Unidad 4

Flujo de inferencia

Bibliografía Unidad 4

- Bishop, C. **Pattern Recognition and Machine Learning**. 2006. (Descargar). (lectura 8.1, 8.2 y 8.4)

- Neal. **Pattern Recognition and Machine Learning**. 2020 (Draft). (Descargar). (lectura capítulo 1 y 3)

Otros:

- Kschischang. *Factor graphs and the sum-product algorithm*; IEEE Transactions on information theory. 2001. (Descargar). (lectura partes del paper)

Probabilidad conjunta

Los modelos causales reducen la dimensionalidad de las distribuciones de probabilidad conjunta.

Probabilidad conjunta

$$P(l, e, t, r, a) =$$

Probabilidad conjunta

$$P(l, e, t, r, a) = P(e)\, P(t|e)\, P(a|t, e)\, P(r|a, t, e)\, P(l|a, r, t, e)$$

Probabilidad conjunta

$$P(l, e, t, r, a) = P(e)\, P(t|\cancel{e})\, P(a|t, e)\, P(r|\cancel{a}, t, \cancel{e})\, P(l|a, \cancel{r}, \cancel{t}, \cancel{e})$$

Probabilidad conjunta

$$P(l, e, t, r, a) = P(e)\, P(t)\, P(a|t, e)\, P(r|t)\, P(l|a)$$

Probabilidad conjunta

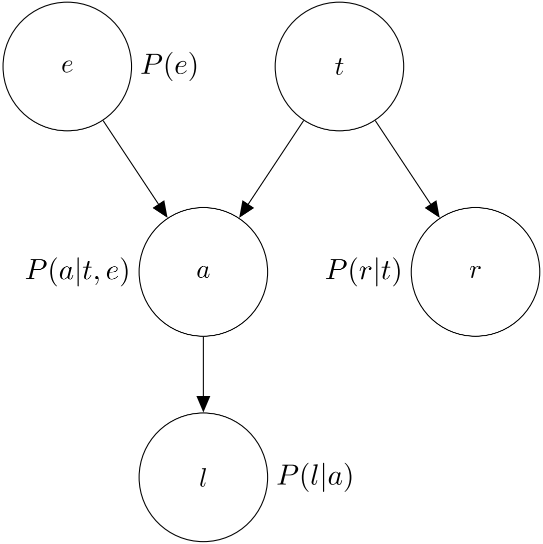
$$\underbrace{P(l, e, t, r, a)}_{\text{Complejidad } 2^n} = P(e)\, P(t)\, P(a|t, e)\, P(r|t)\, P(l|a)$$

Probabilidad conjunta

$$\underbrace{P(l,e,t,r,a)}_{\text{Complejidad } 2^n} = \underbrace{P(e)}_{2}\,\underbrace{P(t)}_{2}\,\underbrace{P(a|t,e)}_{2^3}\,\underbrace{P(r|t)}_{2^2}\,\underbrace{P(l|a)}_{2^2}$$

Mecanismos Causales

Las relaciones causales se expresan mediante distribuciones de probabilidad condicional que relacionan las causas con sus consecuencias.



Mecanismos Causales
Prior Entradera

$$P(e)$$

e^0	e^1

- Una vez cada 3 años
- Una casa cada 1000 por día

Mecanismos Causales
Prior Entradera

$$P(e)$$

e^0	e^1
$999/1000$	$1/1000$

- Una vez cada 3 años
- Una casa cada 1000 por día

Mecanismos Causales
Prior Terremoto

$$P(t)$$

t^0	t^1

- Hay 3 terremotos (leves) por años

Mecanismos Causales
Prior Terremoto

$$P(t)$$

t^0	t^1
$362/365$	$3/365$

- Hay 3 terremotos (leves) por años

Mecanismos Causales
Prior Redes Sociales

$$P(r|t)$$

	r^0	r^1
(t^0)		
(t^1)		

- Siempre que hay un terremoto, en alguna de mis redes sociales (whatsapp, twitter, instagram), se habla del tema.

Mecanismos Causales
Prior Redes Sociales

$$P(r|t)$$

	r^0	r^1
(t^0)	1	0
(t^1)	0	1

- Siempre que hay un terremoto, en alguna de mis redes sociales (whatsapp, twitter, instagram), se habla del tema.

Mecanismos Causales
Prior Redes Sociales

$P(r|t)$

	r^0	r^1
(t^0)	1	0
(t^1)	0	1

- Siempre que hay un terremoto, en alguna de mis redes sociales (whatsapp, twitter, instagram), se habla del tema.
- También puedo mirar mal, o por algún momento no haya nada, o que por algún otro motivo nadie pueda comunicarse

Mecanismos Causales
Prior Redes Sociales

$$P(r|t)$$

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

- Siempre que hay un terremoto, en alguna de mis redes sociales (whatsapp, twitter, instagram), se habla del tema.
- También puedo mirar mal, o por algún momento no haya nada, o que por algún otro motivo nadie pueda comunicarse

Mecanismos Causales
Prior Llamada

$P(l|a)$

	l^0	l^1
(a^0)		
(a^1)		

- Siempre que se activa la alarma, me llaman desde el call center de la empresa (si no pasa nada raro)

Mecanismos Causales
Prior Llamada

$P(l|a)$

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

- Siempre que se activa la alarma, me llaman desde el call center de la empresa (si no pasa nada raro)

Mecanismos Causales

Prior Alarma

$P(a|e,t)$

	a^0	a^1
(e^0, t^0)		
(e^1, t^0)		
(e^0, t^1)		
(e^1, t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

Mecanismos Causales
Prior Alarma

$P(a|e,t)$

	a^0	a^1
(e^0, t^0)		
(e^1, t^0)		
(e^0, t^1)		
(e^1, t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0,t^0)	$P(\overline{\alpha})$	$P(\alpha)$
(e^1,t^0)		
(e^0,t^1)		
(e^1,t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	$P(\overline{\alpha})$	$P(\alpha)$
(e^1, t^0)		
(e^0, t^1)		
(e^1, t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):
  - α Se activa sola: $P(\alpha) = 0.01$
  - $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$
  - $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	$P(\overline{\alpha})$	$P(\alpha)$
(e^1, t^0)	$P(\overline{\alpha \cup \varepsilon})$	$P(\alpha \cup \varepsilon)$
(e^0, t^1)		
(e^1, t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):
  - α Se activa sola: $P(\alpha) = 0.01$
  - $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$
  - $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	$P(\overline{\alpha})$	$P(\alpha)$
(e^1, t^0)	$P(\overline{\alpha} \cap \overline{\varepsilon})$	$P(\alpha \cup \varepsilon)$
(e^0, t^1)		
(e^1, t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0,t^0)	$P(\overline{\alpha})$	$P(\alpha)$
(e^1,t^0)	$P(\overline{\alpha}\cap\overline{\varepsilon})$	$P(\alpha\cup\varepsilon)$
(e^0,t^1)	$P(\overline{\alpha}\cap\overline{\tau})$	$P(\alpha\cup\tau)$
(e^1,t^1)		

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):
  - α Se activa sola: $P(\alpha)=0.01$
  - $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon})=0.01$
  - $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau})=0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	$P(\overline{\alpha})$	$P(\alpha)$
(e^1, t^0)	$P(\overline{\alpha} \cap \overline{\varepsilon})$	$P(\alpha \cup \varepsilon)$
(e^0, t^1)	$P(\overline{\alpha} \cap \overline{\tau})$	$P(\alpha \cup \tau)$
(e^1, t^1)	$P(\overline{\alpha} \cap \overline{\varepsilon} \cap \overline{\tau})$	$P(\alpha \cup \varepsilon \cup \tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	0.99	$P(\alpha)$
(e^1, t^0)	$P(\overline{\alpha} \cap \overline{\varepsilon})$	$P(\alpha \cup \varepsilon)$
(e^0, t^1)	$P(\overline{\alpha} \cap \overline{\tau})$	$P(\alpha \cup \tau)$
(e^1, t^1)	$P(\overline{\alpha} \cap \overline{\varepsilon} \cap \overline{\tau})$	$P(\alpha \cup \varepsilon \cup \tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0,t^0)	0.99	0.01
(e^1,t^0)	$P(\overline{\alpha}\cap\overline{\varepsilon})$	$P(\alpha\cup\varepsilon)$
(e^0,t^1)	$P(\overline{\alpha}\cap\overline{\tau})$	$P(\alpha\cup\tau)$
(e^1,t^1)	$P(\overline{\alpha}\cap\overline{\varepsilon}\cap\overline{\tau})$	$P(\alpha\cup\varepsilon\cup\tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha)=0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon})=0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau})=0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	$0.99 \cdot 0.01$	$P(\alpha \cup \varepsilon)$
(e^0, t^1)	$P(\overline{\alpha} \cap \overline{\tau})$	$P(\alpha \cup \tau)$
(e^1, t^1)	$P(\overline{\alpha} \cap \overline{\varepsilon} \cap \overline{\tau})$	$P(\alpha \cup \varepsilon \cup \tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	≈ 0.01	$P(\alpha \cup \varepsilon)$
(e^0, t^1)	$P(\overline{\alpha} \cap \overline{\tau})$	$P(\alpha \cup \tau)$
(e^1, t^1)	$P(\overline{\alpha} \cap \overline{\varepsilon} \cap \overline{\tau})$	$P(\alpha \cup \varepsilon \cup \tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$P(a|e,t)$

	a^0	a^1
(e^0,t^0)	0.99	0.01
(e^1,t^0)	≈ 0.01	≈ 0.99
(e^0,t^1)	$P(\overline{\alpha}\cap\overline{\tau})$	$P(\alpha\cup\tau)$
(e^1,t^1)	$P(\overline{\alpha}\cap\overline{\varepsilon}\cap\overline{\tau})$	$P(\alpha\cup\varepsilon\cup\tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):
  - α Se activa sola: $P(\alpha)=0.01$
  - $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon})=0.01$
  - $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau})=0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	≈ 0.01	≈ 0.99
(e^0, t^1)	≈ 0.01	≈ 0.99
(e^1, t^1)	$P(\overline{\alpha} \cap \overline{\varepsilon} \cap \overline{\tau})$	$P(\alpha \cup \varepsilon \cup \tau)$

- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):
  - α Se activa sola: $P(\alpha) = 0.01$
  - $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$
  - $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

Mecanismos Causales
Prior Alarma

$$P(a|e,t)$$

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	≈ 0.01	≈ 0.99
(e^0, t^1)	≈ 0.01	≈ 0.99
(e^1, t^1)	≈ 0.0001	≈ 0.9999

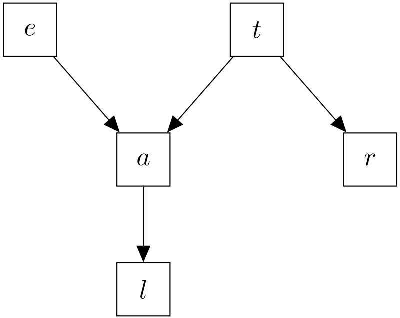
- Siempre que entra alguien a la casa o que hay un terremoto se activa la alarma (si no pasa nada raro):

  α Se activa sola: $P(\alpha) = 0.01$

  $\overline{\varepsilon}$ No se activa a pesar de entradera: $P(\overline{\varepsilon}) = 0.01$

  $\overline{\tau}$ No se activa a pesar de terremoto: $P(\overline{\tau}) = 0.01$

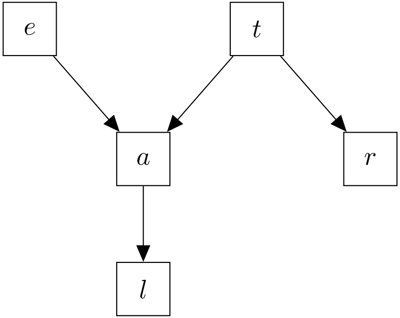
Mecanismos causales



Mecanismos causales

Entradera:

e^0	e^1
0.999	0.001



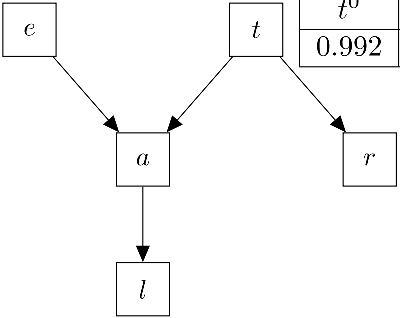
Mecanismos causales

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Mecanismos causales

Entradera:

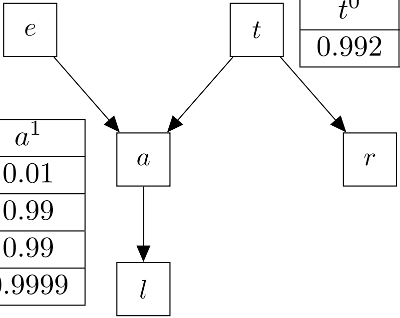
e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999



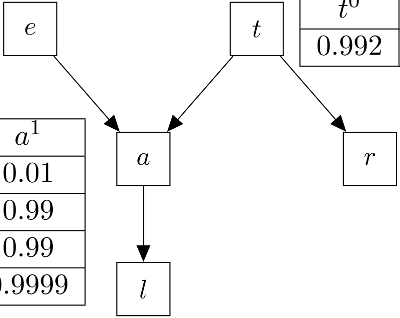
Mecanismos causales

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

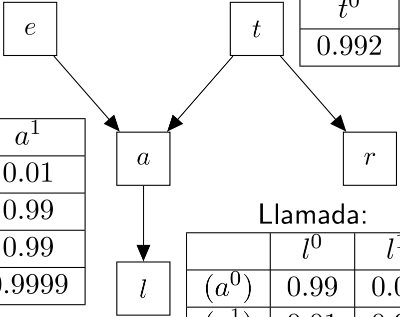
Mecanismos causales

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

Mecanismos causales

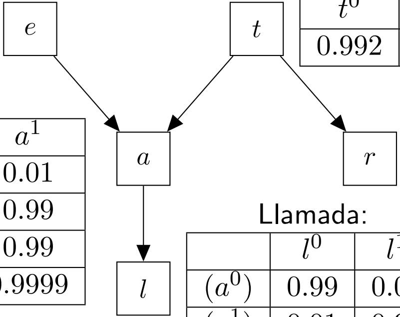
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(e^0, t^0, a^0, r^0, l^0) = P(e^0)P(t^0)P(a^0|t^0, e^0)P(r^0|t^0)P(l^0|a^0)$$

Mecanismos causales

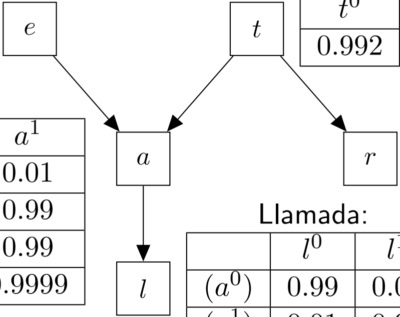
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(e^0, t^0, a^0, r^0, l^0) = 0.999 \cdot P(t^0)P(a^0|t^0, e^0)P(r^0|t^0)P(l^0|a^0)$$

Mecanismos causales

Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

$$\begin{tikzcd} e \arrow[dr] & & t \arrow[dl] \arrow[dr] & \\ & a \arrow[d] & & r \\ & l & & \end{tikzcd}$$

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(e^0, t^0, a^0, r^0, l^0) = 0.999 \cdot 0.992 \cdot 0.99 \cdot 0.99 \cdot 0.99$$

Mecanismos causales

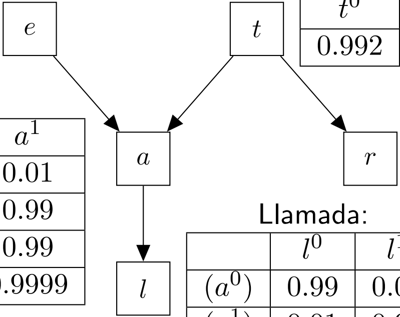
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(e^0, t^0, a^0, r^0, l^0) = 0.999 \cdot 0.992 \cdot 0.99 \cdot 0.99 \cdot 0.99 \approx 0.96$$

Mecanismos causales

Inferencia

$$\begin{tikzcd} e \arrow[dr] & & t \arrow[dl] \arrow[dr] & \\ & a \arrow[d] & & r \\ & l & & \end{tikzcd}$$

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) =$$

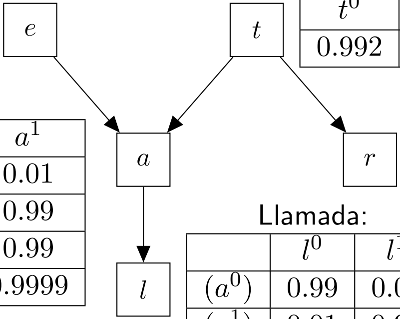
Mecanismos causales
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_e \sum_t \sum_r \sum_l P(e, t, a^1, r, l)$$

Mecanismos causales

Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

$$
\begin{tikzcd}
e \arrow[dr] & & t \arrow[dl] \arrow[dr] & \\
& a \arrow[d] & & r \\
& l & &
\end{tikzcd}
$$

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_{e,t,r,l} P(e, t, a^1, r, l)$$

Mecanismos causales

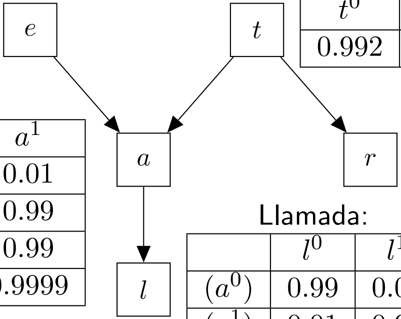
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

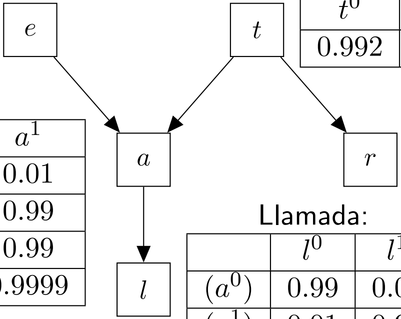
Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_{e,t,r,l} P(e)P(t)P(a^1|t,e)P(r|t)P(l|a^1)$$

Mecanismos causales

Inferencia



Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_{e,t,r,\color{red}{l}} P(e)P(t)P(a^1|t,e)P(r|t)\color{red}{\boldsymbol{P(l|a^1)}}$$

Mecanismos causales

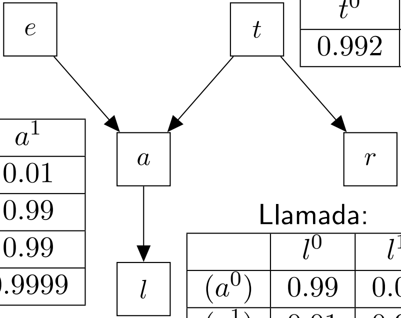
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = P(l^0|a^1)\sum_{e,t,r} P(e)P(t)P(a^1|t,e)P(r|t) + P(l^1|a^1)\sum_{e,t,r} P(e)P(t)P(a^1|t,e)P(r|t)$$

Mecanismos causales

Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

$$\begin{tikzcd} e \arrow[dr] & & t \arrow[dl] \arrow[dr] & \\ & a \arrow[d] & & r \\ & l & & \end{tikzcd}$$

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \Big(\sum_{l} P(l|a^1)\Big)\Big(\sum_{e,t,r} P(e)P(t)P(a^1|t,e)P(r|t)\Big)$$

Mecanismos causales

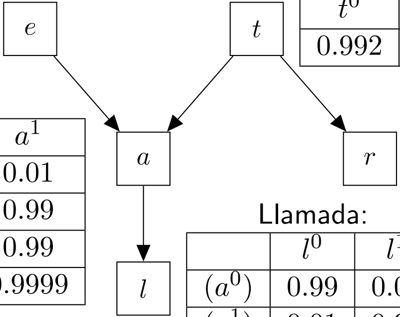
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \Big(\sum_{l} P(l|a^1)\Big)\Big(\sum_{e,\boldsymbol{t},r} P(e)P(t)P(a^1|t,e)P(r|\boldsymbol{t})\Big)$$

Mecanismos causales

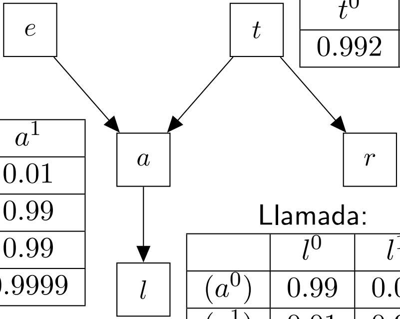
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \Big(\sum_{l} P(l|a^1)\Big)\Big(\sum_{e,t} P(e)P(t)P(a^1|t,e)(\sum_{r} P(r|t))\Big)$$

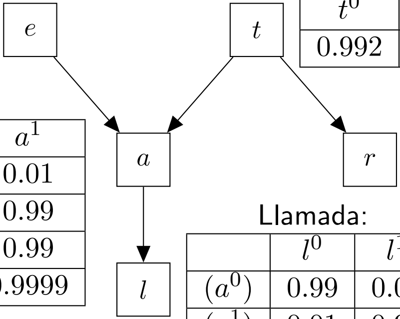
Mecanismos causales
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \Big(\cancel{\sum_l P(l|a^1)}\Big)\Big(\sum_{e,t} P(e)P(t)P(a^1|t,e)(\sum_r P(r|t))\Big)$$

Mecanismos causales

Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

$$\begin{tikzcd}
e \arrow[dr] & & t \arrow[dl] \arrow[dr] & \\
& a \arrow[d] & & r \\
& l & &
\end{tikzcd}$$

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

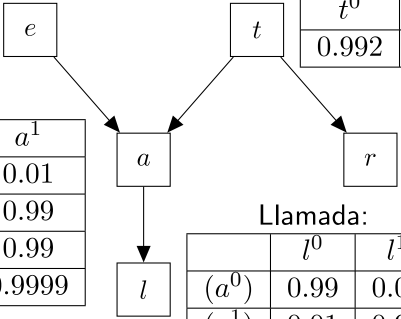
Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \Big(\cancel{\sum_{l} P(l|a^1)}\Big)\Big(\sum_{e,t} P(e)P(t)P(a^1|t,e)(\cancel{\sum_{r} P(r|t)})\Big)$$

Mecanismos causales

Inferencia



Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_{e,t} P(e)P(t)P(a^1|t,e)$$

Mecanismos causales

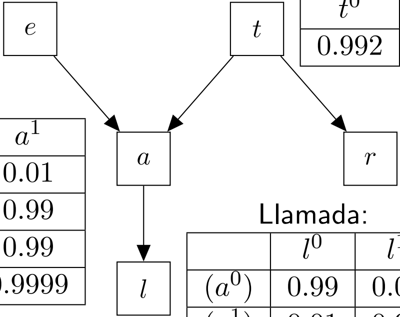
Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008



Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_{e,t} P(e)P(t)P(a^1|t,e) = ?$$

Mecanismos causales

Inferencia

Entradera:

e^0	e^1
0.999	0.001

Terremoto:

t^0	t^1
0.992	0.008

$$\begin{tikzcd} e \arrow[dr] & & t \arrow[dl] \arrow[dr] & \\ & a \arrow[d] & & r \\ & l & & \end{tikzcd}$$

Alarma:

	a^0	a^1
(e^0, t^0)	0.99	0.01
(e^1, t^0)	0.01	0.99
(e^0, t^1)	0.01	0.99
(e^1, t^1)	0.0001	0.9999

Redes:

	r^0	r^1
(t^0)	0.99	0.01
(t^1)	0.01	0.99

Llamada:

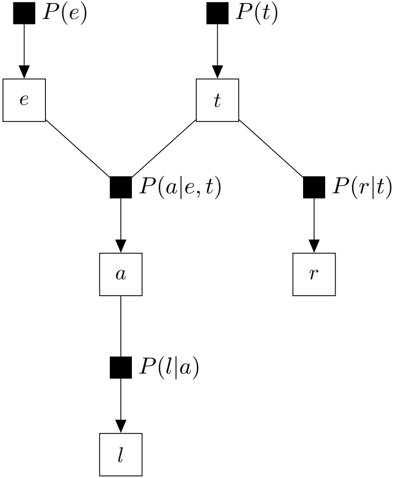
	l^0	l^1
(a^0)	0.99	0.01
(a^1)	0.01	0.99

$$P(a^1) = \sum_{e,t} P(e)P(t)P(a^1|t,e) \approx 0.019$$

Inferencia eficiente por pasaje de mensajes

Método de especificación

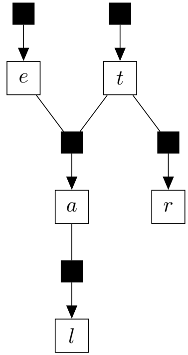
Grafo de factorización (factor graph)



Nodos:
variables y funciones

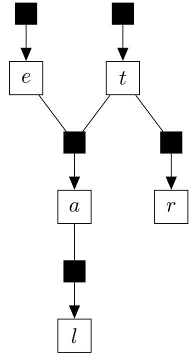
Ejes:
"la variable v es argumento de la función P"

Sum-product algorithm



Algoritmo para calcular cualquier marginal
mediante pasaje de mensajes entre nodos

Sum-product algorithm

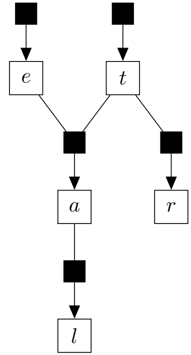


Algoritmo para calcular cualquier marginal mediante pasaje de mensajes entre nodos

$m_{x\rightarrow f}(x)$ **:** Mensaje de variable x a factor f

$m_{f\rightarrow x}(x)$ **:** Mensaje de factor f a variable x

Sum-product algorithm



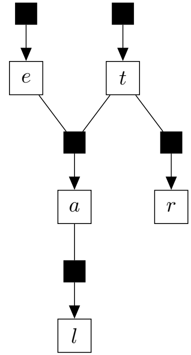
Algoritmo para calcular cualquier marginal mediante pasaje de mensajes entre nodos

$v(n)$ **:** Vecinos del nodo n

$m_{x \to f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \to x}(x)$ **:** Mensaje de factor f a variable x

Sum-product algorithm



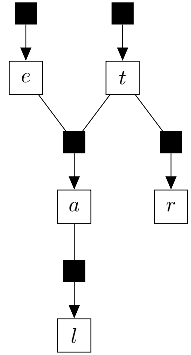
$$P(x) = \prod_{f \in v(x)} m_{f \rightarrow x}(x)$$

$v(n)$ **:** Vecinos del nodo n

$m_{x \rightarrow f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \rightarrow x}(x)$ **:** Mensaje de factor f a variable x

Sum-product algorithm



$$P(x) = \prod_{f \in v(x)} m_{f\rightarrow x}(x)$$

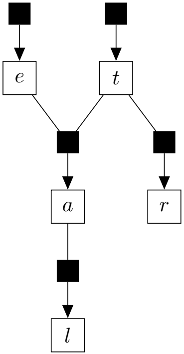
$v(n)$ **:** Vecinos del nodo n

$m_{x\rightarrow f}(x)$ **:** Mensaje de variable x a factor f

$m_{f\rightarrow x}(x)$ **:** Mensaje de factor f a variable x

$$m_{x\rightarrow f}(x) = \prod_{h \in v(x) \setminus \{f\}} m_{h\rightarrow x}(x)$$

Sum-product algorithm



$$P(x) = \prod_{f \in v(x)} m_{f \to x}(x)$$

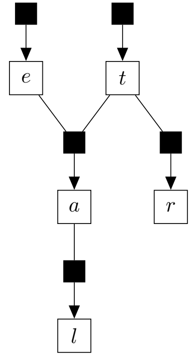
$v(n)$: Vecinos del nodo n

$m_{x \to f}(x)$: Mensaje de variable x a factor f

$m_{f \to x}(x)$: Mensaje de factor f a variable x

$$m_{x \to f}(x) = \prod_{h \in v(x) \setminus \{f\}} m_{h \to x}(x)$$

Sum-product algorithm



$$P(x) = \prod_{f \in v(x)} m_{f\rightarrow x}(x)$$

$v(n)$ **:** Vecinos del nodo n

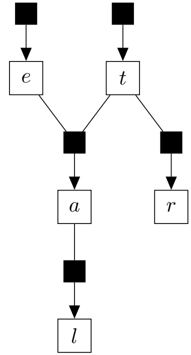
$m_{x\rightarrow f}(x)$ **:** Mensaje de variable x a factor f

$m_{f\rightarrow x}(x)$ **:** Mensaje de factor f a variable x

$$m_{x\rightarrow f}(x) = \prod_{h \in v(x)\setminus\{f\}} m_{h\rightarrow x}(x)$$

$$m_{f\rightarrow x}(x) = \qquad \prod_{h \in v(f)\setminus\{x\}} m_{h\rightarrow f}(h)$$

Sum-product algorithm



$$P(x) = \prod_{f \in v(x)} m_{f \to x}(x)$$

$v(n)$ **:** Vecinos del nodo n

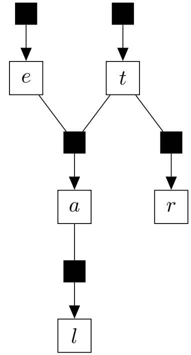
$m_{x \to f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \to x}(x)$ **:** Mensaje de factor f a variable x

$$m_{x \to f}(x) = \prod_{h \in v(x) \setminus \{f\}} m_{h \to x}(x)$$

$$m_{f \to x}(x) = \qquad f(\boldsymbol{h}, x) \prod_{h \in v(f) \setminus \{x\}} m_{h \to f}(h)$$

Sum-product algorithm



$$P(x) = \prod_{f \in v(x)} m_{f \rightarrow x}(x)$$

$v(n)$ **:** Vecinos del nodo n

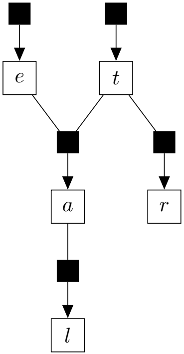
$m_{x \rightarrow f}(x)$ **:** Mensaje de variable x a factor f

$m_{f \rightarrow x}(x)$ **:** Mensaje de factor f a variable x

$$m_{x \rightarrow f}(x) = \prod_{h \in v(x) \setminus \{f\}} m_{h \rightarrow x}(x)$$

$$m_{f \rightarrow x}(x) = \sum_{\boldsymbol{h}} \Big(f(\boldsymbol{h}, x) \prod_{h \in v(f) \setminus \{x\}} m_{h \rightarrow f}(h) \Big)$$

Sum-product algorithm



$$P(x) = \prod_{f \in v(x)} m_{f \to x}(x)$$

$v(n)$ **:** Vecinos del nodo n

$m_{x \to f}(x)$ **:** Mensaje de variable x a factor f

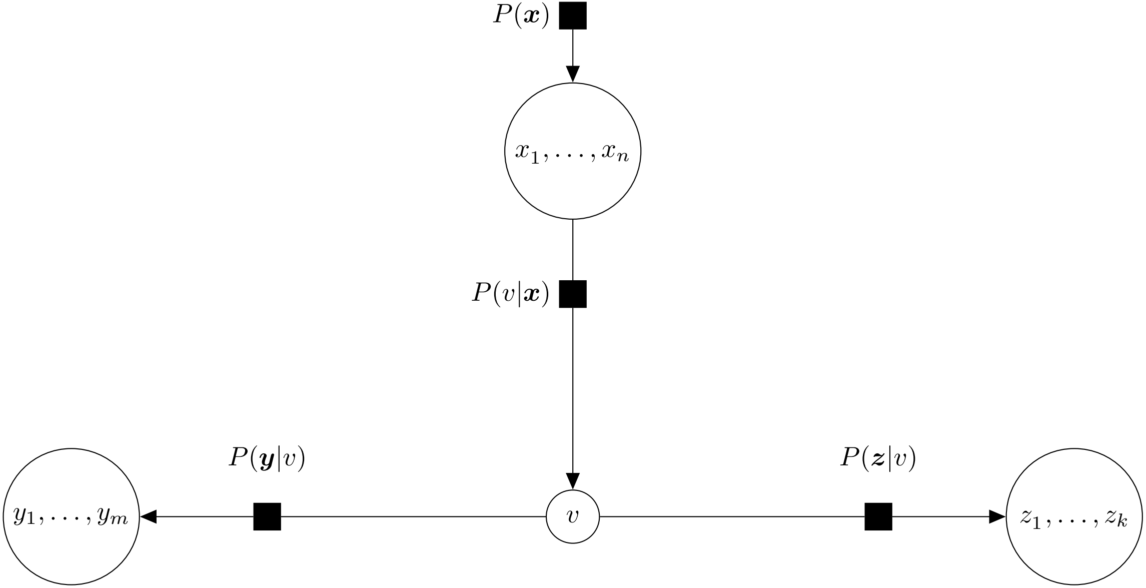
$m_{f \to x}(x)$ **:** Mensaje de factor f a variable x

$$m_{x \to f}(x) = \prod_{h \in v(x) \setminus \{f\}} m_{h \to x}(x)$$

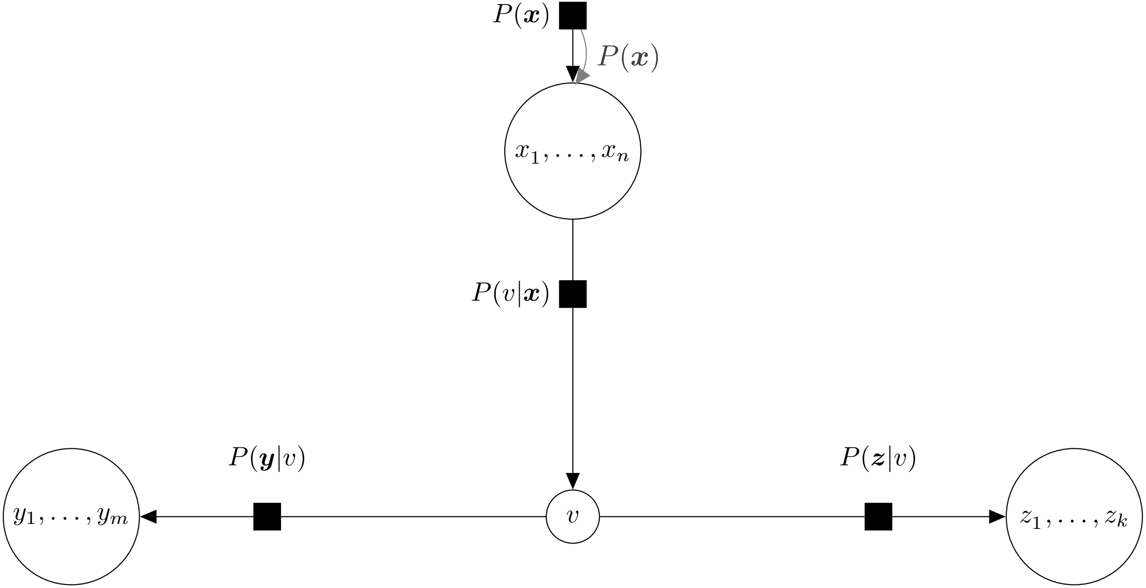
$$m_{f \to x}(x) = \sum_{\boldsymbol{h}} \left(f(\boldsymbol{h}, x) \prod_{h \in v(f) \setminus \{x\}} m_{h \to f}(h) \right)$$

Kschischang FR, Frey BJ, Loeliger HA. Factor graphs and the sum-product algorithm. 2001

