Factorización de árboles de probabilidad y sus aplicaciones

I. Martínez*, S. Moral**, C. Rodríguez*, A. Salmerón*

*Dept. Estadística y Mat. Aplicada
Universidad de Almería

**Dept. Ciencias de la Computación e I.A.
Universidad de Granada

Índice

¿Para qué factorizar?

- En cálculo de probabilidades, FACTORIZAR $\Rightarrow$ MÁS EFICIENCIA

¿Para qué factorizar?

- En cálculo de probabilidades, FACTORIZAR $\Rightarrow$ MÁS EFICIENCIA
- Modelos multivariantes, gran número de variables, f.m.p. sin expresión analítica: MUY FRECUENTES EN LA PRÁCTICA

¿Para qué factorizar?

- En cálculo de probabilidades, FACTORIZAR $\Rightarrow$ MÁS EFICIENCIA
- Modelos multivariantes, gran número de variables, f.m.p. sin expresión analítica: MUY FRECUENTES EN LA PRÁCTICA
- Antes de factorizar: INTRATABLES

¿Para qué factorizar?

- En cálculo de probabilidades, FACTORIZAR $\Rightarrow$ MÁS EFICIENCIA
- Modelos multivariantes, gran número de variables, f.m.p. sin expresión analítica: MUY FRECUENTES EN LA PRÁCTICA
- Antes de factorizar: INTRATABLES
- Después de factorizar: TRATABLES

Niveles de factorización

- PRIMER NIVEL: Red bayesiana $\Leftrightarrow$ Join tree.

Niveles de factorización

- PRIMER NIVEL: Red bayesiana $\Leftrightarrow$ Join tree.
    - Concepto subyacente: Independencia condicional

Niveles de factorización

- PRIMER NIVEL: Red bayesiana $\Leftrightarrow$ Join tree.
  - Concepto subyacente: Independencia condicional
- SEGUNDO NIVEL: Potenciales asociados a los nodos del join tree.

Niveles de factorización

- PRIMER NIVEL: Red bayesiana $\Leftrightarrow$ Join tree.
  - Concepto subyacente: Independencia condicional
- SEGUNDO NIVEL: Potenciales asociados a los nodos del join tree.

Un **potencial** es una función

$$\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}_0^+$$

con soporte finito

Árboles de probabilidad

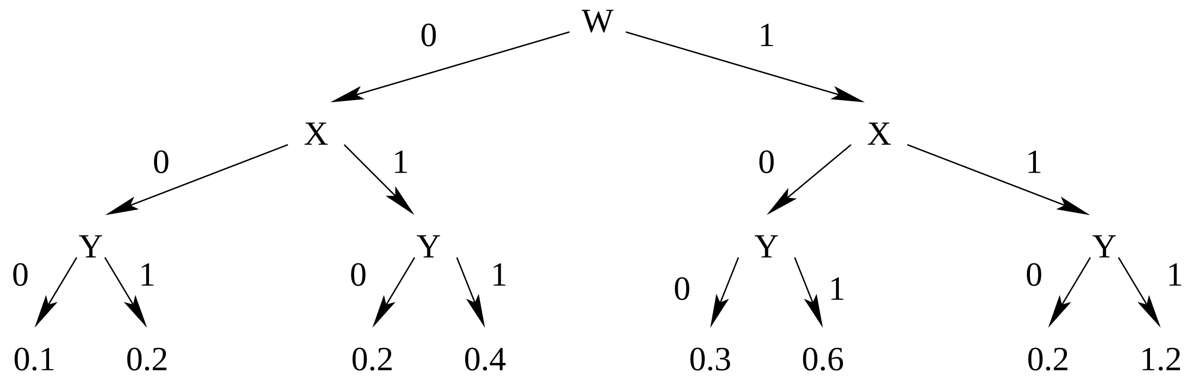
Un potencial se puede representar como un árbol:

w	x	y	$\psi(w,x,y)$
0	0	0	0.1
0	0	1	0.2
0	1	0	0.2
0	1	1	0.4
1	0	0	0.3
1	0	1	0.6
1	1	0	0.6
1	1	1	1.2

Árboles de probabilidad

Un potencial se puede representar como un árbol:

w	x	y	$\psi(w,x,y)$
0	0	0	0.1
0	0	1	0.2
0	1	0	0.2
0	1	1	0.4
1	0	0	0.3
1	0	1	0.6
1	1	0	0.6
1	1	1	1.2



Factorización de árboles de prob.

