UNIVERSITÄT KOBLENZ · LANDAU

Probability Propagation in Petri Nets

Kurt Lautenbach, Alexander Pinl

16/2005

Fachberichte
INFORMATIK

ISSN 1860-4471

Universität Koblenz-Landau
Institut für Informatik, Universitätsstr. 1, D-56070 Koblenz

E-mail: researchreports@uni-koblenz.de,
WWW: http://www.uni-koblenz.de/FB4/

Probability Propagation in Petri Nets

Kurt Lautenbach and Alexander Pinl

University of Koblenz-Landau, Universitätsstr. 1, 56070 Koblenz, Germany,
{laut,apinl}@uni-koblenz.de,
WWW home page: http://www.uni-koblenz.de/∼ag-pn

Abstract A class of high level Petri nets, called "probability propagation nets", is introduced which is particularly useful for modelling probability and evidence propagation. It is demonstrated how propagation of probabilistic Horn abduction and Bayesian networks can be represented and structured by these nets. Their simplest form is a probabilistic extension of place/transition nets for modelling logical formulae and deductions. As a consequence, the main means for structuring are transition invariants that can easily be calculated for all "probability propagation nets" on a place/transition net level.

1 Introduction

This paper deals with the representation of logical-probabilistic structures and processes by Petri nets with the intention to find a transparent and well structured probability propagation model for the processes in probabilistic Horn abductions (PHAs) [Poo93a],[Poo93b],[PT97] and in Bayesian Networks (BNs) [Pea88].

Roughly speaking, BNs are directed acyclic graphs (DAGs) whose nodes are random variables. Starting at the input boundary nodes, prior probabilities flow to the follower nodes in order to influence their probabilities (or beliefs). Likelihoods flow against the arc direction, thus making the probabilities (or beliefs) of the predecessor nodes topical dependent on new evidences about certain nodes.

What makes it desirable to represent BNs by Petri nets is that the structure of BNs is a bit meager for describing bidirectional propagations which thus are "hidden" in the algorithms. Petri nets, on the other hand, are especially developed for representing all sorts of processes.

In contrast to some existing approaches on combining BNs and Petri nets (e.g. [WL99],[VAB04]), we introduce step by step new Petri nets that are particularly well suited for our intentions: they are transparent and structurable.

First, we modify the p/t-nets for representing logical inference [Lau03] by inscribing tokens and transitions with probabilities. These nets ("probability propagation nets", PPNs) allow to represent stochastic inference and PHA.

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

Third, a possibility to cope with loops in BNs is conditioning (see [Pea88]). Caused by cutting off the loops new BNs arise, thus complicating the structure considerably. A further folding, however, makes it possible to represent that whole structure in a nearly unchanged Petri net. The increased complexity will be found in the markings.

Inspite of the considerable complexity, the Petri nets 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). That means that structuring the Petri net representations by t-invariants can be calculated in the underlying p/t-nets.

Our paper is organized as follows: In section 3 a class of Petri nets is introduced that allows the modelling of PHA. This class of nets is generalized by foldings in section 4 such that it is suitable to model the propagation processes in BNs.

2 Preliminaries

As to preliminaries of place/transition nets (p/t-nets) we refer to [Lau02]. Moreover, the p-column-vector **0** stands for the *empty marking*. A Petri net is **0**-*reproducing* iff there exists a firing sequence φ such that $\mathbf{0}\,[\varphi\rangle\,\mathbf{0}$. A transition t is **0**-*firable* iff t can be enabled by some follower marking of **0**. For multisets, we prefer the linear representation: $a + 3b + 2c$ instead of $\{a, b, b, b, c, c\}$. For some t-invariant I $||I|| = \{t | t$ is a transition with $I(t) \neq 0\}$ is the support of I. $\mathbb{M}(A)$ denotes the set of multisets over some finite set A.

3 Propositional Logic and Probabilistic Horn Abduction

In this section, we introduce into the PHA. Furthermore, we generalize a Petri net representation of propositional logic inference by inscribing probabilities on tokens and transitions. On the whole, we are influenced by [Poo93a],[Poo93b], and [PT97].

Definition 1. *The* alphabet *of propositional logic consists of* atoms $a, b, c, \ldots,$ operators $\neg, \vee, \wedge,$ *and* brackets (*and*).

The formulae α *are exactly the words which can be constructed by means of the following rules:*

- *all* atoms *are formulae;*
- *if* α *is a formula, the* negation $\neg\alpha$ *is a formula, too;*
- *if* α *and* β *are formulae, the* conjunction $(\alpha \wedge \beta)$ *and the* disjunction $(\alpha \vee \beta)$ *are formulae, too.*

The implication $(\alpha \to \beta)$ *is an abbreviation for* $((\neg\alpha) \vee \beta)$*; the* biimplication $(\alpha \leftrightarrow \beta)$ *is an abbreviation for* $((\alpha \to \beta) \wedge (\beta \to \alpha))$*.*

For omitting brackets, we stick to the usual operator hierarchy.

$\mathbb{A}(\alpha)$ *denotes the* set of atoms *contained in a formula* α*.*

Definition 2. *A* literal *is an atom or the negation of an atom.*

A clause *is a disjunction of literals. Usually,* $\Box$ *denotes the* empty clause*.*

Let $\tau = \neg a_1 \vee \cdots \vee \neg a_m \vee b_1 \vee \cdots \vee b_n$ *be a clause. Often a set notation is used:* $\tau = \{\neg a_1, \ldots, \neg a_m, b_1, \ldots, b_n\}$ *or* $\tau = \neg A \cup B$ *for* $\neg A = \{\neg a_1, \ldots, \neg a_m\}$ *and* $B = \{b_1, \ldots, b_n\}$*.*

If in $\tau = (\neg A \cup B)$ $\quad \neg A$ *is empty,* τ *is a* fact clause*; if* B *is empty,* τ *is a* goal clause*.*

A formula α *is in* conjunctive normal form *(is a* CNF-formula*) if it is a conjunction of clauses.* $\mathbb{C}(\alpha)$ *denotes the* set of clauses *contained in a CNF-formula* α*.* $\mathbb{F}(\alpha)$ *denotes the* set of fact clauses*.* $\mathbb{G}(\alpha)$ *denotes the* set of goal clauses*.* $\mathbb{R}(\alpha) = \mathbb{C}(\alpha) - (\mathbb{F}(\alpha) \cup \mathbb{G}(\alpha))$ *denotes the* set of rule clauses *in* α*.*

Definition 3. *Let* α *be a formula. A mapping* $I : \mathbb{A}(\alpha) \to \{\mathit{true}, \mathit{false}\}$ *is an* interpretation *of* α*.*

An interpretation of α *such that* α *is true is a* model *of* α*.*

If α *has no models,* α *is* contradictory*. If* α *has at least one model,* α *is* satisfiable*. If all interpretations of* α *are models,* α *is* tautological*.*

If for two formulae α *and* β $\quad (\alpha \leftrightarrow \beta)$ *is tautological,* α *and* β *are* equivalent*.*

If for some formulae $\alpha_1, \ldots, \alpha_n, \beta$ $\quad ((\alpha_1 \wedge \cdots \wedge \alpha_n) \to \beta)$ *is tautological* β *is a* logical consequence *of* $\alpha_1, \ldots, \alpha_n$*.*

Definition 4 (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* ${}^\bullet\tau = \{a_1, \ldots, a_m\}$*,* $\tau^\bullet = \{b_1, \ldots, b_n\}$*.*

Definition 5. *A clause* κ *is a* Horn clause *iff it contains at most one positive literal.*

A CNF-formula is a Horn formula *iff its clauses are Horn clauses.*

The canonical representation of a Horn clause is a Horn transition*.*

Theorem 1. *Let* α *be a Horn formula and* $\mathcal{N}_\alpha = (S_\alpha, T_\alpha, F_\alpha)$ *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 [Lau03].

Definition 6. *Let* α *be a Horn formula and* $P_\alpha : \mathbb{C}(\alpha) \to [0,1]$ *a real function, called a* probability function of α*;*

let $H \subseteq \mathbb{F}(\alpha)$ *be a set of fact clauses, called a set of* "assumable" hypotheses*;*

let $\{D_1, \ldots, D_n\}$ *be a partition of* H *(i.e.* $D_i \cap D_j = \emptyset$ *for* $i \neq j, \bigcup_{i=1}^n D_i = H$*) where for all* $D_i, 1 \leq i \leq n, \sum_{f_i \in D_i} P_\alpha(f) = 1$*; the sets* $D_1, \ldots, D_n$ *are called the* disjoint classes*;*

let be $E \subseteq H$*,* $R \subseteq \mathbb{R}(\alpha) \cup \mathbb{F}(\alpha)$*,* $\gamma \in \mathbb{G}(\alpha)$ *and let* $\varepsilon = \bigwedge_{c \in E} c$*,* $\varrho = \bigwedge_{c \in R} c$ *be the corresponding Horn formulae;*

ε *is an* explanation (diagnosis) of $\neg\gamma$ *iff*

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

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*.*

Example 1. Let be $\alpha = A \wedge B \wedge (\neg A \vee \neg B \vee C) \wedge \neg C$ a Horn formula.

