x) \) is a uniform over \( \mathcal{X} \). See Fig. 1 for a conceptual illustration of the conjunction and disjunction of states of information over some arbitrary densities \( f_a(x) \) and \( f_b(x) \). Moreover, it is worth noting that the conjunction and disjunction operations over states of information can be extended to the case of multiple states of information (e.g. \( f_1(x), f_2(x), \ldots, f_n(x) \)), as follows [18]:

\[
\begin{align*}
(f_1 \lor \cdots \lor f_n)(x) &= \sum_{i=1}^{n} f_i(x) \\
(f_1 \land \cdots \land f_n)(x) &= \prod_{i=1}^{n} f_i(x)
\end{align*}
\]

Finally, observe that the normalization of the density function in (8b) may require the evaluation of an intractable integral such that \( \int_{A} \frac{1}{\mu(x)} f_a(x)f_b(x)dx = 1 \), where \( \alpha \) is a normalizing constant. Moreover, note that there might be situations where the conjunction is conducted with density functions which are not completely known, perhaps because they are defined through samples. Hence, sampling-based algorithms (e.g. particle methods) [20], [21] can be used in those cases to circumvent the evaluation of the normalizing constant. In particle methods, a set of \( N \) samples \( \{x^{(n)}\}_{n=1}^{N} \) with associated weights \( \{w^{(n)}\}_{n=1}^{N} \) are used to obtain an approximation for the required density function [e.g. \( (f_a \land f_b)(x) \)] with a feasible computational cost, as follows:

\[
(f_a \land f_b)(x) \approx \sum_{n=1}^{N} w^{(n)} \delta(x - x^{(n)}) \quad (10)
\]

where \( x^{(n)} \sim (f_a \land f_b)(x) \), and \( \delta \) is the Dirac delta. The particle weight \( w^{(n)} \) represents a likelihood value of the \( n \)-th particle, which can be obtained as follows:

\[
w^{(n)} = \frac{f_a(x^{(n)})f_b(x^{(n)})}{\sum_{n=1}^{N} f_a(x^{(n)})f_b(x^{(n)})}
\]

where \( \mathcal{X} \) is assumed to be a linear space.
III. INFORMATION FLOW DYNAMICS IN PPNs

A. Modelling assumptions

Let \( \{x_k, k \in \mathbb{N}\} \) be a stochastic process taking values in \( \mathcal{X} \), which is manifested through different modes corresponding to each place of the Petri net. Next, let us consider that \( X \) can be partitioned into two disjoint sets: 1) numerical places \( P(\mathcal{X}) \subset \mathbb{N}^\nu \), and 2) symbolic places \( P^{(S)} \subset \mathbb{N}^\nu \), such that \( P(\mathcal{X}) \cup P^{(S)} = P \subset \mathbb{N}^{\nu + \nu'} \), and \( P(\mathcal{X}) \cap P^{(S)} = \emptyset \). Analogously, the set of transitions \( T \) is partitioned into numerical transitions \( T(\mathcal{X}) \subset \mathbb{N}^\nu \), and symbolic transitions \( T^{(S)} \subset \mathbb{N}^\nu' \), where \( T(\mathcal{X}) \cup T^{(S)} = T \subset \mathbb{N}^{\nu + \nu'} \), and \( T(\mathcal{X}) \cap T^{(S)} \neq \emptyset \). Those transitions that belong to \( T(\mathcal{X}) \cap T^{(S)} \) are referred to as mixed transitions. Finally, let us call by \( *P_t \), the subset of places from the pre-set of transition \( t_i \), \( i = 1, \ldots, n_t \). Analogously \( *P_t \) is to denote the subset of places that belongs to the post-set of \( t_i \). From this standpoint, the dynamics of PPNs is formulated under the adoption of the following items:

1) At a certain time \( k \), the place \( p_0(\mathcal{X}) \) encloses a state of information about \( x_k \) given by \( f_k \), where \( f_k : A_j \rightarrow \mathbb{R}^+ \) and \( A_j \) is a subset of \( \mathcal{X} \);

2) Equivalently, any transition \( t_i \in T(\mathcal{X}) \) carries a state of information about \( x_k \) defined over \( A_i \subseteq \mathcal{X} \), namely \( f_k^{A_i} (x_k) \);

