UNIVERSITÄT
KOBLENZ · LANDAU
Institut für Informatik

FB 4
Informatik

Probability Propagation Nets

Kurt Lautenbach
Alexander Pinl

Nr. 20/2007

Arbeitsberichte aus dem Fachbereich Informatik

Arbeitsberichte des Fachbereichs Informatik

ISSN (Print): 1864-0346
ISSN (Online): 1864-0850

Kontaktdaten der Verfasser

Kurt Lautenbach, Alexander Pinl
Institut für Informatik
Fachbereich Informatik
Universität Koblenz-Landau
Universitätsstraße 1
D-56070 Koblenz
EMail: laut@uni-koblenz.de, apinl@uni-koblenz.de,

Probability Propagation Nets

Kurt Lautenbach, Alexander Pinl*

***Abstract*—A class of high level Petri nets, called "probability propagation nets", is introduced which is particularly useful for modeling probability and evidence propagation. These nets themselves are well suited to represent the probabilistic Horn abduction, whereas specific foldings of them will be used for representing the flows of probabilities and likelihoods in Bayesian networks.**

***Index Terms*—Bayes procedures, Horn clauses, Petri nets, Probability, Propagation, Stochastic logic.**

I. Introduction

THIS paper deals with the propagation of probabilities in Petri nets (PNs). That means, first of all, it is a paper about PNs and their ability to represent the dynamic in logical-probabilistic structures.

By far most of the papers about PNs and probabilities are about transitions whose duration is governed by a probability distribution. In contrast to that, we will introduce a class of PNs, called *"Probability Propagation Nets" (PPNs)*, for developing transparent and well structured models in which probabilities are propagated, for example as decision aids or degrees of risk.

We will try out the modeling power of our approach by means of the probabilistic Horn abduction [1]–[3] and Bayesian networks (BNs) [4]. In doing so, we will avoid to give the impression that we are going to improve these approaches. However, using PNs means to work with one of the most famous modeling tools. So the outcome might have positive facets. We think, for example, that our approach is very transparent and structured in a "natural" way. Moreover, PNs for propagating probabilities can be combined or mixed with other types of PNs, for example with representations of biological, medical or technical systems. We think, that this is a very important aspect for the development of tools.

Different from some existing approaches on combining BNs and PNs (e.g. [5], [6]), we introduce step by step new PNs that are particularly well suited for our intentions: they are transparent and structured.

First, we modify the p/t-nets for representing logical inference [7] by inscribing tokens and arcs with probabilities. These nets, the PPNs, allow to represent stochastic inference and probabilistic Horn abduction.

Second, foldings of these nets reduce the net structure and allow the representation of BNs. Fortunately, the inscriptions and the firing rule remain clear.

In spite of the considerable complexity, the PNs are of manageable size. A particular advantage is that all the propagation processes are represented by t-invariants (because of the **0**-reproducibility of the propagations) and that the t-invariants can be calculated in the underlying p/t-nets.

* A. Pinl is supported by the DFG under grant LA 1042/7-2.

Our paper is organized as follows. Section II comprises that part of PN theory which is needed for the subsequent sections. In section III PPNs are introduced on the basis of a PN representation of Horn formulas. The probabilistic Horn abduction is used to exemplify the modeling power of PPNs. Section IV addresses higher PPNs as foldings of PPNs. Here, BNs are used to demonstrate the modeling ability. This is continued in section V by means of popular examples. Section VI and an appendix conclude the paper.

II. Preliminaries

Definition 1 1) *A* place/transition net *(*p/t-net*) is a quadruple $\mathcal{N} = (S, T, F, W)$ where*

(a) *S and T are finite, non empty, and disjoint sets. S is the set of* places *(in the figures represented by circles). T is the set of* transitions *(in the figures represented by boxes).*

(b) *$F \subseteq (S \times T) \cup (T \times S)$ is the set of directed* arcs.

(c) *$W : F \to \mathbb{N} \setminus \{0\}$ assigns a* weight *to every arc. In case of $W : F \to \{1\}$, we will write $\mathcal{N} = (S, T, F)$ as an abridgment.*

2) *The* preset *(*postset*) of a node $x \in S \cup T$ is defined as ${}^\bullet x = \{y \in S \cup T \mid (y, x) \in F\}$ ($x^\bullet = \{y \in S \cup T \mid (x, y) \in F\}$).*
The preset *(*postset*) of a set $H \subseteq S \cup T$ is ${}^\bullet H = \bigcup_{x \in H} {}^\bullet x$ ($H^\bullet = \bigcup_{x \in H} x^\bullet$).*
For all $x \in S \cup T$ it is assumed that $|{}^\bullet x| + |x^\bullet| \geq 1$ holds; i.e. there are no isolated nodes.

3) *A place p (transition t) is* shared *iff $|{}^\bullet p| \geq 2$ or $|p^\bullet| \geq 2$ ($|{}^\bullet t| \geq 2$ or $|t^\bullet| \geq 2$).*

4) *A place p is an* input (output) boundary place *iff ${}^\bullet p = \varnothing$ ($p^\bullet = \varnothing$).*

5) *A transition t is an* input (output) boundary transition *iff ${}^\bullet t = \varnothing$ ($t^\bullet = \varnothing$).*

Definition 2 *Let $\mathcal{N} = (S, T, F, W)$ be a p/t-net.*

1) *A* marking *of $\mathcal{N}$ is a mapping $M : S \to \mathbb{N}$. $M(p)$ indicates the number of* tokens *on p under M. $p \in S$ is* marked by *M iff $M(p) \geq 1$. $H \subseteq S$ is* marked by *M iff at least one place $p \in H$ is marked by M. Otherwise p and H are* unmarked*, respectively.*

2) *A transition $t \in T$ is* enabled by *M, in symbols $M[t\rangle$, iff*

$$\forall p \in {}^\bullet t : M(p) \geq W((p, t)).$$

3) *If $M[t\rangle$, the transition t may* fire *or* occur*, thus leading*

to a new marking M', in symbols $M[t\rangle M'$, with

$$M'(p) := \begin{cases} M(p) - W((p,t)) & \text{if } p \in {}^\bullet t \setminus t^\bullet \\ M(p) + W((t,p)) & \text{if } p \in t^\bullet \setminus {}^\bullet t \\ M(p) - W((p,t)) + W((t,p)) & \text{if } p \in {}^\bullet t \cap t^\bullet \\ M(p) & \text{otherwise} \end{cases}$$

for all $p \in S$.

4) *The set of all markings* reachable *from a marking M_0, in symbols $[M_0\rangle$, is the smallest set such that*

$$M_0 \in [M_0\rangle$$
$$M \in [M_0\rangle \wedge M[t\rangle M' \Rightarrow M' \in [M_0\rangle.$$

$[M_0\rangle$ is also called the set of follower markings *of M_0.*

5) *$\sigma = t_1 \dots t_n$ is a* firing sequence *or* occurrence sequence *for transitions $t_1, \dots, t_n \in T$ iff there exist markings $M_0, M_1, \dots, M_n$ such that*

$$M_0[t_1\rangle M_1[t_2\rangle \dots [t_n\rangle M_n \text{ holds;}$$

in short $M_0[\sigma\rangle M_n$. $M_0[\sigma\rangle$ denotes that σ starts from *M_0. The* firing count *$\bar{\sigma}(t)$ of t in σ indicates how often t occurs in σ. The (column) vector of firing counts is denoted by $\bar{\sigma}$.*

6) *The pair $(\mathcal{N}, M_0)$ for some marking M_0 of $\mathcal{N}$ is a p/t-*system *or a* marked *p/t*-net. *M_0 is the* initial marking.

7) *A marking $M \in [M_0\rangle$ is* reproducible *iff there exists a marking $M' \in [M\rangle, M' \neq M$ s.t. $M \in [M'\rangle$.*

8) *Moreover, the p-column-vector* $\mathbf{0}$ *stands for the* empty marking. *A p/t-net is* $\mathbf{0}$-reproducing *iff there exists a firing sequence φ such that $\mathbf{0}[\varphi\rangle\mathbf{0}$. A transition t is* $\mathbf{0}$-firable *iff t can be enabled by some follower marking of* $\mathbf{0}$.

Definition 3 *Let $\mathcal{N} = (S, T, F, W)$ be a p/t-net;*

1) *$\mathcal{N}$ is* pure *iff $\nexists (x,y) \in (S \times T) \cup (T \times S) : (x,y) \in F \wedge (y,x) \in F$.*
2) *A* place vector *($|S|$-vector) is a column vector $\upsilon : S \to \mathbb{Z}$ indexed by S.*
3) *A* transition vector *($|T|$-vector) is a column vector $\omega : T \to \mathbb{Z}$ indexed by T.*
4) *The* incidence matrix *of $\mathcal{N}$ is a matrix $[\mathcal{N}] : S \times T \to \mathbb{Z}$ indexed by S and T such that*

$$[\mathcal{N}](p,t) = \begin{cases} -W((p,t)) & \text{if } p \in {}^\bullet t \setminus t^\bullet \\ W((t,p)) & \text{if } p \in t^\bullet \setminus {}^\bullet t \\ -W((p,t)) + W((t,p)) & \text{if } p \in {}^\bullet t \cap t^\bullet \\ 0 & \text{otherwise.} \end{cases}$$

υ^t and A^t are the transposes *of a vector υ and a matrix A, respectively. The columns of $[\mathcal{N}]$ are $|S|$-vectors, the rows of $[\mathcal{N}]$ are transposes of $|T|$-vectors. Markings are representable as $|S|$-vectors, firing count vectors as $|T|$-vectors.*

Definition 4 *Let I be a place vector and J a transition vector of $\mathcal{N} = (S, T, F, W)$.*

1) *I is a* place invariant *(p-invariant) iff $I \neq \mathbf{0}$ and $I^t \cdot [\mathcal{N}] = \mathbf{0^t}$*
2) *J is a* transition invariant *(t-invariant) iff $J \neq \mathbf{0}$ and $[\mathcal{N}] \cdot J = \mathbf{0}$*
3) *$\|I\| = \{p \in S \mid I(p) \neq 0\}$ and $\|J\| = \{t \in T \mid J(t) \neq 0\}$ are the* supports *of I and J, respectively.*
4) *A p-invariant I (t-invariant J) is*
   - non-negative *iff $\forall p \in S : I(p) \geq 0$ ($\forall t \in T : J(t) \geq 0$)*
   - positive *iff $\forall p \in S : I(p) > 0$ ($\forall t \in T : J(t) > 0$)*
   - minimal *iff I (J) is non-negative and $\nexists$ p-invariant $I' : \|I'\| \subsetneq \|I\|$ ($\nexists$ t-invariant $J' : \|J'\| \subsetneq \|J\|$) and the greatest common divisor of all entries of I (J) is 1.*
5) *The* net representation *$\mathcal{N}_I = (S_I, T_I, F_I, W_I)$ of a p-invariant I is defined by*

$$S_I := \|I\|$$
$$T_I := {}^\bullet S_I \cup S_I{}^\bullet$$
$$F_I := F \cap ((S_I \times T_I) \cup (T_I \times S_I))$$

W_I is the restriction of W to F_I.

6) *The* net representation *$\mathcal{N}_J = (S_J, T_J, F_J, W_J)$ of a t-invariant J is defined by*

$$T_J := \|J\|$$
$$S_J := {}^\bullet T_J \cup T_J{}^\bullet$$
$$F_J := F \cap ((S_J \times T_J) \cup (T_J \times S_J))$$

W_J is the restriction of W to F_J.

7) *$\mathcal{N}$ is* covered *by a p-invariant I (t-invariant J) iff $\forall p \in S : I(p) \neq 0$ ($\forall t \in T : J(t) \neq 0$).*

Proposition 1 *Let $(\mathcal{N}, M_0)$ be a p/t-system, I a p-invariant; then*

$$\forall M \in [M_0\rangle : I^t \cdot M = I^t \cdot M_0.$$

Proposition 2 *Let $(\mathcal{N}, M_0)$ be a p/t-system, $M_1 \in [M_0\rangle$ a follower marking of M_0, and σ a firing sequence that reproduces $M_1 : M_1[\sigma\rangle M_1$; then the firing count vector $\bar{\sigma}$ of σ is a t-invariant.*

Definition 5 *Let $\mathcal{N} = (S, T, F, W)$ be a p/t-net, M_0 a marking of $\mathcal{N}$, and $r \geq \mathbf{0}$ a $|T|$-vector; r is* realizable in *$(\mathcal{N}, M_0)$ iff there exists a firing sequence σ with $M_0[\sigma\rangle$ and $\bar{\sigma} = r$.*

Proposition 3 *Let $\mathcal{N} = (S, T, F, W)$ be a p/t-net, M_1 and M_2 markings of $\mathcal{N}$, and σ a firing sequence s.t. $M_1[\sigma\rangle M_2$; then the linear relation*

$$M_1 + [\mathcal{N}]\bar{\sigma} = M_2 \text{ holds.}$$

In the above linear relation, the state equation, *the order of transition firings is lost.*

Definition 6 (Natural Multiset) *Let A be a non-empty set;*

- $m : A \to \mathbb{N}$ *is a* natural multiset over A*;*
- $\mathbb{M}(A)$ *is the* set of all natural multisets over A.

III. Probability Propagation Nets

In this section, we introduce probability propagation nets (PPNs) which are a "probabilistic extension" of place/transition nets representing logical formulas (see [7], [8]). Starting with the canonical net representation of Horn formulas in conjunctive normal form, we enrich these formulas by probabilities as in probabilistic Horn abduction [1]–[3]. After that, we introduce an appropriate extension of the canonical (Petri) net representation modeling Horn formulas. The resulting PNs are called "probability propagation nets".