The canonical net representation $\mathcal{N}_\alpha = (S_\alpha, T_\alpha, F_\alpha)$ is shown in figure 1(a), where

- $S_\alpha = \mathbb{A}(\alpha) = \{A, B, C\}$ is the set of atoms and
- $T_\alpha = \mathbb{C}(\alpha) = \{\tau_1, \tau_2, \tau_3, \tau_4, \}$ for $\tau_1 = A$, $\tau_2 = B$, $\tau_3 = \neg A \vee \neg B \vee C$, $\tau_4 = \neg C$ is the set of clauses.

The probability function P_α is given by $P_\alpha(\tau_1) = 0.2$, $P_\alpha(\tau_2) = 0.3$, $P_\alpha(\tau_3) = 0.4$, $P_\alpha(\tau_4) = 1$.

$\mathbb{F}(\alpha) = \{\tau_1, \tau_2\}$ is the set of facts, $\mathbb{R}(\alpha) = \{\tau_3\}$ is the set of rules, $\mathbb{G}(\alpha) = \{\tau_4\}$ is the set of goals. Disjoint classes are not given.

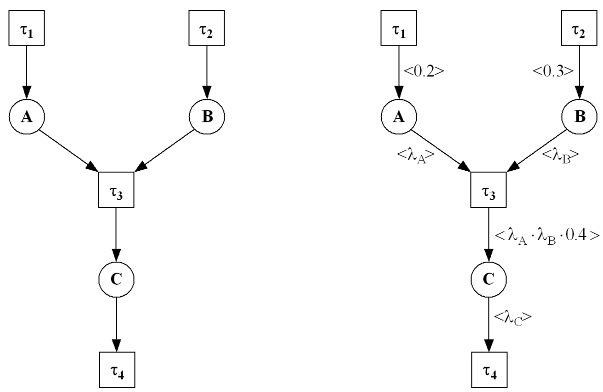
In terms of PHA, $\varepsilon = A \wedge B$ is an explanation of $\neg\gamma = C$ because $\varepsilon \wedge \varrho = A \wedge B \wedge (\neg A \vee \neg B \vee C) = A \wedge B \wedge C$ is not contradictory and $\neg\gamma = C$ is a logical consequence of $\varepsilon \wedge \varrho = A \wedge B \wedge C$.

$I = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{matrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{matrix}$ is the only (minimal) t-invariant of $\mathcal{N}_\alpha$.

$\prod_{\tau \in \|I\| \setminus \{\tau_4\}} P_\alpha(\tau) = 0.2 \cdot 0.3 \cdot 0.4 = 0.024$ is the probability of ε and of $\neg\gamma$ w.r.t. I.
□

Even though the canonical net representation is quite useful for representing and visualizing logical inference, it is not yet suited for modelling the propagation of probabilities. So, we "enrich" the tokens of canonical net representations by probability values. This results in 1-tuples $<p>$, $p \in [0,1]$, for which we postulate the following rule:

Let A be a finite multiset over $[0,1]$; then $\prod_{p\in A} <p> := <\prod_{p\in A} p>$ shall hold, saying that multiplying 1-tuples results in a 1-tuple.



(a) $\mathcal{N}_\alpha$ (b) $\mathcal{PN}_\alpha$

Figure 1. Canonical net representation and corresponding PPN

Definition 7. *Let* α *be a Horn formula;* $\mathcal{PN}_\alpha = (S_\alpha, T_\alpha, F_\alpha, P_\alpha, L_\alpha)$ *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*

$$L_\alpha(f) := \begin{cases} <P_\alpha(\tau)> & \text{if } f = (\tau, b) \in F_\alpha \cap (T_\alpha \times S_\alpha) \\ & \text{and } \tau \in \mathbb{F}(\alpha), \\ <\lambda> & \text{if } f = (a, \tau) \in F_\alpha \cap (S_\alpha \times T_\alpha) \\ & \text{and } \tau \in \mathbb{R}(\alpha) \cup \mathbb{G}(\alpha) \\ & \text{and } \lambda \text{ is a variable ranging over } [0,1], \\ <P_\alpha(\tau)> \cdot \prod_{a \in {}^\bullet\tau} L_\alpha(a, \tau) & \text{if } f = (\tau, b) \in F_\alpha \cap (T_\alpha \times S_\alpha) \\ & \text{and } \tau \in \mathbb{R}(\alpha). \end{cases}$$

Definition 8. *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 $A \subseteq [0,1]$ *be finite, and* *let* $<A> := \{<a> \mid a \in A\}$ *be the corresponding set of 1-tuples*

$M : S_\alpha \to \mathbb{M}(<A>)$ *is a* marking *of* $\mathcal{PN}_\alpha$ *;*

let be $\tau \in T_\alpha$ *and* ${}^\bullet\tau = \{s_1, \dots, s_m\}, \tau^\bullet = \{s_{m+1}\}$*;*

τ *is* enabled *for* $\{<a_1>, \dots, <a_m>\}$ *by* M *iff* $<a_1> \in M(s_1), \dots, <a_m> \in M(s_m)$*,*

the follower marking M' after firing of τ *for* $\{<a_1>, \dots, <a_m>\}$ *is given by*

$$\begin{aligned} M'(s_1) = M(s_1) - <a_1>, \dots, M'(s_m) = M(s_m) - <a_m>, \\ M'(s_{m+1}) = M(s_{m+1}) + <a_1 \cdot a_2 \dots a_m \cdot P_\alpha(\tau)>, \end{aligned}$$

if $<\lambda_1>, \dots, <\lambda_m>$ *are the arc labels of* $(s_1, \tau), \dots, (s_m, \tau) \in F_\alpha$*, we may write*

$$\begin{aligned} M'(s_1) = M(s_1) - <\lambda_1>, \dots, M'(s_m) = M(s_m) - <\lambda_m>, \\ M'(s_{m+1}) = M(s_{m+1}) + <\lambda_1 \dots \lambda_m \cdot P_\alpha(\tau)>, \end{aligned}$$

if the λ_i *are bound by the corresponding* $a_i, 1 \le i \le m$*.*

Example 2 (see example 1).

The PPN $\mathcal{PN}_\alpha = (S_\alpha, T_\alpha, F_\alpha, P_\alpha, L_\alpha)$ (see figure 1(b)) is based on the Horn formula $\alpha = A \wedge B \wedge (\neg A \vee \neg B \vee C) \wedge \neg C$ and the canonical net representation $\mathcal{N}_\alpha$ (see example 1, figure 1(a)). The probability function $P_\alpha(\tau)$ and the arc label function $L_\alpha(f)$ are given by:

$$P_\alpha(\tau) = \begin{cases} 0.2 & \text{if } \tau = \tau_1 \\ 0.3 & \text{if } \tau = \tau_2 \\ 0.4 & \text{if } \tau = \tau_3 \\ 1.0 & \text{if } \tau = \tau_4 \end{cases} \qquad L_\alpha(f) = \begin{cases} <0.2> & \text{if } f = (\tau_1, A) \\ <0.3> & \text{if } f = (\tau_2, B) \\ <\lambda_A> & \text{if } f = (A, \tau_3) \\ <\lambda_B> & \text{if } f = (B, \tau_3) \\ <\lambda_C> & \text{if } f = (C, \tau_4) \\ <\lambda_A \cdot \lambda_B \cdot 0.4> & \text{if } f = (\tau_3, C) \end{cases}$$

$M_0 = \begin{pmatrix} \emptyset \\ \emptyset \\ \emptyset \end{pmatrix} = \mathbf{0}$ is the initial marking by which τ_1 and τ_2 are enabled because they have no input places. The marking after firing of τ_1 and τ_2 is $M_1 = \begin{pmatrix} <0.2> \\ <0.3> \\ \emptyset \end{pmatrix}$ by which now τ_3 is enabled for $\lambda_A = 0.2, \lambda_B = 0.3$. After firing of τ_3 the marking is $M_2 = \begin{pmatrix} \emptyset \\ \emptyset \\ <0.2 \cdot 0.3 \cdot 0.4> \end{pmatrix} = \begin{pmatrix} \emptyset \\ \emptyset \\ <0.024> \end{pmatrix}$. Firing of τ_4 leads back to $M_0 = \mathbf{0}$.

