Kurt Lautenbach laut@uni-koblenz.de

Belief Revision and Petri Nets

February 2011

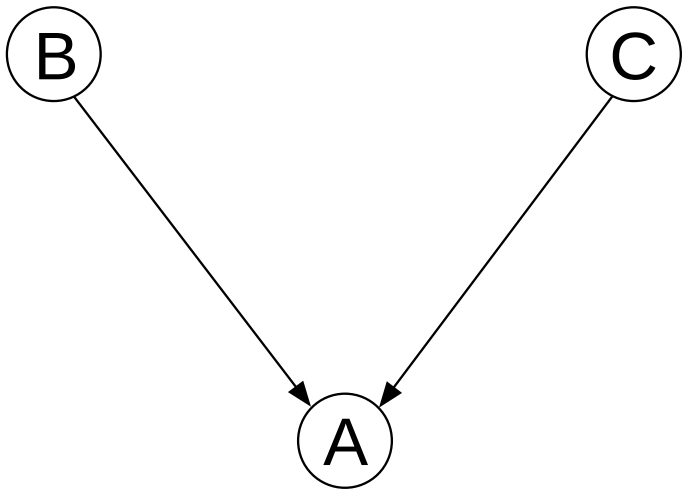
Institut für Softwaretechnik
Universität Koblenz-Landau

Arbeitsgruppe
Petri-Netze

Contents

Contents

Burglar Alarm



A Bayesian network

A_1 = Mr. Holmes' burglar alarm sounds

A_0 = Mr. Holmes' burglar alarm does not sound

B_1 = Mr. Holmes' residence is burglarized

B_0 = Mr. Holmes' residence is not burglarized

C_1 = there is an earthquake

C_0 = there is no earthquake

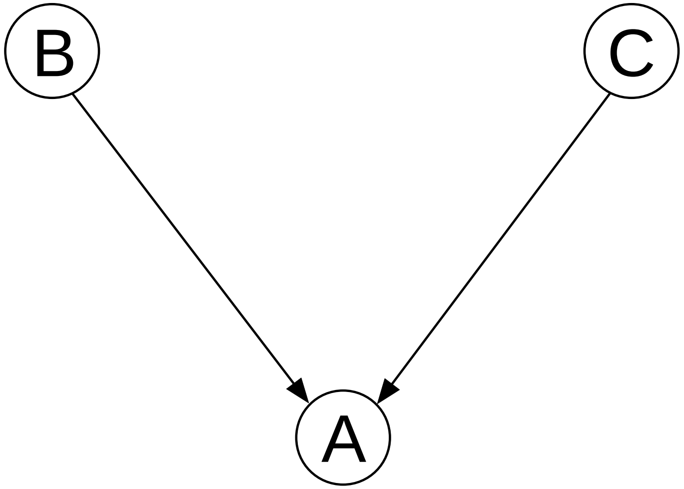
Burglar Alarm

B	1	0
	0.01	0.99

= P(B)

C	1	0
	0.001	0.999

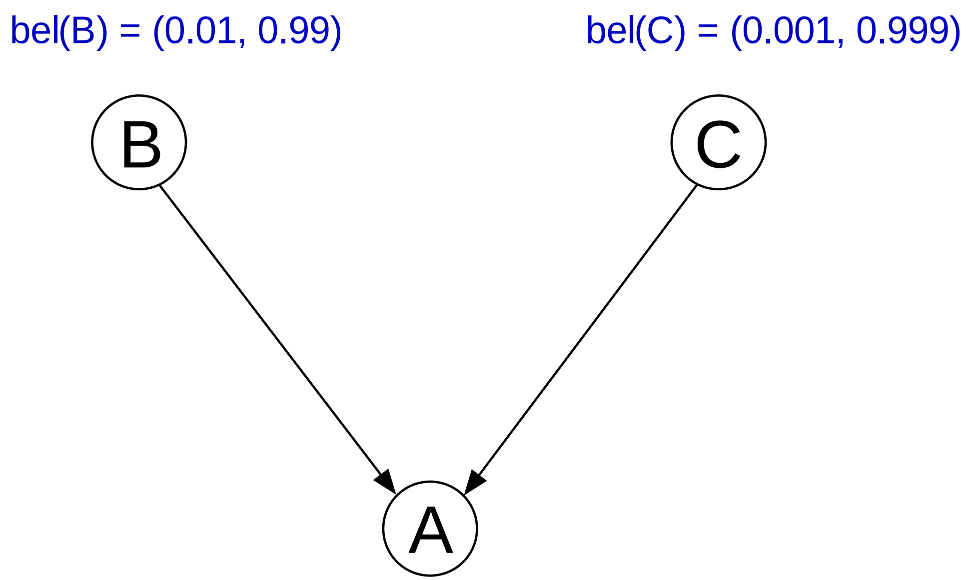
= P(C)



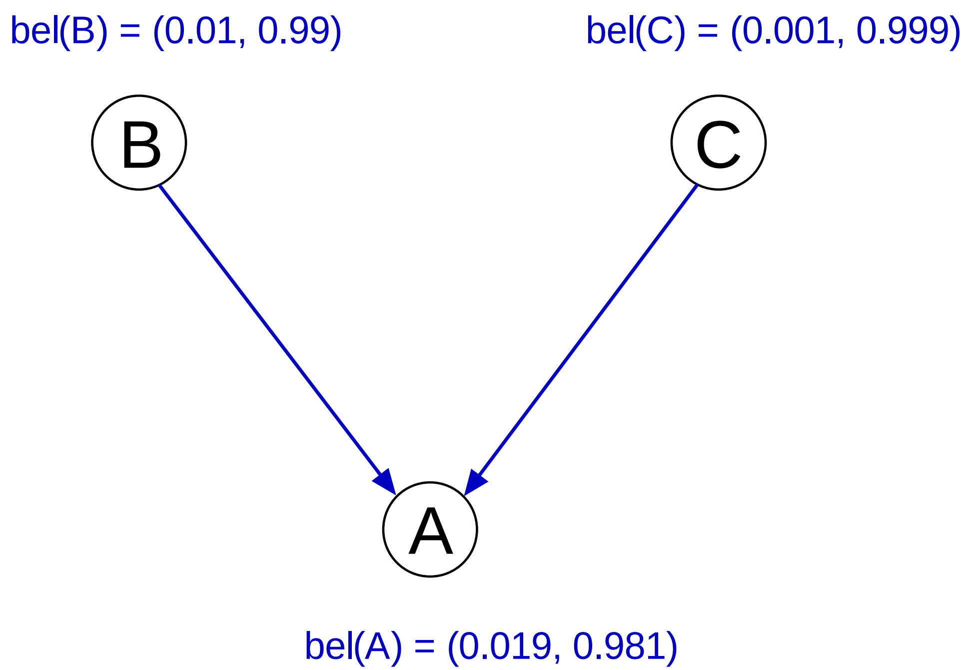
P(A|BC) =

		A	
B	C	1	0
1	1	0.99	0.01
1	0	0.9	0.1
0	1	0.5	0.5
0	0	0.01	0.99

Burglar Alarm



Initialization

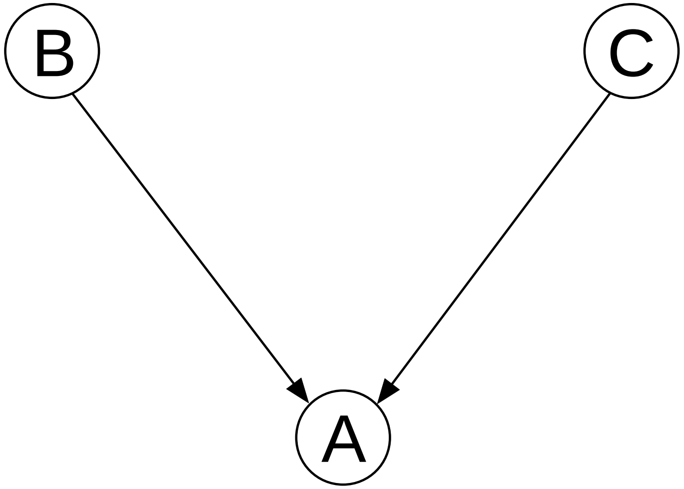


Initialization

Burglar Alarm

bel(B) = (0.01, 0.99)

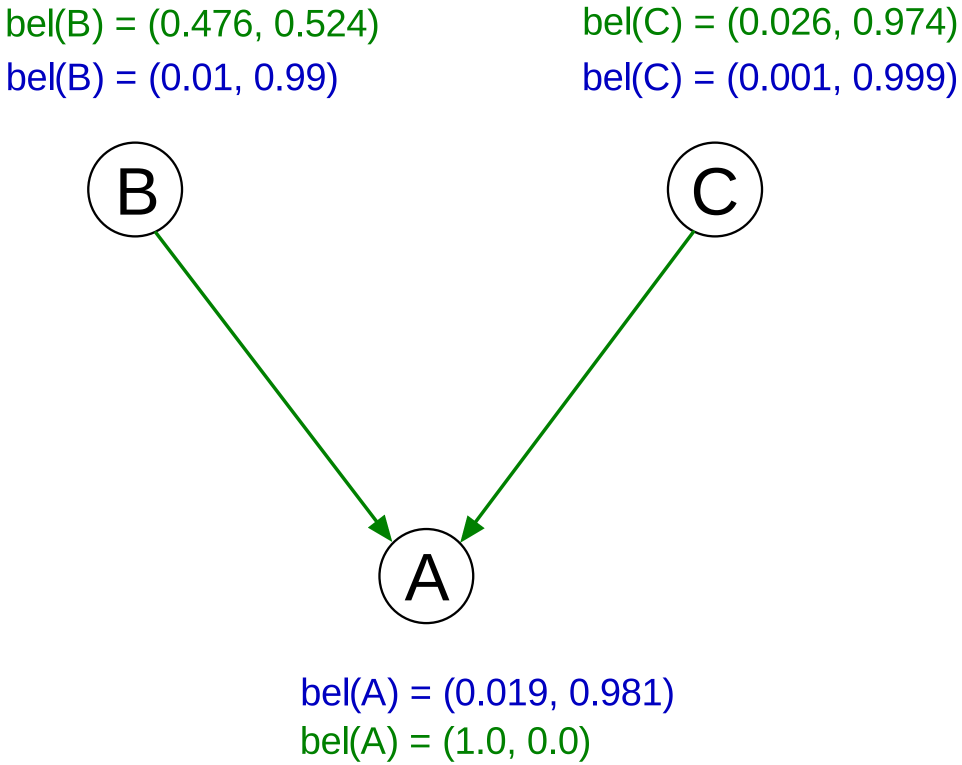
bel(C) = (0.001, 0.999)



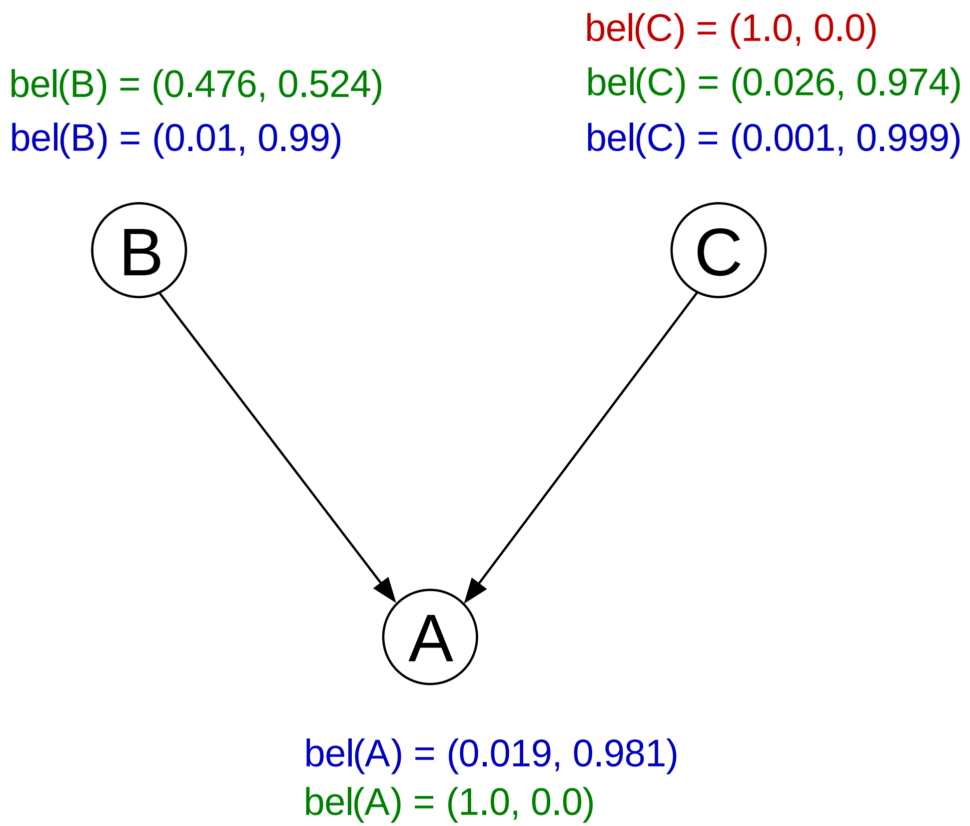
bel(A) = (0.019, 0.981)

bel(A) = (1.0, 0.0)

New Evidence: Mr. Holmes' burglar alarm sounds

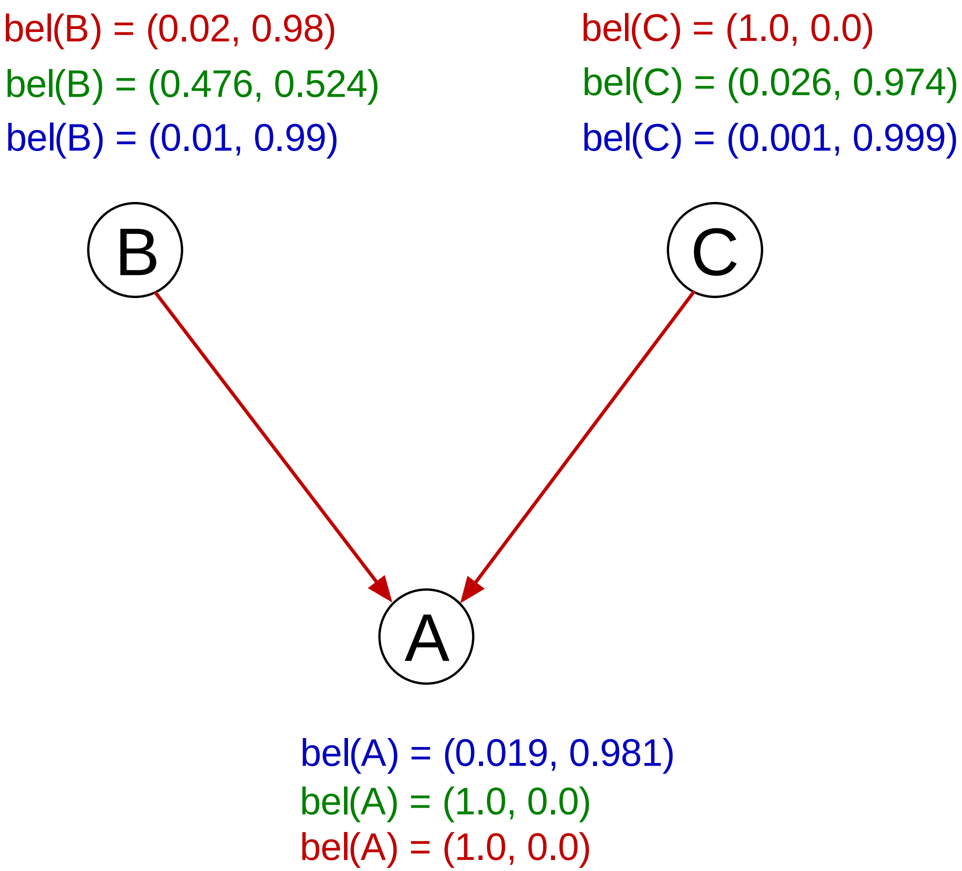


New Evidence: Mr. Holmes' burglar alarm sounds



New Evidence: there was an earth quake

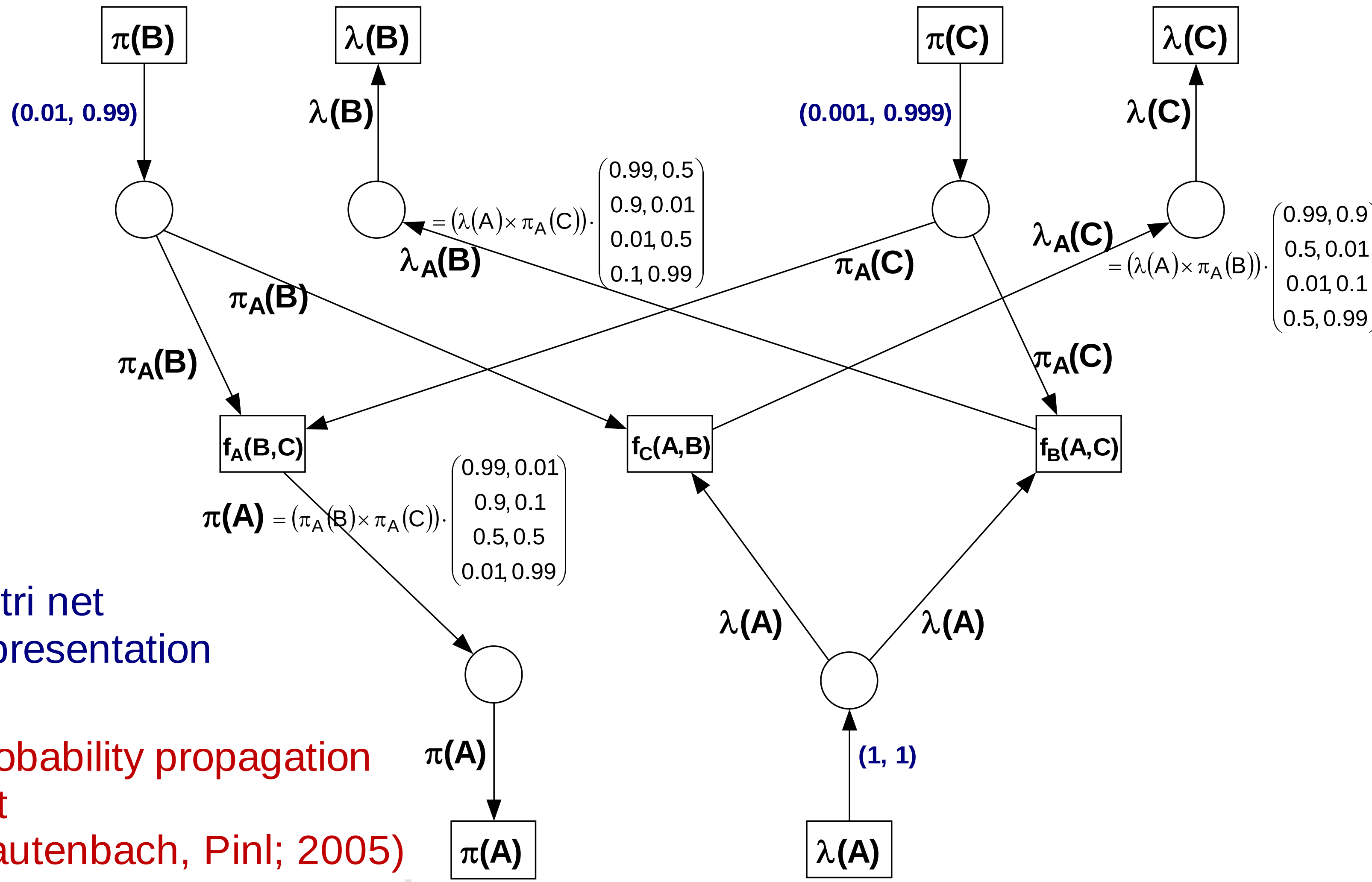
Burglar Alarm



New Evidence: there was an earth quake

Contents

Burglar Alarm



Petri net representation

Probability propagation net (Lautenbach, Pinl; 2005)

Beliefs in Bayesian networks

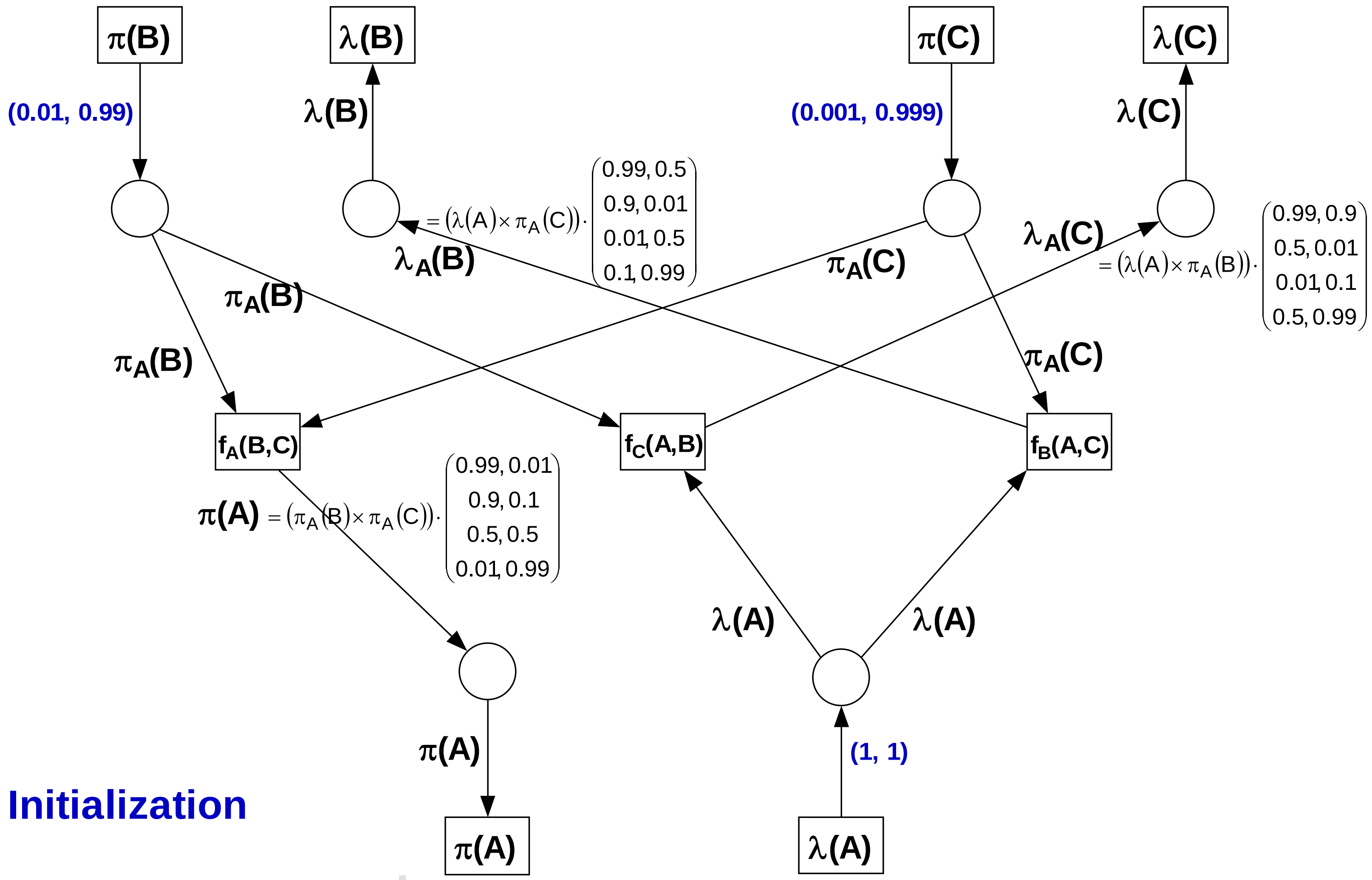
components product

$$
\begin{aligned}
\mathrm{bel}(X) &:= \alpha \cdot (\pi(X) \circ \lambda(X)) \\
&= \alpha \cdot ((\pi_1, \ldots, \pi_n) \circ (\lambda_1, \ldots, \lambda_n)) \\
&= \alpha \cdot (\pi_1\lambda_1, \ldots, \pi_n\lambda_n) = (b_1, \ldots, b_n)
\end{aligned}
$$