The transformation of a logical Horn formula into the canonical net representation is detailedly described in [7]. In order to give a short summary and to introduce the relevant terms, we stick to an example.

Definition 7 *Let $\tau = \neg a_1 \vee \cdots \vee \neg a_m \vee b_1 \vee \cdots \vee b_n$ be a clause;*
in set notation: $\tau = \neg A \cup B$ for $\neg A = \{\neg a_1, \ldots, \neg a_m\}$ and $B = \{b_1, \ldots, b_n\}$;

- τ *is a* fact clause *iff* $\neg A = \varnothing$,
- τ *is a* goal clause *iff* $B = \varnothing$,
- τ *is a* rule clause *iff* $\neg A \neq \varnothing \wedge B \neq \varnothing$,
- τ *is a* Horn clause *iff* $|B| \leq 1$.

Let α be a conjunction of clauses, i.e. α is a conjunctive normal form (CNF) formula;

- $\mathbb{A}(\alpha)$ *denotes the* set of atoms *of* α,
- $\mathbb{C}(\alpha)$ *denotes the* set of clauses *of* α,
- $\mathbb{F}(\alpha)$ *denotes the* set of fact clauses *of* α,
- $\mathbb{G}(\alpha)$ *denotes the* set of goal clauses *of* α,
- $\mathbb{R}(\alpha) := \mathbb{C}(\alpha) \setminus (\mathbb{F}(\alpha) \cup \mathbb{G}(\alpha))$ *denotes the* set of rule clauses *of* α*;*

α *is a* Horn formula *iff its clauses are Horn clauses.*

Definition 8 (Canonical Net Representation) *Let α be a CNF -formula and let $\mathcal{N}_\alpha = (S_\alpha, T_\alpha, F_\alpha)$ be a p/t-net;*
$\mathcal{N}_\alpha$ *is the* canonical p/t-net representation *of* α *iff*

- *$S_\alpha = \mathbb{A}(\alpha)$ (set of atoms of α) and $T_\alpha = \mathbb{C}(\alpha)$ (set of clauses of α)*
- *for all $\tau = \neg a_1 \vee \cdots \vee \neg a_m \vee b_1 \vee \cdots \vee b_n \in \mathbb{C}(\alpha)$, where $\{a_1, \ldots, a_m, b_1, \ldots, b_n\} \subseteq \mathbb{A}(\alpha)$, F_α is determined by*
  ${}^\bullet\tau = \{a_1, \ldots, a_m\}$, $\tau^\bullet = \{b_1, \ldots, b_n\}$, i.e. the atoms $a_1, \ldots, a_m$ which are negated in the clause τ are the input places, the non-negated atoms $b_1, \ldots, b_n$ are the output places of the transition τ.

The transition τ is called fact (goal, rule) transition *iff the clause τ is a fact (goal, rule) clause.*

Remark 1
In non-canonical p/t-net representations, S_α contains negated atoms (see [7]).

TABLE I
Horn Clauses of Example 1

1	$\neg nolo \vee \neg igno \vee acno$	$\mathbb{R}(\alpha)$
2	$\neg nolo \vee \neg igir \vee acno$	
3	$\neg nolo \vee \neg igno \vee acde$	
4	$\neg nolo \vee \neg igir \vee acde$	
5	$\neg lo \vee \neg igno \vee acno$	
6	$\neg lo \vee \neg igir \vee acno$	
7	$\neg lo \vee \neg igno \vee acde$	
8	$\neg lo \vee \neg igir \vee acde$	
9	$\neg nolo \vee owof$	
10	$\neg lo \vee owof$	
11	$\neg nolo \vee owon$	
12	$\neg lo \vee owon$	
13	$igno$	$\mathbb{F}(\alpha)$
14	$igir$	
15	$nolo$	
16	lo	
17	$\neg acno$	$\mathbb{G}(\alpha)$
18	$\neg acde$	
19	$\neg owof$	
20	$\neg owon$	

Theorem 1 Let α be a Horn formula and let $\mathcal{N}_\alpha = (S_\alpha, T_\alpha, F_\alpha)$ be its canonical p/t-representation; then the following statements are equivalent:

(1) α is contradictory.
(2) $\mathcal{N}_\alpha$ is $\mathbf{0}$-reproducing.
(3) $\mathcal{N}_\alpha$ has a t-invariant $R \geq 0$ with $R(g) > 0$ for some goal transition g.
(4) In $\mathcal{N}_\alpha$ a goal transition g is $\mathbf{0}$-firable.
(5) In $\mathcal{N}_\alpha$ there exists a set Y of reverse paths from a goal transition to fact transitions such that with any transition t of a path of Y its incidenting places $p \in {}^\bullet t \cup t^\bullet$ are nodes of a path of Y, too.

Proof: See [7]. ■

Example 1 (cf. [3])
Let α be the Horn formula that is the conjunction of the clauses given in Table I where the atoms are lo (lack of oil), $nolo$ (no lack of oil), $igir$ (ignition irregular), $igno$ (ignition normal), $owon$ (oil warning lamp on), $owof$ (oil warning lamp off), $acde$ (acceleration delayed), and $acno$ (acceleration normal).

Fig. 1 shows the canonical net representation $\mathcal{N}_\alpha$ of α (see Definition 8).

□

Definition 9 *Let α be a Horn formula, let $H \subseteq \mathbb{F}(\alpha)$ be a set of fact clauses called the* "assumable hypotheses", *let be $E \subseteq H$, $R \subseteq \mathbb{R}(\alpha) \cup (\mathbb{F}(\alpha) \setminus E)$, let $\varepsilon = \bigwedge_{\varphi \in E} \varphi$, $\varrho = \bigwedge_{\kappa \in R} \kappa$ be the corresponding Horn formulas, let $\gamma = \neg g_1 \vee \cdots \vee \neg g_m$, $\gamma \in \mathbb{G}(\alpha)$ be a goal clause; then ε is an* explanation (diagnosis) *of $\neg\gamma = g_1 \wedge \cdots \wedge g_m$ iff*

- *$\neg\gamma = g_1 \wedge \cdots \wedge g_m$ is a logical consequence of $\varepsilon \wedge \varrho$ and*
- *$\varepsilon \wedge \varrho$ is not contradictory.*

Remark 2
A minimal t-invariant of $\mathcal{N}_\alpha$ has only one goal transition g, because α is a Horn formula.

Example 2
The t-invariants of $\mathcal{N}_\alpha$ (see Fig. 1 and Table II) are $\mathbf{0}$-reproducing, which can easily be verified by simulation. There are four t-invariants passing through $t_{18} = \neg acde$ (for which $I(t_{18}) > 0$ holds), namely I_5, I_6, I_7, I_8.

According to theorem 1 $((3), (2), (1))$, in the net representation of all these t-invariants (regarded as single canonical net

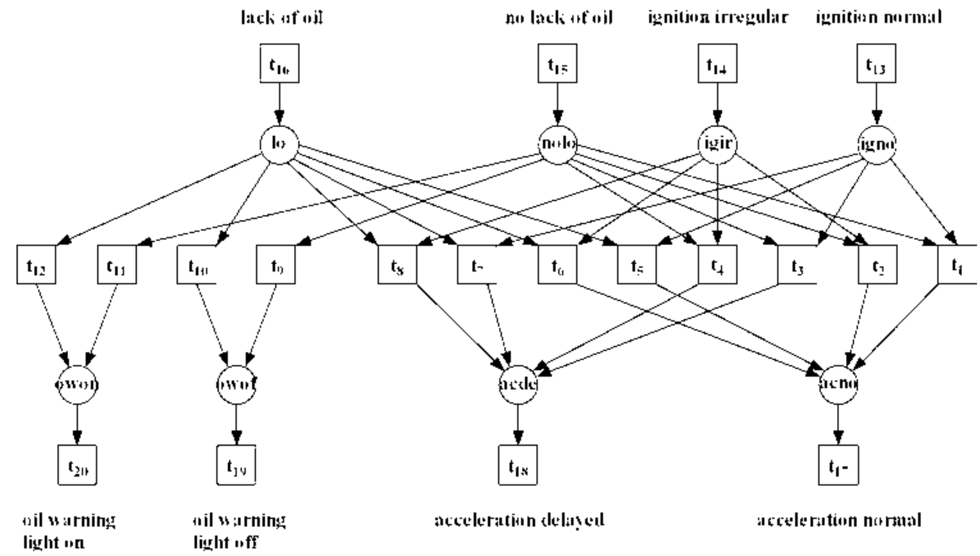


Fig. 1. $\mathcal{N}_\alpha$ of Example 1

TABLE II
T-INVARIANTS OF $\mathcal{N}_\alpha$ (EXAMPLE 1)

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13} $igno$	t_{14} $igir$	t_{15} $nolo$	t_{16} lo	t_{17} $\neg acno$	t_{18} $\neg acde$	t_{19} $\neg owof$	t_{20} $\neg owon$
I_1					1								1			1	1			
I_2	1												1		1		1			
I_3						1								1		1	1			
I_4		1												1	1		1			
I_5							1						1			1		1		
I_6			1										1		1			1		
I_7								1						1		1		1		
I_8				1										1	1			1		
I_9										1						1			1	
I_{10}									1						1				1	
I_{11}												1				1				1
I_{12}											1				1					1

representations), $acde$ is a logical consequence of the other clauses. For example I_7:

$$\gamma_7 = t_{18} = \neg acde$$
$$\varepsilon_7 = \{t_{16}, t_{14}\} = \{lo, igir\} = lo \wedge igir$$
$$\varrho_7 = t_8 = \neg lo \vee \neg igir \vee acde = lo \wedge igir \rightarrow acde$$

so, $\neg\gamma_7 = acde$ is a logical consequence of $\varepsilon_7 \wedge \varrho_7$.

Moreover, $\varepsilon_7 \wedge \varrho_7$ is not contradictory since in its canonical net representation t_{18} is missing such that the empty marking $\mathbf{0}$ is not reproducible (see theorem 1). So ε_7 is an explanation of $\neg\gamma_7 = acde$.

Alltogether,

$$\varepsilon_5 = \{lo, igno\} = lo \wedge igno$$
$$\varepsilon_6 = \{nolo, igno\} = nolo \wedge igno$$
$$\varepsilon_7 = \{lo, igir\} = lo \wedge igir$$
$$\varepsilon_8 = \{nolo, igir\} = nolo \wedge igir$$

are the explanations of $acde$.

□

Definition 10 *Let α be a Horn formula and $P_\alpha : \mathbb{C}(\alpha) \rightarrow [0,1]$ a real function, called a* probability function of α*;*

let $H \subseteq \mathbb{F}(\alpha)$ be a set of fact clauses; let $\{D_1, \ldots, D_n\}$ be a partition of H (i.e. $D_i \cap D_j = \varnothing$ for $i \neq j, \bigcup_{i=1}^n D_i = H$) where for all $D_i, 1 \leq i \leq n, \sum_{\varphi \in D_i} P_\alpha(\varphi) = 1$;

then the sets $D_1, \ldots, D_n$ are called disjoint classes*;*

let be $P_\alpha(\gamma) := 1$ for all goal clauses $\gamma \in \mathbb{G}(\alpha)$, let be $E \subseteq H$, $R \subseteq \mathbb{R}(\alpha) \cup \mathbb{F}(\alpha)$, $\gamma \in \mathbb{G}(\alpha)$ and let $\varepsilon = \bigwedge_{\varphi \in E} \varphi$, $\varrho = \bigwedge_{\kappa \in R} \kappa$ be the corresponding Horn formulas, where ε is an explanation (diagnosis) of $\neg\gamma$.

The probability of ε *is given by $P_\alpha(\varepsilon \wedge \varrho)$. The problem to find explanations is the* probabilistic Horn abduction (PHA).

Let furthermore I be a t-invariant of the canonical net representation $\mathcal{N}_\alpha$ of α such that I performs the $\mathbf{0}$-reproduction, induced by $\varepsilon \wedge \varrho \wedge \gamma$ being contradictory; then $\prod_{t \in \|I\| \setminus \{\gamma\}} P_\alpha(t)$ equals the probabilities of ε and of $\neg\gamma$ *w.r.t. I.*

Remark 3
The atoms of α are now to be interpreted as random variables. The atoms of the fact clauses in a disjoint class D form together with P_α a finite probability space.

Remark 4
For interpreting the probability function P_α, let $\tau = \neg a_1 \vee \cdots \vee \neg a_m \vee b$ be a Horn clause of α where $\neg A = \{\neg a_1, \ldots, \neg a_m\}$, $\quad B = \{b\}$:

- if τ is a fact clause ($\tau \in \mathbb{F}(\alpha), \neg A = \varnothing, B \neq \varnothing$), $P_\alpha(\tau)$ is the prior probability $P(b)$ of b,

TABLE III
PROBABILITY FUNCTION P_α OF EXAMPLE 3

transition	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}	t_{17}	t_{18}	t_{19}	t_{20}
P_α	1.0	0.4	0.0	0.6	0.2	0.0	0.8	1.0	1.0	0.0	0.0	1.0	0.9	0.1	0.6	0.4	1.0	1.0	1.0	1.0
Disjoint class													D_1	D_1	D_2	D_2				

- if τ is a rule clause ($\tau \in \mathbb{R}(\alpha), \neg A \neq \varnothing, B \neq \varnothing$), $P_\alpha(\tau)$ is the conditional probability $P(b \mid a_1, \dots, a_m)$ of b given $a_1, \dots, a_m$,
- if τ is a goal clause ($\tau \in \mathbb{G}(\alpha), \neg A \neq \varnothing, B = \varnothing$), the value of $P_\alpha(\tau)$ is not relevant for any calculation;
from a logical point of view, the value 0 is justified because every **0**-reproduction is an indirect proof and results in a contradiction; the value 1 (see Definition 10) is a very handy compromise.