¿Se puede factorizar un árbol?

Factorización de árboles de prob.

¿Se puede factorizar un árbol? bajo ciertas condiciones, sí

Factorización de árboles de prob.

¿Se puede factorizar un árbol? bajo ciertas condiciones, sí

Definición: Sea ψ un potencial sobre $\mathbf{X} = \{X_1, \ldots, X_n\}$ representado por un árbol $\mathcal{T}$. Sea $\mathbf{X}_C \subseteq \mathbf{X}$ y $(\mathbf{X}_C = \mathbf{x}_C)$ una configuración de variables que llevan desde $\mathcal{T}$ hasta una variable X_i. Decimos que $\mathcal{T}$ es factorizable en X_i para el contexto $(\mathbf{X}_C = \mathbf{x}_C)$, si para todo $x, y \in \Omega_{X_i}$, $\exists \varepsilon > 0$ t.q.

$$\mathcal{T}^{R(\mathbf{X}_C = \mathbf{x}_C, X_i = x)} = \varepsilon \cdot \mathcal{T}^{R(\mathbf{X}_C = \mathbf{x}_C, X_i = y)} \quad .$$

Factorización de árboles de prob.

Definición: Sea $\mathcal{T}$ un árbol de probabilidad. Definimos $\mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i, x, y, \varepsilon)$ como el árbol que se obtiene reemplazando, en $\mathcal{T}$, $\mathcal{T}^{R(\mathbf{X}_C=\mathbf{x}_C, X_i=x)}$ por la constante 1 y $\mathcal{T}^{R(\mathbf{X}_C=\mathbf{x}_C, X_i=y)}$ por ε.

Factorización de árboles de prob.

Definición: Sea $\mathcal{T}$ un árbol de probabilidad. Definimos $\mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i, x, y, \varepsilon)$ como el árbol que se obtiene reemplazando, en $\mathcal{T}$, $\mathcal{T}^{R(\mathbf{X}_C=\mathbf{x}_C, X_i=x)}$ por la constante 1 y $\mathcal{T}^{R(\mathbf{X}_C=\mathbf{x}_C, X_i=y)}$ por ε.

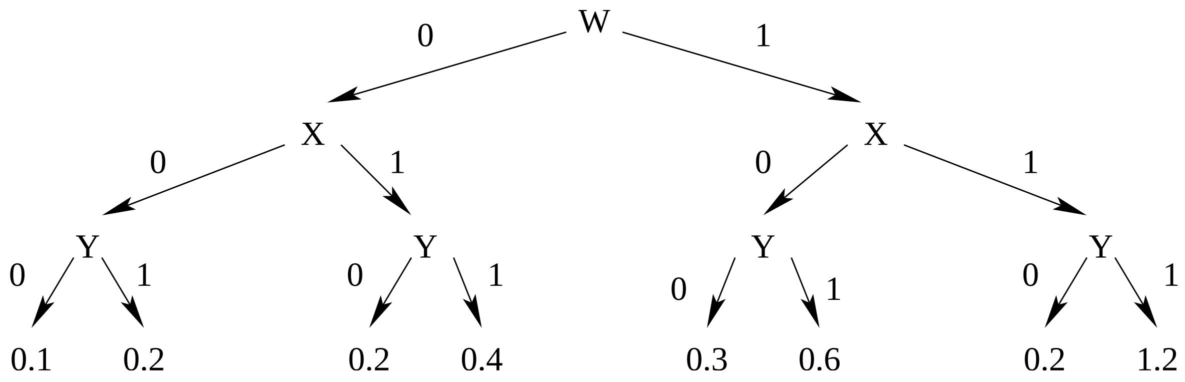
Definición: Sea $\mathcal{T}$ un árbol de probabilidad. Definimos $\mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i, x)$ como el árbol que se obtiene reemplazando, en $\mathcal{T}$, $\mathcal{T}^{R(\mathbf{X}_C=\mathbf{x}_C)}$ por $\mathcal{T}^{R(\mathbf{X}_C=\mathbf{x}_C, X_i=x)}$ y $\mathcal{T}^{R(\mathbf{X}_D=\mathbf{x}_D)}$ por 1 para cualquier contexto $(\mathbf{X}_D = \mathbf{x}_D)$ incompatible con $(\mathbf{X}_C = \mathbf{x}_C)$.