3) As in classical PNs, there exist arc weights for the symbolic places, denoted by \( a_{ij}^+, a_{ij}^- \in W^{(S)} \subset \mathbb{N} \), whereby the incidence matrix \( A^{(S)} \) can be obtained as: \( A^{(S)} = a_{ij}^+-a_{ij}^- \), \( i = 1, \ldots, n'_i \), \( j = 1, \ldots, n'_j \), where \( n'_i, n'_j \) are the number of symbolic places and transitions of the PPNs, respectively. The arc weights for the numerical places are denoted by \( a_{ij}^+, a_{ij}^- \in W^{(X)} \subset \mathbb{R}^+ \), where \( A^{(X)} = a_{ij}^+-a_{ij}^- \), \( i = 1, \ldots, n_t \), \( j = 1, \ldots, n_p \), being in this case \( n_t, n_p \) the amount of numerical transitions and numerical places of the PPN, respectively. These weights provide us with a measure of the importance of the information that flows from/to the corresponding transition;

4) Any transition \( t_i \) always consumes from all its input arcs at the same time. Moreover, an input arc from place \( p_i^{(N)} \) to transition \( t_i \in T^{(N)} \) conveys a state of information given by the conjunction of states of information of \( f_k^P (x_k) \) and \( f_k^{(X)} (x_k) \), namely \( (f_k^P \wedge f_k^{(X)})(x_k) \);

5) Any transition \( t_i \) produces to its output arcs the same amount of information given by the conjunction of states of information between \( f_k^P (x_k) \) and \( f^{(X)} (x_k) \), i.e., \( (f_k^P \wedge f^{(X)})(x_k) \). In the last expression, \( f^{(X)} (x_k) \) results from the disjunction of the states of information of the pre-set of transition \( t_i \), i.e., \( f^{P_i}(x_k) = (f^{P_i} \wedge f^{(X)} \vee \neg f^{P_i}) \)

---

Fig. 1. Illustrative example of the conjunction and disjunction operation over two states of information, namely \( f_a(x) \) (left panels) and \( f_b(x) \) (center panels). The right panels represent the resulting density \( f(x) \) from the conjunction (upper right) and disjunction (lower right) of \( f_a \) and \( f_b \). The resulting PDF is represented superimposed over \( f_a \) and \( f_b \) for better understanding.
\(f^{p_{ij}}_k \lor \cdots \lor f^{p_{in}}_k(x_k)\) (recall \([9a]\)), where places 
\(p_{i1}, p_{i2}, \ldots, p_{in}, m \in \mathbf{P}_k \subset \mathbf{P}(N)\);

6) If numerical place \(p_{ij}^{(N)}\) belongs to the pre-set of \(t_i\), and assuming that \(f^{p_{ij}}_k(x_k) \notin \mathbf{P}_k^*\), \(\forall i = 1, \ldots, n_t\), then the state of information that remains in \(p_{ij}^{(N)}\) at \(k+1\) after firing \(t_i\), is the conjunction of \(f^{p_{ij}}_k(x_k)\) and \(f^{i_k}_j(x_k)\) weighted according to \(a_{ij}\), i.e. \(f^{p_{ij}}_{k+1}(x_{k+1}) = a_{ij}(f^{p_{ij}}_k \land f^{i_k}_j)(x_k)\);

7) Correspondingly, the state of information resulting in place \(p_{ij}^{(N)}\) after firing \(t_i\) given that \(p_{ij}^{(N)} \in \mathbf{P}_k^*\), is the disjunction of the states of information \(f^{p_{ij}}_k(x_k)\) and \(f^{i_k}_j(x_k)\), where the latter is multiplied by its output weight \(a_{ij}\). Mathematically \(f^{p_{ij}}_{k+1}(x_{k+1}) = (f^{p_{ij}}_k \lor a_{ij}(f^{i_k}_j \land f^{p_{ij}}_k)(x_k))\).

Observe that the rules given above, specifically those from items 3) and 5), are in agreement with the rules of classical PNs. In particular, the results in Section III-A, and in particular, results from applying the rules 4) to 7). Both markings, namely \(M_k^{(N)}\) and \(M_k^{(S)}\), evolve in a synchronized manner driven by the firing rule for PPNs, as will be explained below.