Example 3 (see Examples 1 and 2)
The probability function P_α with two disjoint classes is shown in Table III. We want to calculate the probabilities of $acde$ and its explanations. There are four t-invariants passing through $t_{18} = \neg acde$: $\{I_i \mid 1 \leq i \leq 12, I_i(t_{18}) \neq 0\} = \{I_5, I_6, I_7, I_8\}$ (see Table II). The explanations of $acde$ are $\varepsilon_i = \|I_i\| \cap \mathbb{F}(\alpha)$ for $5 \leq i \leq 8$:

$$\varepsilon_5 = \{lo, igno\} = lo \wedge igno$$
$$\varepsilon_6 = \{nolo, igno\} = nolo \wedge igno$$
$$\varepsilon_7 = \{lo, igir\} = lo \wedge igir$$
$$\varepsilon_8 = \{nolo, igir\} = nolo \wedge igir$$

In simple cases like this one, or if it is not necessary to watch the simulation of the (net representation of the) t-invariants, we calculate immediately:
$P(\varepsilon_i) = \prod_{t \in \|I_i\|} P_\alpha(t)$
(Please note that for the goal transitions $P_\alpha(t) = 1.0$ holds.)

$$P(\varepsilon_5) = 0.9 \cdot 0.4 \cdot 0.8 \cdot 1.0 = 0.288 \text{ (max.)}$$
$$P(\varepsilon_6) = 0.9 \cdot 0.6 \cdot 0.0 \cdot 1.0 = 0.0$$
$$P(\varepsilon_7) = 0.1 \cdot 0.4 \cdot 1 \cdot 1.0 = 0.04$$
$$P(\varepsilon_8) = 0.1 \cdot 0.6 \cdot 0.6 \cdot 1.0 = 0.036$$

$P(acde)$ sums up to 0.364. In case of simulating the four t-invariants, transition t_{18} (acceleration delayed) would fire for $ad = 0.288, 0, 0.04,$ and 0.036.

In order to combine the probability aspects with the propagation abilities of PNs, we introduce a new class of nets in two steps. Fig. 2(a) shows the net representation of t-invariant I_5. For the calculation of $P(\varepsilon_5)$ it would be convenient to have the following sequence of markings:

1) M with $M(lo) = M(igno) = M(acde) = \varnothing$ (empty marking)
2) $M'(lo) = (P(lo)) = (P_\alpha(t_{16})) = (0.4)$;
$M'(igno) = (P(igno)) = (P_\alpha(t_{13})) = (0.9)$;
$M'(acde) = \varnothing$ after one subsequent firing of t_{16} and t_{13}
3) $M''(lo) = M''(igno) = \varnothing$
$M''(acde) = (P(\varepsilon_5)) = (P(lo) \cdot P(igno) \cdot P(acde \mid lo \wedge igno)) = (0.4 \cdot 0.9 \cdot P_\alpha(t_7)) = (0.4 \cdot 0.9 \cdot 0.8) = (0.288)$ after one subsequent firing of t_7

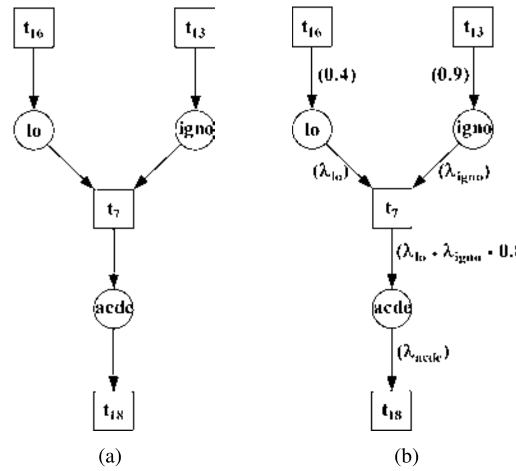


Fig. 2. Invariant I_5 of Example 4

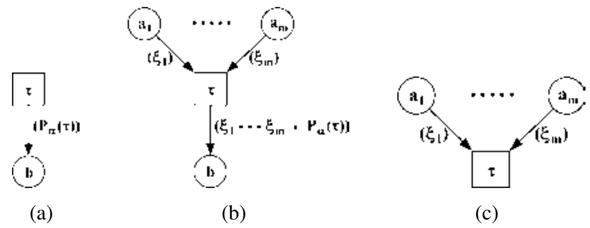


Fig. 3. Arc Label Function Types for a PPN (see Definition 11)

4) $M''' = M$ (empty marking) after one subsequent firing of t_{18}.

To get all that in accordance with the notation of a suitable higher level PN (predicate/transition net notation in this case) we have to complete the net as shown in Fig. 2(b).

□

Definition 11 (Probability Propagation Net, PPN) *Let α be a Horn formula and $\tau = \neg a_1 \vee \cdots \vee \neg a_m \vee b$ a Horn clause of α with $\neg A = \{\neg a_1, \dots, \neg a_m\}, B = \{b\}$;*
$\mathcal{PN}_\alpha = (S_\alpha, T_\alpha, F_\alpha, P_\alpha, L_\alpha)$ is a probability propagation net (PPN) *for α iff*

- *$\mathcal{N}_\alpha = (S_\alpha, T_\alpha, F_\alpha)$ is the canonical net representation of α,*
- *P_α is a probability function for α,*
- *L_α is an* arc label function *for α where for τ the following holds:*
  - *if τ is a fact clause ($\tau \in \mathbb{F}(\alpha), \neg A = \varnothing, B \neq \varnothing$), $L_\alpha(\tau, b) = (P_\alpha(\tau))$, $(\tau, b) \in F_\alpha$ (see Fig. 3(a))*
  - *if τ is a rule clause ($\tau \in \mathbb{R}(\alpha), \neg A \neq \varnothing, B \neq \varnothing$), $L_\alpha(a_i, \tau) = (\xi_i)$ for $1 \leq i \leq m$
$L_\alpha(\tau, b) = (\xi_1 \cdots \xi_m \cdot P_\alpha(\tau))$*

where the ξ_i are variables ranging over $[0,1]$ (see Fig. 3(b))

– *if τ is a goal clause ($\tau \in \mathbb{G}(\alpha), \neg A \neq \varnothing, B = \varnothing$) $L_\alpha(a_i, \tau) = (\xi_i)$ for $1 \leq i \leq m$ (see Fig. 3(c)).*

Definition 12 (PPN Marking) *Let α be a Horn formula and $\mathcal{PN}_\alpha = (S_\alpha, T_\alpha, F_\alpha, P_\alpha, L_\alpha)$ a PPN for α; let W be a finite subset of $[0,1]$, and let $(W) := \{(w) \mid w \in W\}$ be the corresponding set of 1-tuples; let be $\tau \in T_\alpha$ with $^\bullet\tau = \{s_1, \dots, s_m\}, \tau^\bullet = \{s_{m+1}\}$ (i.e. $\tau = \neg s_1 \vee \dots \vee \neg s_m \vee s_{m+1}$); then $M : S_\alpha \to \mathbb{M}((W))$ is a* marking *of $\mathcal{PN}_\alpha$;*

τ is enabled *by M for $\{(w_1), \dots, (w_m)\}$ iff $(w_1) \in M(s_1), \dots, (w_m) \in M(s_m)$,*

the follower marking *M' of M after one firing of τ for $\{(w_1), \dots, (w_m)\}$ is given by*

$$\begin{aligned} M'(s_1) &= M(s_1) - (w_1), \\ &\vdots \\ M'(s_m) &= M(s_m) - (w_m), \\ M'(s_{m+1}) &= M(s_{m+1}) + (w_1 \cdot w_2 \dots w_m \cdot P_\alpha(\tau)); \end{aligned}$$

if $(\xi_1), \dots, (\xi_m)$ are the arc labels of $(s_1, \tau), \dots, (s_m, \tau) \in F_\alpha$, we may write

$$M'(s_1) = M(s_1) - (\xi_1), \dots, M'(s_m) = M(s_m) - (\xi_m),$$
$$M'(s_{m+1}) = M(s_{m+1}) + (\xi_1 \dots \xi_m \cdot P_\alpha(\tau)),$$

if the ξ_i are bound by the corresponding $w_i, 1 \leq i \leq m$.

Example 4 (see Example 3)
The net $\mathcal{PN}_\alpha$ of Fig. 4 is the PPN that combines the net $\mathcal{N}_\alpha$ of Fig. 1 and the probabilities of Table III.

The probabilities of Example 3 will now be calculated by simulating the t-invariants I_5, I_6, I_7, I_8. For example, simulating I_5 yields the maximal probability $P(\varepsilon_5) = 0.288$. Firing t_{13} and t_{16} yields tuples (0.9) and (0.4) on places $igno$ and lo, respectively. Firing t_7 takes away these tuples and puts the tuple $(0.4 \cdot 0.9 \cdot 0.8) = (0.288)$ on place $acde$, from where it is taken away by t_{18} — such completing the reproduction of the empty marking by simulating I_5.

□

A major problem, the "loopiness", arises from the fact that the conjunction operator $\wedge$ is idempotent ($a \wedge a = a$), but the corresponding product of probabilities is not idempotent in general:

$$P(a) \cdot P(a) \begin{cases} = P(a) & \text{if } P(a) = 1 \text{ or } P(a) = 0 \\ \neq P(a) & \text{else} \end{cases}$$

The following example shows a case of loopiness and a method to get over that difficulty.

Example 5 (see Example 4)
We want to calculate the probability of $acde \wedge owon$. For that, we modify the PPN of Fig. 4 in several steps:

- transitions (goal clauses) $t_{18} = \neg acde$ and $t_{20} = \neg owon$ are unified to one transition (goal clause) $t_{20} = \neg acde \vee \neg owon = \neg(acde \wedge owon)$;

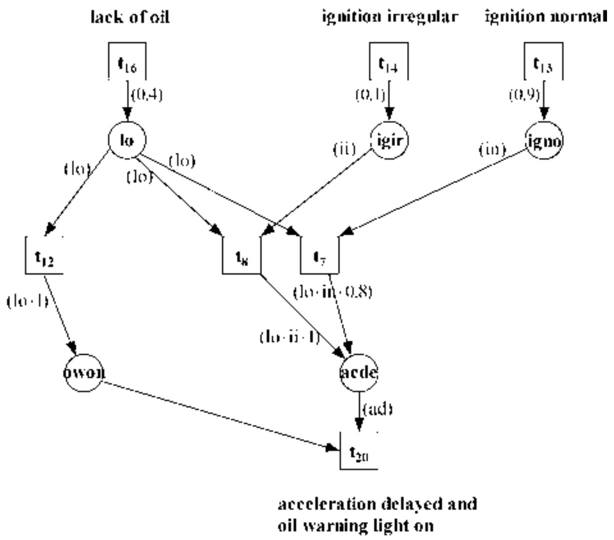


Fig. 5. $\mathcal{PN}_\alpha'$ of Example 5

TABLE IV
T-INVARIANTS OF $\mathcal{PN}_\alpha'$ (EXAMPLE 5)

	t_5	t_7	t_8	t_{12}	t_{13} $igno$	t_{14} $igir$	t_{16} lo	t_{20} $\neg owon$
I_2			1	1		1	2	1
I_3		1		1	1		2	1

- the transitions $t_{19} = \neg owof$ and $t_{17} = \neg acno$ are omitted because they are not needed any more; as a consequence, also $t_9, t_{10}, t_1, t_2, t_5, t_6$ are no longer needed.
- all t-invariants with transitions t, where $P_\alpha(t) = 0$, are omitted: t_{11}, t_3;
- t_{15} and t_4 are omitted because the only t-invariant they belong to contains a factual contradiction: t_{16} (lack of oil) and t_{15} (no lack of oil).

The result is the PPN $\mathcal{PN}_\alpha'$ shown in Fig. 5. From a structural point of view, this net is well suited for solving our problem because its set of t-invariants (see Table IV) is reduced to the relevant ones. From a probabilistic point of view, we first of all have to note that the net is loopy. On the other hand, the net is optimal to apply Pearls's conditioning method [4]. In contrast to his technique to cut the loops, we do not need to cut the net because of the t-invariant structure that forces to fire t_{16} twice in both t-invariants (see Table IV). This, in principle, leads to a double effect of (lo) when t_{20} fires (via $owon$ and via $acde$). For $L_\alpha(t_{16}, lo) = (P(t_{16})) = (1.0)$, however, this effect is neutralized. So, by simulating or simply multiplying the probabilities, we get for the t-invariants the following temporary values:

$$I_1 : P_\alpha(t_{16})^2 \cdot P_\alpha(t_{14}) \cdot P_\alpha(t_{12}) \cdot P_\alpha(t_8) \cdot P_\alpha(t_{20}) = 0.1$$
$$I_2 : P_\alpha(t_{16})^2 \cdot P_\alpha(t_{13}) \cdot P_\alpha(t_{12}) \cdot P_\alpha(t_7) \cdot P_\alpha(t_{20}) = 0.72$$

Finally, both values have to be multiplied by the weight 0.4 which is the original value of $P_\alpha(t_{16})$:

