Curso de Inteligencia Artificial

Unidad V: Incertidumbre

Jose Aguilar
CEMISID, Facultad de Ingeniería
Universidad de los Andes
Mérida, Venezuela
aguilar@ula.ve

¿Cuándo hay incertidumbre?

Se dice que hay incertidumbre cuando no se posee información suficiente o es ambigua como para asignarle un grado de veracidad a una sentencia.

Ejemplo de incertidumbre

Un muchacho desea **vender periódicos** en la cafetería de la universidad y tiene que decidir **cuántos debe comprar**.

La **experiencia indica** que la demanda diaria varia entre 15 a 30 periódicos. Además, debe pagar 2 Bs por cada diario para venderlos en 3 Bs. cada uno.

Los **periódicos que no son vendidos** durante el día **se pierden**.

Determinar el número de diarios que el muchacho debe encargar.

Conceptos vinculados a agentes

- Algunas sentencias pueden ser:
  - Verificadas por percepción
  - Inferidas desde el estado actual y/o previo conocimiento

- Pero, en ciertos casos el agente debe actuar con incertidumbre por:
  - Desconocimiento completo del ambiente
  - No comprensión de la situación
  - Complejidad del problema
  - Poca fiabilidad de la información
  - Imprecisión inherente al modelo mental

Agente actuando bajo incertidumbre

Un problema con lógica de primer orden, es que los **agentes casi nunca tienen acceso a la verdad completa acerca de su entorno**.

En casi todos los casos habrá cuestiones importantes para las cuales el agente **no puede encontrar una respuesta categórica**.

Por lo tanto, el agente debe **actuar bajo incertidumbre**.

Incertidumbre

Debo ir al aeropuerto que esta a 30 Km. ya que debo tomar un avión

- Tengo un Plan que me hará llegar 90 min. antes
- Tengo otro Plan que me hará llegar 240 min. antes y otro 120 min. antes,

¿Cuál escojo?

- Lo correcto es tomar la decisión racional,

En este caso es considerando la importancia relativa de los diferentes objetivos (esperar vs. tomar el avión), y el grado con el cual ellos se pueden lograr

Incertidumbre

- Acción A_t = llegar al aeropuerto t minutos antes del vuelo

Cual es el A_t idóneo?

- Problemas:
  1. Ambiente parcialmente observable (estado de las calles, planes de los otros conductores, etc.)
  2. Ruido en los sensores (reporte del tráfico)
  3. Incertitud sobre el futuro (caucho desinflado, accidente, etc.)
  4. Complejidad en el modelado y predicción del trafico

"A_{25} me permitirá llegar a tiempo si no hay accidentes en la ruta, no se me desinfla el caucho, ..."

"A_{1440} me permitirá llegar a tiempo, pero me obligara a dormir en el aeropuerto…"

Ignorancia

- ***Decisiones bajo ignorancia parcial***: cuando podemos atribuir probabilidades a algunos actos pero a otros no.

- ***Decisiones bajo incertidumbre estructurada***: sabemos los estados del sistema y sus consecuencias, pero no conocemos el estado del sistema en cada momento.

- ***Decisiones bajo incertidumbre no estructurada o bajo ignorancia total***: ni siquiera conozco los estados posibles de un sistema.

Ignorancia

- En las decisiones tomadas en **ignorancia total**, **el comportamiento del decisor se basa puramente en su actitud hacia la incógnita**.

Algunos de estos comportamientos son de forma optimista, pesimista o de arrepentimiento.

- Siempre que un decisor tiene **cierto conocimiento sobre los estados de la naturaleza**, **puede usar una técnica** (probabilidades, lógica difusa) para analizar la ocurrencia de cada estado.

toma de decisiones bajo riesgo.

- El decisor en **caso de incertidumbre** debería **minimizar sus incertidumbres**

Encuestas, estadísticas, análisis de sensibilidad, etc.

Incertidumbre

Ejemplo Conocimiento Incierto

$\forall$p Sintoma(p, dolormuela)=> Enfermedad (p,Caries) no es verdad

$\forall$p Enfermedad (p,Caries) => Sintoma(p, dolormuela) Tampoco!!

Solo haciendolo exhautivo!!!

$\forall$p Sintoma(p, dolormuela)=> Enfermedad (p,Caries) O imposible
Enfermedad (p,Oido) O ...

Incertidumbre

- Lógica primer orden falla por:
  - ***Pereza***. Es mucho trabajo listar el conjunto completo de antecedentes o consecuencias necesarios para asegurar que una regla no tenga excepciones, y se vuelve muy pesado usar las enormes reglas que resultan.
  - ***Ignorancia teórica***. Faltan de avances teóricos en el área (para el ejemplo, la ciencia médica no tiene una teoría completa para el tema).
  - ***Ignorancia práctica***. Podemos tener incertidumbre acerca de un paciente en particular porque todos los tests necesarios no han sido o no pueden ser realizados. (Aunque conozcamos todas las reglas).
- Solución
  - **Teoría Estadística**

    Grados de Creencia => teoría de probabilidades
  - **Lógica Difusa**

    Grados de Verdad => teoría de los conjuntos difusos

Forma fácil con reglas de producción

Incorporar a las reglas el *grado de confianza* o *factor de certeza*

Si umbral_pasado>80% entonces estado_estudiante=raspado con FC=90

Estamos seguro en un 90% que estudiante raspara si ese es el umbral para raspar

Incertidumbre

Probabilidad provee una forma de resumir la incertidumbre que viene de la ignorancia y pereza

- Probabilidad 0.8 significa 80% grado de creencia y no 80% verdad (ver lógica difusa=> conocido como grado de verdad)

- 0.8 significa que el 80% paciente tienen caries al dolerle la muela y el 20% es el resto de casos (pereza o ignorancia)

- En sentencias: 0 => creencia total que será falsa
  1 => creencia total que será verdad

Incertidumbre

La teoría de Probabilidad se puede aplicar a

- **Sucesos o eventos**: probabilidad de ocurrencia (clásico de la teoría de probabilidades)

- **Proposiciones**: probabilidad de que sean o no ciertas (indica grado de credibilidad de las proposiciones basado en las evidencias con las que se cuentan)

=> de interés en I.A, a partir de ellas se pueden hacer inferencias!!

