Tema 5. Razonamiento con incertidumbre

Teorema de Bayes

Razonamiento y Representación del Conocimiento

Índice

- Incertidumbre
- Teoría de la probabilidad
- Teorema de Bayes

Incertidumbre

- Introducción
  - La incertidumbre surge cuando no podemos (o somos demasiado ‘vagos’ para) tener un conocimiento completo del mundo que nos rodea o de las consecuencias que sobre el mismo tendrán nuestras acciones
  - Nos va a permitir dar una mayor expresividad a la LPO para modelar mundos complejos sin necesidad de describir exhaustivamente todas las reglas
  - Vamos a dar un grado de creencia a las proposiciones: valor entre 0 y 1

Incertidumbre

- Grado de creencia
  - Asignar un grado de creencia 0/1 a una oración → creemos que la oración es inequívocamente falsa/verdadera
  - Valores intermedios de creencia corresponden a grados intermedios de creencia en una oración
  - La oración en sí misma es o verdadera o falsa
  - Grado de creencia != grado de pertenencia (fuzzy logic)

Incertidumbre

- Decisiones racionales
  - Agente lógico: ejecuta cualquier acción que le garantice alcanzar su objetivo:
    - Una acción se selecciona o rechaza en función de si nos acerca al objetivo o no
    - Con incertidumbre: esquema anterior no válido
  - Utilidad (calidad de ser útil): establecer preferencias entre las consecuencias de las diversas acciones
    - Teoría de la utilidad: cada estado tiene un grado de utilidad (beneficio) para el agente
    - Preferiremos estados con mayor utilidad
  - Teoría de la Decisión: teoría de la probabilidad + teoría de la utilidad

Teoría de la probabilidad

- Notación básica
  - Variables aleatorias: elementos del mundo sobre los que desconocemos su estado. P.ej: Chocar, Tiempo, ...
  - Dominio: posibles valores que puede tomar una variable
  - Clases de variables aleatorias:
    - Booleanas: dominio <cierto, falso> p.ej: Chocar
    - Discretas: dominio contable→ Tiempo:<soleado, lluvioso, nublado, nevado>
    - Continuas: dominio todos los números reales o algún subconjunto de éstos

Teoría de la probabilidad

- Probabilidad *a priori* o incondicional
  - Para una proposición a: grado de creencia que se le otorga en ausencia de cualquier otra información
  - Proviene de la observación y contabilización de sucesos
  - Escribimos P(a) → para la variable Chocar: P(Chocar=cierto) = 0.2 o P(chocar) = 0.2
  - Usaremos **P**(a) para referirnos a todos los valores posibles de una variable aleatoria:

P(Tiempo = soleado) = 0.7
P(Tiempo = lluvioso) = 0.2
P(Tiempo = nuboso) = 0.08
P(Tiempo = nevado) = 0.02

P(Tiempo) = <0.7, 0.2, 0.08, 0.02>

Teoría de la probabilidad

- Distribución de probabilidad
  - En el ejemplo anterior **P**(Tiempo) define una **distribución de probabilidad** *a priori* para la variable aleatoria Tiempo
  - La suma de los elementos en una distribución de probabilidad siempre será 1
  - **Distribución de probabilidad conjunta**: denota las probabilidades de todas las combinaciones de valores de las variables de un conjunto
  - En el caso continuo hablaremos de funciones de densidad de probabilidad

Teoría de la probabilidad

- Probabilidad condicional
  - Las probabilidades *a posteriori* o condicionales se usan cuando el agente obtiene alguna evidencia referente a las variables aleatorias desconocidas que constituyen el Universo de Discurso
  - Notación P(a|b) y se lee: probabilidad de a dado que conozco b
  - Regla del producto: $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$

$$P(a/b) = \frac{P(a \wedge b)}{P(b)}$$

Teoría de la probabilidad

- Axiomas de la probabilidad

1) Todas las probabilidades están entre 0 y 1. Para cualquier proposición a, $0 \leq P(a) \leq 1$
2) Las proposiciones necesariamente ciertas tienen probabilidad 1 y las necesariamente falsas 0: $P(\text{cierto}) = 1$, $P(\text{falso}) = 0$
3) La probabilidad de una disyunción viene dada por