$P(acde \wedge owon) = 0.04$	w.r.t. I_1
$P(acde \wedge owon) = 0.288$	w.r.t. I_2

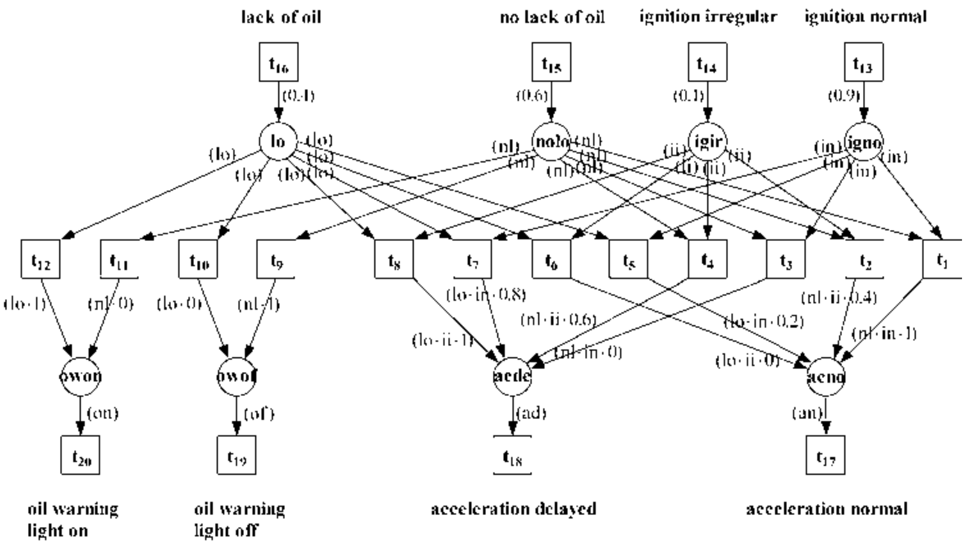


Fig. 4. $\mathcal{PN}_\alpha$ of Example 4

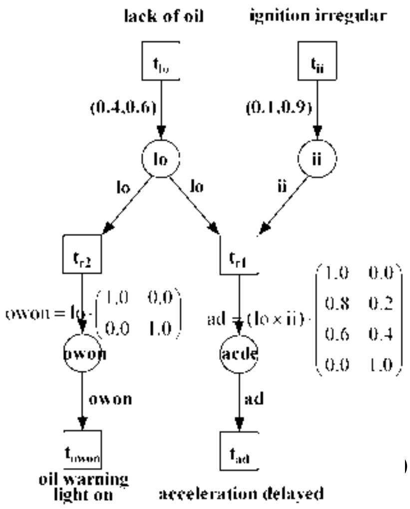


Fig. 6. $\mathcal{FPN}_\alpha$ of Example 6

These values are also the probabilities for the two explanations:

$$\varepsilon_1 : \{lo, igir\} = lo \wedge igir, P(\varepsilon_1) = 0.04$$
$$\varepsilon_2 : \{lo, igno\} = lo \wedge igno, P(\varepsilon_2) = 0.288.$$

Finally, $P(acde \wedge owon) = 0.04 + 0.288 = 0.328$.

□

For the representation of BNs, foldings of PPNs are appropriate. Since we do not need the formal definition, we will be content with an example.

The higher PPN $\mathcal{FPN}_\alpha$ shown in Fig. 6 is a folding of the PPN $\mathcal{PN}_\alpha$ depicted in Fig. 4. The variable mapping of the folding is shown in Table V. Let's assume that the initial marking M_0 is the empty marking. After firing of t_{lo} and t_{ii}, the marking changed into M_1 with

TABLE V
VARIABLE MAPPING OF THE FOLDED NET

$\mathcal{PN}_\alpha$	$\mathcal{FPN}_\alpha$
$lo, nolo$	lo
$igir, igno$	ii
$owon, owof$	$owon$
$acde, acno$	ad
t_{16}, t_{15}	t_{lo}
t_{14}, t_{13}	t_{ii}
t_{20}, t_{19}	t_{owon}
t_{18}, t_{17}	t_{ad}
$t_1, \ldots, t_8$	t_{r_1}
$t_9, \ldots, t_{12}$	t_{r_2}

$$M_1(p) = \begin{cases} (0.4, 0.6) & \text{if } p = lo \\ (0.1, 0.9) & \text{if } p = ii \\ \varnothing & \text{else} \end{cases}$$

If t_{r_1} fires lo and ii are cleared and $ad = (lo \times ii) \cdot \begin{pmatrix} 1.0 & 0.0 \\ 0.8 & 0.2 \\ 0.6 & 0.4 \\ 0.0 & 1.0 \end{pmatrix}$ is put on $acde$;

$$\begin{aligned}(lo \times ii) &= ((0.4, 0.6) \times (0.1, 0.9)) = (0.04, 0.36, 0.06, 0.54) \\ ad &= (0.04, 0.36, 0.06, 0.54) \cdot \begin{pmatrix} 1.0 & 0.0 \\ 0.8 & 0.2 \\ 0.6 & 0.4 \\ 0.0 & 1.0 \end{pmatrix} \\ &= \begin{pmatrix} 0.04 & 0.0 \\ +0.288 & +0.072 \\ +0.036 & +0.024 \\ +0.0 & +0.54 \end{pmatrix} = (0.364, 0.636) \\ &= (P(acde), P(\neg acde)) = (P(acde), P(acno))\end{aligned}$$

(see Example 3).

□

As it is common use in PN theory that foldings of nets of a certain net class are called higher nets, foldings of PPNs are called *higher PPNs*.

IV. Bayesian Networks and Higher Probability Propagation Nets

In this section, we will show how BNs can be represented by higher PPNs. It will turn out that the structure of BNs is a bit meager for modeling directed flows of values (probabilities and likelihoods). Likelihoods are conditional probabilities in a certain interpretation. Let S be a symptom (manifestation) and D be a diagnosis (hypothesis). Then $P(D \mid S)$ is a "diagnostic" probability, $P(S \mid D)$ is a "causal" probability. Bayes' rule combines both probabilities:

$$P(D \mid S) = \frac{P(S \mid D) \cdot P(D)}{P(S)}$$

In case of several conceivable diagnoses $D_i, 1 \leq i \leq n$, $P(S \mid D_i)$ is a measure of how probable it is that D_i causes S. So, $P(S \mid D_i)$ is a degree of confirmation that D_i is the cause for S which is called the *"likelihood of D_i given S"*.

Definition 13 (Bayesian Network) *Let $\mathcal{B} = (R, E)$ be a directed acyclic graph with the set R of nodes and the set E of edges; let for every $r \in R$ $\;par(r)$ be the set of parent nodes of r;*

$\mathcal{B}$ is a Bayesian Network *(*BN*) iff R equals a set of random variables and to every $r \in R$ the table $P(r \mid par(r))$ of conditional probabilities is assigned. $P(r \mid par(r))$ indicates the prior probabilities of r if $par(r) = \emptyset$.*

Definition 14 *Let $A = (a_1, \ldots, a_n), B = (b_1, \ldots, b_n)$ be non-negative real vectors;*

$$\begin{aligned} A \circ B &:= (a_1 \cdot b_1, a_2 \cdot b_2, \ldots, a_n \cdot b_n) \\ A \times B &:= (a_1 \cdot b_1, \ldots, a_1 \cdot b_n, a_2 \cdot b_1, \ldots, \\ &\qquad a_2 \cdot b_n, \ldots, a_n \cdot b_1, \ldots, a_n \cdot b_n) \end{aligned}$$

We will introduce the Petri net representation of BNs by means of examples. The Petri nets are absolutely transparent and reveal the respective situation of algorithms and belief propagation (see [4], [9]).

The following example is a shortened version of the scenario of Example 6.

Example 7
The directed acyclic graph together with the probabilities assigned to the nodes in Fig. 7 is a BN $\mathcal{B}$. Furthermore, it is noted that messages π (probabilities) and λ (likelihoods) flow in both directions via the edges from node to node.

Fig. 8 shows the Petri net representation $\mathcal{PB}$ of the BN $\mathcal{B}$. In order to initialize the net, the transitions π_{lo} and π_{ii} fire, thus putting the tuples $(0.4, 0.6)$ and $(0.1, 0.9)$ on places lo and ii, respectively. $P(lo) = 0.4, P(\neg lo) = 0.6, P(ii) = 0.1, P(\neg ii) = 0.9$ reflect our current feeling about the possibility that lack of oil or irregular ignition happens. Based on these values, the probability of delayed acceleration can be calculated (exactly as in Example 6) by firing of transition

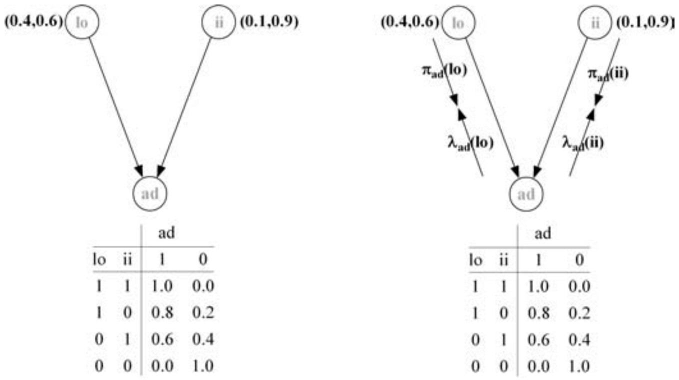


Fig. 7. $\mathcal{B}$ of Example 7

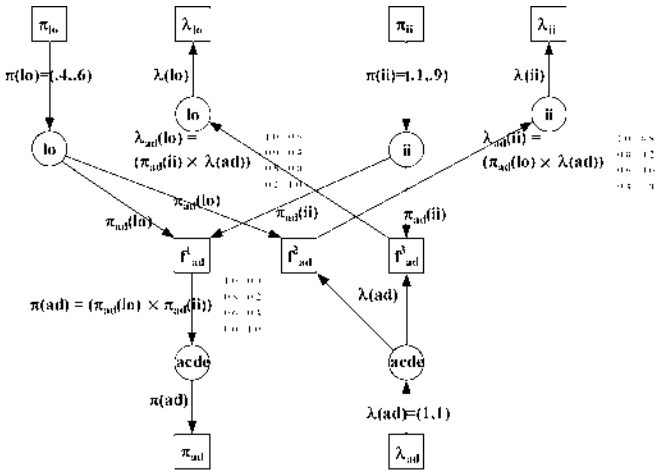


Fig. 8. $\mathcal{PB}$ of Example 7

TABLE VI
THE FUNCTIONS OF $\mathcal{PB}$

$$f^1_{ad} \simeq P(ad \mid lo, ii) = \begin{array}{cc|cc} & & \multicolumn{2}{c}{ad} \\ lo & ii & 1 & 0 \\ \hline 1 & 1 & 1.0 & 0.0 \\ 1 & 0 & 0.8 & 0.2 \\ 0 & 1 & 0.6 & 0.4 \\ 0 & 0 & 0.0 & 1.0 \end{array}$$

$$f^2_{ad} \simeq P^{ii \leftarrow lo,ad}(ad \mid lo, ii) = \begin{array}{cc|cc} & & \multicolumn{2}{c}{ii} \\ lo & ad & 1 & 0 \\ \hline 1 & 1 & 1.0 & 0.8 \\ 1 & 0 & 0.0 & 0.2 \\ 0 & 1 & 0.6 & 0.0 \\ 0 & 0 & 0.4 & 1.0 \end{array}$$

$$f^3_{ad} \simeq P^{lo \leftarrow ii,ad}(ad \mid lo, ii) = \begin{array}{cc|cc} & & \multicolumn{2}{c}{lo} \\ ii & ad & 1 & 0 \\ \hline 1 & 1 & 1.0 & 0.6 \\ 1 & 0 & 0.0 & 0.4 \\ 0 & 1 & 0.8 & 0.0 \\ 0 & 0 & 0.2 & 1.0 \end{array}$$

f^1_{ad}:

$$\begin{aligned} \pi(ad) &= (\pi_{ad}(lo) \times \pi_{ad}(ii)) \cdot \begin{pmatrix} 1.0 & 0.0 \\ 0.8 & 0.2 \\ 0.6 & 0.4 \\ 0.0 & 1.0 \end{pmatrix} \\ &= ((0.4, 0.6) \times (0.1, 0.9)) \cdot \begin{pmatrix} 1.0 & 0.0 \\ 0.8 & 0.2 \\ 0.6 & 0.4 \\ 0.0 & 1.0 \end{pmatrix} \\ &= (0.364, 0.636) \end{aligned}$$

The functions belonging to the respective transitions are shown in Table VI. In the conditional probability table $P(ad \mid lo, ii)$, ad is a function of lo and ii. $P^{ii \leftarrow lo,ad}(ad \mid lo, ii)$ is the re-sorted table $P(ad \mid lo, ii)$ such that now ii is written as a function of lo and ad. In all tables 0.8 for example is the value for $ad = 1, lo = 1, ii = 0$.