where $\sum_{j=1}^{n} b_j = 1$

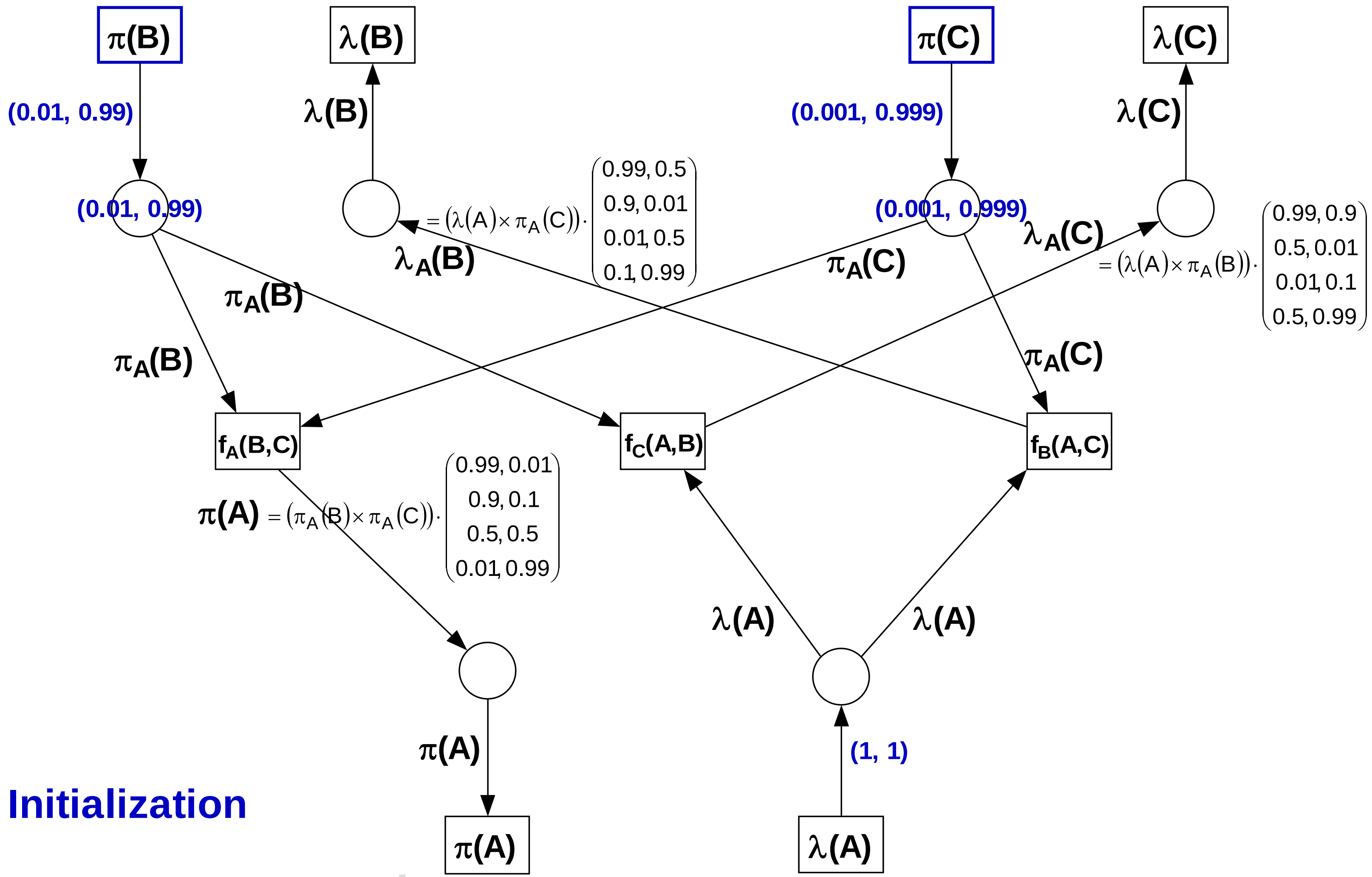
α is a normalization factor

Burglar Alarm



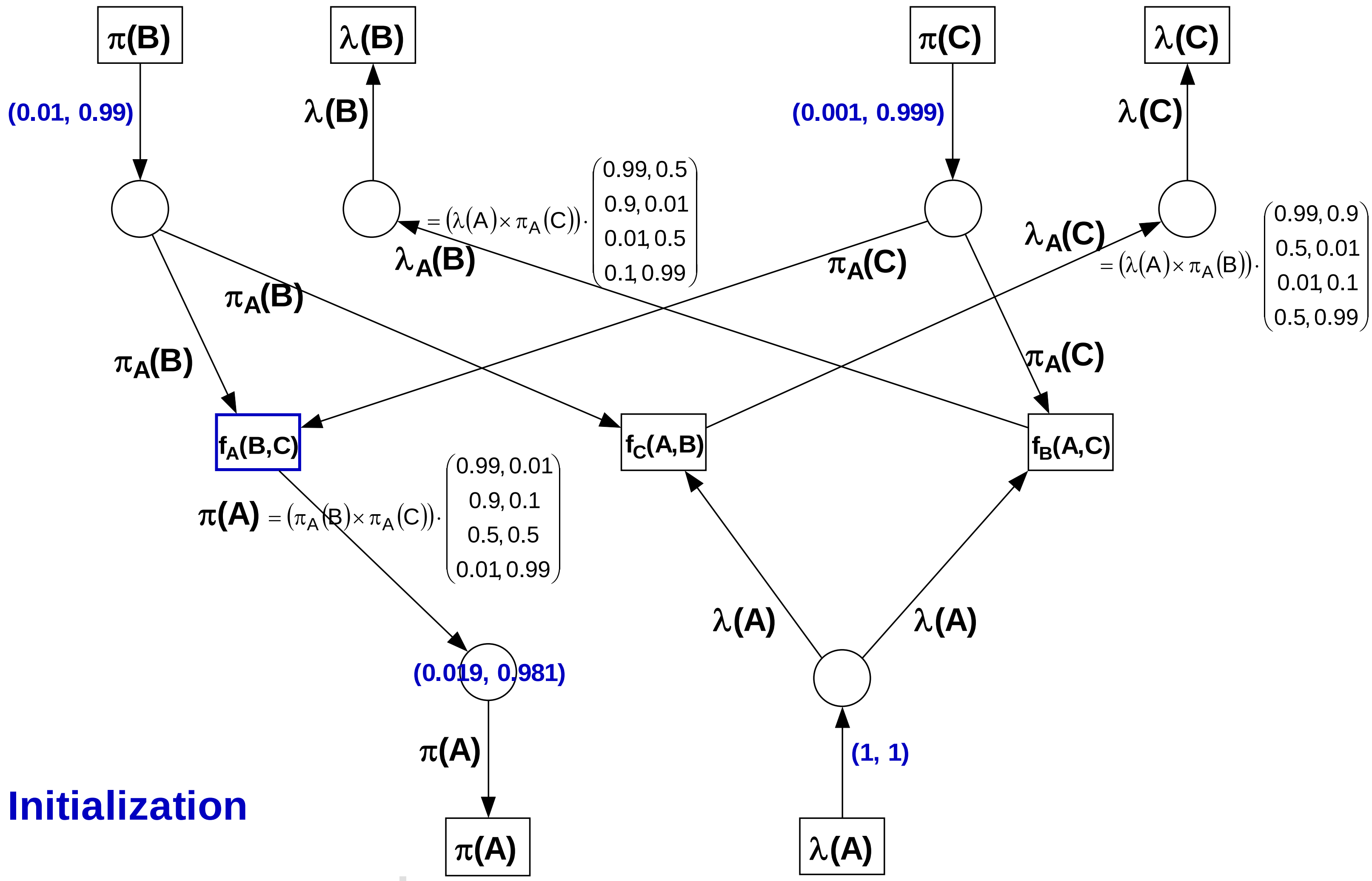
Initialization

Burglar Alarm



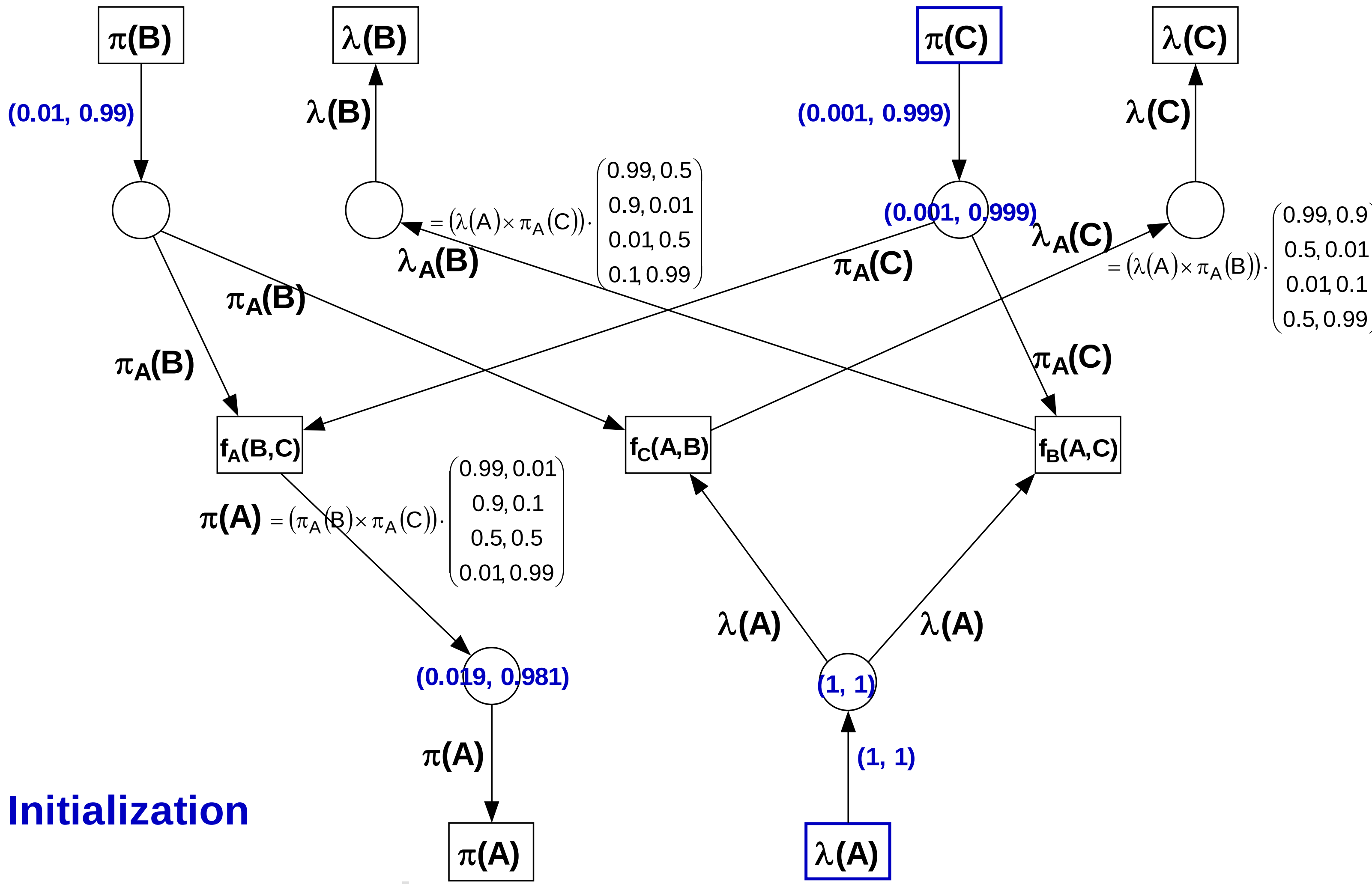
Initialization

Burglar Alarm



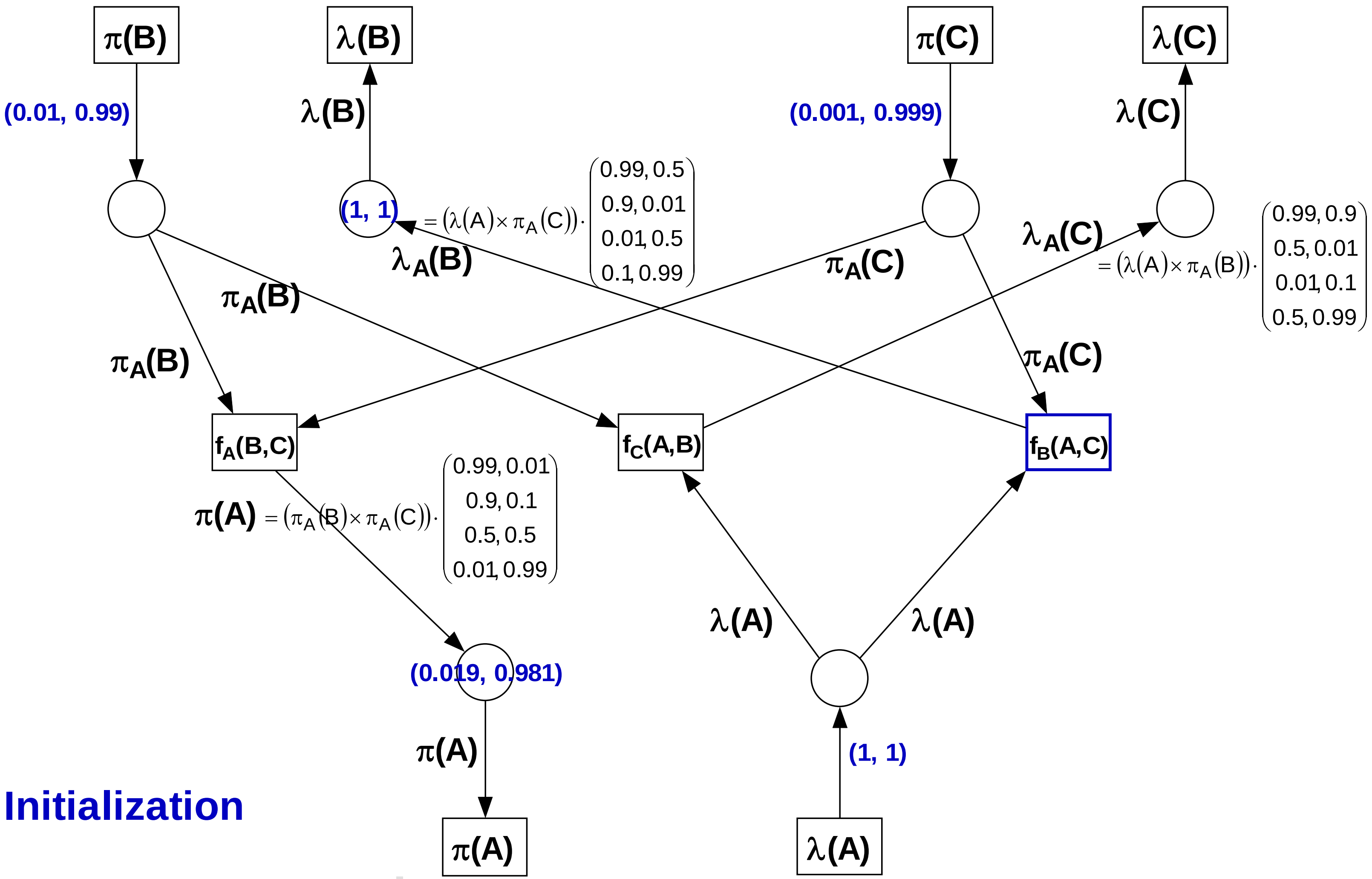
Initialization

Burglar Alarm



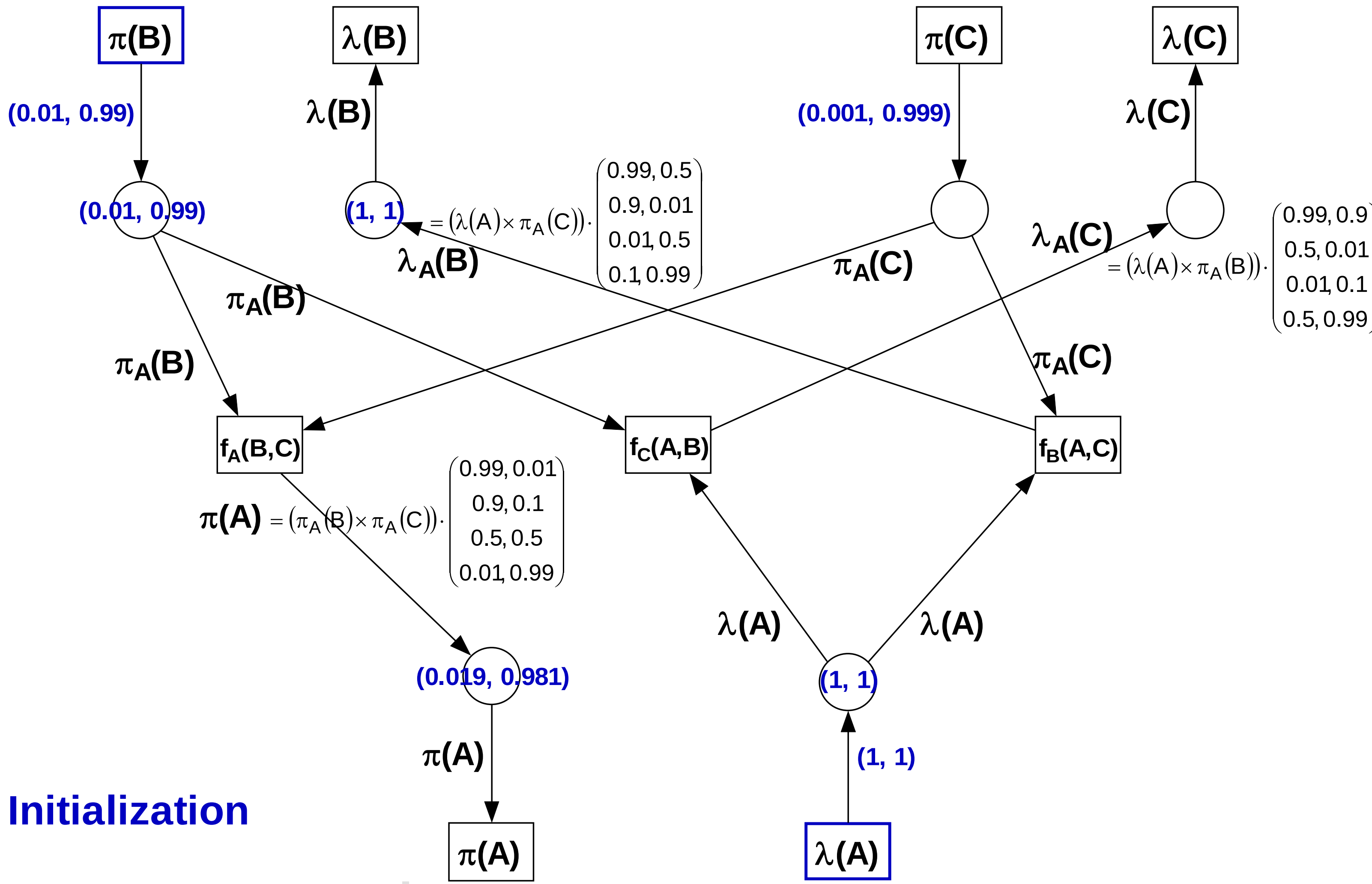
Initialization

Burglar Alarm



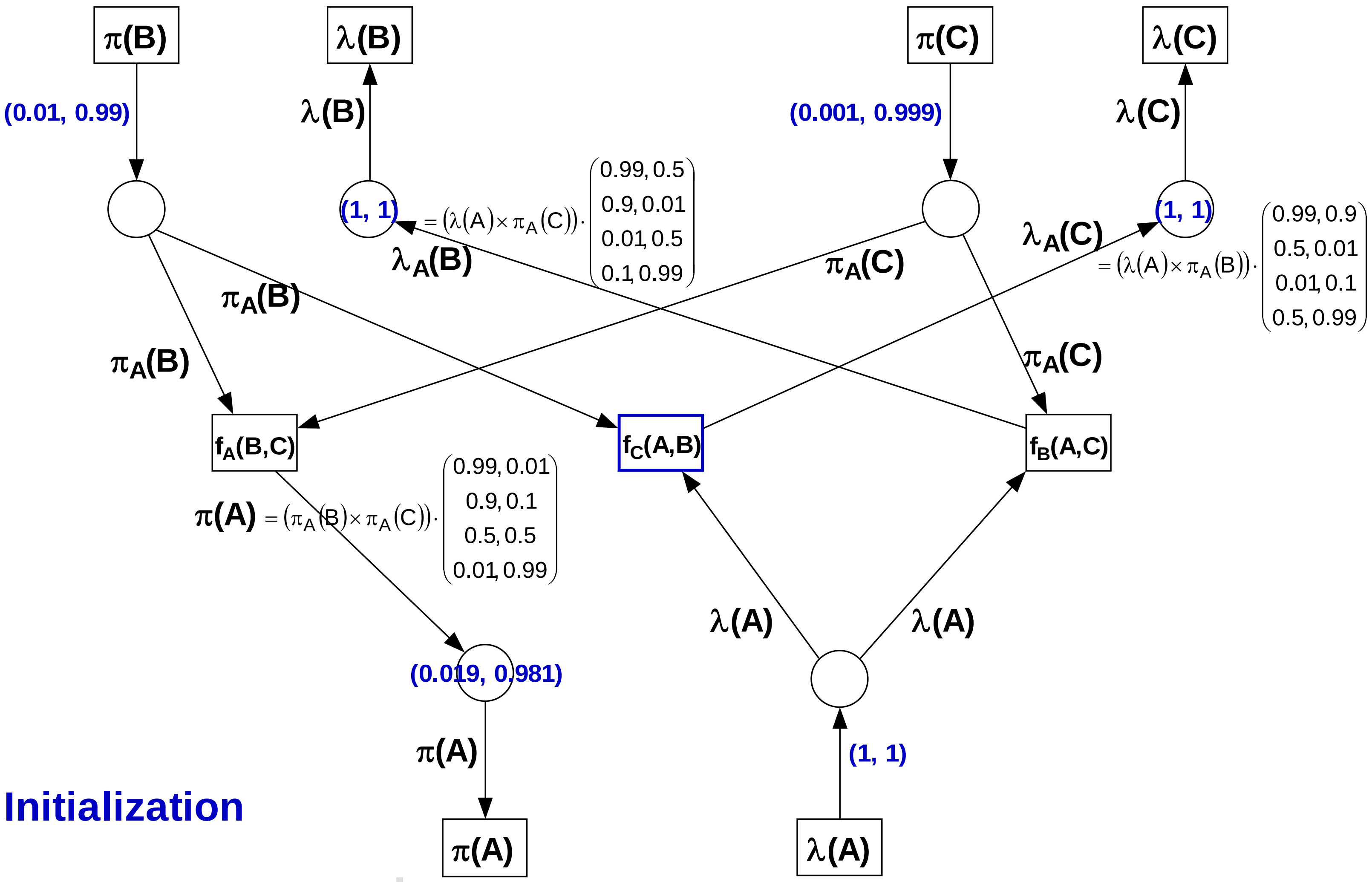
Initialization

Burglar Alarm



Initialization

Burglar Alarm



Initialization

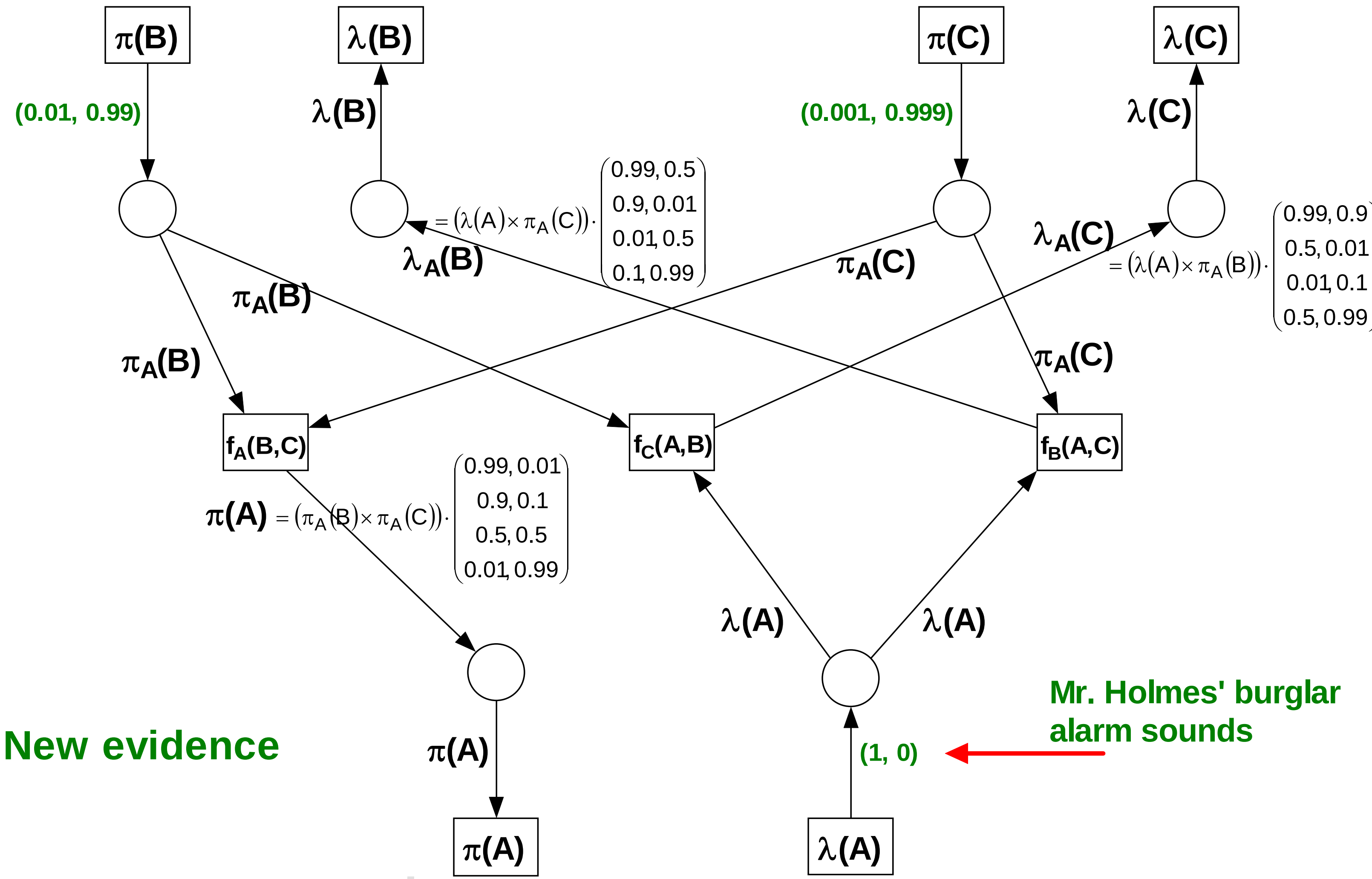
Calculation of beliefs

$bel(B) = \alpha((0.01, 0.99) \circ (1, 1)) = (0.01, 0.99)$

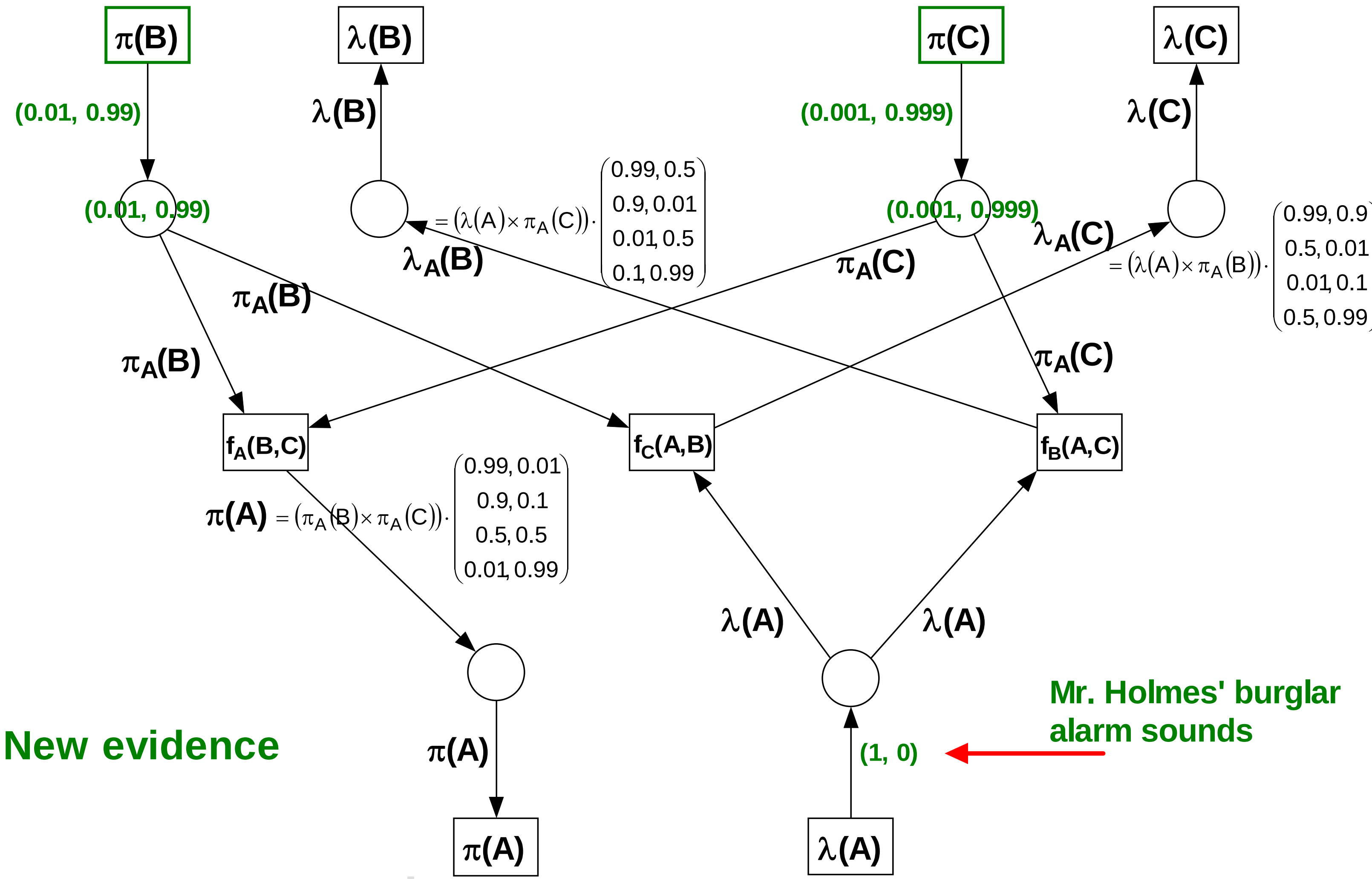
$bel(C) = \alpha((0.001, 0.999) \circ (1, 1)) = (0.001, 0.999)$

$bel(A) = \alpha((0.019, 0.981) \circ (1, 1)) = (0.019, 0.981)$

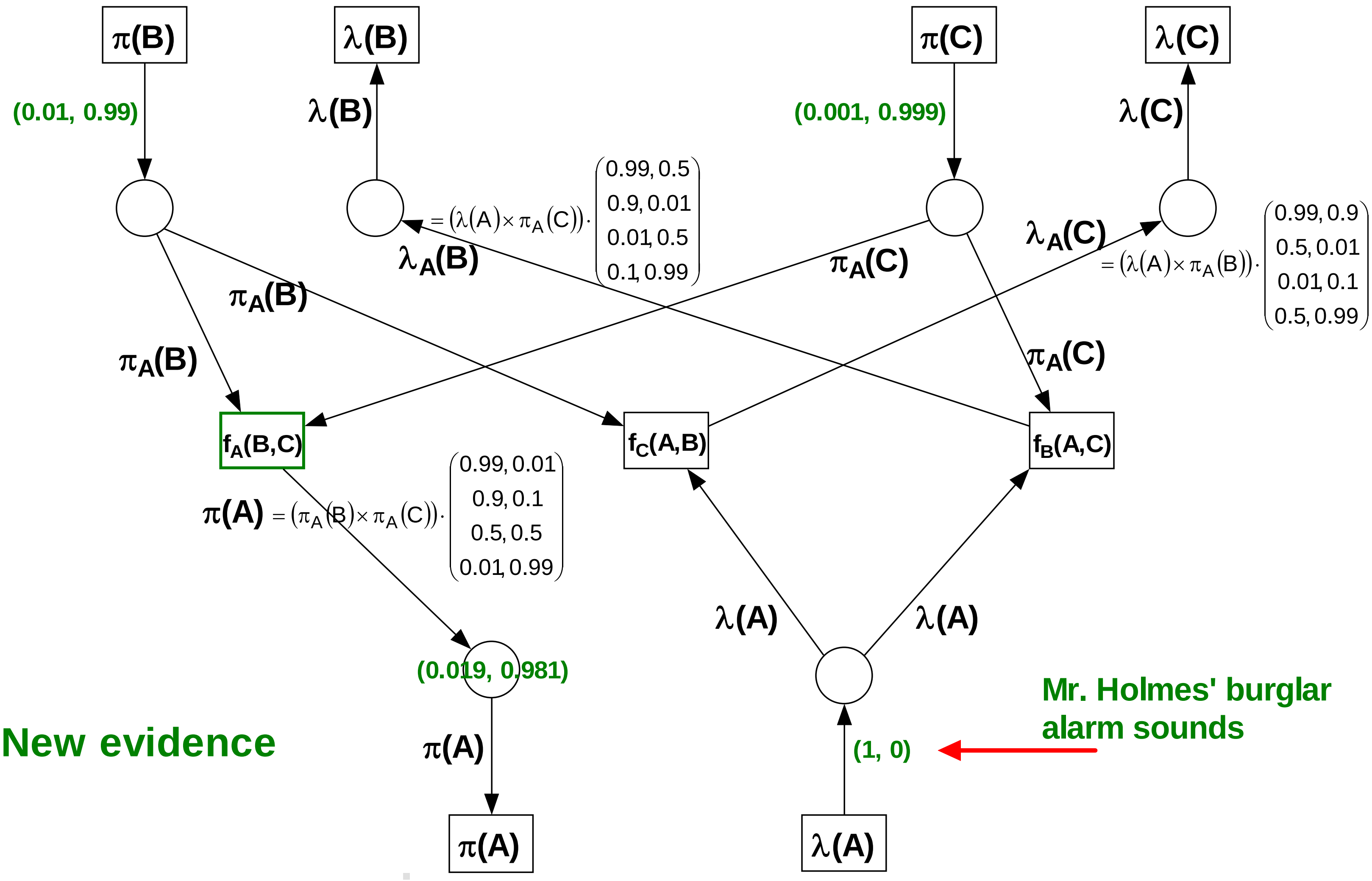
Burglar Alarm



Burglar Alarm

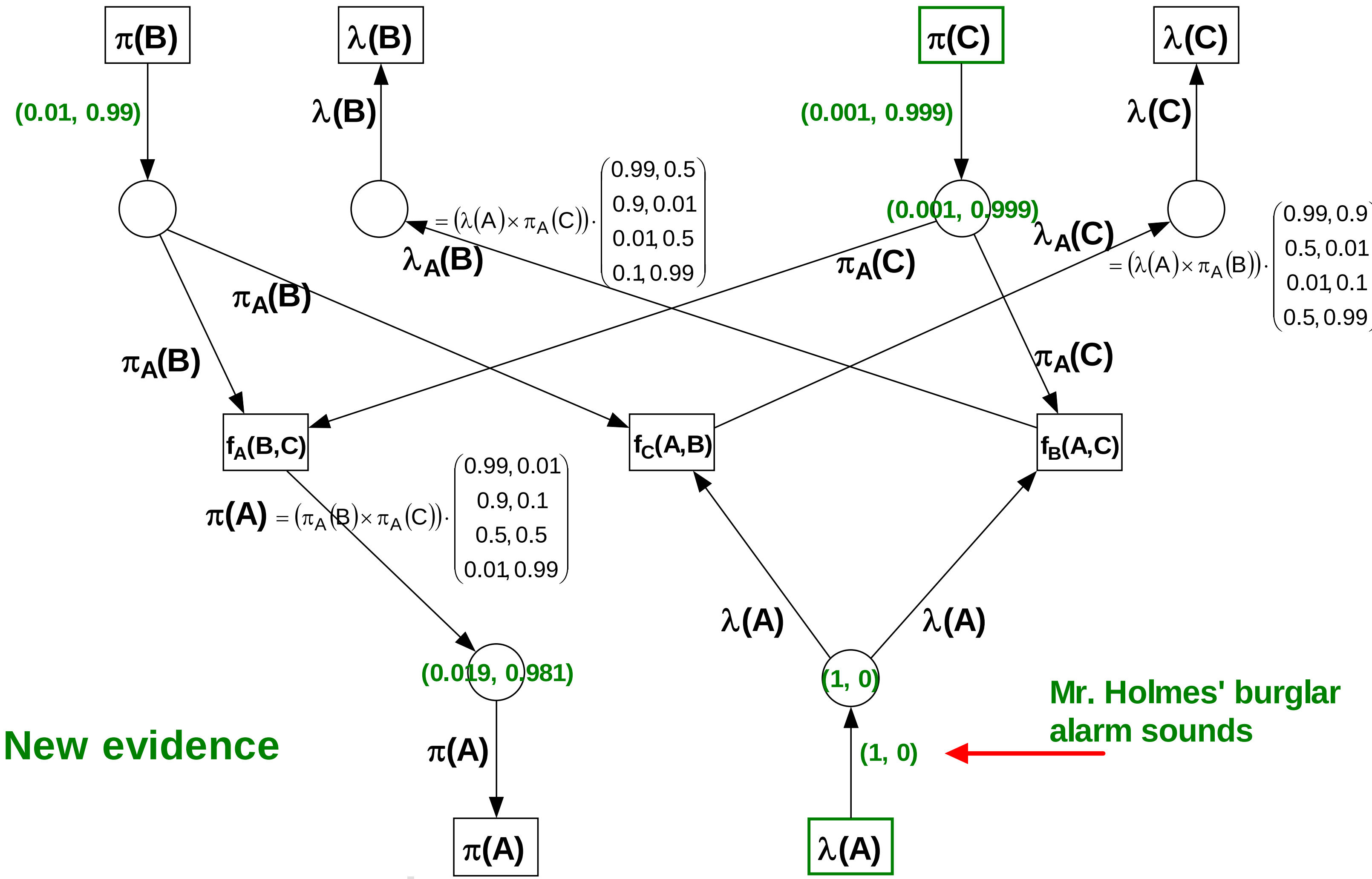


Burglar Alarm

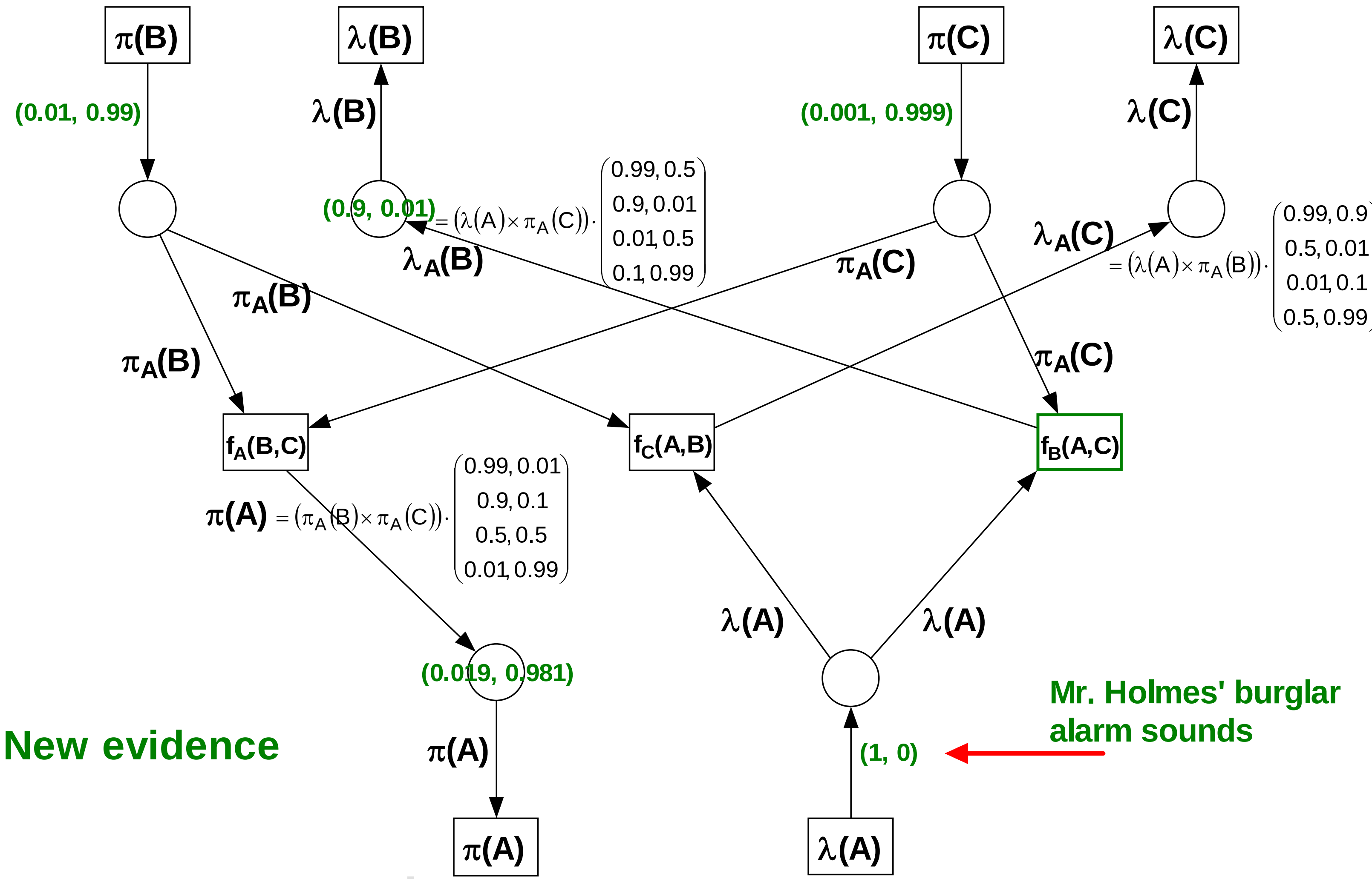


New evidence

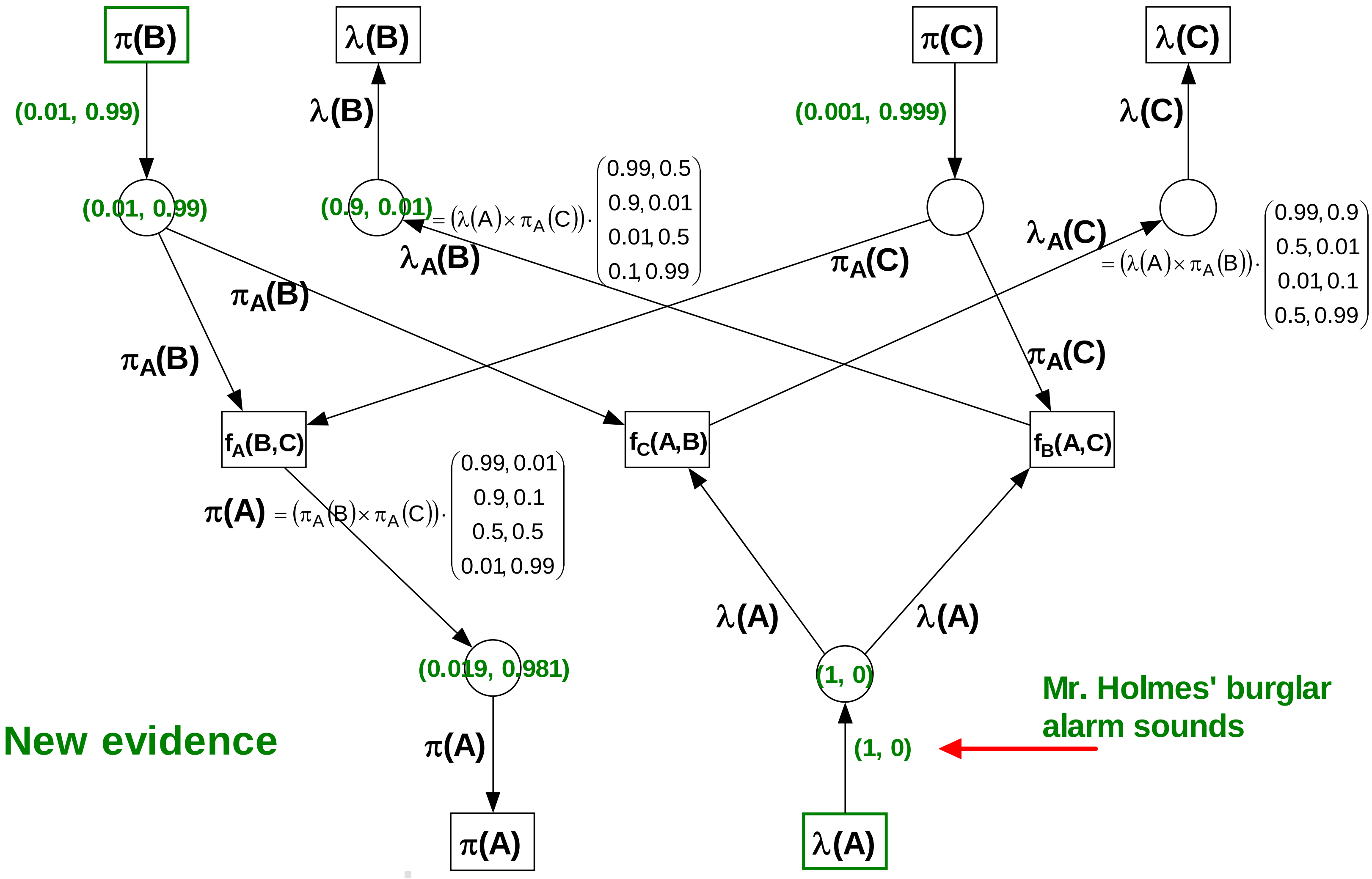
Burglar Alarm

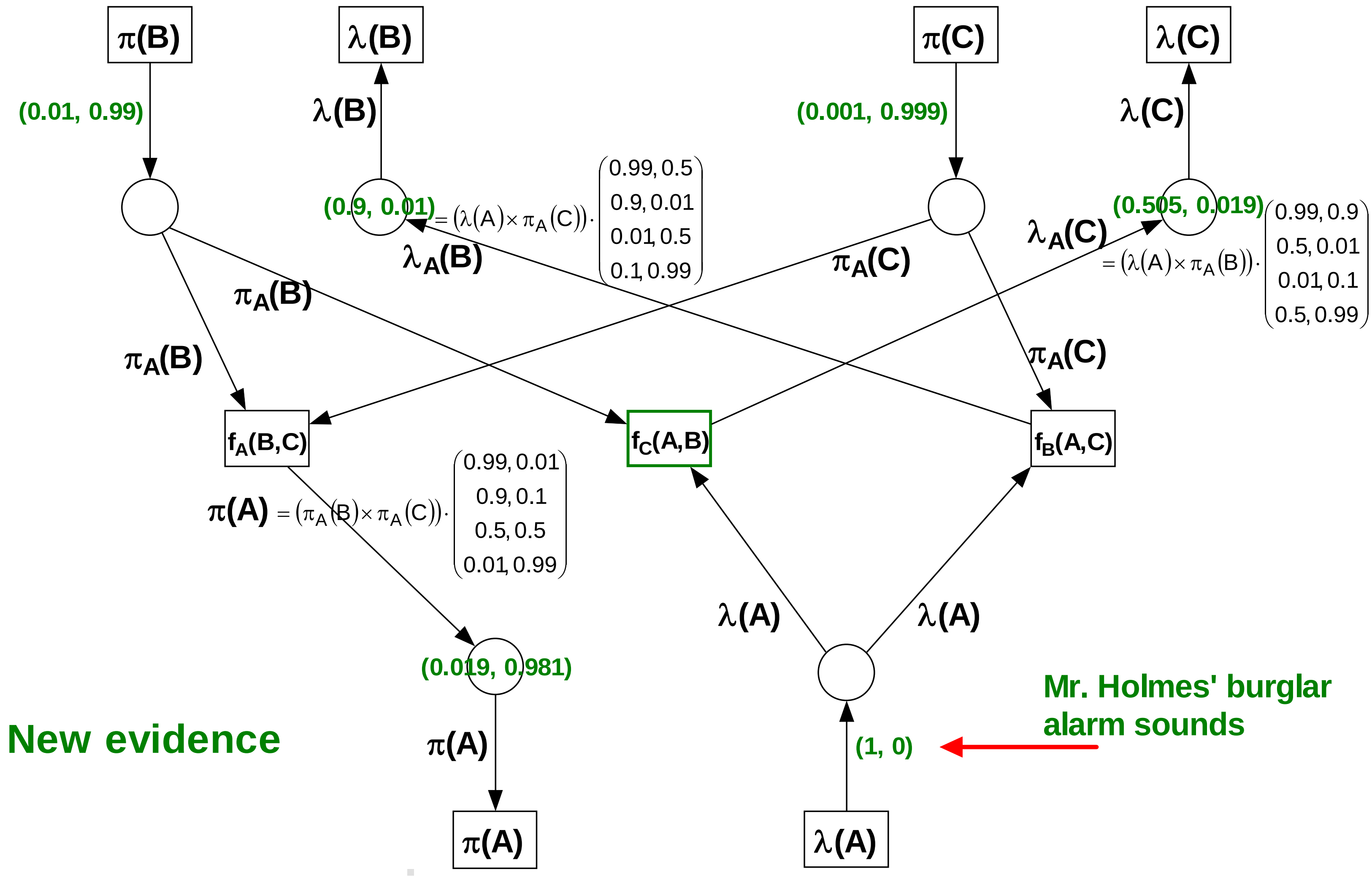


Burglar Alarm



Burglar Alarm




π(B)
λ(B)
π(C)
λ(C)
(0.01, 0.99)
λ(B)
(0.001, 0.999)
λ(C)
(0.9, 0.01)
$= (\lambda(A) \times \pi_A(C)) \cdot \begin{pmatrix} 0.99, 0.5 \\ 0.9, 0.01 \\ 0.01, 0.5 \\ 0.1, 0.99 \end{pmatrix}$
$\lambda_A(B)$
$\pi_A(B)$
$\pi_A(B)$
$\pi_A(C)$
$\lambda_A(C)$
(0.505, 0.019)
$= (\lambda(A) \times \pi_A(B)) \cdot \begin{pmatrix} 0.99, 0.9 \\ 0.5, 0.01 \\ 0.01, 0.1 \\ 0.5, 0.99 \end{pmatrix}$
$\pi_A(C)$
$f_A(B,C)$
$f_C(A,B)$
$f_B(A,C)$
π(A) $= (\pi_A(B) \times \pi_A(C)) \cdot \begin{pmatrix} 0.99, 0.01 \\ 0.9, 0.1 \\ 0.5, 0.5 \\ 0.01, 0.99 \end{pmatrix}$
λ(A)
λ(A)
(0.019, 0.981)
New evidence
π(A)
(1, 0)
Mr. Holmes' burglar alarm sounds
π(A)
λ(A)