$$P(a \lor b) = P(a) + P(b) - P(a \land b)$$

Teoría de la probabilidad

- Axiomas de la probabilidad
  - P(a V ¬a) = P(a) + P(¬a) – P(a Λ ¬a) (por ax. 3 con b=¬a)
  - P(cierto) = P(a) + P(¬a) – P(falso) (por equivalencia lógica)
  - 1 = P(a) + P(¬a) (por ax. 2)
  - P(¬a) = 1 – P(a) (por álgebra)
  - En el caso discreto general tendríamos:

$$\sum_{i=1}^{n} P\left(D=d_i\right)=1$$

Teoría de la probabilidad

- Axiomas de la probabilidad
  - Un agente que viole los axiomas de probabilidad se comportará de forma irracional
  - “Si el Agente 1 expresa un conjunto de grados de creencia que viola los axiomas de la probabilidad entonces hay una combinación de apuestas para un Agente 2 que garantiza que el Agente 1 perderá dinero todas las veces” (Finetti, 1931)

Teoría de la probabilidad

- Inferencia probabilística
  - Distribución conjunta completa como base de conocimiento

	detecta-obstáculo		¬detecta-obstáculo	
	en-mapa	¬en-mapa	en-mapa	¬en-mapa
chocar	0.108	0.012	0.072	0.008
¬chocar	0.016	0.064	0.144	0.576

  - P(chocar **V** detecta-obstáculo) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28
  - Marginalización: $$P(Y) = \sum_{z} P(Y, z)$$
    - P(chocar) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2

Teoría de la probabilidad

- Inferencia probabilística
  - Distribución conjunta completa como base de conocimiento

	detecta-obstáculo		¬detecta-obstáculo	
	en-mapa	¬en-mapa	en-mapa	¬en-mapa
chocar	0.108	0.012	0.072	0.008
¬chocar	0.016	0.064	0.144	0.576

  - Condicionamiento: $\boxed{\mathbf{P}(\mathbf{Y}) = \sum_{z} \mathbf{P}(\mathbf{Y}/\mathbf{z}) P(z)}$

$$P(chocar / detecta - obstáculo) = \frac{P(chocar \wedge detecta - obstáculo)}{P(detecta - obstáculo)}$$

$$\frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

Teoría de la probabilidad

- Independencia
  - Dadas dos proposiciones a y b, son independientes si
    - $P(a|b) = P(a)$
    - $P(b|a) = P(b)$
    - $P(a \wedge b) = P(a)P(b)$
  - Afirmaciones de independencia nos ayudarán a reducir el tamaño de representación del problema

Teorema de Bayes

- Se deriva de la regla del producto
  - P(ab) = P(a|b)*P(b)
  - P(ab) = P(b|a)*P(a)
- Igualando las partes derechas y dividiendo por P(a):

$$P(b/a) = \frac{P(a/b)\,P(b)}{P(a)}$$

- En el caso general

$$\boldsymbol{P}(Y/X) = \frac{\boldsymbol{P}(X/Y)\,\boldsymbol{P}(Y)}{\boldsymbol{P}(X)}$$

- Donde

$$\boldsymbol{P}(X) = \sum_i \boldsymbol{P}(X/Y_i)\,\boldsymbol{P}(Y_i)$$

Teorema de Bayes

- Ejercicio propuesta: queremos determinar la probabilidad de que un robot móvil choque durante su movimiento teniendo la información de un sensor para detectar obstáculos. Sabemos que el sensor detectó el obstáculo el 85% de las veces que el robot chocó, que la probabilidad de que se produzca un choque es de un 20% y que la probabilidad de que el sensor detecte un obstáculo es un 25%.
- ¿Qué probabilidad hay de que el robot choque si el sensor nos da una lectura positiva (detecta obstáculo)?
- ¿Y si si nos da una lectura negativa?

Bibliografía recomendada

- Inteligencia Artificial. Un enfoque Moderno. Stuart Russell, Peter Noving. Ed. Prentice Hall. 2004
- http://www.ecomportamiento.org/blog/2017/12/5/mejorando-nuestras-creencias-el-teorema-de-bayes