C. Firing rule for PPNs

In PPNs, any transition \(t_i \in \mathbf{T}\) is fired at state \(k\) if the delay time has passed and:

1) \(\forall p_{ij}^{(S)} \in \mathbf{P}_k, \ M_k^{(S)}(j) \geq a_{ij}^{-}, \text{ for } t_i \in \mathbf{T}(S), t_i \notin (\mathbf{T}(S) \cap \mathbf{T}(N))\);

2) \(\forall p_{ij}^{(N)} \in \mathbf{P}_k, (f^{p_{ij}}_k \land f^{i_k}_j)(x_k) \neq \emptyset\), which in this case, applies for numerical transitions, i.e., \(t_i \in \mathbf{T}(N), t_i \notin (\mathbf{T}(S) \cap \mathbf{T}(N))\);

3) Condition 1) and 2) are both satisfied for numerical places and symbolic places from the pre-set of \(t_i\), when \(t_i\) is a mixed transition, i.e. \(t_i \in (\mathbf{T}(S) \cap \mathbf{T}(N))\).

Condition 1) means that each symbolic place from the pre-set of \(t_i\) has enough amount of tokens according to their input weight, as in classical PNs. On the other hand, condition 2) allows us to ensure that every conjunction of states of information between \(f^{i_k}_j\) and each of the density functions from the pre-set of \(t_i\), are possible. Note that a conjunction (e.g. \((f^{p_{ij}}_k \land f^{i_k}_j)(x_k)\)) is possible if \((f^{p_{ij}}_k \land f^{i_k}_j)(x_k) \neq \emptyset\) for any subset \(B \in \mathcal{X}\) \([12]\).

IV. APPLICATION EXAMPLE

In this section, a numerical example is provided to illustrate our proposed methodology. To this end, let us consider that there exist a measurement about a bi-dimensional state variable \(x_k = (x_{1,k}, x_{2,k})\), performed using two sensors denoted by \(s_1\) and \(s_2\). The information from \(s_1\) and \(s_2\) is initially represented in places \(p_{1}^{(N)}\) and \(p_{2}^{(N)}\) using probability densities as respective states of information about \(x_k\). Suppose now that the sensors are imperfect so they introduce uncertainty that can modeled as a zero mean Gaussian density function with covariance matrix given by \(\Sigma_1\) and \(\Sigma_2\), respectively. Next, consider that the component-wise measurements are stochastically independent so that \(\Sigma_1 = diag(\sigma_{11}^2, \sigma_{21}^2), \Sigma_2 = diag(\sigma_{12}^2, \sigma_{22}^2)\), where \(\sigma_{11}, \sigma_{22}, \sigma_{12}, \sigma_{21}\) are the standard deviations of measurement errors \(v = (v_1, v_2, v_1, v_2)\) for sensors \(s_1\) and \(s_2\), respectively. The system being modeled is based on an expert system that controls the activation of a discrete-event subsystem (which can represent an automata, machine, or similar, in a real-world application). Once a new data point arrives, then the activation/deactivation of the discrete subsystem occurs conditioned upon the quality of the information coming from the sensors. The differential entropy, as a measure of the data uncertainty, is used in this example as a quality indicator of the information given by the sensors. Fig. 2 provides a geometric description of the PPN being modeled using four numerical places \((p_1^{(N)}, p_2^{(N)}, p_3^{(N)}, p_4^{(N)})\), five symbolic places \((p_1^{(S)}, p_2^{(S)}, p_3^{(S)})\), and three mixed transitions labeled as \(t_1, t_2, t_3\). A cold transition \((e)\) is used to represent the data arrival to the system. In our PPN model, the information from sensors \(s_1\) and \(s_2\), initially represented in places \(p_1\) and \(p_2\), is further gathered in place \(p_3\) provided that transition \(t_1\) is fired. The joint information in place \(p_3\) is then used to activate a sequence of discrete-event system (e.g. a machine) if the differential entropy of \(x_k\) is lower than a certain reference value \(\xi \in \mathbb{R}\), which is taken as \(\xi = 7\) in this example. Next, the state of information about \(x_k\) is finally collected in place \(p_4\). Otherwise, transition \(t_2\) is activated whereby the information in place \(p_2\) is improved using the joint information about \(x_k\) given in place \(p_3\). Note that \(p_2\) helps in synchronizing transitions \(t_1\) and \(t_3\), so as to avoid them to be activated simultaneously.