The constituents of a PHA problem are $\varepsilon = A \wedge B$, $\varrho = \neg A \vee \neg B \vee C$, $\gamma = \neg C$. $\neg\gamma$ is a logical consequence of $\varepsilon \wedge \varrho = A \wedge B \wedge (\neg A \vee \neg B \vee C) = A \wedge B \wedge C$, so ε is an explanation of $\neg\gamma = C$.

$I = (1, 1, 1, 1)^t$ is a t-invariant of $\mathcal{N}_\alpha$, so $P_\alpha(\tau_1) \cdot P_\alpha(\tau_2) \cdot P_\alpha(\tau_3) = 0.024$ is the probability of $\varepsilon = A \wedge B$ and $of \neg\gamma = C$ w.r.t. I.

□

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

Table 1. Horn clauses of example 3

Example 3 (cf. [PT97]). Let α be the Horn formula that is the conjunction of the clauses given in table 1 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).

Figure 2 shows the canonical net representation $\mathcal{N}_\alpha$ of α (see definition 4). The "characteristic" sets are:

- $H = \mathbb{F}(\alpha) = \{igno, igir, nolo, lo\}$,
- $G = \mathbb{G}(\alpha) = \{\neg acno, \neg acde, \neg owof, \neg owon\}$,
- $R = \mathbb{R}(\alpha) \cup \mathbb{F}(\alpha)$.

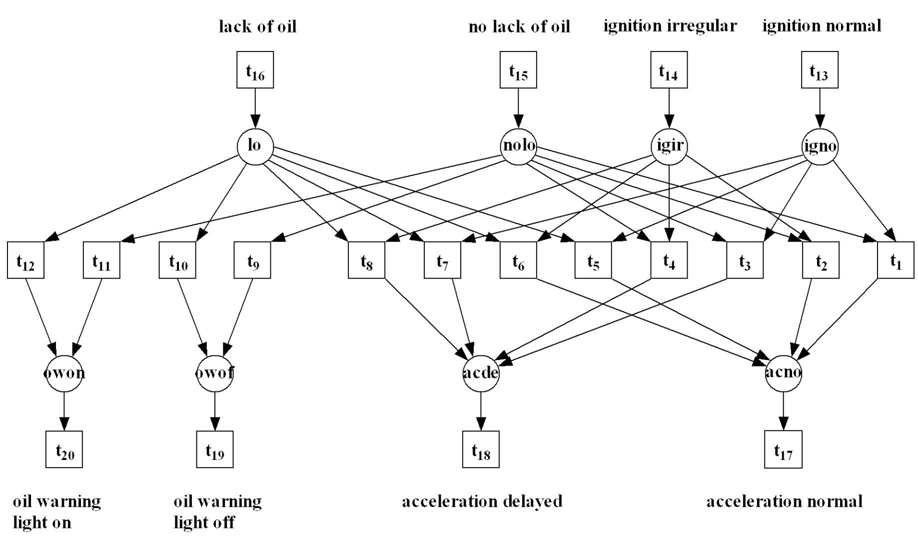


Figure 2. $\mathcal{N}_\alpha$ of example 3

	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

Table 2. T-invariants of $\mathcal{N}_\alpha$ (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	D_3	D_3	D_4	D_4

Table 3. Probability function P_α of example 3

The t-invariants of $\mathcal{N}_\alpha$ and the probability function P_α with four disjoint classes are shown in tables 2 and 3, respectiveley. Finally, the corresponding PPN is given in figure 3.

To demonstrate how the PPNs should be applied, we calculate the probabilities of *acde* and its explanations. A considerable advantage is that the (minimal) t-invariants can be calculated in $\mathcal{N}_\alpha$ instead of $\mathcal{PN}_\alpha$. There are four t-invariants passing through $t_{18} = \neg acde$: $\{I_i | 1 \leq i \leq 12, I_i(t_{18}) \neq 0\} = \{I_5, I_6, I_7, I_8\}$ (see table 2). So, the explanations of *acde* are $\varepsilon_i = \|I_i\| \cap \mathbb{F}(\alpha)$ for $5 \leq i \leq 8$:

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

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 $P_\alpha(t) = 1$ holds.)

$$P(\varepsilon_5) = 0.9 \cdot 0.4 \cdot 0.8 \cdot 1 = 0.288 \qquad P(\varepsilon_7) = 0.1 \cdot 0.4 \cdot 1 \cdot 1 = 0.04$$
$$P(\varepsilon_6) = 0.9 \cdot 0.6 \cdot 0 \cdot 1 = 0 \qquad P(\varepsilon_8) = 0.1 \cdot 0.6 \cdot 0.6 \cdot 1 = 0.036$$

$P(acde)$ 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.

Now, we want to calculate the probabilities of $acde \wedge owon$. For that we modify figure 3 in several steps:

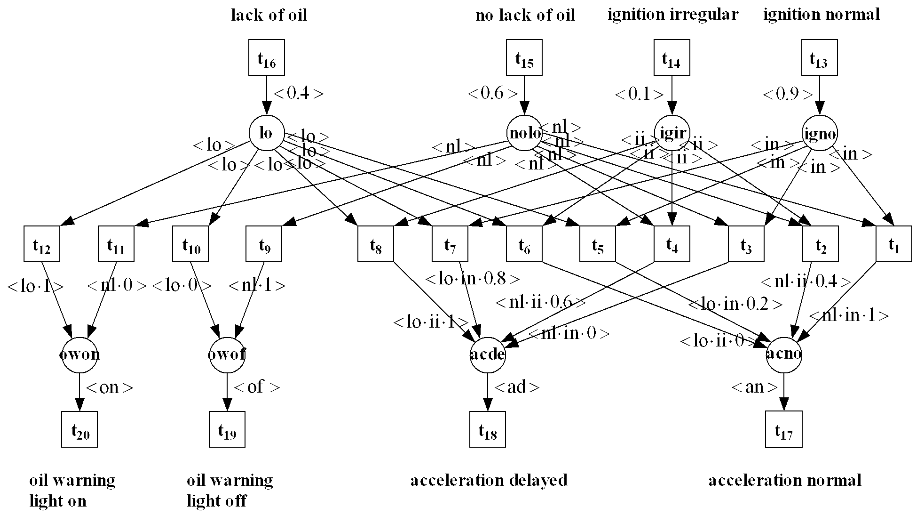


Figure 3. $\mathcal{PN}_\alpha$ of example 3

- transitions (goal clauses) $t_{18} = \neg acde$ and $t_{20} = \neg owon$ are unified to one transition (goal clause) $t_{20} = \neg acde \vee \neg owon = \neg(acde \wedge owon)$;
- the probabilities $P_\alpha(t_{16}) = 0.4$ and $P_\alpha(t_{15}) = 0.6$ are changed into $P_\alpha(t_{16}) = 1, P_\alpha(t_{15}) = 0$;
- all t-invariants with transitions t, where $P_\alpha(t) = 0$ are omitted;
- the transitions $t_{19} = \neg owof$ and $t_{17} = \neg acno$ are omitted because they are not needed any more.

	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

Table 4. T-invariants of $\mathcal{PN}'_\alpha$ (example 3)

The result ist the PPN $\mathcal{PN}'_\alpha$ shown in figure 4. From a structural point of view, this net is well suited for solving our problem because its set of t-invariants (see table 4) 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 [Pea88]. 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 4). This, in principle, leads to a double effect of <*lo*> when t_{20} fires (via *owon* and via *acde*). Since $lo = 1$, however, this effect is neutralized. So, by simulating or simply multiplying the probabilities, we get for the t-invariants:

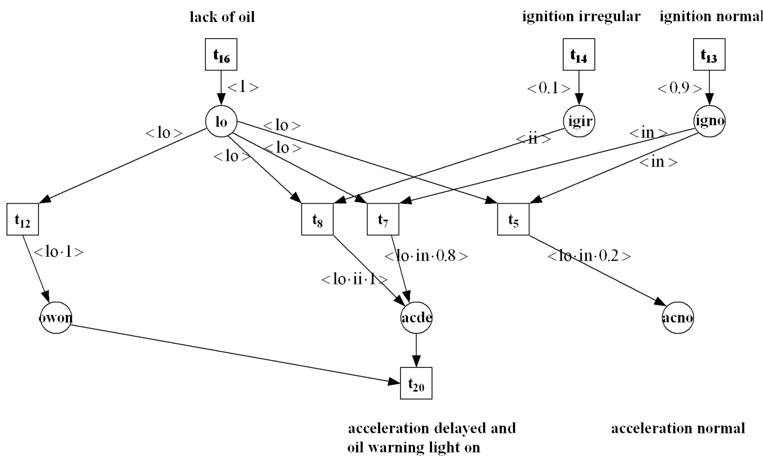


Figure 4. $\mathcal{PN}'_\alpha$ of example 3

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

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

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

□

4 Petri Net Representation of Bayesian Networks

This paper deals with PPNs that we introduced in section 3. We demonstrated their use with help of a PHA-example ([PT97]). In principle, we can apply PPNs also to BNs. For that we simply have to translate the steps of the propagation algorithms (see [Nea90]) into transitions of PPNs. However, in doing so, we produce huge nets which are hard to follow and for which we have not enough space. So, we work with foldings of PPNs (”higher” PPNs) which we demonstrate by means of an example (see [Pea88]). (Of course, one can easily find translation rules that transform BNs into higher PPNs. Since the applicability of PPNs is the real aim of this paper, we postpone these translations to a forthcoming paper.)

