Modelos Probabilísticos Dinámicos

Verónica E. Arriola-Rios

Inteligencia Artificial

10 de agosto de 2020

Temas

Variables plantilla

- Una *variable plantilla* $X(U_1, ..., U_k)$ se puede instanciar varias veces.

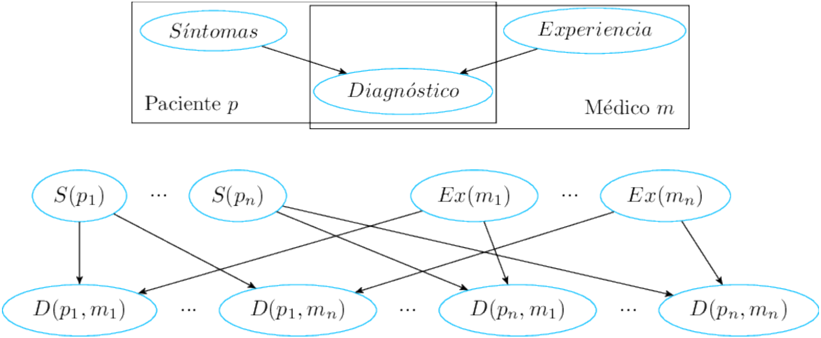
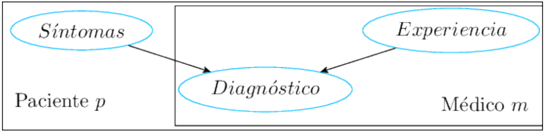


Figura: Cada paciente puede ser diagnosticado por médicos con experiencia diferente.

Otra variante



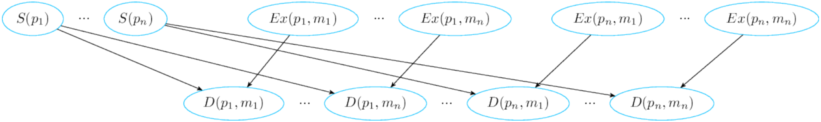


Figura: El diagnóstico de cada médico depende de su experiencia previa con cada paciente.

Características

- Las variables plantilla se utilizan para trabajar con varias instancias de objetos similares.
- Las distribuciones de probabilidad condicional utilizadas para todas las instancias de una variable plantilla son las mismas.

$$\text{Ej}: P(Diagnóstico|Síntomas, Experiencia)$$

- Lo que cambia entre instancias son los valores de las variables.

$$\text{Ej}: Síntomas(Juanito) = tos, Síntomas(Lulú) = fiebre.$$

Temas

1. Modelos de placas

2. Redes Bayesianas Dinámicas
   - Hipótesis
   - Red Bayesiana Dinámica

3. Cadenas de Márkov en tiempo discreto

Secuencia de eventos

Definición

Sea X una variable aleatoria cuyo valor cambia con el tiempo. La probabilidad de una secuencia de eventos que ocurren a tiempos discretos $t = \{0, ..., T\}$ se expresa:

$$P(X^{(0:T)}) = P(X^{(0)}) \prod_{t=0}^{T-1} P(X^{(t+1)} | X^{(0:t)})$$

utilizando la regla de la cadena.

Por ejemplo: sea X la vocal leída de una cadena al tiempo t, entonces la secuencia $X^{(0:9)}$ es $\{c, a, m, a, r, o, n, e, s\}$.

Temas

Hipótesis de Márkov

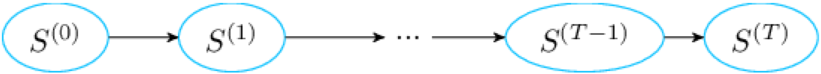
Definición

La *hipótesis de Márkov* asume que:

$$(X^{(t+1)} \perp X^{(0:t-1)} | X^{(t)}) \tag{1}$$

entonces se cumple que:

$$P(X^{(0:T)}) = P(X^{(0)}) \prod_{t=0}^{T-1} P(X^{(t+1)} | X^{(t)}) \tag{2}$$



Ejemplo

Figura: Dado el estado de movimiento (x, v, a) del auto al tiempo t se puede calcular el estado en $t + 1$.

https://www.eluniversal.com.mx/autopistas/manejar-en-curvas-un-arte-que-pocos-dominan

Modelo estacionario

Definición

El *modelo estacionario* asume que la dinámica del sistema no cambia con el tiempo.

$$P(X^{(t+1)}|X^{(t)}) = P(X'|X) \qquad \forall t \tag{3}$$

Figura: En un río el agua se mueve, pero el comportamiento del agua que pasa por cada punto se puede describir con las mismas ecuaciones durante un lapso.