To complete the initialization, by firing of $\pi_{lo}, \lambda_{ad}, f^2_{ad}$ we get

$$\begin{aligned}\lambda(ii) = \lambda_{ad}(ii) &= (\pi_{ad}(lo) \times \lambda(ad)) \cdot \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= ((0.4, 0.6) \times (1.0, 1.0)) \cdot \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= (1.0, 1.0).\end{aligned}$$

Similarly:

$$\begin{aligned}\lambda(lo) = \lambda_{ad}(lo) &= (\pi_{ad}(ii) \times \lambda(ad)) \cdot \begin{pmatrix} 1.0 & 0.6 \\ 0.0 & 0.4 \\ 0.8 & 0.0 \\ 0.2 & 1.0 \end{pmatrix} \\ &= ((0.1, 0.9) \times (1.0, 1.0)) \cdot \begin{pmatrix} 1.0 & 0.6 \\ 0.0 & 0.4 \\ 0.8 & 0.0 \\ 0.2 & 1.0 \end{pmatrix} \\ &= (1.0, 1.0).\end{aligned}$$

The likelihood $\lambda(ad) = (1.0, 1.0)$ and, as a consequence, $\lambda(lo) = \lambda(ii) = (1.0, 1.0)$ indicate that there is no reason or evidence to re-asses $\pi(ad), \pi(lo), \pi(ii)$. So, our initial beliefs are

$$\begin{aligned}BEL(lo) &:= \alpha(\lambda(lo) \circ \pi(lo)) = \pi(lo) = (0.4, 0.6) \\ BEL(ii) &:= \alpha(\lambda(ii) \circ \pi(ii)) = \pi(ii) = (0.1, 0.9) \\ BEL(ad) &:= \alpha(\lambda(ad) \circ \pi(ad)) = \pi(ad) = (0.364, 0.636).\end{aligned}$$

Next, we assume the acceleration to be really delayed as a new evidence. So, we set $\lambda(ad) = (1.0, 0.0)$ which results in

$$\begin{aligned}BEL(ad) &= \alpha(\lambda(ad) \circ \pi(ad)) \\ &= \alpha((1.0, 0.0) \circ (0.364, 0.636)) \\ &= \alpha(0.364, 0.0) = (1.0, 0.0)\end{aligned}$$

with the normalizing constant α. As further consequences, the beliefs of lo and ii change. Firing of $\pi(lo), \lambda(ad), f^2_{ad}$ leads to

$$\begin{aligned}\lambda(ii) = \lambda_{ad}(ii) &= (\pi(lo) \times \lambda(ad)) \cdot \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= ((0.4, 0.6) \times (1.0, 0.0)) \cdot \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= (0.76, 0.32)\end{aligned}$$

Firing of $\pi(ii), \lambda(ad), f^3_{ad}$ leads to

$$\begin{aligned}\lambda(lo) = \lambda_{ad}(lo) &= (\pi(ii) \times \lambda(ad)) \cdot \begin{pmatrix} 1.0 & 0.6 \\ 0.0 & 0.4 \\ 0.8 & 0.0 \\ 0.2 & 1.0 \end{pmatrix} \\ &= ((0.1, 0.9) \times (1.0, 0.0)) \cdot \begin{pmatrix} 1.0 & 0.6 \\ 0.0 & 0.4 \\ 0.8 & 0.0 \\ 0.2 & 1.0 \end{pmatrix} \\ &= (0.82, 0.06).\end{aligned}$$

Then the beliefs of lo and ii are

$$\begin{aligned}BEL(lo) &= \alpha(\lambda(lo) \circ \pi(lo)) \\ &= \alpha((0.82, 0.06) \circ (0.4, 0.6)) \\ &= \alpha(0.328, 0.036) = (0.845, 0.155) \\ BEL(ii) &= \alpha(\lambda(ii) \circ \pi(ii)) \\ &= \alpha((0.76, 0.32) \circ (0.1, 0.9)) \\ &= \alpha(0.076, 0.288) = (0.209, 0.791)\end{aligned}$$

In contrast to the initial beliefs, we now strongly believe in a lack of oil ($0.845 > 0.4$) and a little less in a normal ignition ($0.791 < 0.9$).

Lastly, we assume (after an inspection) that there is definitely no lack of oil. So, in addition to $\lambda(ad) = (1.0, 0.0)$ we set $\pi(lo) = (0.0, 1.0)$. Thus, the new belief of lo is $BEL(lo) = (0.0, 1.0)$, and the belief of ii changes:

$$\begin{aligned}\lambda(ii) = \lambda_{ad}(ii) &= (\pi(lo) \times \lambda(ad)) \cdot \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= ((0.0, 1.0) \times (1.0, 0.0)) \cdot \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= (0.6, 0.0).\end{aligned}$$

$$\begin{aligned}BEL(ii) &= \alpha(\lambda(ii) \circ \pi(ii)) = \alpha((0.6, 0.0) \circ (0.1, 0.9)) \\ &= \alpha(0.06, 0.0) = (1.0, 0.0).\end{aligned}$$

If lack of oil is not the reason for the delayed acceleration, we have to believe in an irregular ignition.

□

In Example 2 we found four explanations of a delayed acceleration:

$$\begin{aligned}\varepsilon_5 &= lo \wedge \neg ii, & \varepsilon_6 &= \neg lo \wedge \neg ii, \\ \varepsilon_7 &= lo \wedge ii, & \varepsilon_8 &= \neg lo \wedge ii.\end{aligned}$$

ε_5 is the most probable explanation with $P(\varepsilon_5) = 0.288$. Since the numbers of explanations (like the t-invariants) might grow exponentially, the calculation of all explanations and their probabilities and then looking for the most probable one is obviously no reasonable approach. Much better is a modification of the Petri net approach shown in Example 8. Instead of using the usual matrix product $(A \cdot B)_{ik} = \sum_{j=1}^{n} a_{ij} \cdot b_{jk}$ (sum-product rule), we will use $(A \bullet B)_{ik} = \max\{a_{ij} \cdot b_{jk} \mid j = 1, \ldots, n\}$ (max-product rule; see [4], [9]).

Example 8 (see Example 7)
We start off with the assumption that the acceleration is delayed, i.e. $\lambda(ad) = (1.0, 0.0)$ and $BEL(ad) = (1.0, 0.0)$. Firing of π_{lo} and λ_{ad} puts the tuples $(0.4, 0.6)$ and $(1.0, 0.0)$ on places lo and ad, respectively. Firing of f^2_{ad} takes them away and puts the following tuple on place ii:

$$\begin{aligned}\lambda(ii) = \lambda_{ad}(ii) &= (\pi(lo) \times \lambda(ad)) \bullet \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= ((0.4, 0.6) \times (1.0, 0.0)) \bullet \begin{pmatrix} 1.0 & 0.8 \\ 0.0 & 0.2 \\ 0.6 & 0.0 \\ 0.4 & 1.0 \end{pmatrix} \\ &= (\max\{0.4, 0.0, 0.36\}, \max\{0.32, 0.0\}) \\ &= (0.4, 0.32).\end{aligned}$$

Correspondingly, firing of $\pi_{ii}, \lambda_{ad}, f^3_{ad}$ leads to the following tuple on place lo:

$$\begin{aligned}
\lambda(lo) &= \lambda_{ad}(lo) = (\pi(ii) \times \lambda(ad)) \bullet \begin{pmatrix} 1.0 & 0.6 \\ 0.0 & 0.4 \\ 0.8 & 0.0 \\ 0.2 & 1.0 \end{pmatrix} \\
&= ((0.1, 0.9) \times (1.0, 0.0)) \bullet \begin{pmatrix} 1.0 & 0.6 \\ 0.0 & 0.4 \\ 0.8 & 0.0 \\ 0.2 & 1.0 \end{pmatrix} \\
&= (\max\{0.1, 0.0, 0.72\}, \max\{0.06, 0.0\}) \\
&= (0.72, 0.06).
\end{aligned}$$

The corresponding beliefs are:

$$\begin{aligned}
BEL(lo) &= \alpha(\lambda(lo) \circ \pi(lo)) \\
&= \alpha((0.72, 0.06) \circ (0.4, 0.6)) \\
&= \alpha(0.288, 0.036) \\
&= (\underline{0.889}, 0.111) \\
BEL(ii) &= \alpha(\lambda(ii) \circ \pi(ii)) \\
&= \alpha((0.4, 0.32) \circ (0.1, 0.9)) \\
&= \alpha(0.04, 0.288) \\
&= (0.122, \underline{0.878})
\end{aligned}$$

The most probable explanation given $BEL(ad) = (1.0, 0.0)$ is $lo \wedge \neg ii$ (or $lo \wedge igno$ in the notation of Example 2).

□

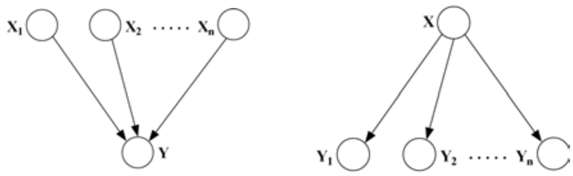
V. Translating Bayesian Networks into Higher Probability Propagation Nets

In this section, we will continue with representing BNs by PNs. In the previous examples we have shown the way higher PPNs work and that they generate the proper values. The next examples are to show that higher PPNs also satisfy the propagation formulas of BNs. We will show that only by means of these examples because the equivalence of the propagation in both approaches is easy to recognize. A formal proof would not be harder, but would also not facilitate the understanding of the equivalence.

We would like to point out that the explicit concept of a situation (marking) in PPNs is a helpful completion to the implicit concept of a situation in BNs.

In essence, there are two structural elements in BNs which are shown in Fig. 9. Fig. 9(a) indicates the probabilistic dependence of Y given $X_1, \ldots, X_n$. So, the probability of Y is defined as $P(Y \mid X_1, \ldots, X_n)$. Case $n = 1$ can be found in Example 9, cases $n = 1$ and $n = 2$ in Example 10.

There are transitions in the PPNs which are closely related to these conditional probabilities. $f^1_Y \simeq P(Y \mid X_1, , \ldots X_n)$ indicates f^1_Y to be the transition that calculates $P(Y \mid X_1, \ldots, X_n)$. The superscript 1 points to a conditional probability. Missing superscripts denote prior probabilities. Superscripts ≥ 2 point to (generalized) transposes of the probability tables (see Tables VIII and XI). For example f^3_A is the transition that belongs to the transpose of ($f^1_A \simeq$) $P(A \mid BC)$ where C is written as depending on A and B, in symbols $P^{C \leftarrow AB}(A \mid BC)$, (e.g. the value belonging to $A = 2, B = 1, C = 2$ is equal to 0.1 in both tables). For $n = 1$ the transposes coincide with the normal transpose of a matrix (see Table XI). Example 10 is to shed light on this substructure and its PN representation.



(a) (b)

Fig. 9. Basic Structures in Bayesian Networks

TABLE VII
Random Variables of Example 9

a_1	spouse is cheating
a_2	spouse is not cheating
b_1	spouse dines with another
b_2	spouse does not dine with another
c_1	spouse is reported seen dining with another
c_2	spouse is not reported seen dining with another
d_1	strange man/lady calls on the phone
d_2	no strange man/lady calls on the phone

The transitions in the PPNs representing structure elements according to Fig. 9(b) cause a component-wise multiplication of vectors with equal length. Let the vectors representing $X, Y_1, \ldots, Y_n$ have length m; then the transition m^1_X transforms the vectors $\lambda_X(Y_1), \ldots, \lambda_X(Y_n)$ into

$$\begin{aligned}
\lambda(X) &= \lambda_X(Y_1) \circ \cdots \circ \lambda_X(Y_n) \\
&= \Big(\prod_{i=1}^{n} (\lambda_X(Y_i))_1, \ldots, \prod_{i=1}^{n} (\lambda_X(Y_i))_m\Big)
\end{aligned}$$

(see Definition 14). m^1_x is the only transition that causes a product of only λ-factors. Superscripts ≥ 2 belong to mixtures of λ- and π-factors. Example 9 is to throw light on this structural element (with $n = m = 2$) and its PN representation.

The next two examples are borrowed from [9] as well as the definitions, propagation formulas and statements we will apply in the sequel. They all are collected in the appendix.

Example 9 (Cheating spouse)
The scenario consists of a spouse and a strange man/lady. It has to be reported that spouse might be cheating. As a consequence, there are four important random variables: spouse is cheating (A), spouse dines with another (B), spouse is reported seen dining with another (C), strange man/lady calls on the phone (D).

The BN with prior and conditional probabilities is shown in Fig. 10. The random variables A, B, C, D have two attributes ($\_1$ and $\_2$ meaning "yes" and "no") which are listed in Table VII.

The functions belonging to the respective transitions are shown in Table VIII. As stated above, the structure of BNs is a bit meager and the actual steps of probability propagation are "hidden" in the algorithms. In contrast to that, the PN representation detailedly shows the probability propagation, and the algorithms are distributed over the net such that each transition's share is of manageable size. In spite of the

TABLE VIII
TRANSITION FUNCTIONS OF EXAMPLE 9

$f_A \simeq P(A) =$

A: 1	A: 2
0.1	0.9

$f_B^1 \simeq P(B \mid A) =$

A	B: 1	B: 2
1	0.7	0.3
2	0.2	0.8

$f_C^1 \simeq P(C \mid B) =$

B	C: 1	C: 2
1	0.4	0.6
2	0.001	0.999

$f_D^1 \simeq P(D \mid A) =$

A	D: 1	D: 2
1	0.8	0.2
2	0.4	0.6

$f_B^2 \simeq P^{A \leftarrow B}(B \mid A) =$

B	A: 1	A: 2
1	0.7	0.2
2	0.3	0.8

$f_C^2 \simeq P^{B \leftarrow C}(C \mid B) =$

C	B: 1	B: 2
1	0.4	0.001
2	0.6	0.999

$f_D^2 \simeq P^{A \leftarrow D}(D \mid A) =$

D	A: 1	A: 2
1	0.8	0.4
2	0.2	0.6

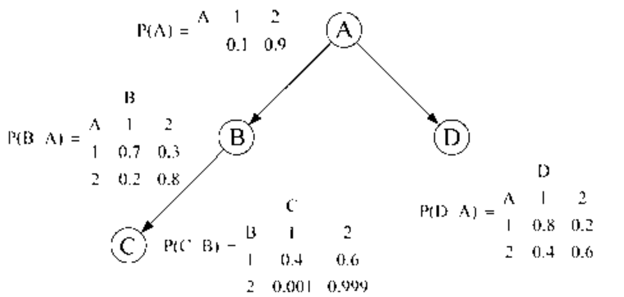


Fig. 10. $\mathcal{B}$ of Example 9

exactness of the PN representation, one might consider the size of the net a disadvantage. On the other hand, the size of the PPNs might indicate that BNs are indeed a little understructured. Moreover, the specific structure of PPNs causes an absolutely appropriate partition into propagation processes. The minimal t-invariants (precisely their net representations) describe exactly the paths of probabilities and likelihoods. The t-invariants need not to be calculated on the higher level. It is sufficient to calculate them on the "black" net (i.e. the underlying place/transition net without arc labels).