The reason for using PPNs to represent the flow of probabilities and evidences is that BNs have no clearly defined concept of situation. Consequently, the flows of probabilities and evidences, their influences on each other, and other effects are not clearly visible. They are "hidden" in the algorithms. Petri nets, on the other hand, are highly qualified and especially developed to model flows. Of course, there are powerful tools for working with BNs. Nevertheless, Petri nets have specific advantages. First of all, all flows are realizations of t-invariants, which can be easily calculated in the underlying p/t-nets. By that, the flows are transparent and can be followed. Second, in connection with loopy BNs there exist two structured approaches, namely clustering and conditioning, which both are due to Pearl [Pea88]. We will not deal with clustering, because after clustering the BNs are loop-free and lead to the "normal" Petri net representation. Conditioning, however, needs to cut the loops, thus modifying the BNs. The complexity, that grows exponentially with the number of nodes required for cutting the loops, will be found in the size of the probability/evidence-objects flowing through the Petri net.

In order to facilitate understanding, we start off with a famous example [Pea88].

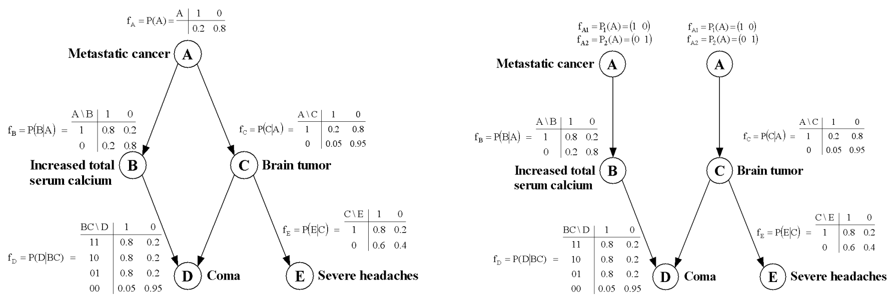
Example 4. The BN of figure 5(a) indicates that "metastatic cancer" (A) has two consequences: "increased total serum calcium" (B) and "brain tumor" (C). The probabilities of A and $\neg A$ are 0.2 and 0.8, respectively. The relevant probabilities of $B, \neg B, C, \neg C$ are the conditional probabilities given A. Similarly, B and C have the consequence "coma" (D). Moreover, C has the consequence "severe headaches" (E). Again, the relevant probabilities are the conditional probabilities given the predecessors (parents).

□

In a BN, the nodes can be interpreted as stochastic or random variables. It is assumed that probabilities flow along the arcs (in arc direction), thus determining the probabilities of the successor nodes by multiplying the arriving probabilities by the assigned conditional probabilities. Initially, the source of flows are the given probabilities of A and $\neg A$. In addition, by outer events or evidences, certain probabilities can change, which enforces re-calculating the probabilities of the successor nodes — and also of the predecessor nodes (against the arc direction). Both is to keep the network consistent with the new evidences.

Roughly spoken, the values flowing in arc direction are probabilities π; the values flowing against arc direction are likelihoods λ.

By definition, BNs have no directed cycles. But they can have loops with (at least) two undesirable effects. The BN in figure 5(a) has the loop $ABDCA$. First of all, there is a double influence of A on D. From a logical point of view this does not matter because of $A \wedge A = A$. However, from a probabilistic point of view this double influence leads to wrong values. Since the probability $P(A)$ of A is a factor in $P(B)$ and in $P(C)$, $(P(A))^2$ instead of $P(A)$ becomes a factor in $P(D)$. With the exception of $P(A) = 1$ and $P(A) = 0$, $(P(A))^2 \neq P(A)$ causes the wrong values. On the other hand, this is the reason for the specific part



(a) $\mathcal{B}$ (b) $\mathcal{B}_{cut}$

Figure 5. Bayesian network and conditioned version of example 5

$P(A) = 1$ and $P(A) = 0$ play in the "conditioning approach" [Pea88] to cope with loops.

Example 5 (see example 4). The BN $\mathcal{B}_{cut}$ of figure 5(b) is a modification of the BN $\mathcal{B}$ of figure 5(a) where the loop has been cut off at A. Moreover, the probabilities of A and $\neg A$ $(0.2, 0.8)$ have been replaced by two pairs $f_{A_1} = (1, 0)$ and $f_{A_2} = (0, 1)$. Strictly speaking, figure 5(b) shows two BNs, where the above mentioned values 0 and 1 are the initial probabilities for A and $\neg A$. After evaluating both BNs, the actual probabilities of all nodes are determined by calculating the weighted sum of the respective probabilities in both BNs, where the weights are the actual probabilities of A and $\neg A$; initially $P(A) = 0.2, P(\neg A) = 0.8$. These weights may change by outer events or evidences:

$$P(A|E = 1), P(\neg A|E = 1)$$
$$P(A|E = 1, D = 0), P(\neg A|E = 1, D = 0)$$

Figure 6 shows the Petri net representation $\mathcal{PB}_{cut}$ of the BN in figure 5(b). In order not to overload the net, all functions are collected in table 5. Figures 7 and 8 contain the net representations of the t-invariants of $\mathcal{PB}_{cut}$, which are calculated as t-invariants of the p/t-net belonging to $\mathcal{PB}_{cut}$.

Roughly spoken, in Petri net $\mathcal{PB}_{cut}$ (figure 6) to every directed arc of $\mathcal{B}_{cut}$ (figure 5(b)) two paths exist. The paths in arc direction (w.r.t. $\mathcal{B}_{cut}$) have arc labels $\pi_{\ldots}$, the paths in opposite direction have arc labels $\lambda_{\ldots}$. (Except some Petri net technicalities, we stick to the nomenclature and the algorithms of [Nea90].) The transitions are "functional" $f_{\ldots}$ or "multiplicative" $m_{\ldots}$. Joins and splits are constructed of functional and multiplicative transitions, respectively.

In order to show how to work with the Petri net we start reproducing the empty marking in t-invariant 1 (figure 7). The (constant) matrix $\binom{1,0}{0,1}$ is the arc label of both arcs (f_a^1, a). So, firing both transitions f_A^1 puts $\pi(A) = \binom{1,0}{0,1}$ on both

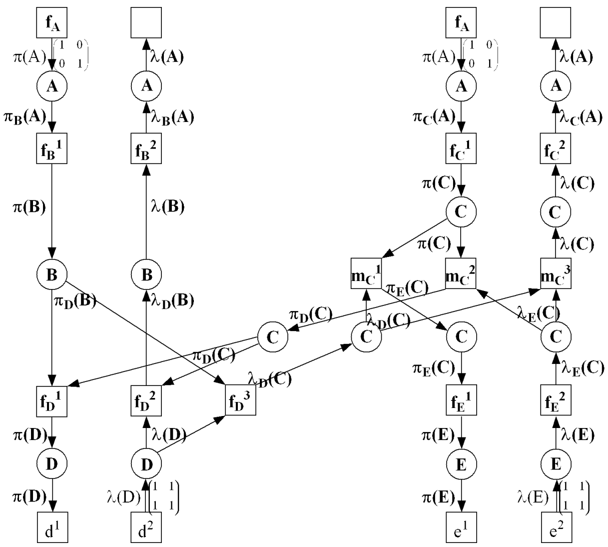


Figure 6. $\mathcal{PB}_{cut}$

places A. This is the beginning of a simultaneous evaluation of both BNs in figure 5(b). Similarly, firing e^2 puts $\lambda(E) = \binom{1,1}{1,1}$ on place E. This indicates in both evaluations that nothing has happened or is known that forces to depart from the relation $P(E) = P(\neg E)$. Next, firing of f_B^1, f_C^1, f_E^2 takes $\pi_B(A) = \binom{1,0}{0,1}, \pi_C(A) = \binom{1,0}{0,1}, \lambda(E) = \binom{1,1}{1,1}$ from A, A, E and puts