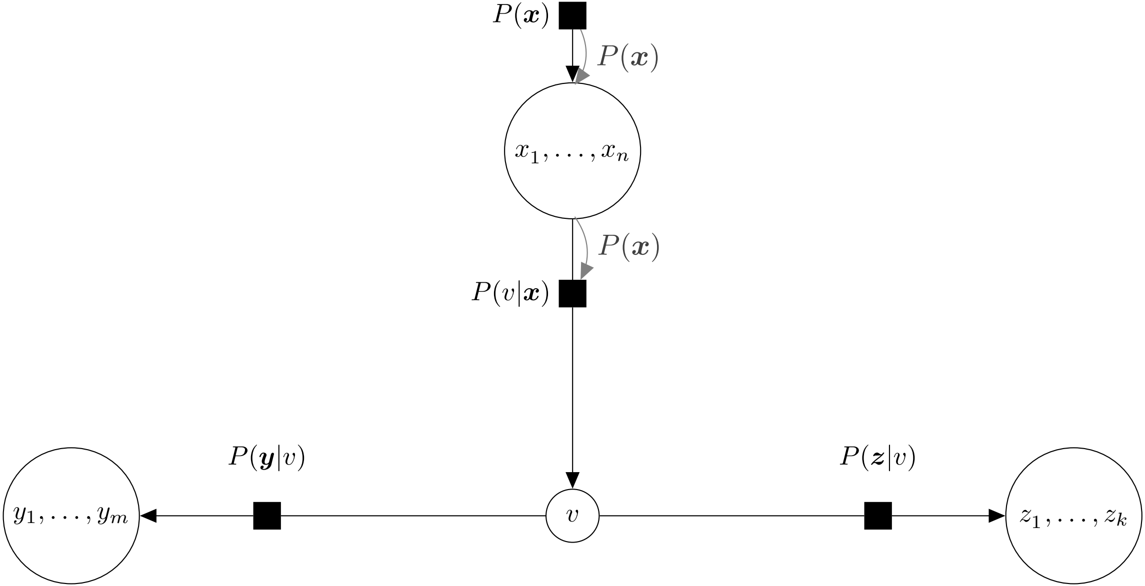
Sum-product algorithm



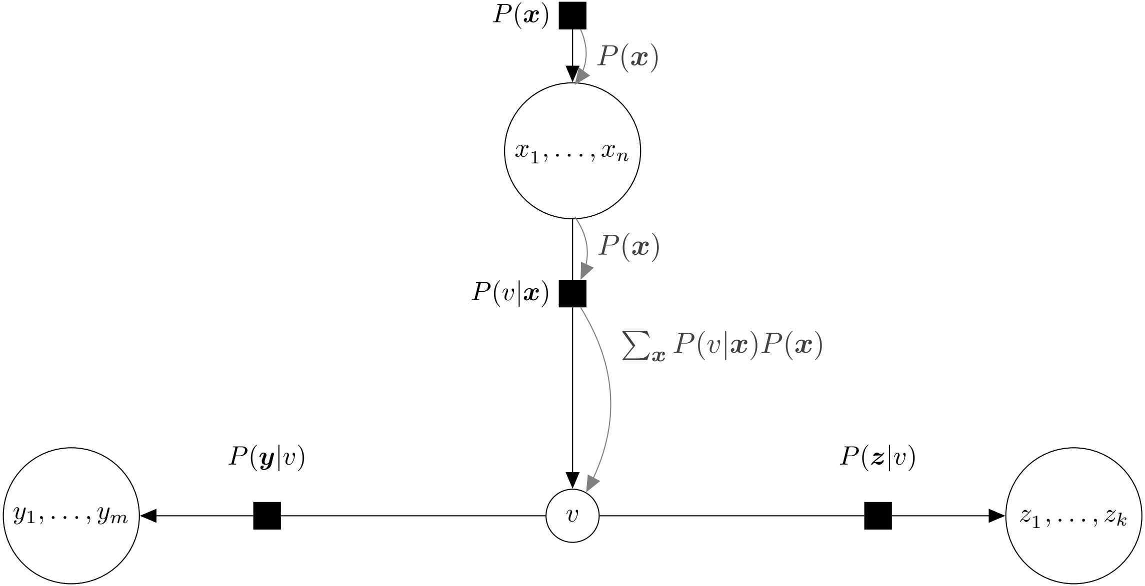
Sum-product algorithm



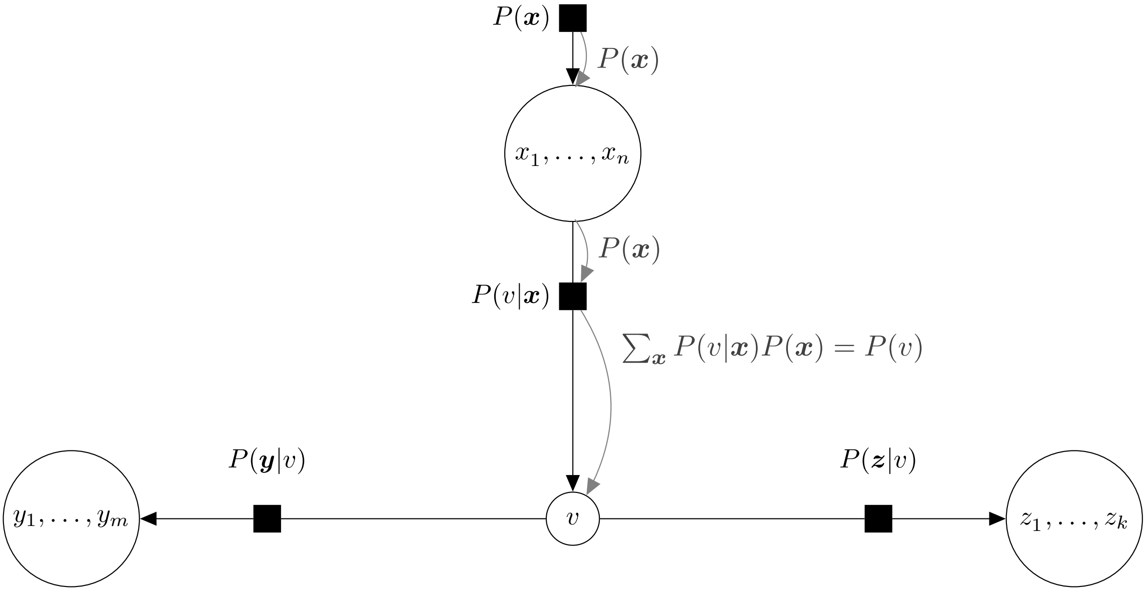
Sum-product algorithm



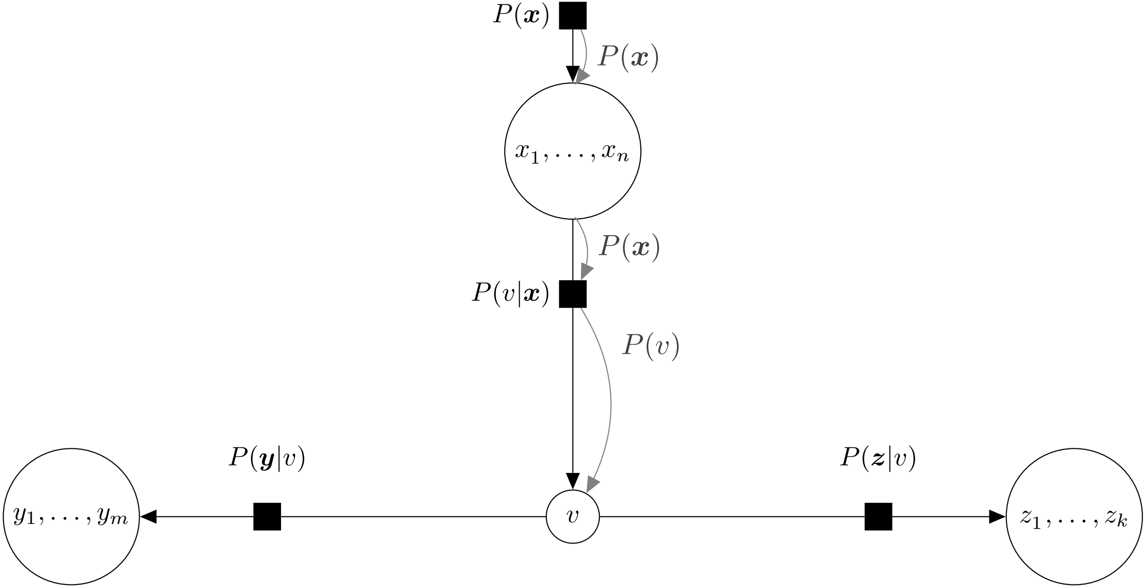
Sum-product algorithm



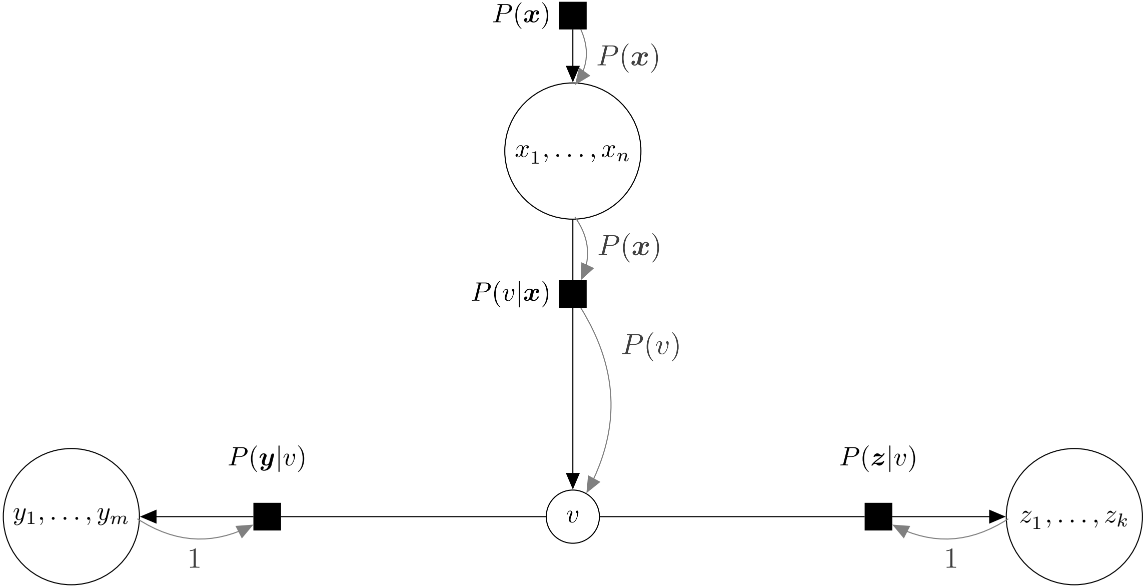
Sum-product algorithm



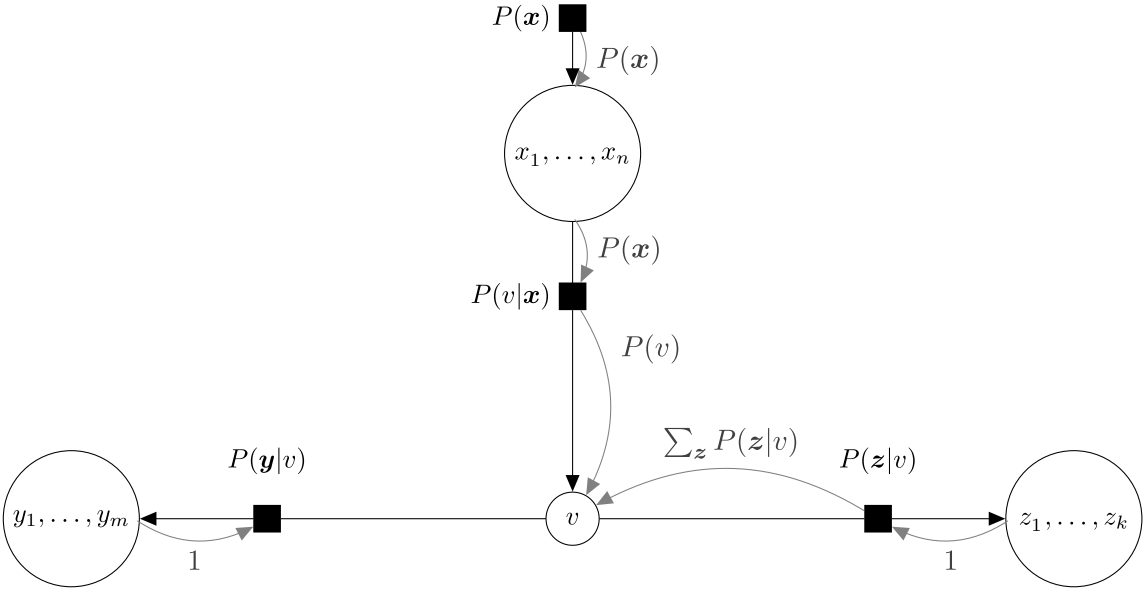
Sum-product algorithm



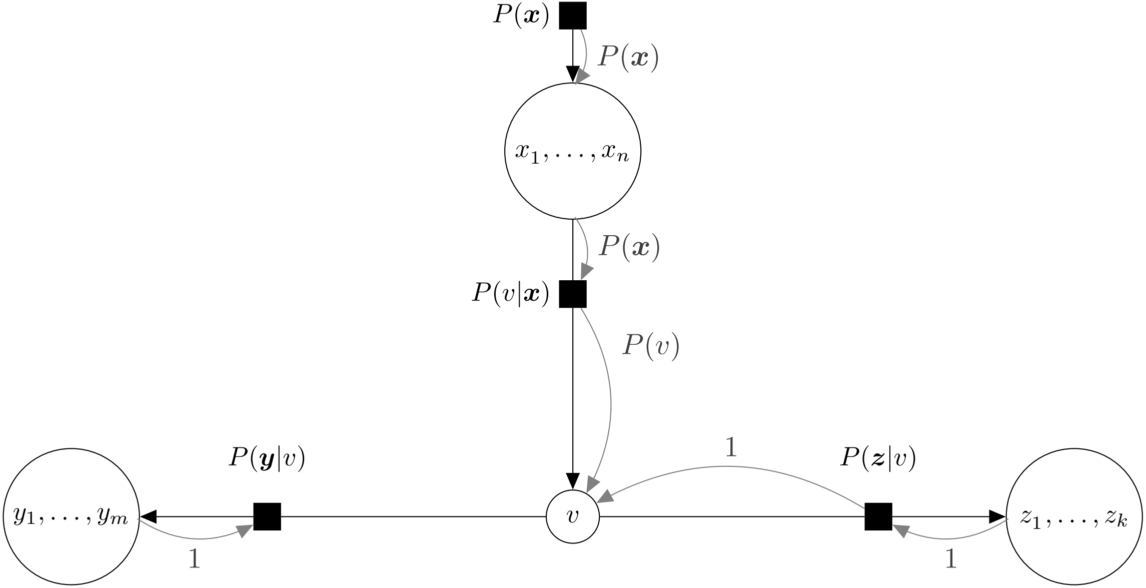
Sum-product algorithm



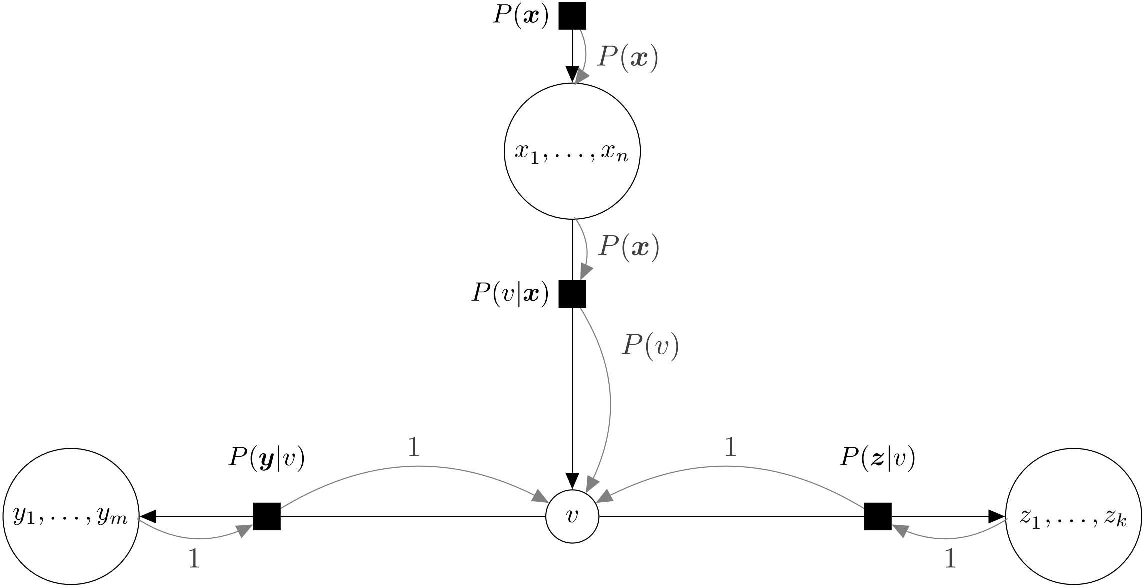
Sum-product algorithm



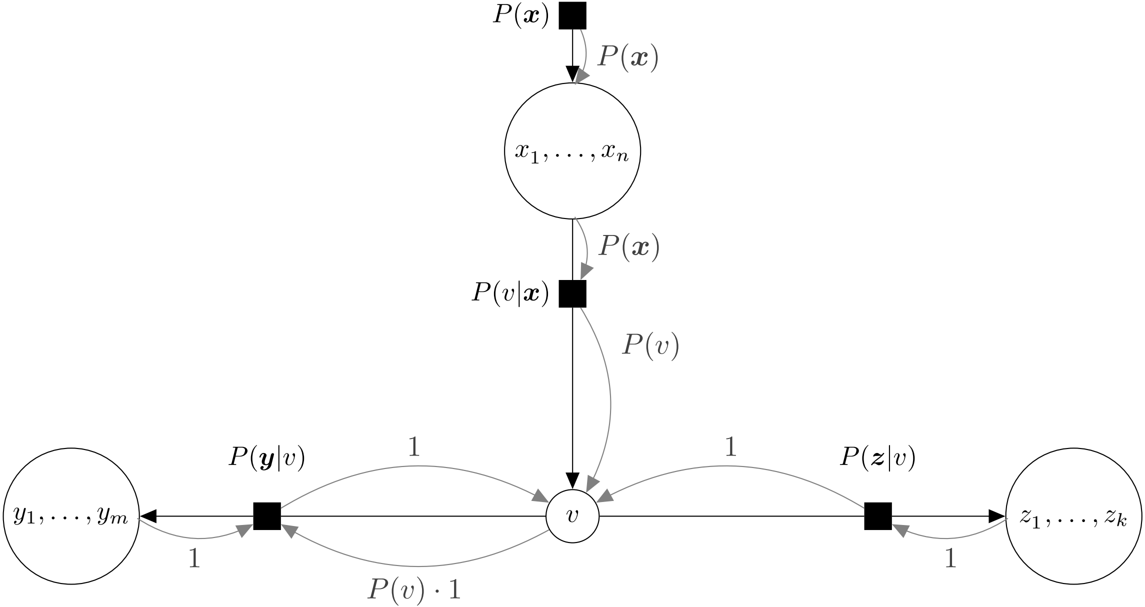
Sum-product algorithm



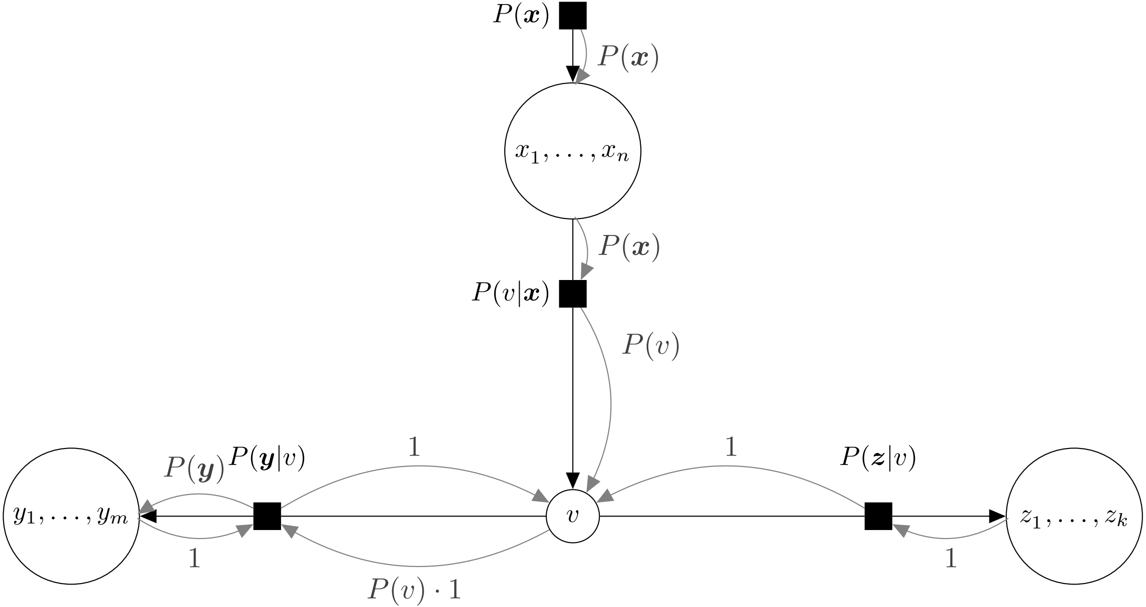
Sum-product algorithm



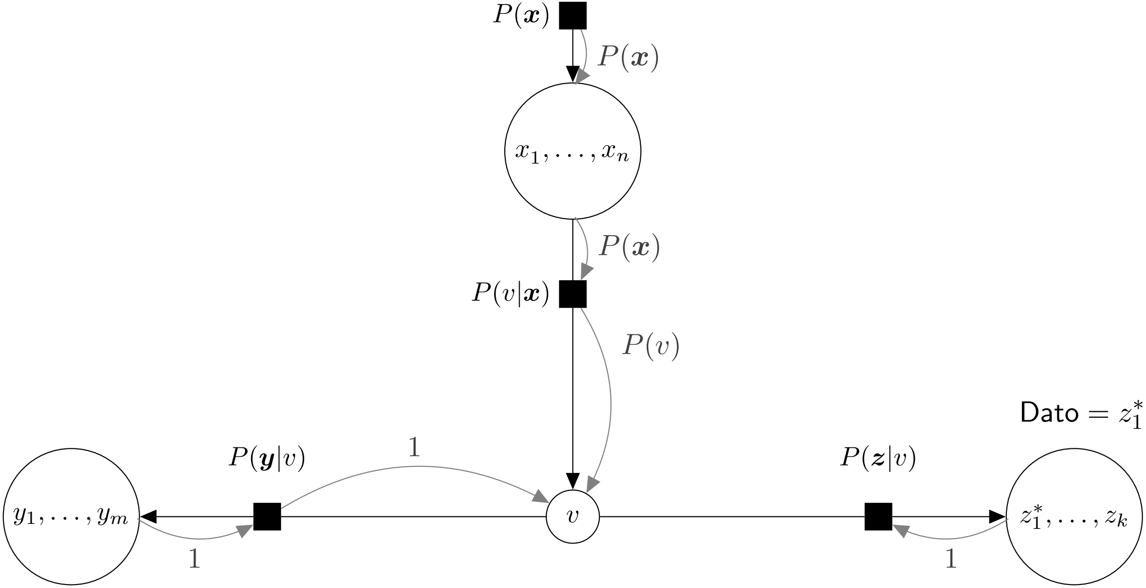
Sum-product algorithm



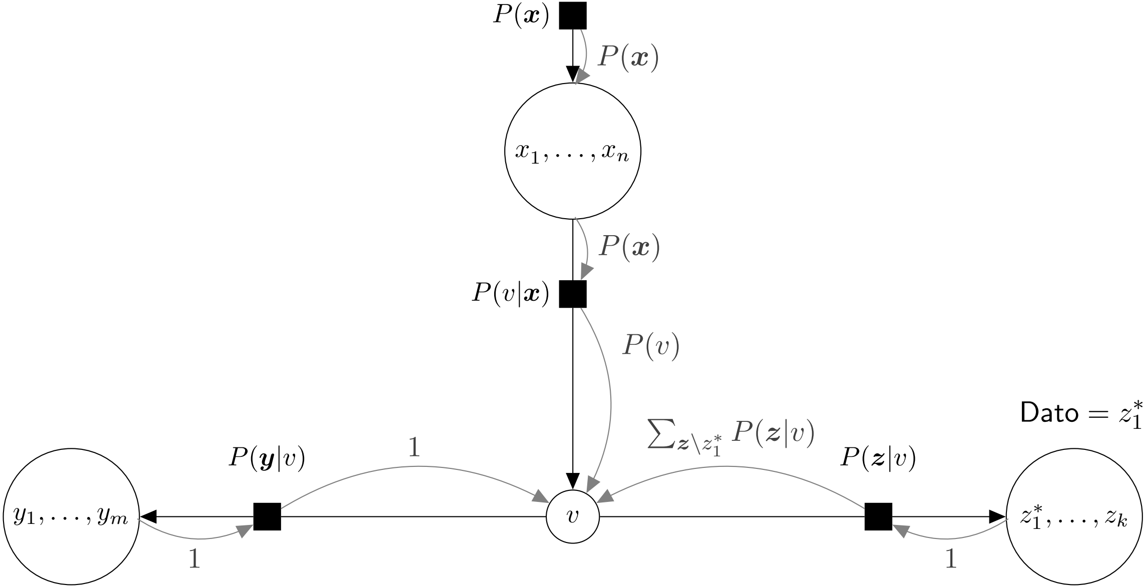
Sum-product algorithm



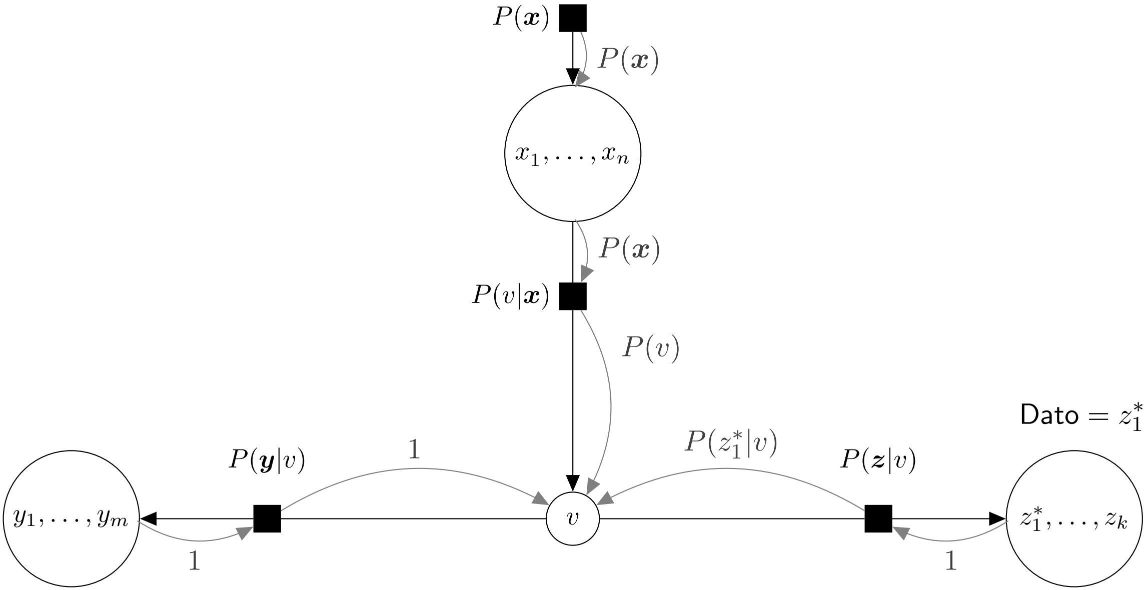
Sum-product algorithm



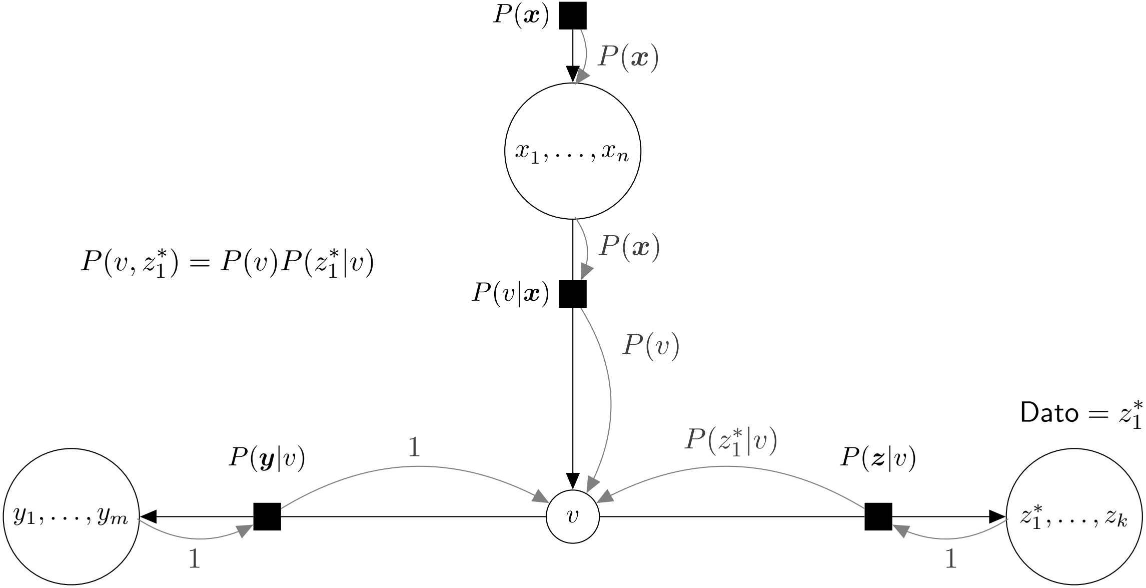
Sum-product algorithm



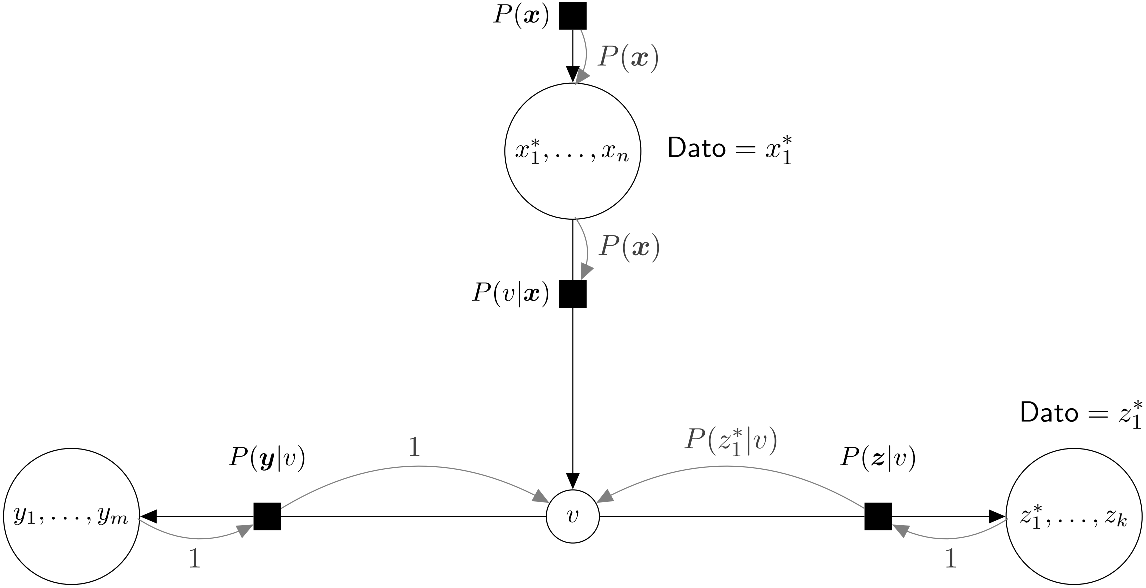
Sum-product algorithm



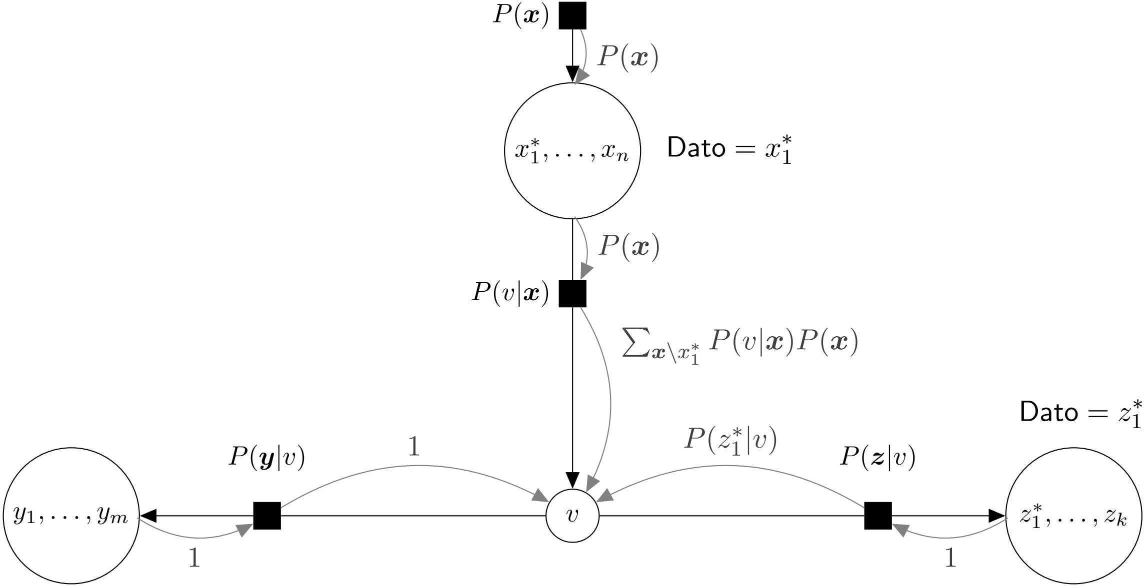
Sum-product algorithm



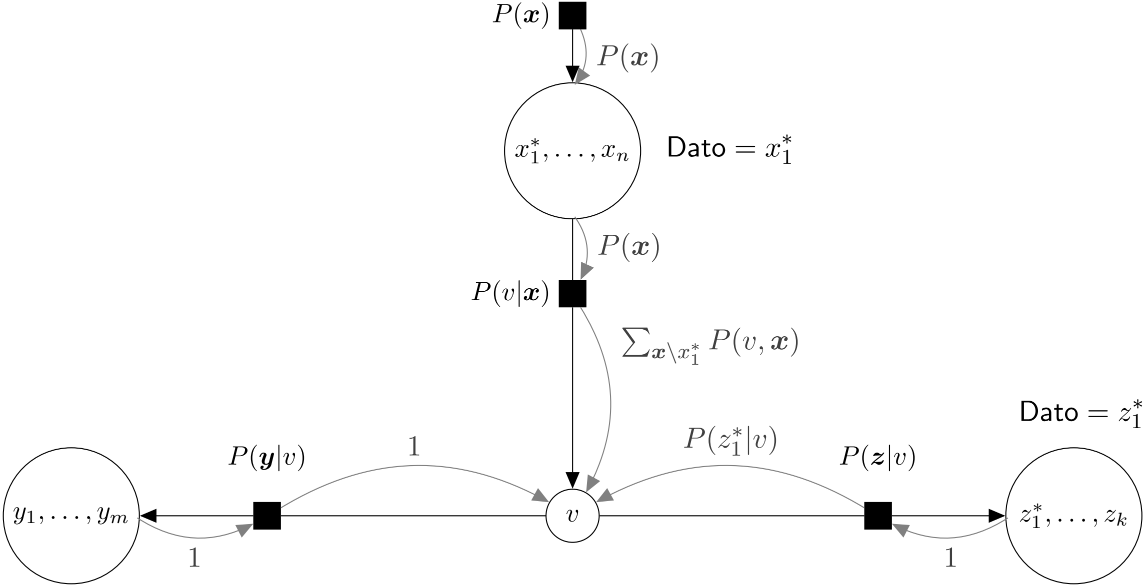
Sum-product algorithm



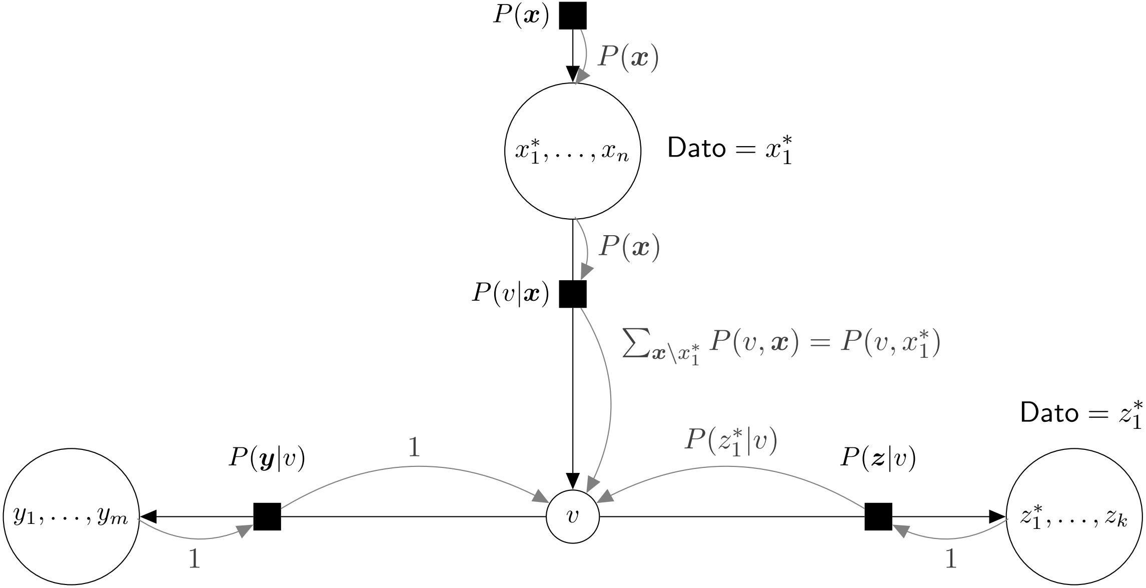
Sum-product algorithm



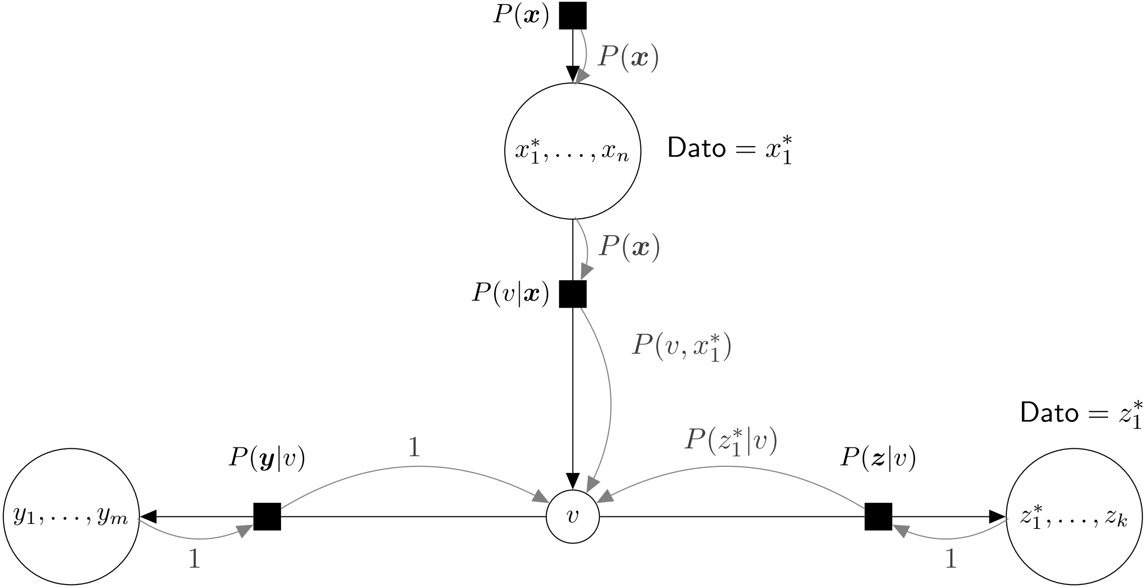
Sum-product algorithm



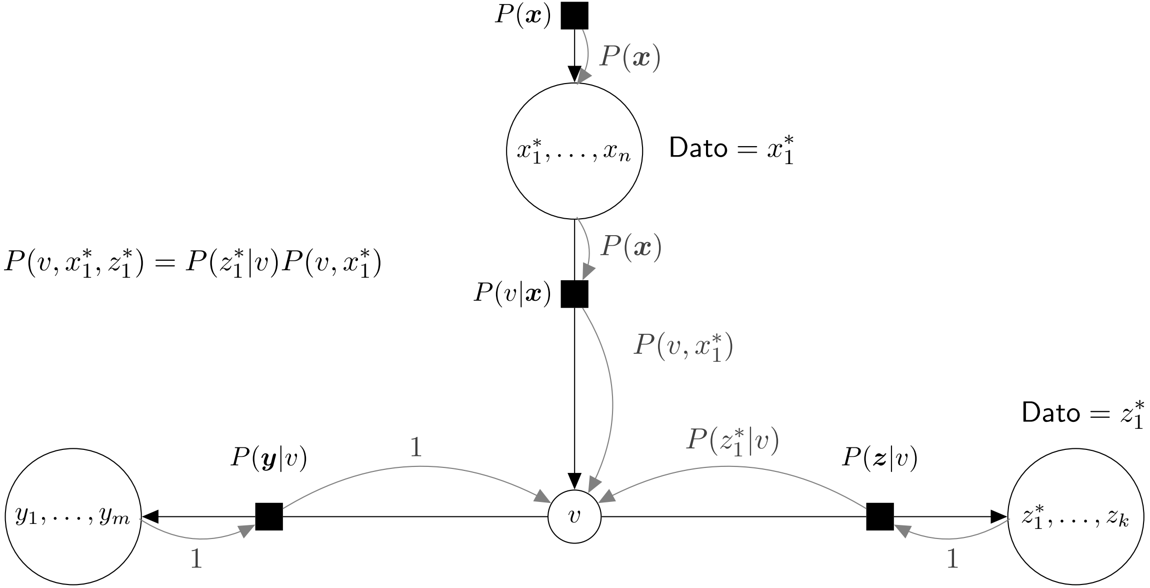
Sum-product algorithm



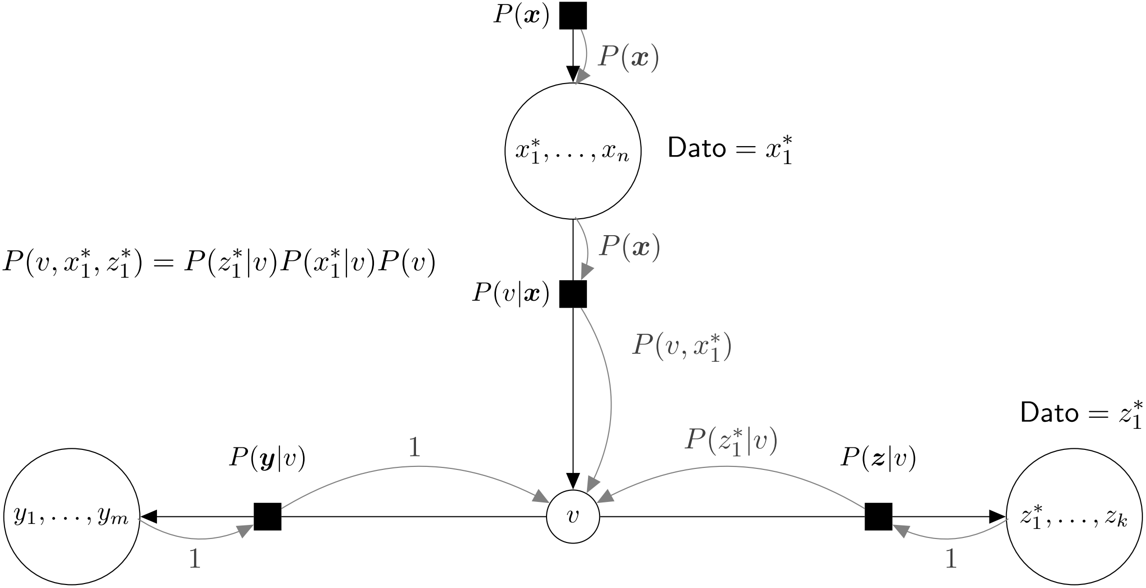
Sum-product algorithm



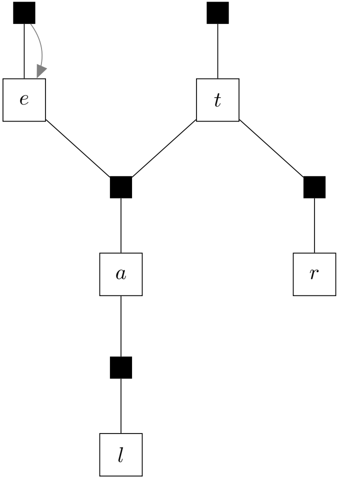
Sum-product algorithm



Sum-product algorithm

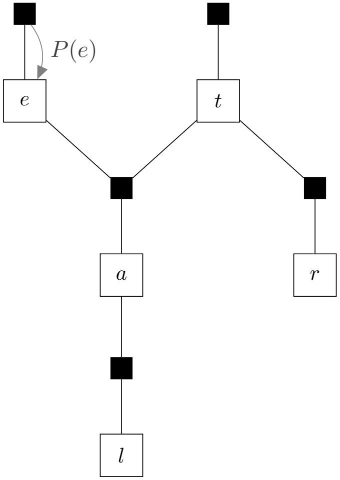


Modelo sin observables



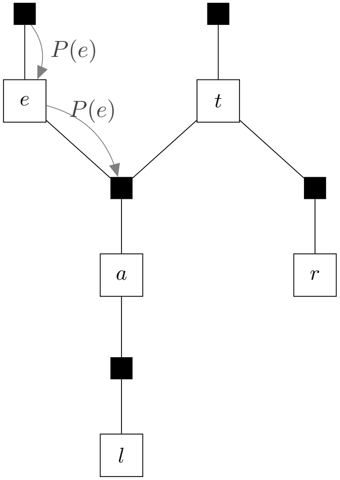
$$m_{f_e \to e}(e) =$$

Modelo sin observables



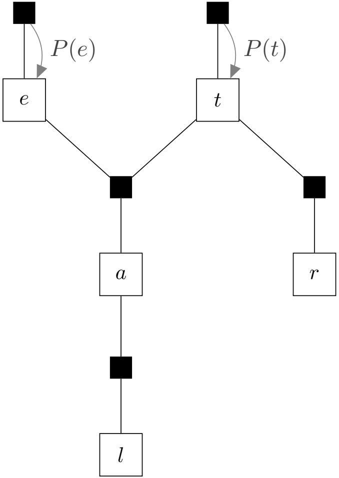
$$m_{f_e \to e}(e) = P(e)$$

Modelo sin observables



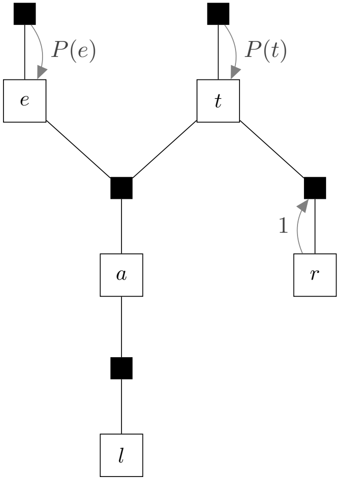
$$m_{f_e \to e}(e) = P(e) = m_{e \to f_a}(e)$$

Modelo sin observables



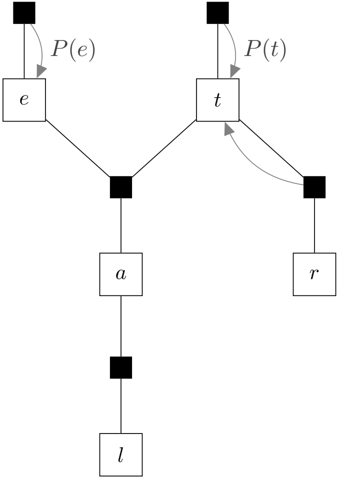
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$

Modelo sin observables



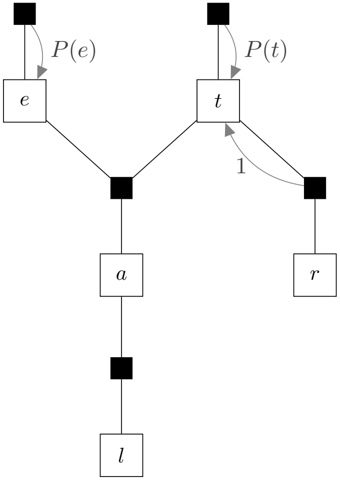
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{r \to f_r}(r) = 1$$

Modelo sin observables



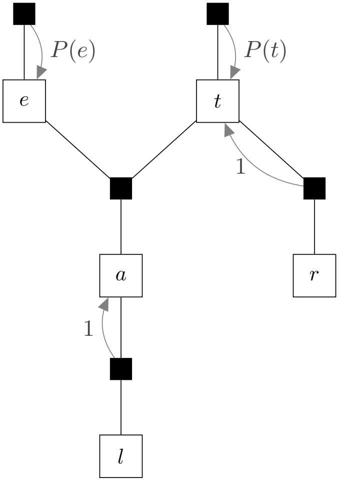
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t)$$

Modelo sin observables



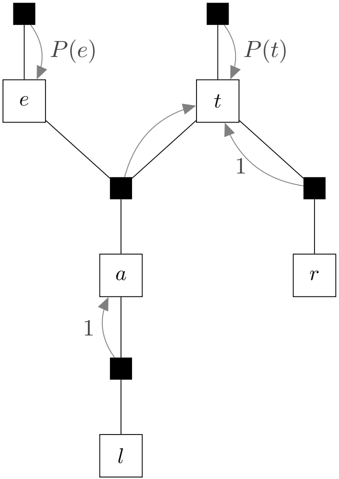
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$

Modelo sin observables



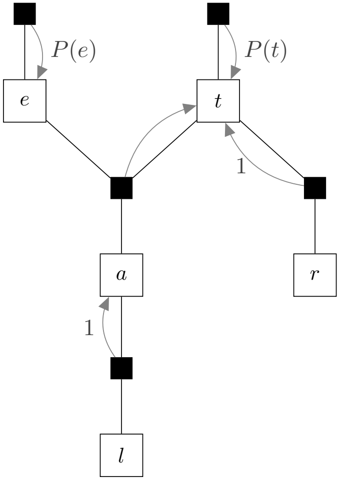
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$

Modelo sin observables



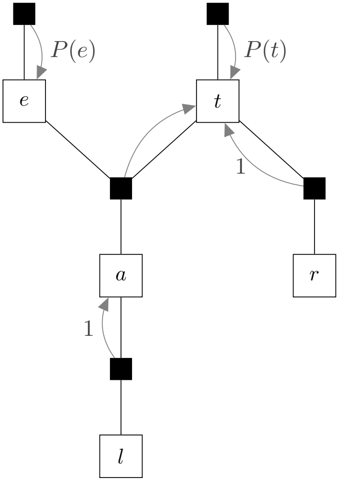
$$m_{f_e \rightarrow e}(e) = P(e)$$
$$m_{f_t \rightarrow t}(t) = P(t)$$
$$m_{f_r \rightarrow t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \rightarrow a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \rightarrow t}(t) = \sum_{ea} P(e)P(a|e,t)$$

Modelo sin observables



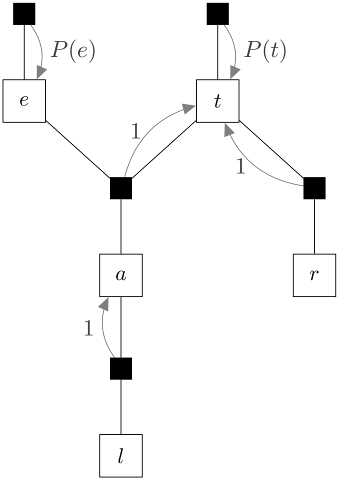
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = \sum_e P(e) \sum_a P(a|e,t)$$

Modelo sin observables



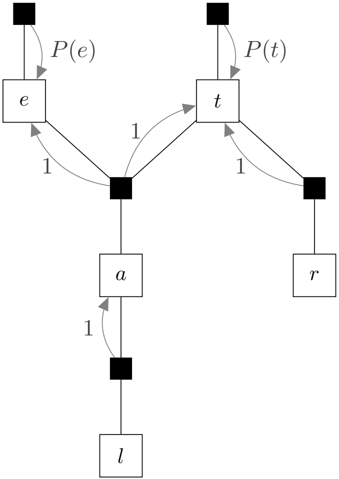
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = \sum_e P(e)$$

Modelo sin observables



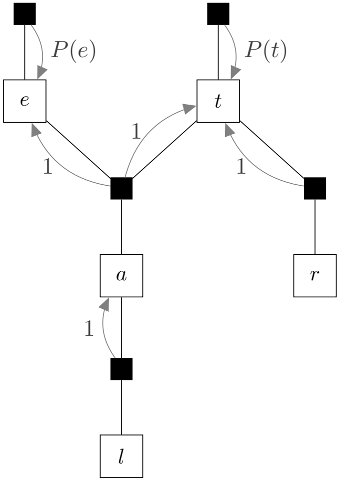
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$

Modelo sin observables



$$m_{f_e \rightarrow e}(e) = P(e)$$
$$m_{f_t \rightarrow t}(t) = P(t)$$
$$m_{f_r \rightarrow t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \rightarrow a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \rightarrow t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \rightarrow e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$

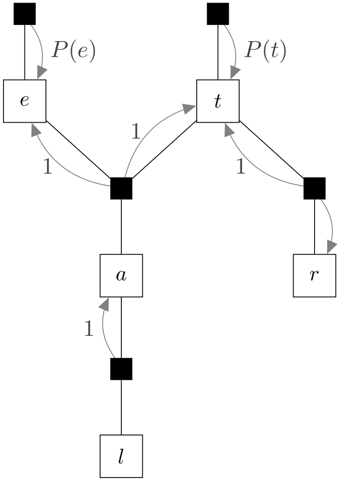
Modelo sin observables



$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$

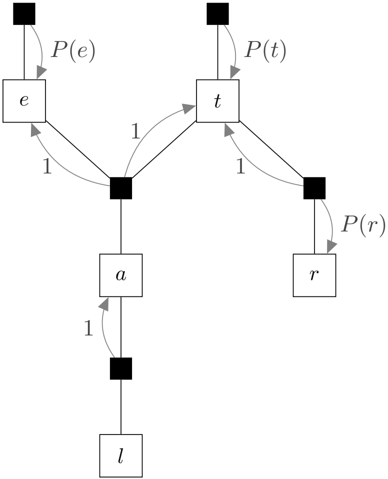
Todos los mensajes
que suben son 1

Modelo sin observables



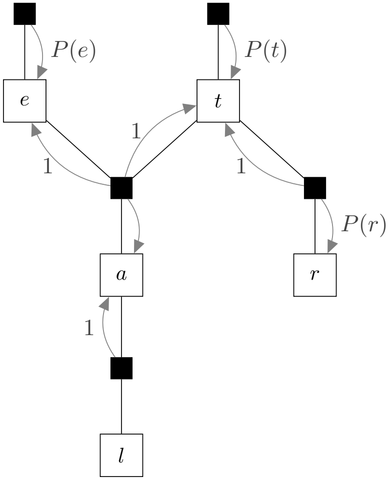
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \textstyle\sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \textstyle\sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \textstyle\sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \textstyle\sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \textstyle\sum_t P(t)P(r|t)$$

Modelo sin observables



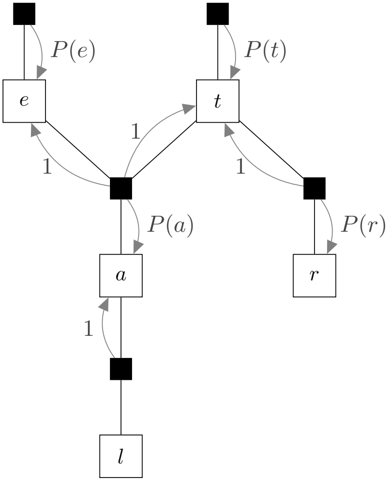
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \sum_t P(t)P(r|t) = P(r)$$

Modelo sin observables



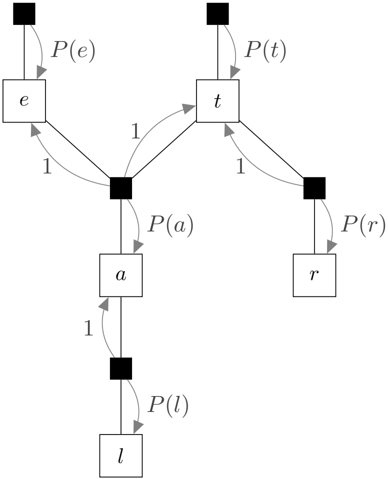
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \sum_t P(t)P(r|t) = P(r)$$
$$m_{f_a \to a}(a) = \sum_{et} P(e)P(t)P(a|e,t)$$

Modelo sin observables



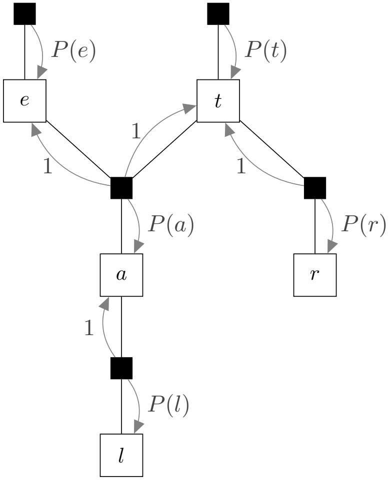
$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \sum_t P(t)P(r|t) = P(r)$$
$$m_{f_a \to a}(a) = \sum_{et} P(e)P(t)P(a|e,t) = P(a)$$

Modelo sin observables



$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \sum_t P(t)P(r|t) = P(r)$$
$$m_{f_a \to a}(a) = \sum_{et} P(e)P(t)P(a|e,t) = P(a)$$
$$m_{f_l \to l}(l) = \sum_a P(a)P(l|a) = P(l)$$

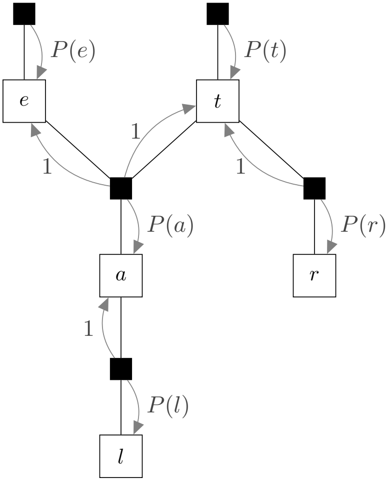
Modelo sin observables



$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \sum_t P(t)P(r|t) = P(r)$$
$$m_{f_a \to a}(a) = \sum_{et} P(e)P(t)P(a|e,t) = P(a)$$
$$m_{f_l \to l}(l) = \sum_a P(a)P(l|a) = P(l)$$

Todos los mensajes que bajan son marginales

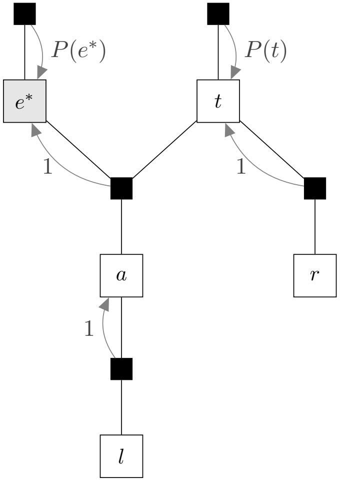
Modelo sin observables



$$m_{f_e \to e}(e) = P(e)$$
$$m_{f_t \to t}(t) = P(t)$$
$$m_{f_r \to t}(t) = \sum_r P(r|t) = 1$$
$$m_{f_l \to a}(a) = \sum_l P(l|a) = 1$$
$$m_{f_a \to t}(t) = \sum_{ea} P(e)P(a|e,t) = 1$$
$$m_{f_a \to e}(e) = \sum_{ta} P(t)P(a|e,t) = 1$$
$$m_{f_r \to r}(r) = \sum_t P(t)P(r|t) = P(r)$$
$$m_{f_a \to a}(a) = \sum_{et} P(e)P(t)P(a|e,t) = P(a)$$
$$m_{f_l \to l}(l) = \sum_a P(a)P(l|a) = P(l)$$

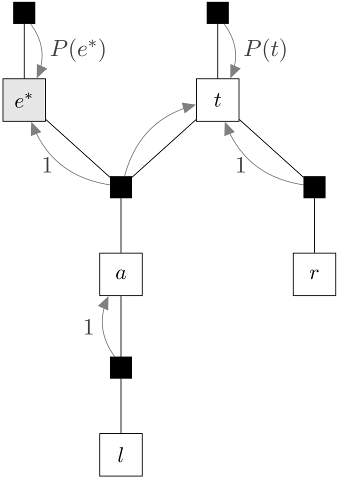
$$P(l=1) \cdot 365 = ?$$

Modelo con observables



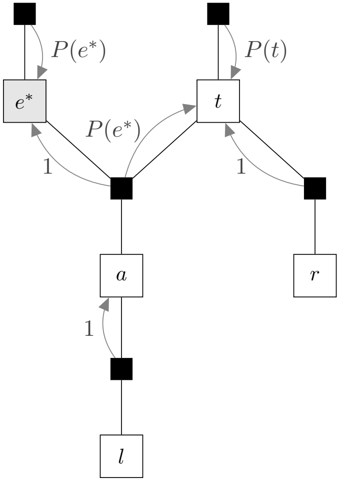
$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

Modelo con observables



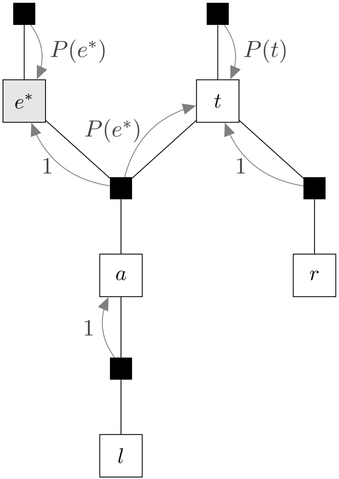
$$
\begin{aligned}
m_{f_e \rightarrow e} &= P(e^*) \\
m_{f_t \rightarrow t}(t) &= P(t) \\
m_{f_r \rightarrow t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \rightarrow a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \rightarrow t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) \\
m_{f_a \rightarrow e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \rightarrow r}(r) &= \\
m_{f_a \rightarrow a}(a) &= \\
m_{f_l \rightarrow l}(l) &=
\end{aligned}
$$

Modelo con observables



$$
\begin{aligned}
m_{f_e \to e} \quad &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*,t) = P(e^*) \\
m_{f_a \to e} \quad &= \textstyle\sum_{ta} P(t)P(a|e^*,t) = 1 \\
m_{f_r \to r}(r) &= \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

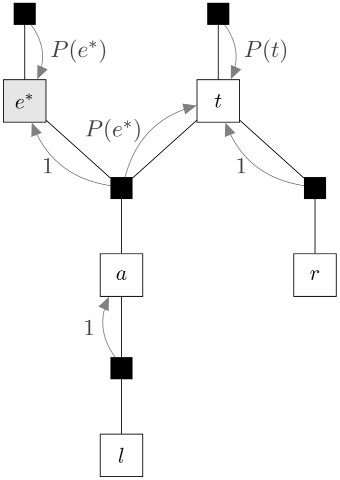
Modelo con observables



$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*,t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*,t) = 1 \\