  – Ejemplo: Juan se referirá a Caracas en su discurso según la información con la que contamos

Teoría de Probabilidad e Incertidumbre

- Elemento básico: variable aleatoria

- Similar a lógica proposicional: posibles mundos definidos asignando valores a variables aleatorias.
  - **Variables Aleatorias lógicas**
    - *Caries* (Tengo una carié?)
  - **Variables Aleatorias Discretas**
    - *Clima* es *<soleado, lluvioso, nublado>*

    Dominio de valores debe ser exhausto y mutuamente excluyente

- Proposición se construye asignando valores a las variables:
  - *Clima = soleado*, *Caries = falsa*

- Proposiciones complejas se forman desde proposiciones elementales conectadas por operadores lógicos
  - *Clima = soleado* $\vee$ *Caries = falsa*

Sintaxis

- Evento Atómico: Una completa especificación del estado del mundo sobre el que el agente esta incierto

  Si el mundo se describe por dos variables lógicas *Caries* y *dolormuela*, hay 4 eventos atómicos:

  $Caries = false \wedge dolormuela = false$

  $Caries = false \wedge dolormuela = true$

  $Caries = true \wedge dolormuela = false$

  $Caries = true \wedge dolormuela = true$

- Eventos atómicos son también mutuamente excluyentes y exhaustivos

Incertidumbre

- Evidencias (percepción): usada para evaluar las probabilidades en las sentencias
  - Antes de la evidencia => a priori o **probabilidad incondicional**
  - Después de la evidencia => a posteriori o **probabilidad condicional**

- Teoría de decisión:
  - Teoría de Utilidad es usada para representar e inferir preferencias
  - Teoría de Decisión = teoría de probabilidad + teoría de utilidad

Un agente es racional si escoge la acción que le de mas utilidad, según la teoría de la decisión

Manejando conocimiento incierto

La probabilidad que un agente asigna a una proposición depende de las percepciones que ha recibido hasta el momento

la *evidencia*

Así como el estado de una implicación puede cambiar cuando se agregan más sentencias a la BC, las probabilidades pueden cambiar cuando se adquiere más evidencia.

- Antes de obtener la evidencia, hablamos de

  probabilidad incondicional o previa;

- luego de obtener la evidencia, hablamos de

  probabilidad condicional o posterior.

Incertidumbre

Agente de decisión

1. Calcular las **probabilidades de ocurrencia de cada acción**, dada las descripciones de las acciones y probabilidades desde el estado actual

2. **Seleccionar acción con mayor utilidad**, dada las probabilidades de su ocurrencia

Notación básica de probabilidad

Probabilidad incondicional. Ej. P(Caries) = 0.1 Si Caries denota la proposición que un paciente particular tenga caries, significa **que *en ausencia de cualquier otra información***, el agente asignará una probabilidad de 0.1 al evento de un paciente con caries.

Las proposiciones están formadas por las llamadas **variables aleatorias o al azar**:

P(Clima = Soleado) = 0.7

P(Clima = Lluvia) = 0.2

P(Clima = Nublado) = 0.08

P(Clima = Nieve) = 0.02

P(Clima) representa la probabilidad de todos los valores posibles de Clima, en este caso:

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

Incertidumbre

- A priori o probabilidad incondicional: no se tiene evidencias

P(caries)=0.1

P(clima=soleado)=0.72 P(clima=lluvioso)=0.08

P(clima=nublado)=0.1 P(clima=con nieve)=0.1

  - Distribución de Probabilidad: da valores a todas las asignaciones posibles

    P(clima)=(0.72,0.1,0.08,0.1)

  - Distribución de Probabilidad Conjunta: para un conjunto de variables aleatorias da la probabilidad de cada evento atómica de ellas

    $P(A \text{ y } B) = P(A \cap B) = P(A,B) = P(A)P(B)$ ya que A y B son independientes

P(*Clima*,*Caries*) = una matriz de 4 × 2

Regla del Producto

Caries vs Clima	soleado	lluvioso	nublado	con nieve
V	0.072	0.008	0.01	0.01
F	0.648	0.072	0.09	0.09

Notación básica de probabilidad

Probabilidad condicional. Cuando el agente ha obtenido alguna **evidencia concerniente a las proposiciones previamente desconocidas**.

Probabilidad condicional, con notación $(A \mid B)$ se lee '*probabilidad de* A *dado que todo lo que conocemos es* B'

$P(A \mid B)$ solo puede ser usado cuando todo lo que se conoce es B. Tan pronto se conozca C se debe calcular $P(A \mid B \wedge C)$.

$\mathbf{P}(X \mid Y)$ es una tabla bidimensional que da los valores de $P(X=x_i \mid Y=y_j)$ para cada par posible i, j.

Incertidumbre

A posteriori o probabilidad condicional

P(caries|dolormuela)=0.8

"probabilidad de caries a sabiendas que tiene dolor de muela"

P(A|B)=P(AYB)/P(B)

– **Nueva evidencia puede ser irrelevante:**

P(caries|dolormuela, *soleado*) = P(caries|dolormuela)= 0.8

– **Pero si conocemos más, mejor:**

P(*caries* | *dolormuela*, *infección bucal*) = 0.9

Incertidumbre

- Probabilidad de llegar a tiempo si conozco algo:

  p.e., $P(A_{25}$ | hay accidentes reportados$) = 0.06$

- Pero como dijimos antes, probabilidad cambia con mas evidencias:

  p.e., $P(A_{25}$ | hay accidentes reportados, 5 a.m.$) = 0.25$

$P(A_{25}$ estaré a tiempo $| \ldots) = 0.04$
$P(A_{90}$ estaré a tiempo $| \ldots) = 0.70$
$P(A_{120}$ estaré a tiempo $| \ldots) = 0.95$
$P(A_{1440}$ estaré a tiempo $| \ldots) = 0.9999$

Escogencia depende entre **mi prioridad** por perder vuelo vs. cuanto esperare

Axiomas de probabilidad

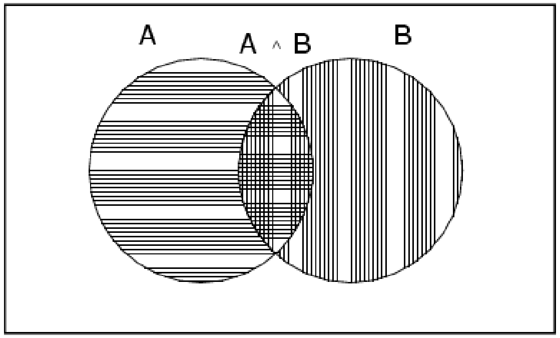
- **A posteriori o probabilidad condicional**

P(A Y B)= $P(A \cap B)$ =P(A|B)P(B)= P(B|A)P(A) Regla del Producto
para A y B dependientes

P(A O B)=P(A) + P(B) - P(A Y B)

$P(A)+P(\neg A)=1$

$0 \leq P(A) \leq 1$



- **Distribución de Probabilidades**

	dolormuela	¬dolormuela
Caries	0.04	0.06
Caries	0.01	0.89

P(Caries o dolormuela)=0.04+0.01+0.06=0.11

diferente a: P(caries|dolormuela)=P(caries Y dolormuela)/P(dolormuela)=0.04/(0.04+0.01)=0.8

Incertidumbre: Ejemplo Regla del producto

- Existen dos candidatos A y B a gerentes en una empresa.
- Las probabilidades de ganar de cada uno son 0,7 y 0,3, respectivamente.
- Si gana A, la probabilidad de introducir nuevos productos es 0,8; si gana B, la probabilidad es 0,4.