Calculation of beliefs

$bel(B) = \alpha((0.01,\ 0.99) \circ (1, 1)) = (0.01,\ 0.99)$

$bel(C) = \alpha((0.001, 0.999) \circ (1, 1)) = (0.001, 0.999)$

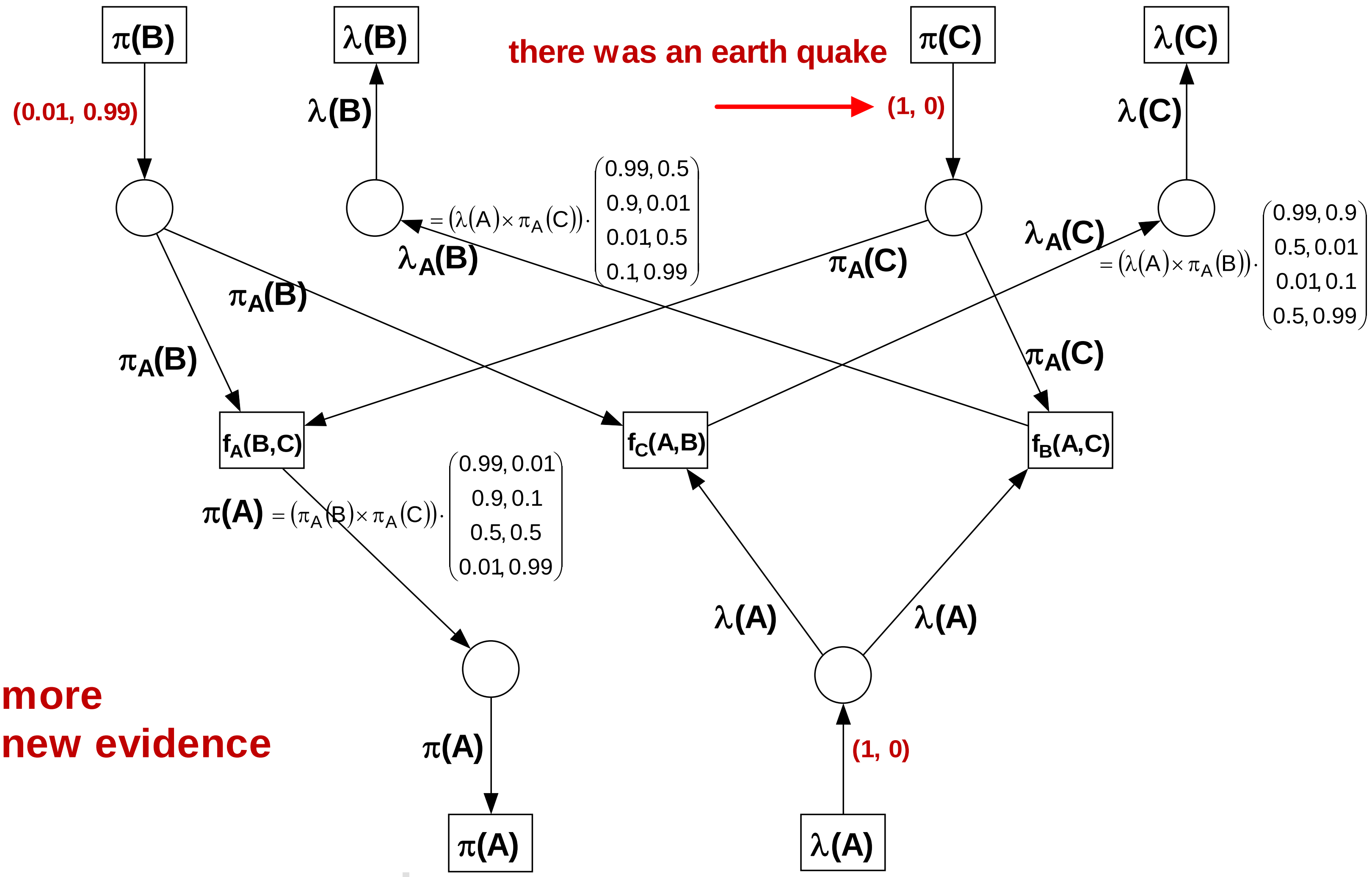
$bel(A) = \alpha((0.019, 0.981) \circ (1, 1)) = (0.019, 0.981)$

$bel(B) = \alpha((0.01,\ 0.99) \circ (0.9, 0.01)) = (0.476, 0.524)$

$bel(C) = \alpha((0.001, 0.999) \circ (0.505, 0.019)) = (0.026, 0.974)$

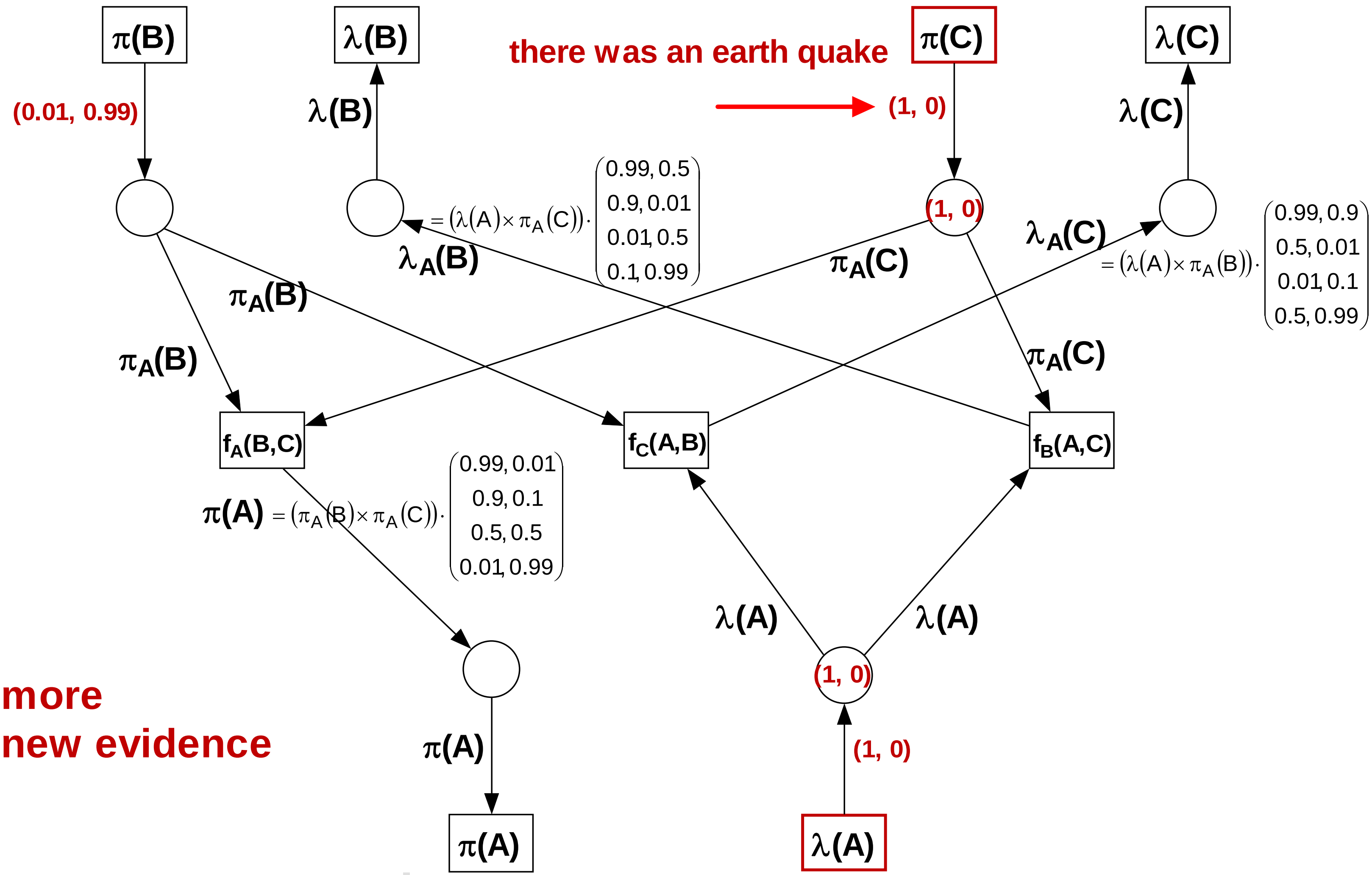
$bel(A) = \alpha((0.019, 0.981) \circ (1, 0)) = (1, 0)$

Burglar Alarm



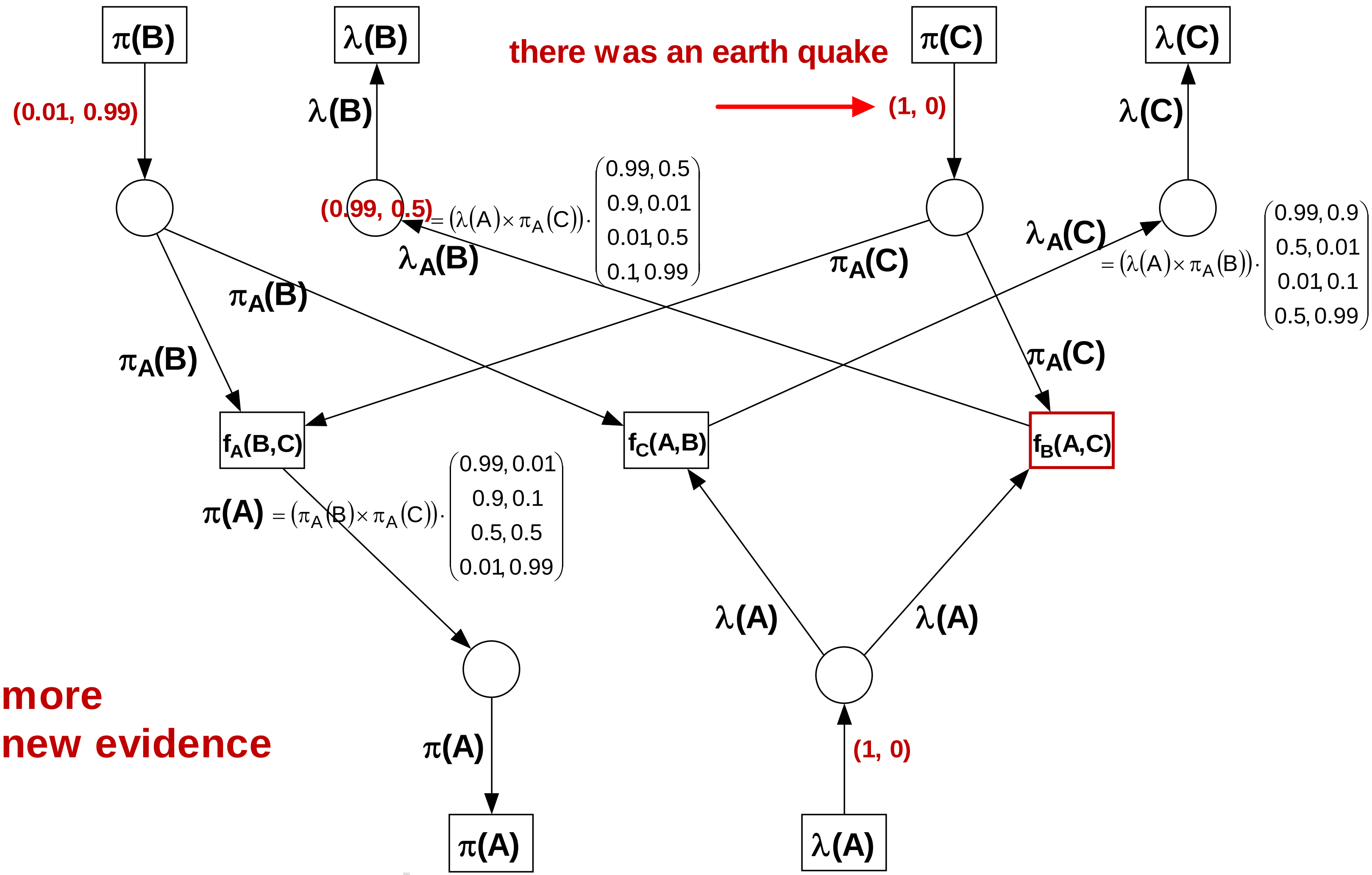
more new evidence

Burglar Alarm



more
new evidence

Burglar Alarm



more
new evidence

Calculation of beliefs

$bel(B) = \alpha((0.01,\ 0.99) \circ (1, 1)) = (0.01,\ 0.99)$

$bel(C) = \alpha((0.001, 0.999) \circ (1, 1)) = (0.001, 0.999)$

$bel(A) = \alpha((0.019, 0.981) \circ (1, 1)) = (0.019, 0.981)$

$bel(B) = \alpha((0.01,\ 0.99) \circ (0.9, 0.01)) = (0.476, 0.524)$

$bel(C) = \alpha((0.001, 0.999) \circ (0.505, 0.019)) = (0.026, 0.974)$

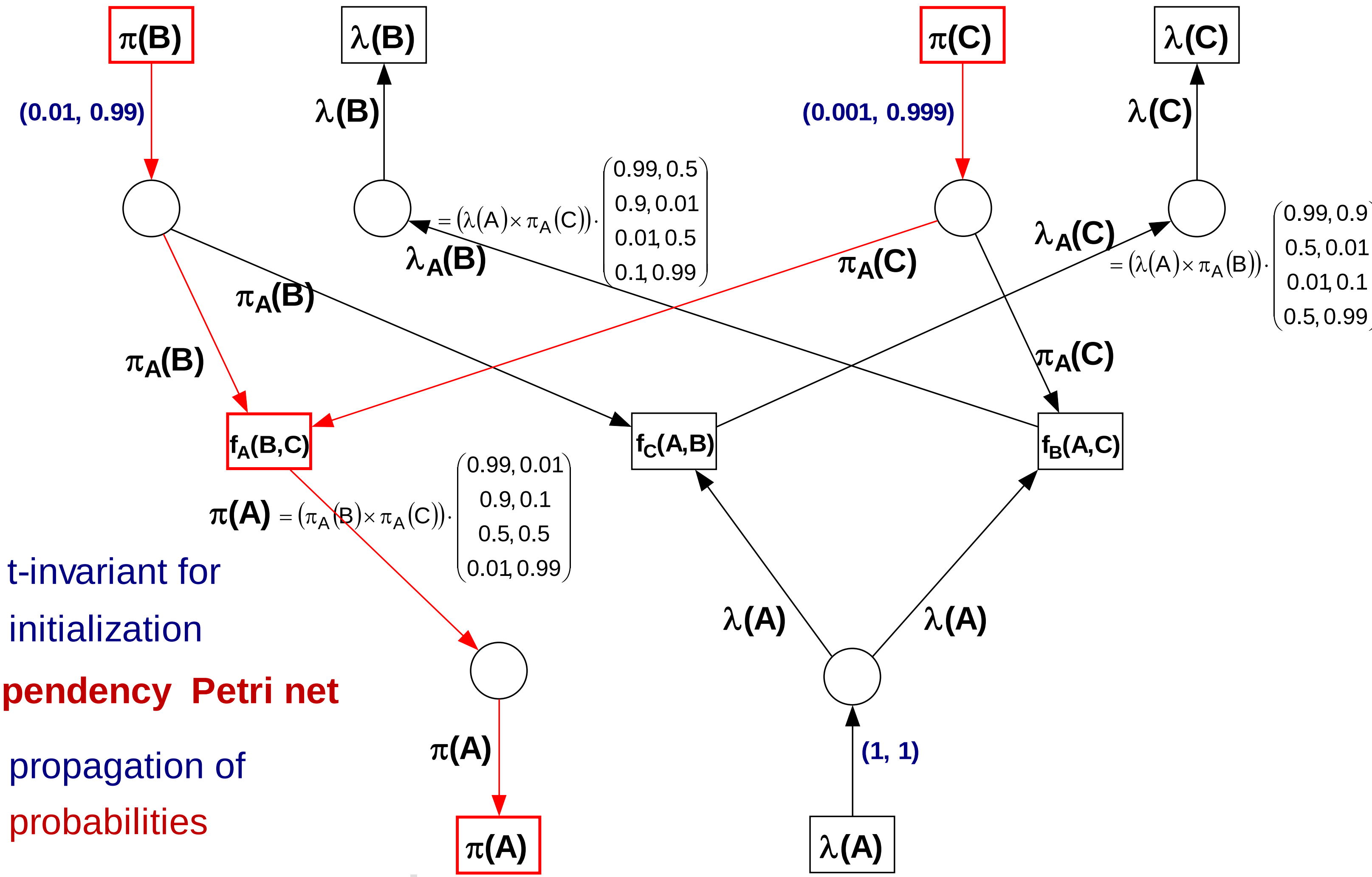
$bel(A) = \alpha((0.019, 0.981) \circ (1, 0)) = (1, 0)$

$bel(B) = \alpha((0.01,\ 0.99) \circ (0.99, 0.5)) = (0.02, 0.98)$

$bel(C) = \alpha((0.001, 0.999) \circ (1, 0)) = (1, 0)$

$bel(A) = \alpha((0.019, 0.981) \circ (1, 0)) = (1, 0)$

Burglar Alarm



t-invariant for

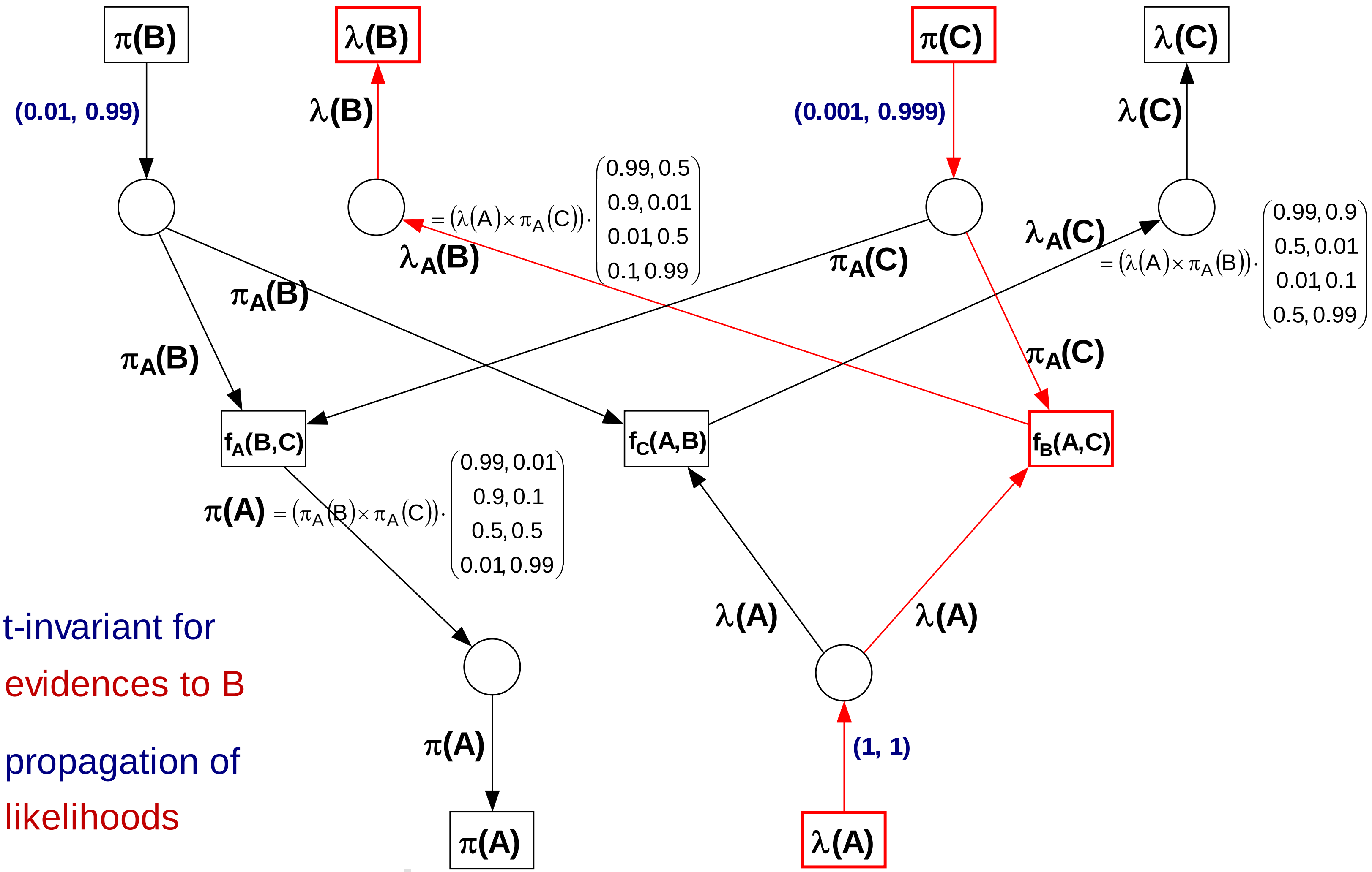
initialization

dependency Petri net

propagation of

probabilities

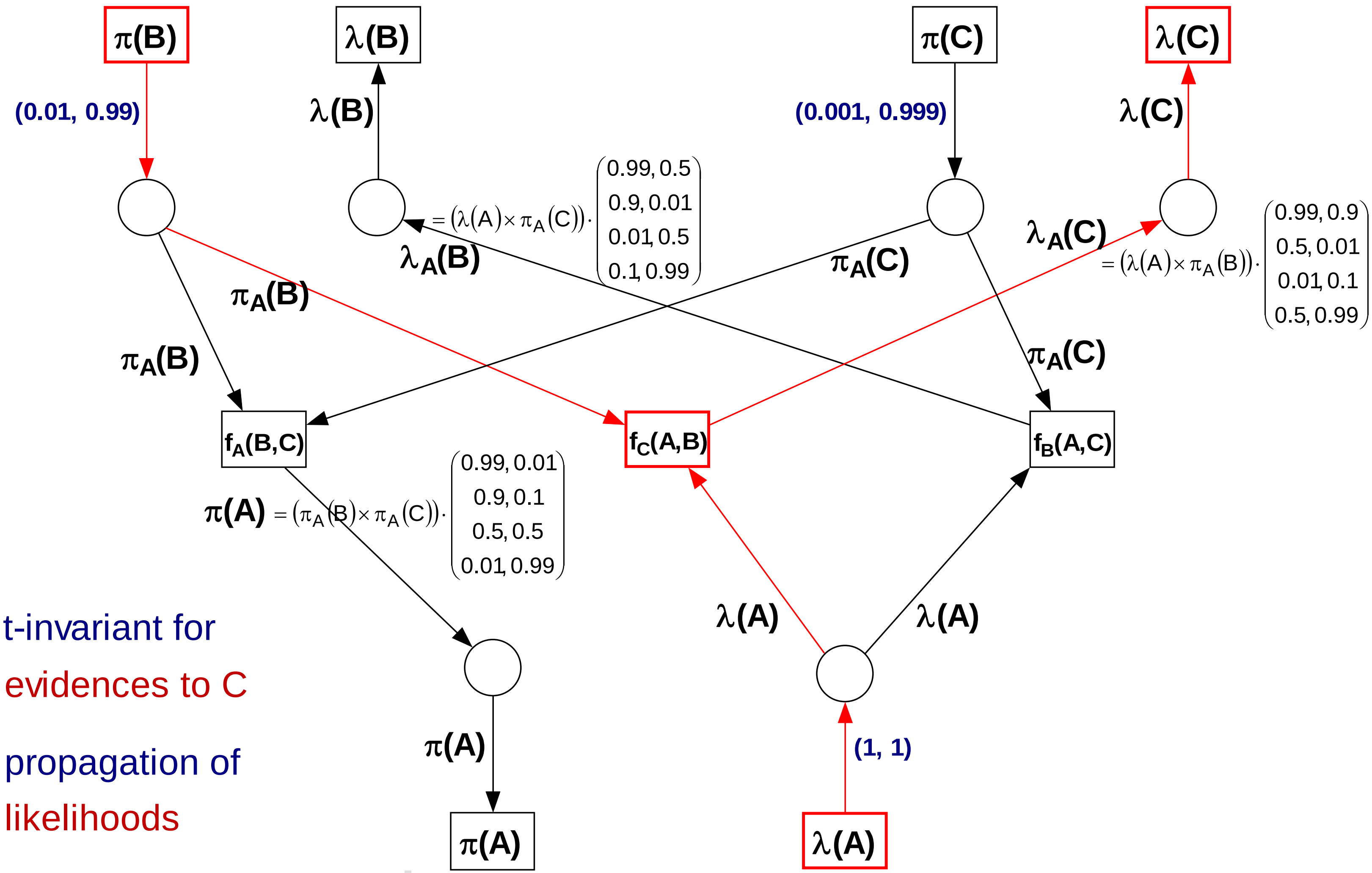
Burglar Alarm



t-invariant for
evidences to B

propagation of
likelihoods

Burglar Alarm



t-invariant for evidences to C

propagation of likelihoods

A likelihood is a conditional probability in a specific interpretation.

Let S be a symptom and D a diagnosis (of kind disease), then

$P(S|D)$ is a "causal" (conditional) probability for S

$P(D|S)$ is a "diagnostic" (conditional) probability for D

In case of several conceivable diagnoses D_j ,
$P(S|D_j)$ is a measure for S indicating how probable it is that D_j causes S;

on the other hand, $P(S|D_j)$ is a degree of confirmation that D_j is the cause for S and as such, it is a measure for D_j,

which is called the "likelihood" $N(D_j|S) = P(S|D_j)$ of D_j given S

Bayesian networks are due to Judea Pearl

Books:

Judea Pearl;

Probabilistic Reasoning in Intelligent Systems:
Networks of Plausible Inference;
Morgan Kaufmann Publishers, Inc.,
San Francisco, 1988.

Richard E. Neapolitan;

Probabilistic Reasoning in Expert Systems:
Theory and Algorithms;
John Wiley & Sons, Inc.,
New York, 1990.

Reverend Thomas Bayes
1701? - 1761

Student of de Moivre

Bayes' theorem (1750): $p(A|B) \cdot p(B) = p(B|A) \cdot p(A) = p(A \wedge B)$

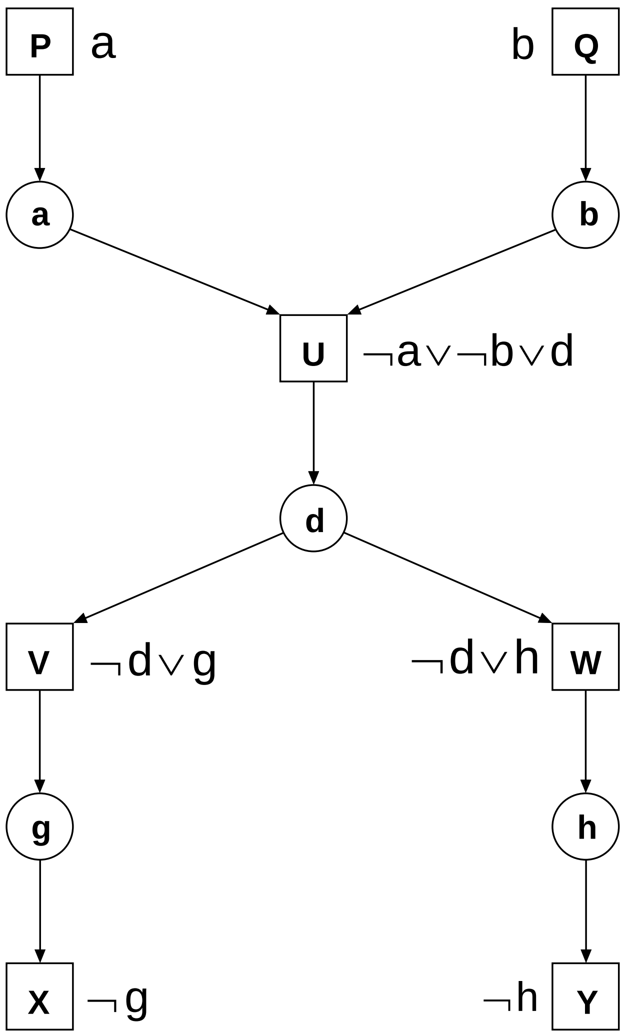
Contents

This net is an "overlay" of two Horn nets (Lautenbach; 2002, 2003)

The underlying Horn formulae are

$a \wedge b \wedge (\neg a \vee \neg b \vee d) \wedge (\neg d \vee g) \wedge \neg g$

$a \wedge b \wedge (\neg a \vee \neg b \vee d) \wedge (\neg d \vee h) \wedge \neg h$

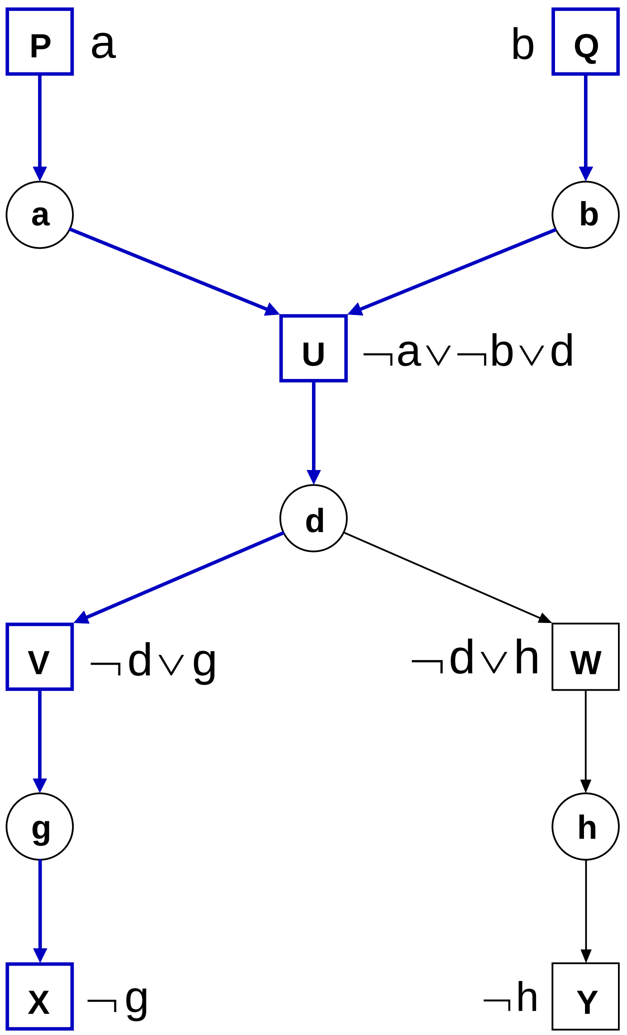


This net is an "overlay" of two Horn nets

The underlying Horn formulae are

$a \wedge b \wedge (\neg a \vee \neg b \vee d) \wedge (\neg d \vee g) \wedge \neg g$

$a \wedge b \wedge (\neg a \vee \neg b \vee d) \wedge (\neg d \vee h) \wedge \neg h$

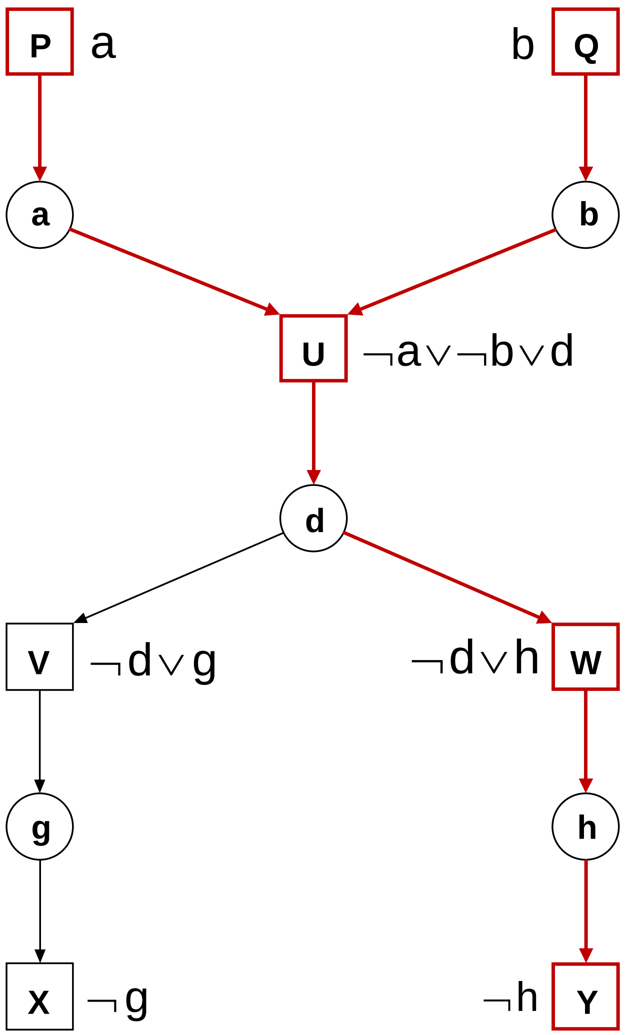


This net is an "overlay" of two Horn nets

The underlying Horn formulae are

$a \wedge b \wedge (\neg a \vee \neg b \vee d) \wedge (\neg d \vee g) \wedge \neg g$

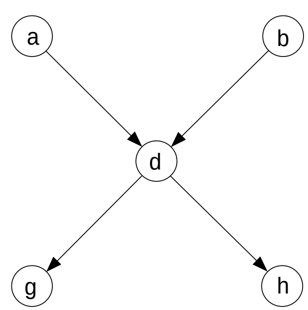
$a \wedge b \wedge (\neg a \vee \neg b \vee d) \wedge (\neg d \vee h) \wedge \neg h$