m_{f_r \to r}(r) &= \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

$$P(t, e^*) = P(e^*)P(t)$$

Modelo con observables

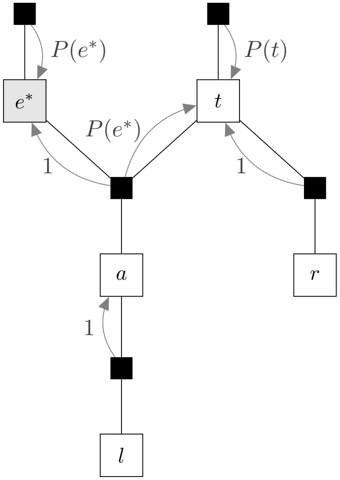


$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*,t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*,t) = 1 \\
m_{f_r \to r}(r) &= \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

$$P(t, e^*) = P(e^*)P(t)$$

$$P(t|e^*) = \frac{P(e^*)P(t)}{P(e^*)}$$

Modelo con observables

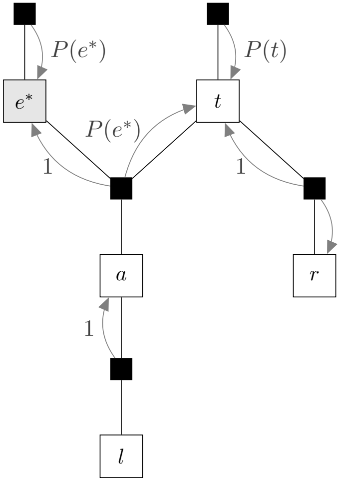


$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*,t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*,t) = 1 \\
m_{f_r \to r}(r) &= \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

$$P(t, e^*) = P(e^*)P(t)$$

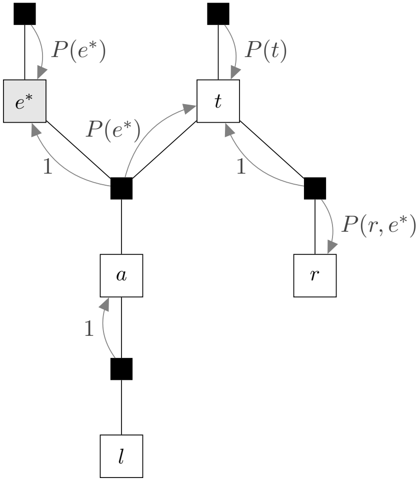
$$P(t|e^*) = \frac{P(e^*)P(t)}{P(e^*)} = P(t)$$

Modelo con observables



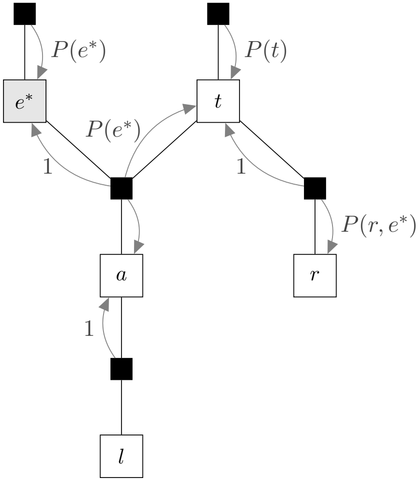
$$
\begin{aligned}
m_{f_e \to e} \quad &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*,t) = P(e^*) \\
m_{f_a \to e} \quad &= \textstyle\sum_{ta} P(t)P(a|e^*,t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

Modelo con observables



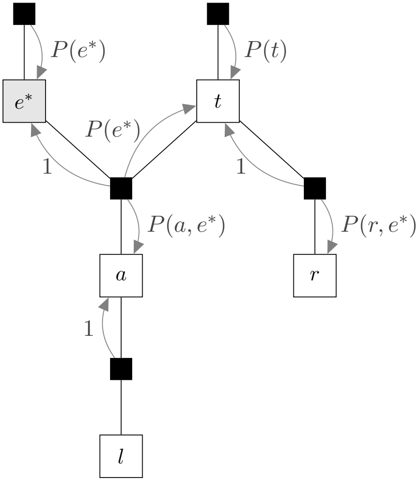
$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

Modelo con observables



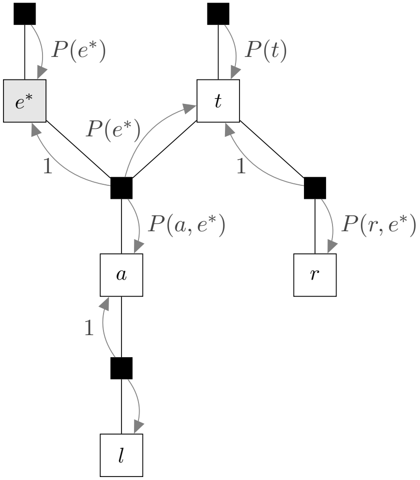
$$
\begin{aligned}
m_{f_e \to e} \quad &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} \quad &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

Modelo con observables



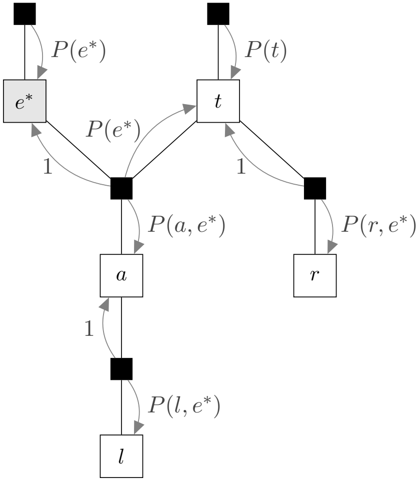
$$
\begin{aligned}
m_{f_e \to e} \quad &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} \quad &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) = P(a, e^*) \\
m_{f_l \to l}(l) &=
\end{aligned}
$$

Modelo con observables



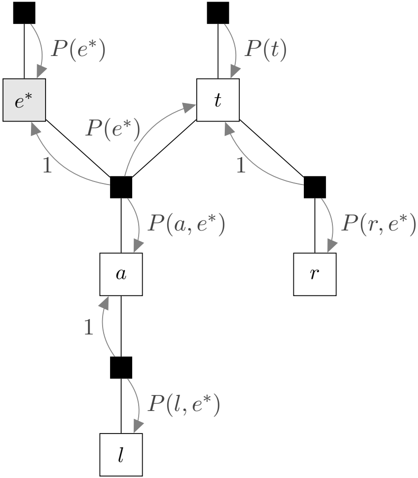
$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) = P(a, e^*) \\
m_{f_l \to l}(l) &= \textstyle\sum_a P(a, e^*)P(l|a)
\end{aligned}
$$

Modelo con observables



$$
\begin{aligned}
m_{f_e \to e} \quad &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} \quad &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) = P(a, e^*) \\
m_{f_l \to l}(l) &= \textstyle\sum_a P(a, e^*)P(l|a) = P(l, e^*)
\end{aligned}
$$

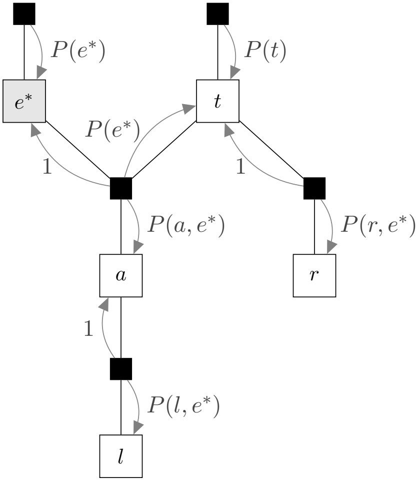
Modelo con observables



$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) = P(a, e^*) \\
m_{f_l \to l}(l) &= \textstyle\sum_a P(a, e^*)P(l|a) = P(l, e^*)
\end{aligned}
$$

$$P(l|e^*) = P(l, e^*)/P(e^*)$$

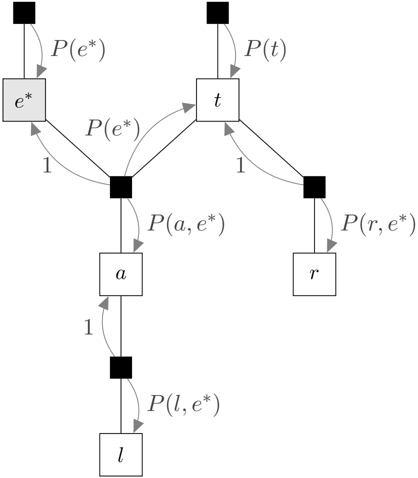
Modelo con observables



$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) = P(a, e^*) \\
m_{f_l \to l}(l) &= \textstyle\sum_a P(a, e^*)P(l|a) = P(l, e^*)
\end{aligned}
$$

$$P(l|e^*) = P(l, e^*)/P(e^*) \overset{?}{=} P(l)$$

Modelo con observables



$$
\begin{aligned}
m_{f_e \to e} &= P(e^*) \\
m_{f_t \to t}(t) &= P(t) \\
m_{f_r \to t}(t) &= \textstyle\sum_r P(r|t) = 1 \\
m_{f_l \to a}(a) &= \textstyle\sum_l P(l|a) = 1 \\
m_{f_a \to t}(t) &= \textstyle\sum_a P(e^*)P(a|e^*, t) = P(e^*) \\
m_{f_a \to e} &= \textstyle\sum_{ta} P(t)P(a|e^*, t) = 1 \\
m_{f_r \to r}(r) &= \textstyle\sum_t P(t)P(e^*)P(r|t) = P(r)P(e^*) = P(r, e^*) \\
m_{f_a \to a}(a) &= \textstyle\sum_t P(e^*)P(t)P(a|e^*, t) = P(a, e^*) \\
m_{f_l \to l}(l) &= \textstyle\sum_a P(a, e^*)P(l|a) = P(l, e^*)
\end{aligned}
$$

$$\underbrace{P(l|e^*)}_{\substack{\text{que llamen} \\ \text{cuando entran}}} = P(l, e^*)/P(e^*) \overset{?}{=} \underbrace{P(l)}_{\text{que llamen}}$$

Ejercicio
Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \rightarrow a \rightarrow l$		
$l \leftarrow a \leftarrow t$		
$a \leftarrow t \rightarrow r$		
$\begin{tikzcd} e \arrow[r] & a \arrow[d] & t \arrow[l] \\ & l & \end{tikzcd}$		

Ejercicio

Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \to a \to l$	$P(l) \stackrel{?}{=} P(l\|e)$	
$l \leftarrow a \leftarrow t$		
$a \leftarrow t \to r$		
$\begin{array}{c} e \to a \leftarrow t \\ \downarrow \\ l \end{array}$		

Ejercicio

Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e,a)$
$l \leftarrow a \leftarrow t$		
$a \leftarrow t \rightarrow r$		
$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$		

Ejercicio
Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \to a \to l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e, a)$
$l \leftarrow a \leftarrow t$	$P(t) \stackrel{?}{=} P(t\|l)$	
$a \leftarrow t \to r$		
$\begin{array}{ccccc} e & \to & a & \leftarrow & t \\ & & \downarrow & & \\ & & l & & \end{array}$		

Ejercicio

Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \to a \to l$	$P(l) \overset{?}{=} P(l\|e)$	$P(l\|a) \overset{?}{=} P(l\|e,a)$
$l \leftarrow a \leftarrow t$	$P(t) \overset{?}{=} P(t\|l)$	$P(t\|a) \overset{?}{=} P(t\|a,l)$
$a \leftarrow t \to r$		
$\begin{array}{c} e \to a \leftarrow t \\ \downarrow \\ l \end{array}$		

Ejercicio
Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \to a \to l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e, a)$
$l \leftarrow a \leftarrow t$	$P(t) \stackrel{?}{=} P(t\|l)$	$P(t\|a) \stackrel{?}{=} P(t\|a, l)$
$a \leftarrow t \to r$	$P(r) \stackrel{?}{=} P(r\|a)$	
$\begin{array}{ccccc} e & \to & a & \leftarrow & t \\ & & \downarrow & & \\ & & l & & \end{array}$		

Ejercicio

Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \to a \to l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e, a)$
$l \leftarrow a \leftarrow t$	$P(t) \stackrel{?}{=} P(t\|l)$	$P(t\|a) \stackrel{?}{=} P(t\|a, l)$
$a \leftarrow t \to r$	$P(r) \stackrel{?}{=} P(r\|a)$	$P(r\|t) \stackrel{?}{=} P(r\|t, a)$
$\begin{array}{c} e \to a \leftarrow t \\ \downarrow \\ l \end{array}$		

Ejercicio
Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \to a \to l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e,a)$
$l \leftarrow a \leftarrow t$	$P(t) \stackrel{?}{=} P(t\|l)$	$P(t\|a) \stackrel{?}{=} P(t\|a,l)$
$a \leftarrow t \to r$	$P(r) \stackrel{?}{=} P(r\|a)$	$P(r\|t) \stackrel{?}{=} P(r\|t,a)$
$\begin{array}{c} e \to a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	

Ejercicio

Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l \mid e)$	$P(l \mid a) \stackrel{?}{=} P(l \mid e, a)$
$l \leftarrow a \leftarrow t$	$P(t) \stackrel{?}{=} P(t \mid l)$	$P(t \mid a) \stackrel{?}{=} P(t \mid a, l)$
$a \leftarrow t \rightarrow r$	$P(r) \stackrel{?}{=} P(r \mid a)$	$P(r \mid t) \stackrel{?}{=} P(r \mid t, a)$
$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t \mid e)$	$P(t \mid a) \stackrel{?}{=} P(t \mid e, a)$

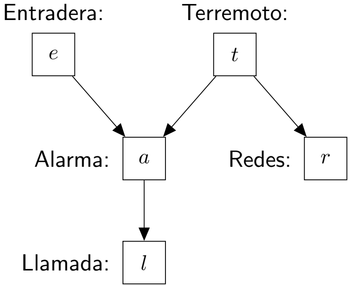
Ejercicio

Flujos de inferencia

	Intermedio no observable	Intermedio observable
$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e,a)$
$l \leftarrow a \leftarrow t$	$P(t) \stackrel{?}{=} P(t\|l)$	$P(t\|a) \stackrel{?}{=} P(t\|a,l)$
$a \leftarrow t \rightarrow r$	$P(r) \stackrel{?}{=} P(r\|a)$	$P(r\|t) \stackrel{?}{=} P(r\|t,a)$
$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