Factorización de árboles de prob.

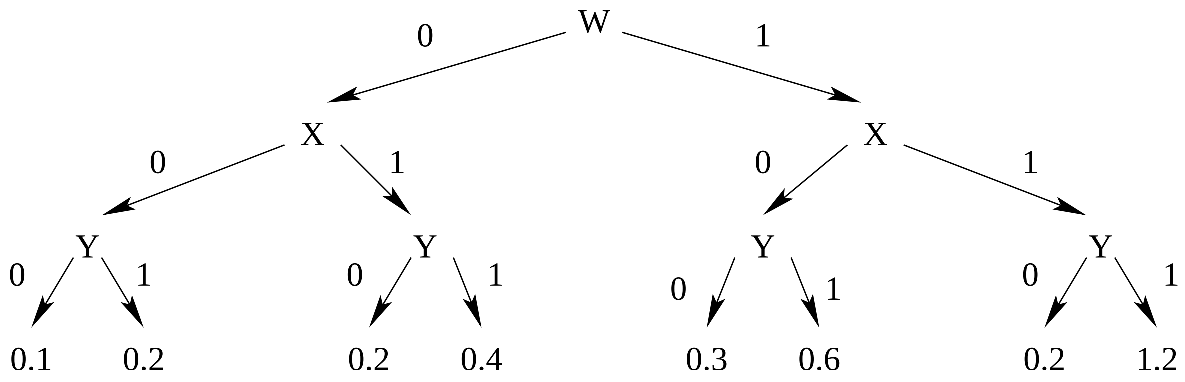
Proposición: Sea $\mathcal{T}$ un árbol de probabilidad factorizable en X_i con factor de proporcionalidad ε para el contexto $(\mathbf{X}_C = \mathbf{x}_C)$. Entonces,

$$\mathcal{T} = \mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i, x, y, \varepsilon) \otimes \mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i = x)$$