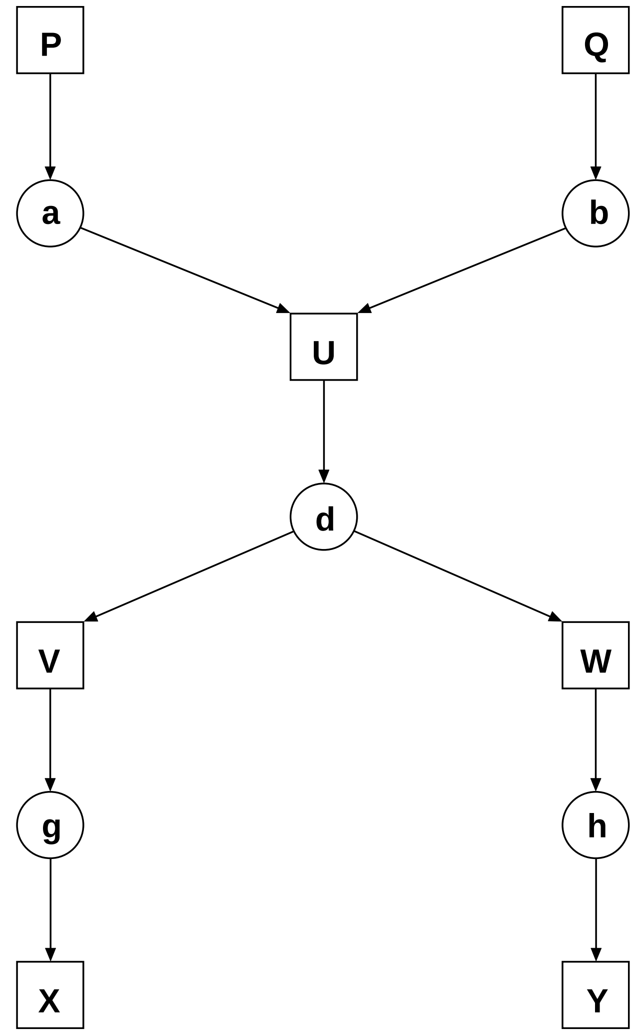
Dependency Nets

A dependency Petri net is a net
N=(P, T, F) where the following holds:

- N has a transition boundary
- $\forall_{k \in P \cup T} \; (|{}^{\bullet}k| \geq 2 \Rightarrow k \in T) \wedge (|k^{\bullet}| \geq 2 \Rightarrow k \in P)$
- N is connected and circle free



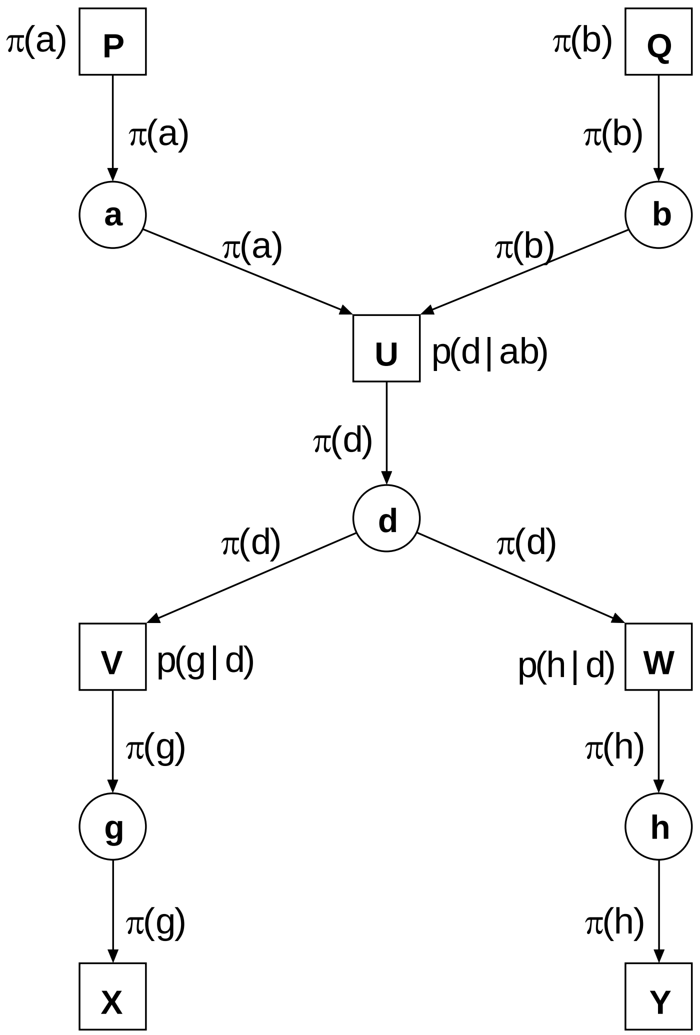
the corresponding Bayesian network



Dependency nets

- have nice net properties e.g. liveness in both directions;

- their t-invariants represent the initializing flows and as such the dependency structure of Bayesian networks;

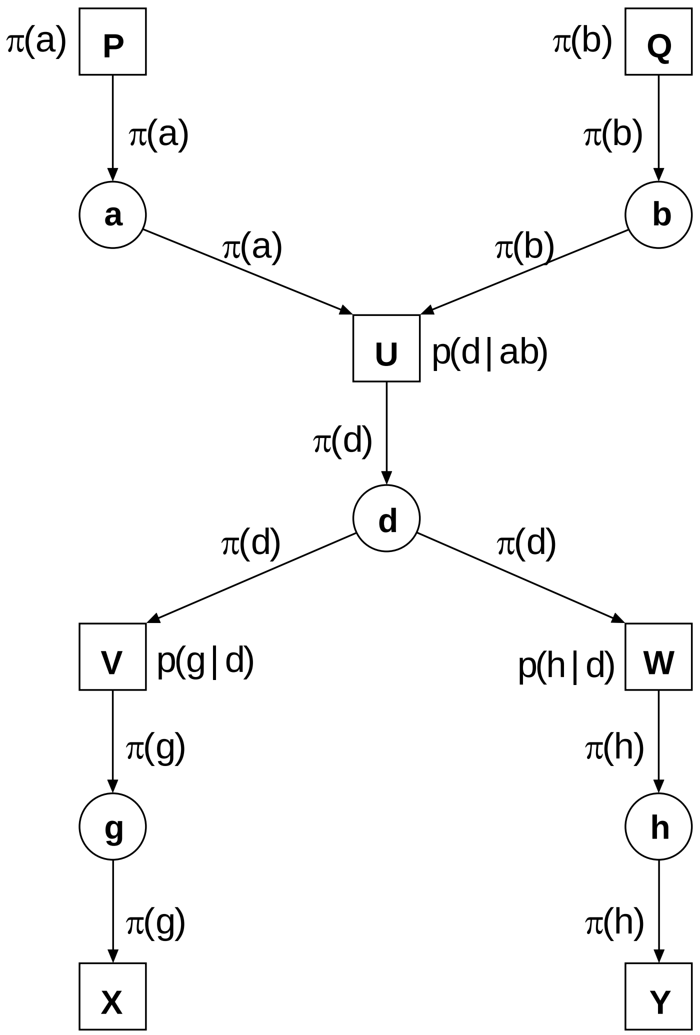


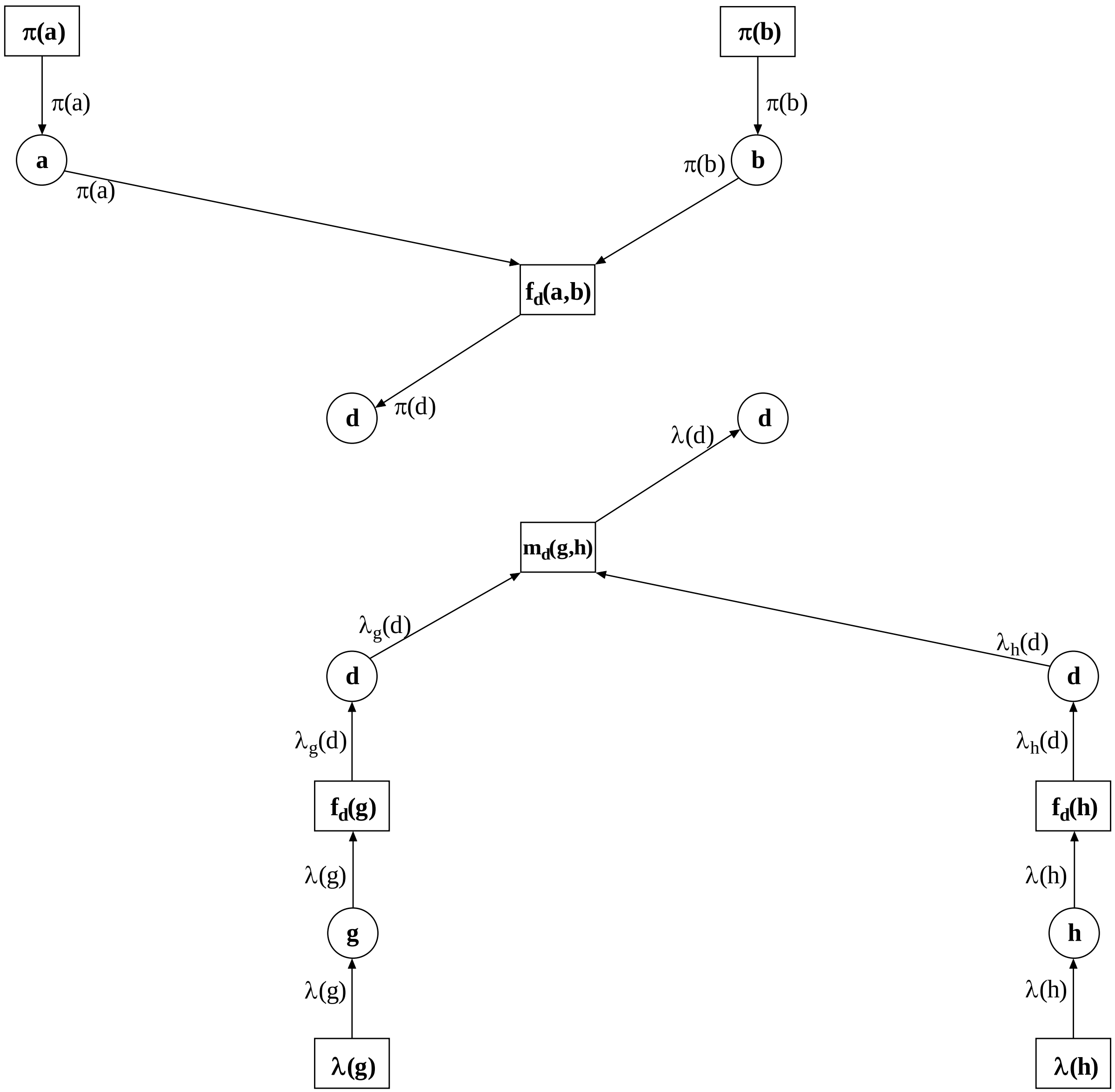
Dependency Nets

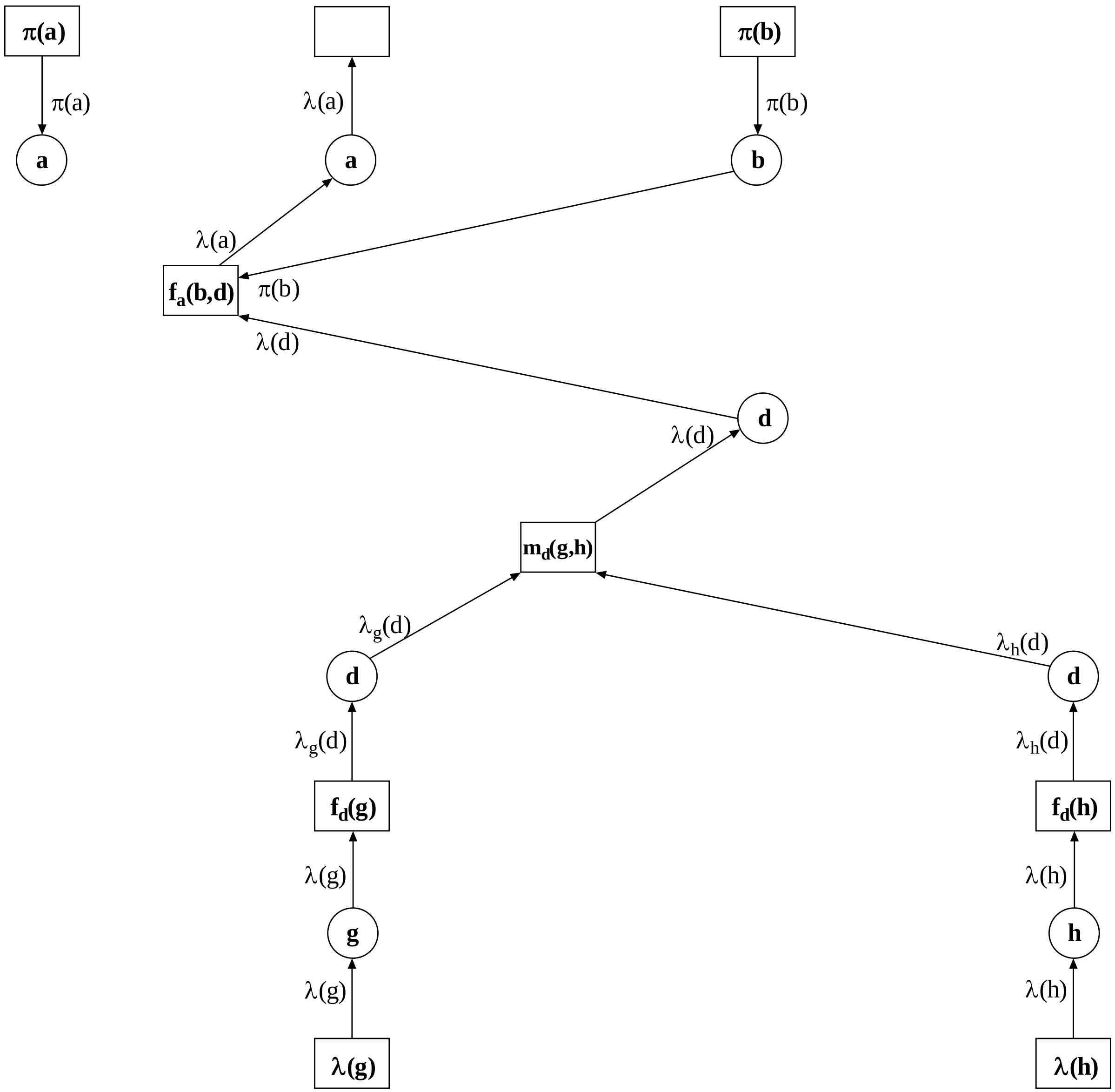
Dependency nets with inscriptions can be transformed into probability propagation nets by means of the following formula

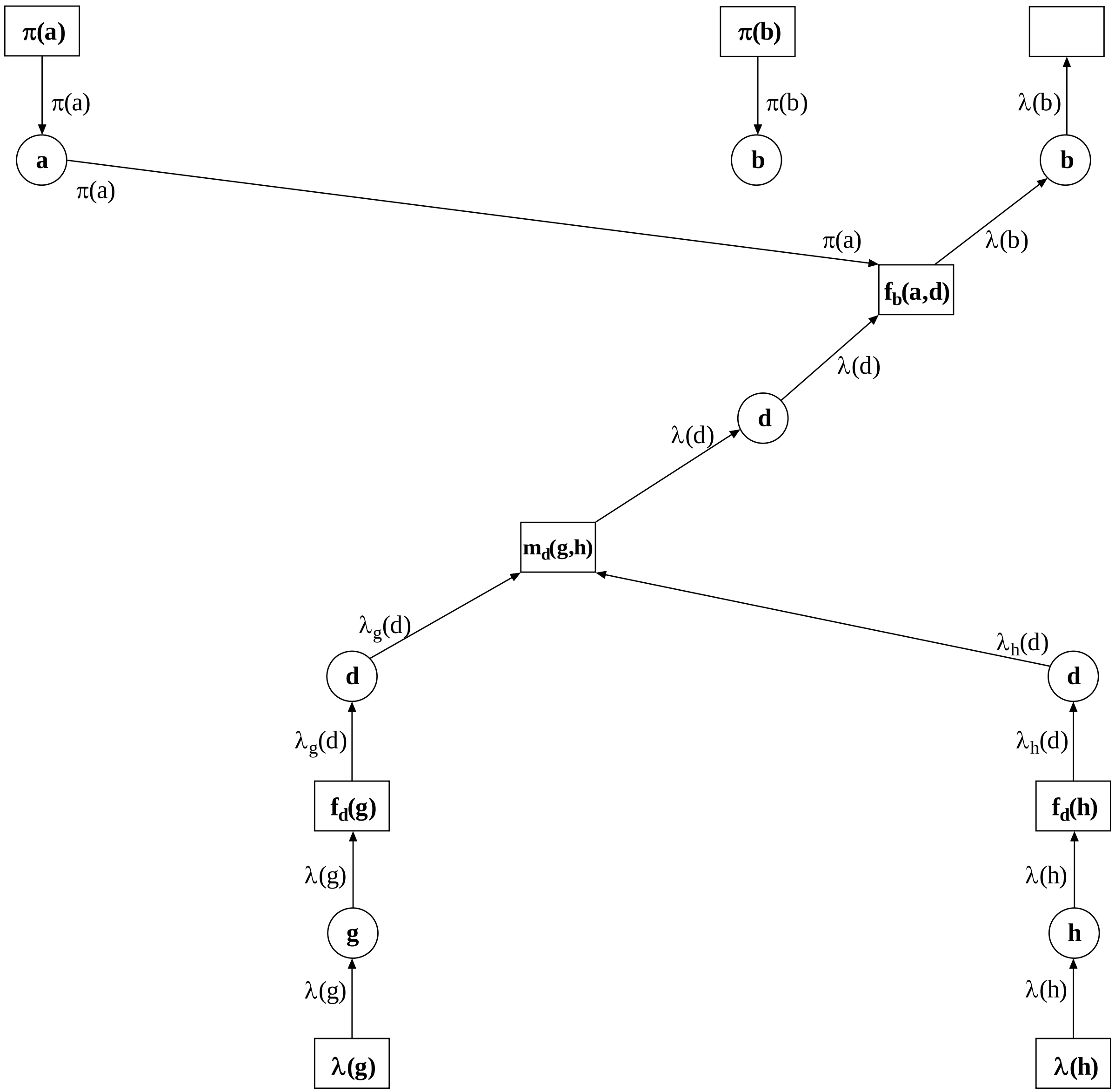
$$\pi(p\,|\,E) = \sum_{p\downarrow} \sum_{R\uparrow} [\lambda(p\downarrow) \cdot \pi(p\downarrow|\,p) \cdot \pi(p\,|\,ab) \cdot \pi(R\uparrow)]$$

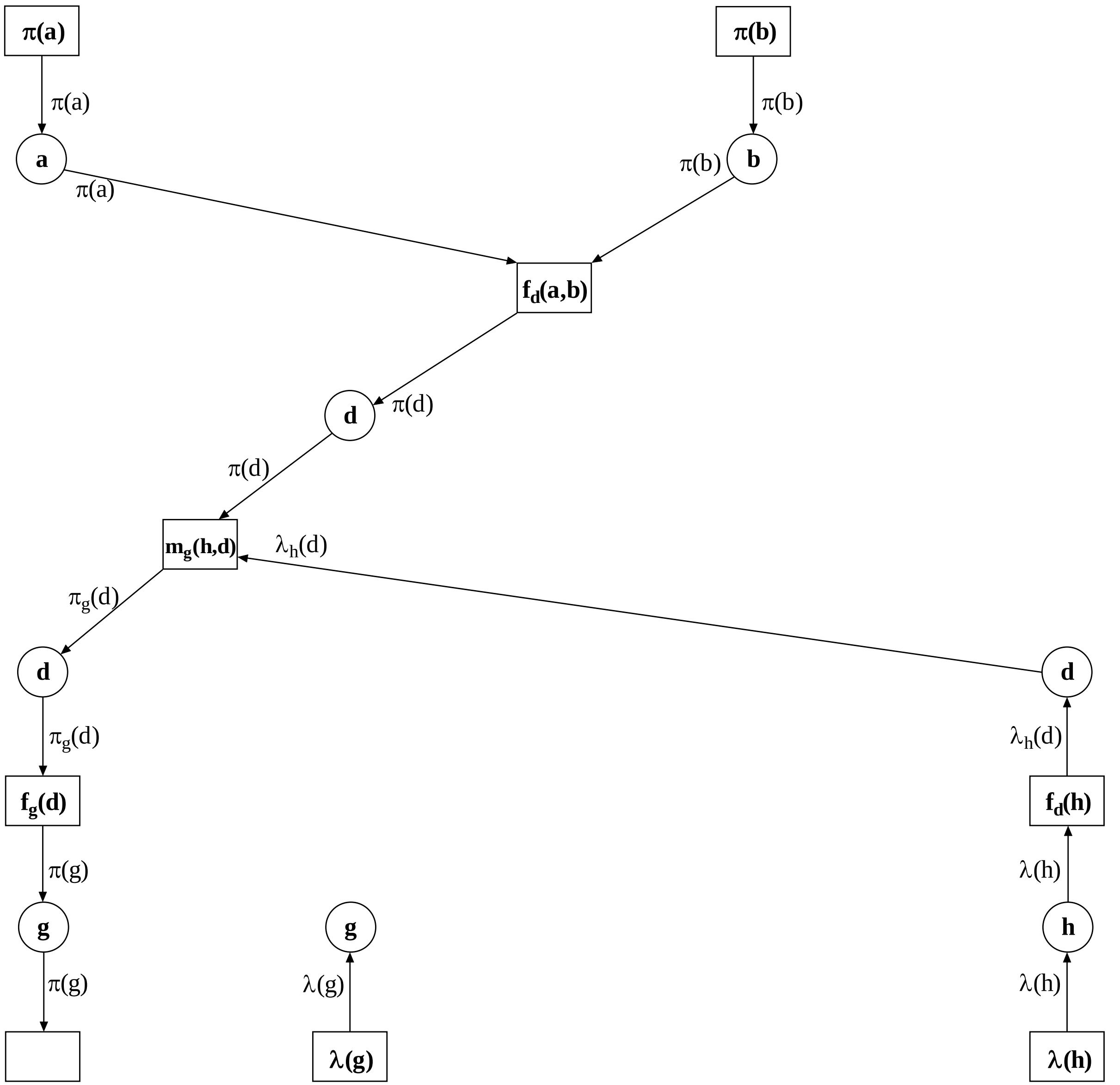
where (R,p) runs over the set {(P,a), (Q,b), (U,d), (V,g), (W,h)}

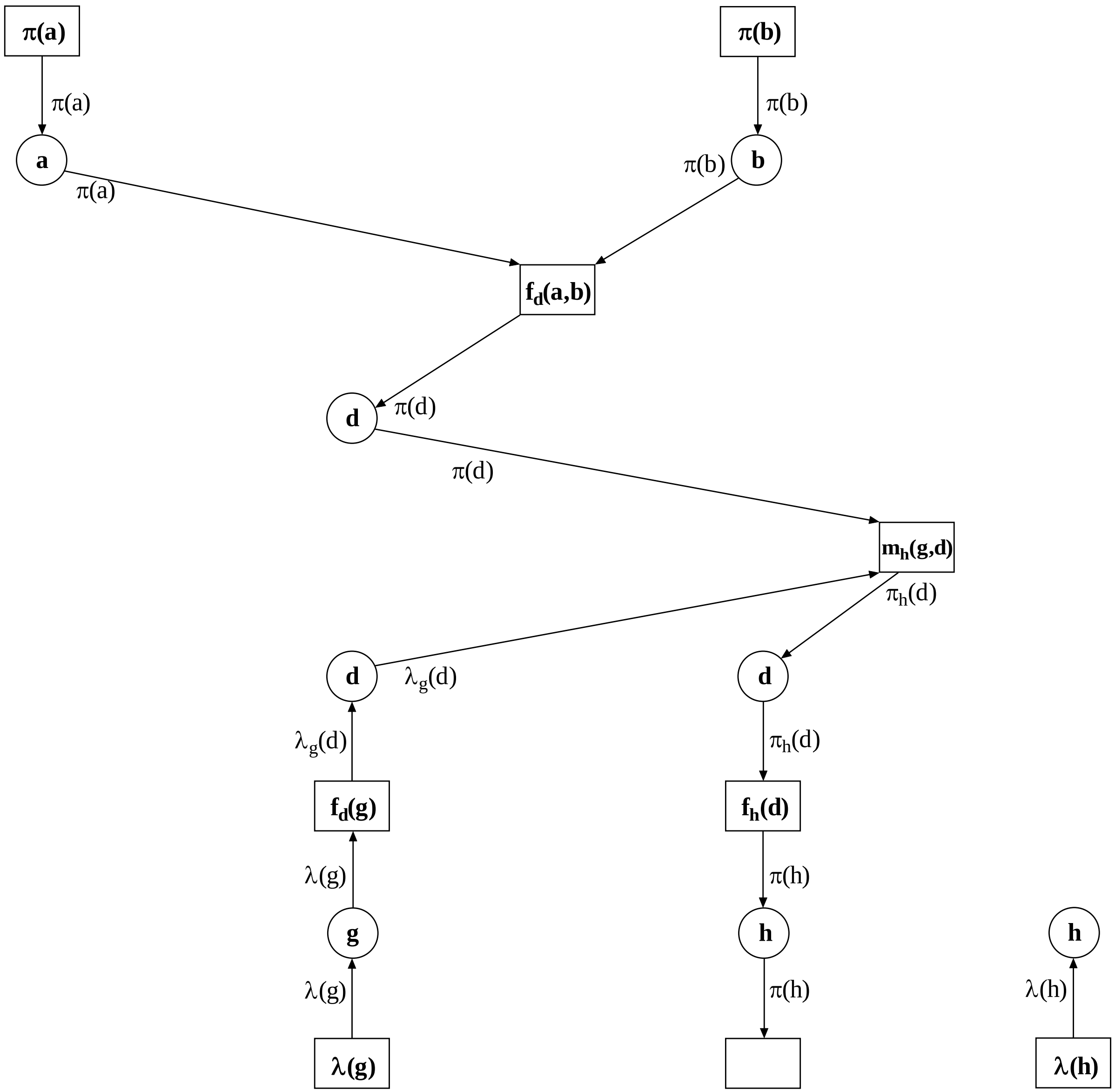


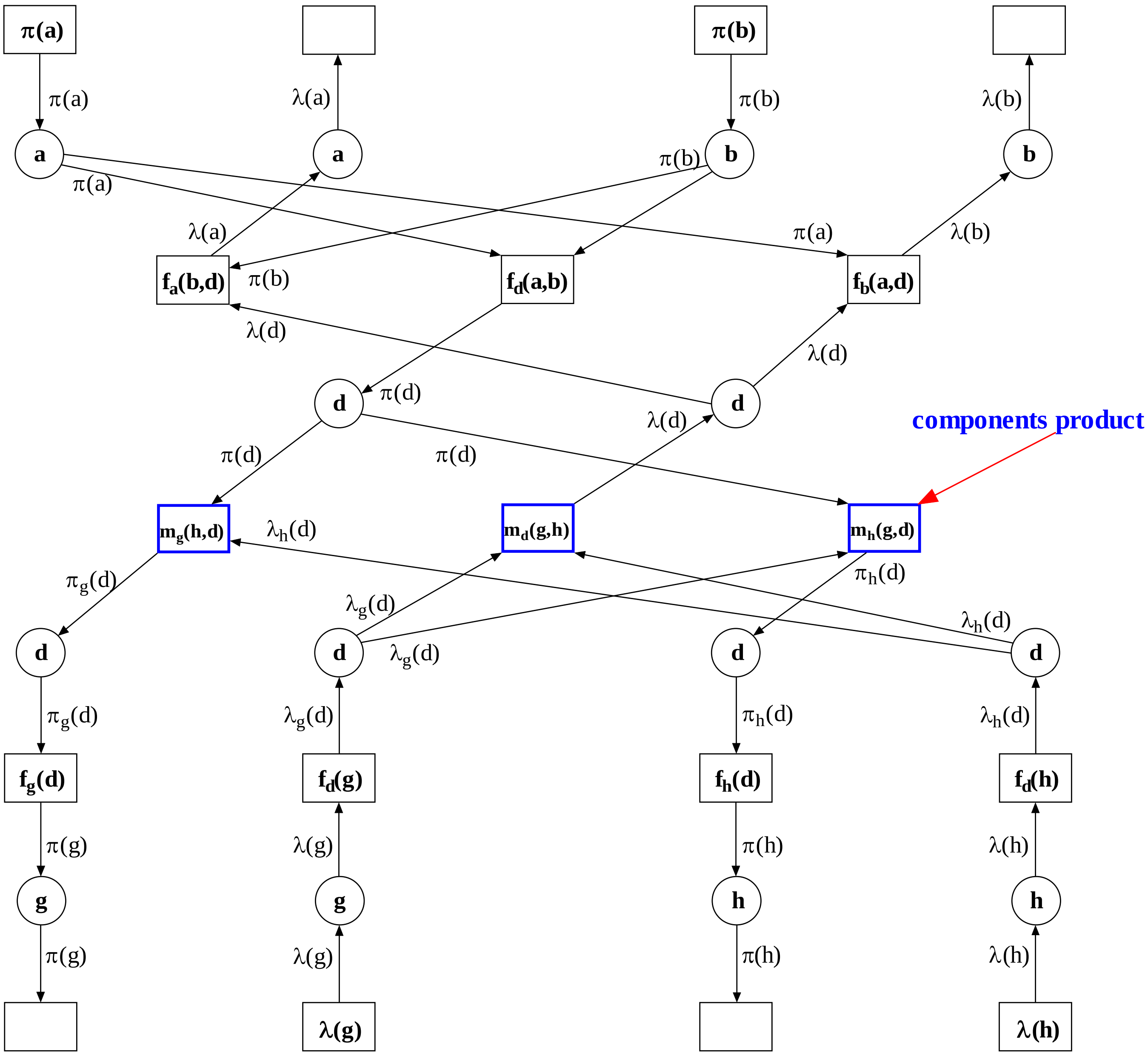

π(a)
π(b)
π(a)
a
π(a)
π(b)
b
π(b)
f_d(a,b)
d
π(d)
d
λ(d)
m_d(g,h)
λ_g(d)
λ_h(d)
d
d
λ_g(d)
λ_h(d)
f_d(g)
f_d(h)
λ(g)
λ(h)
g
h
λ(g)
λ(h)
λ(g)
λ(h)


π(a)
π(a)
a
λ(a)
a
π(b)
π(b)
b
λ(a)
f_a(b,d)
π(b)
λ(d)
d
λ(d)
m_d(g,h)
λ_g(d)
λ_h(d)
d
d
λ_g(d)
λ_h(d)
f_d(g)
f_d(h)
λ(g)
λ(h)
g
h
λ(g)
λ(h)
λ(g)
λ(h)


π(a)
π(b)
π(a)
λ(b)
a
b
b
π(a)
π(a)
λ(b)
f_b(a,d)
λ(d)
d
λ(d)
m_d(g,h)
λ_g(d)
λ_h(d)
d
d
λ_g(d)
λ_h(d)
f_d(g)
f_d(h)
λ(g)
λ(h)
g
h
λ(g)
λ(h)
λ(g)
λ(h)


π(a)
π(b)
π(a)
π(b)
a
b
π(b)
π(a)
f_d(a,b)
d
π(d)
π(d)
m_g(h,d)
λ_h(d)
π_g(d)
d
d
π_g(d)
λ_h(d)
f_g(d)
f_d(h)
π(g)
λ(h)
g
g
h
π(g)
λ(g)
λ(h)
λ(g)
λ(h)


π(a)
π(b)
π(a)
π(b)
a
b
π(b)
π(a)
fd(a,b)
d
π(d)
π(d)
mh(g,d)
πh(d)
d
λg(d)
d
λg(d)
πh(d)
fd(g)
fh(d)
λ(g)
π(h)
g
h
h
λ(g)
π(h)
λ(h)
λ(g)
λ(h)


π(a)
π(b)
a
a
b
b
λ(a)
λ(b)
$f_a(b,d)$
$f_d(a,b)$
$f_b(a,d)$
d
d
π(d)
λ(d)
components product
$m_g(h,d)$
$m_d(g,h)$
$m_h(g,d)$
$\lambda_h(d)$
$\pi_g(d)$
$\lambda_g(d)$
$\pi_h(d)$
$f_g(d)$
$f_d(g)$
$f_h(d)$
$f_d(h)$
π(g)
λ(g)
π(h)
λ(h)
g
g
h
h
λ(g)
λ(h)

Ignorance

We distinguish between two types of ignorance: uncertainty and vagueness

uncertainty reflects a human being's faith or trust in a data source;
the state of data or the data generating process is unknown
or not fully understood,
but we rely on our own or some other person's experience;
"tossing a coin", "playing dice",
random mechanisms in general.

can be captured by probabilistic measures

vagueness is a lack of precision without doubt of the meaning;
"a thick book", "the distance Berlin-Cologne is about 600 km"

cannot be captured by probabilistic measures.

Contents

Mass Distribution

Let Ω be a non-empty set, (usually called the frame of discernment)

every function $m: 2^{\Omega} \rightarrow [0,1]$ is a mass distribution

iff $m(\emptyset) = 0$ and $\sum_{A \subseteq \Omega} m(A) = 1$;

$A \subseteq \Omega$ is a focal element of m iff $m(A) > 0$.

m is a probabilistic basic measure: not additive, not monotonic!

m(A) concerns nothing but A; no subset, no superset,

not the complement, etc.

m("the weather is fine tomorrow") = 0.8 does not mean that

m("the weather is not fine tomorrow") = 0.2 holds.

m is a measure for modeling uncertainties.

Let Ω be a frame of discernment, let m_I and m_B be two mass distributions;

complete ignorance is expressed by the mass distribution $m_I(A) = \begin{cases} 0 & \text{if} \quad A \subset \Omega \\ 1 & \text{if} \quad A = \Omega \end{cases}$

a Bayesian mass distribution m_B is defined by $m_B(A) > 0 \quad \text{iff} \quad |A| = 1$

for $\Omega = \{\omega_1, ..., \omega_n\}$, $P(\omega_i) \coloneqq m_B(\omega_i), 1 \leq i \leq n$,
defines a discrete probability distribution.

Mass Distribution

Let m be a mass distribution on Ω;

$Bel_m : 2^{\Omega} \rightarrow [0,1]$ $\quad Bel_m(A) = \sum_{B \subseteq A} m(B)$ is the belief function based on m;

$Pl_m : 2^{\Omega} \rightarrow [0,1]$ $\quad Pl_m(A) = \sum_{B \cap A \neq \emptyset} m(B)$ is the plausibility function based on m;

$Dbt_m : 2^{\Omega} \rightarrow [0,1]$ $\quad Dbt_m(A) = \sum_{B \cap A = \emptyset} m(B)$ is the doubt function based on m;

$Unc_m : 2^{\Omega} \rightarrow [0,1]$ $\quad Unc_m(A) = \sum_{B \cap A \neq \emptyset \wedge B \not\subset A} m(B)$ is the uncertainty function based on m;

Let m be a mass distribution on Ω; then the following holds

$$
\begin{aligned}
Pl_m(A) &= 1 - Bel_m(\overline{A}) \\
Bel_m(A) &\leq Pl_m(A) \\
Bel_m(A) + Dbt_m(A) + Unc_m(A) &= 1
\end{aligned}
$$

For every Bayesian mass distribution m_B the self-duality $Bel_{m_B} = Pl_{m_B}$ holds

and $P := Bel_{m_B} = Pl_{m_B}$ is a discrete probability distribution.

Compared to mass distributions, probability distributions are full of "syntactic sugar": additivity, monotony, Bayes' theorem, etc.

Question: is this "syntactic sugar" necessary for propagation nets working well or are "mass propagation nets" conceivable?