¿Cuál es la probabilidad, antes de las elecciones, de que sea introducido un nuevo producto?

- **Solución:** Si llamamos N al nuevo producto, será:

$$P(N\ Y\ A) + P(N\ Y\ B) = P(N \cap A) + P(N \cap B) =$$
$$= P(A) * P(N|A) + P(B) * P(N|B) =$$
$$= 0,7 * 0,8 + 0,3 * 0,4 = 0,56 + 0,12 = 0,68.$$

Incertidumbre

- **Regla de la Cadena** es derivada por la aplicación sucesiva de la regla del producto :

$$
\begin{aligned}
\mathbf{P}(X_1, \ldots, X_n) &= \mathbf{P}(X_1,\ldots,X_{n-1})\ \mathbf{P}(X_n \mid X_1,\ldots,X_{n-1}) \\
&= \mathbf{P}(X_1,\ldots,X_{n-2})\ \mathbf{P}(X_{n-1} \mid X_1,\ldots,X_{n-2})\ \mathbf{P}(X_n \mid X_1,\ldots,X_{n-1}) \\
&= \ldots \\
&= \pi_{i=1}\text{^}n\ \mathbf{P}(X_i \mid X_1, \ldots, X_{i-1})
\end{aligned}
$$

- *Recordemos que A* y *B* son independientes ssi

$$\mathbf{P}(A|B) = \mathbf{P}(A) \quad \text{o} \quad \mathbf{P}(B|A) = \mathbf{P}(B) \quad \text{o} \quad \mathbf{P}(\text{A Y B}) = \mathbf{P}(A)\ \mathbf{P}(B)$$

Por ejemplo:

P(Dolormuela, *Caries*, *Clima*) = **P**(Dolormuela, *Caries*) **P**(*Clima*)

Independencia Condicional

- Si yo tengo caries, la probabilidad que este hinchado no depende si tengo o no dolor de muela:

$$\mathbf{P}(\mathit{hinchado} \mid \mathit{dolormuela}, \mathit{caries}) = \mathbf{P}(\text{hinchado} \mid \mathit{caries})$$

Hinchado es condicionalmente independiente de *dolormuela* dado *Caries*

- Por otro lado, son *Equivalentes*:

$$\mathbf{P}(\text{Dolormuela}, \mathit{Hinchado} \mid \mathit{Caries}) = \mathbf{P}(\text{Dolormuela} \mid \mathit{Caries})\ \mathbf{P}(\text{Hinchado} \mid \mathit{Caries})$$

Independencia Condicional

- Usando regla de la cadena:

$\mathbf{P}(Caries, dolormuela, hinchado)$

$= \mathbf{P}(Caries, dolormuela)\ \mathbf{P}(\text{hinchado} \mid Caries, dolormuela)$

$= \mathbf{P}(Caries)\ \mathbf{P}(dolormuela \mid Caries)\ \mathbf{P}(hinchado \mid Caries, dolormuela)$

Si *hinchado* condicionalmente independiente de *dolormuela* :

$= \mathbf{P}(Caries)\ \mathbf{P}(dolormuela \mid Caries)\ \mathbf{P}(hinchado \mid Caries)$

Durante la resolución de la regla de la cadena puedo hacer simplificaciones por relaciones de equivalencia, de dependiencia, etc.

Incertidumbre

Regla de Bayes:

- **Vincula la probabilidad de A dado B con la probabilidad de B dado A**

$$P(B|A)=P(A|B)P(B)/P(A)$$

- Excelente para **diagnóstico** desde una posible **causa**:

$$P(Causa|Efecto) = P(Efecto|Causa)\ P(Causa) / P(Efecto)$$

Ejemplo: sabiendo la probabilidad de tener nariz húmeda dado que se tiene meningitis, se podría saber la probabilidad de tener meningitis si se tiene nariz húmeda (si se tiene algún dato más)

N:narizhúmeda M:Meningitis

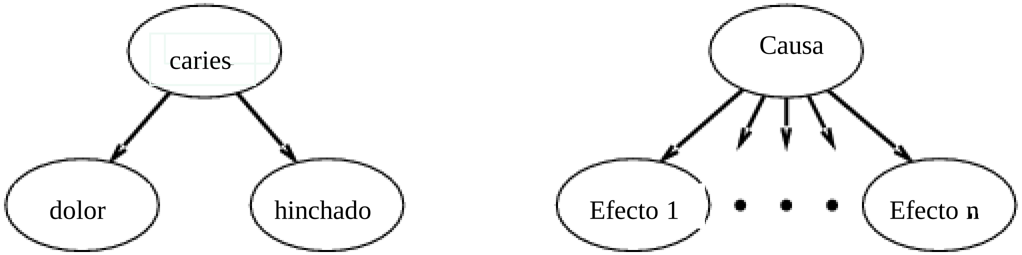
$P(N|M)=0.5$ $P(M)=1/5000$ $P(N)=1/20$

$$P(M|N)=P(N|M)P(M)/P(N)=0.002$$

Incertidumbre

$$\mathbf{P}(\text{Causa},\text{Efecto}_1, \ldots ,\text{Efecto}_n) = \mathbf{P}(\text{Causa}) \prod_i \mathbf{P}(\text{Efecto}_i|\text{Causa})$$

Si condicionalmente independencia entre los efectos!!!



Regla de la suma:

$$P(X+Y|Z)=P(X|Z)+P(Y|Z)+P(X,Y|Z)$$

Para proposiciones mutuamente exclusivas: $P(X1+X2+\ldots+Xk)= \sum i\ P(Xi|Z)$

Incertidumbre

- Comparación

$$P(B|A)/P(C|A)=P(A|B)P(B)/P(A|C)P(C)$$

probabilidad que ocurra C más que B dado que ocurre A

Ejemplo:

G: Gripe

$$P(M|N)/P(G|N)=1/80$$

significa que gripe es 80 veces mas probable que meningitis al tener nariz húmeda

Resume

- Regla del producto:

$$\mathbf{P}(X, Y) = \mathbf{P}(X|Y)\,\mathbf{P}(Y) = \mathbf{P}(Y|X)\,\mathbf{P}(X)$$

- **Regla de Bayes**:

$$\mathbf{P}(Y|X) = \mathbf{P}(X|Y)\,\mathbf{P}(Y) \,/\, \mathbf{P}(X)$$

Esta regla y la propiedad de **independencia** son el fundamento del razonamiento probabilístico y permiten relacionar las probabilidades de unas evidencias con otras.