Ejemplo

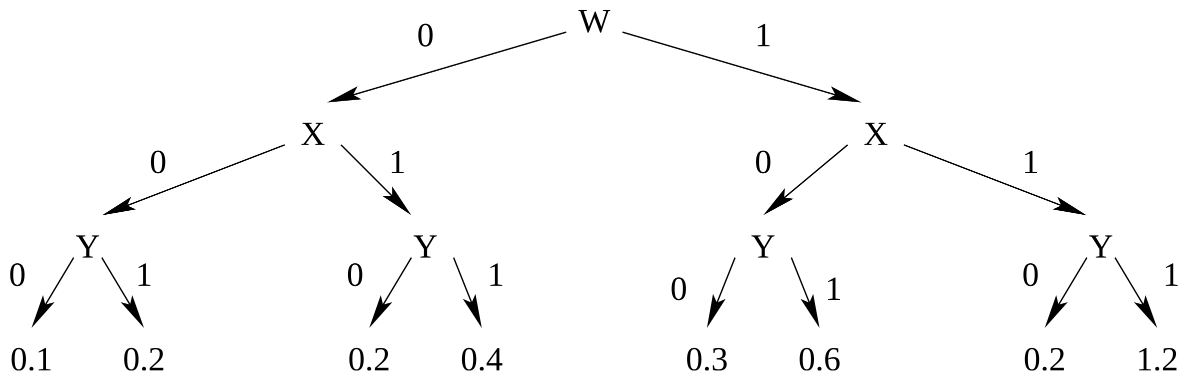


Ejemplo

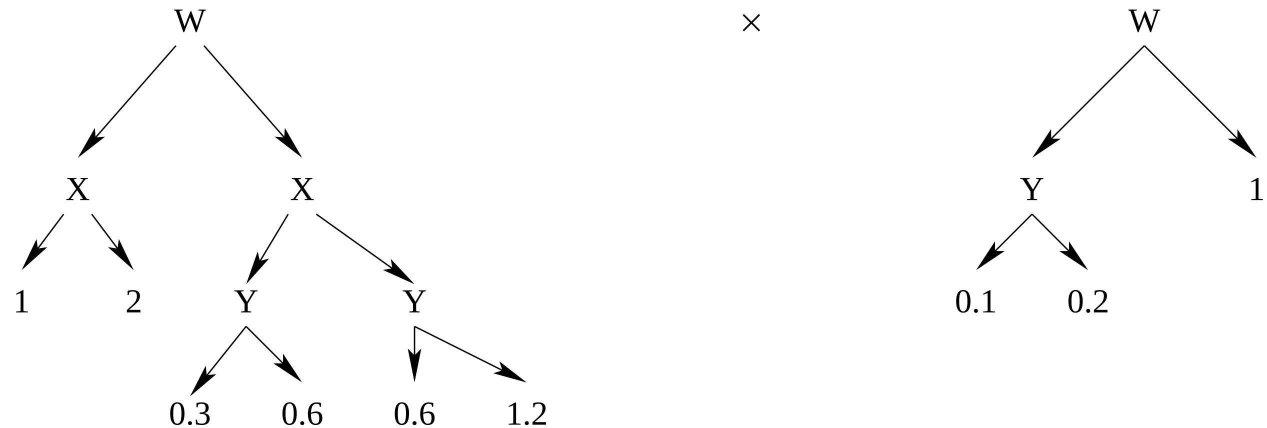


$\mathcal{T}$ es factorizable en X para el contexto $(W = 0)$

Ejemplo



$\mathcal{T}$ es factorizable en X para el contexto $(W = 0)$



Aplicaciones: propagación perezosa

La factorización es útil cuando:

- Se va a borrar una variable X que está en más de un árbol.

Aplicaciones: propagación perezosa

La factorización es útil cuando:

- Se va a borrar una variable X que está en más de un árbol.
- Alguno de los árboles es factorizable en X para algún contexto.

Aplicaciones: propagación perezosa

La factorización es útil cuando:

- Se va a borrar una variable X que está en más de un árbol.
- Alguno de los árboles es factorizable en X para algún contexto.
- Es tanto más efectiva cuanto más cerca de la raíz esté X.

Justificación

Proposición: Ninguna variable aparece más de una vez en una misma rama de un árbol de probabilidad.

Justificación

Proposición: Ninguna variable aparece más de una vez en una misma rama de un árbol de probabilidad.

Corolario: $\mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i)$ no contiene a X_i.

Justificación

Proposición: Ninguna variable aparece más de una vez en una misma rama de un árbol de probabilidad.

Corolario: $\mathcal{T}(\mathbf{X}_C = \mathbf{x}_C, X_i)$ no contiene a X_i.

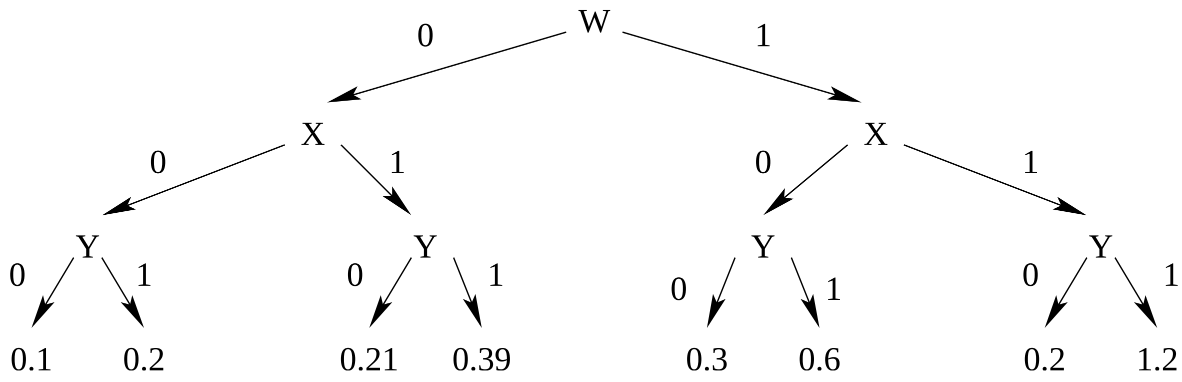
PROBLEMA: Puede ser difícil que se dé la proporcionalidad en un árbol