Contents

Specialization

$$
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}
\end{array}
\cdot
\begin{array}{c|cccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
1 & 1 & & & & & \\
2 & & 1 & & & & \\
3 & & & 1 & & & \\
4 & & & & 1 & & \\
5 & & & & & 1 & \\
6 & & & & & & 1
\end{array}
=
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}
\end{array}
$$

$$
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}
\end{array}
\cdot
\begin{array}{c|cccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
1 & 1 & & & & & \\
2 & & 1 & & & & \\
3 & & & 1 & & & \\
4 & & & & 1 & & \\
5 & & & & & 1 & \\
6 & & & & & & 1
\end{array}
=
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}
\end{array}
$$

$$
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}
\end{array}
\cdot
\begin{array}{c|cccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
1 & & 1 & & & & \\
2 & & 1 & & & & \\
3 & & & & 1 & & \\
4 & & & & 1 & & \\
5 & & & & & & 1 \\
6 & & & & & & 1
\end{array}
=
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3}
\end{array}
$$

$$
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{array}
\cdot
\begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & & 1 & & & & \\ 2 & & 1 & & & & \\ 3 & & & & 1 & & \\ 4 & & & & 1 & & \\ 5 & & & & & & 1 \\ 6 & & & & & & 1 \end{array}
=
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 3 & 6 \\ \hline 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{array}
$$

$$
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{array}
\cdot
\begin{array}{c|ccc} & 2 & 4 & 6 \\ \hline 1 & 1 & & \\ 2 & 1 & & \\ 3 & & 1 & \\ 4 & & 1 & \\ 5 & & & 1 \\ 6 & & & 1 \end{array}
=
\begin{array}{ccc} 2 & 4 & 6 \\ \hline \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}
$$

$$
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{array}
\cdot
\underbrace{\begin{array}{c|ccc} & 2 & 4 & 6 \\ \hline 1 & 1 & & \\ 2 & 1 & & \\ 3 & & 1 & \\ 4 & & 1 & \\ 5 & & & 1 \\ 6 & & & 1 \end{array}}
=
\begin{array}{ccc} 2 & 4 & 6 \\ \hline \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}
$$

The matrix is the conditional probability table for playing (fair) dice in view of the knowledge that an even number had been thrown.

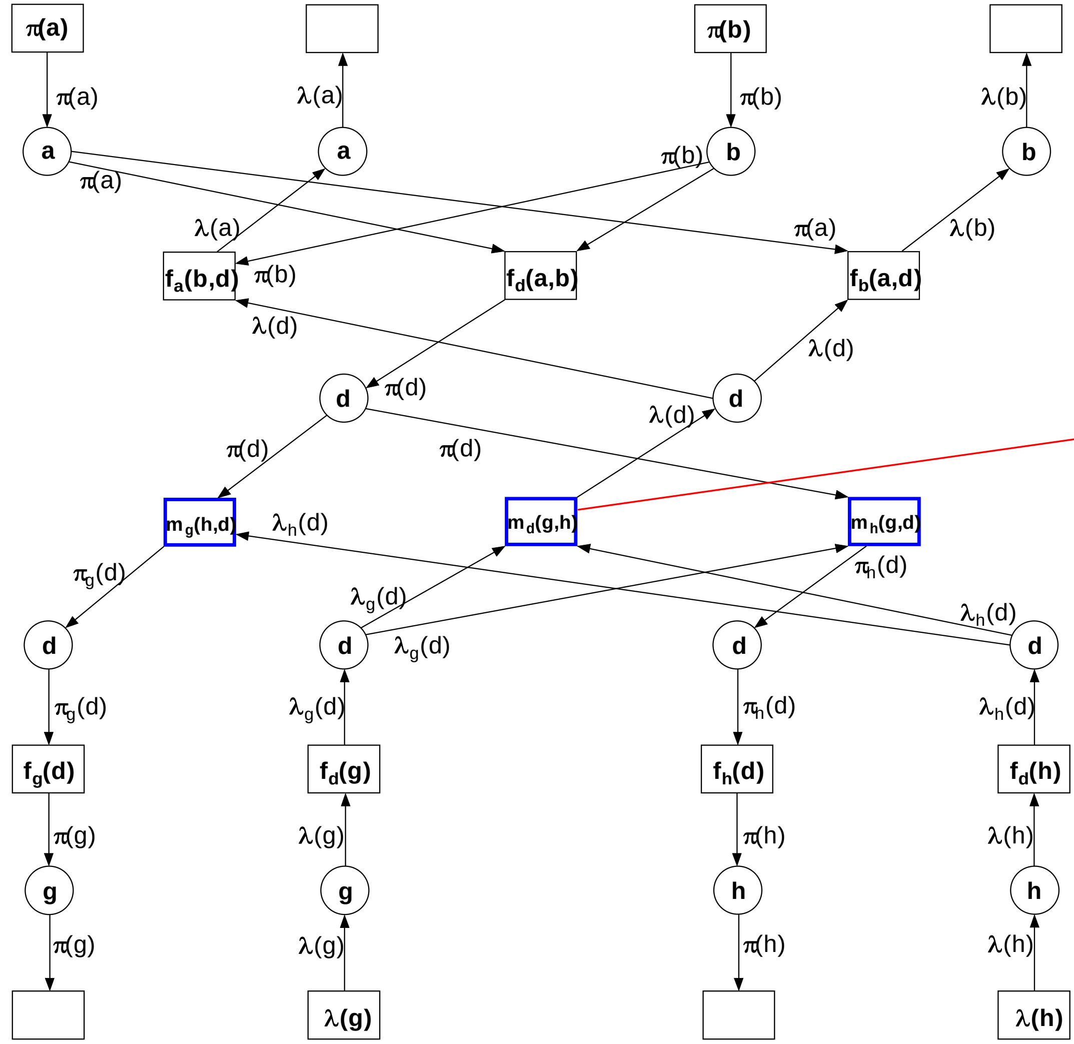
The probabilities of 1, 3, 5 are distributed on 2, 4, 6, respectively; this can be interpreted as an internal flow of probabilities.

Specialization

$$\begin{array}{cc} d & \neg d \\ \hline 0.019 & 0.981 \end{array} \cdot \begin{array}{c|cc} & g & \neg g \\ \hline d & 1 & 0 \\ \neg d & 0.2 & 0.8 \end{array} = \begin{array}{cc} g & \neg g \\ \hline 0.019 & 0.0 \\ +0.1962 & +0.7848 \end{array} = \begin{array}{cc} g & \neg g \\ \hline 0.2152 & 0.7848 \end{array}$$

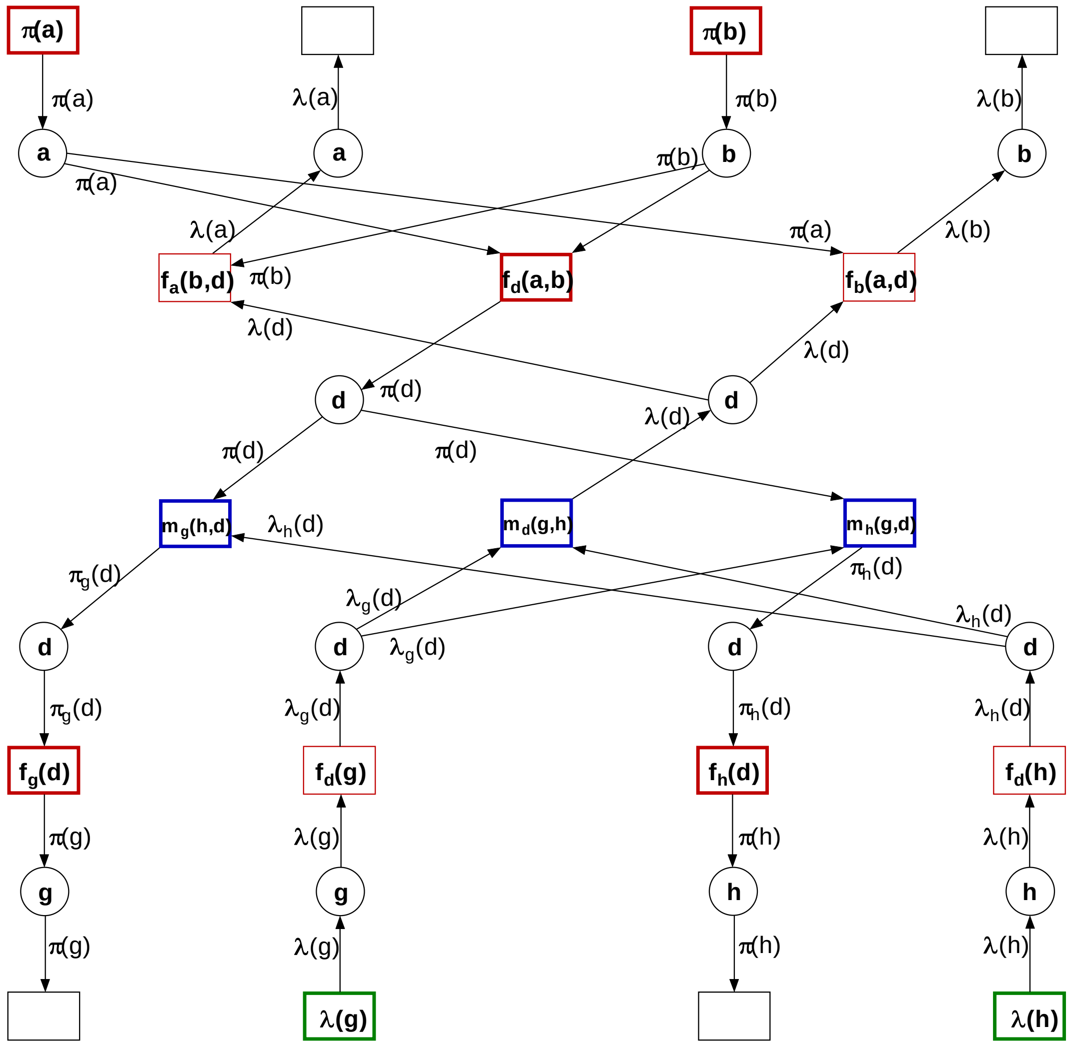
Here the probability $p(\neg d) = 0.981$ is distributed in a ratio of $0.2 : 0.8$ onto g and $\neg g$

For mass distributions, Rudolf Kruse (Magdeburg) et al. call this distribution or flow a specialization and define specialization matrices (comparable to conditional probability tables).



$m_d(g,h): \quad \lambda(d) := \lambda_g(d) \circ \lambda_h(d)$

Specialization



$\pi(a)$

$\pi(b)$

$f_d(a,b)$

$f_g(d)$

$f_h(d)$

$\lambda(g)$

$\lambda(h)$

Probabilities

$\pi(a) = 0.01$ $\qquad \pi(b) = 0.001$

$\pi(\neg a) = 0.99$ $\qquad \pi(\neg b) = 0.999$

$f_d(a,b) =$

		d	¬d
a	b	0.99	0.01
a	¬b	0.9	0.1
¬a	b	0.5	0.5
¬a	¬b	0.01	0.99

$f_g(d) =$

	g	¬g
d	1.0	0.0
¬d	0.2	0.8

$f_h(d) =$

	h	¬h
d	1.0	0.0
¬d	0.2	0.8

Specialization

Probabilities

$\pi(a) = 0.01$

$\pi(\neg a) = 0.99$

$\pi(b) = 0.001$

$\pi(\neg b) = 0.999$

Masses

$\Omega^a = \{a, \neg a\}$

$$\begin{aligned}
\pi(a)[\varnothing^a] &= 0.0 \\
\pi(a)[\{a\}] &= 0.01 \\
\pi(a)[\{\neg a\}] &= 0.99 \\
\pi(a)[\Omega^a] &= 0.0
\end{aligned}$$

$\Omega^b = \{b, \neg b\}$

$$\begin{aligned}
\pi(b)[\varnothing^b] &= 0.0 \\
\pi(b)[\{b\}] &= 0.001 \\
\pi(b)[\{\neg b\}] &= 0.999 \\
\pi(b)[\Omega^b] &= 0.0
\end{aligned}$$

Probabilities

$$f_g(d) = \begin{array}{c|cc} & g & \neg g \\ \hline d & 1.0 & 0.0 \\ \neg d & 0.2 & 0.8 \end{array}$$

$$\Omega^d = \{d, \neg d\}$$

$$f_h(d) = \begin{array}{c|cc} & h & \neg h \\ \hline d & 1.0 & 0.0 \\ \neg d & 0.2 & 0.8 \end{array}$$

$$\Omega^d = \{d, \neg d\}$$

Masses

$$\Omega^g = \{g, \neg g\}$$

$f_g(d)$	$\varnothing^g$	$\{g\}$	$\{\neg g\}$	Ω^g
$\varnothing^d$	0	0	0	0
$\{d\}$	0	1.0	0.0	0
$\{\neg d\}$	0	0.2	0.8	0
Ω^d	0	0	0	1

$$\Omega^h = \{h, \neg h\}$$

$f_h(d)$	$\varnothing^h$	$\{h\}$	$\{\neg h\}$	Ω^h
$\varnothing^d$	0	0	0	0
$\{d\}$	0	0.2	0.8	0
$\{\neg d\}$	0	1.0	0.0	0
Ω^d	0	0	0	1

Specialization

Probabilities

$f_d(a,b) =$

		d	¬d
a	b	0.99	0.01
a	¬b	0.9	0.1
¬a	b	0.5	0.5
¬a	¬b	0.01	0.99

Masses

$f_d(a,b)$	$\varnothing^d$	$\{d\}$	$\{\neg d\}$	Ω^d
$\varnothing^a \times \varnothing^b$	0	0	0	0
$\varnothing^a \times \{b\}$	0	0	0	0
$\varnothing^a \times \{\neg b\}$	0	0	0	0
$\varnothing^a \times \Omega^b$	0	0	0	0
$\{a\} \times \varnothing^b$	0	0	0	0
$\{a\} \times \{b\}$	0	0.99	0.01	0
$\{a\} \times \{\neg b\}$	0	0.9	0.1	0
$\{a\} \times \Omega^b$	0	0	0	0
$\{\neg a\} \times \varnothing^b$	0	0	0	0
$\{\neg a\} \times \{b\}$	0	0.5	0.5	0
$\{\neg a\} \times \{\neg b\}$	0	0.01	0.99	0
$\{\neg a\} \times \Omega^b$	0	0	0	0
$\Omega^a \times \varnothing^b$	0	0	0	0
$\Omega^a \times \{b\}$	0	0	0	0
$\Omega^a \times \{\neg b\}$	0	0	0	0
$\Omega^a \times \Omega^b$	0	0	0	0

With these larger - but sparse - vectors and matrices the propagation nets work for mass distributions in the same way as for probabilities.

The specialization approach is superior to Dempster's rule of combination which from case to case yields absurd results.

Contents

Incidence Calculi

An incidence calculus is a quintuple $< W, \mu, P, A, i >$ where

W is a finite set of possible worlds (primitives , not defined in detail)

for all $w \in W$: $\mu(w)$ is the probability of w, with $\mu(W) = 1$ and

$$\text{for all } I \subseteq W: \quad \mu(I) = \sum_{w \in I} \mu(w)$$

P is a set of atomic propositions, L(P) is the set of wwfs over P

$A \subseteq L(P)$ is the set of axioms

$i: A \rightarrow 2^W$ is the incidence function, $i(\phi)$ is the set of possible worlds in which ϕ is true

i must satisfy the following (truth functionality):

$$i(\text{true}) = W, \qquad i(\text{false}) = \emptyset,$$

$$i(\neg\phi) = W \setminus i(\phi), \qquad i(\phi \wedge \varphi) = i(\phi) \cap i(\varphi), \qquad i(\phi) \cap i(\neg\varphi) = \emptyset,$$

$$i(\phi \vee \varphi) = i(\phi) \cup i(\varphi), \qquad i(\phi \rightarrow \varphi) = i(\neg\phi) \cup i(\varphi), \qquad i(\phi) \cup i(\neg\phi) = W$$

Incidence Calculi

A general incidence calculus is a quintuple $< W, \mu, P, A, i >$ where

W is a finite set of possible worlds (primitives , not defined in detail)

for all $w \in W$: $\mu(w)$ is the probability of w, with $\mu(W) = 1$ and

for all $I \subseteq W$: $\mu(I) = \sum_{w \in I} \mu(w)$

P is a set of atomic propositions, L(P) is the set of wwfs over P

$A \subseteq L(P)$ is the set of axioms

$i: A \rightarrow 2^W$ is the incidence function, $i(\phi)$ is the set of possible worlds in which ϕ is true

i must satisfy the following (no truth functionality):

$$i(true) = W, \qquad i(false) = \emptyset,$$
$$i(\phi \wedge \varphi) = i(\phi) \cap i(\varphi), \qquad i(\phi) \cap i(\neg\varphi) = \emptyset,$$
$$i(\phi \vee \varphi) \supseteq i(\phi) \cup i(\varphi), \qquad i(\phi) \cup i(\neg\phi) \subseteq W$$

Example: A meeting will be held next week on a day which is preferred by most of the 10 delegates.

Delegates d_1 to d_4 prefer Monday
d_5 prefers Monday or Tuesday
d_6 to d_{10} prefer Tuesday

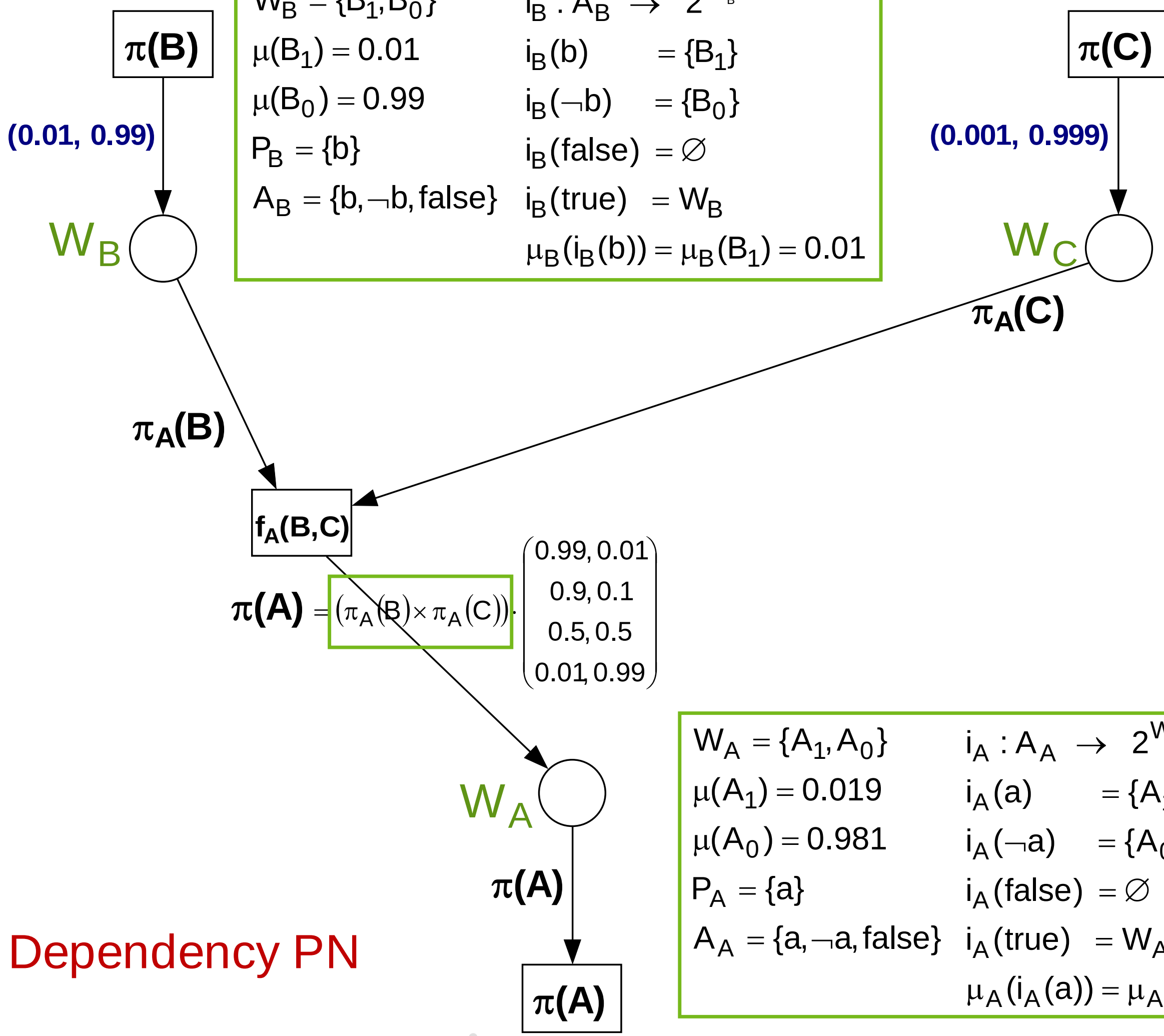
q_1 stands for "The meeting is held on Monday"
q_2 stands for "The meeting is held on Tuesday"

$$i(q_1) = \{d_1, d_2, d_3, d_4\}, \quad i(q_1) \cup i(q_2) = \{d_1, d_2, d_3, d_4, d_6, d_7, d_8, d_9, d_{10}\},$$

$$i(q_2) = \{d_6, d_7, d_8, d_9, d_{10}\}, \quad i(q_1 \vee q_2) = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_4, d_{10}\}$$

$$\boxed{i(q_1) \cup i(q_2) \subseteq i(q_1 \vee q_2)}$$

Incidence Calculi



$$\pi(A) = (\pi_A(B) \times \pi_A(C)) \cdot \begin{pmatrix} 0.99, 0.01 \\ 0.9, 0.1 \\ 0.5, 0.5 \\ 0.01, 0.99 \end{pmatrix}$$

$$W_B = \{B_1, B_0\} \qquad i_B : A_B \to 2^{W_B}$$
$$\mu(B_1) = 0.01 \qquad i_B(b) = \{B_1\}$$
$$\mu(B_0) = 0.99 \qquad i_B(\neg b) = \{B_0\}$$
$$P_B = \{b\} \qquad i_B(\text{false}) = \emptyset$$
$$A_B = \{b, \neg b, \text{false}\} \qquad i_B(\text{true}) = W_B$$
$$\mu_B(i_B(b)) = \mu_B(B_1) = 0.01$$

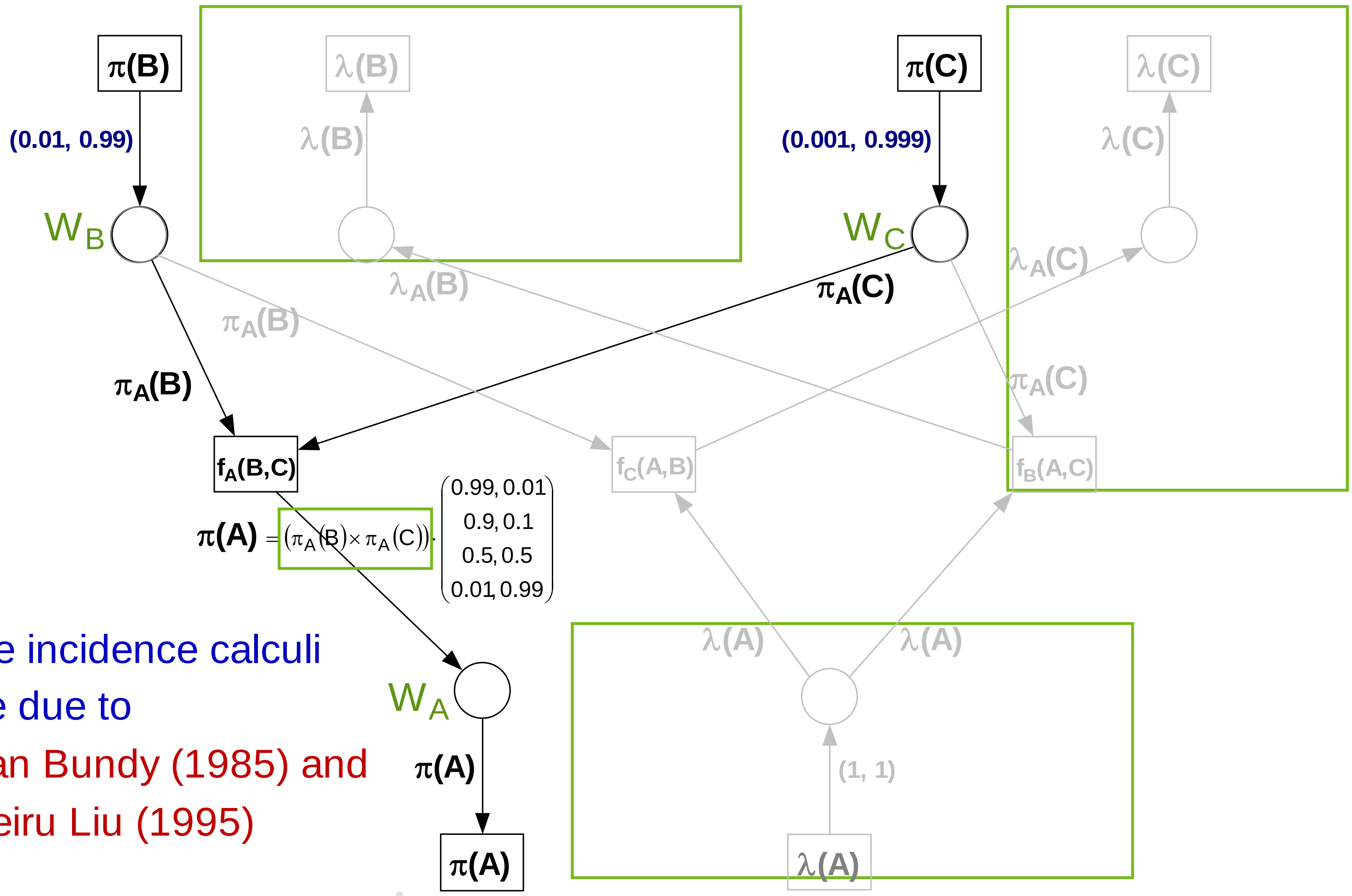
$$W_C = \{C_1, C_0\}$$
$$\mu(C_1) = 0.001$$
$$\mu(C_0) = 0.999$$
$$P_C = \{c\}$$
$$A_C = \{c, \neg c, \text{false}\}$$

$$i_C : A_C \to 2^{W_C}$$
$$i_C(c) = \{C_1\}$$
$$i_C(\neg c) = \{C_0\}$$
$$i_C(\text{false}) = \emptyset$$
$$i_C(\text{true}) = W_C$$
$$\mu_C(i_C(c)) = \mu_C(C_1) = 0.001$$

$$W_A = \{A_1, A_0\} \qquad i_A : A_A \to 2^{W_A}$$
$$\mu(A_1) = 0.019 \qquad i_A(a) = \{A_1\}$$
$$\mu(A_0) = 0.981 \qquad i_A(\neg a) = \{A_0\}$$
$$P_A = \{a\} \qquad i_A(\text{false}) = \emptyset$$
$$A_A = \{a, \neg a, \text{false}\} \qquad i_A(\text{true}) = W_A$$
$$\mu_A(i_A(a)) = \mu_A(A_1) = 0.019$$

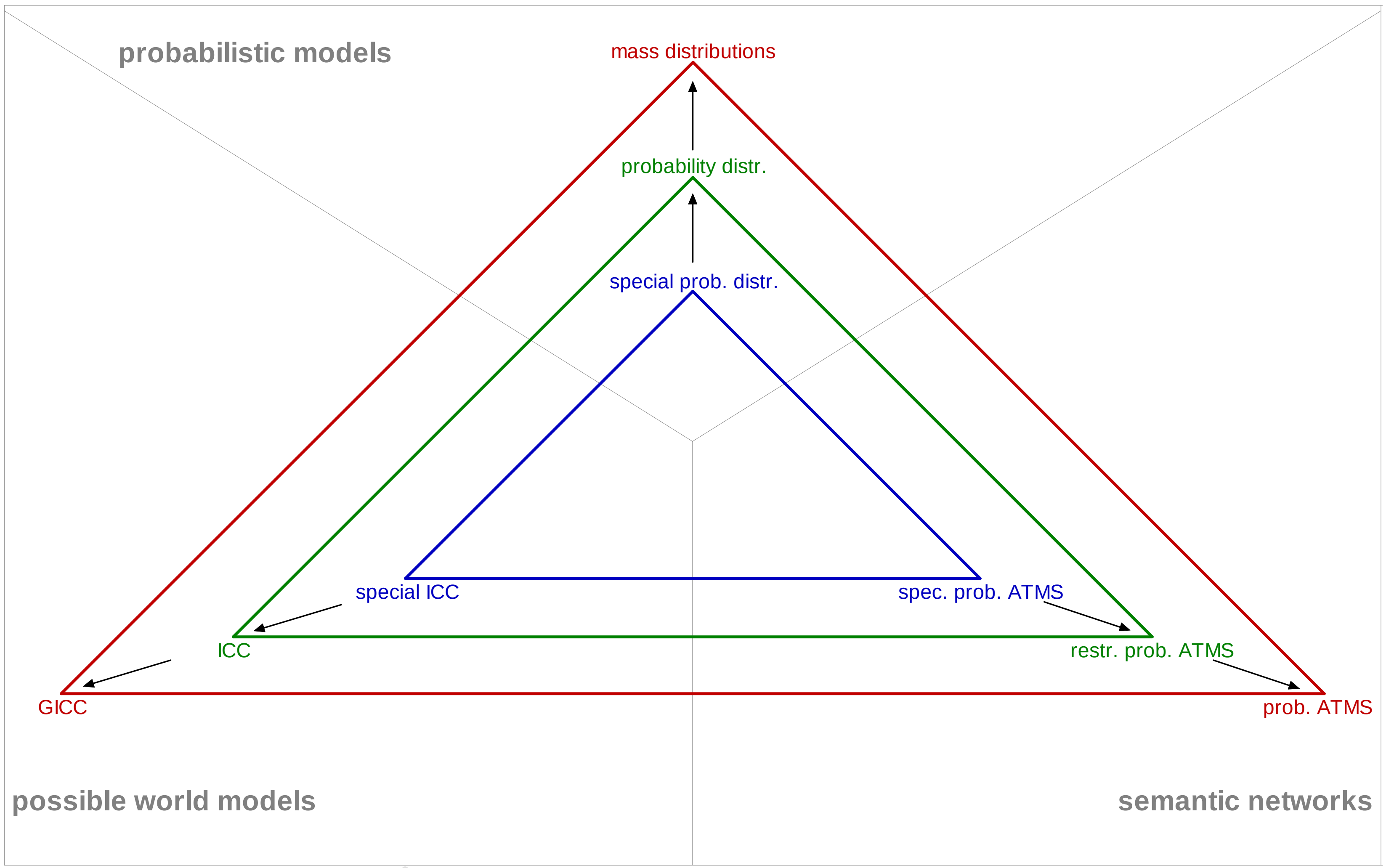
Dependency PN

Incidence Calculi



The incidence calculi are due to Alan Bundy (1985) and Weiru Liu (1995)