Flujo de inferencia

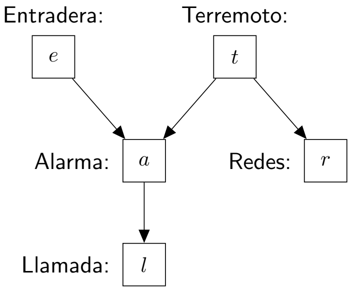
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \stackrel{?}{=} P(r\|a)$	$P(r\|t) \stackrel{?}{=} P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

Flujo de inferencia

Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \overset{?}{=} P(r \mid a)$	$P(r \mid t) \overset{?}{=} P(r \mid t, a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \overset{?}{=} P(l \mid e)$	$P(l \mid a) \overset{?}{=} P(l \mid e, a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \overset{?}{=} P(t \mid e)$	$P(t \mid a) \overset{?}{=} P(t \mid e, a)$ $P(t \mid l) \overset{?}{=} P(t \mid e, l)$

Flujo de inferencia

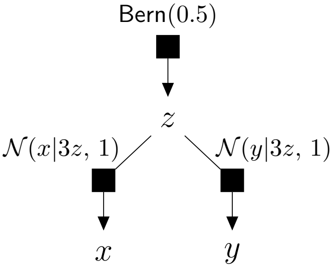
Fork

$$\begin{tikzcd} x & z \arrow[l] \arrow[r] & y \end{tikzcd}$$

Flujo de inferencia

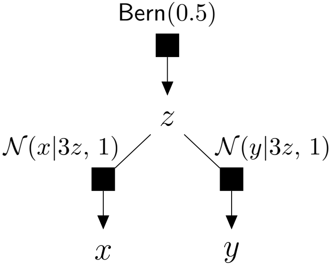
Fork

$$x \longleftarrow z \longrightarrow y$$

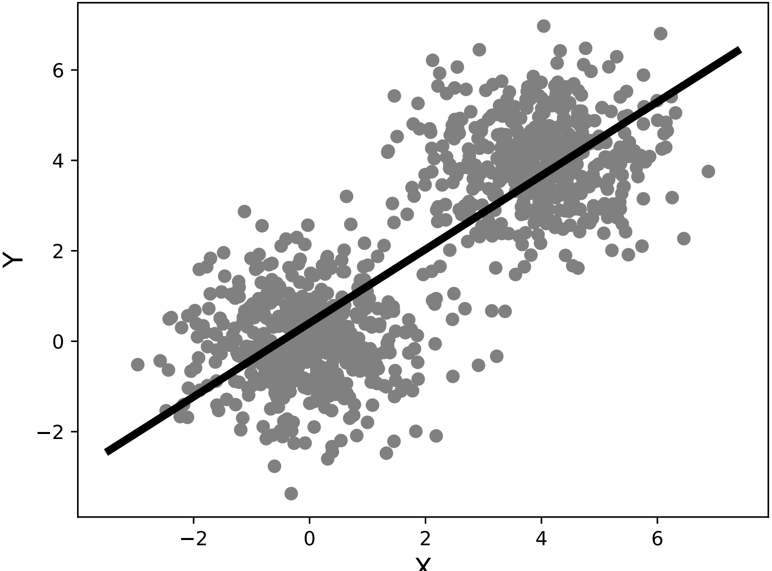


Flujo de inferencia

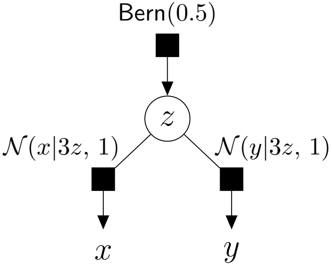
Fork



$$X \not\perp\!\!\!\perp Y|\emptyset$$

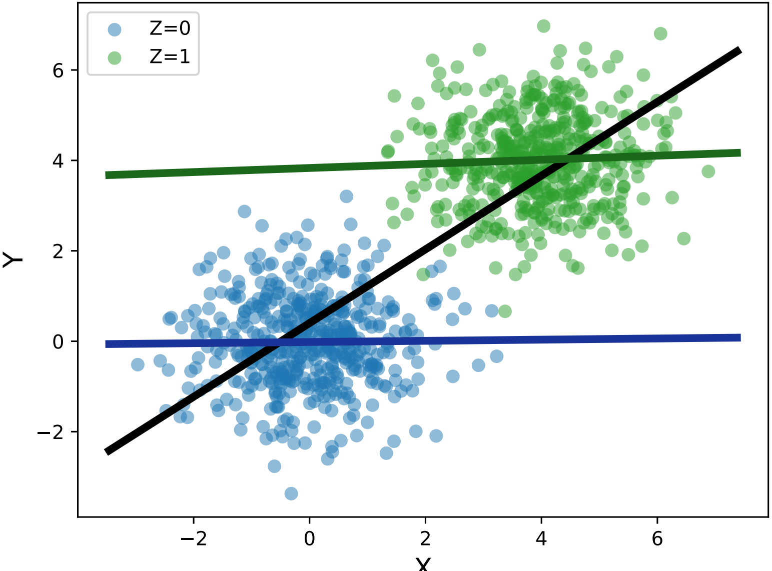


Flujo de inferencia

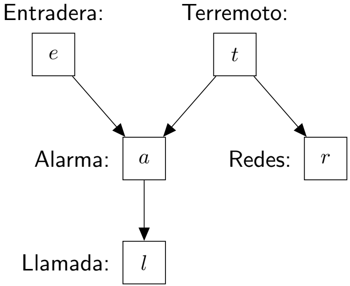


$$X \perp\!\!\!\perp Y | Z$$

Fork

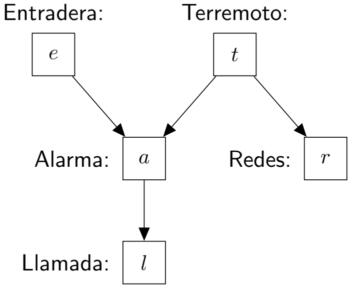


Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \overset{?}{=} P(r\|a)$	$P(r\|t) \overset{?}{=} P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \overset{?}{=} P(l\|e)$	$P(l\|a) \overset{?}{=} P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \overset{?}{=} P(t\|e)$	$P(t\|a) \overset{?}{=} P(t\|e,a)$ $P(t\|l) \overset{?}{=} P(t\|e,l)$

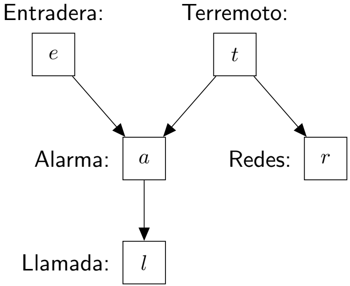
Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r \mid a)$	$P(r \mid t) \stackrel{?}{=} P(r \mid t, a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l \mid e)$	$P(l \mid a) \stackrel{?}{=} P(l \mid e, a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t \mid e)$	$P(t \mid a) \stackrel{?}{=} P(t \mid e, a)$ $P(t \mid l) \stackrel{?}{=} P(t \mid e, l)$

$$A \not\perp\!\!\!\perp R \mid \emptyset$$

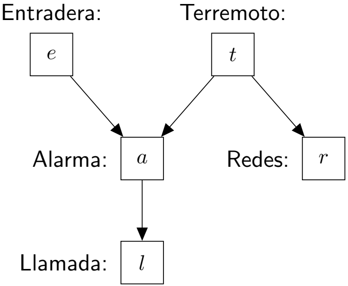
Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$
		$A \not\perp\!\!\!\perp R\|\emptyset$	$A \perp\!\!\!\perp R\|T$

Flujo de inferencia

Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \overset{?}{=} P(l\|e)$	$P(l\|a) \overset{?}{=} P(l\|e,a)$
Collider:	$e \rightarrow a \leftarrow t$, $a \rightarrow l$	$P(t) \overset{?}{=} P(t\|e)$	$P(t\|a) \overset{?}{=} P(t\|e,a)$ $P(t\|l) \overset{?}{=} P(t\|e,l)$

Flujo de inferencia

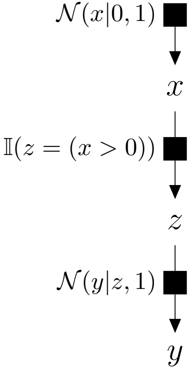
Pipe

$$x \longrightarrow z \longrightarrow y$$

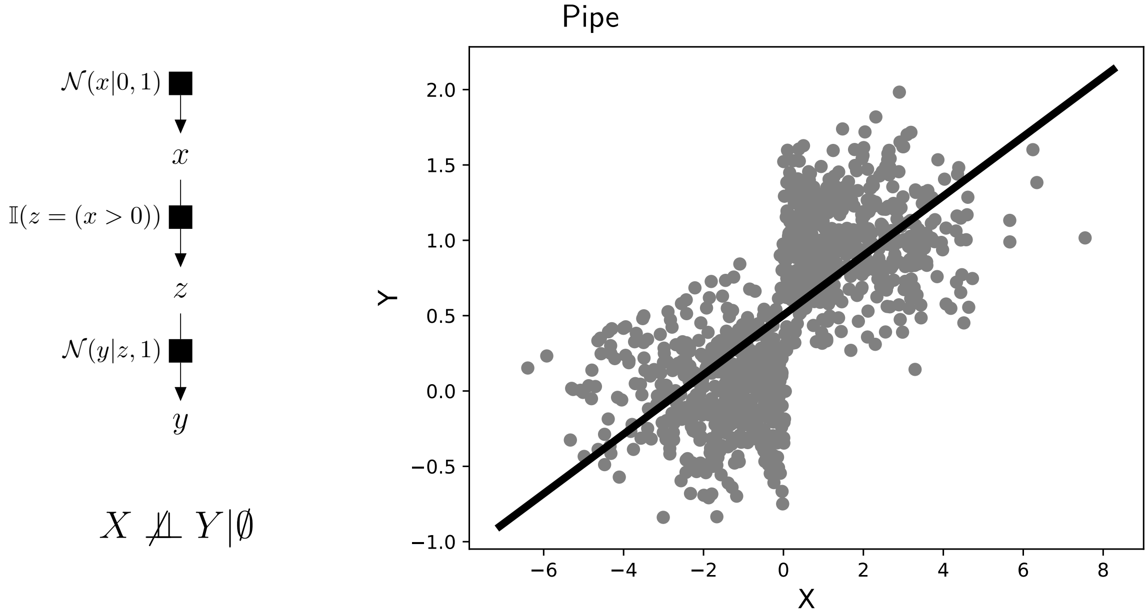
Flujo de inferencia

Pipe

$$x \longrightarrow z \longrightarrow y$$

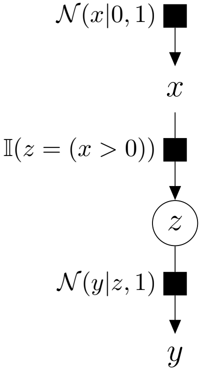


Flujo de inferencia

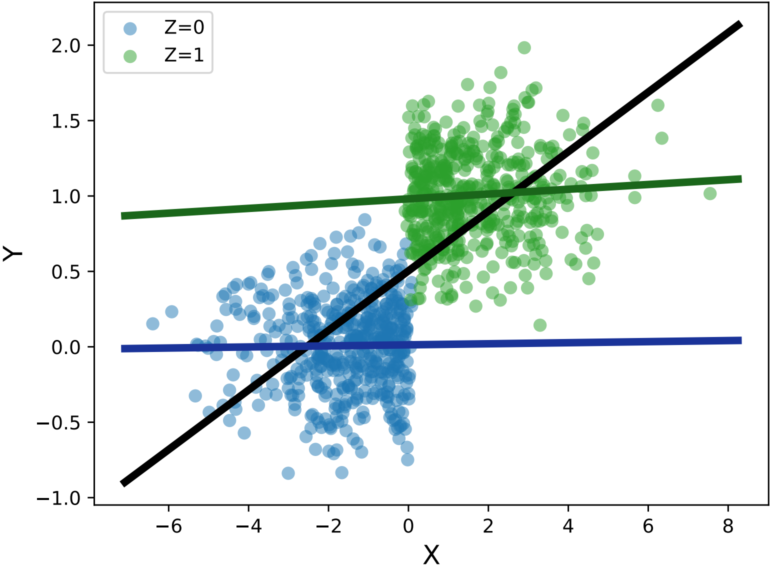


Flujo de inferencia

Pipe

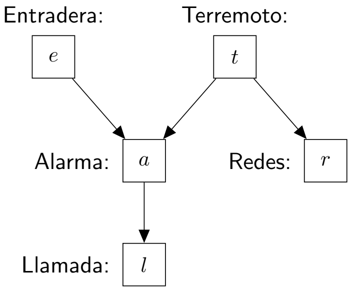


$$X \perp\!\!\!\perp Y | Z$$



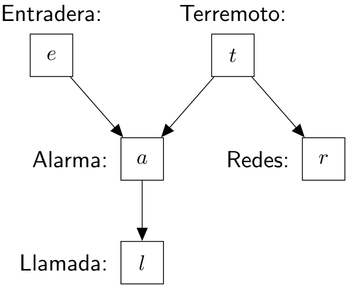
Flujo de inferencia

Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \stackrel{?}{=} P(l\|e)$	$P(l\|a) \stackrel{?}{=} P(l\|e,a)$
Collider:	$\begin{array}{ccccc} e & \rightarrow & a & \leftarrow & t \\ & & \downarrow & & \\ & & l & & \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

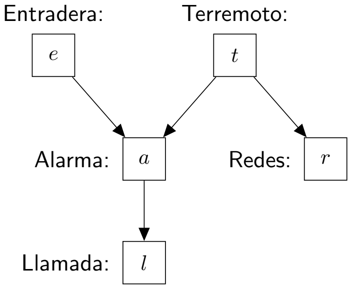
Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r \vert a)$	$P(r \vert t) = P(r \vert t, a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l \vert e)$	$P(l \vert a) \overset{?}{=} P(l \vert e, a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \overset{?}{=} P(t \vert e)$	$P(t \vert a) \overset{?}{=} P(t \vert e, a)$ $P(t \vert l) \overset{?}{=} P(t \vert e, l)$

$$E \not\perp\!\!\!\perp L | \emptyset$$

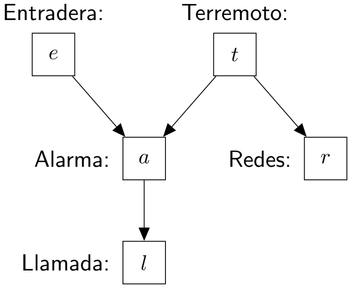
Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l\|e)$	$P(l\|a) = P(l\|e,a)$
Collider:	$e \rightarrow a \leftarrow t$, $a \rightarrow l$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

$$E \not\perp\!\!\!\perp L|\emptyset \qquad E \perp\!\!\!\perp L|A$$

Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l\|e)$	$P(l\|a) = P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

Flujo de inferencia

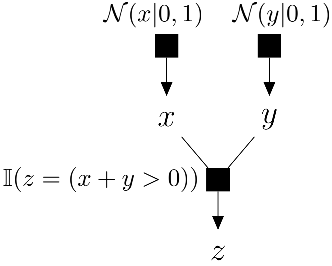
Collider

$$x \longrightarrow z \longleftarrow y$$

Flujo de inferencia

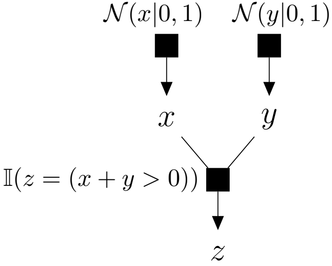
Collider

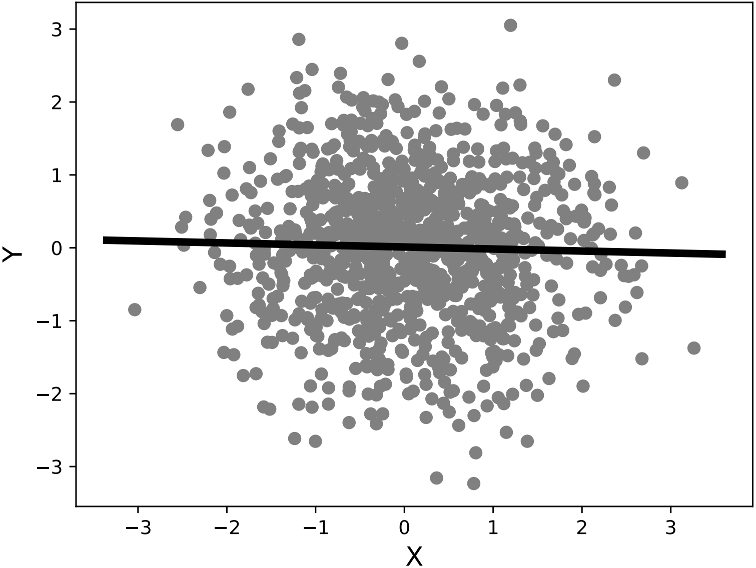
$$x \longrightarrow z \longleftarrow y$$



Flujo de inferencia

Collider

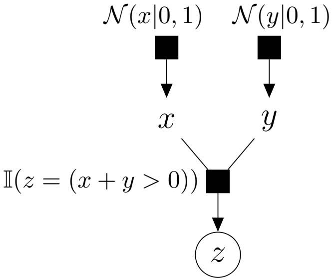




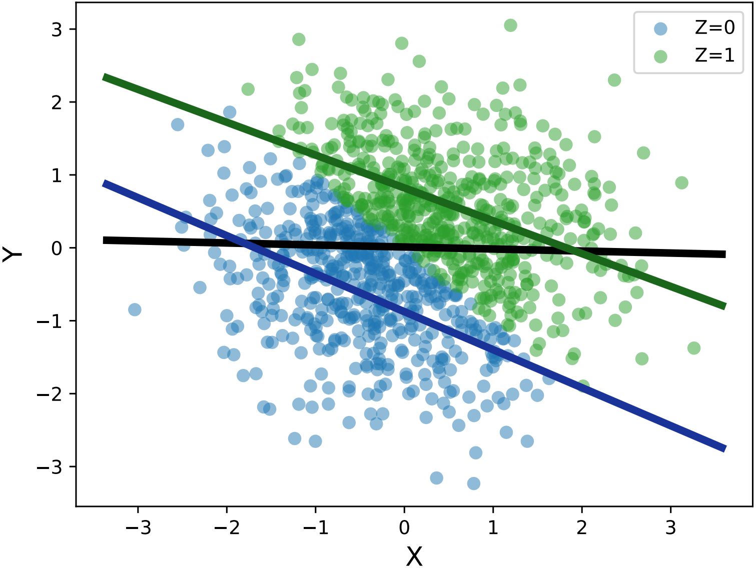
$$X \perp\!\!\!\perp Y|\emptyset$$

Flujo de inferencia

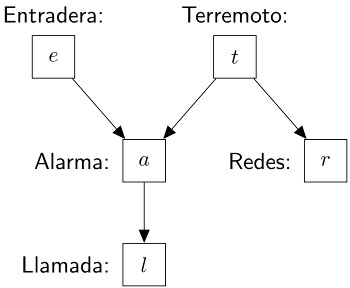
Collider



$$X \not\perp\!\!\!\perp Y | Z$$



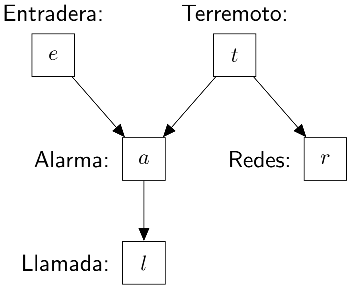
Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l\|e)$	$P(l\|a) = P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) \stackrel{?}{=} P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

Flujo de inferencia

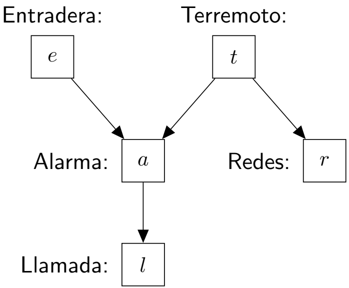
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l\|e)$	$P(l\|a) = P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) = P(t\|e)$	$P(t\|a) \stackrel{?}{=} P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$

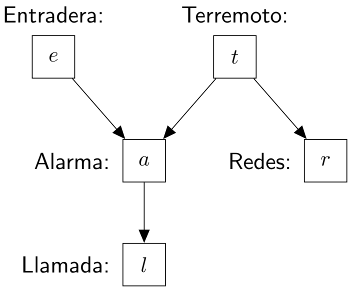
$$T \perp\!\!\!\perp E | \emptyset$$

Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l\|e)$	$P(l\|a) = P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) = P(t\|e)$	$P(t\|a) \neq P(t\|e,a)$ $P(t\|l) \stackrel{?}{=} P(t\|e,l)$
		$T \perp\!\!\!\perp E \| \emptyset$	$T \not\perp\!\!\!\perp E \| A$

Flujo de inferencia
Estructuras básicas



		Intermedio no observable	Intermedio observable
Fork:	$a \leftarrow t \rightarrow r$	$P(r) \neq P(r\|a)$	$P(r\|t) = P(r\|t,a)$
Pipe:	$e \rightarrow a \rightarrow l$	$P(l) \neq P(l\|e)$	$P(l\|a) = P(l\|e,a)$
Collider:	$\begin{array}{c} e \rightarrow a \leftarrow t \\ \downarrow \\ l \end{array}$	$P(t) = P(t\|e)$	$P(t\|a) \neq P(t\|e,a)$ $P(t\|l) \neq P(t\|e,l)$
		$T \perp\!\!\!\perp E \| \emptyset$	$T \not\!\perp\!\!\!\perp E \| L$

Flujo de inferencia

Estructuras complejas (caminos)

Flujo de inferencia

Estructuras complejas (caminos)

$$X \longleftarrow Z \longrightarrow Y'$$

Fork

$$X \longleftarrow A \longrightarrow C \longrightarrow Y'$$

Fork + **Pipe**

$$X \longleftarrow A \longrightarrow C \longleftarrow B \longrightarrow Y'$$

Fork + **Collider** + **Fork**

Flujo de inferencia
Estructuras complejas (caminos)

Hay flujo de inferencia entre los extremos de una cadena si:
(camino *d-conectado*)

- Todas las consecuencias comunes (o sus descendientes) son observables
- Ninguna otra variable es observable

Flujo de inferencia

Estructuras complejas (caminos)

Hay flujo de inferencia entre los extremos de una cadena si:

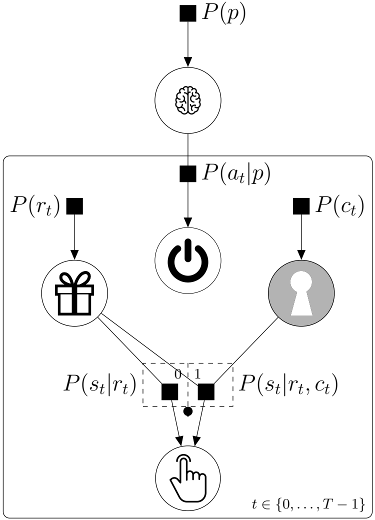
(camino *d-conectado*)

- Todas las consecuencias comunes (o sus descendientes) son observables
- Ninguna otra variable es observable

Se cierra el flujo si está $\underbrace{\text{no } \textit{d-conectado}}_{\text{d-separado}}$

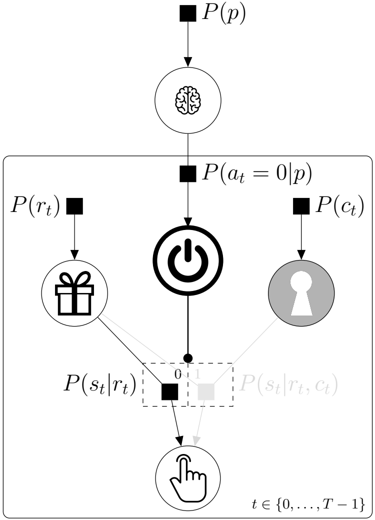
Mecanismos causales dinámicos

Gates



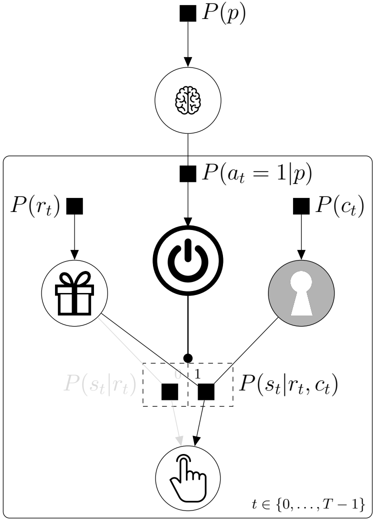
Mecanismos causales dinámicos

Gates



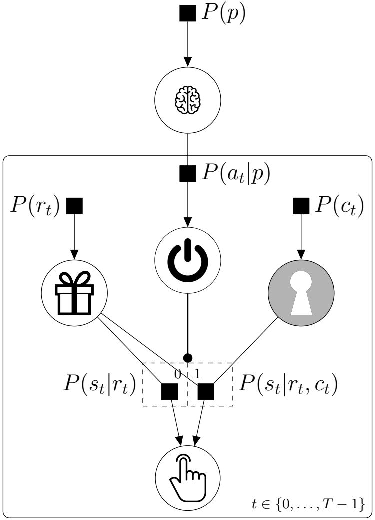
Mecanismos causales dinámicos

Gates



Mecanismos causales dinámicos

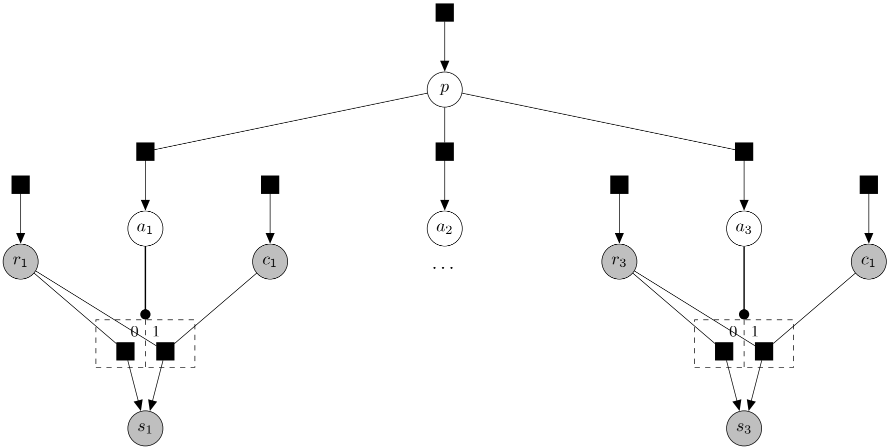
Gates



$$P(s_t|r_t, c_t, a_t) = P(s_t|r_t)^{\mathbb{I}(a_t=0)} P(s_t|r_t, c_t)^{\mathbb{I}(a_t=1)}$$

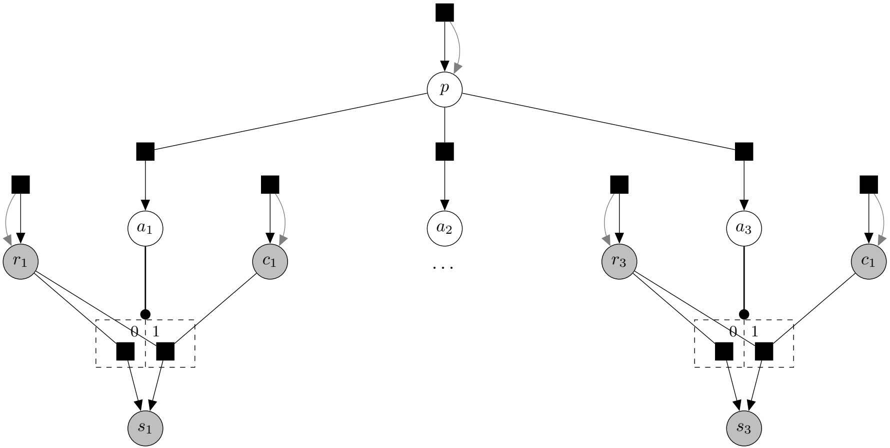
Flujo de inferencia

estructuras causales dinámicas



Flujo de inferencia

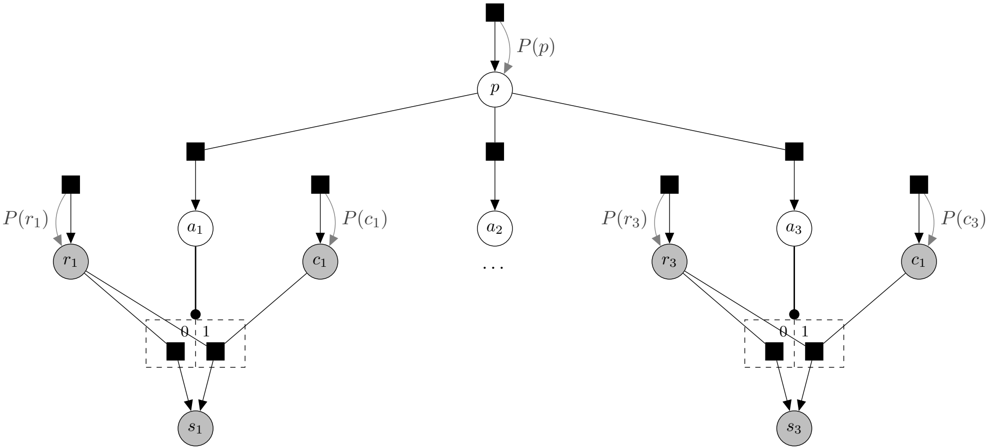
estructuras causales dinámicas



$m_{f_r \to r}(r) =$ $\quad$ $m_{f_p \to p}(p) =$ $\quad$ $m_{f_c \to c}(c) =$

Flujo de inferencia

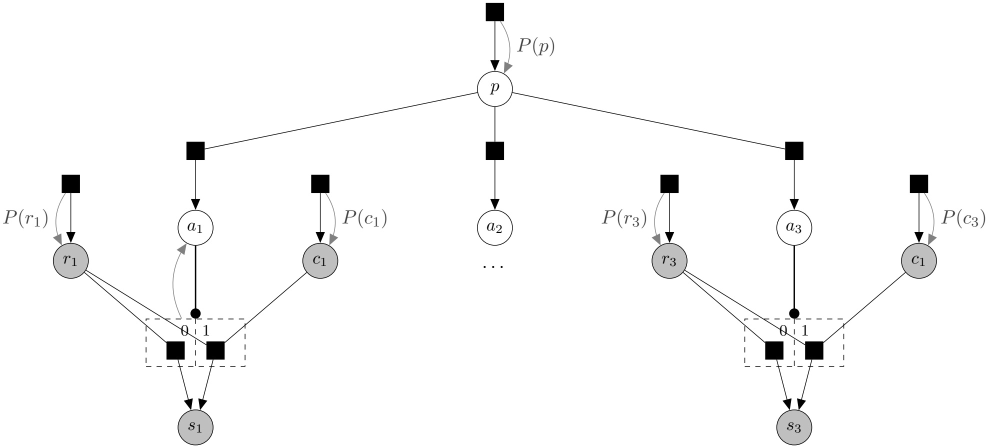
estructuras causales dinámicas



$$m_{f_r \to r}(r) = P(r) \qquad m_{f_p \to p}(p) = P(p) \qquad m_{f_c \to c}(c) = P(c)$$

Flujo de inferencia

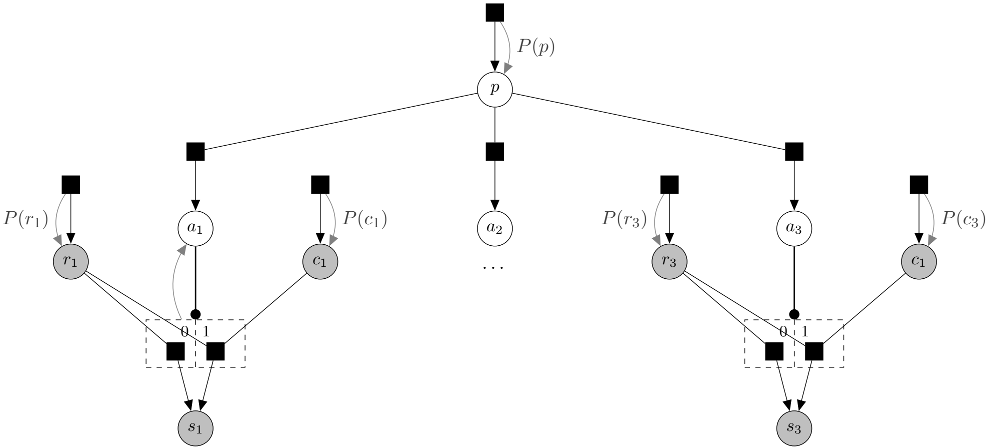
estructuras causales dinámicas



$$m_{f_s \rightarrow a}(a) =$$

Flujo de inferencia

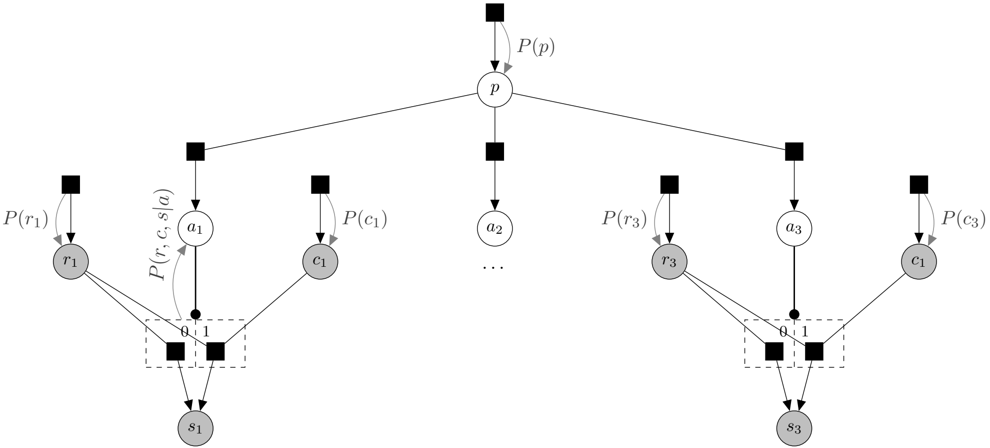
estructuras causales dinámicas



$$m_{f_s \to a}(a) = P(r)P(c)P(s|r)^{\mathbb{I}(a=0)}P(s|r,c)^{\mathbb{I}(a=1)}$$

Flujo de inferencia

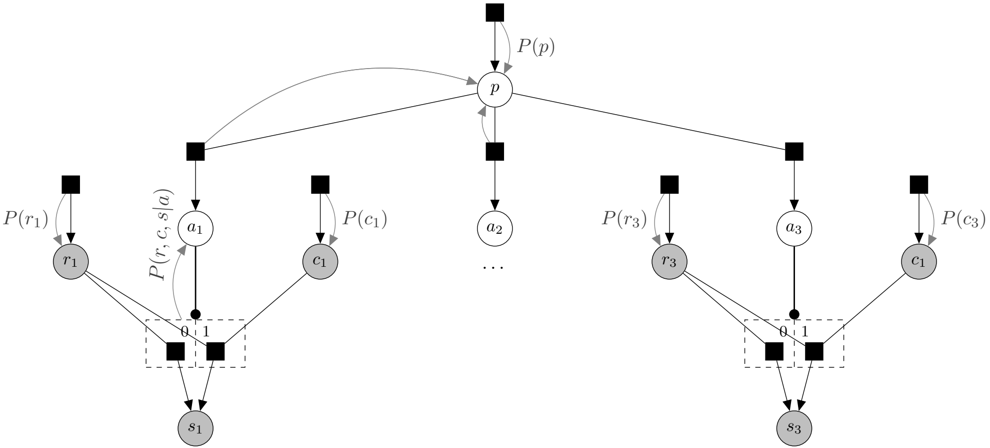
estructuras causales dinámicas



$$m_{f_s \to a}(a) = P(r)P(c)P(s|r)^{\mathbb{I}(a=0)}P(s|r,c)^{\mathbb{I}(a=1)} = P(r,c,s|a)$$