\(^{1}\)The assumption implies that place \(p_{ij}^{(N)}\) does not belong to the post-set of any transition.
The initial marking for the numerical places is $M_0^{(N)} = \left( f_0^{p_1^N}(x_0), f_0^{p_2^N}(x_0), f_0^{p_3^N}(x_0), f_0^{p_4^N}(x_0) \right)^T$, where $f_0^{p_i} \sim \mathcal{N}(\hat{x}, \Sigma_1)$, $f_0^{p_3} \sim \mathcal{N}(\hat{x} + \beta, \Sigma_2)$, $f_0^{p_4} = f_0^{p_5} = \emptyset$. In the last expressions, $\hat{x}$ is the data point that arrives to the system, and $\beta$ is a bias that applies to the measurements from sensor $s_2$. In this example, $\hat{x} = (20, 15)$, and $\beta = (5, 5)$, both expressed in arbitrary units. The covariance matrices for $f_0^{p_1}(x_0)$ and $f_0^{p_3}(x_0)$ are given by $\Sigma_1 = \text{diag}(4^2, 3^2)$ and $\Sigma_2 = \text{diag}(3^2, 3^2)$, respectively. See Fig. 3 for a particle-based representation of the states of information for places $p_1^{(N)}$ and $p_2^{(N)}$ at $k = 0$.

On the other side, the initial marking for symbolic places is given by $M_0^{(S)} = (0, 0, 0, 0, 1)^T$. The mixed transitions $t_i$, $i = \{1, 2, 3\}$ are defined using Dirac Delta density functions, i.e., $f_i^{T}(x_k) \sim \mathcal{I}_{B_i}(x_k)$, thus their firing is prescribed for the state variable $x_k$ on fulfilling the condition $x_k \in B_i$, where:

- $B_1 = \{ x_k \in \mathcal{X} : (x_{1,k} - 20)^2 + (x_{2,k} - 20)^2 \leq 10 \}$ (12a)
- $B_2 = \{ x_k \in \mathcal{X} : H(x_k) \geq \xi \}$ (12b)
- $B_3 = \{ x_k \in \mathcal{X} : H(x_k) \leq \xi \}$ (12c)

In (12b) and (12c), $H : \mathcal{X} \rightarrow \mathbb{R}$ denotes the differential entropy of $x_k$, that can be obtained by calculating $1/2 \ln(2\pi e)^d \det(\text{cov}(x_k))$ as a measure quantifying the uncertainty of $x_k$. Here, we assume for simplicity that the time spent by the machine in performing transitions $t_4$ and $t_5$ is negligible. The rules for the information flow dynamics of PPNs [recall rules 4) to 7) in Section III-A] along with (2), are applied in confluence with the firing rule for the system state evolution described through the marking $M_k$. The particle approximation for conjunction of states of information described in Section II, has been applied to evaluate the conjunction of states of information resulting from applying rules 4) to 7) from the information flow dynamics for the numerical subnet. The disjunction of states of information are straightforwardly evaluated using samples by just joining the samples from the component-wise density functions, and affecting their particle weights using an appropriate normalizing constant so as to obtain a bone fide density. The results for the numerical places are presented for states $k = 0$ and $k = 2$ in Fig. 4 (the states of information about $x_k$ remain unchanged in $P^{(N)}$ after $k = 2$, hence they are not represented). A summary of the results for the analysis of the PPN is provided in Table I. Note from Table I that the results for symbolic marking and active transitions are column vectors, although they are not explicitly reflected as so in this table for clarity. Also, the fourth column represents the continuum entropy of $f_k^{p_i}(x_k), k = 1, 2, 3$. The incidence matrices used for calculations of the numerical and symbolic

---

\[ \text{This expression for the differential entropy is actually an upper-bound approximation to the actual differential entropy, where the exactness is achieved when the density function is Gaussian.} \]
A novel hybrid approach for PNs has been proposed in this paper by combining the algebra of classical PNs with the basis of information theory. As in classical PNs, there exist a set of places and transitions (labeled here as “symbolic”) to model discrete-event systems using regular tokens as moving units. In addition, a set of numerical places and transitions are considered to model knowledge representation about a system state variable, which interact with the regular tokens in a synchronized manner. The simulated results demonstrate that PPNs are a versatile tool to fuse uncertain information (which can be diverse through the different numerical nodes) about the state variable with sequences of Boolean events, provided that an adequate net architecture is adopted. This fact makes them useful for analyzing hybrid systems with interaction of diverse sources of information, like in expert systems. One future research direction is to explore their ability to model cybersystems, since they can receive, store, exchange, and process information so as to use it for control. Moreover, further research effort is needed to investigate formal aspects of the resulting hybrid system, as well as to explore efficient implementations using a variety of examples of application.