- Suponiendo que podemos estimar las probabilidades que involucran todos los valores de la variable Y podemos reescribir la formula de Bayes como:

$$\mathbf{P}(Y|X) = \mathbf{P}(X|Y)\,\mathbf{P}(Y) \,/\, \mathbf{P}(X) = \alpha\mathbf{P}(X|Y)\,\mathbf{P}(Y)$$

- Y suponiendo independencia condicional entre dos variables podremos escribir:

$$\mathbf{P}(X,Y|Z) = \mathbf{P}(X|Z)\,\mathbf{P}(Y|Z)$$

$$\mathbf{P}(Z|X,Y) = \alpha\mathbf{P}(X|Z)\,\mathbf{P}(Y|Z)\,\mathbf{P}(Z)$$

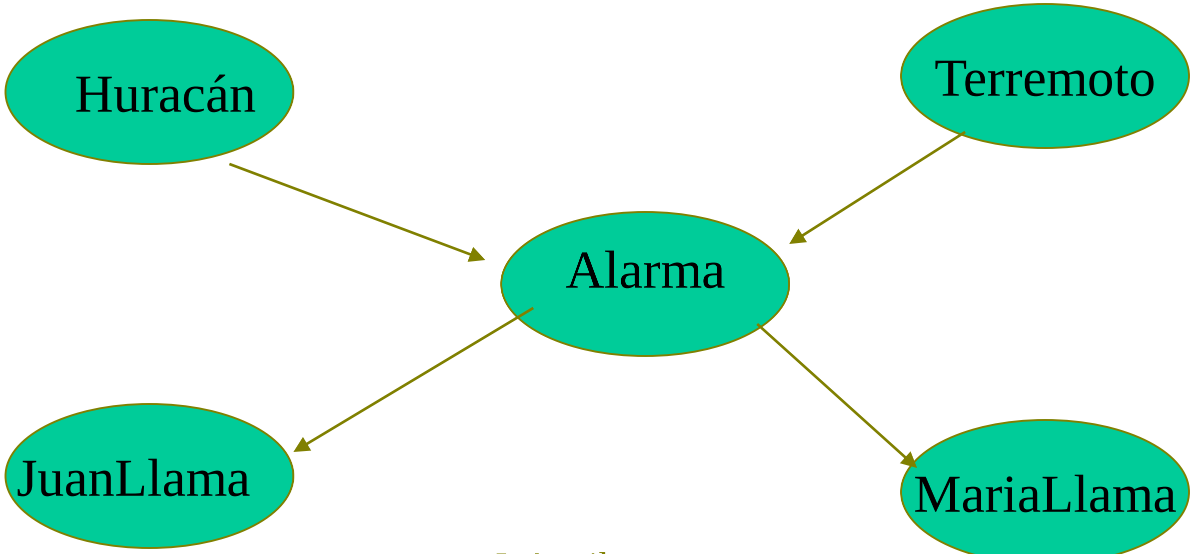
Resumen razonamiento estocástico

- Probabilidad es un formalismo para manejo del conocimiento incierto
- Distribución de probabilidad conjunta especifica la probabilidad de eventos atómicos juntos
- Consultas pueden ser respondidas verificando probabilidad ocurrencia de eventos
- Conocimiento parcial se puede incluir usando cálculos probabilísticos:

Independencia , dependencia condicional, etc. proveen las herramientas para ello

Red de Creencia

- Grafo Aciclico
- Nodos tienen una probabilidad condicional según sus padres (nodos que lo apuntan)
- Arcos representan la influencia de un nodo (padre) sobre el otro



Red de Creencias

- Probabilidad de Conjunción (P(x1, ...xn)):

P(X1=x1 Y ... Y Xn=xn)= Π P(xi|Padres(xi))

- Construir Red de Creencias

P(x1, ...xn)= Π P(xi|xi-1, ... , x1) (regla de la cadena)

Π P(Xi|Padres(Xi)) ya que padres son c.i. entre ellos!!

1. Escoger orden de variables $X_1, \ldots, X_n$
2. Para $i = 1$ a n

   Seleccionar padre desde $X_1, \ldots, X_{i-1}$ tal que

   $\boldsymbol{P}(X_i \mid Padre(X_i)) = \boldsymbol{P}(X_i \mid X_1, \ldots X_{i-1})$

Red de Creencias

Huracán	Terremoto	P(Alarma| Huracán,Terremoto)	
		V	F
V	V	.95	.05
V	F	.95	.05
F	V	.29	.71
F	F	.001	.999

P(Huracán)=0.001 P(Terremoto)=0.02

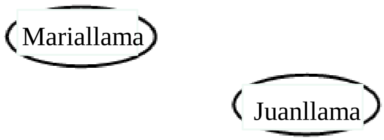
A	P(JuanLlama|A)
V	.9
F	.1

A	p(MariaLlama|A)
V	.7
F	.3

Ejemplo de mal orden

- Suponga orden *M, J, A, H, T*

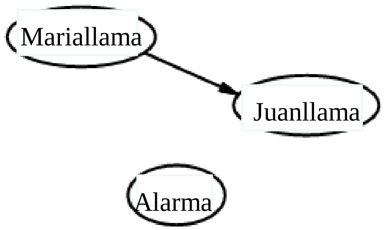
Evaluemos influencia casual



$\boldsymbol{P}(J \mid M) = \boldsymbol{P}(J)?$

Ejemplo de mal orden

- Orden: *M, J, A, H, T*

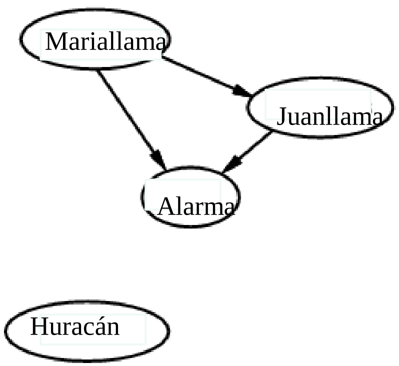


$\boldsymbol{P}(J \mid M) = \boldsymbol{P}(J)$? **No**

$\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A \mid M)$? $\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A)$?

Ejemplo de buen orden

- Orden: *M*, *J*, *A*, *H*, *T*



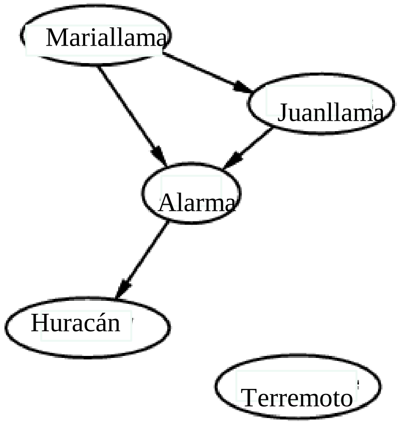
$\boldsymbol{P}(J \mid M) = \boldsymbol{P}(J)$? **No**

$\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A \mid M)$? $\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A)$? **No**

$\boldsymbol{P}(H \mid A, J, M) = \boldsymbol{P}(H \mid A)$?