$$\pi_B(A) \cdot f_B^1 = \begin{pmatrix} 1,0 \\ 0,1 \end{pmatrix} \cdot \begin{pmatrix} 0.8, 0.2 \\ 0.2, 0.8 \end{pmatrix} = \begin{pmatrix} 0.8, 0.2 \\ 0.2, 0.8 \end{pmatrix} = \pi(B)$$

$$\pi_C(A) \cdot f_C^1 = \begin{pmatrix} 1,0 \\ 0,1 \end{pmatrix} \cdot \begin{pmatrix} 0.2, 0.8 \\ 0.05, 0.95 \end{pmatrix} = \begin{pmatrix} 0.2, 0.8 \\ 0.05, 0.95 \end{pmatrix} = \pi(C)$$

$$\lambda(E) \cdot f_E^2 = \begin{pmatrix} 1,1 \\ 1,1 \end{pmatrix} \cdot \begin{pmatrix} 0.8, 0.6 \\ 0.2, 0.4 \end{pmatrix} = \begin{pmatrix} 1,1 \\ 1,1 \end{pmatrix} = \lambda_E(C)$$

on B, C, C, respectively.

Now, transition m_C^2 is enabled, takes $\pi(C)$ and $\lambda_E(C)$ from places C, and puts

$$\pi_D(C) = \pi(C) \circ \lambda_E(C) = \begin{pmatrix} p_{11}, p_{12} \\ p_{21}, p_{22} \end{pmatrix} \circ \begin{pmatrix} l_{11}, l_{12} \\ l_{21}, l_{22} \end{pmatrix}$$
$$:= \begin{pmatrix} p_{11} \cdot l_{11}, p_{12} \cdot l_{12} \\ p_{21} \cdot l_{21}, p_{22} \cdot l_{22} \end{pmatrix} = \begin{pmatrix} 0.2, 0.8 \\ 0.05, 0.95 \end{pmatrix}$$

on C, thus enabling transition f_D^1.

Firing f_D^1 consists of clearing the input places B and C and putting $\pi(D)$ on D which is calculated as follows:

$$f_B^1 = P(BA) = \begin{array}{c|cc} A\backslash B & 1 & 0 \\ \hline 1 & 0.8 & 0.2 \\ 0 & 0.2 & 0.8 \end{array} \qquad f_B^2 = P'(BA) = \begin{array}{c|cc} B\backslash A & 1 & 0 \\ \hline 1 & 0.8 & 0.2 \\ 0 & 0.2 & 0.8 \end{array}$$

$$f_C^1 = P(C|A) = \begin{array}{c|cc} A\backslash C & 1 & 0 \\ \hline 1 & 0.2 & 0.8 \\ 0 & 0.05 & 0.95 \end{array} \qquad f_C^2 = P'(CA) = \begin{array}{c|cc} C\backslash A & 1 & 0 \\ \hline 1 & 0.2 & 0.05 \\ 0 & 0.8 & 0.95 \end{array}$$

$$f_D^1 = P(DBC) = \begin{array}{c|cc} BC\backslash D & 1 & 0 \\ \hline 11 & 0.8 & 0.2 \\ 10 & 0.8 & 0.2 \\ 01 & 0.8 & 0.2 \\ 00 & 0.05 & 0.95 \end{array} \qquad f_D^2 = P(B|DC) = \begin{array}{c|cc} CD\backslash B & 1 & 0 \\ \hline 11 & 0.8 & 0.8 \\ 10 & 0.2 & 0.2 \\ 01 & 0.8 & 0.05 \\ 00 & 0.2 & 0.95 \end{array}$$

$$f_D^3 = P(CBD) = \begin{array}{c|cc} BD\backslash C & 1 & 0 \\ \hline 11 & 0.8 & 0.8 \\ 10 & 0.2 & 0.2 \\ 01 & 0.8 & 0.05 \\ 00 & 0.2 & 0.95 \end{array}$$

$$f_E^1 = P(EC) = \begin{array}{c|cc} C\backslash E & 1 & 0 \\ \hline 1 & 0.8 & 0.2 \\ 0 & 0.6 & 0.4 \end{array} \qquad f_E^2 = P'(E|C) = \begin{array}{c|cc} E\backslash C & 1 & 0 \\ \hline 1 & 0.8 & 0.6 \\ 0 & 0.2 & 0.4 \end{array}$$

Table 5. Probabilities in the Petri net representation of the BN

$$\begin{aligned}
\pi(D) &= (\pi_D(B) \otimes \pi_D(C)) \cdot f_D^1 \\
&= \left(\begin{pmatrix} b_{11}, b_{12} \\ b_{21}, b_{22} \end{pmatrix} \otimes \begin{pmatrix} c_{11}, c_{12} \\ c_{21}, c_{22} \end{pmatrix} \right) \cdot f_D^1 \\
:&= \begin{pmatrix} (b_{11}\ b_{12}) \times (c_{11}\ c_{12}) \\ (b_{21}\ b_{22}) \times (c_{21}\ c_{22}) \end{pmatrix} \cdot f_D^1 \\
&= \begin{pmatrix} b_{11}c_{11}\ b_{11}c_{12}\ b_{12}c_{11}\ b_{12}c_{12} \\ b_{21}c_{21}\ b_{21}c_{22}\ b_{22}c_{21}\ b_{22}c_{22} \end{pmatrix} \cdot f_D^1 \\
&= \begin{pmatrix} 0.8 \cdot 0.2 & 0.8 \cdot 0.8 & 0.2 \cdot 0.2 & 0.2 \cdot 0.8 \\ 0.2 \cdot 0.05 & 0.2 \cdot 0.95 & 0.8 \cdot 0.05 & 0.8 \cdot 0.95 \end{pmatrix} \cdot f_D^1 \\
&= \begin{pmatrix} 0.16 & 0.64 & 0.04 & 0.16 \\ 0.01 & 0.19 & 0.04 & 0.76 \end{pmatrix} \cdot \begin{pmatrix} 0.8 & 0.2 \\ 0.8 & 0.2 \\ 0.8 & 0.2 \\ 0.05 & 0.95 \end{pmatrix} \\
&= \begin{pmatrix} 0.68 & 0.32 \\ 0.23 & 0.77 \end{pmatrix}
\end{aligned}$$

In completing the **0**-reproduction, transition d^1 empties the net representation of t-invariant 1. The results so far are the probabilities $\pi(B), \pi(C), \pi(D)$. By

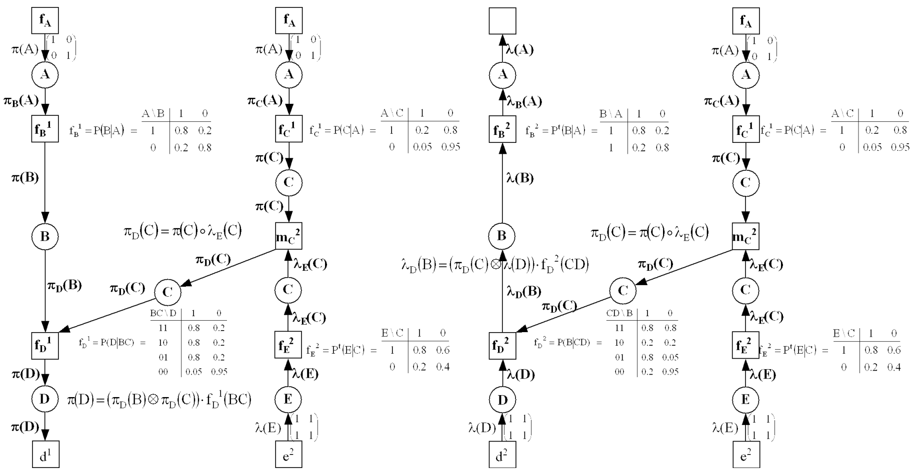


Figure 7. T-invariant 1 and t-invariant 2

0-reproduction also in the remaining three t-invariants, we get altogether:

$$\pi(A) = \begin{pmatrix} 1, 0 \\ 0, 1 \end{pmatrix} \qquad \lambda(A) = \lambda(B) = \lambda(C) = \lambda(D) = \lambda(E) = \begin{pmatrix} 1, 1 \\ 1, 1 \end{pmatrix}$$

$$\pi(B) = \begin{pmatrix} 0.8, 0.2 \\ 0.2, 0.8 \end{pmatrix}$$

$$\pi(C) = \begin{pmatrix} 0.2, \;\; 0.8 \\ 0.05, 0.95 \end{pmatrix}$$

$$\pi(D) = \begin{pmatrix} 0.68, 0.32 \\ 0.23, 0.77 \end{pmatrix}$$

$$\pi(E) = \begin{pmatrix} 0.64, 0.36 \\ 0.61, 0.39 \end{pmatrix}$$

In general, the "belief" of a node χ is calculated by $\alpha(\pi(\chi) \circ \lambda(\chi))$ where α normalizes the product such that the constituents sum up to 1. For the following, we need the beliefs of B and C. We find

$$BEL_1(B) = (0.8, 0.2) \circ (1, 1) = (0.8, 0.2)$$
$$BEL_1(C) = (0.2, 0.8) \circ (1, 1) = (0.2, 0.8)$$

for the first BN of figure 5(b), and

$$BEL_0(B) = (0.2, 0.8) \circ (1, 1) = (0.2, 0.8)$$
$$BEL_0(C) = (0.05, 0.95) \circ (1, 1) = (0.05, 0.95)$$