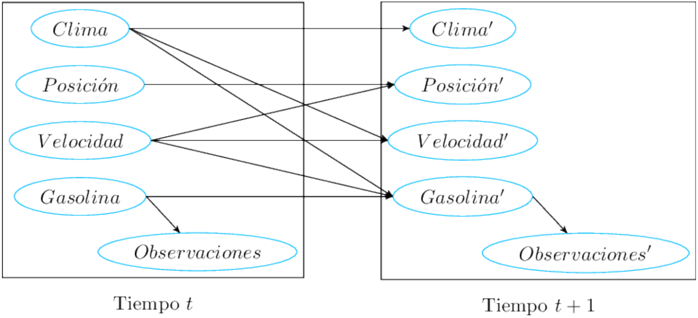
https://centrourbano.com/2018/02/09/rescatan-unico-rio-vivo-la-cdmx/

Temas

1. Modelos de placas
2. Redes Bayesianas Dinámicas
   - Hipótesis
   - Red Bayesiana Dinámica
3. Cadenas de Márkov en tiempo discreto

Placa para el modelo de la transición

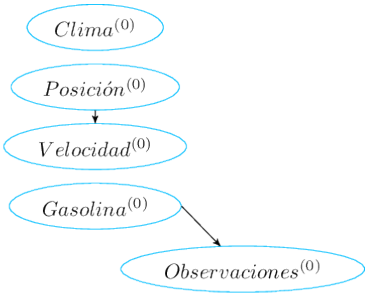
- Se utilizan variables pantillas dentro de placas para modelar la probabilidad de transición.



$$P(C', P', V', G', O'|C, P, V, G) = P(C'|C)P(P'|P, V)P(V'|C, V)P(G'|C, G)P(O'|G')$$

Distribución para el estado inicial

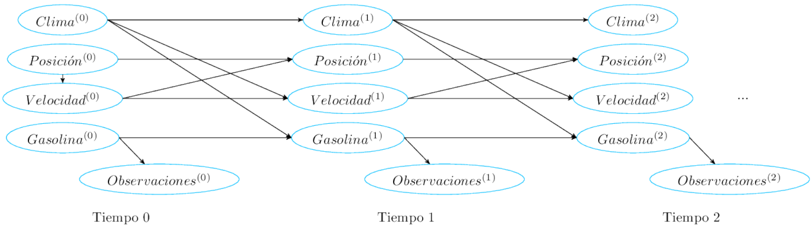
- Se debe especificar la distribución de probabilidad para los valores iniciales de las variables plantilla.
- Las dependencias internas pueden ser ligeramente diferentes a las otras placas.



$$P(C^{(0)}, P^{(0)}, V^{(0)}, G^{(0)}, O^{(0)}) = P(C^{(0)})P(V^{(0)}|P^{(0)})P(P^{(0)})P(O^{(0)}|G^{(0)})P(G^{(0)})$$

Red bayesiana aterrizada

- La red bayesiana sobre la que se está trabajando en un modelo dinámico, se obtiene al expandir el contenido de la distribución inicial y las placas.



Temas

Cadena de Márkov en tiempo discreto

- Una cadena de Márkov es un modelo bayesiano dinámico donde cada nodo sólo tiene un padre: el estado en el tiempo anterior.

$$S^{(0)} \longrightarrow S^{(1)} \longrightarrow \cdots \longrightarrow S^{(T-1)} \longrightarrow S^{(T)}$$

- El sistema tiene un conjunto finito de estados discretos $S = \{s_1, ..., s_n\}$.
- Frecuentemente, la probabilidad de transición $P(s^{(t+1)}|s^{(t)})$ de un estado a otro se puede expresar con un autómata.

Representación

Basada en el ejemplo en GEO Tutoriales 2020.

- La distribución de probabilidad inicial se especifica como un vector donde cada entrada es la probabilidad de iniciar en ese estado.

$$f^0 = \begin{bmatrix} P(S^{(0)} = 1) & = 0.45 \\ P(S^{(0)} = 2) & = 0.25 \\ P(S^{(0)} = 3) & = 0.30 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.25 \\ 0.30 \end{bmatrix}$$

- El modelo de transición se puede expresar con una matriz de conectividad donde, para cada arista en un autómata, se indica la probabilidad de transitar del estado i (renglón) al estado j (columna), $P(S^{(t+1)}|S^{(t)})$. A esta se le llama *matriz de probabilidades*.

$$P = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.03 & 0.95 & 0.02 \\ 0.2 & 0.05 & 0.75 \end{bmatrix}$$

Consultas

- Dado lo anterior, calcular la probabilidad de cada estado al tiempo $t = n$ es tan sencillo como multiplicar:

$$f^1 = P^T f^0 \tag{4}$$

$$f^2 = P^T P^T f^0 = (P^T)^2 f^0 \tag{5}$$

$$\ldots \tag{6}$$

$$f^n = (P^T)^n f^0 \tag{7}$$

Distribución límite

- Cuando todos los estados de la cadena de Márkov se comunican entre sí
- con estados recurrentes positivos y aperiódicos (su evolución no cae en ciclos)
- es posible calcular la *distribución límite* f^* de la cadena, aquella para la cual:

$$f^* = P^T f^* \tag{8}$$

- Esto se resuelve, dado P, encontrando las componentes del vector f tales que la ecuación anterior se cumple, sujeta al hecho de que f es una distribución de probabilidad y, por lo tanto, la suma de sus componentes debe ser 1.

$$f^* - P^T f^* = 0$$

$$f^* = (1 - P^T)^{-1}$$

Referencias I

GEO Tutoriales (15 de jun. de 2020). *Cadenas de Markov (Ejercicios resueltos)*. URL: https://www.gestiondeoperaciones.net/cadenas-de-markov/cadenas-de-markov-ejercicios-resueltos/.

Koller, Daphne y Nir Friedman (2009). *Probabilistic Graphical Models, Principles and Techniques*. MIT Press Cambridge.