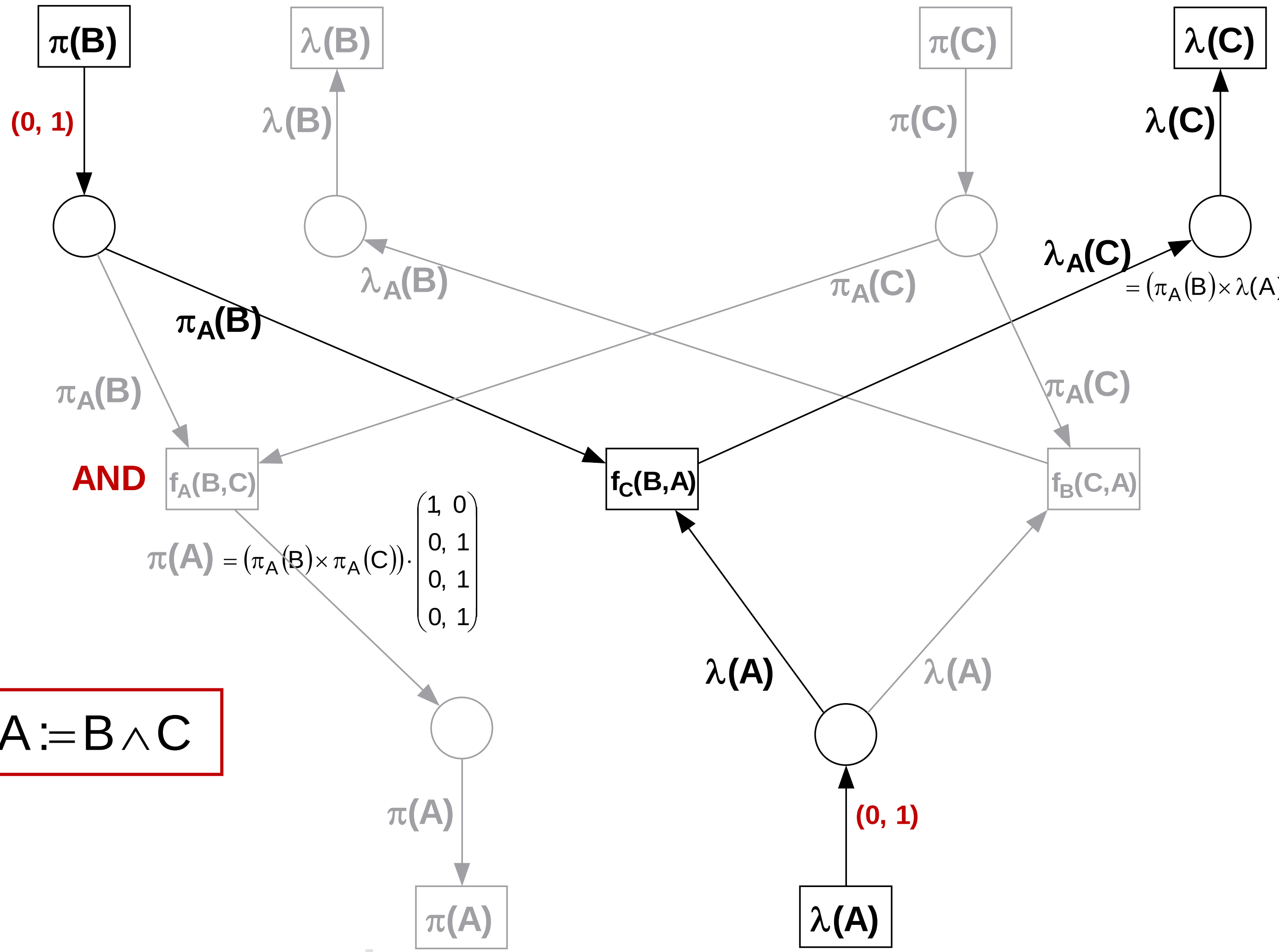
Belief Revision



Why propagation nets work well in evidential reasoning

Contents

Logical Propagation Net

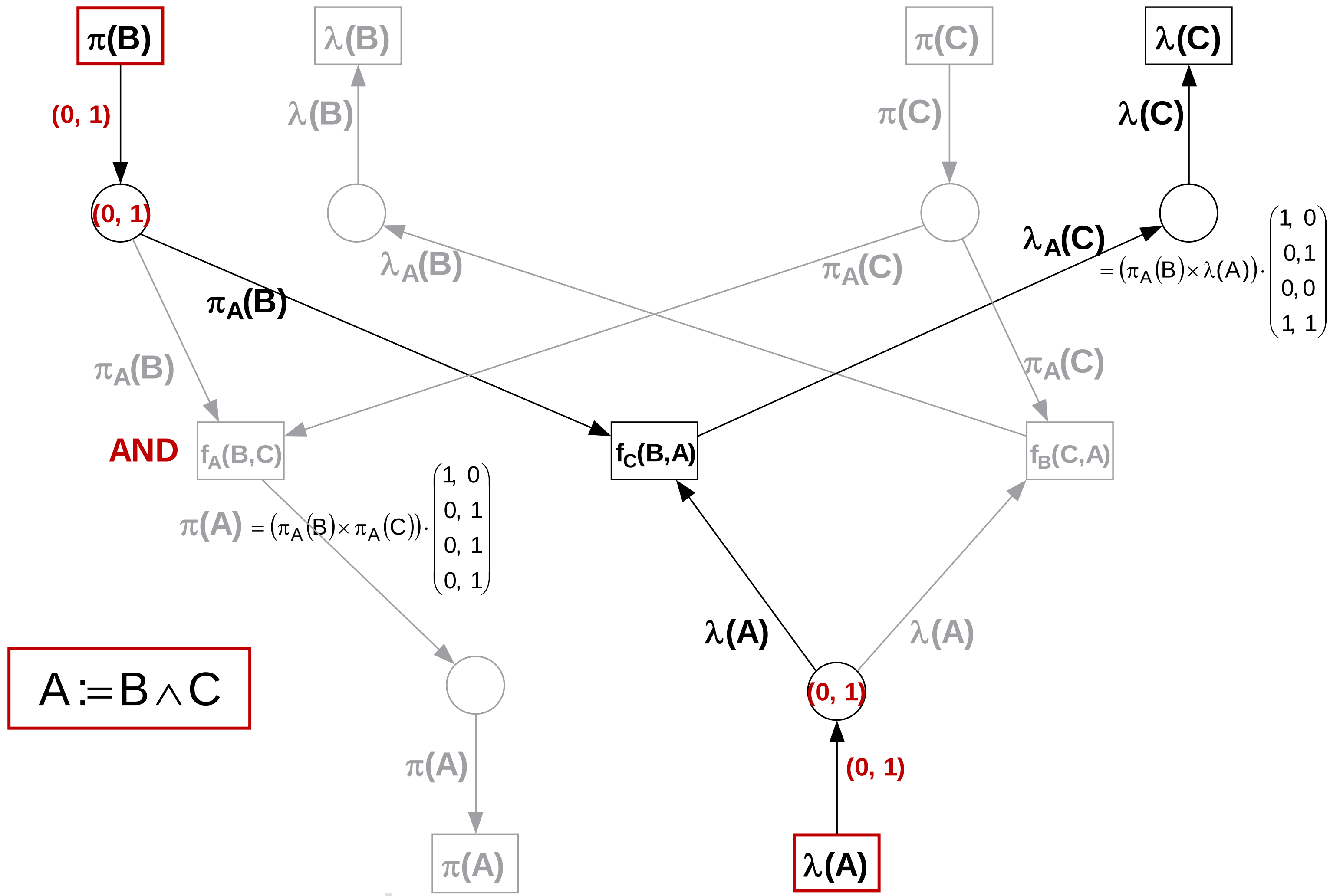


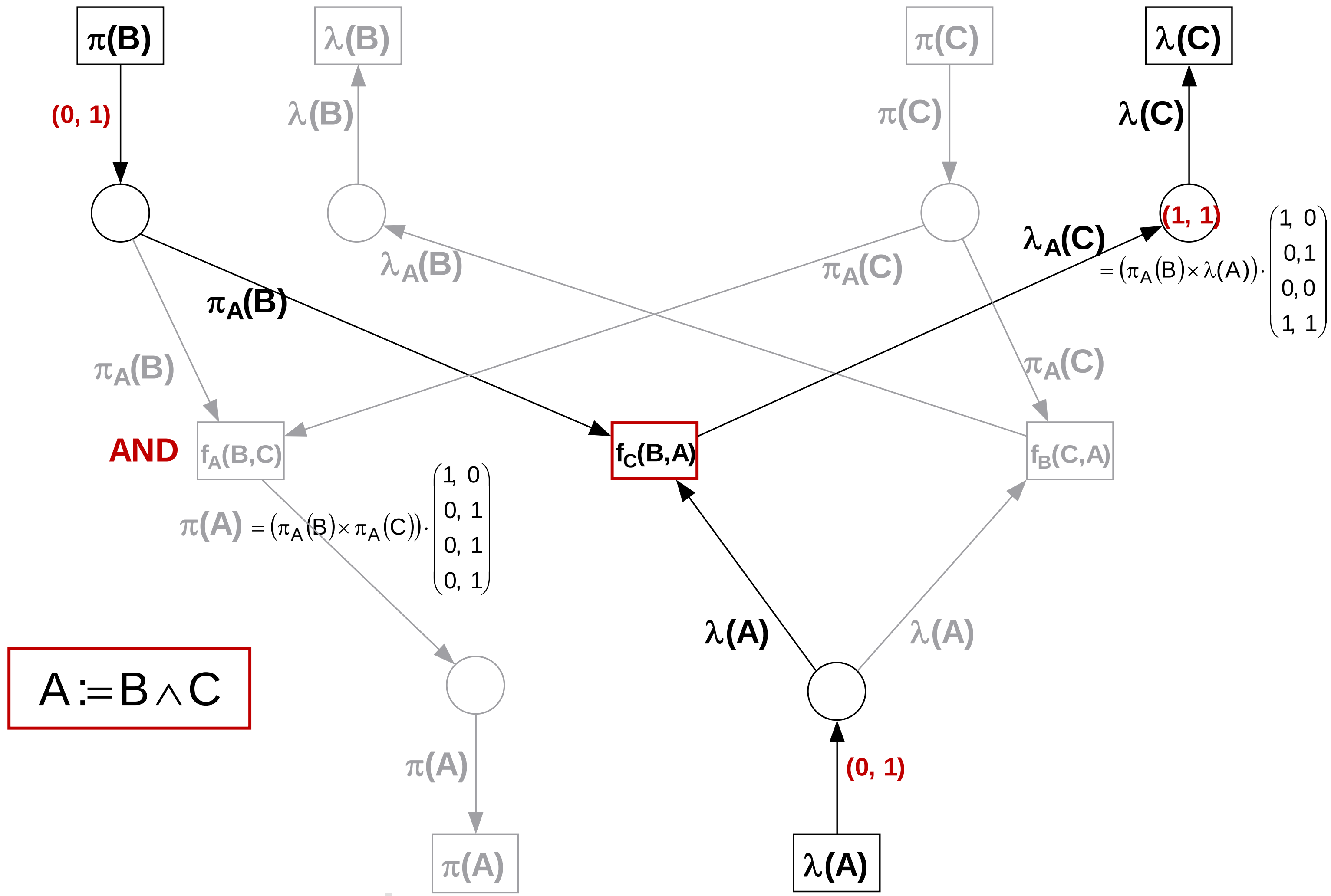
$$\lambda_A(C) = \left(\pi_A(B) \times \lambda(A)\right) \cdot \begin{pmatrix} 1, & 0 \\ 0, & 1 \\ 0, & 0 \\ 1, & 1 \end{pmatrix}$$

$$\pi(A) = \left(\pi_A(B) \times \pi_A(C)\right) \cdot \begin{pmatrix} 1, & 0 \\ 0, & 1 \\ 0, & 1 \\ 0, & 1 \end{pmatrix}$$

$$A := B \wedge C$$

Logical Propagation Net





π(B)
λ(B)
π(C)
λ(C)
(0, 1)
λ(B)
π(C)
λ(C)
(1, 1)
λA(C)
λA(B)
πA(C)
πA(B)
= (πA(B) × λ(A)) · (1, 0; 0, 1; 0, 0; 1, 1)
πA(B)
πA(C)
AND
fA(B,C)
fC(B,A)
fB(C,A)
π(A) = (πA(B) × πA(C)) · (1, 0; 0, 1; 0, 1; 0, 1)
λ(A)
λ(A)
A := B ∧ C
π(A)
(0, 1)
π(A)
λ(A)

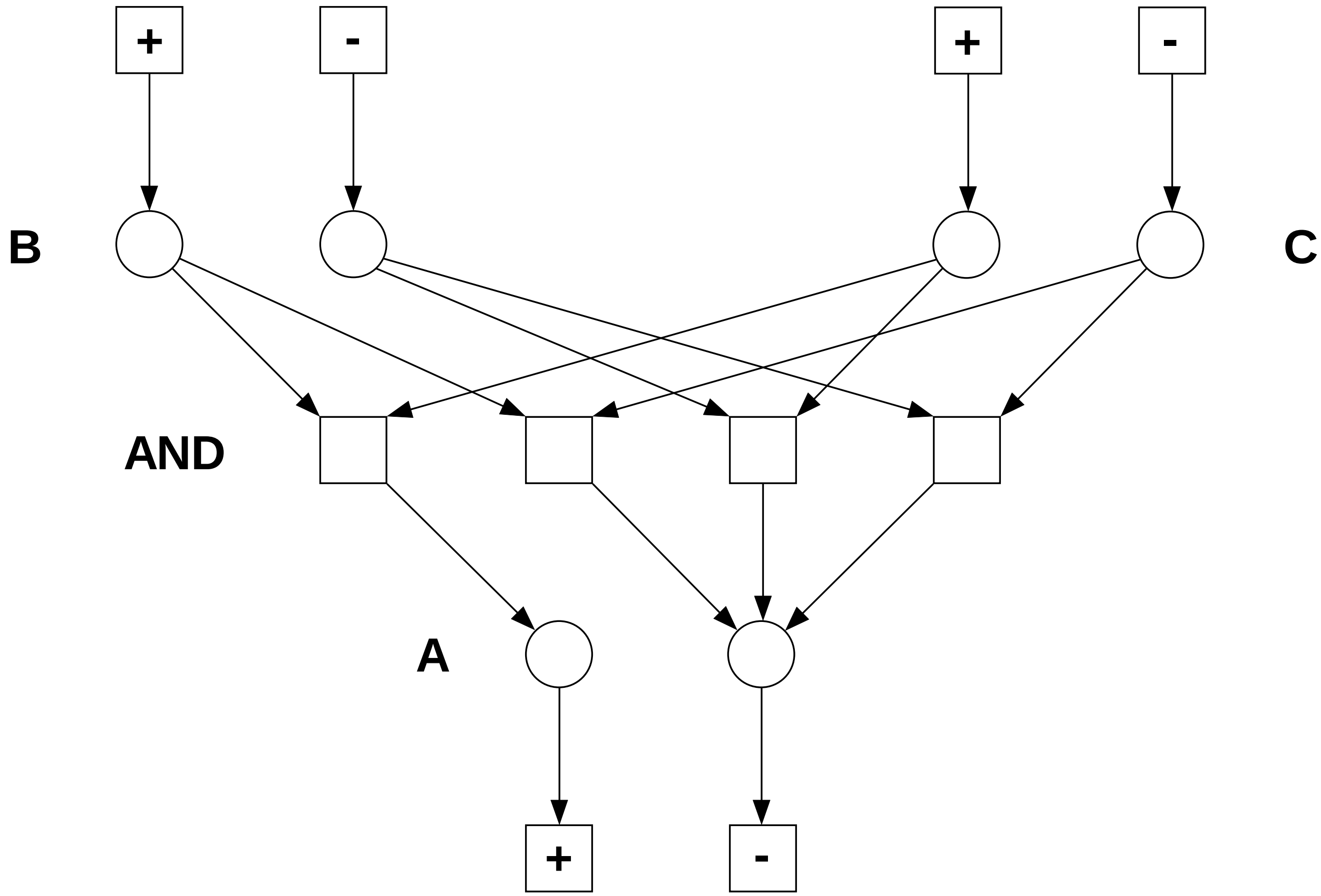
Logical Propagation Net

B	A	C
+(1, 0)	(0, 0) !	(0, 0) !
+(1, 0)	(0, 1) -	(0, 1) -
+(1, 0)	(1, 0) +	(1, 0) +
+(1, 0)	(1, 1) ?	(1, 1) ?

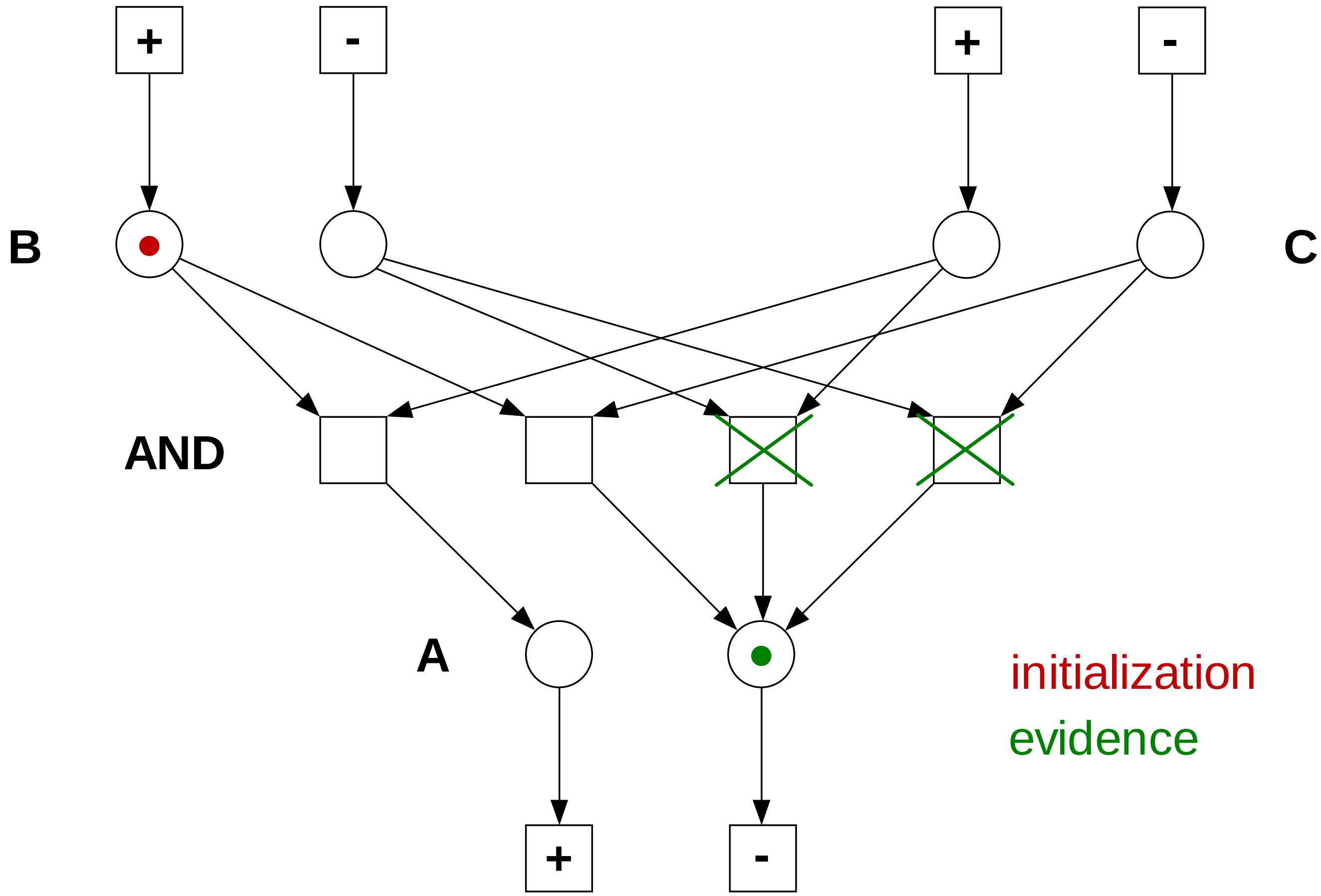
B	A	C
- (0, 1)	(0, 0) !	(0, 0) !
- (0, 1)	(0, 1) -	(1, 1) ?
- (0, 1)	(1, 0) +	(0, 0) !
- (0, 1)	(1, 1) ?	(1, 1) ?

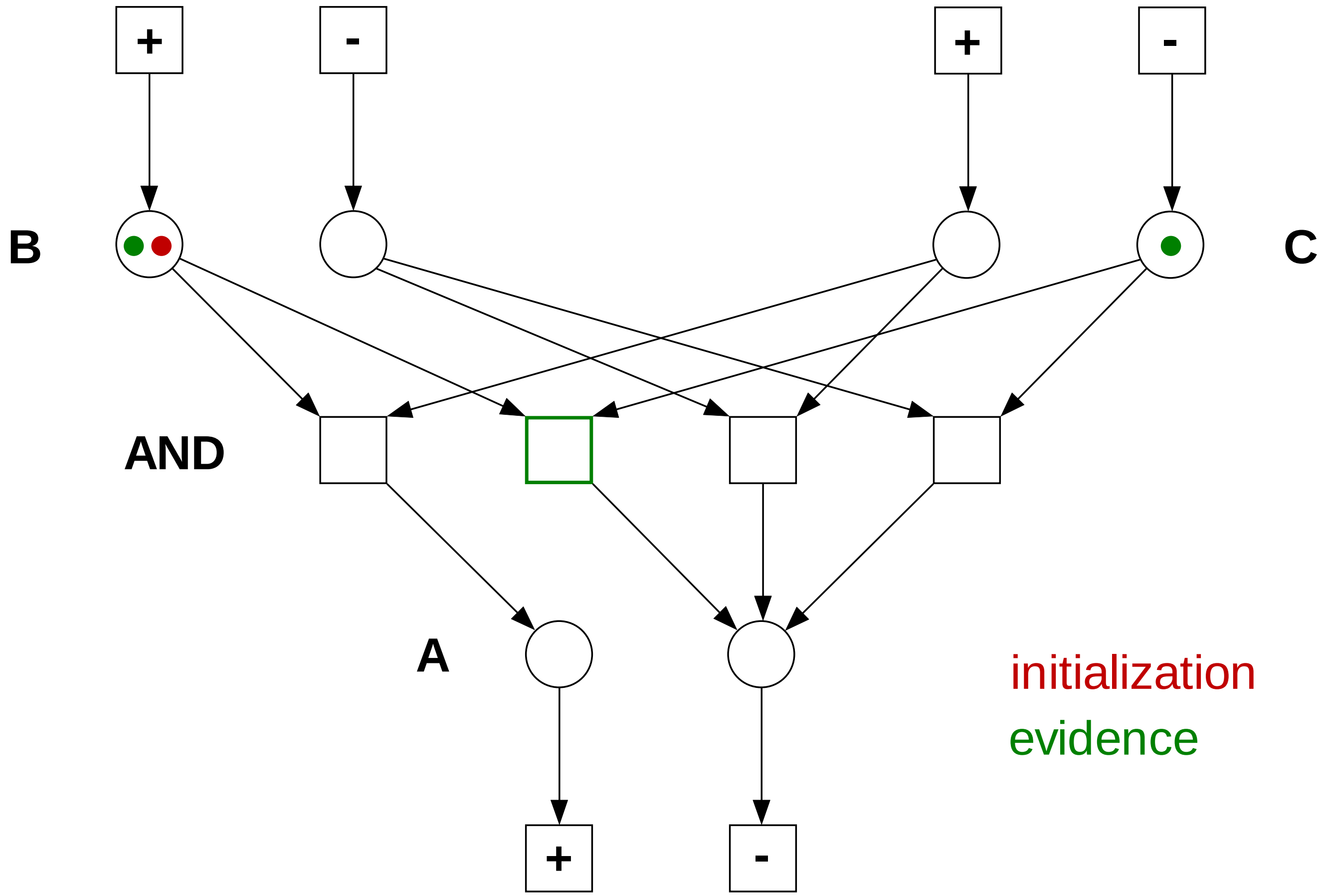
!	contradiction
-	negated
+	non-negated
?	no information

Logical Propagation Net

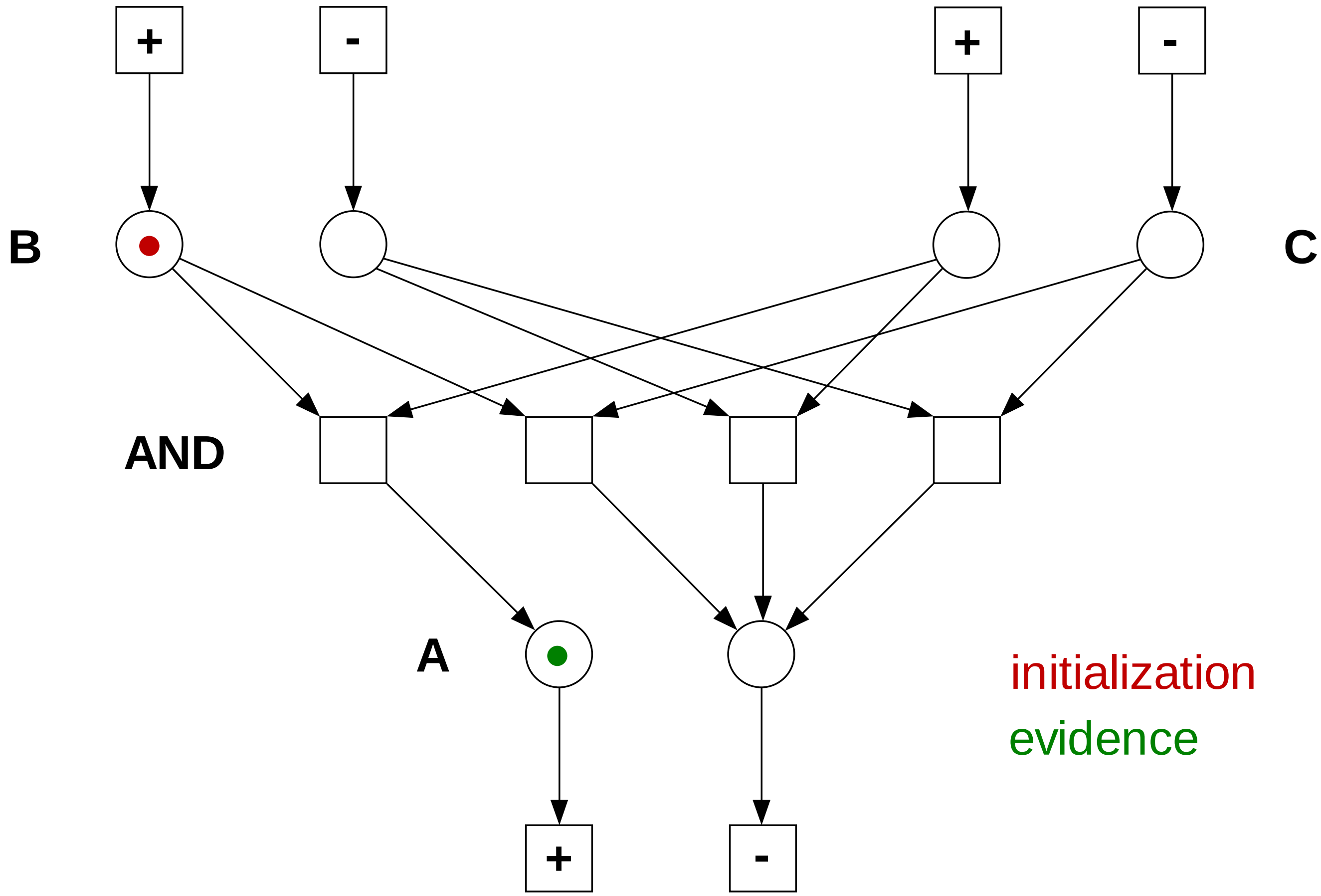


Logical Propagation Net

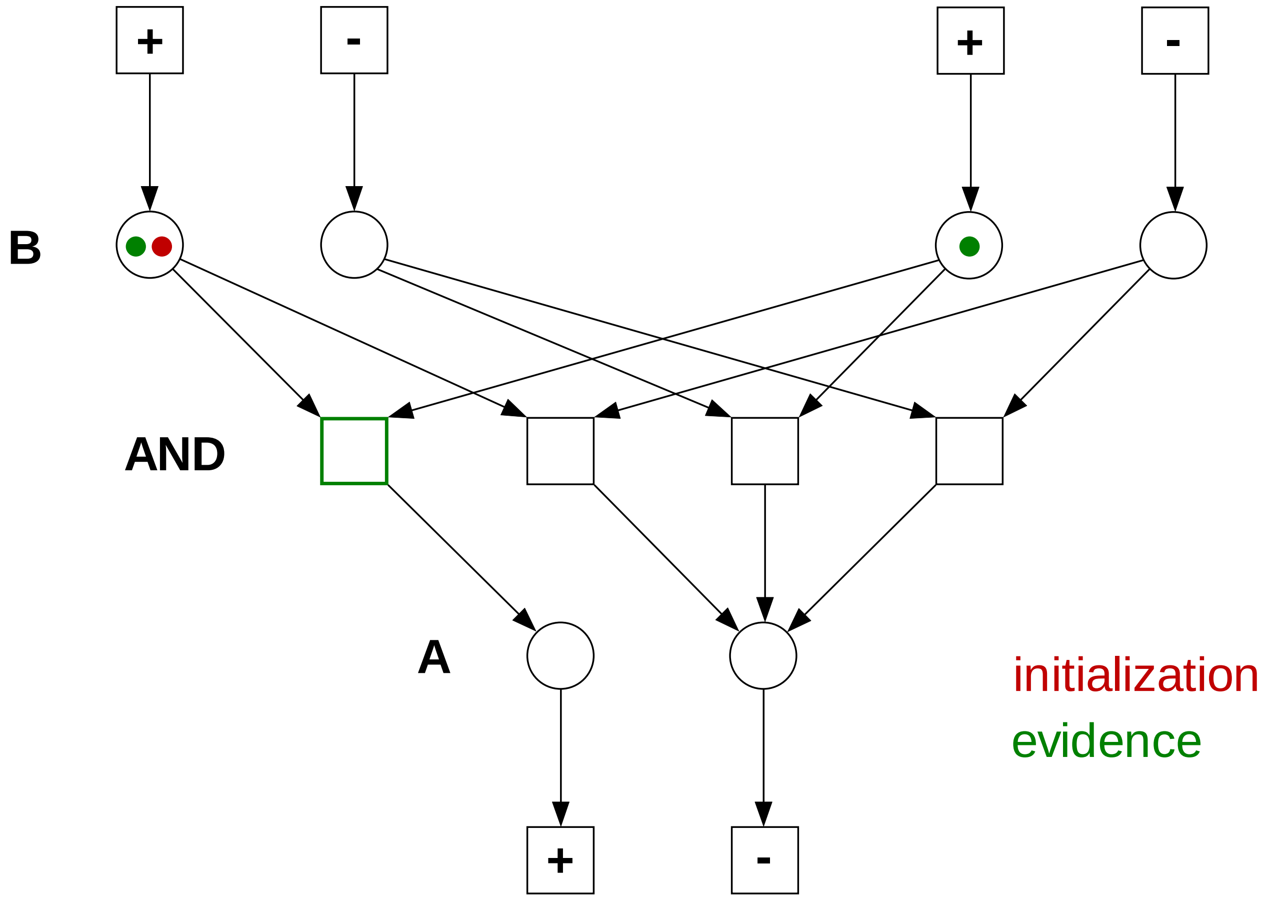



+
-
+
-
B
C
AND
A
+
-
initialization
evidence

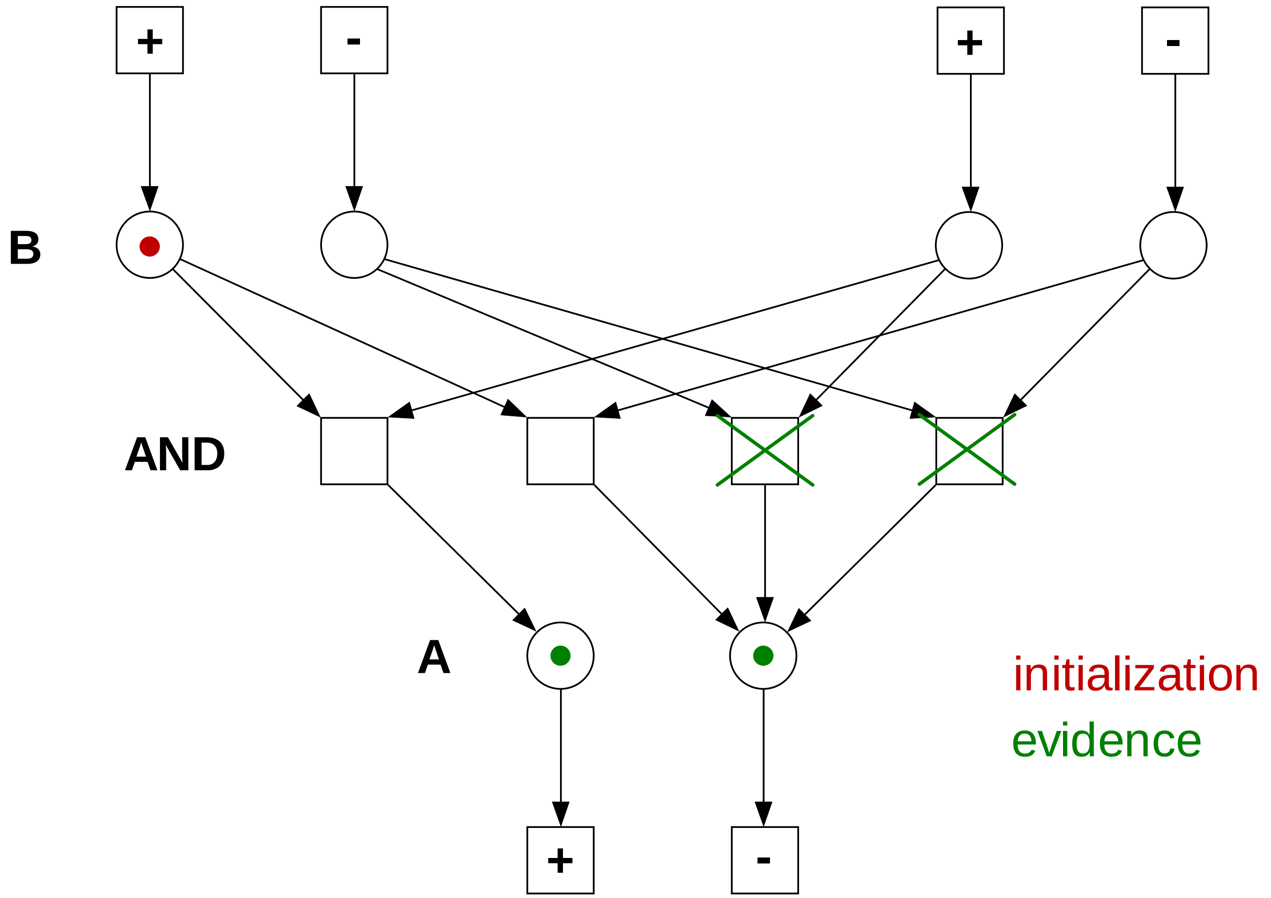
Logical Propagation Net



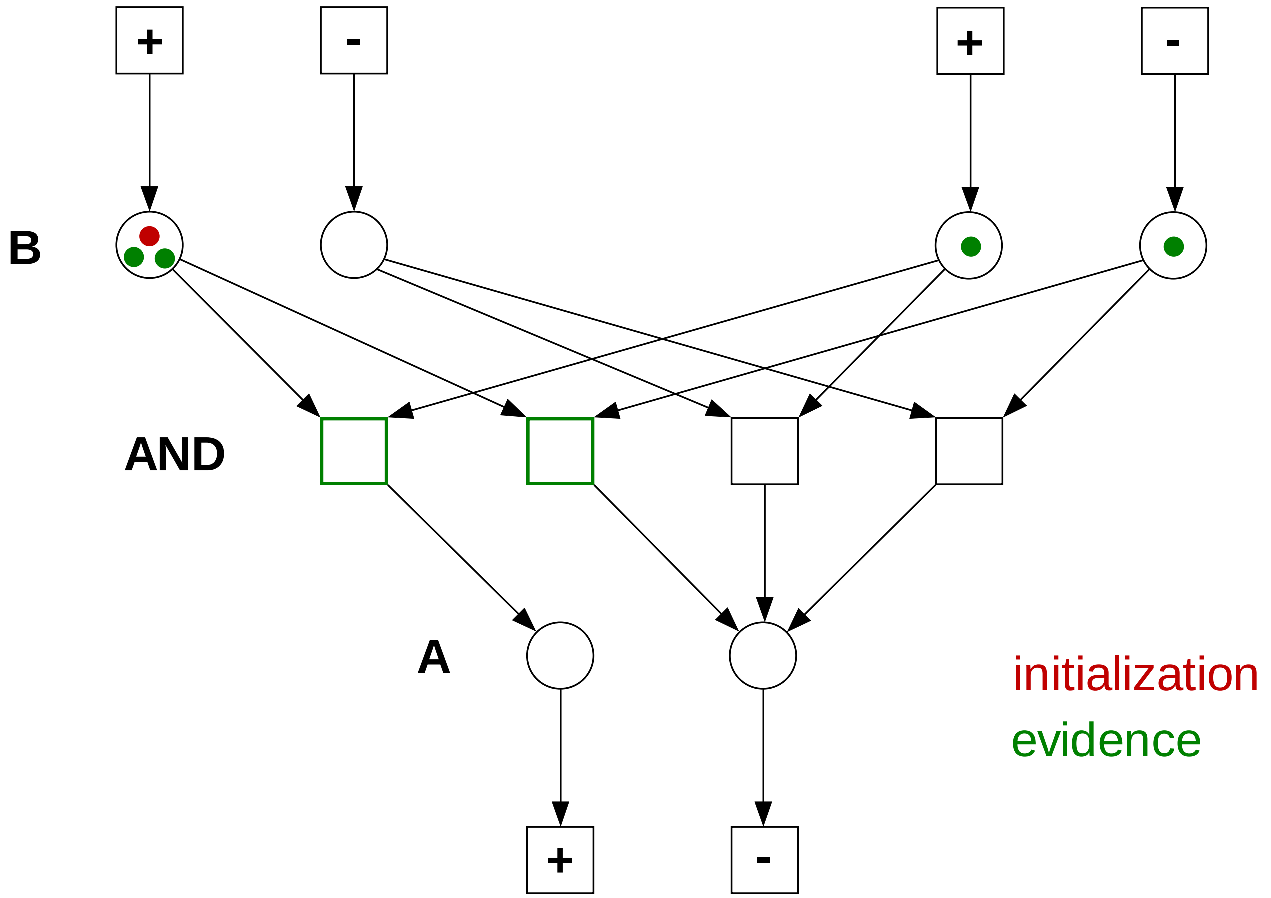
Logical Propagation Net



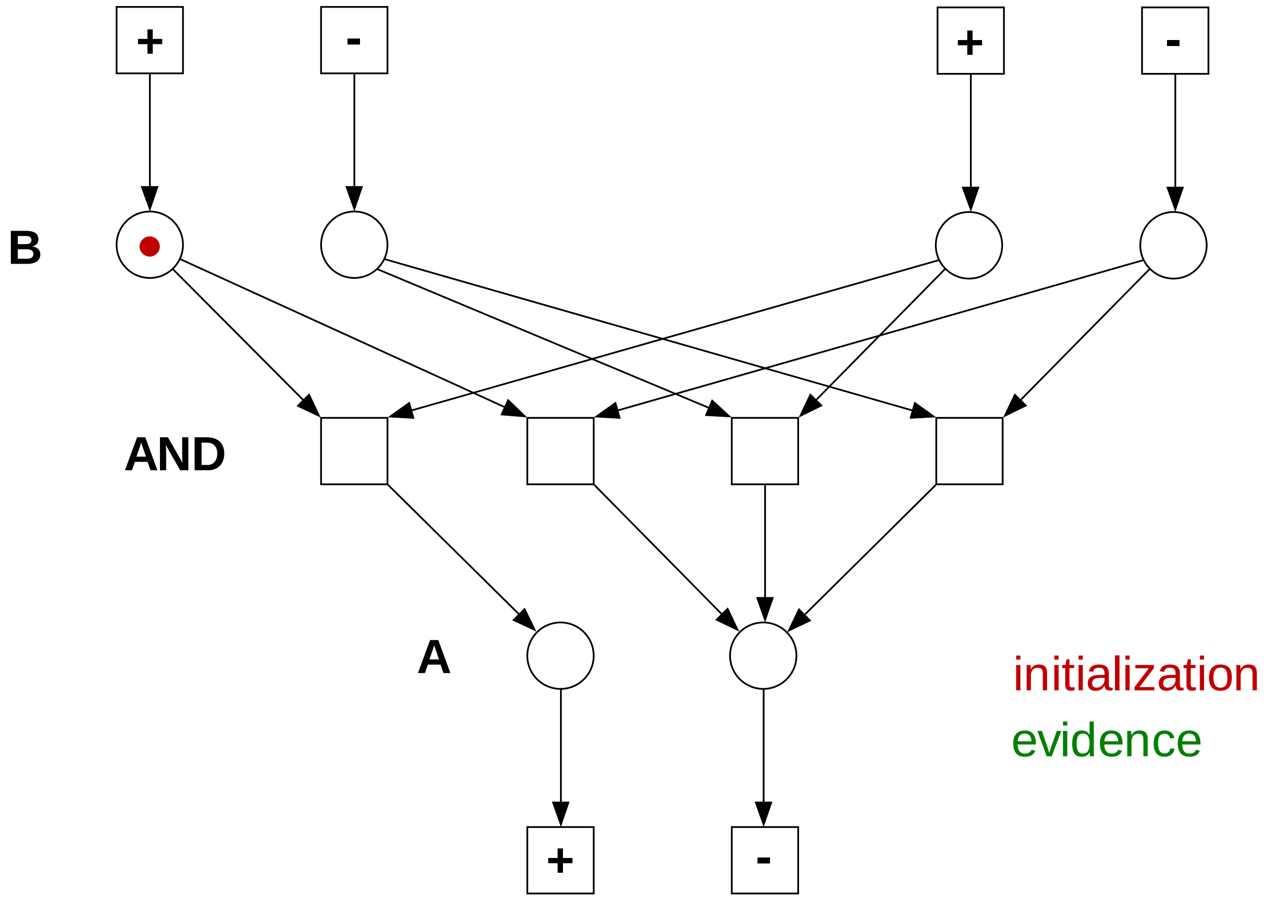
Logical Propagation Net

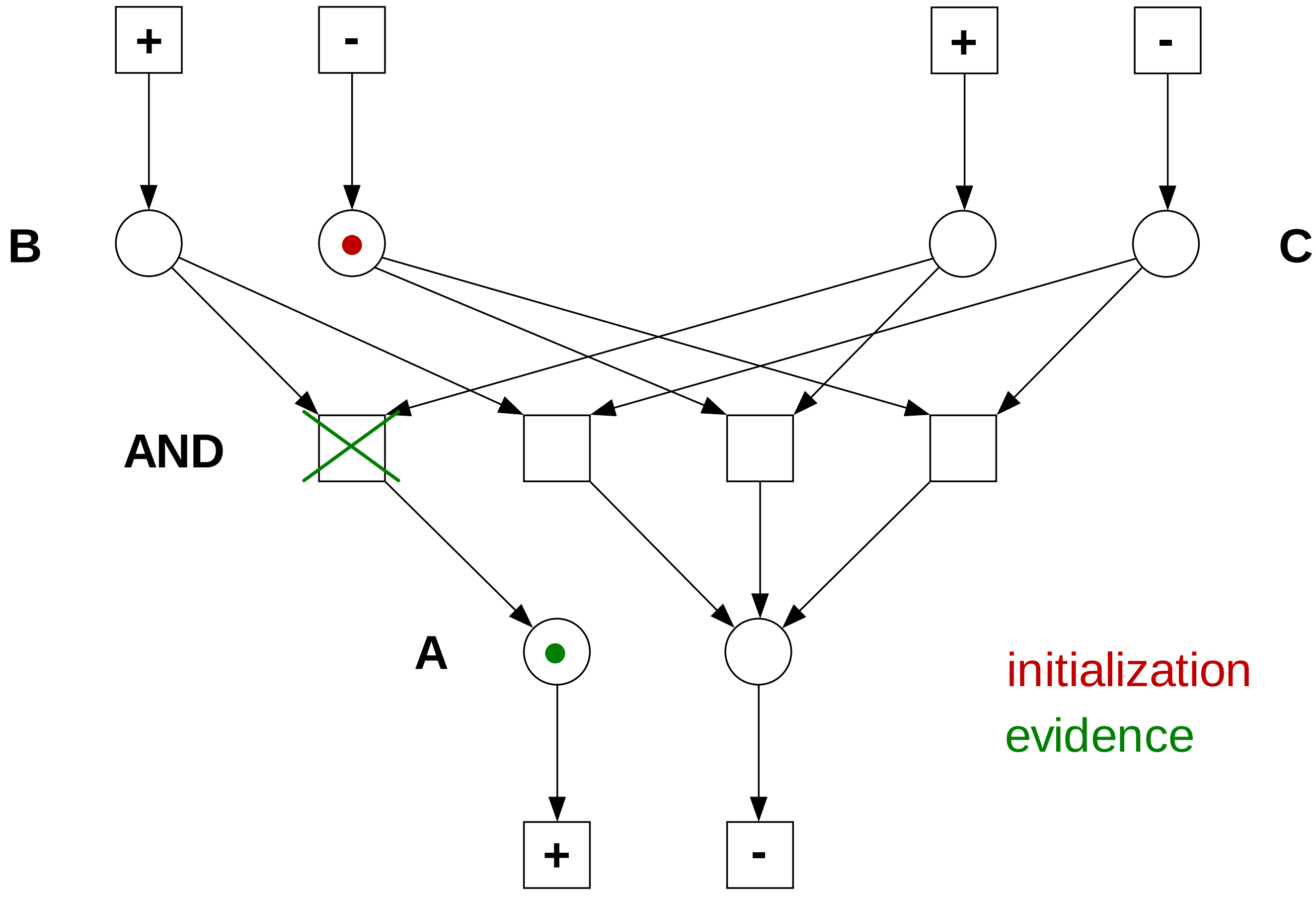


Logical Propagation Net



Logical Propagation Net




+
-
+
-
B
C
AND
A
+
-
initialization
evidence

Definition:

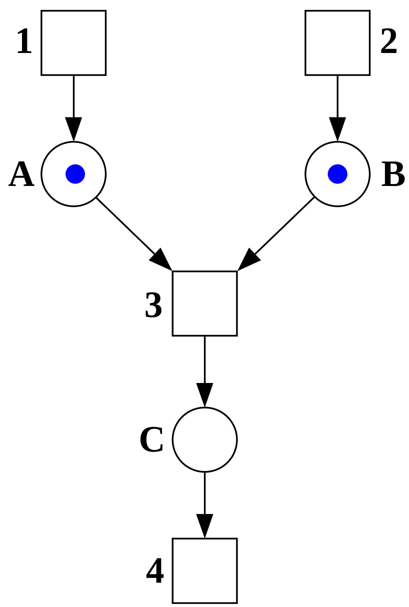
Let N = (S, T, F) be a p/t-net and M a marking of N;
the dual net $N_d = (S_d, T_d, F_d)$ is defined by $[N_d] := [N]^t$.