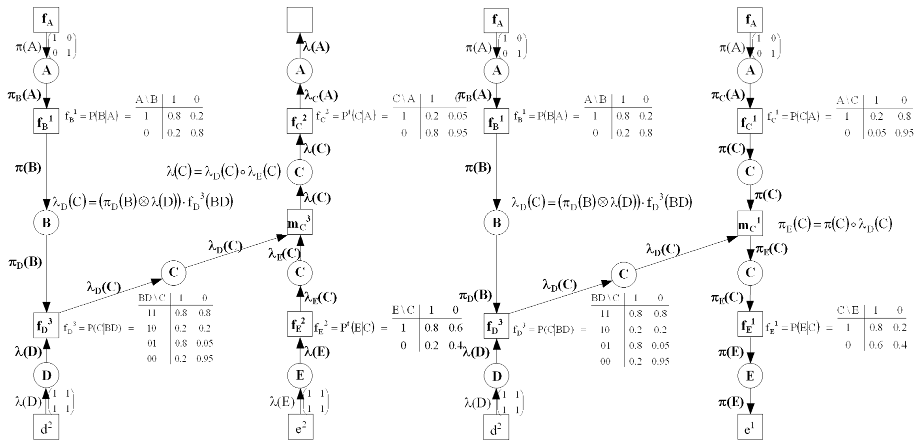


Figure 8. T-invariant 3 and t-invariant 4

for the second one. The actual probability $P(A) = (0.2, 0.8)$ provides the weights for

$$P(B) = BEL_1(B) \cdot 0.2 + BEL_0(B) \cdot 0.8 = (0.32, 0.68)$$
$$P(C) = BEL_1(C) \cdot 0.2 + BEL_0(C) \cdot 0.8 = (0.08, 0.92).$$

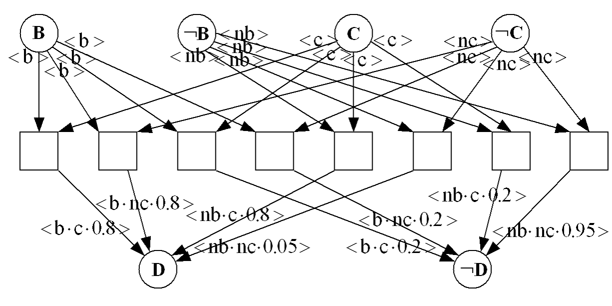


Figure 9. $\mathcal{PN}$

That means:

Given the probability for metastatic cancer $P(A) = (0.2, 0.8)$, the probabilities for increased total serum calcium and brain tumor are $P(B) = (0.32, 0.68)$ and $P(C) = (0.08, 0.92)$, respectively.

□

Before continuing that example, we want to show that the Petri nets of figures 6 to 8 are (higher level) PPNs. For example, the functional transition f_D^1 plus adjacent arcs and places B, C, D is a folding of the PPN of figure 9. In detail:

- $B, \neg B$ in figure 9 are folded to B in figure 7
- $C, \neg C$ in figure 9 are folded to C in figure 7
- $D, \neg D$ in figure 9 are folded to D in figure 7
- all transitions in figure 9 are folded to f_D^1 in figure 7.

Consequently, the arc labels $\pi_D(B), \pi_D(C)$, and $\pi(D)$ denote first of all pairs of complementary probability values. In addition the net represents (a folding of) two BNs (figure 5(b)) which leads to arc labels in form of (2×2)-matrices.

The multiplicative transition m_C^2 is a folding of the two PPNs of figure 10, again modified for (2×2)-matrices as arc lables.

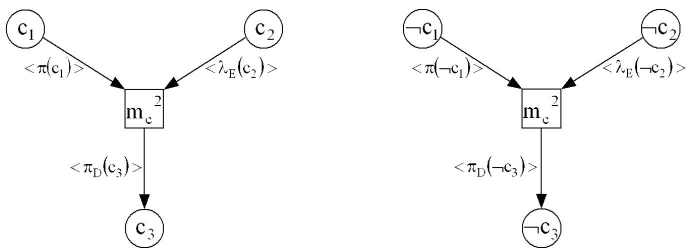


Figure 10. Multiplicative transition m_C^2

Example 6 (see example 5). After initializing the Petri net representation $\mathcal{PB}_{cut}$ (figure 6), we now observe that severe headaches set in. That requires to change the label $\lambda(E) = \binom{1,1}{1,1}$ of arc (e^2, E) into $\binom{1,0}{1,0}$. That means to change a "non-evidence" (equal weights 1 for $P(E)$ and $P(\neg E)$) into a "hard evidence" (weight 1 for $P(E)$, weight 0 for $P(\neg E)$). Now, the simulation leads to the following probabilities, evidences and beliefs:

$$\pi(B) = \begin{pmatrix} 0.8, 0.2 \\ 0.2, 0.8 \end{pmatrix} \qquad \lambda(B) = \begin{pmatrix} 1, 1 \\ 1, 1 \end{pmatrix}$$

$$\pi(C) = \begin{pmatrix} 0.2, \;\; 0.8 \\ 0.05, 0.95 \end{pmatrix} \qquad \lambda(C) = \begin{pmatrix} 0.8, 0.6 \\ 0.8, 0.6 \end{pmatrix}$$

$$\pi(E) = \begin{pmatrix} 0.64, 0.36 \\ 0.61, 0.39 \end{pmatrix}$$

$$BEL_1(B) = (0.8, 0.2) \circ (1, 1) = (0.8, 0.2)$$
$$BEL_0(B) = (0.2, 0.8) \circ (1, 1) = (0.2, 0.8)$$
$$BEL_1(C) = (0.2, 0.8) \circ (0.8, 0.6) = \alpha(0.16, 0.48) = (0.25, 0.75)$$
$$BEL_0(C) = (0.05, 0.95) \circ (0.8, 0.6) = \alpha(0.04, 0.57) = (0.0656, 0.9344).$$

To get the weights, we need the actual probability of A, namely $P(A|E=1)$. What we know is:

$$\pi(E) = \begin{pmatrix} P(E=1|A=1),\, P(E=0|A=1) \\ P(E=1|A=0),\, P(E=0|A=0) \end{pmatrix} = \begin{pmatrix} 0.64,\, 0.36 \\ 0.61,\, 0.39 \end{pmatrix},$$

$$P(A) = (0.2, 0.8), \text{ and}$$

$$\begin{aligned} P(A|E=1) &= P(E=1|A)\cdot P(A) = \alpha\cdot(0.64\cdot 0.2, 0.61\cdot 0.8) \\ &= \alpha\cdot(0.128, 0.488) \\ &= (0.208, 0.792) \text{ (the actual probability of } A). \end{aligned}$$

So,

$$\begin{aligned} P(B) &= BEL_1(B)\cdot 0.208 + BEL_0(B)\cdot 0.792 \\ &= (0.8, 0.2)\cdot 0.208 + (0.2, 0.8)\cdot 0.792 \\ &= (0.1664, 0.0416) + (0.1584, 0.6336) \\ &= (0.3248, 0.6752) \approx (0.32, 0.68) \end{aligned}$$

and

$$\begin{aligned} P(C) &= BEL_1(C)\cdot 0.208 + BEL_0(C)\cdot 0.792 \\ &= (0.25, 0.75)\cdot 0.208 + (0.0656, 0.9344)\cdot 0.792 \\ &= (0.052, 0.156) + (0.052, 0.740) = (0.104, 0.896) \end{aligned}$$

This shows that $\lambda(E) = \left(\begin{smallmatrix} 1,0 \\ 1,0 \end{smallmatrix}\right)$ (severe headaches set in) does not influence $P(B)$. On the other hand, the probability $P(C)$ of a brain tumor is increased from $(0.08, 0.92)$ to $(0.104, 0.896)$.