Flujo de inferencia

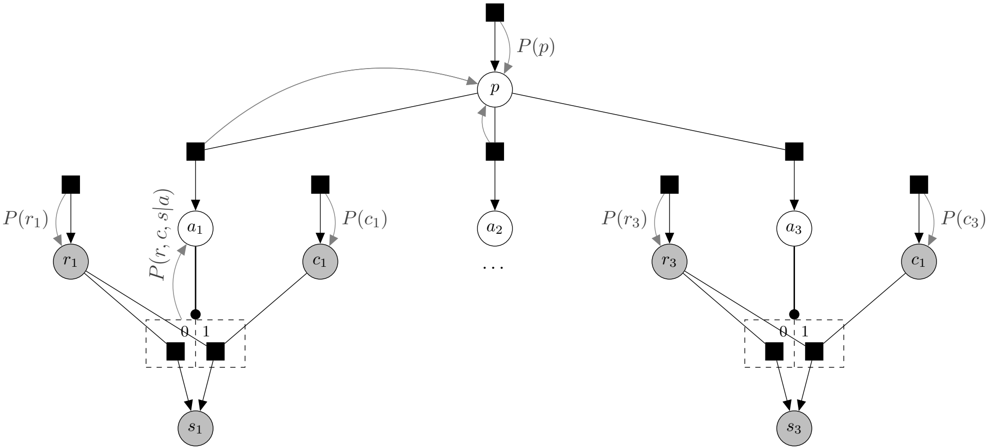
estructuras causales dinámicas



$$m_{f_a \to p}(p) =$$

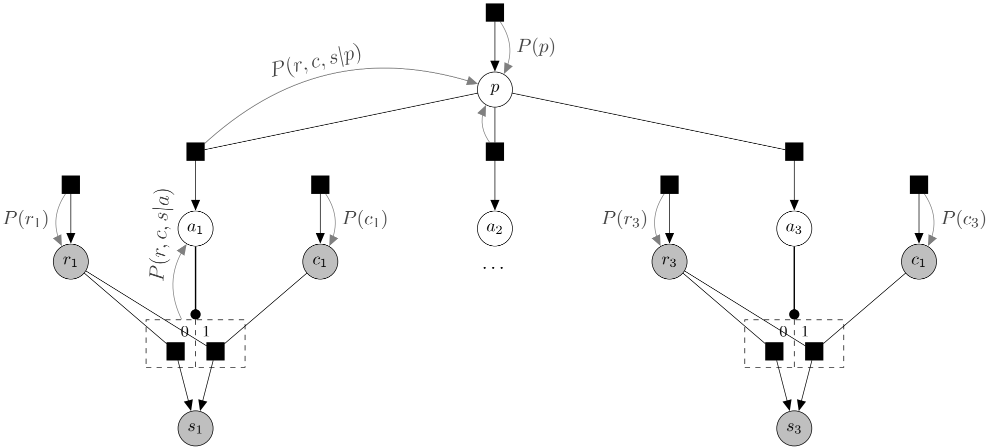
Flujo de inferencia

estructuras causales dinámicas



$$m_{f_a \to p}(p) = \sum_a P(r, c, s|a) P(a|p)$$

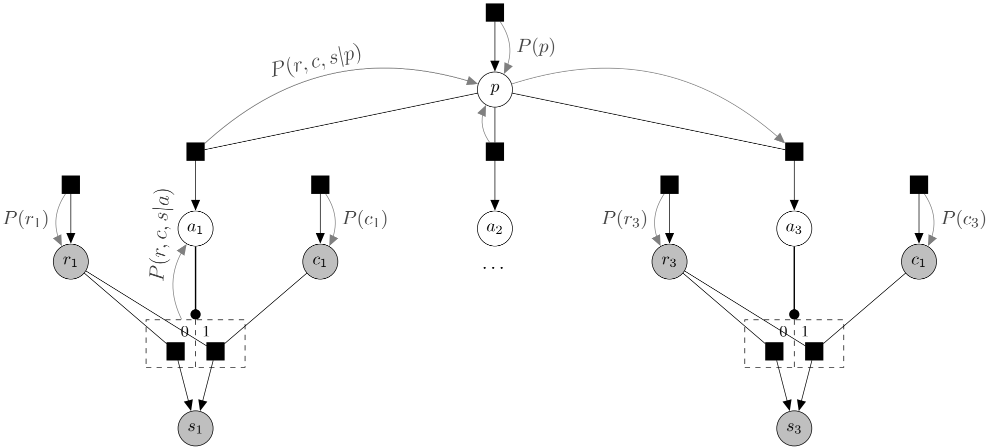
Flujo de inferencia
estructuras causales dinámicas



$$m_{f_a \to p}(p) = \sum_a P(r,c,s|a)P(a|p) = P(r,c,s|p)$$

Flujo de inferencia

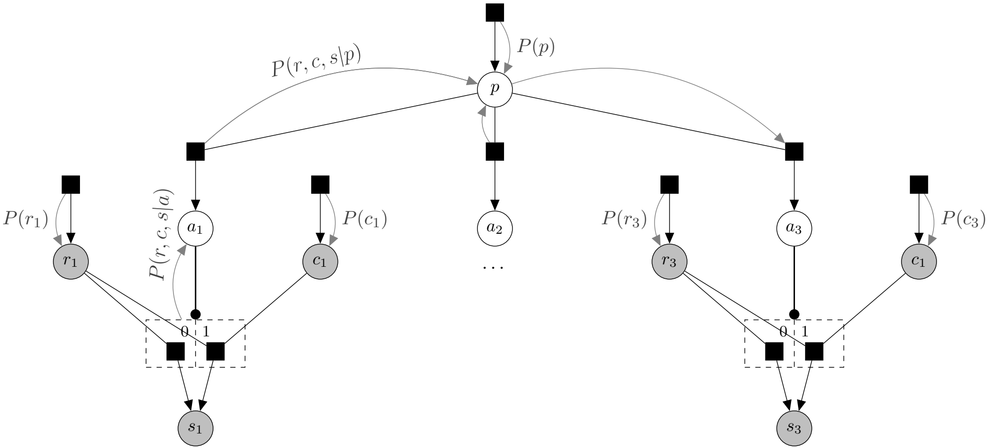
estructuras causales dinámicas



$$m_{p\rightarrow f_{a_{T+1}}}(p) =$$

Flujo de inferencia

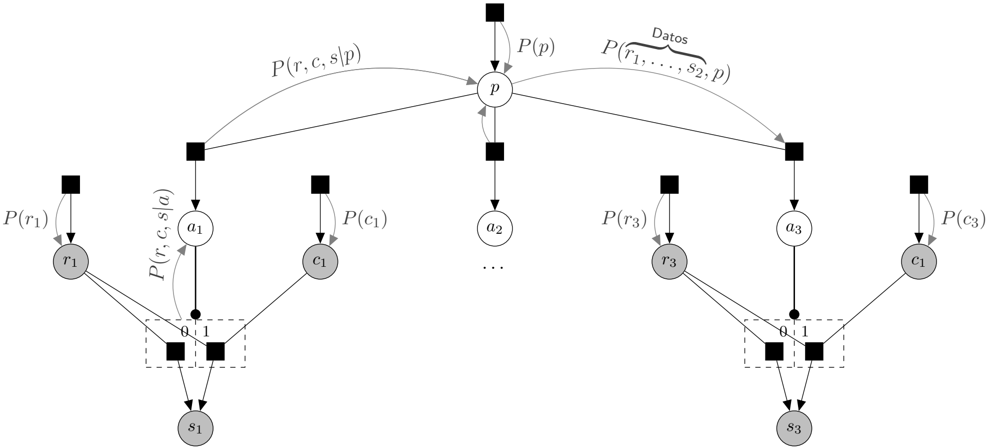
estructuras causales dinámicas



$$m_{p\rightarrow f_{a_{T+1}}}(p) = P(p)\prod_{i=1}^{T} P(r_i, c_i, s_i|p)$$

Flujo de inferencia

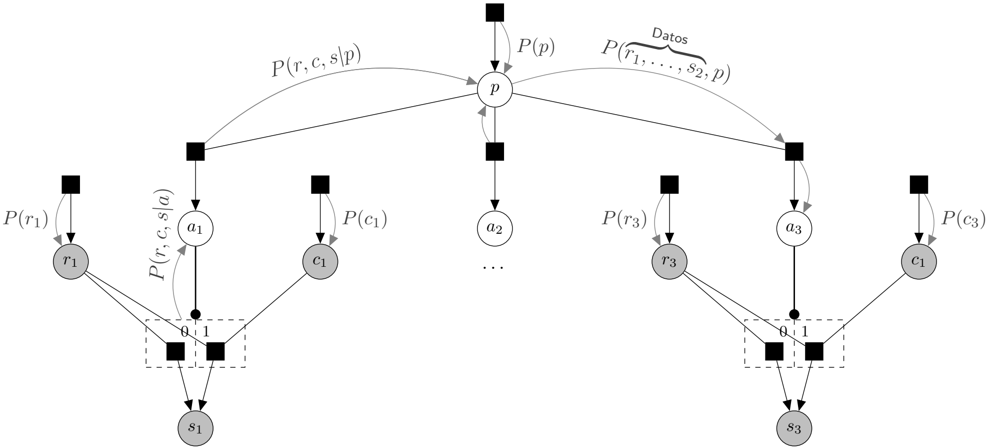
estructuras causales dinámicas



$$m_{p\rightarrow f_{a_{T+1}}}(p) = P(p)\prod_{i=1}^{T} P(r_i, c_i, s_i|p) = P(r_1, c_1, \dots, c_T, s_T, p)$$

Flujo de inferencia

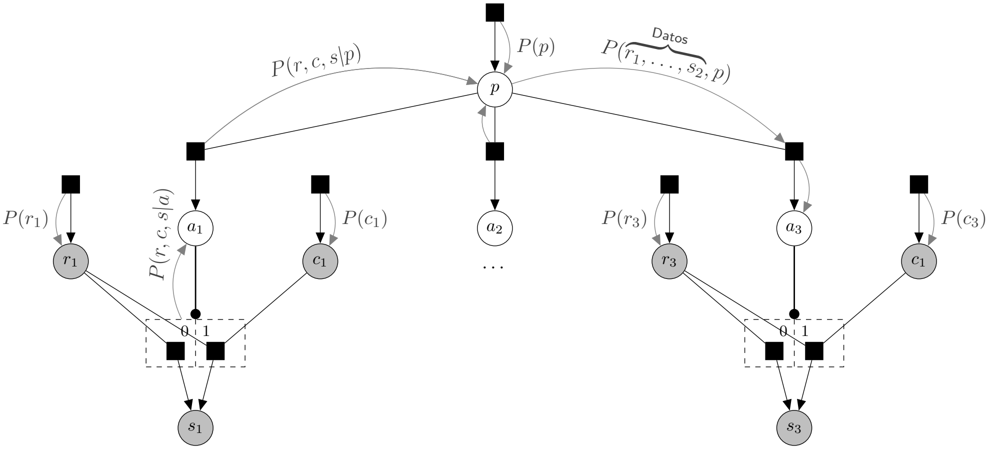
estructuras causales dinámicas



$$m_{f_{a_3}\to a_3}(a_3) =$$

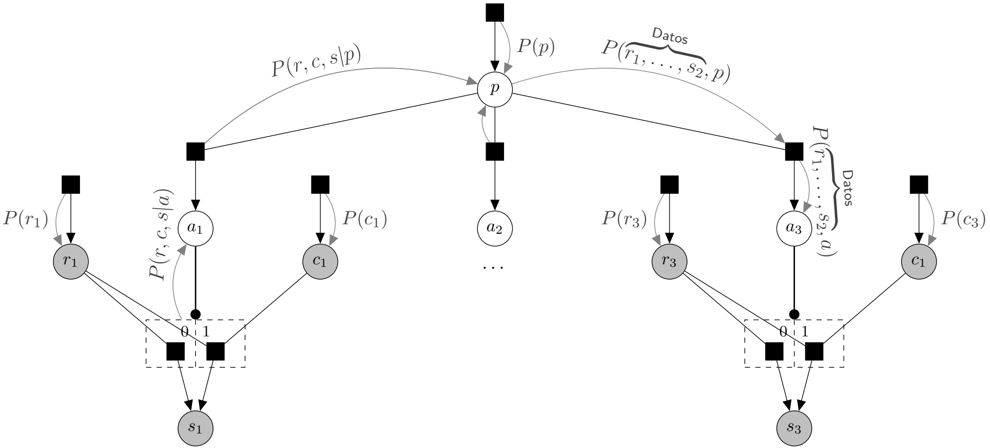
Flujo de inferencia

estructuras causales dinámicas



$$m_{f_{a_3}\rightarrow a_3}(a_3) = \sum_p P(a_3|p)P(r_1,c_1,s_1,r_2,c_2,s_2,p)$$

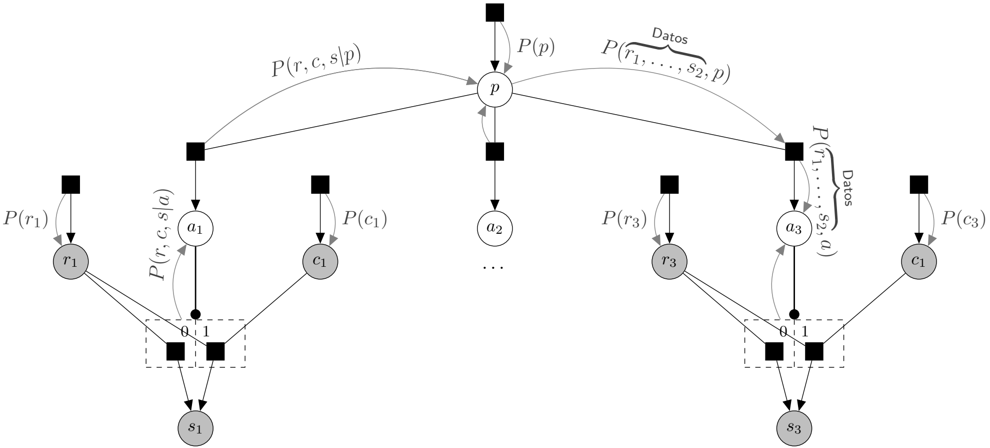
Flujo de inferencia
estructuras causales dinámicas



$$m_{f_{a_3}\to a_3}(a_3) = \sum_p P(a_3|p)P(r_1,c_1,s_1,r_2,c_2,s_2,p) = P(r_1,\ldots,s_2,a)$$

Flujo de inferencia

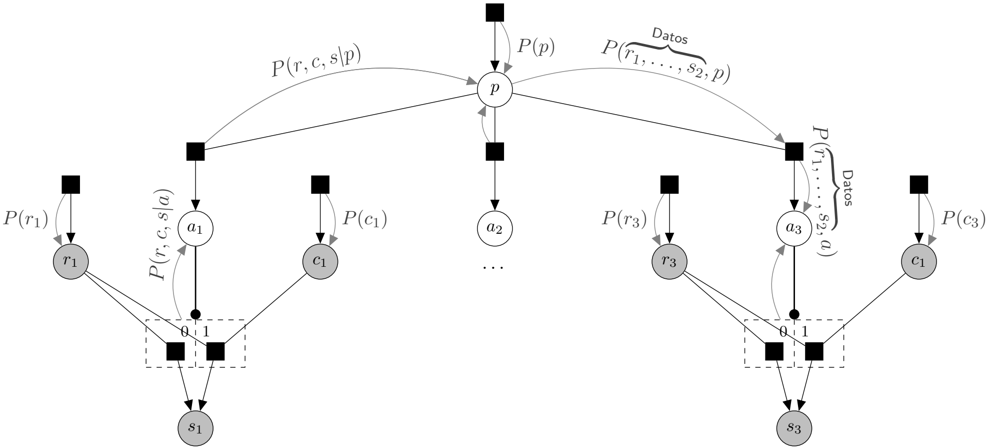
estructuras causales dinámicas



$$P(a_3|\overbrace{r_1,\ldots,s_3}^{\text{Datos}}) \propto$$

Flujo de inferencia

estructuras causales dinámicas



$$P(a_3|\overbrace{r_1, \dots, s_3}^{\text{Datos}}) \propto P(r_1, c_1, s_1, r_2, c_2, s_2, a_3)P(r_3, c_3, s_3|a_3)$$

Laboratorios de
Métodos Bayesianos

Bibliografía Unidad 5

- Neal. **Pattern Recognition and Machine Learning**. 2020 (Draft). (Descargar). (lectura capítulo 2, 4, 5 y 6)

Unidad 3
Parte 2

Estimación de habilidad por pasaje de mensajes

Unidad 3
Parte 2

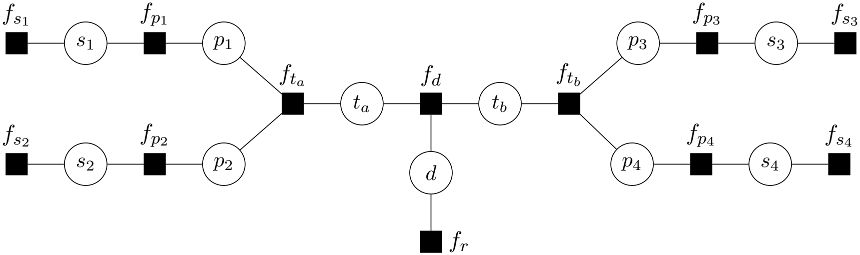
Estimación de habilidad por pasaje de mensajes

- Herbrich et al. **TrueSkill: A Bayesian Skill Rating System**. Advances in Neural Information Processing Systems. 2006. (Descargar). Lectura: paper completo.

Más adelante (Unidad 7):

- Dangauthier et al. *Trueskill through time: Revisiting the history of chess*. Advances in Neural Information Processing Systems. 2008. (Descargar). Lectura: paper completo.

TrueSkill



$f_{s_i} = \mathcal{N}(s_i|\mu_i, \sigma^2) \quad f_{p_i} = \mathcal{N}(p_i|s_i, \beta^2) \quad f_{t_e} = \mathbb{I}(t_e = \sum_{i \in A_e} p_i) \quad f_d = \mathbb{I}(d = t_a - t_b) \quad f_r = \mathbb{I}(d > 0)$

Propiedades

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\mu|x,\sigma^2) = \mathcal{N}(-\mu|-x,\sigma^2) = \mathcal{N}(-x|-\mu,\sigma^2)$$

Propiedades

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\mu|x,\sigma^2) = \mathcal{N}(-\mu|-x,\sigma^2) = \mathcal{N}(-x|-\mu,\sigma^2)$$

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\frac{X-\mu}{\sigma}|0,1)$$

Propiedades

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\mu|x,\sigma^2) = \mathcal{N}(-\mu|-x,\sigma^2) = \mathcal{N}(-x|-\mu,\sigma^2)$$

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\frac{X-\mu}{\sigma}|0,1)$$

$$\frac{\partial}{\partial x}\Phi(x|\mu,\sigma^2) = \mathcal{N}(x|\mu,\sigma^2)$$

Propiedades

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\mu|x,\sigma^2) = \mathcal{N}(-\mu|-x,\sigma^2) = \mathcal{N}(-x|-\mu,\sigma^2)$$

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\frac{X-\mu}{\sigma}|0,1)$$

$$\frac{\partial}{\partial x}\Phi(x|\mu,\sigma^2) = \mathcal{N}(x|\mu,\sigma^2)$$

$$\iint_{-\infty}^{\infty} \mathbb{I}(x = h(y,z))\, f(x)\, g(y)\, dx\, dy = \int_{-\infty}^{\infty} f(h(y,z))\, g(y)\, dy$$

Propiedades

$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\mu|x,\sigma^2) = \mathcal{N}(-\mu|-x,\sigma^2) = \mathcal{N}(-x|-\mu,\sigma^2)$$

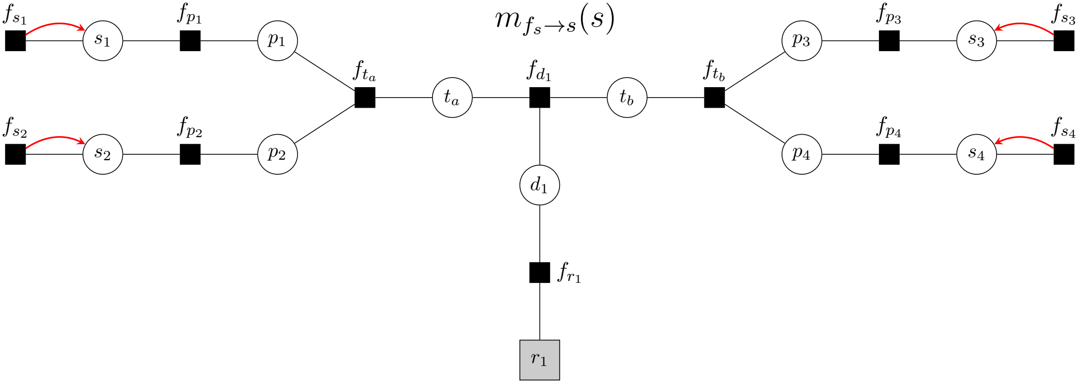
$$\mathcal{N}(x|\mu,\sigma^2) = \mathcal{N}(\frac{X-\mu}{\sigma}|0,1)$$

$$\frac{\partial}{\partial x}\Phi(x|\mu,\sigma^2) = \mathcal{N}(x|\mu,\sigma^2)$$

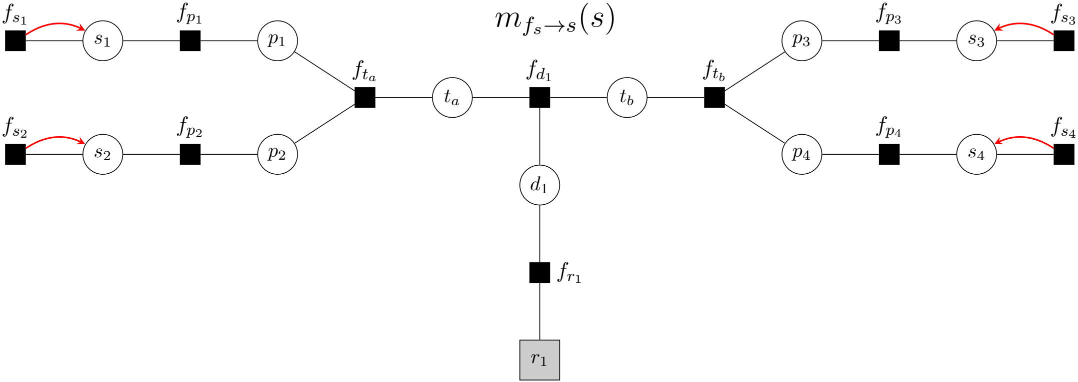
$$\iint_{-\infty}^{\infty} \mathbb{I}(x = h(y,z))\, f(x)\, g(y)\, dx\, dy = \int_{-\infty}^{\infty} f(h(y,z))\, g(y)\, dy$$

$$\int_{-\infty}^{\infty} N(x|\mu_x,\sigma_x^2)N(x|\mu_y,\sigma_y^2)\, dx \overset{*}{=} \int_{-\infty}^{\infty} \underbrace{N(\mu_x|\mu_y,\sigma_x^2+\sigma_y^2)}_{\text{constante}}\, \underbrace{N(x|\mu_*,\sigma_*^2)dx}_{\text{integra 1}}$$

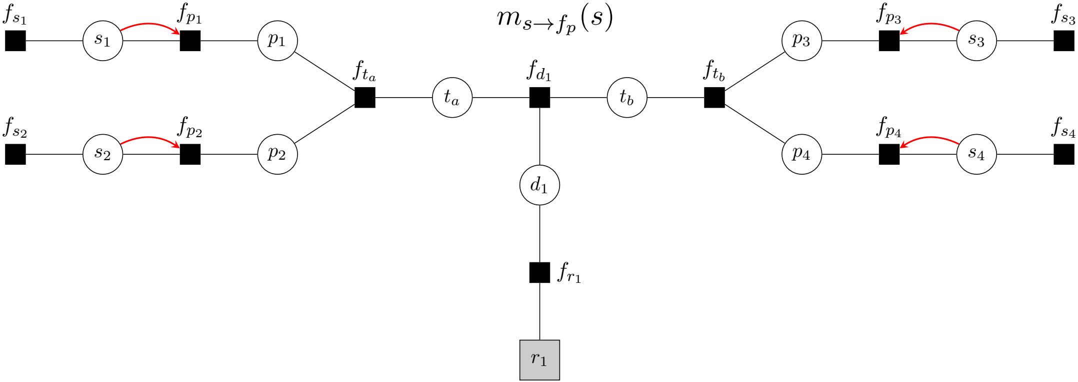
$$m_{f_s \to s}(s)$$



$$m_{f_{s_i} \to s_i}(s_i) =$$

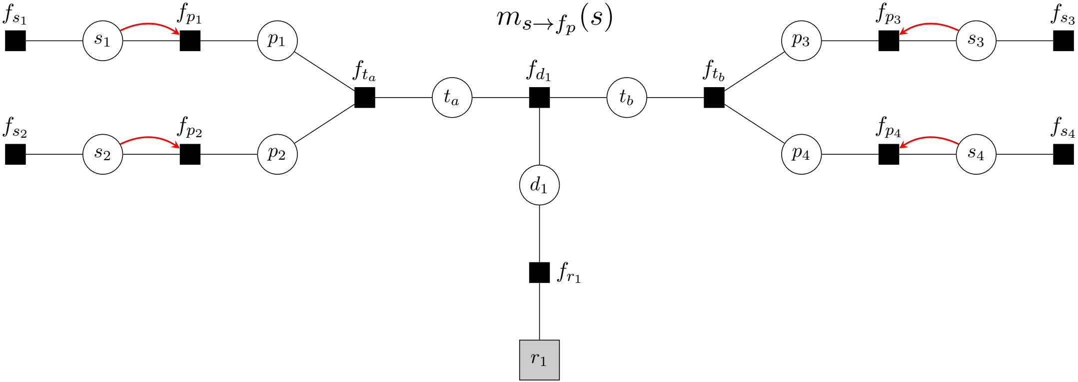


$$m_{f_{s_i} \to s_i}(s_i) = \mathcal{N}(s_i | \mu_i, \sigma_i^2)$$


$m_{s \to f_p}(s)$
f_{s_1}
s_1
f_{p_1}
p_1
f_{t_a}
t_a
f_{d_1}
t_b
f_{t_b}
p_3
f_{p_3}
s_3
f_{s_3}
f_{s_2}
s_2
f_{p_2}
p_2
p_4
f_{p_4}
s_4
f_{s_4}
d_1
f_{r_1}
r_1

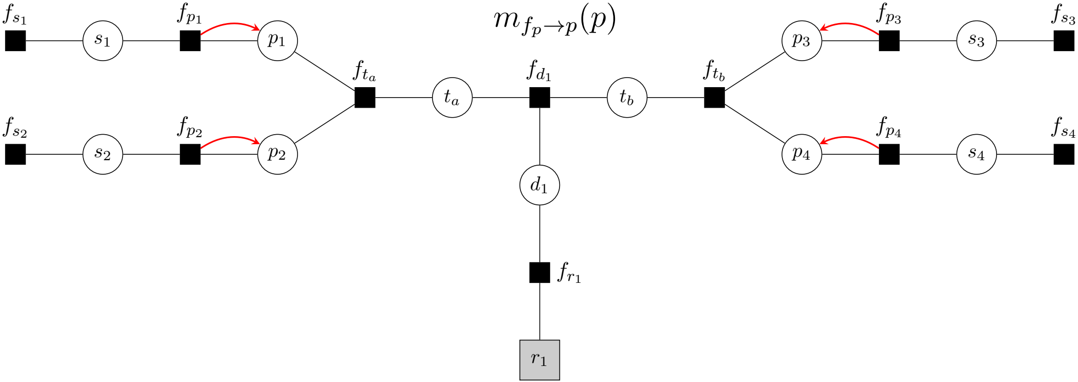
$$m_{s_i \to f_{p_i}}(s_i) =$$

$$m_{s \to f_p}(s)$$



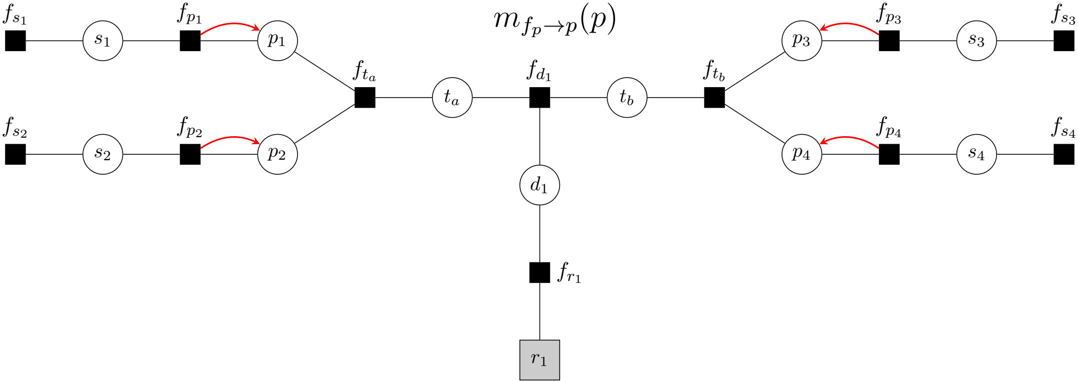
$$m_{s_i \to f_{p_i}}(s_i) = \mathcal{N}(s_i | \mu_i, \sigma_i^2)$$

$$m_{f_p \to p}(p)$$

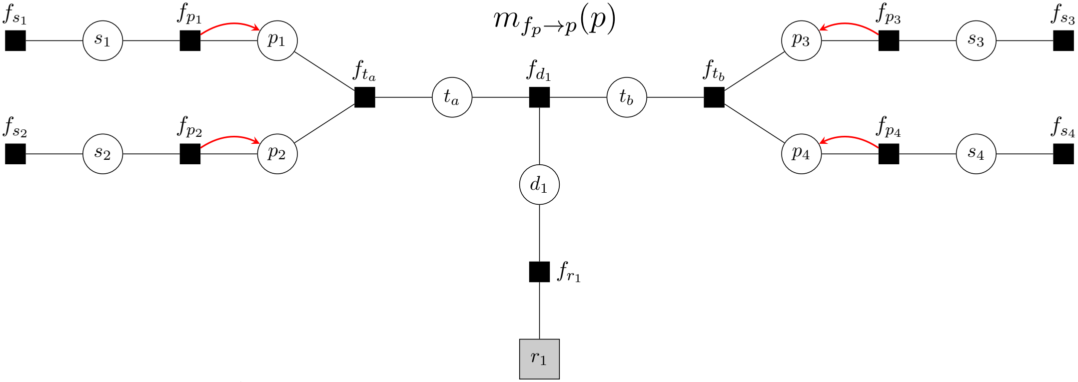


$$m_{f_{p_i} \to p_i}(p_i) =$$

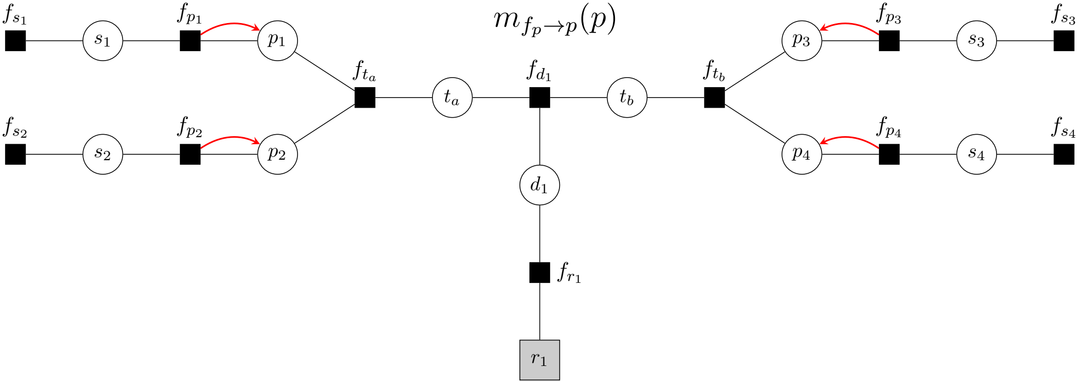
$$m_{f_p \to p}(p)$$



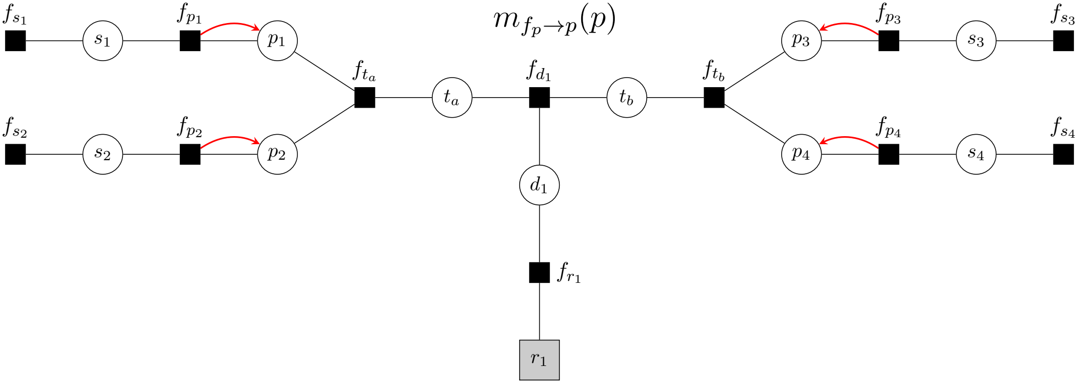
$$m_{f_{p_i} \to p_i}(p_i) =$$



$$m_{f_{p_i} \to p_i}(p_i) = \int \mathcal{N}(p_i | s_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i$$

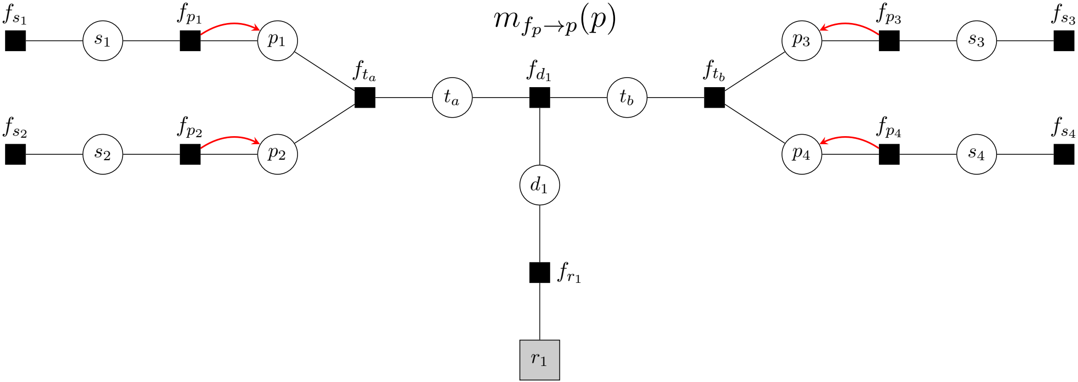


$$m_{f_{p_i} \to p_i}(p_i) = \int \mathcal{N}(p_i | s_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i = \int \mathcal{N}(s_i | p_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i$$



$$m_{f_{p_i} \to p_i}(p_i) = \int \mathcal{N}(p_i | s_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i = \int \mathcal{N}(s_i | p_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i$$

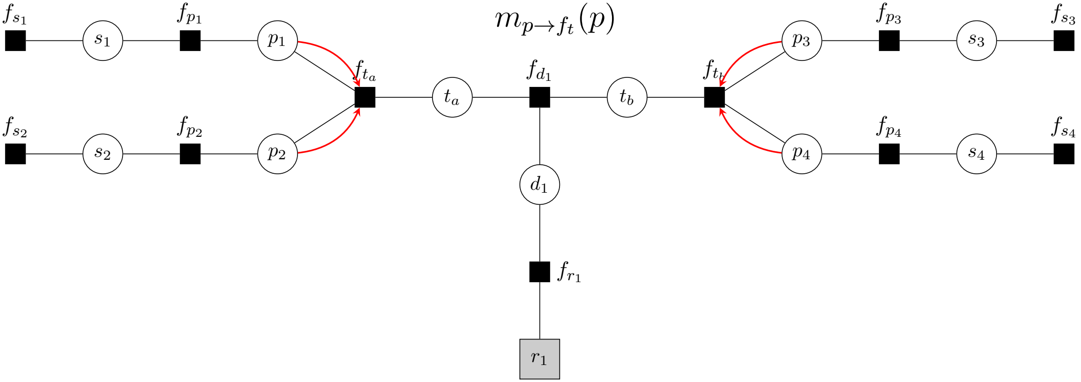
$$= \int \underbrace{\mathcal{N}(p_i | \mu_i, \beta^2 + \sigma_i^2)}_{\text{const.}} \underbrace{\mathcal{N}(s_i | \mu_*, \sigma_*^2)\, ds_i}_{1}$$



$$m_{f_{p_i} \to p_i}(p_i) = \int \mathcal{N}(p_i | s_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i = \int \mathcal{N}(s_i | p_i, \beta^2)\, \mathcal{N}(s_i | \mu_i, \sigma_i^2)\, ds_i$$