Ejemplo de mal orden

- Orden: *M, J, A, H, T*



$\boldsymbol{P}(J \mid M) = \boldsymbol{P}(J)$? **No**

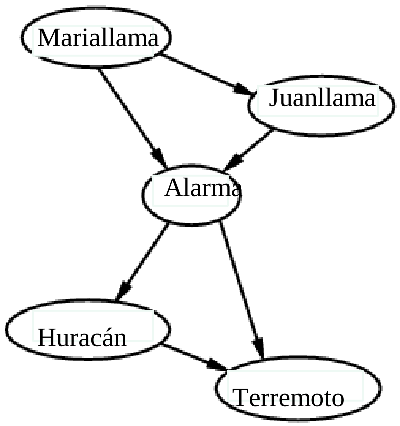
$\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A \mid M)$? $\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A)$? **No**

$\boldsymbol{P}(H \mid A, J, M) = \boldsymbol{P}(H \mid A)$? **Si**

$\boldsymbol{P}(T \mid H, A, J, M) = \boldsymbol{P}(T \mid A)$?

Ejemplo de buen orden

- M, J, A, H, T



$\boldsymbol{P}(J \mid M) = \boldsymbol{P}(J)$? **No**

$\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A \mid M)$? $\boldsymbol{P}(A \mid J, M) = \boldsymbol{P}(A)$? **No**

$\boldsymbol{P}(H \mid A, J, M) = \boldsymbol{P}(H \mid A)$? **Si**

$\boldsymbol{P}(T \mid H, A ,J, M) = \boldsymbol{P}(T \mid A)$? **No**

$\boldsymbol{P}(T \mid H, A, J, M) = \boldsymbol{P}(T \mid A, H)$? **Si**

Red de Creencias

- Inferencia de Diagnostico

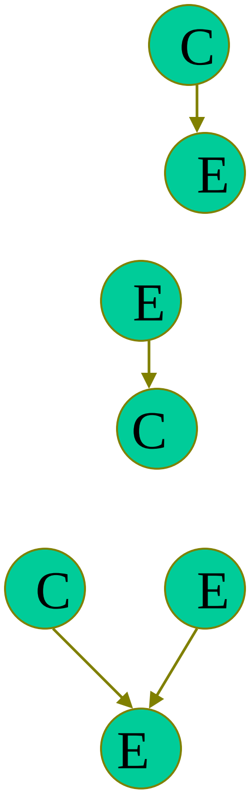
P(H|JuanLlama)

- Inferencia causal

P(JuanLlama|H)

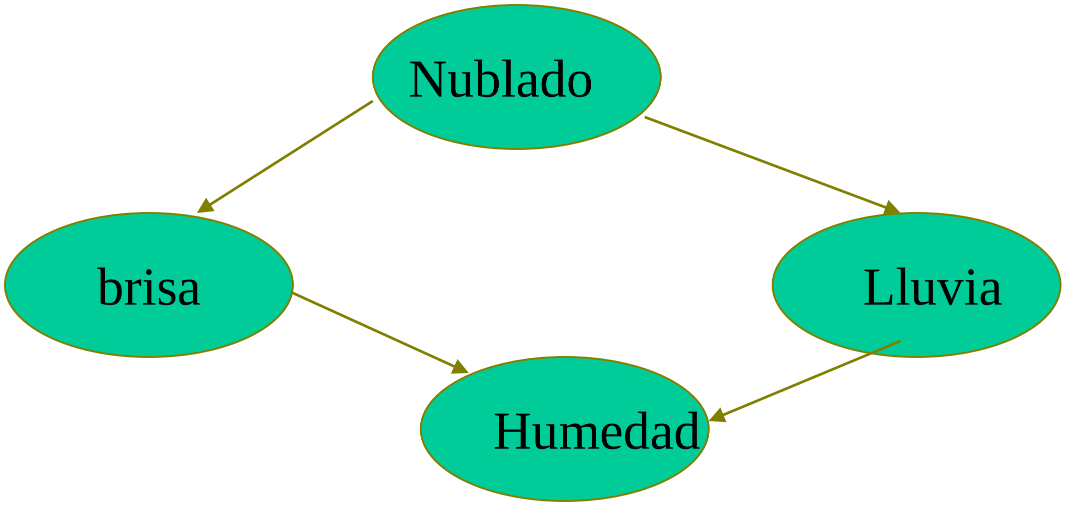
- Inferencia intercausal

P(H|Alarma Y T)



Red de Creencia

Método de Agrupamiento condicional



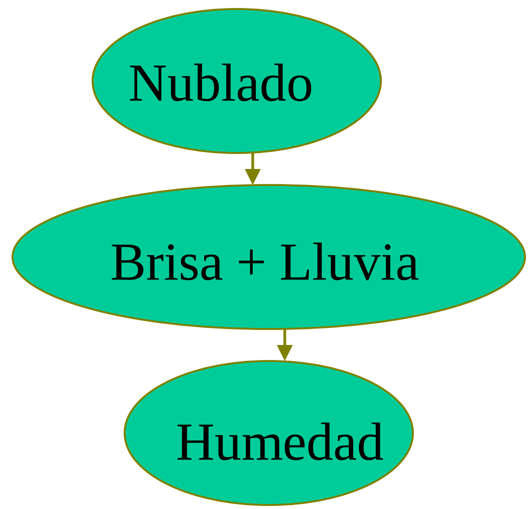
P(N)=.5

N	P(Bri)	P(Llu)
V	.5	.8
F	.5	.2

Bri	Llu	P(H)
V	V	.99
V	F	.9
F	V	.9
F	F	0

Red de Creencia

Método de Agrupamiento condicional



P(N)=.5

N	P(Bri + Llu)			
	VV	VF	FV	FF
V	.08	.02	.72	.18
F	.4	.1	.4	.1

Bri	+	Llu	P(H)
V		V	.99
V		F	.9
F		V	.9
F		F	0

Incertidumbre y decisiones racionales

Para tomar decisión, un agente debe tener preferencias entre los diferentes resultados posibles.

La teoría de utilidad representa y razona con preferencias.