The observation, that no coma sets in (in addition to the severe headaches), leads to changing the label $\lambda(D) = \left(\begin{smallmatrix} 1,1 \\ 1,1 \end{smallmatrix}\right)$ of the arc (d^2, D) into $\left(\begin{smallmatrix} 0,1 \\ 0,1 \end{smallmatrix}\right)$. The simulation yields now

$$\pi(B) = \begin{pmatrix} 0.8,\, 0.2 \\ 0.2,\, 0.8 \end{pmatrix}, \qquad \lambda(B) = \begin{pmatrix} 0.2,\, 0.7625 \\ 0.2,\, 0.9005 \end{pmatrix}$$

$$\pi(C) = \begin{pmatrix} 0.2,\quad 0.8 \\ 0.05,\, 0.95 \end{pmatrix}, \qquad \lambda(C) = \begin{pmatrix} 0.16,\, 0.21 \\ 0.16,\, 0.48 \end{pmatrix}$$

$$\pi(D) = \begin{pmatrix} 0.6875,\, 0.3125 \\ 0.2396,\, 0.7604 \end{pmatrix}$$

$$
\begin{aligned}
BEL_1(B) &= (0.8, 0.2) \circ (0.2, 0.7625) = \alpha(0.16, 0.1525) = (0.512, 0.488)\\
BEL_0(B) &= (0.2, 0.8) \circ (0.2, 0.9005) = \alpha(0.04, 0.7204) = (0.0526, 0.9474)\\
BEL_1(C) &= (0.2, 0.8) \circ (0.16, 0.21) = \alpha(0.032, 0.168) = (0.16, 0.84)\\
BEL_0(C) &= (0.05, 0.95) \circ (0.16, 0.48) = \alpha(0.008, 0.456) = (0.0172, 0.9828)\\
\pi(D) &= \begin{pmatrix} P(D=1|A=1, E=1),\ P(D=0|A=1, E=1) \\ P(D=1|A=0, E=1),\ P(D=0|A=0, E=1) \end{pmatrix}\\
&= \begin{pmatrix} 0.6875,\ 0.3125 \\ 0.2396,\ 0.7604 \end{pmatrix}\\
P(A|E=1) &= (0.208, 0.792) \qquad \text{(see example 5).}
\end{aligned}
$$

The actual probability of A is

$$
\begin{aligned}
P(A|E=1, D=0) &= \alpha P(D=0|A, E=1) \circ P(A|E=1)\\
&= \alpha(0.3125 \cdot 0.208, 0.7604 \cdot 0.792)\\
&= \alpha(0.065, 0.602) = (0.0975, 0.9025).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
P(B) &= BEL_1(B) \cdot 0.0975 + BEL_0(B) \cdot 0.9025\\
&= (0.512, 0.488) \cdot 0.0975 + (0.0526, 0.9474) \cdot 0.9025\\
&= (0.0499, 0.0476) + (0.0475, 0.8550)\\
&= (0.0974, 0.9026)\\
P(C) &= BEL_1(C) \cdot 0.0975 + BEL_0(C) \cdot 0.9025\\
&= (0.16, 0.84) \cdot 0.0975 + (0.0172, 0.9828) \cdot 0.9025\\
&= (0.0156, 0.0819) + (0.0155, 0.0887)\\
&= (0.0311, 0.9689).
\end{aligned}
$$

This indicates that the absence of a coma decreases $P(B)$ and $P(C)$.

□

5 Conclusion

We introduced probability propagation nets (PPNs) as an extension of canonical net representations of Horn formulae [Lau03]. First, we applied these nets to a specific sort of diagnosis, namely to probabilistic Horn abduction (PHA). Second, we used PPNs to model probability and evidence propagations in Bayesian networks (BNs). In fact, we used foldings of PPNs to keep the net structures manageable. Since these "higher level" PPNs yield quite a clear representation of algorithmic and of propagation aspects, we are planning to show their applicability, in particular to BNs, in a forthcoming paper.

References

[Jen96] Finn V. Jensen. *Introduction to Bayesian Networks*. UCL Press, London, 1996.

[Lau02] Kurt Lautenbach. Reproducibility of the empty marking. In Javier Esparza and Charles Lakos, editor, *Applications and Theory of Petri Nets 2002, 23rd International Conference, ICATPN 2002, Adelaide, Australia, June 24–30, 2002, Proceedings*, volume 2360 of *Lecture Notes in Computer Science*, pages 237–253. Springer, 2002.

[Lau03] Kurt Lautenbach. Logical reasoning and petri nets. In Wil M. P. van der Aalst and Eike Best, editors, *Applications and Theory of Petri Nets 2003, 24th International Conference, ICATPN 2003, Eindhoven, Netherlands, June 23–27, 2003, Proceedings*, volume 2679 of *Lecture Notes in Computer Science*, pages 276–295. Springer, 2003.

[Nea90] Richard E. Neapolitan. *Probabilistic reasoning in expert systems: theory and algorithms*. John Wiley & Sons, Inc., New York, NY, USA, 1990.

[Nea03] Richard E. Neapolitan. *Learning Bayesian Networks*. Prentice Hall, Upper Saddle River, NJ, 2003.

[Pea88] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.

[Poo93a] David Poole. Logic programming, abduction and probability: a top-down anytime algorithm for estimating prior and posterior probabilities. *New Generation Computing*, 11(3-4):377–400, 1993.

[Poo93b] David Poole. Probabilistic horn abduction and bayesian networks. *Artificial Intelligence*, 64(1):81–129, 1993.

[PT97] Luigi Portinale and Pietro Torasso. A comparative analysis of horn models and bayesian networks for diagnosis. In Maurizio Lenzerini, editor, *Advances in Artificial Intelligence 1997, Rome, Italy*, volume 1321 of *Lecture Notes in Computer Science*, pages 254–265. Springer, 1997.

[VAB04] Somsak Vanit-Anunchai and Jonathan 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*, pages 117–136, 2004.

[VAB05] Somsak Vanit-Anunchai and Jonathan 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*, pages 98–106, 2005.

[WL99] Lee W. Wagenhals and Alexander 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.

Available Research Reports (since 2000):

2005

16/2005 *Kurt Lautenbach, Alexander Pinl.* Probability Propagation in Petri Nets.

15/2005 *Rainer Gimnich, Uwe Kaiser, Andreas Winter.* 2. Workshop "Reengineering Prozesse" — Software Migration —.

14/2005 *Jan Murray, Frieder Stolzenburg, Toshiaki Arai.* Hybrid State Machines with Timed Synchronization for Multi-Robot System Specification.

13/2005 *Reinhold Letz.* FTP 2005 — Fifth International Workshop on First-Order Theorem Proving.

12/2005 *Bernhard Beckert.* TABLEAUX 2005 — Position Papers and Tutorial Descriptions.

11/2005 *Dietrich Paulus, Detlev Droege.* Mixed-reality as a challenge to image understanding and artificial intelligence.

10/2005 *Jürgen Sauer.* 19. Workshop Planen, Scheduling und Konfigurieren / Entwerfen.

9/2005 *Pascal Hitzler, Carsten Lutz, Gerd Stumme.* Foundational Aspects of Ontologies.

8/2005 *Joachim Baumeister, Dietmar Seipel.* Knowledge Engineering and Software Engineering.

7/2005 *Benno Stein, Sven Meier zu Eißen.* Proceedings of the Second International Workshop on Text-Based Information Retrieval.

6/2005 *Andreas Winter, Jürgen Ebert.* Metamodel-driven Service Interoperability.

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

4/2005 *Torsten Gipp, Jürgen Ebert.* Web Engineering does profit from a Functional Approach.

3/2005 *Oliver Obst, Anita Maas, Joschka Boedecker.* HTN Planning for Flexible Coordination Of Multiagent Team Behavior.

2/2005 *Andreas von Hessling, Thomas Kleemann, Alex Sinner.* Semantic User Profiles and their Applications in a Mobile Environment.

1/2005 *Heni Ben Amor, Achim Rettinger.* Intelligent Exploration for Genetic Algorithms – Using Self-Organizing Maps in Evolutionary Computation.

2004

12/2004 *Manfred Rosendahl.* Objektorientierte Implementierung einer Constraint basierten geometrischen Modellierung.

11/2004 *Urs Kuhlmann, Harry Sneed, Andreas Winter.* Workshop Reengineering Prozesse (RePro 2004) — Fallstudien, Methoden, Vorgehen, Werkzeuge.

10/2004 *Bernhard Beckert, Gerd Beuster.* Formal Specification of Security-relevant Properties of User-Interfaces.