The PN representation of the BN in Fig. 10 is shown in Fig. 11. The three minimal t-invariants are shown in Table IX, their net representations in Fig. 12–14. Due to lack of space, the rules for calculating the output tuples of the transitions are missing but they are specified in the net representations of all t-invariants (Fig. 12,13,14). The three minimal t-invariants in vector form are given in Table IX. The t-invariants as

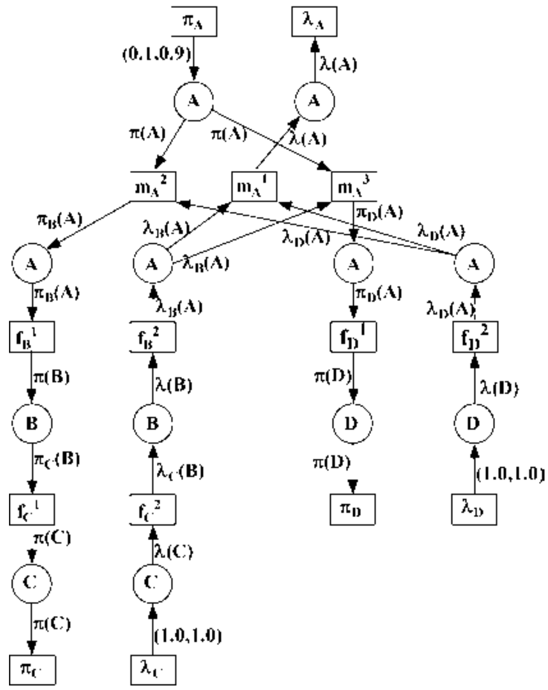


Fig. 11. $\mathcal{PB}$ of Example 9

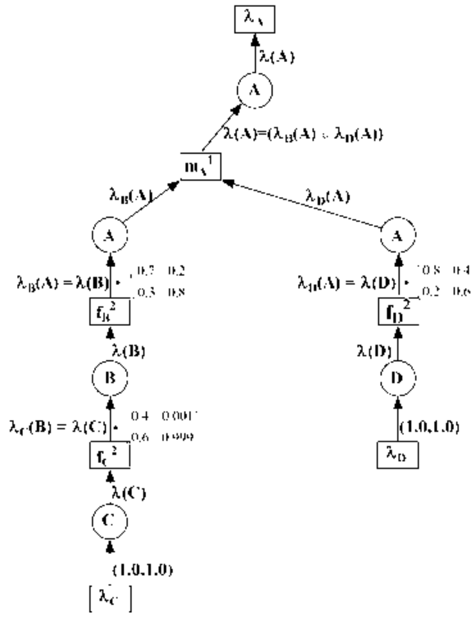


Fig. 12. $\lambda(A)$-t-invariant of $\mathcal{PB}$

solutions of a homogeneous linear equation system result in net representations in which markings are reproducible. In our case, the PNs are cycle-free and have a transition boundary. This implies the reproducibility of the empty marking by every t-invariant as a flow of tuples from input to output boundary. Fortunately, these flows describe exactly the flow of λ- and π-messages. So, the net representations of t-invariants are a framework for the propagation of λ- and π-tuples. We will show that by explaining how the PNs work and refer at each step to the corresponding definitions, lemmas, and propagation formulas of [9] which we collected in the appendix.

TABLE IX
T-INVARIANTS OF $\mathcal{PB}$ (SEE EXAMPLE 9)

	π_A	λ_A	π_C	λ_C	π_D	λ_D	f_B^1	f_B^2	f_C^1	f_C^2	f_D^1	f_D^2	m_A^1	m_A^2	m_A^3
$\lambda(A)$		1		1		1		1		1		1	1		
$\pi(C)$	1		1			1	1		1			1		1	
$\pi(D)$	1			1	1			1		1	1				1

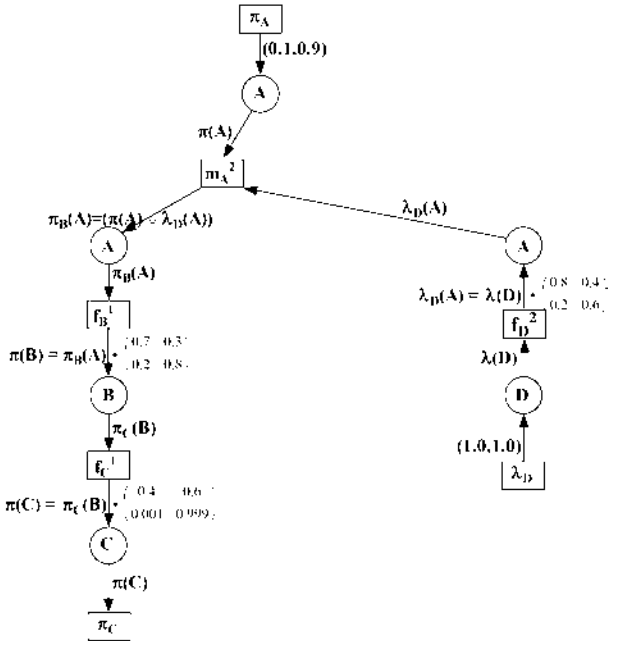


Fig. 13. $\pi(C)$-t-invariant of $\mathcal{PB}$

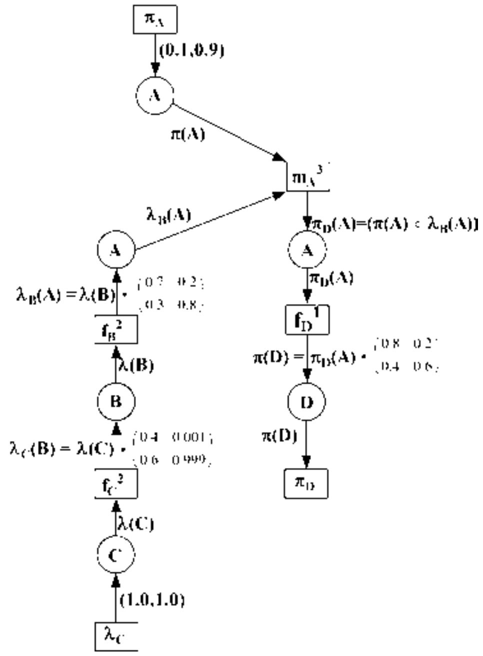


Fig. 14. $\pi(D)$-t-invariant of $\mathcal{PB}$

In the initialization phase, we first set all λ-tuples to $(1.0, 1.0)$ (initialization rule (1), see Appendix, section B), thus expressing that we have no external evidence concerning the random variables. Next, we set $\pi(A) = P(A) = (0.1, 0.9)$ (initialization rule (2)), thus preparing the propagation of $P(a_1) = 0.1, P(a_2) = 0.9$ as a π-message. Finally, we have to calculate $P(C)$ and $P(D)$. For that we will use the $\pi(C)$- and $\pi(D)$-t-invariant (Fig. 13 and 14). In Fig. 13, the boundary transitions π_A and λ_D are enabled because they have no input places. When firing, π_A puts the tuple $\pi(A) = (0.1, 0.9)$ on place A (initialization rule (2)), λ_D puts $\lambda(D) = (1.0, 1.0)$ on place D (initialization rule (1)). Now, transition f_D^2 is enabled and removes $(1.0, 1.0)$ from place D and puts $\lambda_D(A) = \lambda(D) \cdot \left(\begin{smallmatrix}0.8 & 0.4\\ 0.2 & 0.6\end{smallmatrix}\right) = (1.0, 1.0)$ on place A (operative formulas (1), (3), see Appendix, section A). (The boundary transitions π_A and λ_D are permanently enabled. But for reproducing the empty marking they have to fire only once. See Table IX where their entries are 1 in the corresponding row.)

The next transition to fire is m_A^2. It takes $\pi(A) = (0.1, 0.9)$ and $\lambda_D(A) = (1.0, 1.0)$ from the respective places A and puts $\pi_B(A) = \pi(A) \circ \lambda_D(A) = (0.1, 0.9)$ on place A (Definition 16, see Appendix, section C).

(When a tuple is removed from a place X, that does not mean that the values of the tuple are no longer valid for X. It simply says that they have been used and that they can be re-generated any time.)

Next, the transitions f_B^1, f_C^1, π_C are enabled and fire in that sequence. By firing of f_B^1 the tuple $(0.1, 0.9)$ is taken from A, and the tuple $\pi(B) = (0.1, 0.9) \cdot \left(\begin{smallmatrix}0.7 & 0.3\\ 0.2 & 0.8\end{smallmatrix}\right) = (0.25, 0.75)$ is put on B (operative formula (4)). This tuple is taken away by firing of f_C^1 and the tuple $\pi(C) = (0.25, 0.75) \cdot \left(\begin{smallmatrix}0.4 & 0.6\\ 0.001 & 0.999\end{smallmatrix}\right) = (0.10075, 0.89925)$ is put on C (operative formula (4)) from where it is removed by transition π_C, thus completing the reproduction of the empty marking.

The probabilities $P(B)$ and $P(C)$ are calculated as follows:

$$
\begin{aligned}
P(B) &= \alpha\left(\lambda(B) \circ \pi(B)\right) = \alpha\left((1.0, 1.0) \circ (0.25, 0.75)\right)\\
&= (0.25, 0.75)\\
P(C) &= \alpha\left(\lambda(C) \circ \pi(C)\right)\\
&= \alpha\left((1.0, 1.0) \circ (0.10075, 0.89925)\right)\\
&= (0.10075, 0.89925) \text{ (operative formula (5)).}
\end{aligned}
$$

Similarly $P(D) = (0.44, 0.56)$ is calculated on the basis of the $\pi(D)$-invariant (Fig. 14).

Now, we assume that B is instantiated: $\lambda(B) = (1.0, 0.0)$ which means that spouse dines with another. The consequence for the PNs in Fig. 12–14 is quite simple: the variable arc labels $\lambda_C(B)$ and $\pi(B)$ are replaced by the constant tuple $(1.0, 0.0)$. That means whatever the input values enabling f_B^1 and f_C^2 are, both transitions put $(1.0, 0.0)$ on their respective output places B (operative formulas (2), (3)).

The changes of the probabilities are as follows:

- In Fig. 12:

$$\lambda_D(A) = (1.0, 1.0) \text{ as before;}$$
$$\lambda_B(A) = \lambda(B) \cdot \left(\begin{smallmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{smallmatrix}\right) = (1.0, 0.0) \cdot \left(\begin{smallmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{smallmatrix}\right)$$
$$= (0.7, 0.2) \text{ by firing of } f_B^2;$$
$$\lambda(A) = \lambda_B(A) \circ \lambda_D(A) = (0.7, 0.2) \circ (1.0, 1.0)$$
$$= (0.7, 0.2).$$
$$P(A) = \alpha\,(\pi(A) \circ \lambda(A)) = \alpha\,((0.1, 0.9) \circ (0.7, 0.2))$$
$$= \alpha(0.07, 0.18) = (0.28, 0.72) \text{ for } \alpha = \tfrac{1}{0.25}$$

- In Fig. 13: f_B^1 puts the tuple (1.0,1.0) on place B; then f_C^1 is enabled and puts

$$\pi(C) = (1.0, 0.0) \cdot \left(\begin{smallmatrix} 0.4 & 0.6 \\ 0.001 & 0.999 \end{smallmatrix}\right)$$
$$= (0.4, 0.6) \text{ on place } C;$$
$$P(C) = \alpha\,(\lambda(C) \circ \pi(C)) = \alpha\,((1.0, 1.0) \circ (0.4, 0.6))$$
$$= (0.4, 0.6)$$

- In Fig. 14: $\lambda_B(A) = (0.7, 0.2)$ (see above). Firing of m_A^3 puts $\pi_D(A) = \alpha\,(\pi(A) \circ \lambda_B(A)) = \alpha\,((0.1, 0.9) \circ (0.7, 0.2)) = (0.28, 0.72)$ on place A; after firing of f_D^1

$$\pi(D) = \pi_D(A) \cdot \left(\begin{smallmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{smallmatrix}\right) = (0.28, 0.72) \cdot \left(\begin{smallmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{smallmatrix}\right)$$
$$= (0.512, 0.488)$$
$$P(D) = \alpha\,(\lambda(D) \circ \pi(D))$$
$$= \alpha\,((1.0, 1.0) \circ (0.512, 0.488)) = (0.512, 0.488)$$

The interpretation is that after spouse dines with another ($\lambda(B) = (1.0, 0.0)$) the probabilities (our beliefs) that

- spouse is cheating
- spouse is reported seen dining with another
- strange man/lady calls on the phone

are increased, now.

If we, in addition, assume that no strange man/lady calls on the phone ($\lambda(D) = (0.0, 1.0)$) the constant arc weight $(1.0, 1.0)$ at the output arc of transition λ_D has to be changed into $(0.0, 1.0)$. This does not change $P(B)$ and $P(C)$. It changes $P(D)$ and $P(A)$:

$$P(D) = \alpha\,(\lambda(D) \circ \pi(D)) = \alpha\,((0.0, 1.0) \circ (0.512, 0.488))$$
$$= \alpha(0.0, 4.88) = (0.0, 1.0) \text{ for } \alpha = \tfrac{1}{0.488}$$

In the $\lambda(A)$-t-invariant of Fig. 12 we find $\lambda_D(A) = \lambda(D) \cdot f_C^2 = (0.0, 1.0) \cdot \left(\begin{smallmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{smallmatrix}\right) = (0.2, 0.6)$ on place A after firing of transition F_D^2. Then, after firing of transition m_A^1 $\lambda(A) = \lambda_B(A) \circ \lambda_D(A) = (0.7, 0.2) \circ (0.2, 0.6) = (0.14, 0.12)$ was put on A. So, $P(A) = \alpha\,(\pi(A) \circ \lambda(A)) = \alpha\,((0.1, 0.9) \circ (0.14, 0.12)) = \alpha(0.014, 0.108) = (0.1148, 0.8852)$ for $\alpha = \frac{1}{0.122}$.