La teoría general de decisiones racionales combina preferencias con probabilidades:

Teoría de Decisión=Teoría de Probabilidades+Teoría de Utilidad

Utilidad esperada máxima*:* la idea fundamental de la teoría de decisión es que un agente es racional si y solo si elige la acción que produce la utilidad más grande esperada,

Utilidad

Utilidad Esperada por realizar una acción A:

$$UE(A|E) = \Sigma i\ P(RESULT_i(A)|E, Do(A))\ U(RESULT_i(A))$$

⇒ un agente tiene todos los posibles resultados de ejecutar la acción no deterministica A en **$RESULT_i(A)$**, donde **U** da la utilidad respectiva

⇒ **Do(A)** es la proposición de que A es ejecutada en el actual momento

⇒ **E** son las evidencias actuales que tiene el agente disponible

$P(RESULT_i(A)|E, Do(A))$ => modelo probabilístico

$RESULT_i(A)$ => búsqueda o planificación

Agente escogerá la acción *A* que maximice la utilidad esperada (MUE)

Utilidad

$$MUE= \max A\ (\Sigma_i\ P(Result_i(A)|E, Do(A))\ U(Result_i(A)))$$

Trabajar con evidencia que descubrimos que es verdad (VI).

Si tenemos evidencia que Ej ocurre, entonces mejor acción actual α es:

$$\mathbf{MUE(\alpha \mid Ej)= \max \alpha\ (\Sigma_i\ P(Result_i(\alpha)|\ Do(\alpha), Ej)\ U(Result_i(\alpha))}$$

Como E es una variable aleatoria cuyo valor es desconocido, se debe estimar el valor de descubrir su valor Ej desde todos los valores posibles e_{jk} que conocemos de E usando nuestras creencias sobre esos valores

$$\mathbf{VI(Ej)= \Sigma_k\ P(Ej=e_{jk}|E)\ UE(\alpha \mid Ej=e_{jk})}$$

Utilidad

Agente que recoge información

Percibir D

Con D calcular VI(Ej) y Costo(Ej)

SI VI(Ej) > Costo(Ej)

  Hacer acción según MUE(α | Ej)

de lo contrario

  buscar otra acción desde otros valores de E

Donde: Costo(Ej) es el costo por obtener la evidencia Ej a través de test, cuestionarios, etc.

Redes Bayesianas

- Representación de las relaciones de independencia entre variable aleatorias.
- **Grafo dirigido acíclico** que tiene información probabilística en sus nodos indicando cual es la influencia de los padres sobre un nodo del grafo:

$$(P(X_i \,|padres(X_i\,)))$$

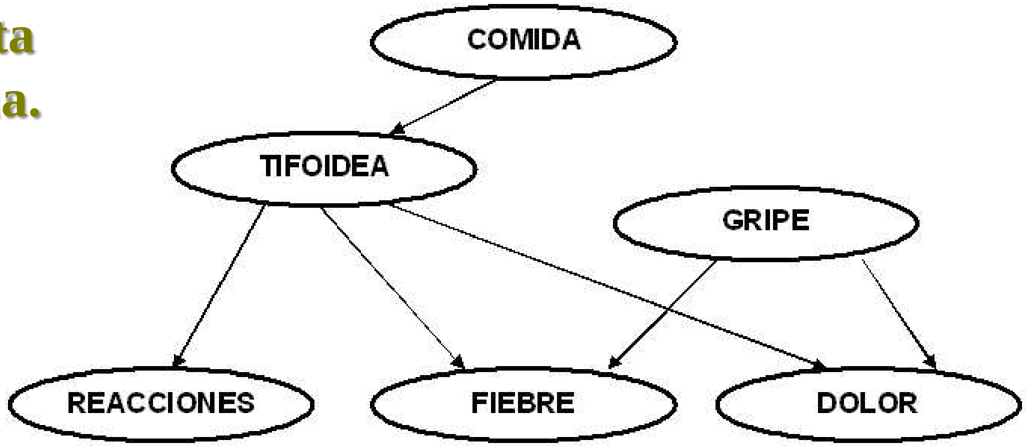
- El conjunto de probabilidades representadas en la red describe la distribución de probabilidad conjunta de todas las variables.
- La distribución conjunta completa se define como el producto de las distribuciones condicionales locales:

$$P\ (X1, \ldots ,Xn) = \pi i = 1\ P\ (Xi \mid Parents(Xi))$$

REDES BAYESIANAS

Representación grafica de dependencias para razonamiento probabilístico, en la cual los nodos representan variables aleatorias y los arcos representan relaciones de dependencia directa entre las variables.

Red bayesiana (RB) que representa cierto conocimiento sobre medicina.



En este caso, los nodos representan enfermedades, síntomas y factores que causan algunas enfermedades. La variable a la que apunta un arco es dependiente de la que esta en el origen de este.

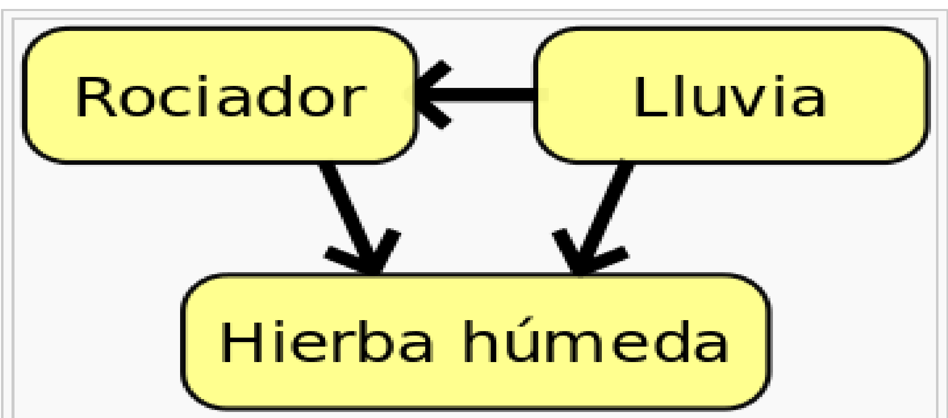
Los nodos representan variables aleatorias y los arcos relaciones de dependencia.

Dados los síntomas, la red puede ser usada para computar las probabilidad de la presencia de varias enfermedades.

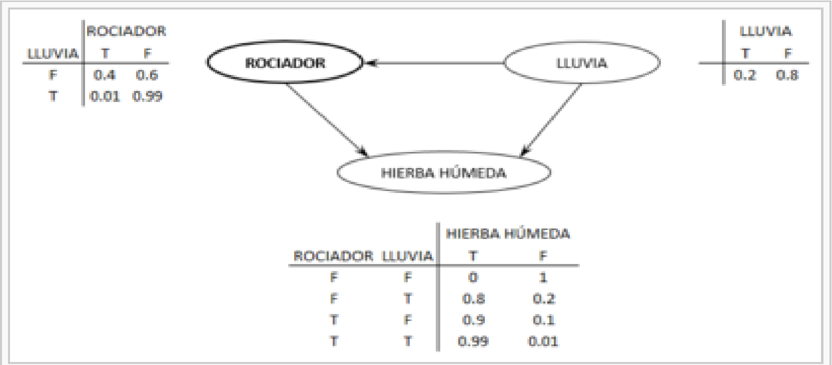
Existen algoritmos eficientes que llevan a cabo la inferencia.