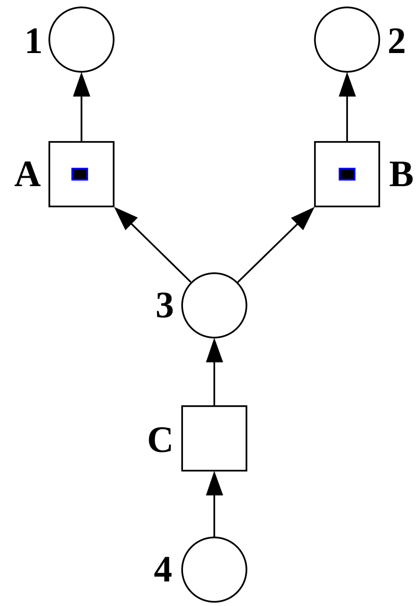
So, for N_d $S_d = T, T_d = S, F_d = F^{-1}$ holds;

$M_d := M$, i.e. $\forall x \in T_d = S: M_d(x) = M(x)$

Now, the tokens are located on transitions.

Duality of Marked Nets

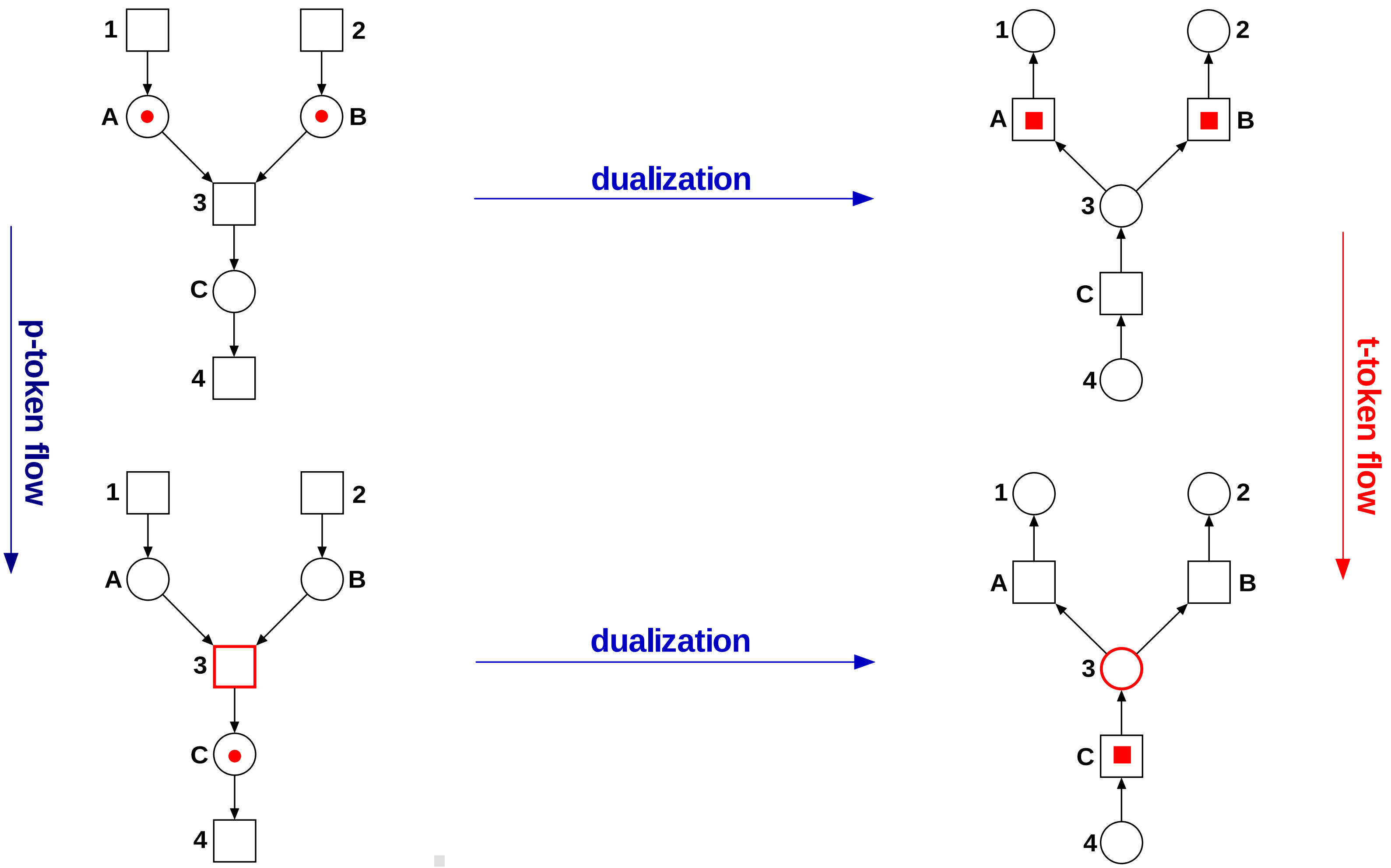




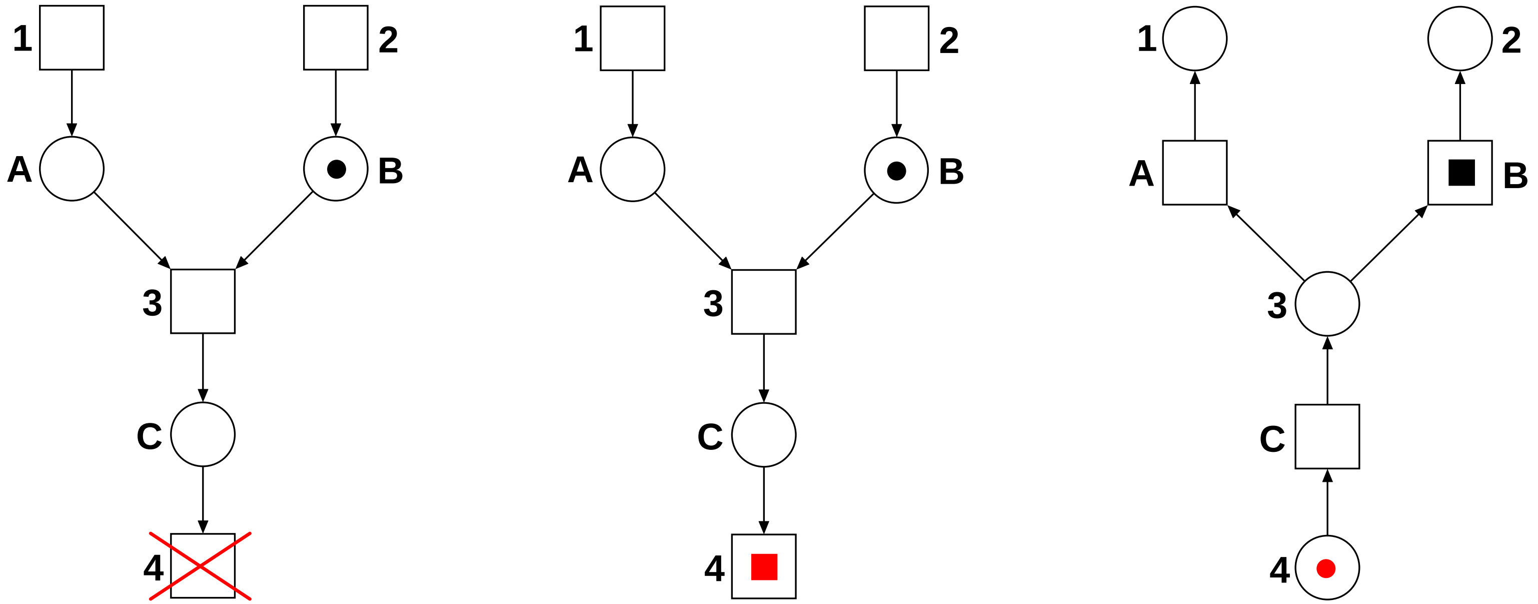
[N]	1	2	3	4	M
A	1		-1		1
B		1	-1		1
C			1	-1	

$[N_d]$	A	B	C
1	1		
2		1	
3	-1	-1	1
4			-1
M_d^t	1	1	

Duality of Marked Nets



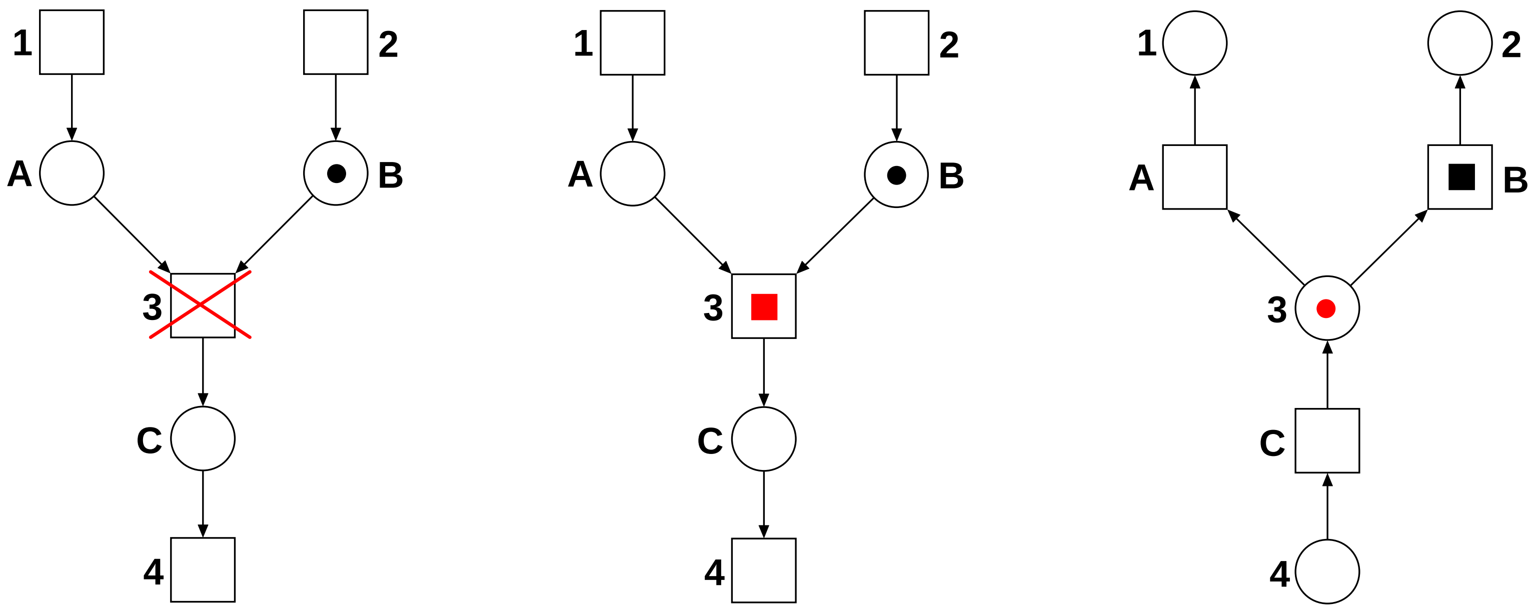
Duality of Marked Nets



if transition 4 did not fire then transition 3 did not

if transition 4 must not fire then transition 3 must not

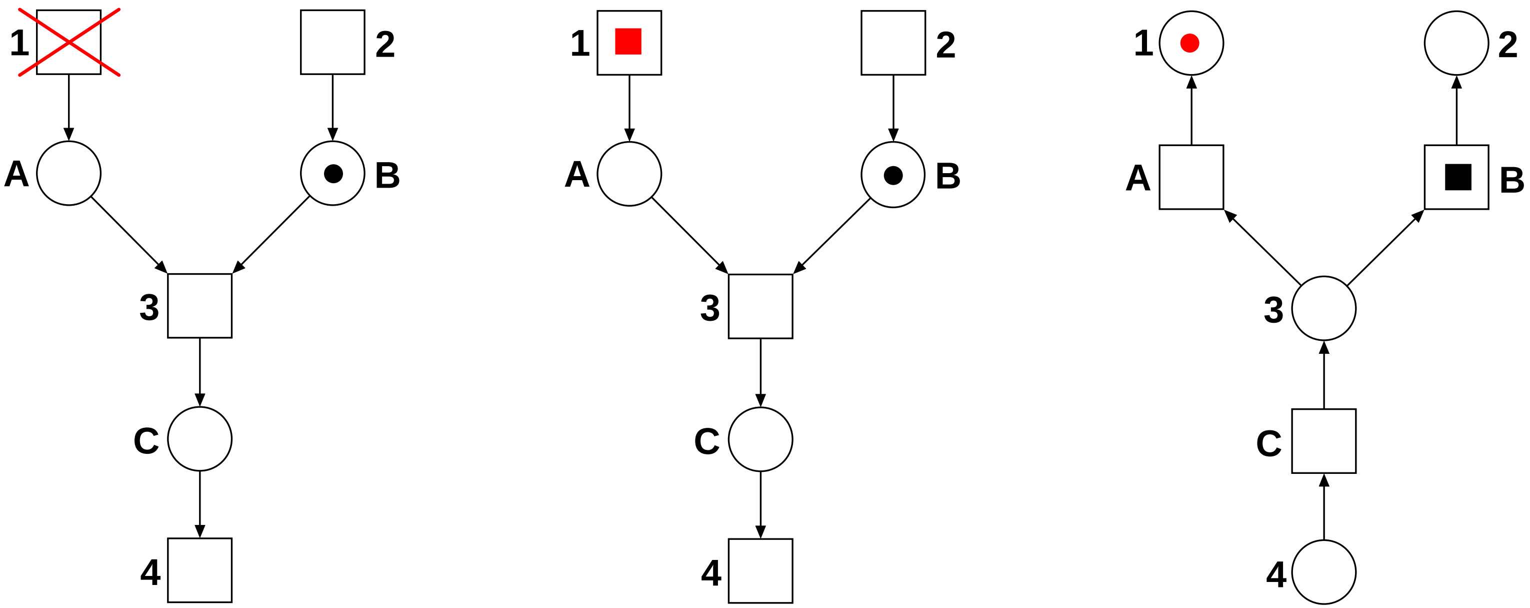
Duality of Marked Nets



if transition 3 did not fire then transition 1 did not

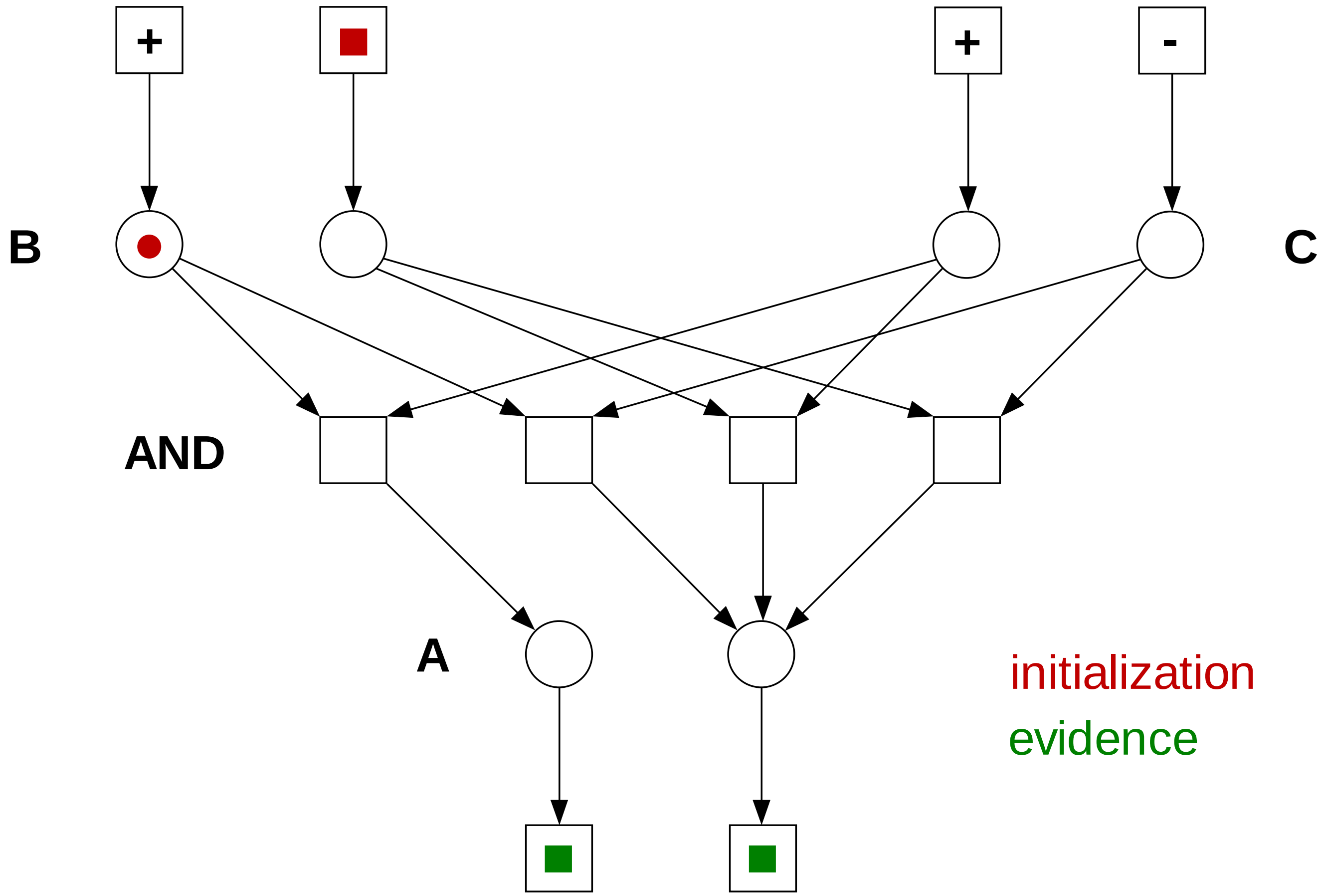
if transition 3 must not fire then transition 1 must not