9/2004 *Bernhard Beckert, Martin Giese, Elmar Habermalz, Reiner Hähnle, Andreas Roth, Philipp Rümmer, Steffen Schlager.* Taclets: A New Paradigm for Constructing Interactive Theorem Provers.

8/2004 *Achim Rettinger.* Learning from Recorded Games: A Scoring Policy for Simulated Soccer Agents.

7/2004 *Oliver Obst, Markus Rollmann.* Spark — A Generic Simulator for Physical Multi-agent Simulations.

6/2004 *Frank Dylla, Alexander Ferrein, Gerhard Lakemeyer, Jan Murray, Oliver Obst, Thomas Röfer, Frieder Stolzenburg, Ubbo Visser, Thomas Wagner.* Towards a League-Independent Qualitative Soccer Theory for RoboCup.

5/2004 *Peter Baumgartner, Ulrich Furbach, Margret Groß-Hardt, Thomas Kleemann.* Model Based Deduction for Database Schema Reasoning.

4/2004 *Lutz Priese.* A Note on Recognizable Sets of Unranked and Unordered Trees.

3/2004 *Lutz Priese.* Petri Net DAG Languages and Regular Tree Languages with Synchronization.

2/2004 *Ulrich Furbach, Margret Groß-Hardt, Bernd Thomas, Tobias Weller, Alexander Wolf.* Issues Management: Erkennen und Beherrschen von kommunikativen Risiken und Chancen.

1/2004 *Andreas Winter, Carlo Simon.* Exchanging Business Process Models with GXL.

2003

18/2003 *Kurt Lautenbach.* Duality of Marked Place/Transition Nets.

17/2003 *Frieder Stolzenburg, Jan Murray, Karsten Sturm.* Multiagent Matching Algorithms With and Without Coach.

16/2003 *Peter Baumgartner, Paul A. Cairns, Michael Kohlhase, Erica Melis (Eds.).* Knowledge Representation and Automated Reasoning for E-Learning Systems.

15/2003 *Peter Baumgartner, Ulrich Furbach, Margret Gross-Hardt, Thomas Kleemann, Christoph Wernhard.* KRHyper Inside — Model Based Deduction in Applications.

14/2003 *Christoph Wernhard.* System Description: KRHyper.

13/2003 *Peter Baumgartner, Ulrich Furbach, Margret Gross-Hardt, Alex Sinner.* 'Living Book' :- 'Deduction', 'Slicing', 'Interaction'..

12/2003 *Heni Ben Amor, Oliver Obst, Jan Murray.* Fast, Neat and Under Control: Inverse Steering Behaviors for Physical Autonomous Agents.

11/2003 *Gerd Beuster, Thomas Kleemann, Bernd Thomas.* MIA - A Multi-Agent Location Based Information Systems for Mobile Users in 3G Networks.

10/2003 *Gerd Beuster, Ulrich Furbach, Margret Groß-Hardt, Bernd Thomas.* Automatic Classification for the Identification of Relationships in a Metadata Repository.

9/2003 *Nicholas Kushmerick, Bernd Thomas.* Adaptive information extraction: Core technologies for information agents.

8/2003 *Bernd Thomas.* Bottom-Up Learning of Logic Programs for Information Extraction from Hypertext Documents.

7/2003 *Ulrich Furbach.* AI - A Multiple Book Review.

6/2003 *Peter Baumgartner, Ulrich Furbach, Margret Groß-Hardt.* Living Books.

5/2003 *Oliver Obst.* Using Model-Based Diagnosis to Build Hypotheses about Spatial Environments.

4/2003 *Daniel Lohmann, Jürgen Ebert.* A Generalization of the Hyperspace Approach Using Meta-Models.

3/2003 *Marco Kögler, Oliver Obst.* Simulation League: The Next Generation.

2/2003 *Peter Baumgartner, Margret Groß-Hardt, Alex Sinner.* Living Book – Deduction, Slicing and Interaction.

1/2003 *Peter Baumgartner, Cesare Tinelli.* The Model Evolution Calculus.

2002

12/2002 *Kurt Lautenbach.* Logical Reasoning and Petri Nets.

11/2002 *Margret Groß-Hardt.* Processing of Concept Based Queries for XML Data.

10/2002 *Hanno Binder, Jérôme Diebold, Tobias Feldmann, Andreas Kern, David Polock, Dennis Reif, Stephan Schmidt, Frank Schmitt, Dieter Zöbel.* Fahrassistenzsystem zur Unterstützung beim Rückwärtsfahren mit einachsigen Gespannen.

9/2002 *Jürgen Ebert, Bernt Kullbach, Franz Lehner.* 4. Workshop Software Reengineering (Bad Honnef, 29./30. April 2002).

8/2002 *Richard C. Holt, Andreas Winter, Jingwei Wu.* Towards a Common Query Language for Reverse Engineering.

7/2002 *Jürgen Ebert, Bernt Kullbach, Volker Riediger, Andreas Winter.* GUPRO – Generic Understanding of Programs, An Overview.

6/2002 *Margret Groß-Hardt.* Concept based querying of semistructured data.

5/2002 *Anna Simon, Marianne Valerius.* User Requirements – Lessons Learned from a Computer Science Course.

4/2002 *Frieder Stolzenburg, Oliver Obst, Jan Murray.* Qualitative Velocity and Ball Interception.

3/2002 *Peter Baumgartner.* A First-Order Logic Davis-Putnam-Logemann-Loveland Procedure.

2/2002 *Peter Baumgartner, Ulrich Furbach.* Automated Deduction Techniques for the Management of Personalized Documents.

1/2002 *Jürgen Ebert, Bernt Kullbach, Franz Lehner.* 3. Workshop Software Reengineering (Bad Honnef, 10./11. Mai 2001).

2001

13/2001 *Annette Pook.* Schlussbericht "FUN - Funkunterrichtsnetzwerk".

12/2001 *Toshiaki Arai, Frieder Stolzenburg.* Multiagent Systems Specification by UML Statecharts Aiming at Intelligent Manufacturing.

11/2001 *Kurt Lautenbach.* Reproducibility of the Empty Marking.

10/2001 *Jan Murray.* Specifying Agents with UML in Robotic Soccer.

9/2001 *Andreas Winter.* Exchanging Graphs with GXL.

8/2001 *Marianne Valerius, Anna Simon.* Slicing Book Technology — eine neue Technik für eine neue Lehre?.

7/2001 *Bernt Kullbach, Volker Riediger.* Folding: An Approach to Enable Program Understanding of Preprocessed Languages.

6/2001 *Frieder Stolzenburg.* From the Specification of Multiagent Systems by Statecharts to their Formal Analysis by Model Checking.

5/2001 *Oliver Obst.* Specifying Rational Agents with Statecharts and Utility Functions.

4/2001 *Torsten Gipp, Jürgen Ebert.* Conceptual Modelling and Web Site Generation using Graph Technology.

3/2001 *Carlos I. Chesñevar, Jürgen Dix, Frieder Stolzenburg, Guillermo R. Simari.* Relating Defeasible and Normal Logic Programming through Transformation Properties.

2/2001 *Carola Lange, Harry M. Sneed, Andreas Winter.* Applying GUPRO to GEOS – A Case Study.

1/2001 *Pascal von Hutten, Stephan Philippi.* Modelling a concurrent ray-tracing algorithm using object-oriented Petri-Nets.

2000

8/2000 *Jürgen Ebert, Bernt Kullbach, Franz Lehner (Hrsg.).* 2. Workshop Software Reengineering (Bad Honnef, 11./12. Mai 2000).

7/2000 *Stephan Philippi.* AWPN 2000 - 7. Workshop Algorithmen und Werkzeuge für Petrinetze, Koblenz, 02.-03. Oktober 2000 .

6/2000 *Jan Murray, Oliver Obst, Frieder Stolzenburg.* Towards a Logical Approach for Soccer Agents Engineering.

5/2000 *Peter Baumgartner, Hantao Zhang (Eds.).* FTP 2000 – Third International Workshop on First-Order Theorem Proving, St Andrews, Scotland, July 2000.

4/2000 *Frieder Stolzenburg, Alejandro J. García, Carlos I. Chesñevar, Guillermo R. Simari.* Introducing Generalized Specificity in Logic Programming.

3/2000 *Ingar Uhe, Manfred Rosendahl.* Specification of Symbols and Implementation of Their Constraints in JKogge.

2/2000 *Peter Baumgartner, Fabio Massacci.* The Taming of the (X)OR.

1/2000 *Richard C. Holt, Andreas Winter, Andy Schürr.* GXL: Towards a Standard Exchange Format.