Las RBs que pueden representar y resolver problemas de decisión bajo incertidumbre son llamados diagramas de influencia.

Supongamos que hay dos eventos los cuales pueden causar que la hierba esté húmeda: que el rociador esté activado o que esté lloviendo. También supongamos que la lluvia tiene un efecto directo sobre el uso del rociador (usualmente cuando llueve el rociador se encuentra apagado).



¿Cuál es la probabilidad de que esté lloviendo dado que la hierba está húmeda?



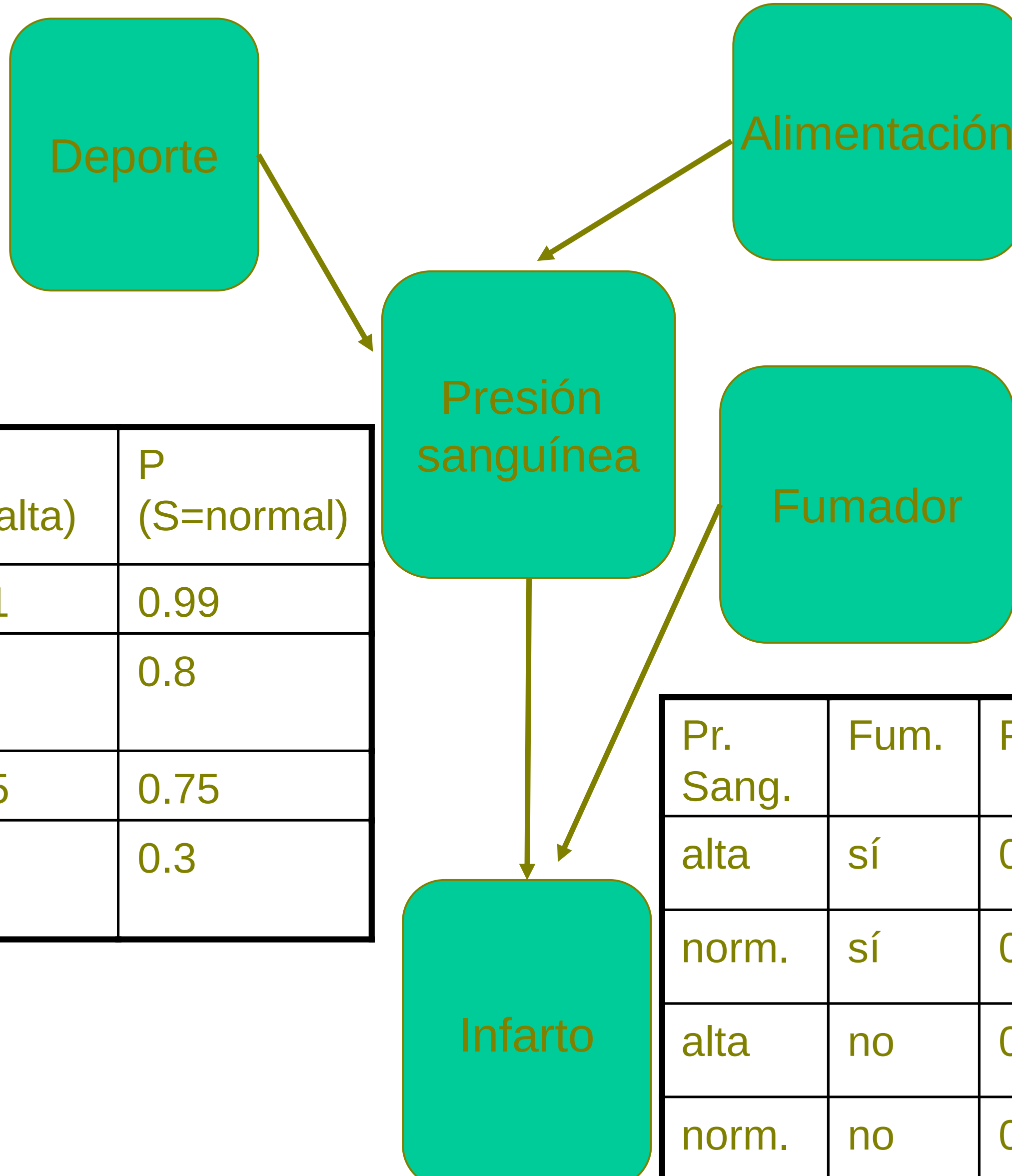
Ejemplo de una red bayesiana simple.

$$\mathrm{P}(G,S,R) = \mathrm{P}(G|S,R)\mathrm{P}(S|R)\mathrm{P}(R)$$

Ejemplo

Deporte	P(D)
sí	0.1
no	0.9

Alimentación	P(A)
equilibrada	0.4
no equilibrada	0.6



Alim.	Deporte	P (S=alta)	P (S=normal)
eq.	sí	0.01	0.99
no eq.	sí	0.2	0.8
eq.	no	0.25	0.75
no eq.	no	0.7	0.3

Fumador	P(F)
sí	0.4
no	0.6

Pr. Sang.	Fum.	P(I=sí)	P(I=no)
alta	sí	0.8	0.2
norm.	sí	0.6	0.4
alta	no	0.7	0.3
norm.	no	0.3	0.7

Redes Bayesianas - distribución conjunta

P(Infarto = sí ∧ Presión = alta ∧ Fumador = sí ∧ Deporte = si ∧ Alimentación = equil.)

=P(Infarto = sí | Presión = alta, Fumador = sí) P(Presión = alta | Deporte = sí, Alimentación = equil.) P(Fumador = sí) P(Deporte = sí) P(Alimentación = equil.)

= 0,8 × 0,01 × 0,4 × 0,1 × 0,4

= 0,000128

- **Inferencia por enumeración**: Cualquier probabilidad condicionada se puede calcular como la suma de todos los posibles casos a partir de la distribución de probabilidad conjunta:

$$P(X|\mathbf{e}) = \alpha P(X, \mathbf{e}) = \alpha \sum_{y} P(X, \mathbf{e}, \mathbf{y})$$

- Si enumeramos todas las posibilidades y las sumamos de acuerdo con la distribución de probabilidad conjunta tenemos que:

$$P(Fumador | Infarto = si, Deporte = no)$$

$$= \alpha \langle 0{,}9 \cdot 0{,}4 \cdot (0{,}4 \cdot (0{,}25 \cdot 0{,}8 + 0{,}75 \cdot 0{,}6) + 0{,}6 \cdot (0{,}7 \cdot 0{,}8 + 0{,}3 \cdot 0{,}6))$$

$$0{,}9 \cdot 0{,}6 \cdot (0{,}4 \cdot (0{,}25 \cdot 0{,}7 + 0{,}75 \cdot 0{,}3) + 0{,}6 \cdot (0{,}7 \cdot 0{,}7 + 0{,}3 \cdot 0{,}3)) \rangle$$