Duality of Marked Nets



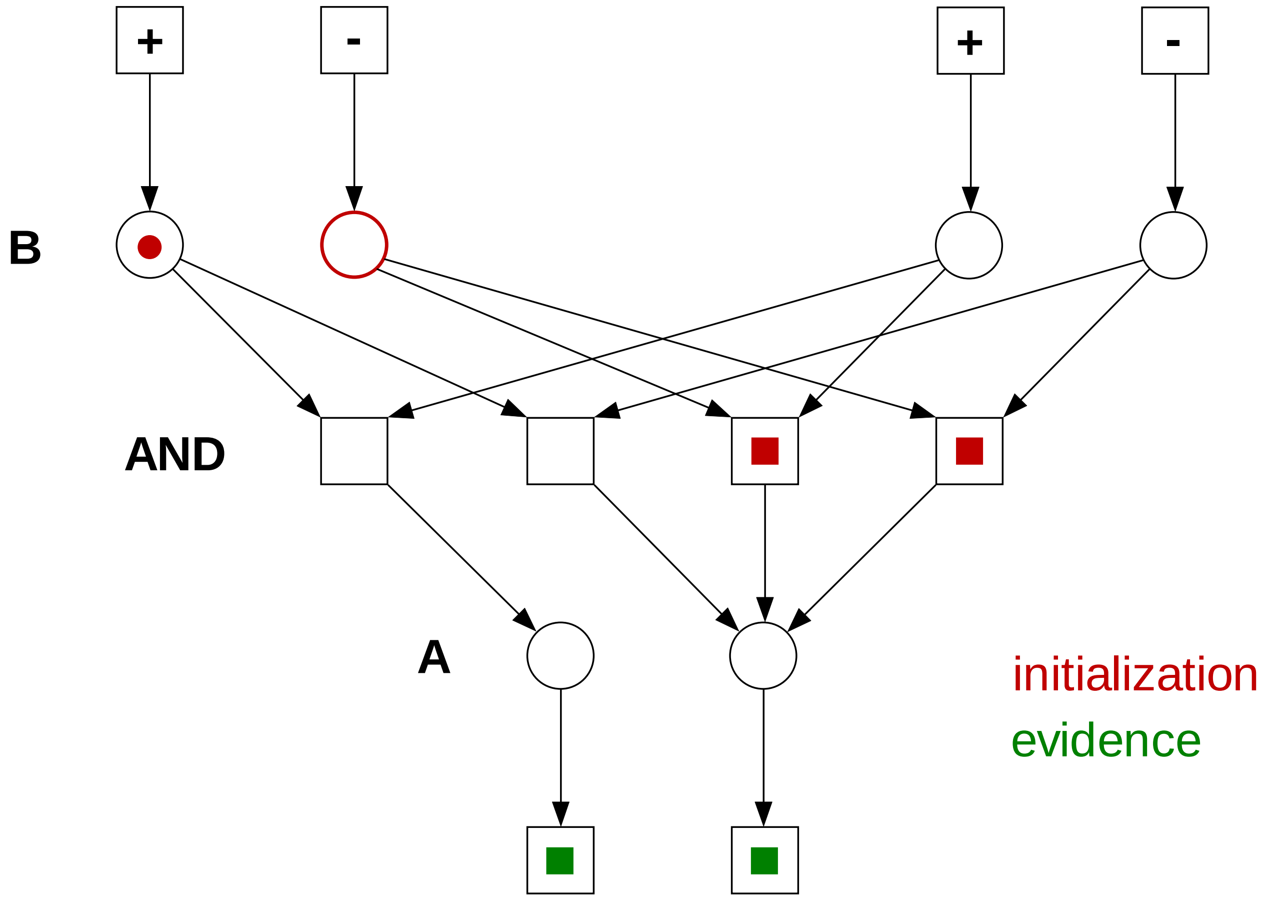
transition 1 did not fire

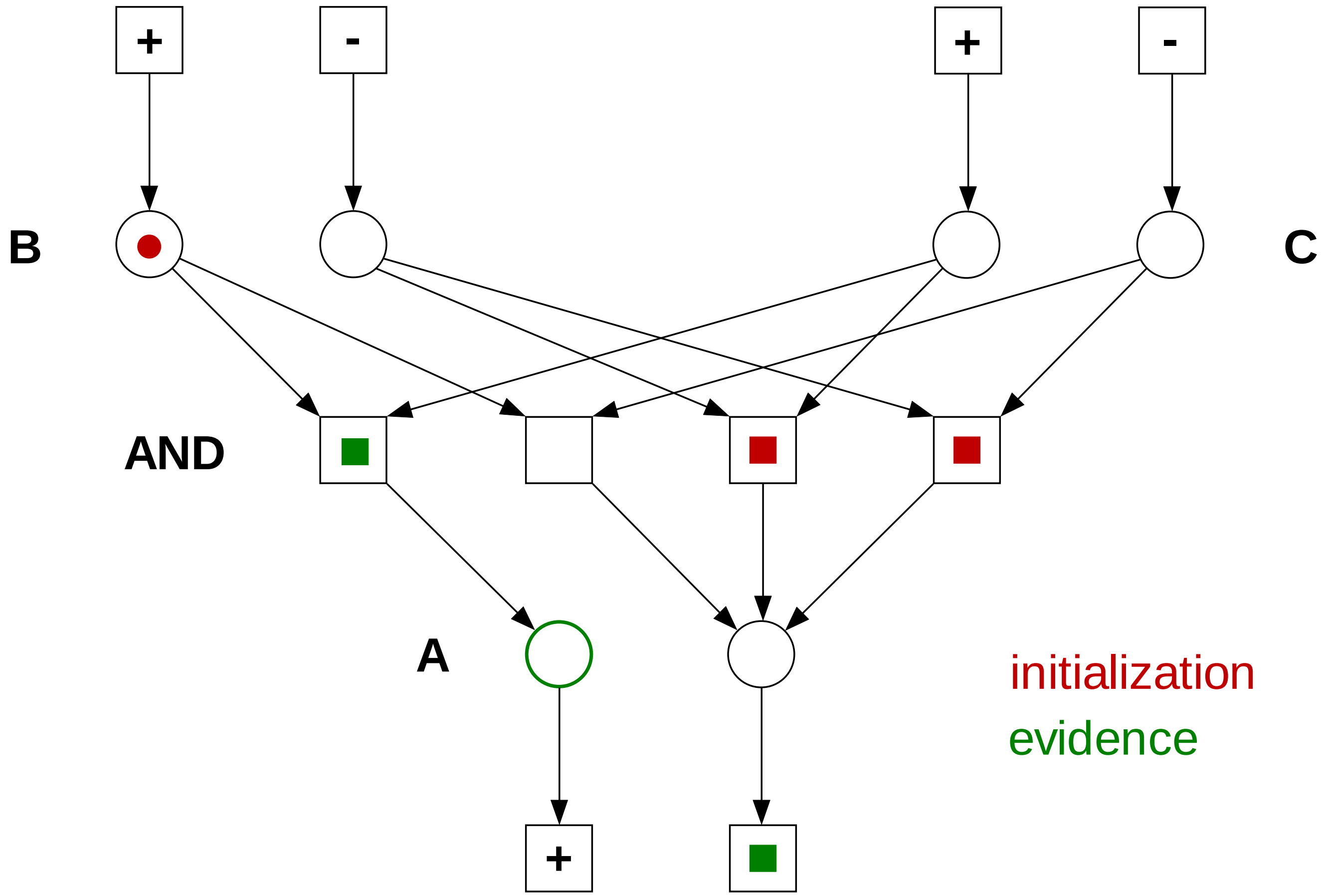
transition 1 must not fire



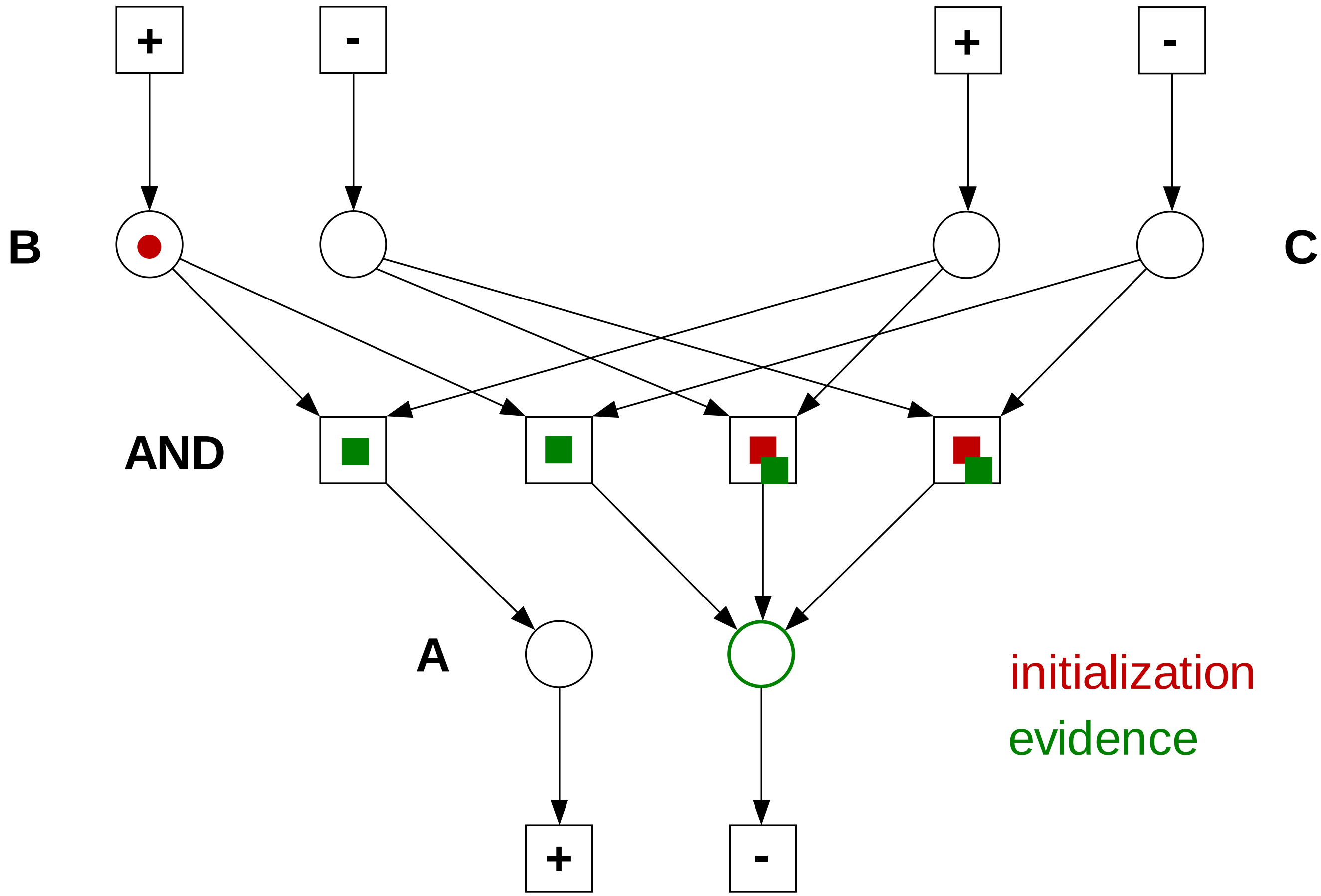
initialization
evidence

Logical Propagation Net

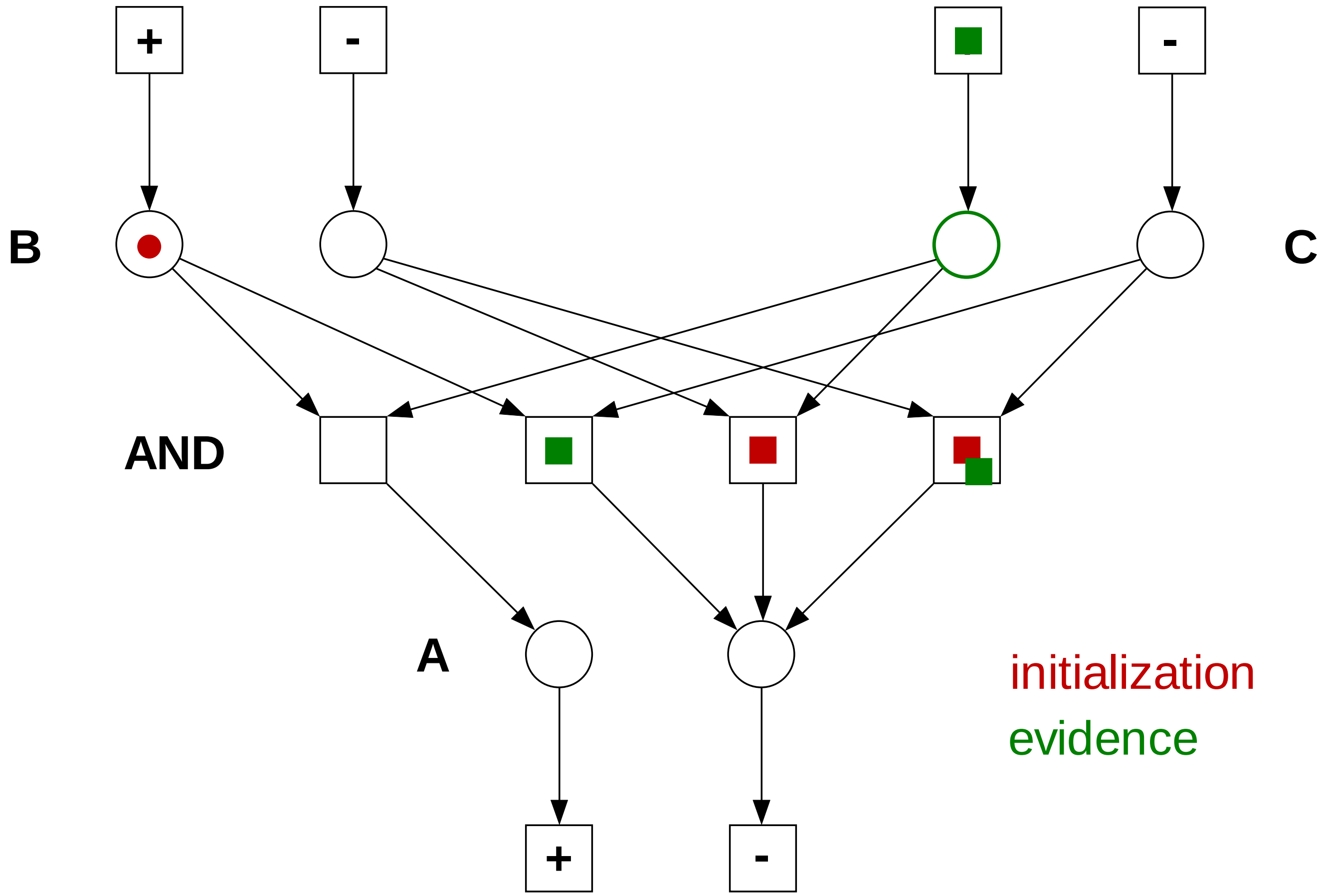



+
-
+
-
B
C
AND
A
+
initialization
evidence

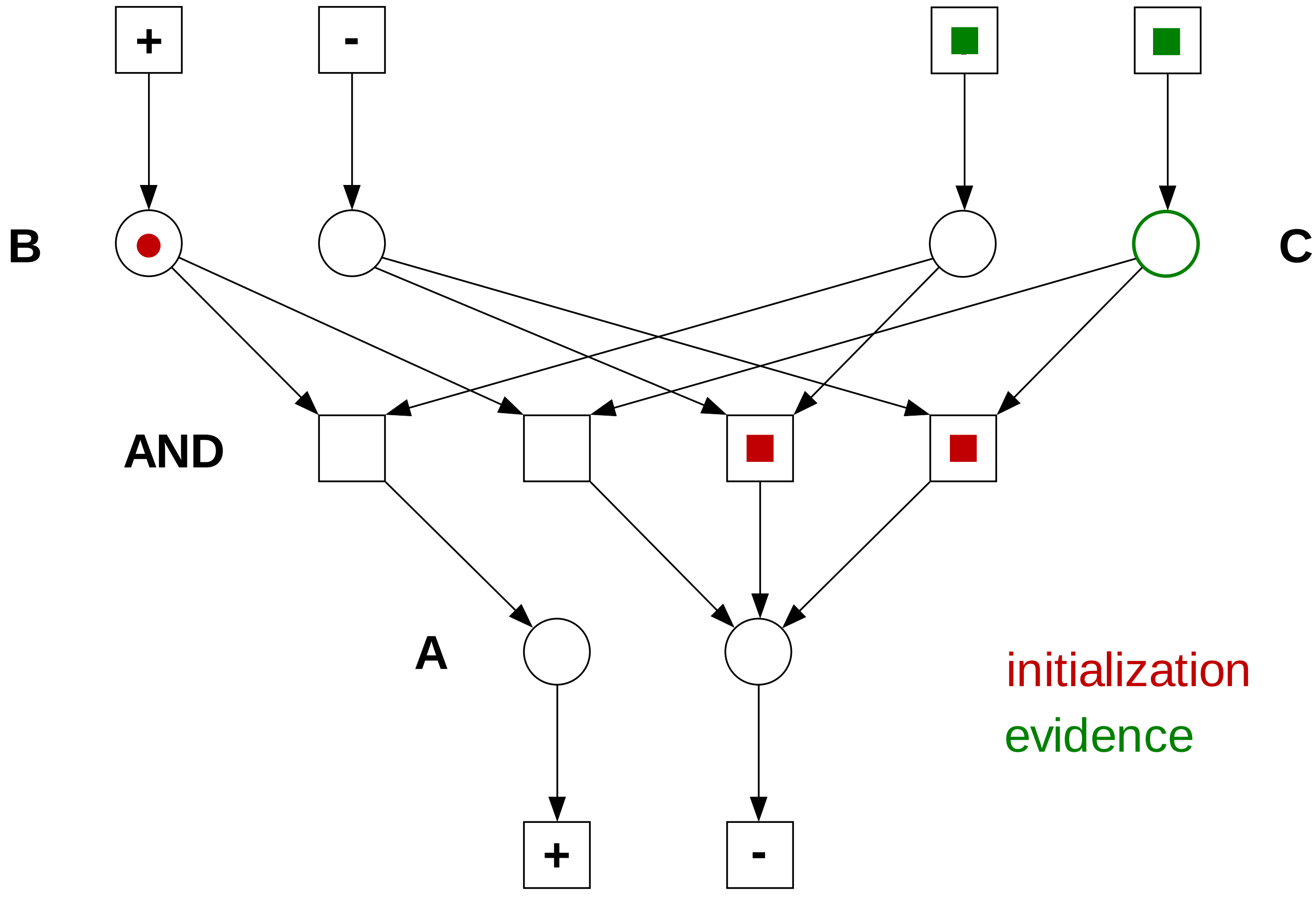
Logical Propagation Net



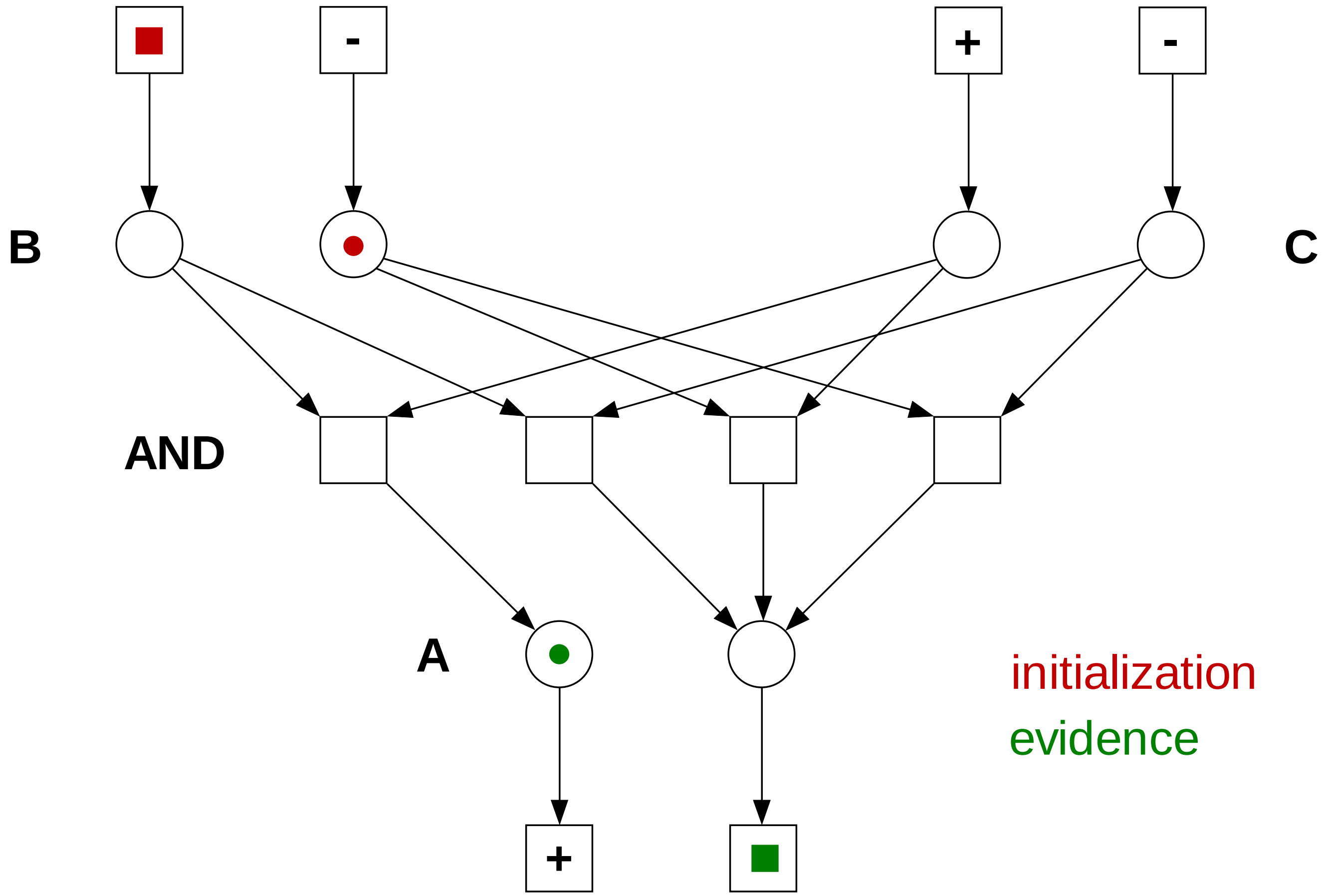
Logical Propagation Net



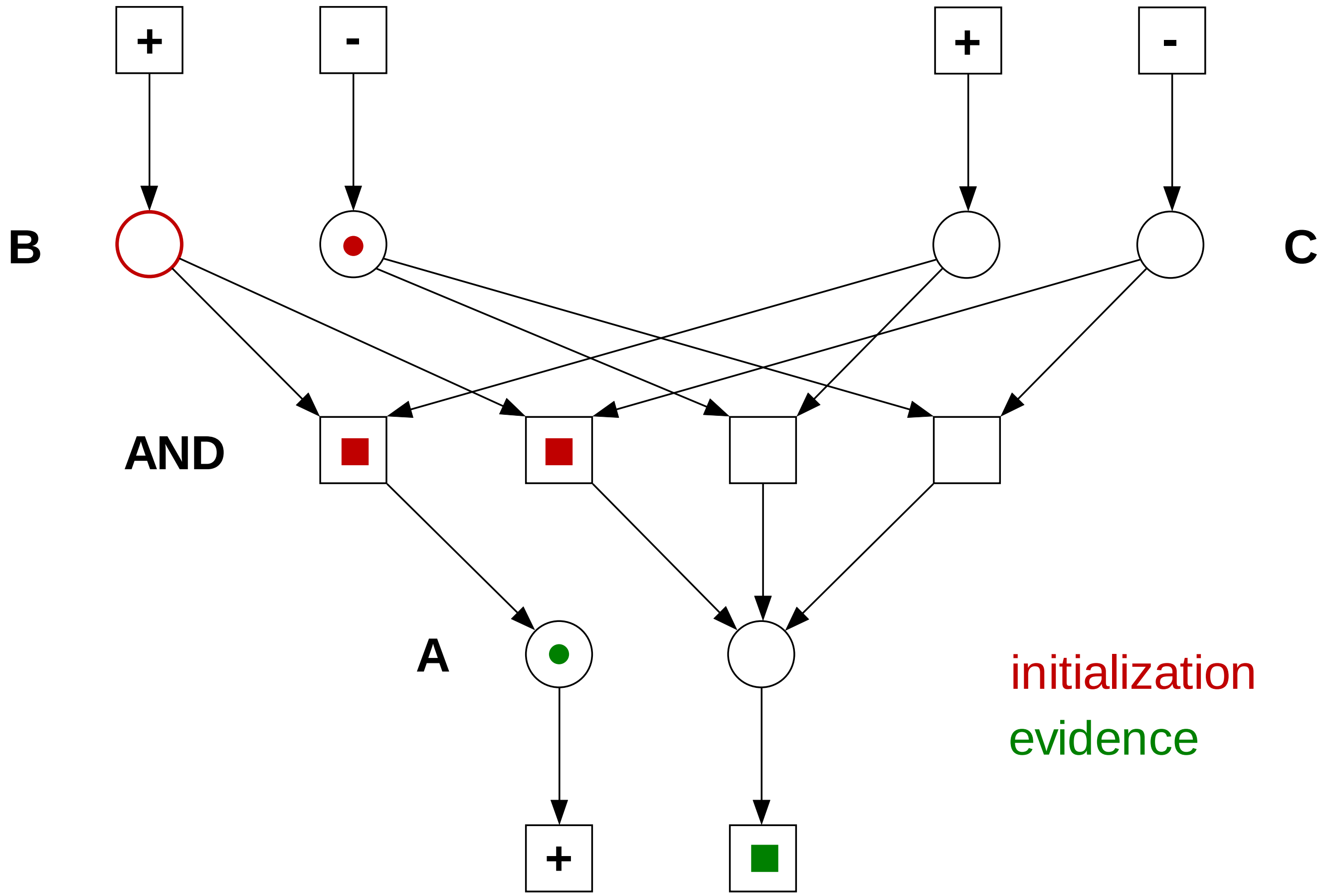
Logical Propagation Net



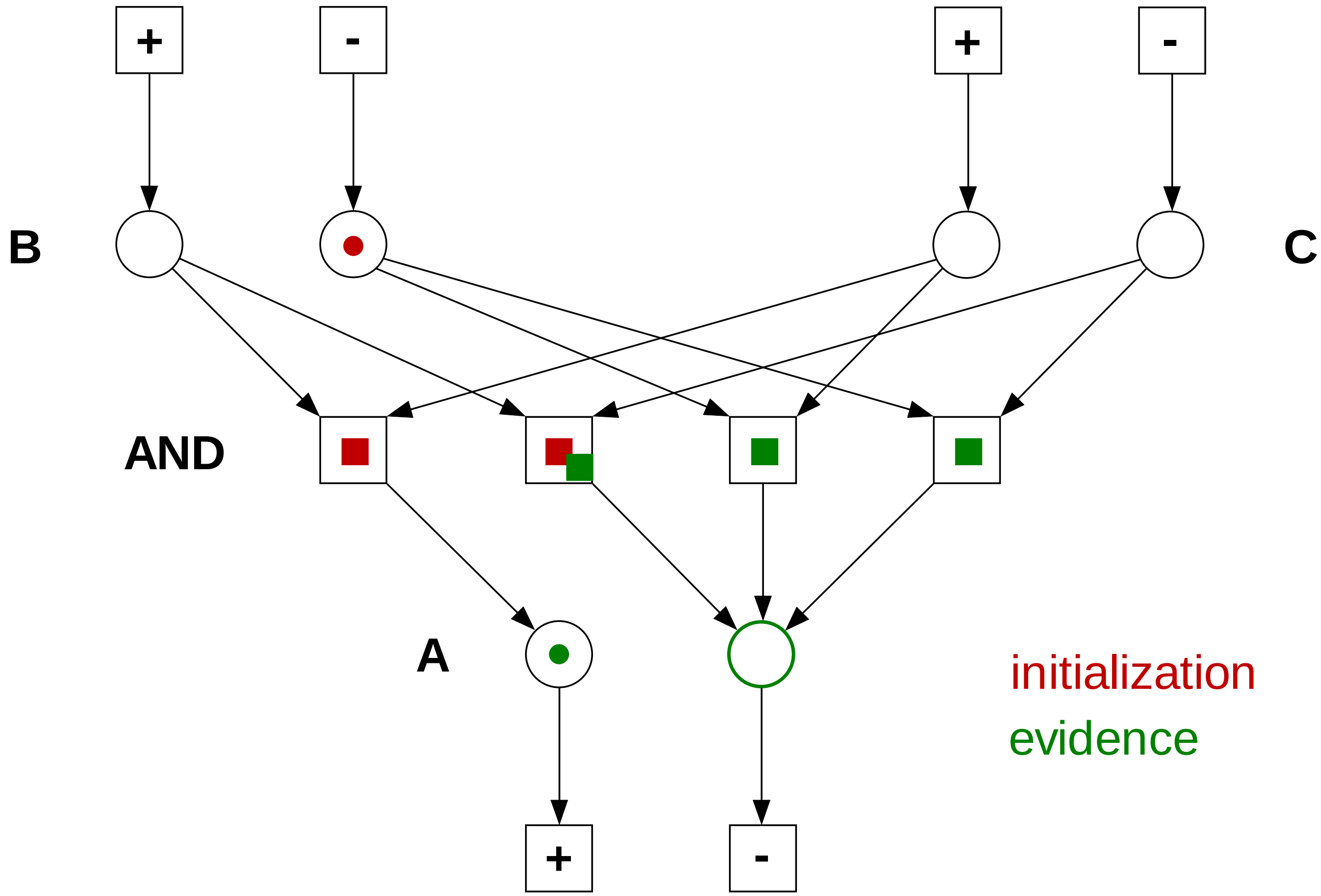
Logical Propagation Net



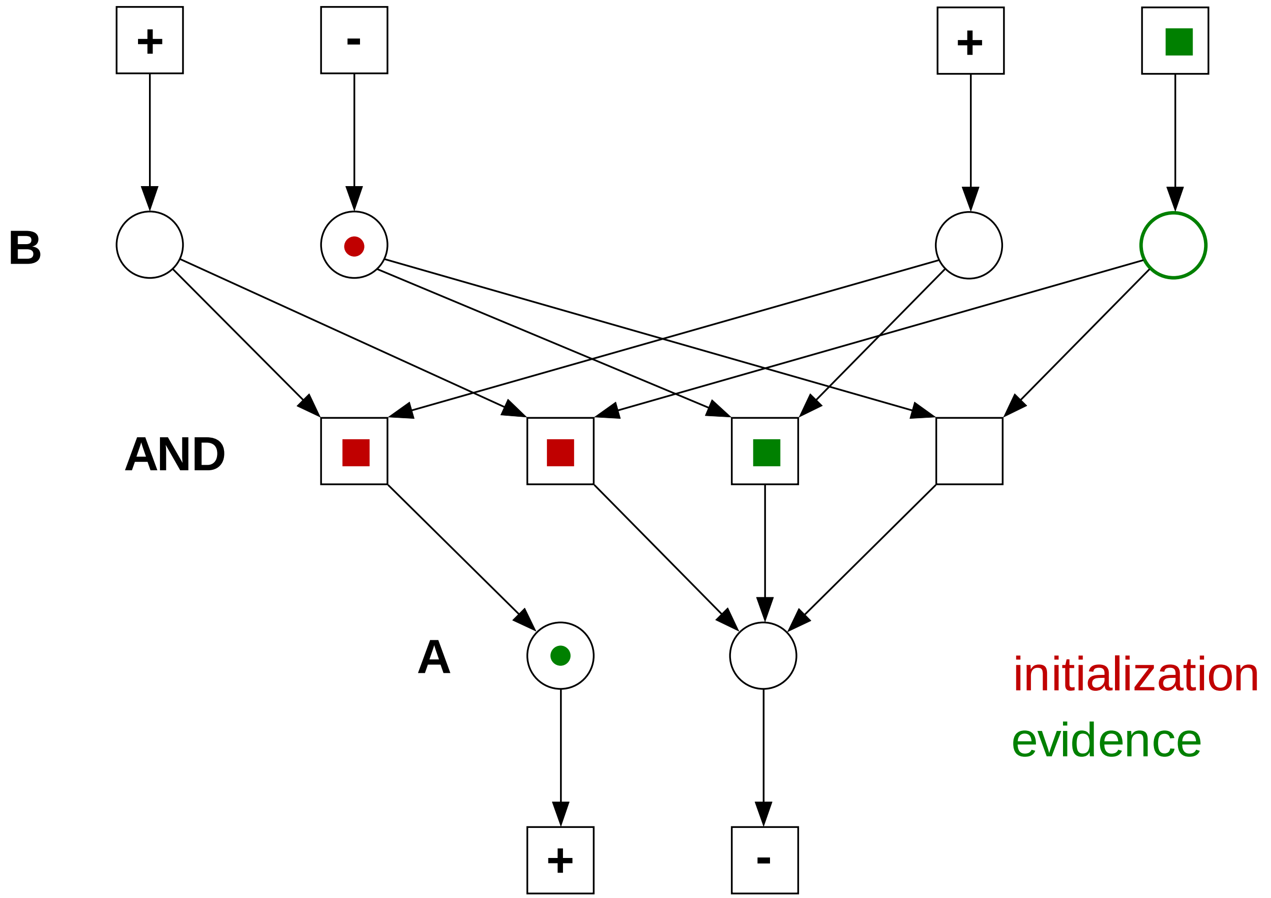
Logical Propagation Net



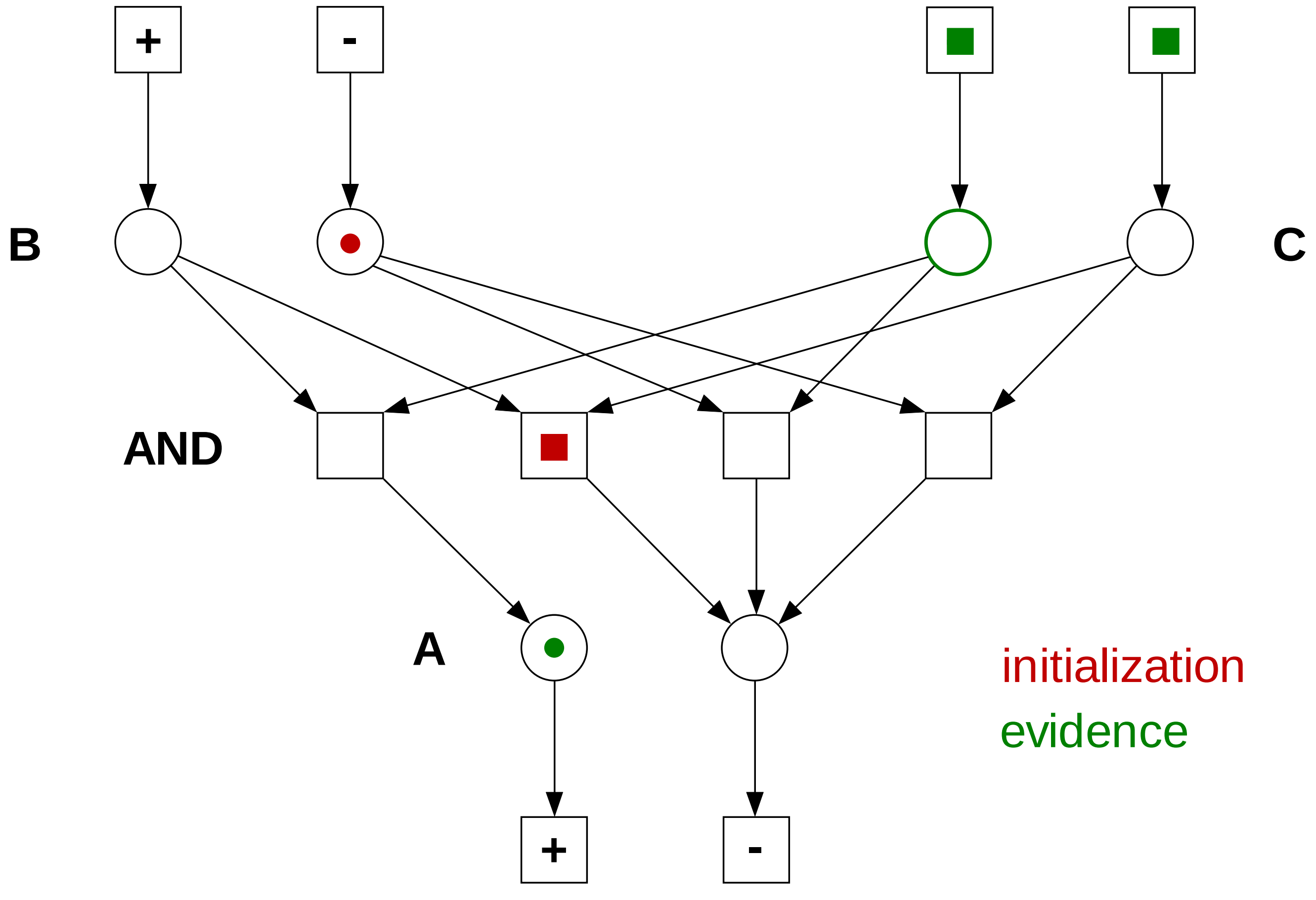
Logical Propagation Net

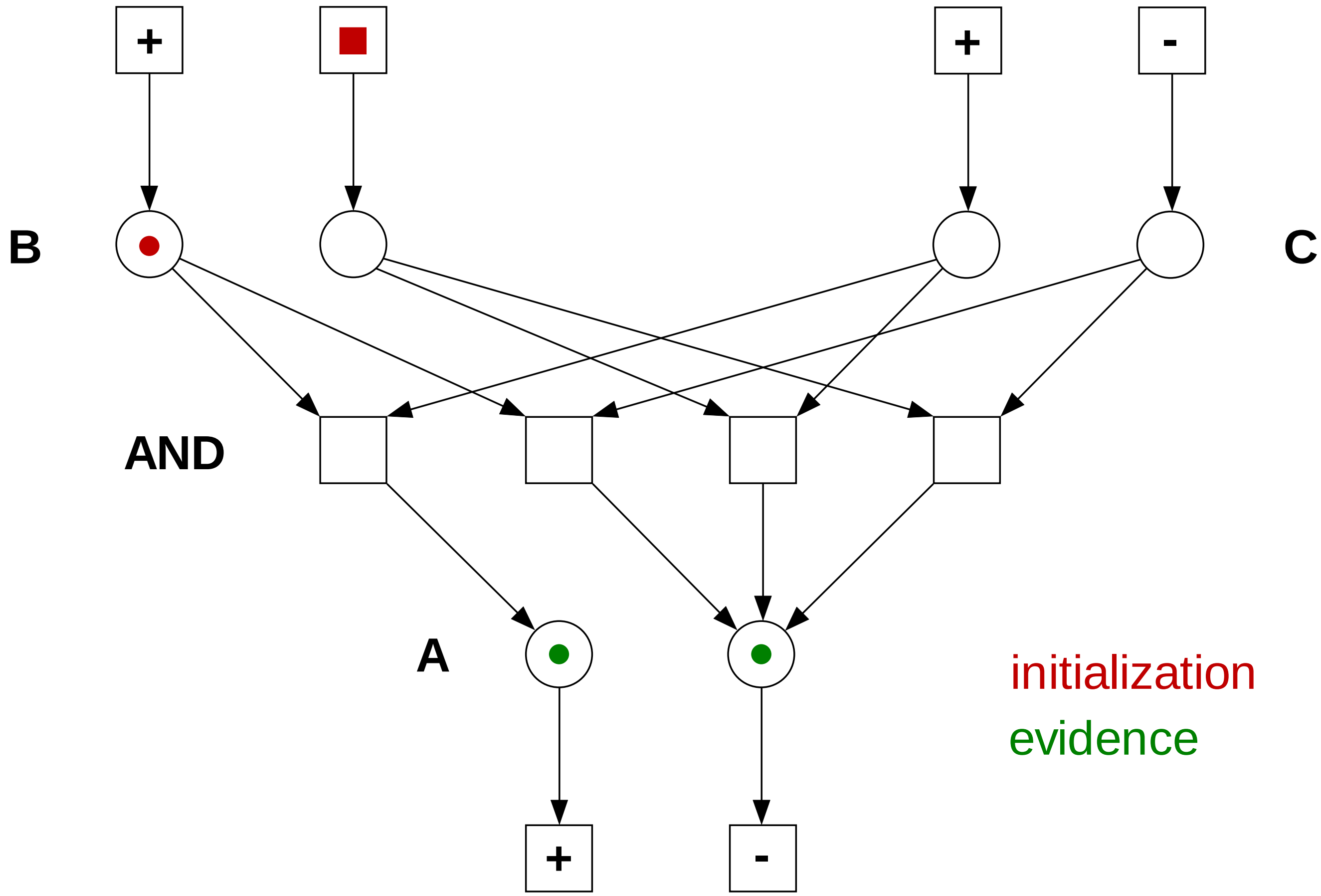


Logical Propagation Net

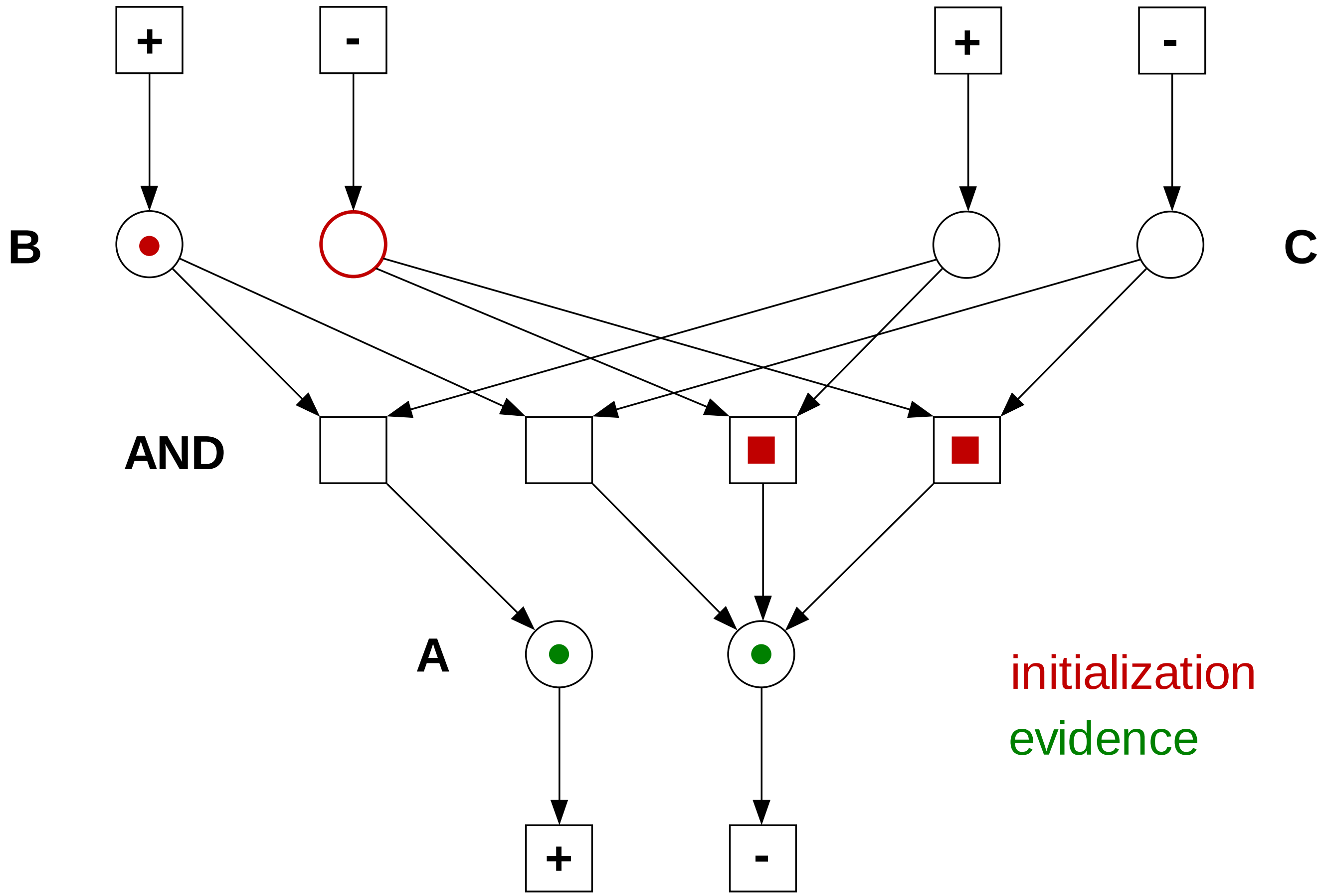


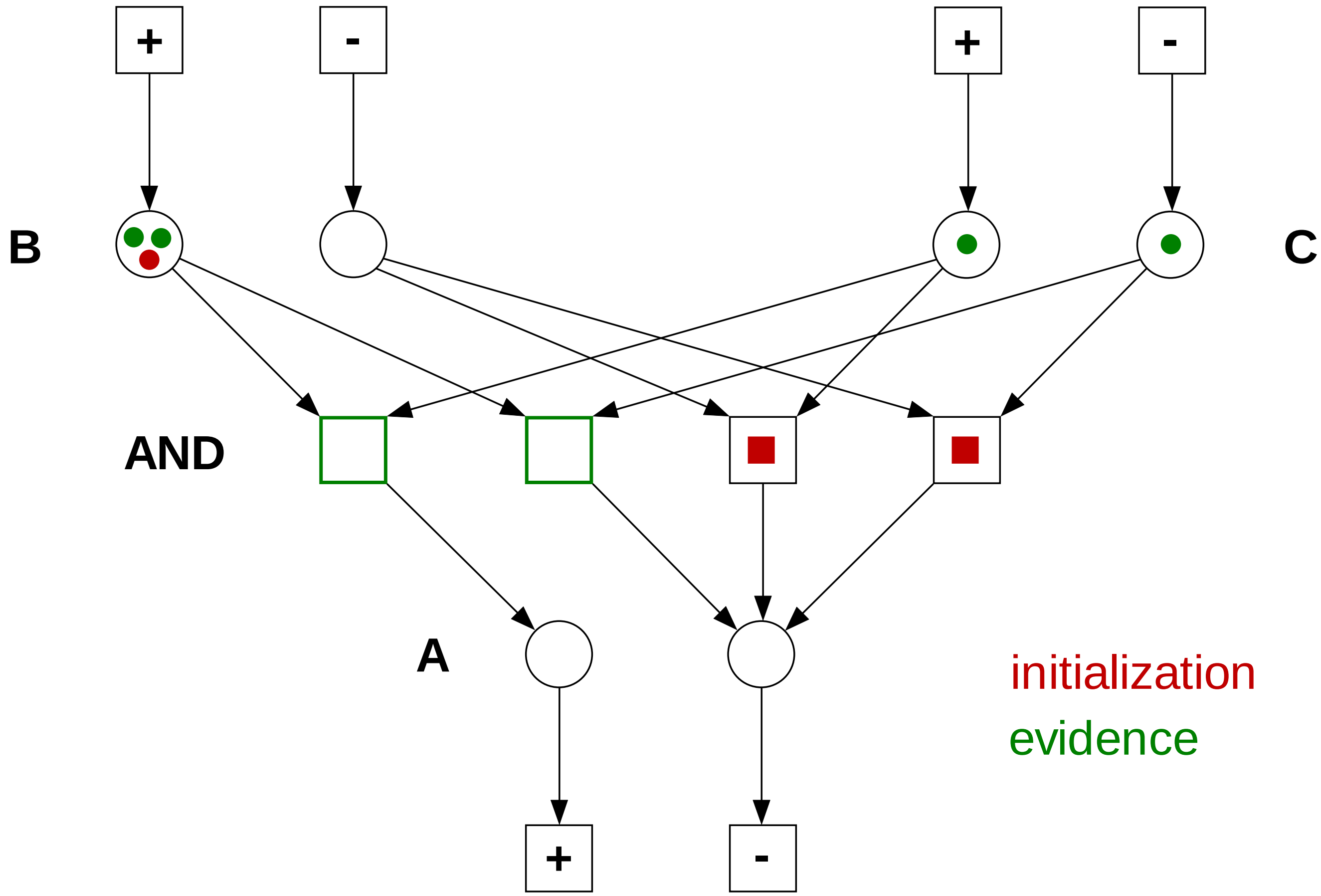
Logical Propagation Net




+
+
-
B
C
AND
A
+
-
initialization
evidence

Logical Propagation Net




+
-
+
-
B
C
AND
A
+
-
initialization
evidence

Evidential reasoning works well in propagation nets because non-events are added in the dependency nets - in a dualized form.

In short:
propagation nets are the "dual completion" of dependency nets.

Contents

C.E. Alchourrón, P. Gärdenfors, D. Makinson developed a series of postulates for belief revision, the AGM-axioms.

They distinguish between three types of belief change:

Expansion: A new piece of information is compatible with the present knowledge and can be added.

Revision: A new piece of information is inconsistent with the present knowledge and cannot be added without eliminating parts of knowledge.

Contraction: The new information means to give up parts of knowledge.

My aims with regard to propagation nets:

(1) to test whether propagation nets are compatible with the AGM-axioms (which is presumably the case).

(2) to check whether propagation nets inscribed with probabilities, masses, time, costs etc. can be a sustainable base for a "sufficiently" general belief revision tool.

(3) to develop a modal logical extension of Bayesian networks and propagation nets.