Solución: factorización aproximada

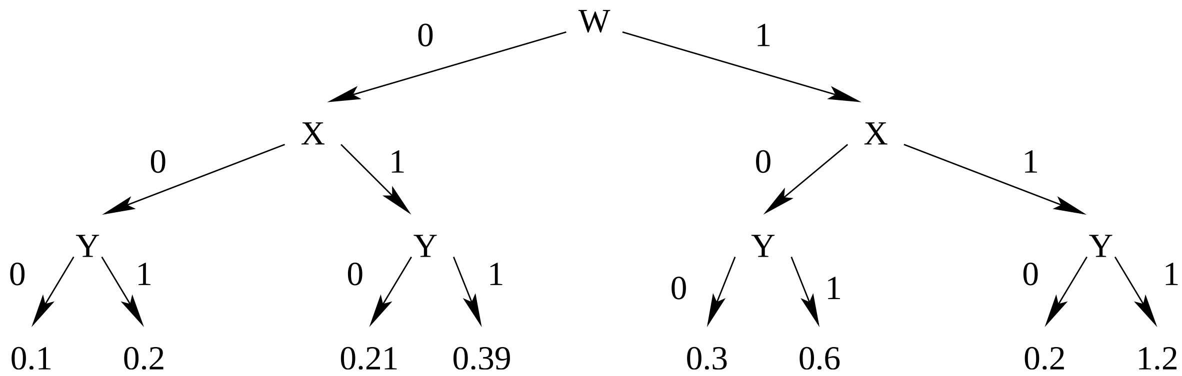
Definición: Sea ψ un potencial sobre $\mathbf{X} = \{X_1, \ldots, X_n\}$ representado por un árbol $\mathcal{T}$. Sea $\mathbf{X}_C \subseteq \mathbf{X}$ y $(\mathbf{X}_C = \mathbf{x}_C)$ una configuración de variables que llevan desde $\mathcal{T}$ hasta una variable X_i. Decimos que $\mathcal{T}$ es δ-factorizable en X_i para el contexto $(\mathbf{X}_C = \mathbf{x}_C)$, con $\delta > 0$, si para todo $x, y \in \Omega_{X_i}$, $\exists \varepsilon > 0$ t.q.

$$\left| \mathcal{T}^{R(\mathbf{X}_C = \mathbf{x}_C, X_i = x)} - \varepsilon \cdot \mathcal{T}^{R(\mathbf{X}_C = \mathbf{x}_C, X_i = y)} \right| \leq \delta \ .$$

Ejemplo

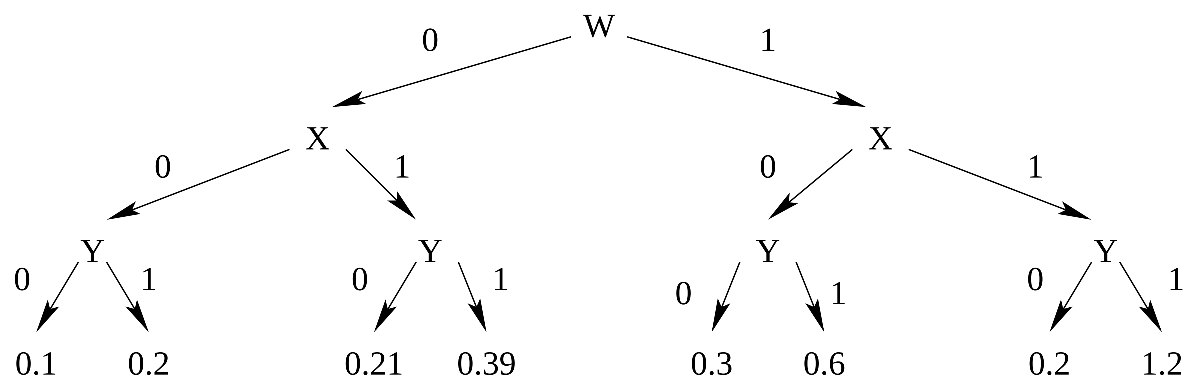


Ejemplo



$\mathcal{T}$ es 0.1-factorizable en X para el contexto $(W = 0)$

Ejemplo



$\mathcal{T}$ es 0.1-factorizable en X para el contexto $(W = 0)$