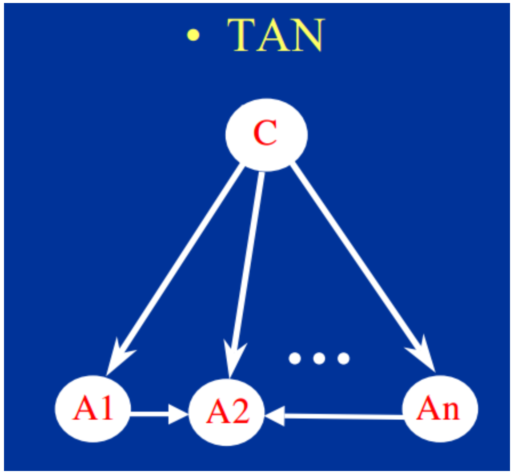
Clasificador bayesiano simple

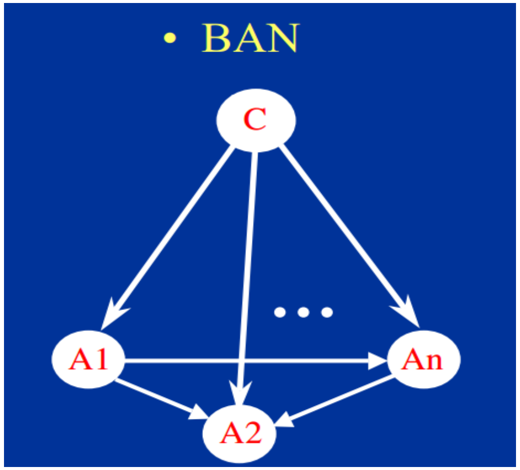
- Un clasificador bayesiano obtiene la probabilidad posterior de cada clase, Ci, usando la regla de Bayes:

- El clasicador bayesiano simple (naive Bayes classier, NBC) asume que los atributos son independientes entre si dada la clase,

¿TAN; es un clasificador bayesiano simple aumentado con un árbol.

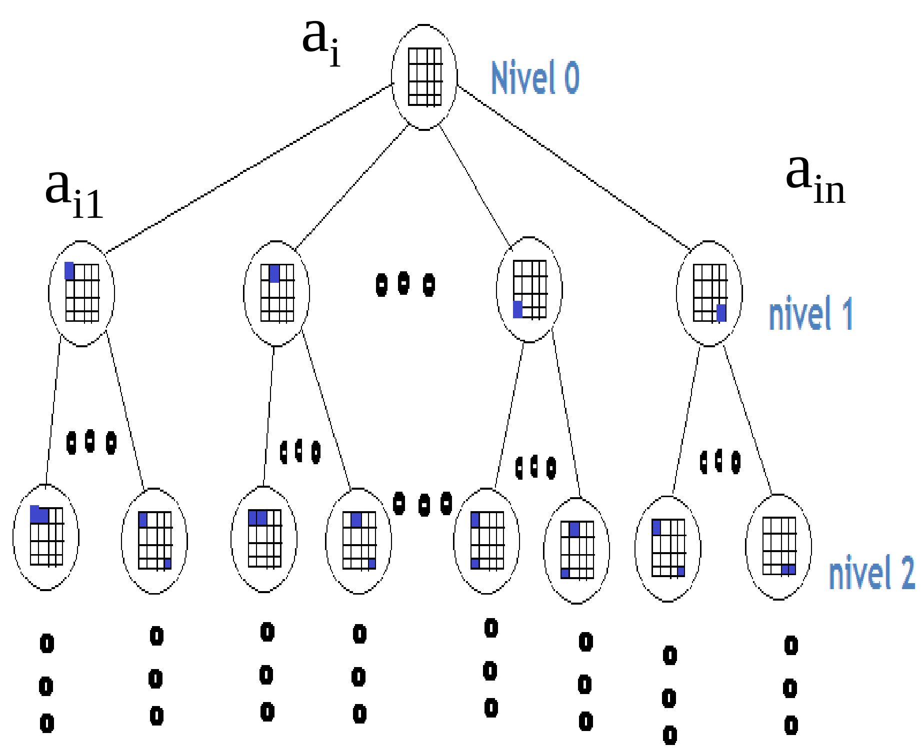
BAN: clasificador bayesiano simple aumentado con una red.





Red bayesiana para el manejo de incertidumbre

Según MUE, la mejor acción será entre todas las posibles jugadas desde a_i, aquella en la cual se maximice la razón de la utilidad de la acción y la probabilidad de que ella ocurra dado que ocurrió a_i

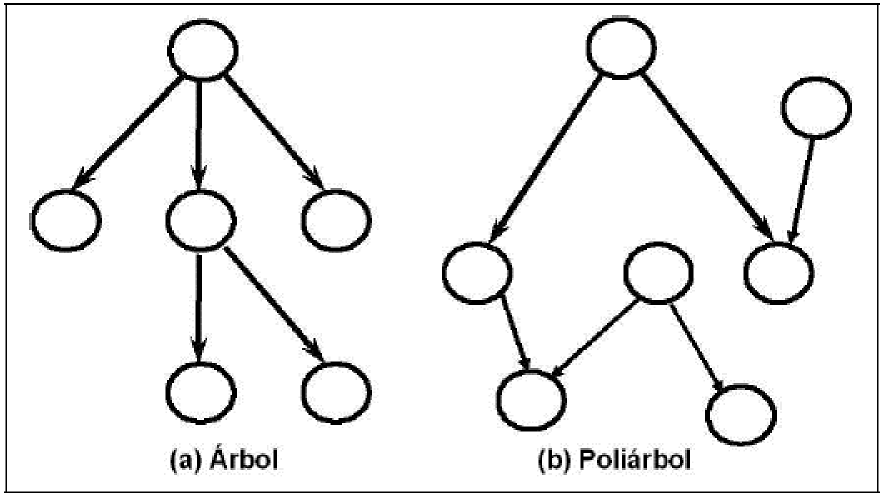


se puede aplicar desde nodo hoja recursivamente

$$MUE = \max_j \left(U(a_{ij}) * P(a_{ij} | a_i) \right) \quad \forall j = 1, n$$

Propagación en Arboles

Este algoritmo se aplica a estructuras de tipo árbol, y se puede extender a poliarboles (un nodo puede tener mas de un padre).



En un árbol, cualquier nodo (B) divide la red en dos subgrafos condicionalmente independientes, E+ y E-, por lo que podemos dividir la evidencia en dos grupos:
E-: Datos en el árbol que cuya raíz es B.
E+: Datos en el resto del árbol.