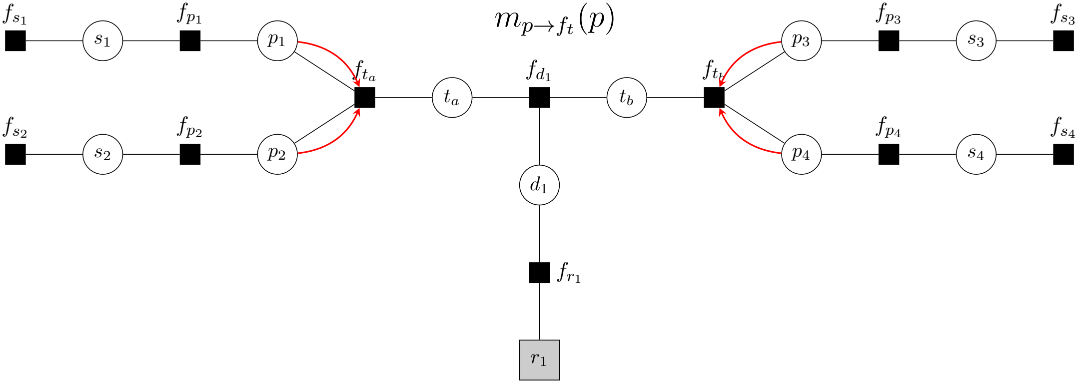
$$= \int \underbrace{\mathcal{N}(p_i | \mu_i, \beta^2 + \sigma_i^2)}_{\text{const.}} \underbrace{\mathcal{N}(s_i | \mu_*, \sigma_*^2)\, ds_i}_{1} = \mathcal{N}(p_i | \mu_i, \beta^2 + \sigma_i^2)$$

$$m_{p \to f_t}(p)$$



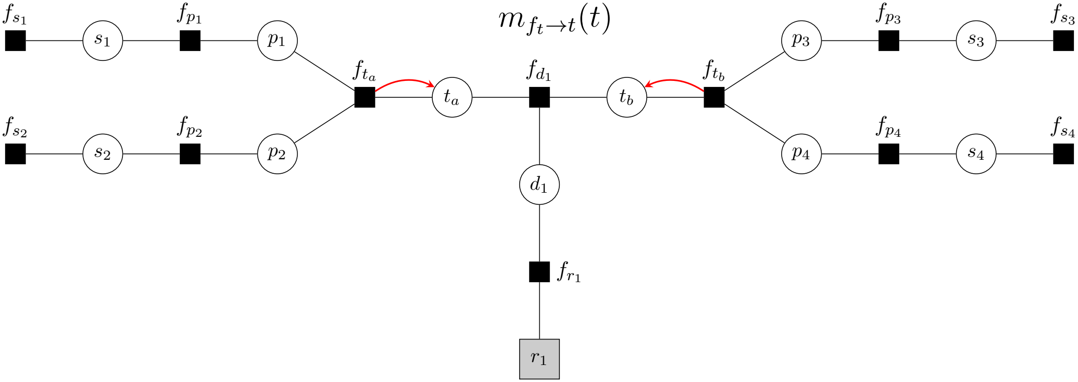
$$m_{p_i \to f_{t_e}}(p_i) =$$

$$m_{p \to f_t}(p)$$

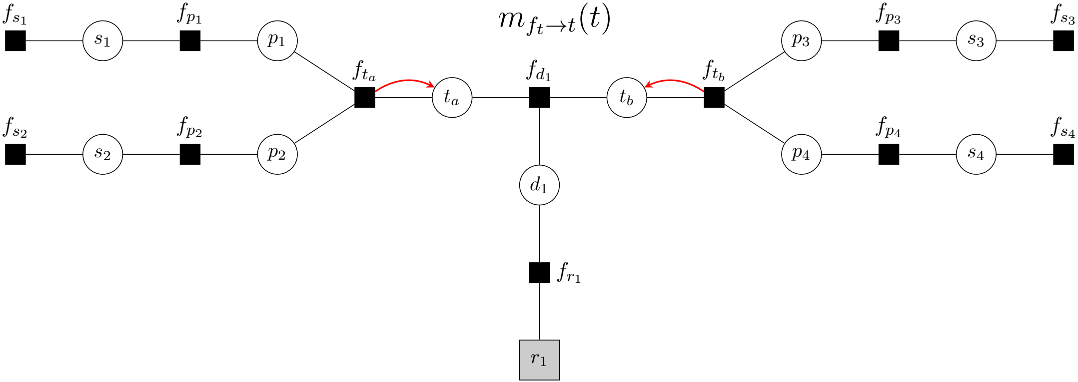


$$m_{p_i \to f_{t_e}}(p_i) = \mathcal{N}(p_i | \mu_i, \beta^2 + \sigma_i^2)$$

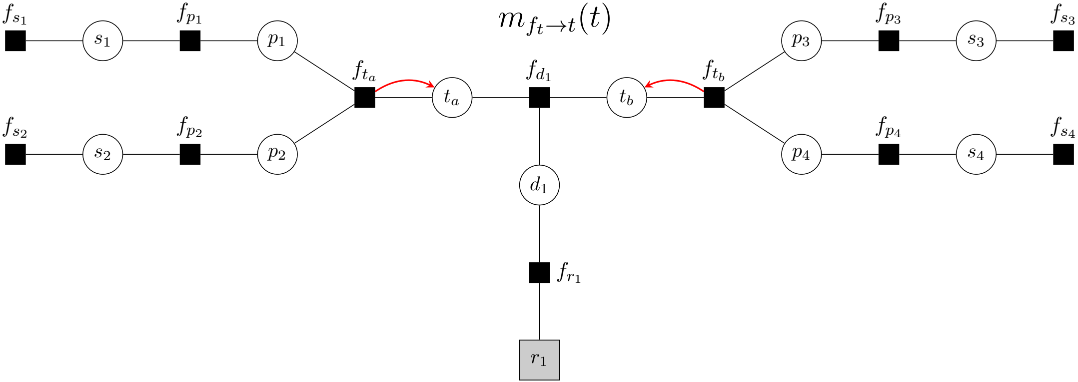
$$m_{f_t \to t}(t)$$



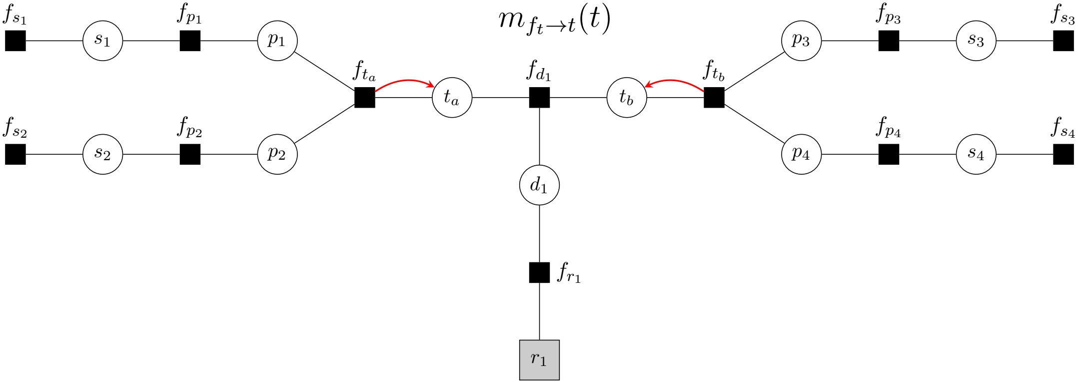
$$m_{f_{t_e} \to t_e}(t_e) =$$



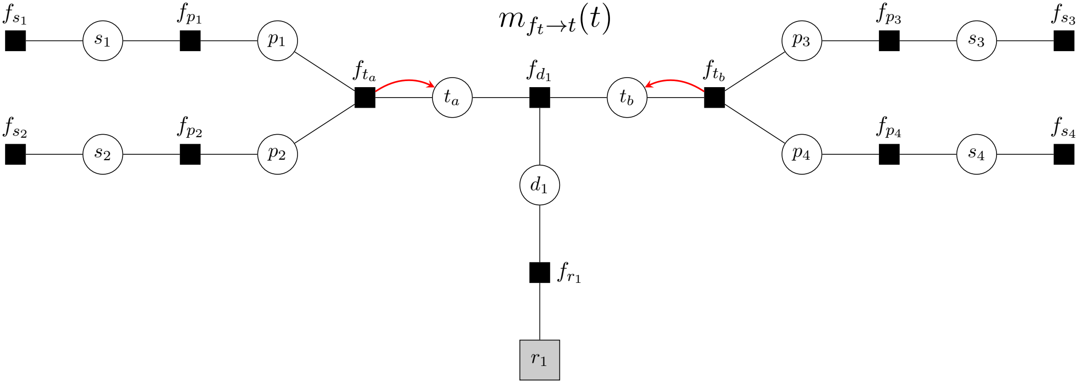
$$m_{f_{t_e} \to t_e}(t_e) = \iint \mathbb{I}(t_e = p_i + p_j)\mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(p_j|\mu_j, \beta^2 + \sigma_j^2)dp_i dp_j$$



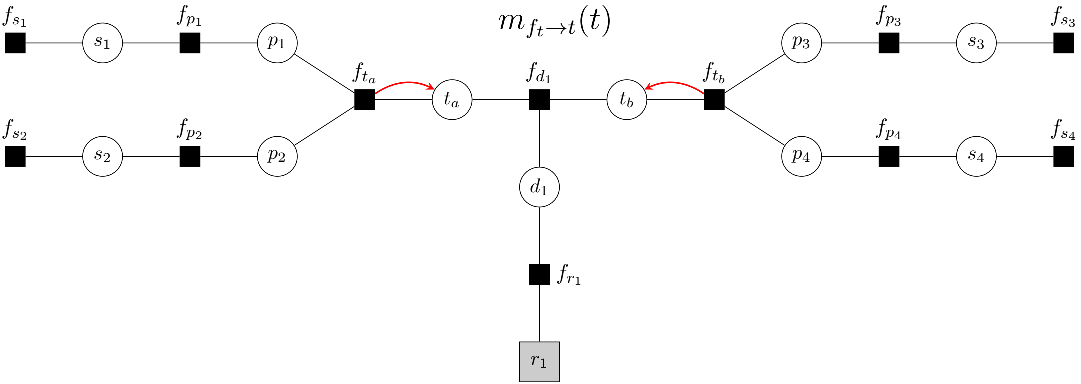
$$\begin{aligned} m_{f_{t_e} \to t_e}(t_e) &= \iint \mathbb{I}(t_e = p_i + p_j)\mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(p_j|\mu_j, \beta^2 + \sigma_j^2)dp_i dp_j \\ &= \int \mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(t_e - p_i|\mu_j, \beta^2 + \sigma_j^2)dp_i \end{aligned}$$



$$
\begin{aligned}
m_{f_{t_e} \to t_e}(t_e) &= \iint \mathbb{I}(t_e = p_i + p_j)\mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(p_j|\mu_j, \beta^2 + \sigma_j^2)dp_i dp_j \\
&= \int \mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(t_e - p_i|\mu_j, \beta^2 + \sigma_j^2)dp_i \\
&= \int \underbrace{\mathcal{N}(t_e|\,\mu_i + \mu_j\,,\, 2\beta^2 + \sigma_i^2 + \sigma_j^2\,)}_{\text{const.}} \underbrace{\mathcal{N}(p_i|\mu_*, \sigma_*^2)dp_i}_{1}
\end{aligned}
$$

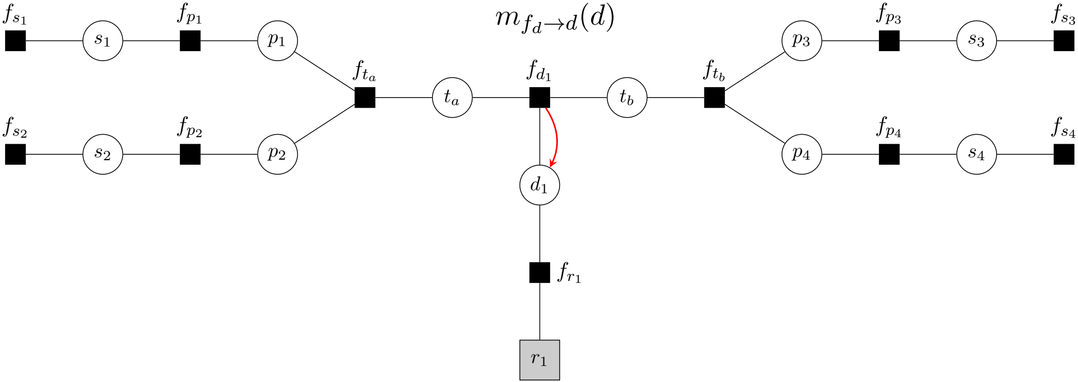


$$
\begin{aligned}
m_{f_{t_e} \to t_e}(t_e) &= \iint \mathbb{I}(t_e = p_i + p_j)\mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(p_j|\mu_j, \beta^2 + \sigma_j^2)dp_i dp_j \\
&= \int \mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(t_e - p_i|\mu_j, \beta^2 + \sigma_j^2)dp_i \\
&= \mathcal{N}(t_e|\,\mu_i + \mu_j\,,\,2\beta^2 + \sigma_i^2 + \sigma_j^2\,)
\end{aligned}
$$

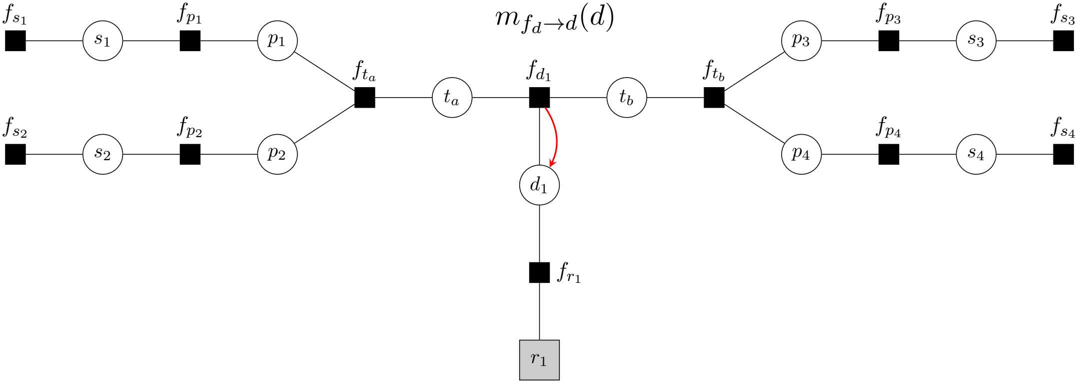


$$
\begin{aligned}
m_{f_{t_e} \to t_e}(t_e) &= \iint \mathbb{I}(t_e = p_i + p_j)\mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(p_j|\mu_j, \beta^2 + \sigma_j^2)dp_i dp_j \\
&= \int \mathcal{N}(p_i|\mu_i, \beta^2 + \sigma_i^2)\mathcal{N}(t_e - p_i|\mu_j, \beta^2 + \sigma_j^2)dp_i \\
&= \mathcal{N}(t_e|\underbrace{\mu_i + \mu_j}_{\mu_e}, \underbrace{2\beta^2 + \sigma_i^2 + \sigma_j^2}_{\sigma_e^2})
\end{aligned}
$$

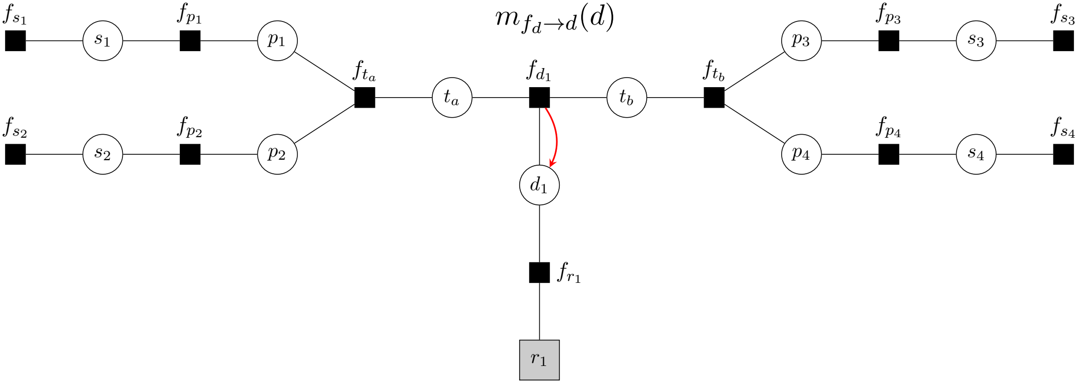
$m_{f_d \to d}(d)$



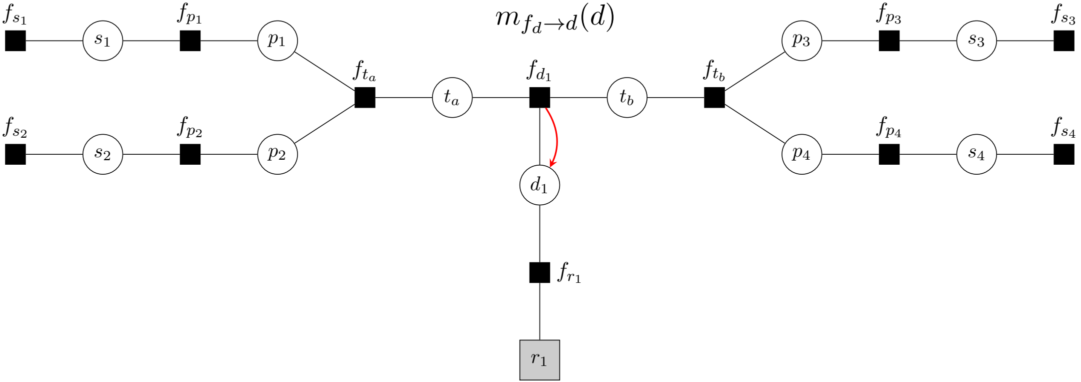
$$m_{f_d \to d}(d) =$$



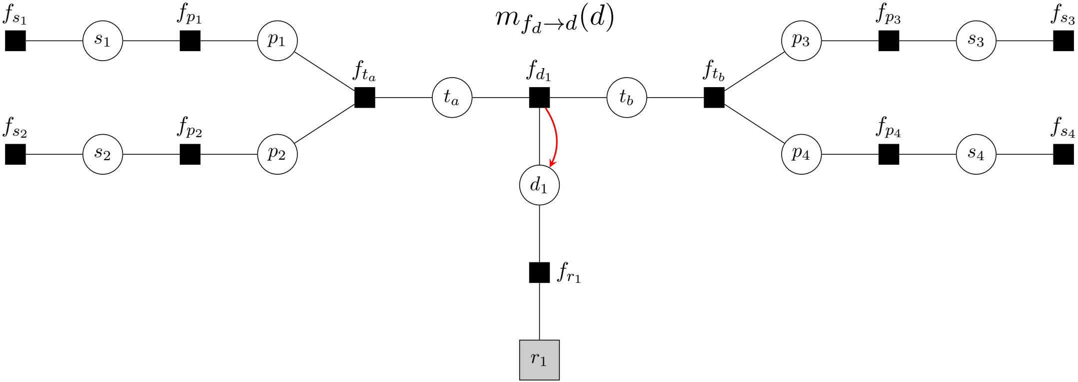
$$m_{f_d \to d}(d) = \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b$$



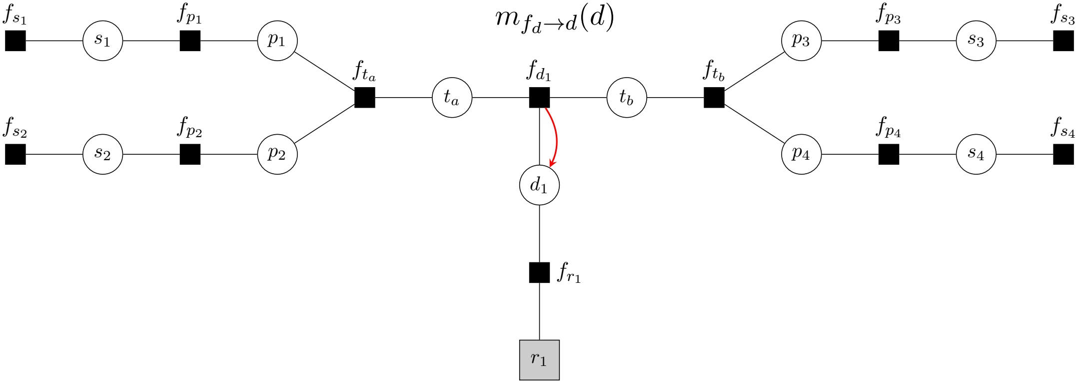
$$\begin{aligned} m_{f_d \to d}(d) &= \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b \\ &= \int \mathcal{N}(d + t_b|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b \end{aligned}$$



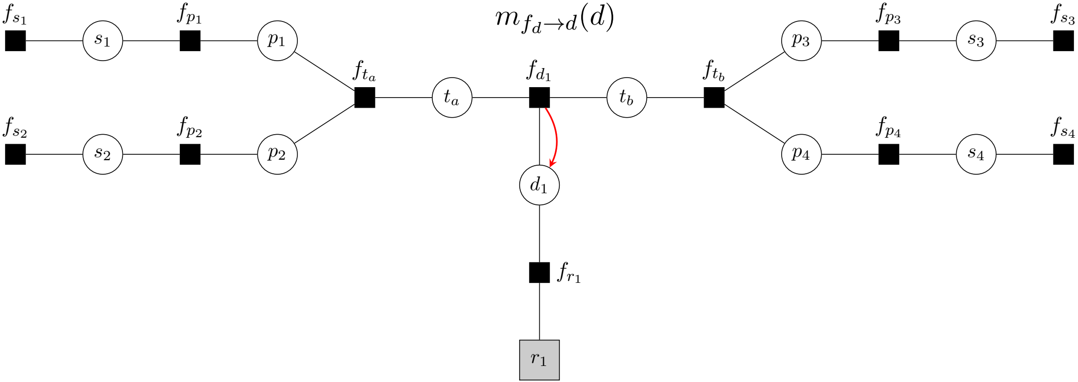
$$
\begin{aligned}
m_{f_d \to d}(d) &= \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b \\
&= \int \mathcal{N}(t_b|\mu_a - d, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b
\end{aligned}
$$



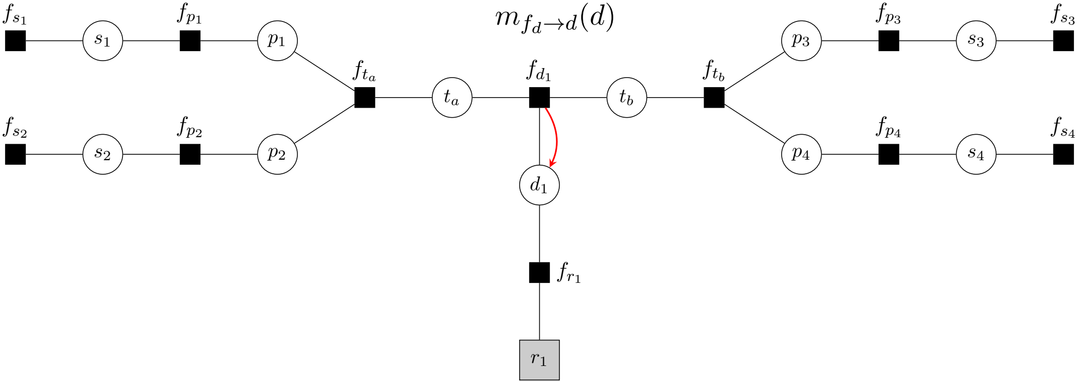
$$\begin{aligned}
m_{f_d \to d}(d) &= \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b \\
&= \int \mathcal{N}(t_b|\mu_a - d, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b \\
&= \int \underbrace{\mathcal{N}(d|\mu_a - \mu_b, \sigma_a^2 + \sigma_b^2)}_{\text{const.}}\ \underbrace{\mathcal{N}(t_b|\mu_*, \sigma_*^2)\, dt_b}_{1}
\end{aligned}$$



$$
\begin{aligned}
m_{f_d \to d}(d) &= \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b \\
&= \int \mathcal{N}(t_b|\mu_a - d, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b \\
&= \quad \mathcal{N}(d|\, \mu_a - \mu_b, \sigma_a^2 + \sigma_b^2)
\end{aligned}
$$

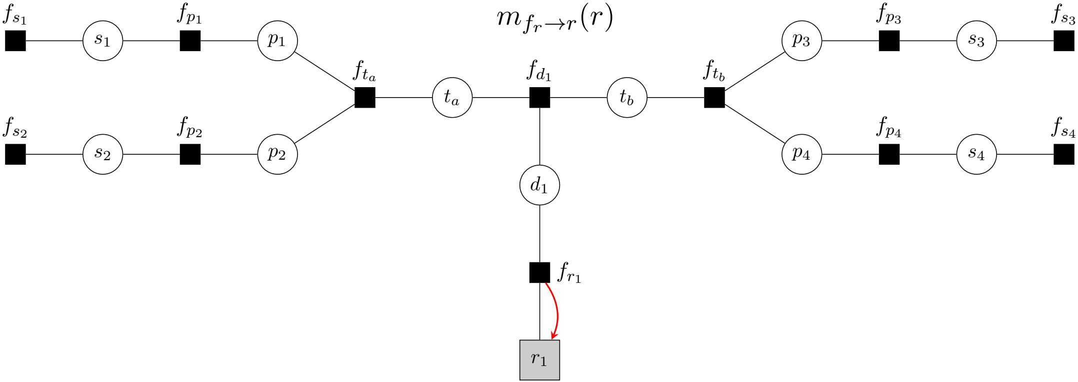


$$
\begin{aligned}
m_{f_d \to d}(d) &= \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b \\
&= \int \mathcal{N}(t_b|\mu_a - d, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b \\
&= \quad \mathcal{N}(d|\underbrace{\mu_a - \mu_b}_{\text{Diferencia esperada } \delta}, \underbrace{\sigma_a^2 + \sigma_b^2}_{\text{Varianza total } \vartheta})
\end{aligned}
$$

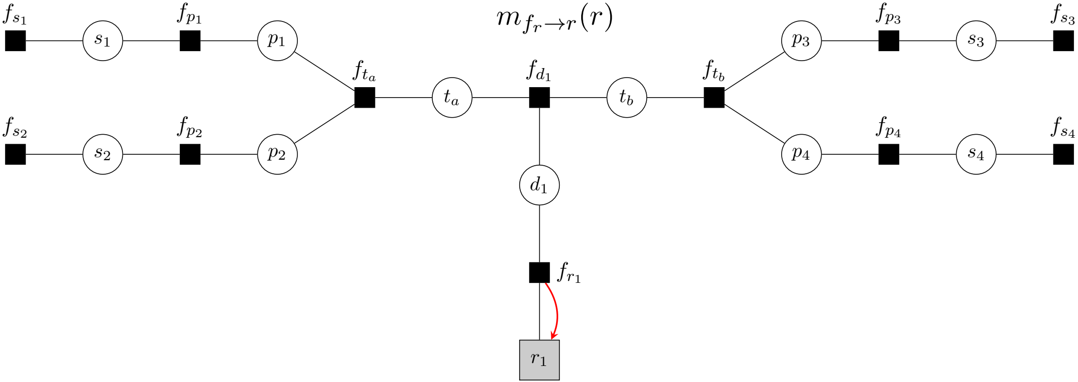


$$
\begin{aligned}
m_{f_d \to d}(d) &= \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(t_a|\mu_a, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_a dt_b \\
&= \int \mathcal{N}(t_b|\mu_a - d, \sigma_a^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b \\
&= \quad \mathcal{N}(d|\underbrace{\mu_a - \mu_b}_{\substack{\text{Diferencia} \\ \text{esperada}}\ \delta}, \underbrace{\sigma_a^2 + \sigma_b^2}_{\substack{\text{Varianza} \\ \text{total}}\ \vartheta}) = \mathcal{N}(d|\delta, \vartheta^2)
\end{aligned}
$$

$m_{f_r \to r}(r)$

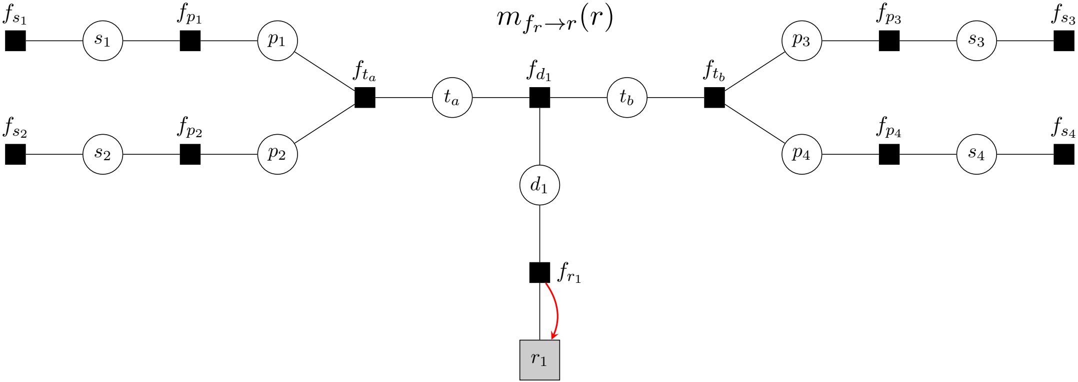


$$m_{f_r \to r}(r) =$$

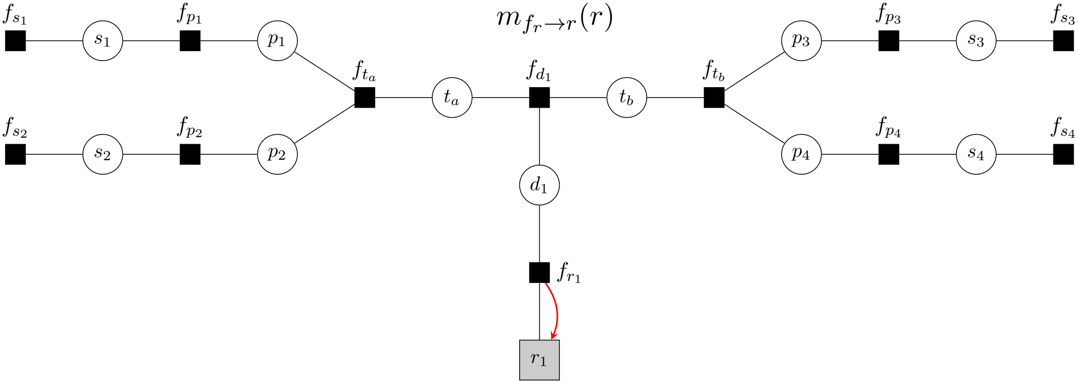


$$m_{f_r \to r}(r) = \int \mathbb{I}(d > 0)\mathcal{N}(d|\delta, \vartheta^2)\, dd$$

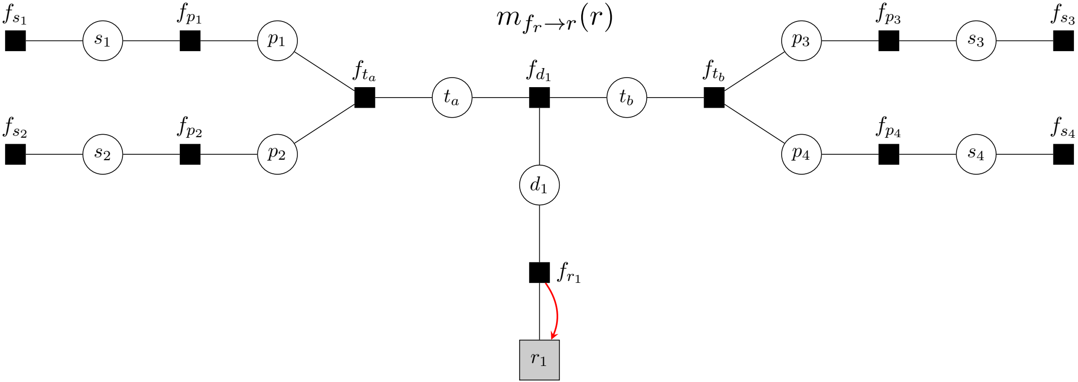
$m_{f_r \to r}(r)$



$$
\begin{aligned}
m_{f_r \to r}(r) &= \int \mathbb{I}(d > 0)\mathcal{N}(d|\delta, \vartheta^2)\, dd \\
&= \int_0^\infty \mathcal{N}(d|\delta, \vartheta^2)\, dd
\end{aligned}
$$



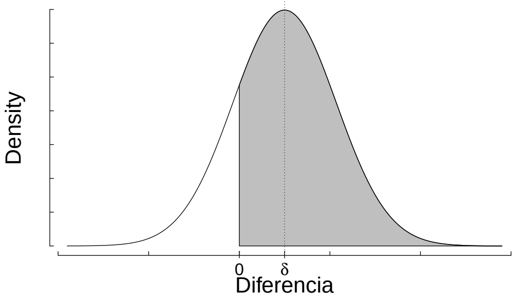
$$
\begin{aligned}
m_{f_r \to r}(r) &= \int \mathbb{I}(d > 0)\mathcal{N}(d|\delta, \vartheta^2)\, dd \\
&= \int_0^\infty \mathcal{N}(d|\delta, \vartheta^2)\, dd \\
&= 1 - \Phi(0|\delta, \vartheta^2)
\end{aligned}
$$



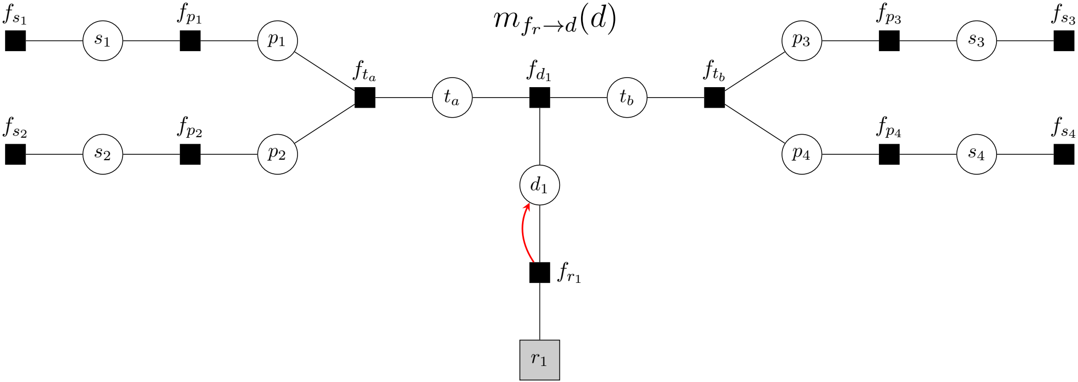
$$
\begin{aligned}
m_{f_r \to r}(r) &= \int \mathbb{I}(d > 0)\mathcal{N}(d|\delta, \vartheta^2)\, dd \\
&= \int_0^\infty \mathcal{N}(d|\delta, \vartheta^2)\, dd \\
&= 1 - \Phi(0|\delta, \vartheta^2) = \Phi\left(\frac{\delta}{\vartheta}\right)
\end{aligned}
$$

Evidencia

$$P(\text{Resultado}|\text{Modelo}) = 1 - \Phi\Big(0 \,|\, \underbrace{(\overbrace{\mu_1 + \mu_2}^{\mu_a}) - (\overbrace{\mu_3 + \mu_4}^{\mu_b})}_{\text{Diferencia esperada: } \delta},\ \underbrace{\overbrace{2\beta^2 + \sigma_1^2 + \sigma_2^2}^{\sigma_a^2} + \overbrace{2\beta^2 + \sigma_3^2 + \sigma_4^2}^{\sigma_b^2}}_{\text{incertidumbre total: } \vartheta^2}\Big)$$

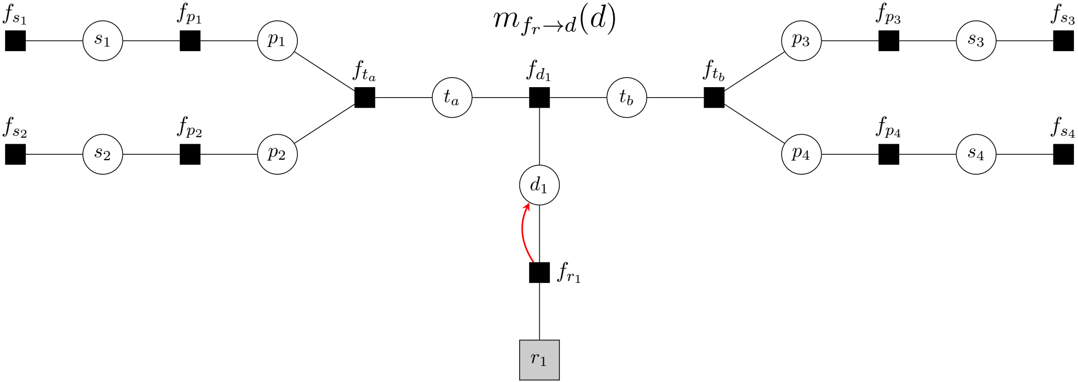


Mensajes ascendentes


f_{s_1}
s_1
f_{p_1}
p_1
$m_{f_r \to d}(d)$
p_3
f_{p_3}
s_3
f_{s_3}
f_{t_a}
t_a
f_{d_1}
t_b
f_{t_b}
f_{s_2}
s_2
f_{p_2}
p_2
p_4
f_{p_4}
s_4
f_{s_4}
d_1
f_{r_1}
r_1

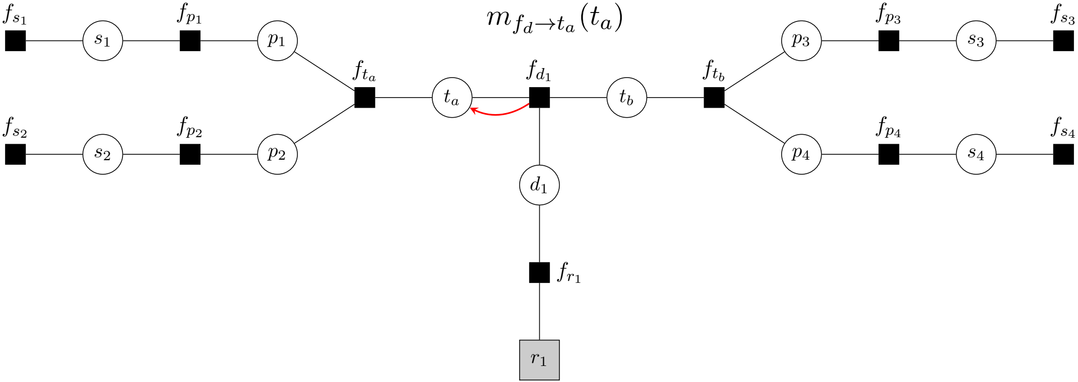
$$m_{f_r \to d}(d) =$$

$$m_{f_r \to d}(d)$$

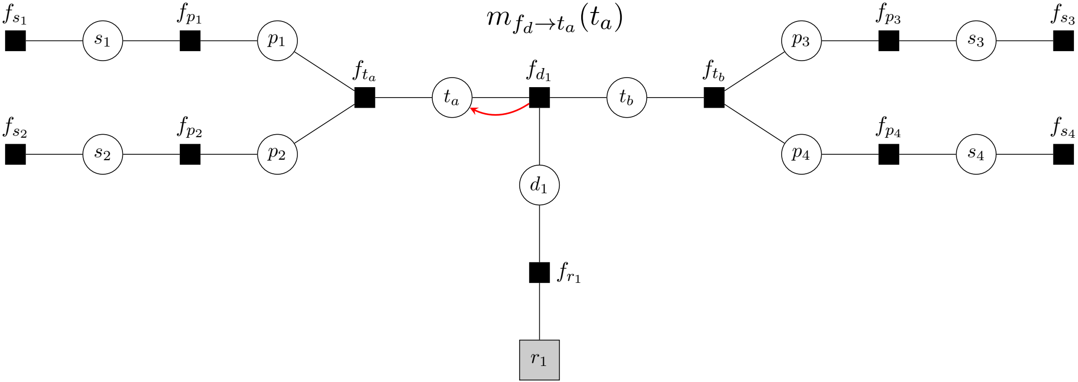


$$m_{f_r \to d}(d) = \mathbb{I}(d > 0)$$

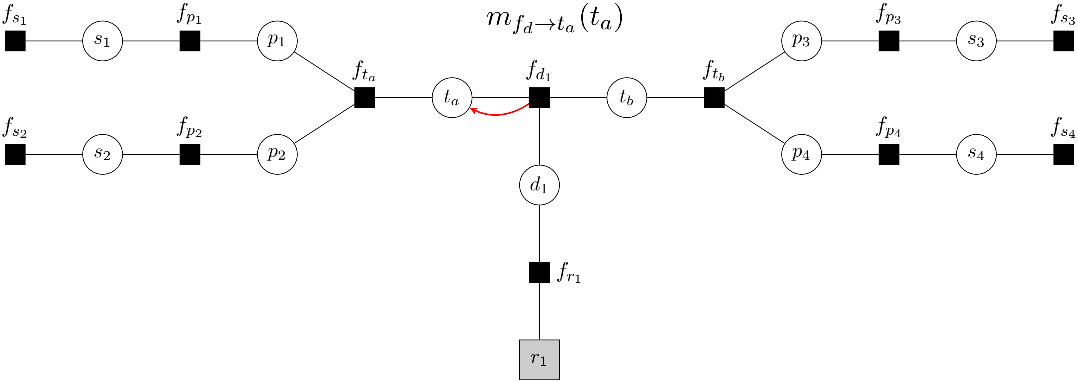
$$m_{f_d \to t_a}(t_a)$$



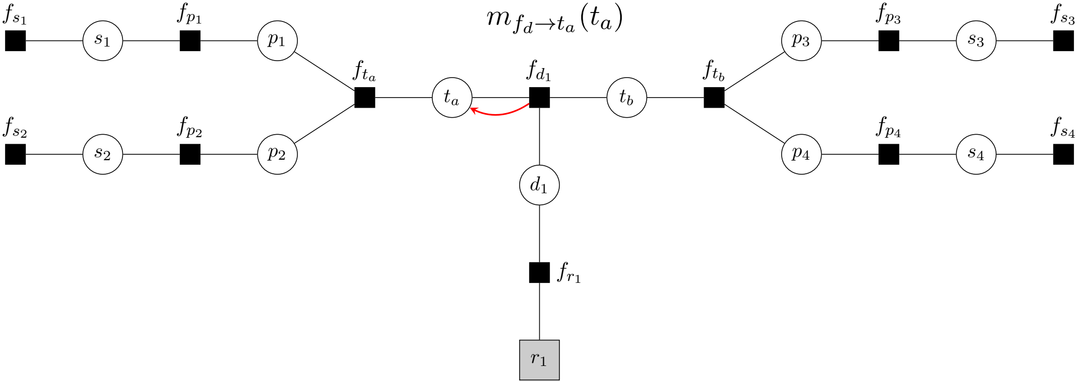
$$m_{f_d \to t_a}(t_a) =$$



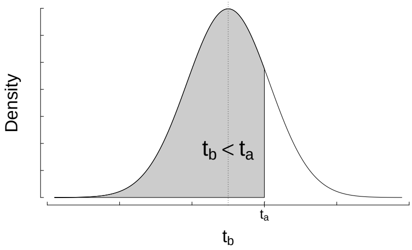
$$m_{f_d \to t_a}(t_a) = \iint \mathbb{I}(d = t_a - t_b)\mathbb{I}(d > 0)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b\, dd$$