So, the probability (our belief) that spouse is cheating is decreased.

□

Example 10 (Burglar Alarm)
In this equally very popular example [9], Mr. Holmes is sitting in his office when he gets a call that his burglar alarm is sounding (A). Of course, he suspects a burglary (B) (even

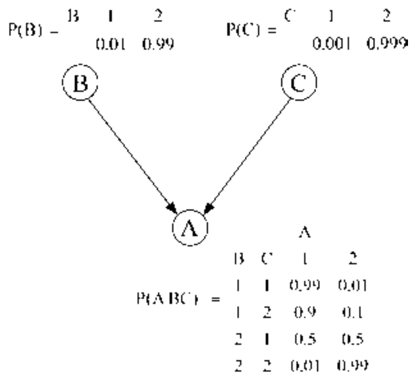


Fig. 15. BN $\mathcal{B}$ of Example 10

TABLE X
RANDOM VARIABLES OF EXAMPLE 10

a_1	Mr. Holmes' burglar alarm sounds
a_2	Mr. Holmes' burglar alarm does not sound
b_1	Mr. Holmes' residence is burglarized
b_2	Mr. Holmes' residence is not burglarized
c_1	there is an earthquake
c_2	there is no earthquake

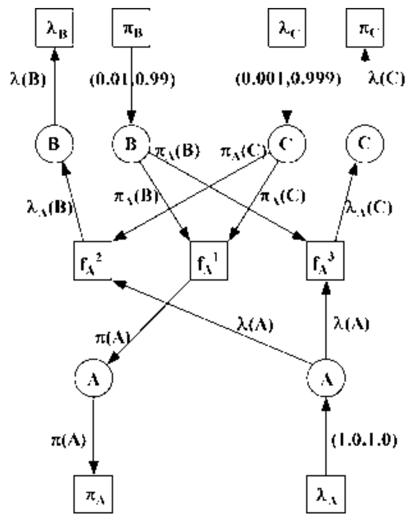


Fig. 16. $\mathcal{PB}$ of Example 10

though there might be other reasons for activating the alarm, e.g. an earthquake (C)). On his ride home, he hears on the radio an announcement about some earthquake. How do the phone call and the radio announcement influence his belief about getting burglarized?

The BN $\mathcal{B}$ with prior and conditional probabilities is shown in Fig. 15. The random variables A, B, C have two attributes ($\_1$ and $\_2$ meaning "yes" and "no") listed in Table X. The PN version $\mathcal{PB}$ of $\mathcal{B}$ is shown in Fig. 16, the corresponding transitions functions in Table XI. Due to lack of space, the rules for calculating $\pi(A), \lambda_A(B), \lambda_A(C)$ are missing in Fig. 16, but they are specified in the t-invariants (Fig. 17–19).

To open the initialization phase, we set all λ-tuples to $(1.0, 1.0)$. In doing so, we state that there is no evidence to change the prior probabilities of B and C (initialization rule (1)). Moreover, we set $\pi(B) = P(B) = (0.01, 0.99)$

TABLE XI
TRANSITION FUNCTIONS OF EXAMPLE 10

$f_B \simeq P(B) =$

B = 1	B = 2
0.01	0.99

$f_C \simeq P(C) =$

C = 1	C = 2
0.001	0.999

$f_A^1 \simeq P(A \mid BC) =$

B	C	A = 1	A = 2
1	1	0.99	0.01
1	2	0.9	0.1
2	1	0.5	0.5
2	2	0.01	0.99

$f_A^2 \simeq P^{B \leftarrow AC}(A \mid BC) =$

A	C	B = 1	B = 2
1	1	0.99	0.5
1	2	0.9	0.01
2	1	0.01	0.5
2	2	0.1	0.99

$f_A^3 \simeq P^{C \leftarrow AB}(A \mid BC) =$

A	B	C = 1	C = 2
1	1	0.99	0.9
1	2	0.5	0.01
2	1	0.01	0.1
2	2	0.5	0.99

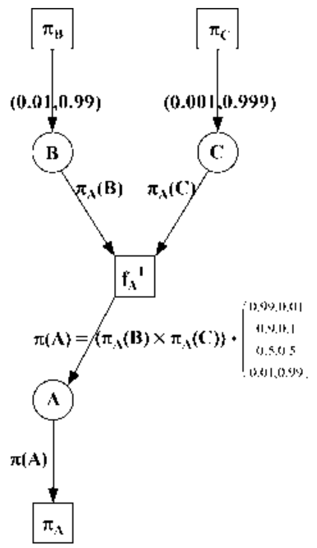


Fig. 17. $\pi(A)$-t-invariant of $\mathcal{PB}$

and $\pi(C) = P(C) = (0.001, 0.999)$ (initialization rule (2)). Next, we calculate $\pi(A)$ by reproducing the empty marking in the $\pi(A)$-t-invariant of Fig. 17. When firing, π_B and π_C put tokens $(0.01, 0.99)$ and $(0.001, 0.999)$ on the places B and C, respectively. Then f_A^1 is activated and fires. By that, the tuples are removed from places B and C and the following tuple is put on place A (operative formula (4)):

$$\begin{aligned}\pi(A) &= (\pi_A(B) \times \pi_A(C)) \cdot \begin{pmatrix} 0.99 & 0.01 \\ 0.9 & 0.1 \\ 0.5 & 0.5 \\ 0.01 & 0.99 \end{pmatrix} \\ &= ((0.01, 0.99) \times (0.001, 0.999)) \cdot \begin{pmatrix} 0.99 & 0.01 \\ 0.9 & 0.1 \\ 0.5 & 0.5 \\ 0.01 & 0.99 \end{pmatrix} \\ &= (0.019, 0.982).\end{aligned}$$

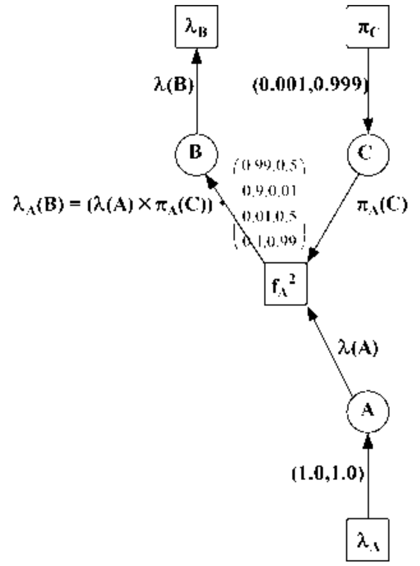


Fig. 18. $\lambda(B)$-t-invariant of $\mathcal{PB}$

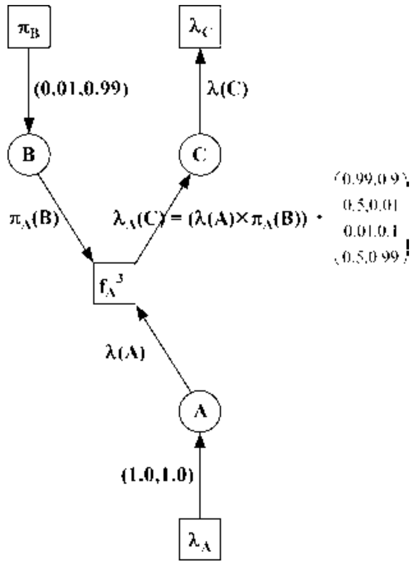


Fig. 19. $\lambda(C)$-t-invariant of $\mathcal{PB}$

So,

$$\begin{aligned}P(A) &= \alpha \cdot (\lambda(A) \circ \pi(A)) \\ &= \alpha\,((1.0, 1.0) \circ (0.019, 0.981)) \\ &= (0.019, 0.982) \text{ for } \alpha = 1\end{aligned}$$

Firing of π_A completes reproducing the empty marking.

Now, we assume that Mr. Holmes got the call and knows that his alarm sounds. That means A has to be instantiated for a_1. Consequently, we have to change the arc label $(1.0, 1.0)$ of arc (λ_A, A) to $(1.0, 0.0)$ in Fig. 16–19. In order to calculate Holmes' present beliefs about B (being burglarized) and C (earthquake), we reproduce the empty marking in the $\lambda(B)$- and $\lambda(C)$-t-invariant (Fig. 18 and 19).

In Fig. 18, after firing λ_A and π_C, tuples $(1.0, 0.0)$ and $(0.001, 0.999)$ are lying on places A and C, respectively. Now, f_A^2 is activated and fires. After that the tuples are removed from places A and C and the following tuple is put on place B (operative formula (1)):

$$\begin{aligned}\lambda_A(B) &= (\lambda(A) \times \lambda_A(C)) \cdot \begin{pmatrix} 0.99 & 0.5 \\ 0.9 & 0.01 \\ 0.01 & 0.5 \\ 0.1 & 0.99 \end{pmatrix} \\ &= ((1.0, 0.0) \times (0.001, 0.999)) \cdot \begin{pmatrix} 0.99 & 0.5 \\ 0.9 & 0.01 \\ 0.01 & 0.5 \\ 0.1 & 0.99 \end{pmatrix} \\ &= (0.9, 0.01).\end{aligned}$$

Removing that tuple by λ_B completes the reproduction of the empty marking.

In the same way, we get $\lambda_A(C) = (0.505, 0.019)$ in the $\lambda(C)$-t-invariant of Fig. 19. This leads to

$$\begin{aligned}P(B) &= \alpha\left((\lambda(B) \circ \pi(B)\right) = \alpha\left((0.9, 0.01) \circ (0.01, 0.99)\right) \\ &= \alpha(0.009, 0.0099) = (0.476, 0.524) \text{ for } \alpha = \tfrac{1}{0.0189} \\ P(C) &= \alpha\left((\lambda(C) \circ \pi(C)\right) \\ &= \alpha\left((0.505, 0.019) \circ (0.001, 0.999)\right) \\ &= \alpha(0.000505, 0.018981) = (0.026, 0.974) \\ &\text{for } \alpha = \tfrac{1}{0.019486}\end{aligned}$$

Holmes' belief in being burglarized has increased from 0.01 to 0.476. Even his belief in an earthquake has increased from 0.001 to 0.026.

Next, we assume that Mr. Homes heard the announcement of an earthquake on the radio. Now, we have to change the arc label $(0.001, 0.99)$ of arc (π_c, C) to $(1.0, 0.0)$ in Fig. 16–19 since C has to be instantiated for c_1. To calculate Mr. Holmes' belief about B (being burglarized) we again reproduce the empty marking in the $\lambda(B)$-t-invariant (Fig. 18). After firing λ_A and π_C, tuples $(1.0, 0.0)$ are lying on places A and C, respectively. After firing f_A^2 these tuples are removed and the following tuple is put on place B:

$$\begin{aligned}\lambda_A(B) &= ((\lambda(A) \times \pi_A(C)) \cdot \begin{pmatrix} 0.99 & 0.5 \\ 0.9 & 0.01 \\ 0.01 & 0.5 \\ 0.1 & 0.99 \end{pmatrix} \\ &= ((1.0, 0.0) \times (1.0, 0.0)) \cdot \begin{pmatrix} 0.99 & 0.5 \\ 0.9 & 0.01 \\ 0.01 & 0.5 \\ 0.1 & 0.99 \end{pmatrix} \\ &= (1.0, 0.0, 0.0, 0.0) \cdot \begin{pmatrix} 0.99 & 0.5 \\ 0.9 & 0.01 \\ 0.01 & 0.5 \\ 0.1 & 0.99 \end{pmatrix} \\ &= (0.99, 0.5).\end{aligned}$$

λ_B fires and removes that tuple from B, thus completing the reproduction of the empty marking. Mr. Holmes' new belief concerning B is

$$\begin{aligned}P(B) &= \alpha\left((\lambda(B) \circ \pi(B)\right) = \alpha\left((0.99, 0.5) \circ (0.01, 0.99)\right) \\ &= \alpha(0.0099, 0.495) = (0.02, 0.98) \text{ for } \alpha = \tfrac{1}{0.5049}.\end{aligned}$$

So, Holmes' belief in being burglarized has changed from 0.01 via 0.476 to 0.02. Holmes was worried after the phone call and calmed down after the announcement.

□

VI. Conclusion and Future Work

We introduced probability propagation nets (PPNs) on the basis of PN representation of propositional Horn clauses. This makes it possible to represent deduction (and abduction) processes as reproduction of the empty marking [7]. Touchstones for our approach are the representation of probabilistic Horn abduction and the propagation of λ- and π-messages in BNs.

In our opinion, it is valuable to introduce specific PN concepts into the field of propagations. In particular t-invariants as an elementary means to structure PNs turned out to be quite fruitful. The minimal t-invariants, on the one hand when reproducing the empty marking describe exactly the flows of λ- and π-messages, thus structuring the PNs (and so the BNs) in a very natural way. On the other hand, they reveal the true complexity behind the simply structured BNs.

Also the markings turned out to be useful insofar as they clearly (and completely) partition all flows in easily observable situations. Altogether, the PPNs are an additional means for describing the flows of probability and evidence that yields a lot of clarity.

In the near future, we aim at integrating PN representations of technical processes and probability propagations.

Appendix

A. Operative Formulas in Bayesian Networks

The following formulas used in chapter V are taken from [9].