**ACKNOWLEDGMENT**

Manuel Chiachío is Research Fellow of the Lloyd’s Register Foundation (LRF), a charitable foundation in the U.K. helping to protect the life and property by supporting engineering-related education, public engagement, and the application of research. Juan Chiachío and Darren Prescott are funded by Engineering and Physical Sciences Research Council (UK) under grant EP/M023028/1). John Andrews is the Director of the Resilience Engineering Research Group and also the Network Rail Professor of Infrastructure Asset Management at the University of Nottingham. The authors gratefully acknowledge the support of these organizations.
Fig. 4. PPN output for numerical places $p_1^{(x)}$ (most-left panel) to $p_4^{(x)}$ (most-right panel). Each subplot represents the state of information about $x_k$, using samples (circles) in the state space $X$.

REFERENCES

[1] C. A. Petri, “Kommunikation mit automaten,” Ph.D. dissertation, Institut für Instrumentelle Mathematik an der Universität Bonn, 1962.
[2] T. Murata, “Petri nets: Properties, analysis and applications,” Proceedings of the IEEE, vol. 77, no. 4, pp. 541–580, 1989.
[3] J. Cardoso, R. Valette, and D. Dubois, “Possibilistic Petri nets,” IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 29, no. 5, pp. 573–582, 1999.
[4] Z. Ding, Y. Zhou, and M. Zhou, “Modeling self-adaptive software systems with learning Petri nets,” in Companion Proceedings of the 36th International Conference on Software Engineering. ACM, 2014, pp. 464–467.
[5] G. Horton, V. G. Kulkarni, D. M. Nicol, and K. S. Trivedi, “Fluid stochastic Petri nets: Theory, applications, and solution techniques,” European Journal of Operational Research, vol. 105, no. 1, pp. 184–201, 1998.
[6] M. Silva, J. Júlvez, C. Mahulea, and C. R. Vázquez, “On fluidization of discrete event models: observation and control of continuous Petri nets,” Discrete Event Dynamic Systems, vol. 21, no. 4, pp. 427–497, 2011.
[7] S.-M. Chen, J.-S. Ke, and J.-F. Chang, “Knowledge representation using fuzzy Petri nets,” Knowledge and Data Engineering, IEEE Transactions on, vol. 2, no. 3, pp. 311–319, 1990.
[8] C. Lesire and C. Tessier, “Particle Petri nets for aircraft procedure monitoring under uncertainty,” in Applications and Theory of Petri Nets 2005: 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005. Proceedings, G. Ciardo, Ed. Springer, 2005, pp. 329–348.
[9] L. Zouaghi, A. Alexopoulos, A. Wagner, and E. Badreddin, “Modified particle Petri nets for hybrid dynamical systems monitoring under environmental uncertainties,” in System Integration (SII), 2011 IEEE/SICE International Symposium on. IEEE, 2011, pp. 497–502.
[10] R. T. Cox, “Probability, frequency and reasonable expectation,” American journal of physics, vol. 14, p. 1, 1946.
[11] E. Jaynes, Papers on probability, statistics and statistical physics. (Ed. R.D Rosenkrantz), Kluwer Academic Publishers, 1983.
[12] A. Tarantola and B. Valette, “Inverse problems = quest for information,”
Journal of Geophysics, vol. 50, no. 3, pp. 159–170, 1982.

[13] I. M. Copi, *Introduction to logic.* New York, Macmillan, 1953.

[14] R. L. Goodstein, *Boolean algebra.* Dover Publications, Inc., 2007.

[15] A. Kolmogorov, *Foundations of the Theory of Probability (Translation of 1933 original in German).* Chelsea Publishing: New York, 1950.

[16] G. Rus, J. Chiachío, and M. Chiachío, “Logical inference for inverse problems,” *Inverse Problems in Science and Engineering,* vol. 24, no. 3, pp. 448–464, 2016.

[17] A. Tarantola, *Inverse problem theory and methods for model parameters estimation.* SIAM, 2005.

[18] K. Mosegaard and A. Tarantola, “Probabilistic approach to inverse problems,” in *International Handbook of Earthquake and Engineering Seismology, Part A,* ser. International Geophysics. Academics Press Ltd, 2002, vol. 81, pp. 237–265.

[19] E. T. Jaynes, “Prior probabilities,” *Systems Science and Cybernetics, IEEE Transactions on,* vol. 4, no. 3, pp. 227–241, 1968.

[20] M. Arumulpalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking,” *IEEE Transactions on Signal Processing,* vol. 50, no. 2, pp. 174–188, 2002.

[21] A. Doucet, N. De Freitas, and N. Gordon, “An introduction to sequential Monte Carlo methods,” in *Sequential Monte Carlo methods in practice,* A. Doucet, N. De Freitas, and N. Gordon, Eds. Springer, 2001, pp. 3–14.