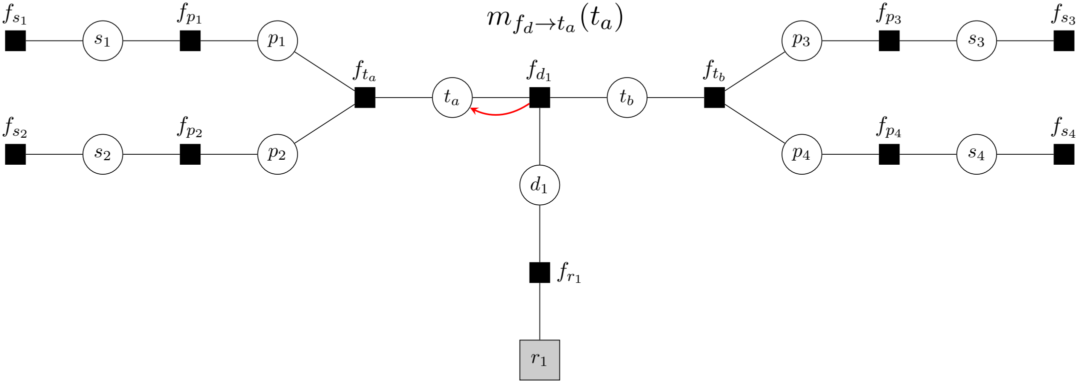


$$m_{f_d \to t_a}(t_a) = \int \mathbb{I}(t_a > t_b)\mathcal{N}(t_b | \mu_b, \sigma_b^2)\, dt_b$$



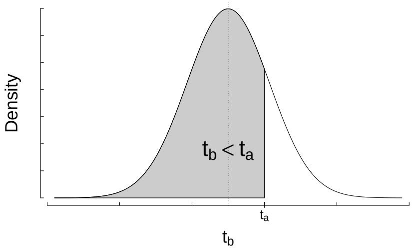
$$m_{f_d \to t_a}(t_a) = \int \mathbb{I}(t_a > t_b)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b$$



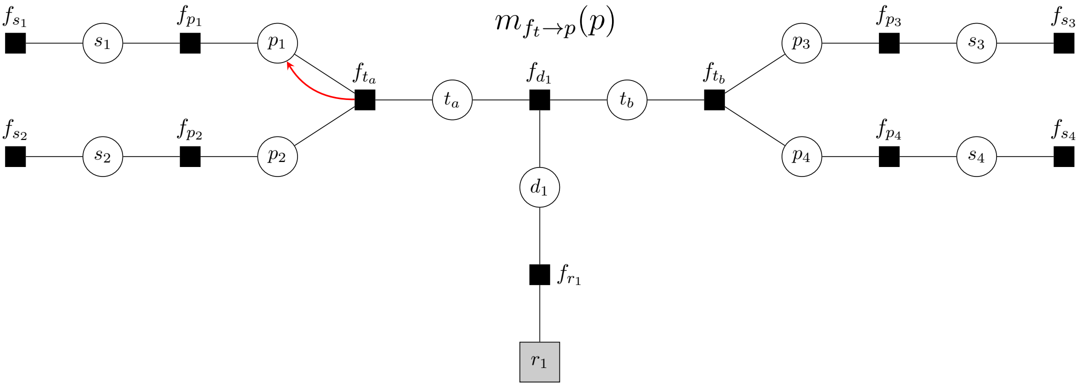


$$m_{f_d \to t_a}(t_a) = \int \mathbb{I}(t_a > t_b)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, dt_b$$

$$= \Phi(t_a|\mu_b, \sigma_b^2)$$

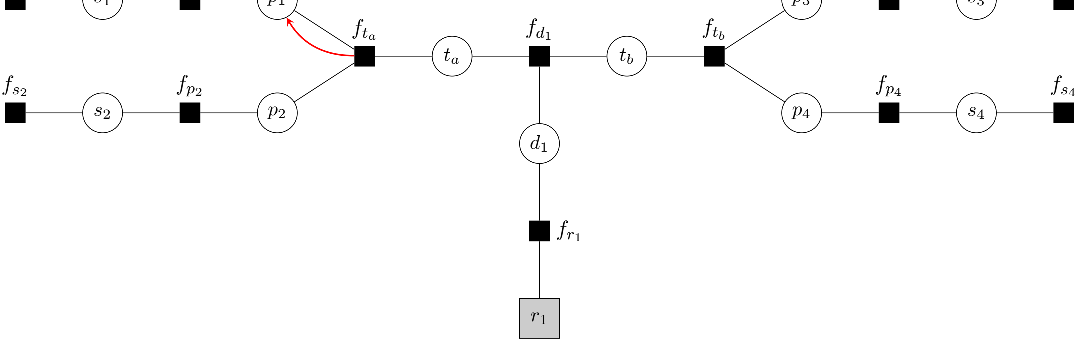


$$m_{f_t \to p}(p)$$



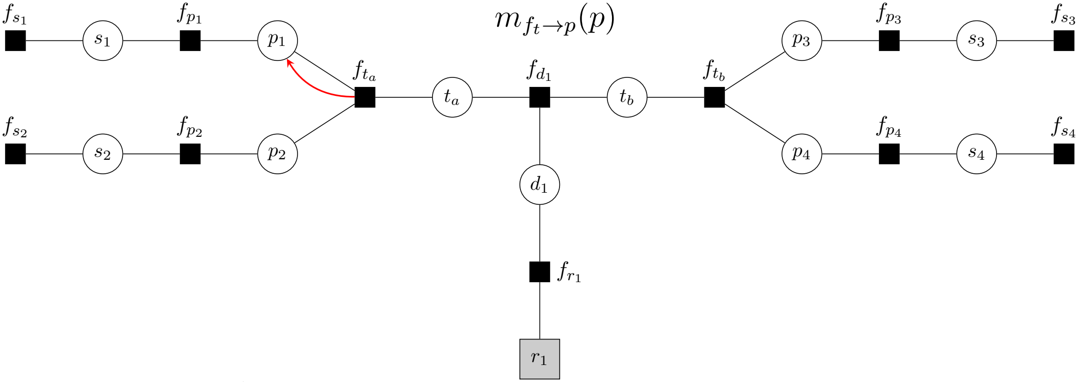
$$m_{f_{t_a} \to p_1}(p_1) =$$

$$m_{f_t \to p}(p)$$



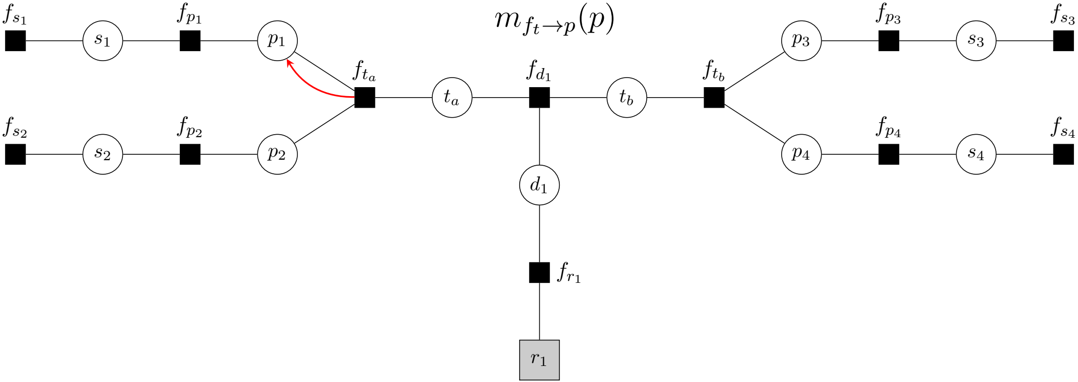
$$m_{f_{t_a} \to p_1}(p_1) = \iint \mathbb{I}(t_a = p_1 + p_2)\mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(t_a | \mu_b, \sigma_b^2)\, dt_a\, dp_2$$

$$m_{f_t \to p}(p)$$



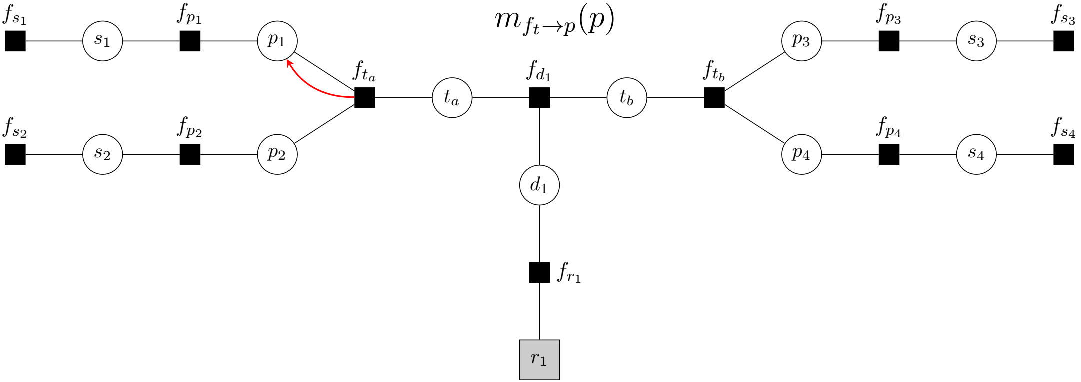
$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 + p_2 | \mu_b, \sigma_b^2)\, dp_2$$

$$m_{f_t \to p}(p)$$



$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

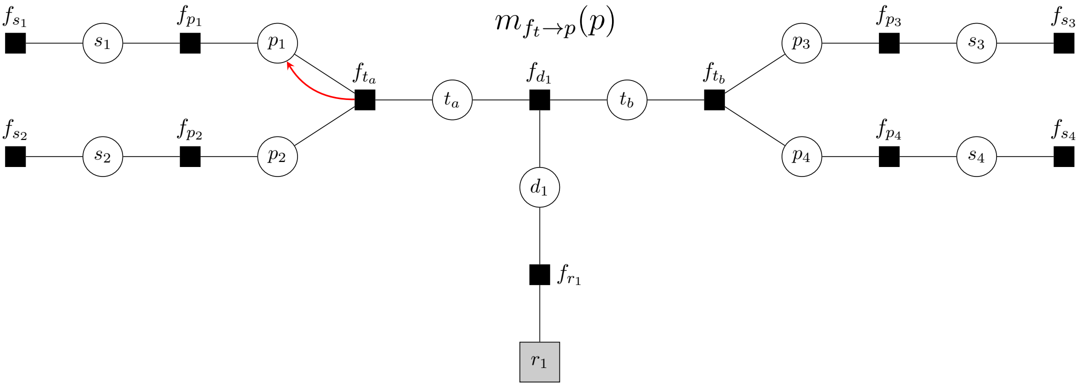
$$m_{f_t \to p}(p)$$



$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

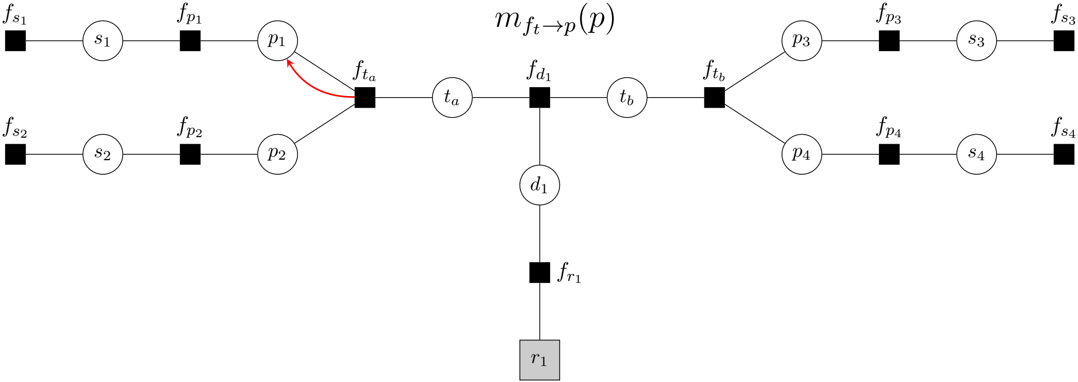
$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \frac{\partial}{\partial p_1} \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

$$m_{f_t \to p}(p)$$



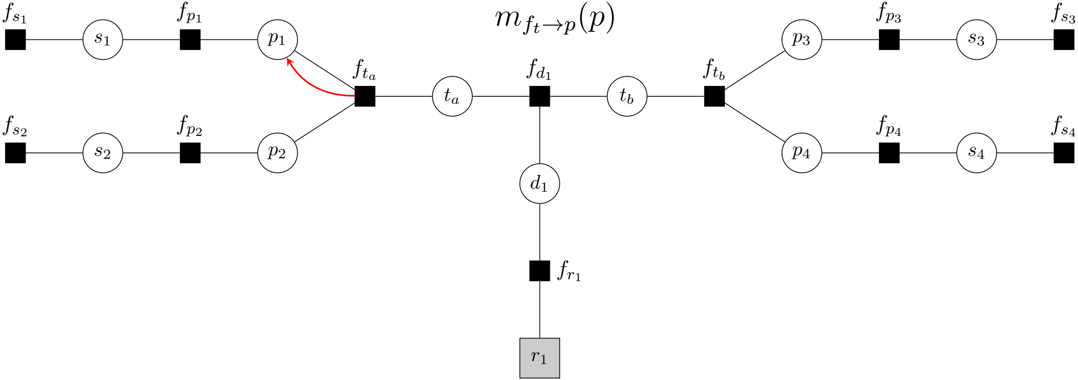
$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \frac{\partial}{\partial p_1} \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$



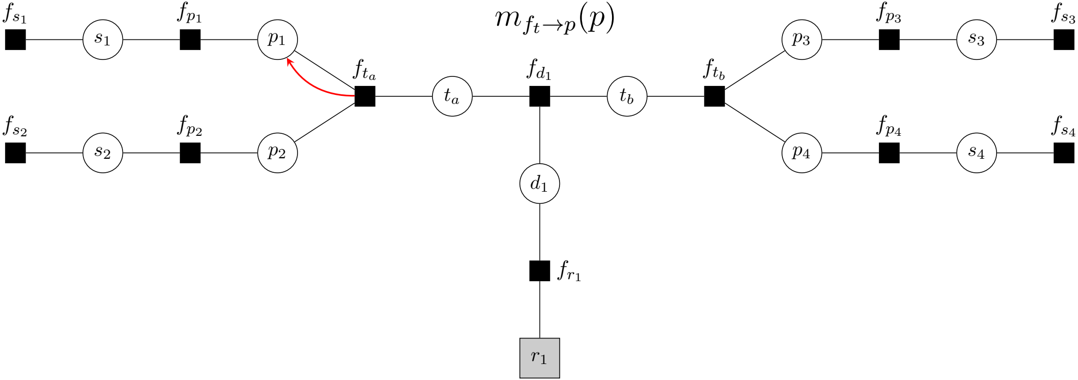
$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \mathcal{N}(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$



$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

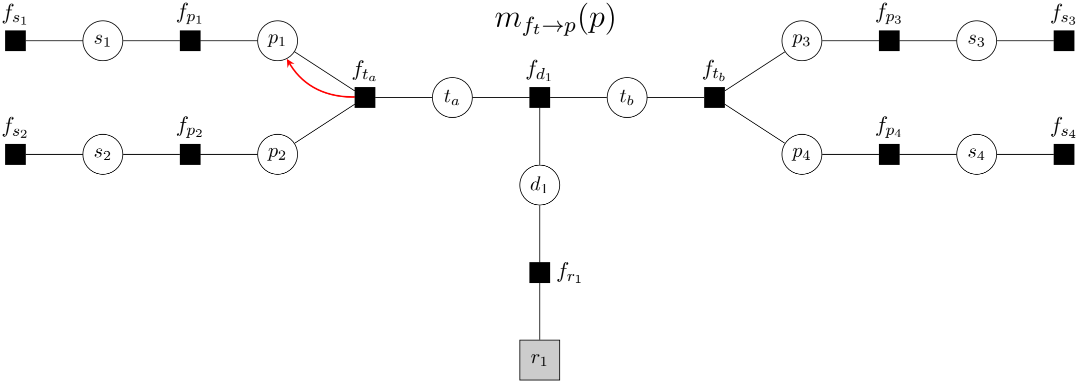
$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \mathcal{N}(p_2 | \mu_b - p_1, \sigma_b^2)\, dp_2$$



$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

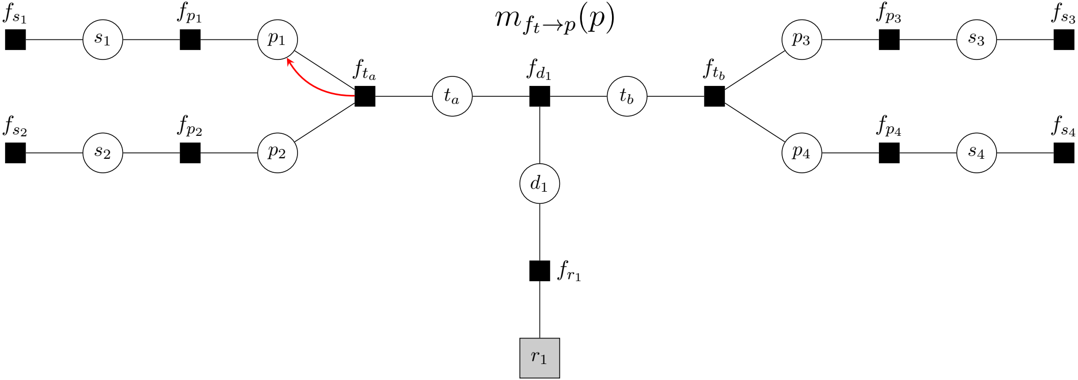
$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \int \underbrace{\mathcal{N}(\mu_2 | \mu_b - p_1, \sigma_b^2 + \sigma_2^2 + \beta^2)}_{\text{const.}} \underbrace{\mathcal{N}(p_2 | \mu_*, \sigma_*^2) dp_2}_{1}$$

$$m_{f_t \to p}(p)$$



$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \mathcal{N}(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)$$

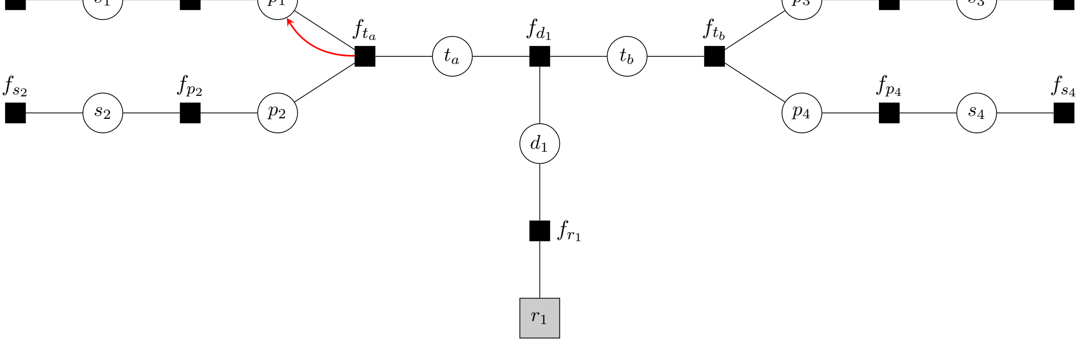


$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \mathcal{N}(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)$$

$$m_{f_{t_a} \to p_1}(p_1) =$$

$m_{f_t \to p}(p)$

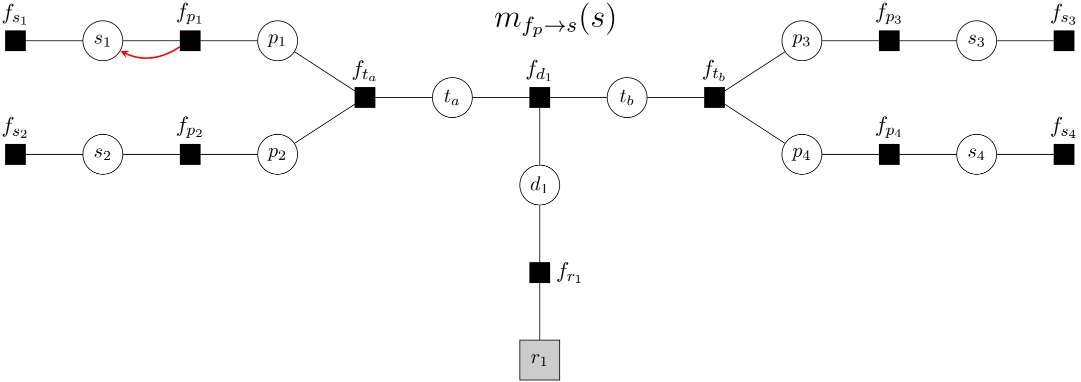


$$m_{f_{t_a} \to p_1}(p_1) = \int \mathcal{N}(p_2 | \mu_2, \sigma_2^2 + \beta^2)\, \Phi(p_1 | \mu_b - p_2, \sigma_b^2)\, dp_2$$

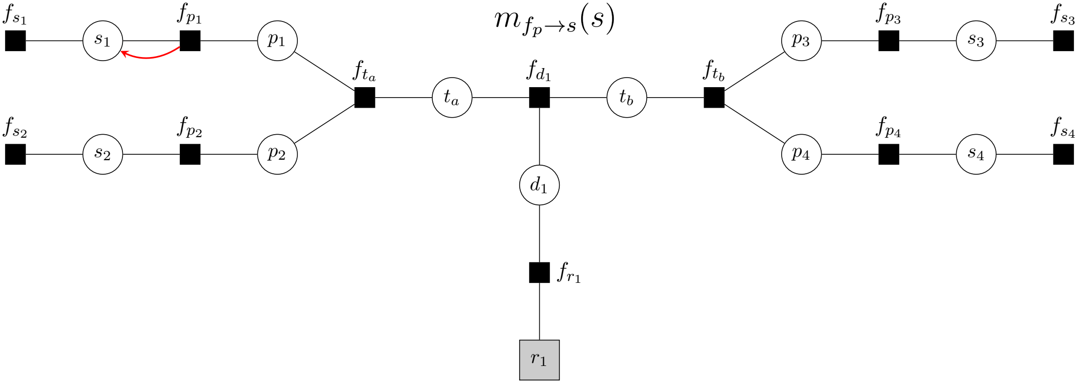
$$\frac{\partial}{\partial p_1} m_{f_{t_a} \to p_1}(p_1) = \mathcal{N}(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)$$

$$m_{f_{t_a} \to p_1}(p_1) = \Phi(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)$$

$$m_{f_p \to s}(s)$$

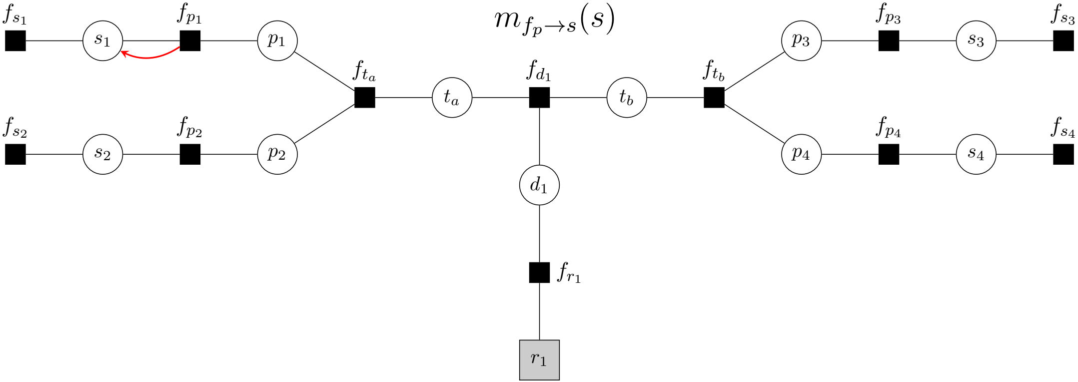


$$m_{f_{p_1} \to s_1}(s_1) =$$



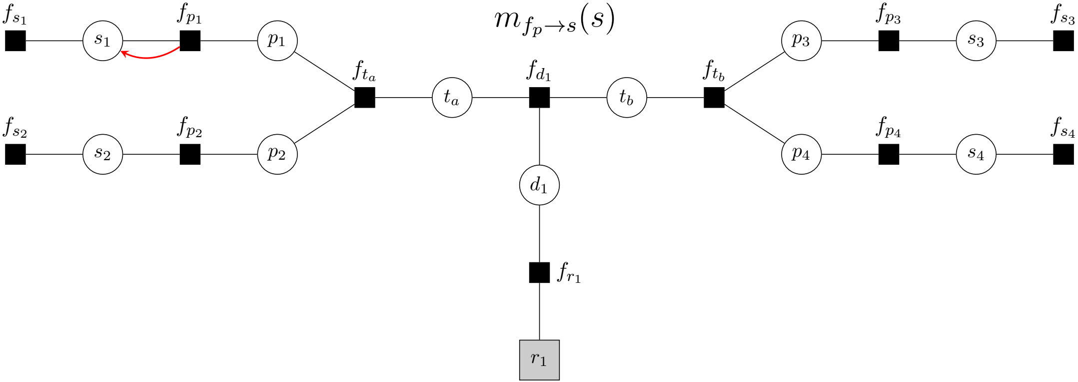
$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \Phi(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$

$$m_{f_p \to s}(s)$$



$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \Phi(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$

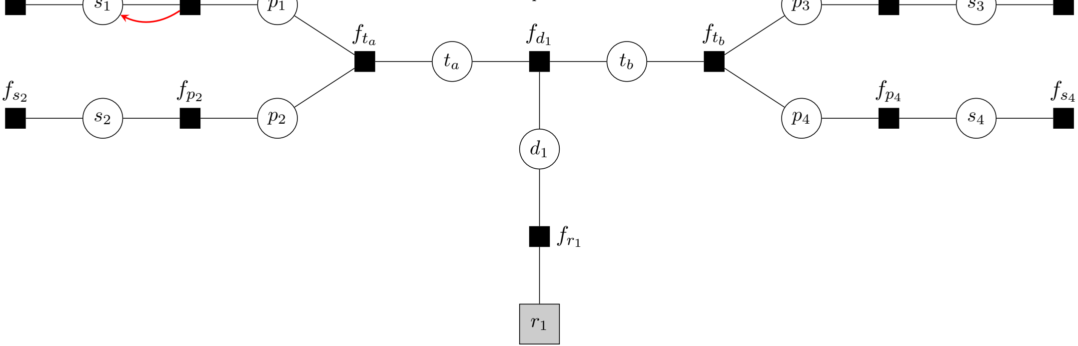
¿Sobre cuál variable derivamos?



$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2) \, \Phi(\mu_2 | \mu_b - p_1, \sigma_b^2 + \sigma_2^2 + \beta^2) \, dp_1$$

¿Sobre cuál variable derivamos?

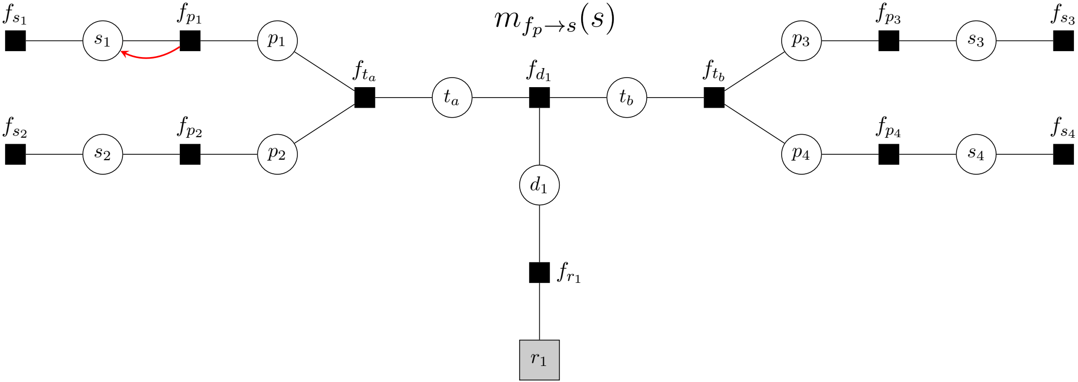
$$m_{f_p \to s}(s)$$



$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \Phi(\mu_2 | \mu_b - p_1, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$

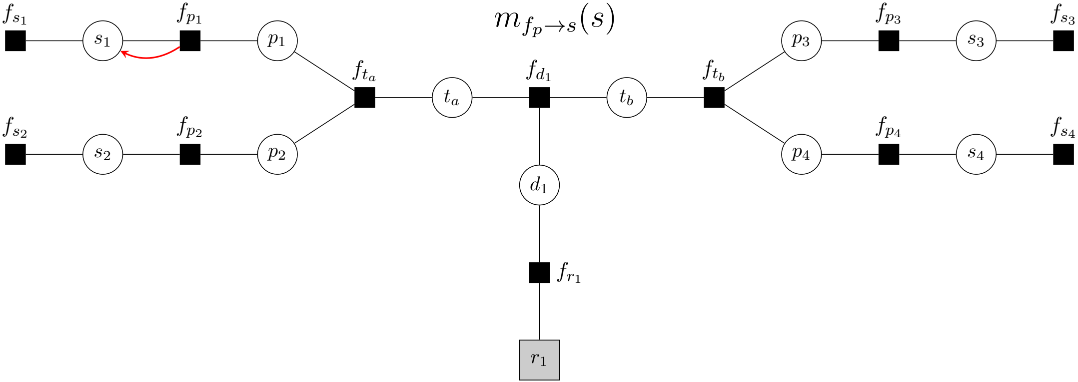
$$\frac{\partial}{\partial \mu_2} m_{f_{s_1} \to s_1}(s_1) =$$

$$m_{f_p \to s}(s)$$



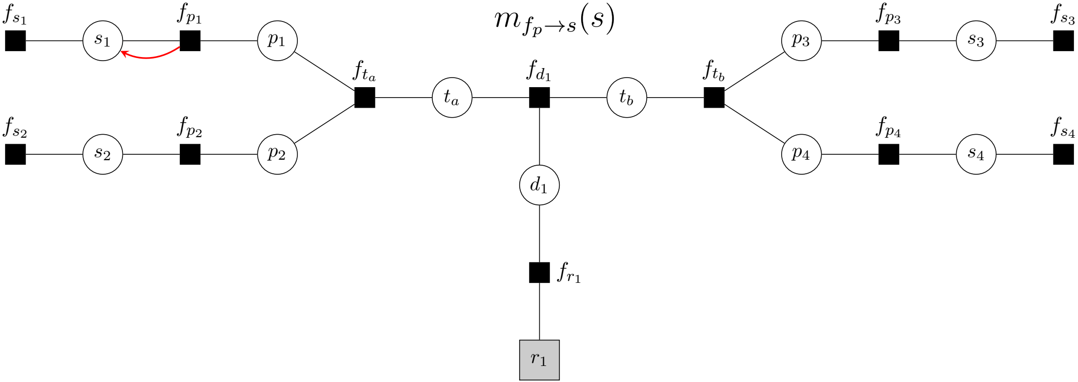
$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \Phi(\mu_2 | \mu_b - p_1, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$

$$\frac{\partial}{\partial \mu_2} m_{f_{s_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \mathcal{N}(p_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$



$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \Phi(\mu_2 | \mu_b - p_1, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$

$$\frac{\partial}{\partial \mu_2} m_{f_{s_1} \to s_1}(s_1) = \mathcal{N}(s_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

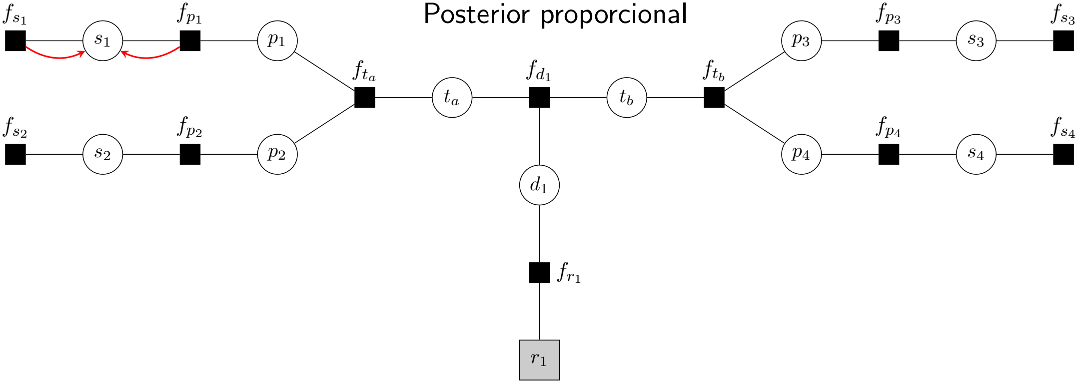


$$m_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\, \Phi(\mu_2 | \mu_b - p_1, \sigma_b^2 + \sigma_2^2 + \beta^2)\, dp_1$$

$$\frac{\partial}{\partial \mu_2} m_{f_{s_1} \to s_1}(s_1) = \mathcal{N}(s_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

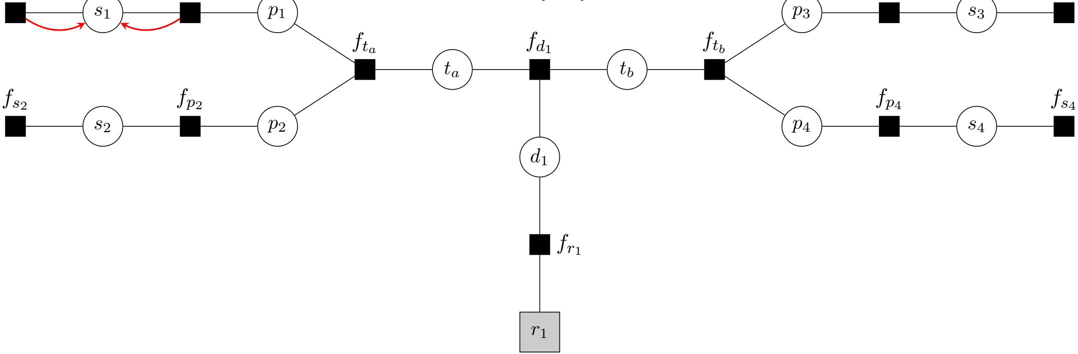
$$m_{f_{p_1} \to s_1}(s_1) = \Phi(s_1 | \mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

Posterior proporcional



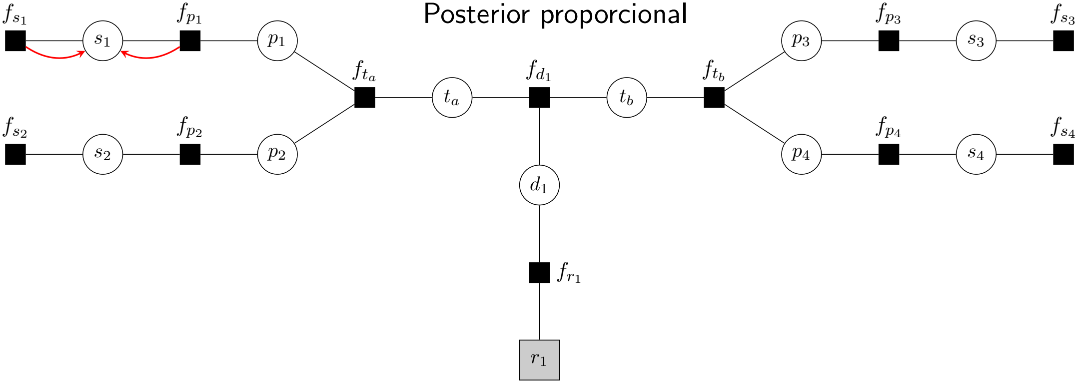
$$p(s_1|r) \propto p(s_1, r) =$$

Posterior proporcional



$$p(s_1|r) \propto p(s_1, r) = \mathcal{N}(s_1|\mu_1, \sigma_1^2)\, \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