1) If B is a child of A, B has k possible values, A has m possible values, and B has one other parent D, with n possible values, then for $1 \leq j \leq m$ the λ message from B to A is given by

$$\lambda_B(a_j) = \sum_{p=1}^{n} \pi_B(d_P)\Big(\sum_{i=1}^{k} P(b_i|a_j, d_p)\lambda(b_i)\Big). \quad (1)$$

2) If B is a child of A and A has m possible values, then for $1 \leq j \leq m$ the π message from A to B is given by

$$\pi_B(a_j) = \begin{cases} 1 & \text{if } A \text{ is instantiated for } a_j \\ 0 & \text{if } A \text{ is instantiated,} \\ & \text{but not for } a_j \\ \frac{P'(a_j)}{\lambda_B(a_j)} & \text{if } A \text{ is not instantiated,} \end{cases} \quad (2)$$

where $P'(a_j)$ is defined to be the current conditional probability of a_j based on the variables thus far instantiated.

3) If B is a variable with k possible values, $s(B)$ is the set of B's children, then for $1 \leq i \leq k$ the λ value of B is given by

$$\lambda(b_i) = \begin{cases} \prod_{C \in s(B)} \lambda_C(b_i) & \text{if } B \text{ is not} \\ & \text{instantiated} \\ 1 & \text{if } B \text{ is instantiated} \\ & \text{for } b_i \\ 0 & \text{if } B \text{ is instantiated,} \\ & \text{but not for } b_i. \end{cases} \quad (3)$$

4) If B is a variable with k possible values and exactly two parents, A and D, A has m possible values, and D has n possible values, then for $1 \leq i \leq k$ the π value of B is given by

$$\pi(b_i) = \sum_{j=1}^{m} \sum_{p=1}^{n} P(b_i|a_j, d_p)\pi_B(a_j)\pi_B(d_p). \quad (4)$$

5) If B is a variable with k possible values, then for $1 \leq i \leq k$, $P'(B_i)$, the conditional probability of b_i based on the variables thus far instantiated, is given by

$$P'(b_i) = \alpha\lambda(b_i)\pi(b_i). \quad (5)$$

B. Initialization in Bayesian Networks

The following rules taken from [9] describe the initialization phase for BNs:

1) Set all λ values, λ messages and π messages to 1.
2) For all roots A, if A has m possible values, then for $1 \leq j \leq m$, set $\pi(a_j) = P(a_j)$.

C. Definition of λ and π Messages

The following definitions are, again, taken from [9].

Definition 15 (λ message) *Let $C = (V, E, P)$ be a causal network in which the graph is a tree, W a subset of instantiated variables, $B \in V$ a variable with k possible values, and $C \in s(B)$ a child of B with m possible values. Then for $1 \leq i \leq k$, we define*

$$\lambda_C(b_i) = \sum_{j=1}^{m} P(c_j|b_i)\lambda(c_j).$$

The entire vector of values $\lambda_C(b_i)$ for $1 \leq i \leq k$, is called the λ message *from C* to *B and is denoted $\lambda_C(B)$.*

Definition 16 (π message) *Let $C = (V, E, P)$ be a causal network in which the graph is a tree, W a subset of instantiated variables, $B \in V$ a variable which is not the root, and $A \in V$ the father of B. Suppose A has m possible values. Then we define for $1 \leq j \leq m$*

$$\pi_B(a_j) = \begin{cases} 1 & \textit{if } A \textit{ is instantiated for } a_j \\ 0 & \textit{if } A \textit{ is instantiated, but not for } a_j \\ \pi(a_j) \prod_{\substack{C \in s(A) \\ C \neq B}} \lambda_C(a_j) & \textit{if } A \textit{ is not instantiated,} \end{cases}$$

where $\lambda_C(a_j)$ is defined in definition 15. Again, if there are no terms in the product, it is meant to represent the value 1. The entire vector of values, $\pi_B(a_j)$ for $1 \leq j \leq m$, is called the π message *from A* to *B and is denoted $\pi_B(A)$.*

REFERENCES

[1] D. Poole, "Logic programming, abduction and probability: a top-down anytime algorithm for estimating prior and posterior probabilities," *New Generation Computing*, vol. 11, no. 3-4, pp. 377–400, 1993.

[2] ——, "Probabilistic Horn abduction and Bayesian networks," *Artificial Intelligence*, vol. 64, no. 1, pp. 81–129, 1993.

[3] L. Portinale and P. Torasso, "A comparative analysis of Horn models and Bayesian networks for diagnosis," in *Advances in Artificial Intelligence 1997, Rome, Italy*, ser. Lecture Notes in Computer Science, M. Lenzerini, Ed., vol. 1321. Springer, 1997, pp. 254–265.

[4] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988.

[5] L. W. Wagenhals and A. H. Lewis, "Converting influence nets with timing information to a discrete event system model, a coloured Petri net," in *Second Workshop on Practical Uses of Coloured Petri Nets and Design/CPN, Aarkus, Denmark, Proceedings*, 1999.

[6] S. Vanit-Anunchai and J. Billington, "Modelling probabilistic inference using coloured Petri nets and factor graphs," in *Fifth Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools, Aarkus, Denmark, October 8–11, 2004, Proceedings*, 2004, pp. 117–136.

[7] K. Lautenbach, "Logical reasoning and Petri nets." in *Applications and Theory of Petri Nets 2003, 24th International Conference, ICATPN 2003, Eindhoven, Netherlands, June 23–27, 2003, Proceedings*, ser. Lecture Notes in Computer Science, W. M. P. van der Aalst and E. Best, Eds., vol. 2679. Springer, 2003, pp. 276–295.

[8] K. Lautenbach, S. Philippi, and A. Pinl, "Bayesian networks and Petri nets," in *Proceedings of the workshop "Entwurf komplexer Automatisierungssysteme" (EKA) 2006*, Braunschweig, 2006.

[9] R. E. Neapolitan, *Probabilistic reasoning in expert systems: theory and algorithms*. New York, NY, USA: John Wiley & Sons, Inc., 1990.

[10] F. V. Jensen, *Introduction to Bayesian Networks*. London: UCL Press, 1996.

[11] R. E. Neapolitan, *Learning Bayesian Networks*. Upper Saddle River, NJ: Prentice Hall, 2003.

[12] S. Vanit-Anunchai and J. Billington, "Modelling dynamic influence diagrams using coloured Petri nets," in *Design, Analysis and Simulation of Distributed Systems Symposiumm, The Spring Multiconference, San Diego, USA, April 3–7, 2005*, 2005, pp. 98–106.

Kurt Lautenbach studied mathematics at the university of Cologne, Germany, and got his PhD degree at the university of Bonn, Germany. From 1979 till 1988, he was a professor for computer science at the university of Bonn. Since 1988 he is a professor for computer science at the university of Koblenz-Landau, Germany.

Alexander Pinl received his diploma in information technology (similar to MSc degree) at the University of Koblenz-Landau, Germany, in 2004. After collecting some experiences as a project manager for an enterprise specialized in networking and content management topics, he returned to the university and started research and implementation work for a DFG-funded project covering safety and reliability modeling in railway applications. In addition, he is currently working toward the PhD degree.

Bisher erschienen

Arbeitsberichte aus dem Fachbereich Informatik
(http://www.uni-koblenz.de/fb4/publikationen/arbeitsberichte)

Kurt Lautenbach, Alexander Pinl: Probability Propagation Nets, Arbeitsberichte aus dem Fachbereich Informatik 20/2007

Rüdiger Grimm, Farid Mehr, Anastasia Meletiadou, Daniel Pähler, Ilka Uerz: SOA-Security, Arbeitsberichte aus dem Fachbereich Informatik 19/2007

Christoph Wernhard: Tableaux Between Proving, Projection and Compilation, Arbeitsberichte aus dem Fachbereich Informatik 18/2007

Ulrich Furbach, Claudia Obermaier: Knowledge Compilation for Description Logics, Arbeitsberichte aus dem Fachbereich Informatik 17/2007

Fernando Silva Parreiras, Steffen Staab, Andreas Winter: TwoUse: Integrating UML Models and OWL Ontologies, Arbeitsberichte aus dem Fachbereich Informatik 16/2007

Rüdiger Grimm, Anastasia Meletiadou: Rollenbasierte Zugriffskontrolle (RBAC) im Gesundheitswesen, Arbeitsberichte aud dem Fachbereich Informatik 15/2007

Ulrich Furbach, Jan Murray, Falk Schmidsberger, Frieder Stolzenburg: Hybrid Multiagent Systems with Timed Synchronization-Specification and Model Checking, Arbeitsberichte aus dem Fachbereich Informatik 14/2007

Björn Pelzer, Christoph Wernhard: System Description:“E-KRHyper“, Arbeitsberichte aus dem Fachbereich Informatik, 13/2007

Ulrich Furbach, Peter Baumgartner, Björn Pelzer: Hyper Tableaux with Equality, Arbeitsberichte aus dem Fachbereich Informatik, 12/2007

Ulrich Furbach, Markus Maron, Kevin Read: Location based Informationsystems, Arbeitsberichte aus dem Fachbereich Informatik, 11/2007

Philipp Schaer, Marco Thum: State-of-the-Art: Interaktion in erweiterten Realitäten, Arbeitsberichte aus dem Fachbereich Informatik, 10/2007

Ulrich Furbach, Claudia Obermaier: Applications of Automated Reasoning, Arbeitsberichte aus dem Fachbereich Informatik, 9/2007

Jürgen Ebert, Kerstin Falkowski: A First Proposal for an Overall Structure of an Enhanced Reality Framework, Arbeitsberichte aus dem Fachbereich Informatik, 8/2007

Lutz Priese, Frank Schmitt, Paul Lemke: Automatische See-Through Kalibrierung, Arbeitsberichte aus dem Fachbereich Informatik, 7/2007

Rüdiger Grimm, Robert Krimmer, Nils Meißner, Kai Reinhard, Melanie Volkamer, Marcel Weinand, Jörg Helbach: Security Requirements for Non-political Internet Voting, Arbeitsberichte aus dem Fachbereich Informatik, 6/2007

Daniel Bildhauer, Volker Riediger, Hannes Schwarz, Sascha Strauß, „grUML – Eine UML-basierte Modellierungssprache für T-Graphen“, Arbeitsberichte aus dem Fachbereich Informatik, 5/2007

Richard Arndt, Steffen Staab, Raphaël Troncy, Lynda Hardman: Adding Formal Semantics to MPEG-7: Designing a Well Founded Multimedia Ontology for the Web, Arbeitsberichte aus dem Fachbereich Informatik, 4/2007

Simon Schenk, Steffen Staab: Networked RDF Graphs, Arbeitsberichte aus dem Fachbereich Informatik, 3/2007

Rüdiger Grimm, Helge Hundacker, Anastasia Meletiadou: Anwendungsbeispiele für Kryptographie, Arbeitsberichte aus dem Fachbereich Informatik, 2/2007

Anastasia Meletiadou, J. Felix Hampe: Begriffsbestimmung und erwartete Trends im IT-Risk-Management, Arbeitsberichte aus dem Fachbereich Informatik, 1/2007

„Gelbe Reihe“

(http://www.uni-koblenz.de/fb4/publikationen/gelbereihe)

Lutz Priese: Some Examples of Semi-rational and Non-semi-rational DAG Languages. Extended Version, Fachberichte Informatik 3-2006

Kurt Lautenbach, Stephan Philippi, and Alexander Pinl: Bayesian Networks and Petri Nets, Fachberichte Informatik 2-2006

Rainer Gimnich and Andreas Winter: Workshop Software-Reengineering und Services, Fachberichte Informatik 1-2006

Kurt Lautenbach and Alexander Pinl: Probability Propagation in Petri Nets, Fachberichte Informatik 16-2005

Rainer Gimnich, Uwe Kaiser, and Andreas Winter: 2. Workshop ''Reengineering Prozesse'' – Software Migration, Fachberichte Informatik 15-2005

Jan Murray, Frieder Stolzenburg, and Toshiaki Arai: Hybrid State Machines with Timed Synchronization for Multi-Robot System Specification, Fachberichte Informatik 14-2005

Reinhold Letz: FTP 2005 – Fifth International Workshop on First-Order Theorem Proving, Fachberichte Informatik 13-2005

Bernhard Beckert: TABLEAUX 2005 – Position Papers and Tutorial Descriptions, Fachberichte Informatik 12-2005

Dietrich Paulus and Detlev Droege: Mixed-reality as a challenge to image understanding and artificial intelligence, Fachberichte Informatik 11-2005

Jürgen Sauer: 19. Workshop Planen, Scheduling und Konfigurieren / Entwerfen, Fachberichte Informatik 10-2005

Pascal Hitzler, Carsten Lutz, and Gerd Stumme: Foundational Aspects of Ontologies, Fachberichte Informatik 9-2005

Joachim Baumeister and Dietmar Seipel: Knowledge Engineering and Software Engineering, Fachberichte Informatik 8-2005

Benno Stein and Sven Meier zu Eißen: Proceedings of the Second International Workshop on Text-Based Information Retrieval, Fachberichte Informatik 7-2005

Andreas Winter and Jürgen Ebert: Metamodel-driven Service Interoperability, Fachberichte Informatik 6-2005

Joschka Boedecker, Norbert Michael Mayer, Masaki Ogino, Rodrigo da Silva Guerra, Masaaki Kikuchi, and Minoru Asada: Getting closer: How Simulation and Humanoid League can benefit from each other, Fachberichte Informatik 5-2005

Torsten Gipp and Jürgen Ebert: Web Engineering does profit from a Functional Approach, Fachberichte Informatik 4-2005

Oliver Obst, Anita Maas, and Joschka Boedecker: HTN Planning for Flexible Coordination Of Multiagent Team Behavior, Fachberichte Informatik 3-2005

Andreas von Hessling, Thomas Kleemann, and Alex Sinner: Semantic User Profiles and their Applications in a Mobile Environment, Fachberichte Informatik 2-2005

Heni Ben Amor and Achim Rettinger: Intelligent Exploration for Genetic Algorithms – Using Self-Organizing Maps in Evolutionary Computation, Fachberichte Informatik 1-2005