Posterior proporcional



$$p(s_1|r) \propto p(s_1, r) = \underbrace{\mathcal{N}(s_1|\mu_1, \sigma_1^2)}_{\text{Prior}} \underbrace{\Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)}_{\text{Verosimilitud}}$$

Verosimilitud

$$p(r|s_1) = \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

Verosimilitud

$$
\begin{aligned}
p(r|s_1) &= \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2) \\
&= \Phi(\,0\,|\mu_b - \mu_2 - s_1, \sigma_b^2 + \sigma_2^2 + 2\beta^2)
\end{aligned}
$$

Verosimilitud

$$p(r|s_1) = \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

$$= \Phi(\,0\,|\boldsymbol{\mu_b} - \mu_2 - s_1, \boldsymbol{\sigma_b^2} + \sigma_2^2 + 2\beta^2)$$

Verosimilitud

$$p(r|s_1) = \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2)$$

$$= \Phi(\,0\,|\boldsymbol{\mu_3} + \boldsymbol{\mu_4} - \mu_2 - s_1, \boldsymbol{\sigma_3^2} + \boldsymbol{\sigma_4^2} + \boldsymbol{2\beta^2} + \sigma_2^2 + 2\beta^2)$$

Verosimilitud

$$
\begin{aligned}
p(r|s_1) &= \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2) \\
&= \Phi(\,0\,|(\mu_3 + \mu_4) - (s_1 + \mu_2), 4\beta^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2) \\
&=
\end{aligned}
$$

Verosimilitud

$$
\begin{aligned}
p(r|s_1) &= \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2) \\
&= \Phi(\,0\,|(\mu_3 + \mu_4) - (s_1 + \mu_2), 4\beta^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2) \\
&= 1 - \Phi(\,0\,|\,(s_1 + \mu_2) - (\mu_3 + \mu_4), 4\beta^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2)
\end{aligned}
$$

Verosimilitud

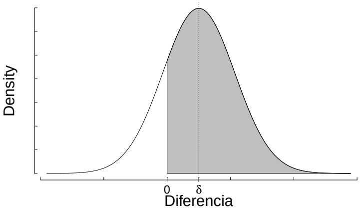
$$
\begin{aligned}
p(r|s_1) &= \Phi(s_1|\mu_b - \mu_2, \sigma_b^2 + \sigma_2^2 + 2\beta^2) \\
&= \Phi(\,0\,|(\mu_3 + \mu_4) - (s_1 + \mu_2), 4\beta^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2) \\
&= 1 - \Phi(\,0\,|\underbrace{(s_1 + \mu_2) - (\mu_3 + \mu_4)}_{\substack{\text{Diferencia esperada} \\ \text{si la habilidad fuera } s_1}}, \quad \underbrace{4\beta^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2}_{\substack{\text{Incertidumbre total} \\ \text{salvo la de la habilidad } s_1}} \quad)
\end{aligned}
$$

Evidencia vs Verosimilitud

$$p(r) = 1 - \Phi\Big(0 \,|\, \underbrace{(\boldsymbol{\mu_1} + \mu_2) - (\mu_3 + \mu_4)}_{\text{Diferencia esperada}}, \underbrace{4\beta^2 + \boldsymbol{\sigma_1^2} + \sigma_2^2 + \sigma_3^2 + \sigma_4^2}_{\text{Incertidumbre total}}\Big)$$

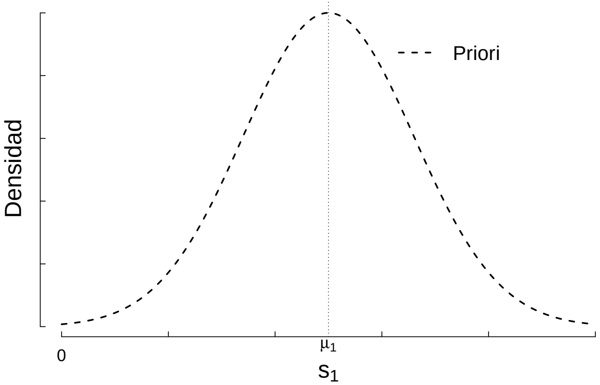
$$p(r|s_1) = 1 - \Phi(\,0 \,|\, \underbrace{(s_1 + \mu_2) - (\mu_3 + \mu_4)}_{\substack{\text{Diferencia esperada}\\ \text{si la habilidad fuera } s_1}}, \quad \underbrace{4\beta^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2}_{\substack{\text{Incertidumbre total}\\ \text{salvo la de la habilidad } s_1}} \quad)$$

Predicciones del resultado
usando todas las hipótesis
pesadas por la creencia a prior

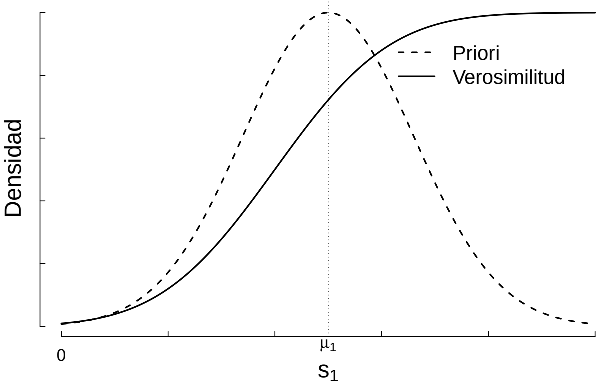


$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \; \overbrace{1 - \Phi(\, 0 \,|\, \delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$

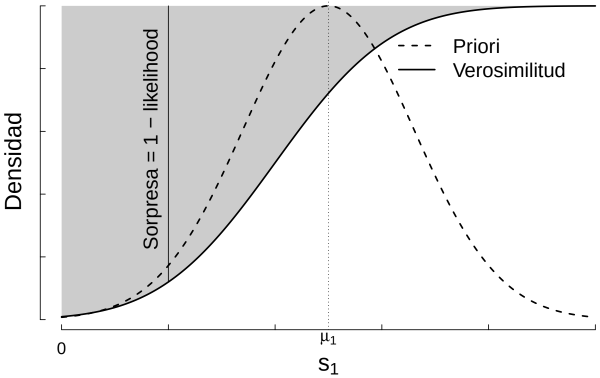
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \; \overbrace{1 - \Phi(\, 0 \,|\, \delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



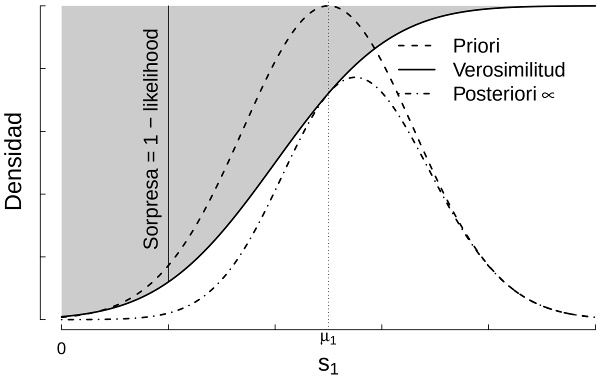
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \ \overbrace{1 - \Phi(\,0\,|\,\delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



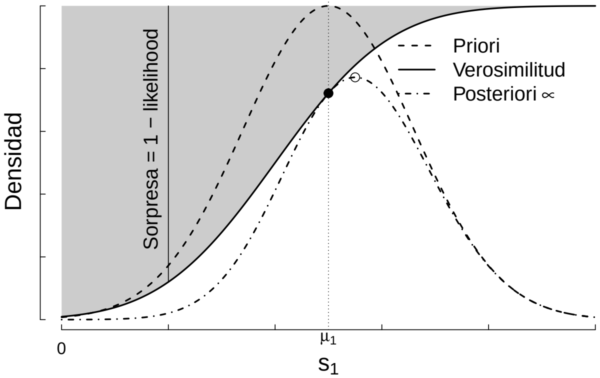
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \; \overbrace{1 - \Phi(\,0\,|\,\delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



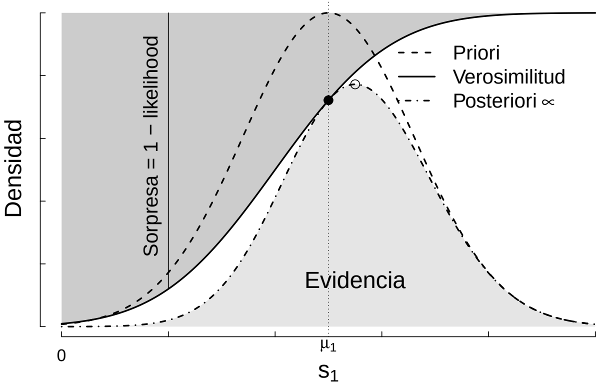
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \; \overbrace{1 - \Phi(\,0\,|\,\delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



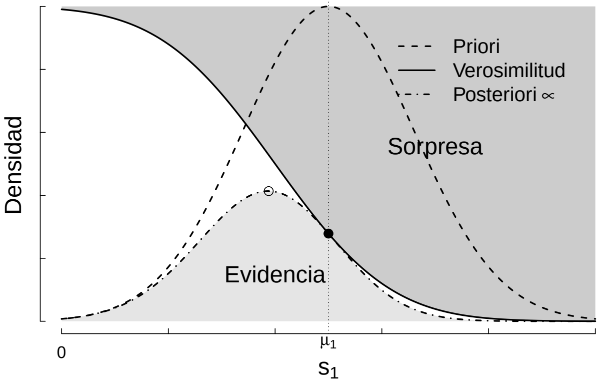
$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \; \overbrace{1 - \Phi(\,0 \,|\, \delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



$$\overbrace{P(s_1 \mid r, \text{Modelo})}^{\text{Posterior}} \propto \overbrace{\mathcal{N}(s_1 \,|\, \mu_1, \sigma_1^2)}^{\text{Prior}} \; \overbrace{1 - \Phi(\, 0 \,|\, \delta - \mu_1 + s_1, \vartheta^2 - \sigma_1^2)}^{\text{Verosimilitud}} \quad \text{Caso ganador}$$



$$\overbrace{P(s_3 \mid r, \text{Modelo})}^{\text{Posteriori}} \propto \overbrace{N(s_3 \,|\, \mu_3, \sigma_2^2)}^{\text{Priori}} \; \overbrace{\Phi(0 \,|\delta - \mu_3 - s_3, \vartheta^2 - \sigma_3^2)}^{\text{Verosimilitud}} \quad \text{Caso perdedor}$$



TrueSkill

Posterior aproximado

$$\underbrace{\widehat{p}(s_a|\text{Gana}, M)}_{\text{Aproximado}} = \underset{\mu,\sigma}{\text{arg min}}\ \mathcal{D}_{KL}(\ \underbrace{p(s_a|\text{Gana}, M)}_{\text{Exacto}}\ ||\ \underbrace{\mathcal{N}(s_a|\mu, \sigma^2)}_{\text{Familia}}\ )$$

Mínima Divergencia

Posterior aproximado

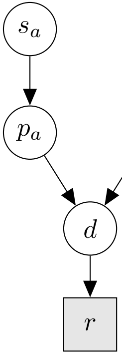
$$\underbrace{\widehat{p}(s_a|\text{Gana}, M)}_{\text{Aproximado}} = \underset{\mu,\sigma}{\text{arg min}}\ \mathcal{D}_{KL}(\ \underbrace{p(s_a|\text{Gana}, M)}_{\text{Exacto}}\ ||\ \underbrace{\mathcal{N}(s_a|\mu,\sigma^2)}_{\text{Familia}}\ )$$

$$\mathcal{D}_{KL}(\text{Exacta}||\text{Familia}) = \sum_{s_a} p(s_a|\text{Gana}, M) \cdot \big(\log p(s_a|\text{Gana}, M) - \log \mathcal{N}(s_a|\mu,\sigma^2)\big)$$

Mínima Divergencia

Posterior aproximado

$$\underbrace{\widehat{p}(s_a|\text{Gana}, M)}_{\text{Aproximado}} = \underset{\mu,\sigma}{\text{arg min}}\ \mathcal{D}_{KL}(\ \underbrace{p(s_a|\text{Gana}, M)}_{\text{Exacto}}\ ||\ \underbrace{\mathcal{N}(s_a|\mu,\sigma^2)}_{\text{Familia}}\ )$$



Habilidad

$$p(s_i) = \mathcal{N}(s_i|\mu_i,\sigma_i^2)$$

Desempeño

$$P(p_i|s_i) = \mathcal{N}(s_i,\beta^2)$$

Diferencia

$$P(d|p_a,p_b) = \delta(d = p_i - p_j)$$

Resultado

$$P(r|d) = \mathbb{I}(r = d > 0)$$

Mínima Divergencia
Posterior aproximado

$$\underbrace{\widehat{p}(s_a|\text{Gana}, M)}_{\text{Aproximado}} = \underset{\mu,\sigma}{\text{arg min}}\ \mathcal{D}_{KL}(\ \underbrace{p(s_a|\text{Gana}, M)}_{\text{Exacto}}\ ||\ \underbrace{\mathcal{N}(s_a|\mu, \sigma^2)}_{\text{Familia}}\ )$$

$$p(d, \text{Gana} \mid \text{Modelo}) =$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{\text{Marginal exacta}}$$

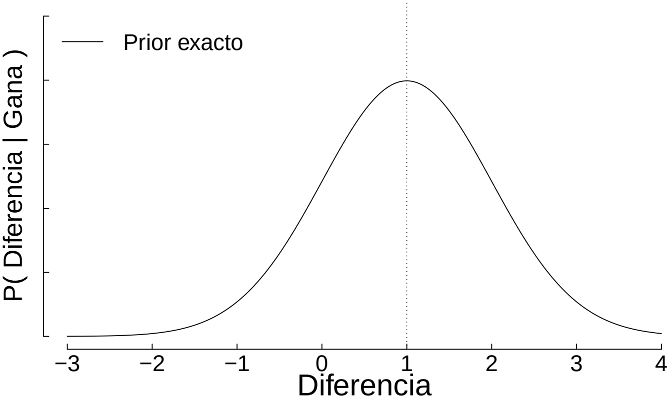
Mínima Divergencia
Posterior aproximado

$$\underbrace{\widehat{p}(s_a|\text{Gana}, M)}_{\text{Aproximado}} = \underset{\mu,\sigma}{\text{arg min}}\ \mathcal{D}_{KL}(\ \underbrace{p(s_a|\text{Gana}, M)}_{\text{Exacto}}\ ||\ \underbrace{\mathcal{N}(s_a|\mu,\sigma^2)}_{\text{Familia}}\ )$$

$$\underbrace{p(d, \text{Gana} \mid \text{Modelo})}_{\text{Marginal exacta}} = \underbrace{\int_{s_a}\int_{p_a}\int_{s_b}\int_{p_b} p(s_a)p(s_b)p(p_a|s_a)p(p_b|s_b)p(d|p_a,p_b)}_{\text{Prior exacto } p(d) = \mathcal{N}(d|\ \underbrace{\mu_a - \mu_b}_{\text{Diferencia de medias}}, \underbrace{\sigma_a^2 + \sigma_b^2 + 2\beta^2}_{\text{Suma de incertidumbres}})}\ P(\text{Gana}|d)$$

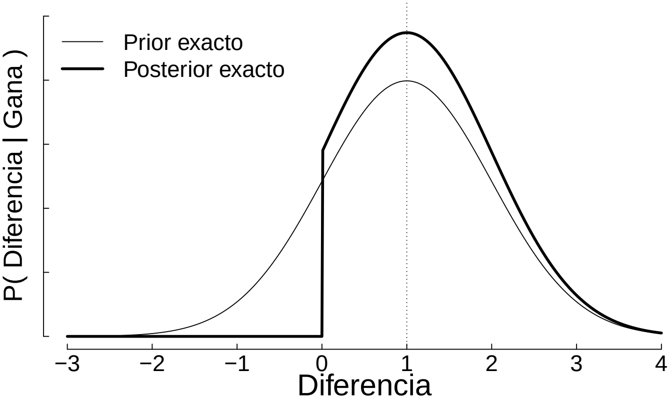
Mínima Divergencia
Posterior aproximado

$$p(d, \text{Gana}) = \mathcal{N}(d|\mu_a - \mu_b, \sigma_a^2 + \sigma_b^2 + 2\beta^2)\mathbb{I}(d > 0)$$



Mínima Divergencia
Posterior aproximado

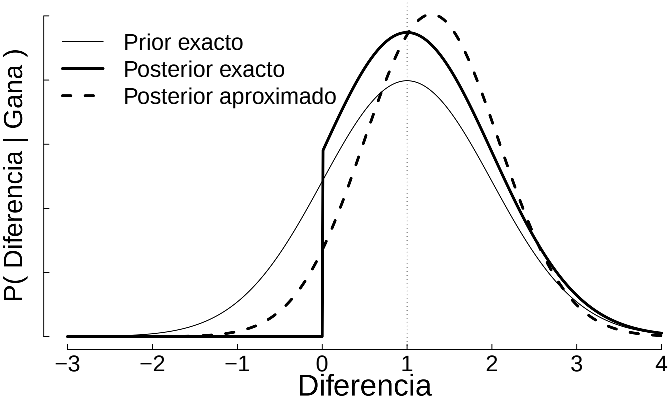
$$p(d, \mathsf{Gana}) = \mathcal{N}(d|\mu_a - \mu_b, \sigma_a^2 + \sigma_b^2 + 2\beta^2)\mathbb{I}(d > 0)$$



Mínima Divergencia

Posterior aproximado

$$p(d, \mathsf{Gana}) = \mathcal{N}(d|\mu_a - \mu_b, \sigma_a^2 + \sigma_b^2 + 2\beta^2)\mathbb{I}(d > 0)$$



Mínima Divergencia

Posterior aproximado

$$p(d, \textsf{Gana}) = \mathcal{N}(d|\mu_a - \mu_b, \sigma_a^2 + \sigma_b^2 + 2\beta^2)\mathbb{I}(d > 0)$$

$$\widehat{p}(d, \textsf{Gana}) \propto \mathcal{N}\Big(d \,|\, \underbrace{E(d, \textsf{Gana}|d > 0)}_{\substack{\textsf{Misma}\\ \textsf{media}}}, \underbrace{V(d, \textsf{Gana}|d > 0)}_{\substack{\textsf{Misma}\\ \textsf{varianza}}}\Big)$$

Mínima Divergencia
Posterior aproximado

$$p(d, \textsf{Gana}) = \mathcal{N}(d|\mu_a - \mu_b, \sigma_a^2 + \sigma_b^2 + 2\beta^2)\mathbb{I}(d > 0)$$

$$\widehat{p}(d, \textsf{Gana}) \propto \mathcal{N}\Big(d \mid \underbrace{E(d, \textsf{Gana}|d > 0)}_{\substack{\text{Misma}\\\text{media}}}, \underbrace{V(d, \textsf{Gana}|d > 0)}_{\substack{\text{Misma}\\\text{varianza}}}\Big)$$

La divergencia es mínima cuando tienen mismos momentos.
(Expectation Propagation)

Mínima Divergencia

Aproximación del posterior de la diferencia

$$p(d, \mathsf{Gana}) = m_{f_d \rightarrow d}(d)\, m_{f_r \rightarrow d}(d)$$

Mínima Divergencia
Aproximación del posterior de la diferencia

$$p(d, \mathsf{Gana}) = m_{f_d \rightarrow d}(d)\, m_{f_r \rightarrow d}(d)$$

$$m_{f_r \rightarrow d}(d) = m_{d \rightarrow f_d}(d)$$

Mínima Divergencia
Aproximación del posterior de la diferencia

$$p(d, \text{Gana}) = m_{f_d \rightarrow d}(d) \, m_{d \rightarrow f_d}(d)$$

Mínima Divergencia
Aproximación del posterior de la diferencia

$$p(d, \mathsf{Gana}) = \underbrace{m_{f_d \rightarrow d}(d)}_{P(d)} m_{d \rightarrow f_d}(d)$$

Mínima Divergencia
Aproximación del posterior de la diferencia

$$p(d, \mathsf{Gana}) = \underbrace{m_{f_d \rightarrow d}(d)}_{P(d)} m_{d \rightarrow f_d}(d)$$

$$m_{d \rightarrow f_d}(d) = \frac{p(d, \mathsf{Gana})}{p(d)}$$

Mínima Divergencia
Aproximación del posterior de la diferencia

$$p(d, \text{Gana}) = \underbrace{m_{f_d \rightarrow d}(d)}_{P(d)} m_{d \rightarrow f_d}(d)$$

$$m_{d \rightarrow f_d}(d) = \frac{p(d, \text{Gana})}{p(d)} \approx \frac{\widehat{p}(d, \text{Gana})}{p(d)}$$

Mínima Divergencia
Aproximación del posterior de la diferencia

$$p(d, \textsf{Gana}) = \underbrace{m_{f_d\rightarrow d}(d)}_{P(d)} m_{d\rightarrow f_d}(d)$$

$$m_{d\rightarrow f_d}(d) = \frac{p(d, \textsf{Gana})}{p(d)} \approx \frac{\widehat{p}(d, \textsf{Gana})}{p(d)}$$

$$= \frac{\mathcal{N}(d \,|\, \widehat{\delta},\, \widehat{\vartheta}^{\,2})}{\mathcal{N}(d|\delta, \vartheta^2)}$$

Mínima Divergencia

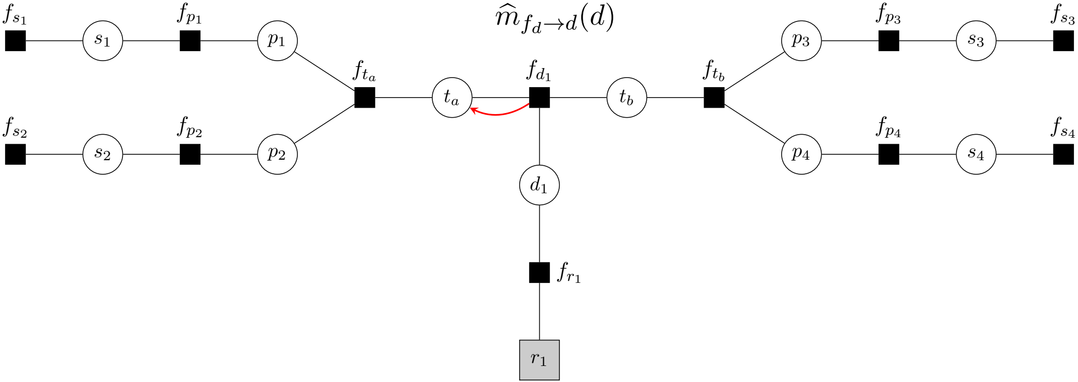
Aproximación del posterior de la diferencia

$$p(d, \text{Gana}) = \underbrace{m_{f_d \rightarrow d}(d)}_{P(d)} m_{d \rightarrow f_d}(d)$$

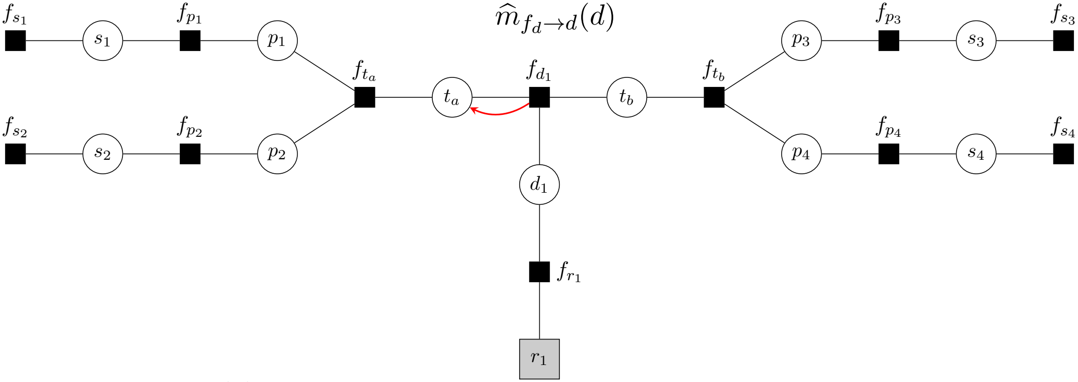
$$m_{d \rightarrow f_d}(d) = \frac{p(d, \text{Gana})}{p(d)} \approx \frac{\widehat{p}(d, \text{Gana})}{p(d)}$$

$$= \frac{\mathcal{N}(d \,|\, \widehat{\delta}, \, \widehat{\vartheta}^{\,2})}{\mathcal{N}(d|\delta, \vartheta^2)} \propto \mathcal{N}(d|\delta_{\div}, \vartheta^2_{\div})$$

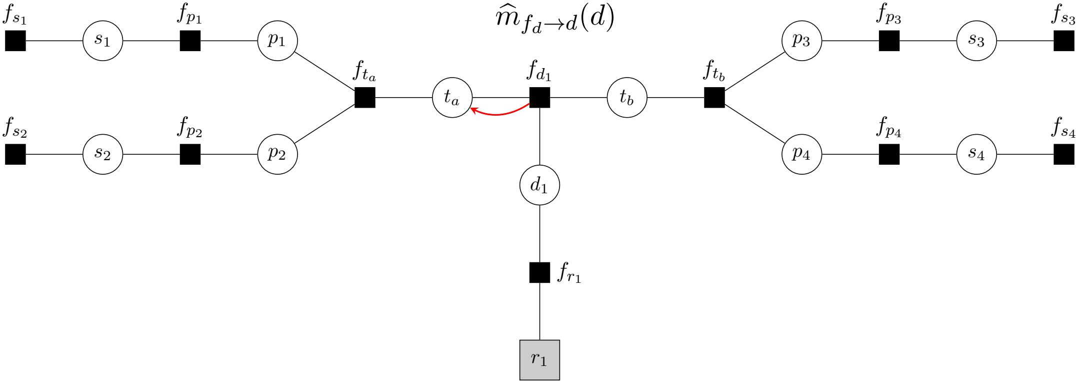
con $\delta_{\div} = \frac{\widehat{\delta}}{\widehat{\vartheta}^2} - \frac{\delta}{\vartheta^2}$ y $\vartheta^2_{\div} = (\frac{1}{\widehat{\vartheta}^2} - \frac{1}{\vartheta^2})^{-1}$



$$\widehat{m}_{f_d \to t_a}(t_a) \approx \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(d|\delta_{\div}, \vartheta_{\div}^2)\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, ddd_{t_b}$$

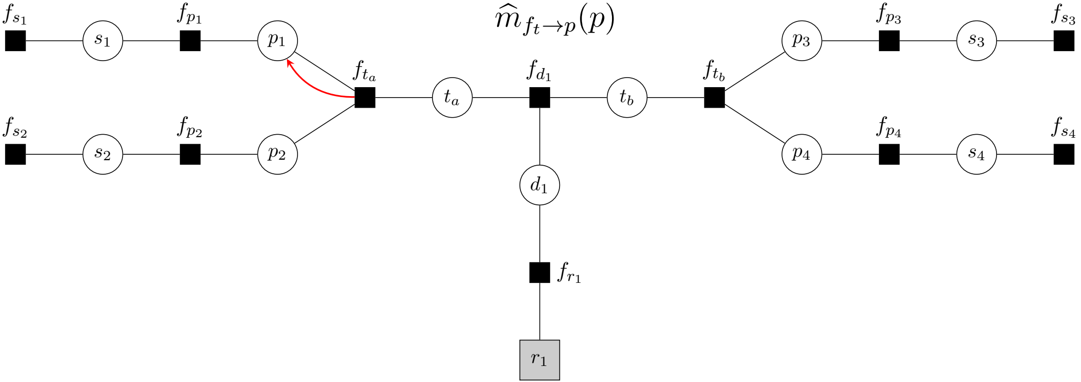


$$
\begin{aligned}
\widehat{m}_{f_d \to t_a}(t_a) &\approx \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(d|\delta_{\div}, \vartheta^2_{\div})\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, ddd_{t_b} \\
&= \int \mathcal{N}(t_a - t_b|\delta_{\div}, \vartheta^2_{\div})\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, d_{t_b}
\end{aligned}
$$



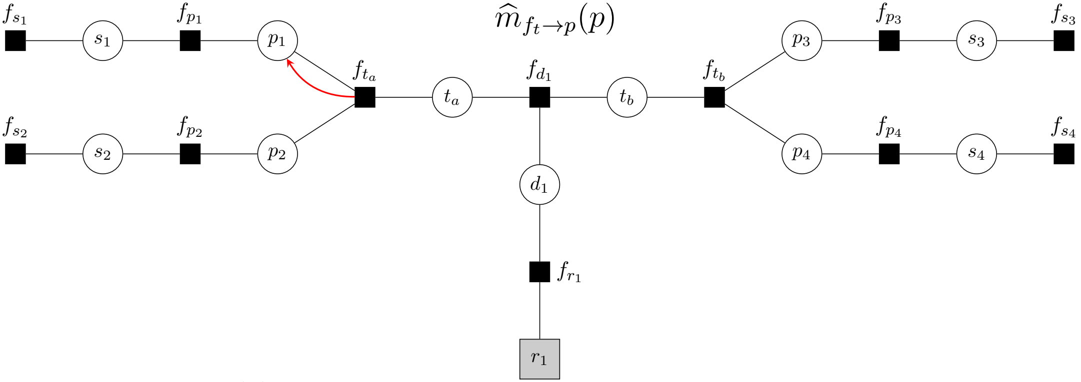
$$\begin{aligned}
\widehat{m}_{f_d \to t_a}(t_a) &\approx \iint \mathbb{I}(d = t_a - t_b)\mathcal{N}(d|\delta_{\div}, \vartheta^2_{\div})\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, d d d_{t_b} \\
&= \int \mathcal{N}(t_a - t_b|\delta_{\div}, \vartheta^2_{\div})\mathcal{N}(t_b|\mu_b, \sigma_b^2)\, d_{t_b} \\
&= \mathcal{N}(t_a \mid \mu_b + \delta_{\div}\, ,\, \vartheta^2_{\div} + \sigma_b^2)
\end{aligned}$$

$$\widehat{m}_{f_t \to p}(p)$$

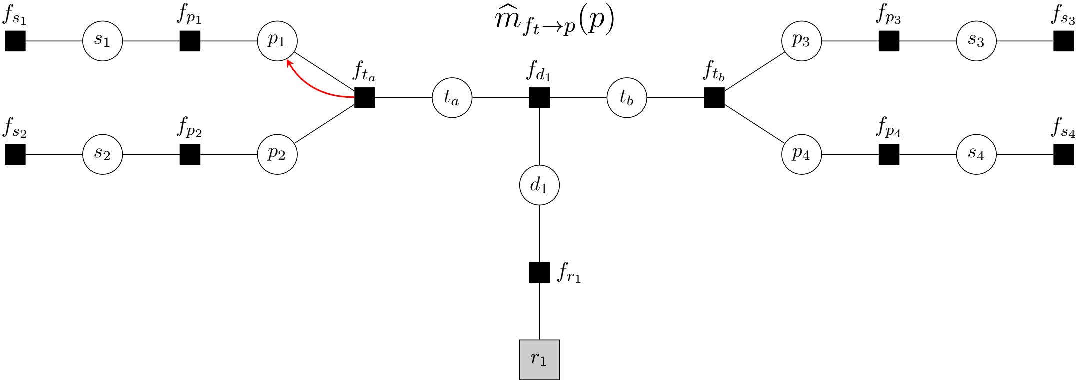


$$\widehat{m}_{f_{t_a} \to p_1}(p_1) = \iint \mathbb{I}(t_a = p_1 + p_2)\mathcal{N}(t_a \,|\, \mu_b + \delta_{\div}\,,\, \vartheta_{\div}^2 + \sigma_b^2)\mathcal{N}(p_2|\mu_2, \sigma_2^2 + \beta^2)\, dt_a d_{p_2}$$

$$\widehat{m}_{f_t \to p}(p)$$

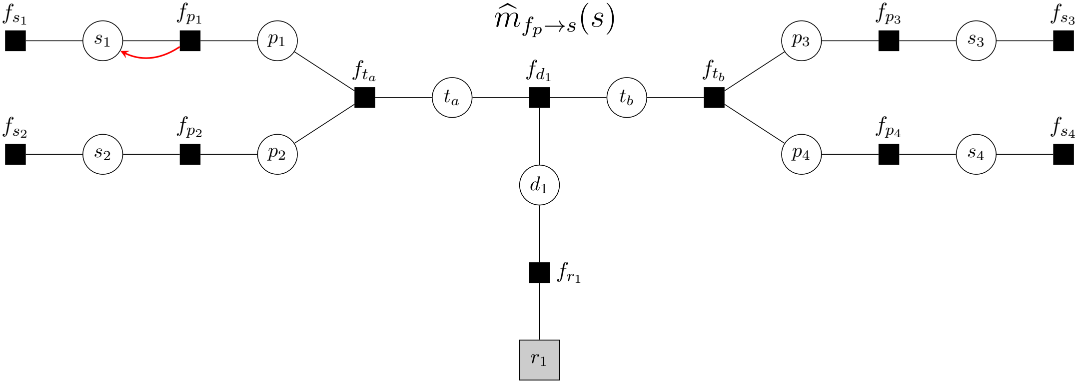


$$\begin{aligned}
\widehat{m}_{f_{t_a} \to p_1}(p_1) &= \iint \mathbb{I}(t_a = p_1 + p_2)\mathcal{N}(t_a \,|\, \mu_b + \delta_{\div}\,,\, \vartheta^2_{\div} + \sigma_b^2)\mathcal{N}(p_2|\mu_2, \sigma_2^2 + \beta^2)\, dt_a d_{p_2} \\
&= \int \mathcal{N}(p_1 + p_2 \,|\, \mu_b + \delta_{\div}\,,\, \vartheta^2_{\div} + \sigma_b^2)\mathcal{N}(p_2|\mu_2, \sigma_2^2 + \beta^2)\, d_{p_2}
\end{aligned}$$

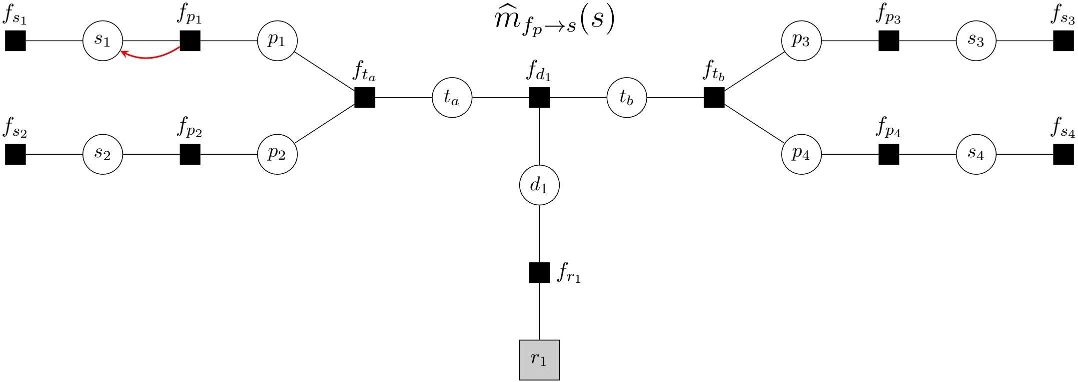


$$
\begin{aligned}
\widehat{m}_{f_{t_a} \to p_1}(p_1) &= \iint \mathbb{I}(t_a = p_1 + p_2)\mathcal{N}(t_a \mid \mu_b + \delta_{\div}\,,\ \vartheta^2_{\div} + \sigma_b^2)\mathcal{N}(p_2|\mu_2, \sigma_2^2 + \beta^2)\, dt_a d_{p_2} \\
&= \int \mathcal{N}(p_1 + p_2 \mid \mu_b + \delta_{\div}\,,\ \vartheta^2_{\div} + \sigma_b^2)\mathcal{N}(p_2|\mu_2, \sigma_2^2 + \beta^2)\, d_{p_2} \\
&= \mathcal{N}(p_1 \mid \underbrace{\mu_b - \mu_2}_{\mu_1 - \delta} + \delta_{\div}\,,\ \vartheta^2_{\div} + \underbrace{\sigma_b^2 + \sigma_2^2 + \beta^2}_{\vartheta^2 - (\sigma_1^2 + \beta^2)})
\end{aligned}
$$

$$\widehat{m}_{f_p \to s}(s)$$



$$\widehat{m}_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1 | s_1, \beta^2)\mathcal{N}(p_1 | \mu_1 - \delta + \delta_{\div}, \vartheta^2_{\div} + \vartheta^2 - \sigma_1^2 - \beta^2) dp_1$$



$$\widehat{m}_{f_{p_1} \to s_1}(s_1) = \int \mathcal{N}(p_1|s_1, \beta^2)\mathcal{N}(p_1|\mu_1 - \delta + \delta_{\div}, \vartheta_{\div}^2 + \vartheta^2 - \sigma_1^2 - \beta^2)dp_1$$
$$= \mathcal{N}(s_1|\mu_1 - \delta + \delta_{\div}, \vartheta_{\div}^2 + \vartheta^2 - \sigma_1^2)$$

Posterior aproximado

$$\widehat{p}(s_1, r) = \mathcal{N}(s_1|\mu_1, \sigma_1^2)\mathcal{N}(s_1|\mu_1 - \delta + \delta_{\div}, \vartheta_{\div}^2 + \vartheta^2 - \sigma_1^2)$$

Laboratorios de
Métodos Bayesianos

Bibliografía Unidad 5

- Neal. **Pattern Recognition and Machine Learning**. 2020 (Draft). (Descargar). (lectura capítulo 2, 4, 5 y 6)