# CI5438. Inteligencia Artificial II
# Clase 11: Decisiones Racionales - Utilidad
# Cap 16 Russel & Norvig

Ivette C. Martínez

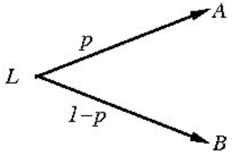
Universidad Simón Bolívar

26 de noviembre de 2008

# Preferencias

Un agente escoge entre premios ($A$, $B$, etc.) y loterías, i.e., situaciones con premios inciertos.

Lotería $L = [p, A;\ (1-p), B]$




Notación:

| | |
|---|---|
| $A \succ B$ | $A$ se prefiere sobre $B$ |
| $A \sim B$ | indiferencia entre $A$ y $B$ |
| $A \succsim B$ | $B$ no se prefiere sobre $A$ |

# Preferencias Racionales

Idea: Las preferencia de un agente racional deben obedecer ciertas restricciones.

Preferencia racionales $\implies$ Comportamiento que se puede describir como la maximización de la utilidad esperada

Restricciones:

Ordenabilidad

$(A \succ B) \vee (B \succ A) \vee (A \sim B)$

Transitividad

$(A \succ B) \wedge (B \succ C) \implies (A \succ C)$

Continuidad

$A \succ B \succ C \implies \exists\, p \;\; [p, A;\; 1-p, C] \sim B$

Substituibilidad

$A \sim B \implies [p, A;\; 1-p, C] \sim [p, B; 1-p, C]$

Monotonicidad

$A \succ B \implies (p \geq q \;\Leftrightarrow\; [p, A;\; 1-p, B] \succsim [q, A;\; 1-q, B])$

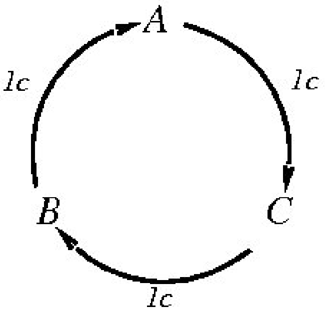
# Preferencias racionales (cont.)

La violación de estas restricciones conlleva a una irracionalidad evidente.
Ejemplo: Un agente con preferencias no-transitivas puede ser inducido a regalar todo su dinero.

Si $B \succ C$, entonces un agente que tiene $C$ pagaría 1 centavo para obtener $B$

Si $A \succ B$, entonces un agente que tiene $B$ pagaría 1 centavo para obtener $A$

Si $C \succ A$, entonces un agente que tiene $A$ pagaría 1 centavo para obtener $C$



# Maximizando la utilidad esperada

*Teorema* (Ramsey, 1931; von Neumann and Morgenstern, 1944):
Dadas preferencias que satisfacen las restricciones
existe una función en los reales $U$ tal que:

$$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$$

$$U([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i p_i U(S_i)$$

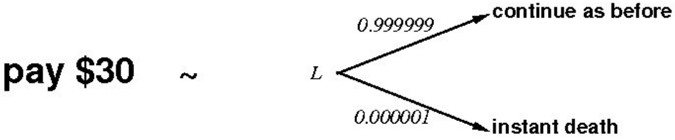
Principio de la MUE:
Escoger la acción que maximiza la utilidad esperada.

Nota: Un agente puede ser completamente racional (consistente con la MUE) sin tener que representar utilidades y probabilidades
Ejm: Una tabla de "lookup" para un juego de "la vieja" perfecto.

# Utilidades

Las Utilidades mapean estado a números reales. Cuáles números?

La aproximación estandard para calcular la utilidades humanas:
Comparar un estado dado $A$ con una lotería estadard $L_p$ que tenga:
"el mejor precio posible" $u_\top$ con probabilidad $p$
"la peor catástrofe posible" $u_\perp$ con probabilidad $(1 - p)$
ajustar la probabilidad de la lotería $p$ hasta que $A \sim L_p$



# Escalas de Utilidad

Utilidades Normalizadas: $u_\top = 1{,}0$, $u_\perp = 0{,}0$

Micromortalidad: una muerte entre un millón
Útil para la "Ruleta Rusa", pagar para reducir los riesgos de un producto, etc.

AVAC: Año de vida ajustado a la calidad
Útil para decisiones médicas que involucran riesgos sustanciales

Nota: El comportamiento es *invariante*

$$U'(x) = k_1 U(x) + k_2 \quad \text{donde } k_1 > 0$$

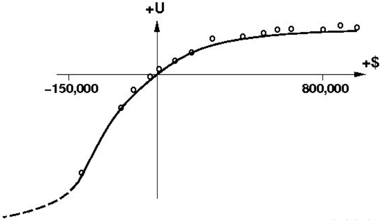
Con sólo premios determinísticos (sin opciones de lotería), sólo de puede determinar la utilidad ordinal, i.e., el ordenamiento total sobre los premios.

# Dinero

El dinero *no* se comporta como una función de utilidad.

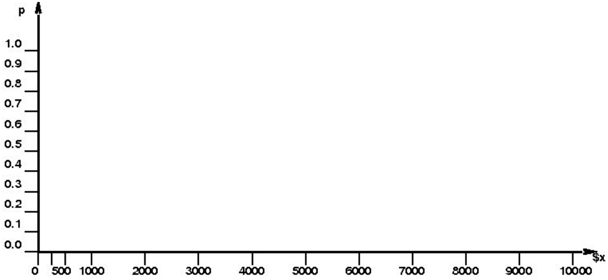
Dada una lotereia $L$ con valor monetario esperado $EMV(L)$, usualmente $U(L) < U(EMV(L))$, i.e., la gente es adversa-al-riesgo.

Curva de Utilidad: Para que probabilidad $p$ soy indiferente entre un premio $x$ y una lotería $[p, \$M;\ (1-p), \$0]$ para un $M$ grande?
Los datos empíricos típicos se extrapolan con un comportamiento risk-prone:



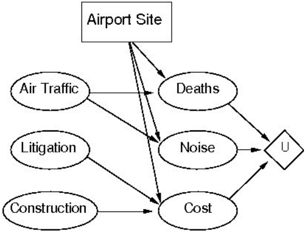
# Utilidad del grupo de estudiantes

Para cada $x$, ajustar $p$ hasta que la mitad de la clase vote por la lotería ($M = 10000$)



# Redes de Decisión

Añadir nodos de acción y nodos de utilidad a las redes de creencias para permitir la toma de decisiones racionales.




Algoritmo:

```
    calcular el valor esperado de la utilidad dada la acci
y la evidencia
retornar la accion con MUE.
```

# Utilidades Multiatributos

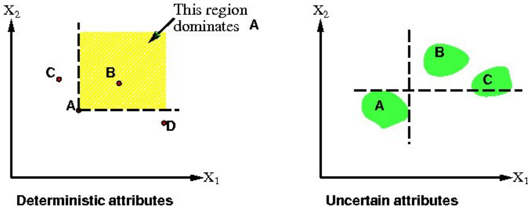
- Como podemos manejar funciones de utilidad de varias variables $X_1 \dots X_n$?
  Ej., cuál es $U(Deaths, Noise, Cost)$?
- Cómo podemos calcular funciones de utlidad complejas a partir de comportamientos de preferencias.
- Idea 1: Identificar las condiciones bajo las cuales las decisiones pueden tomarse sin identificar completamente a $U(x_1, \dots, x_n)$
- Identificar varios tipos de *independencia* en las preferencias y derivar las formas canónicas correspondientes para $U(x_1, \dots, x_n)$

# Dominancia Estricta

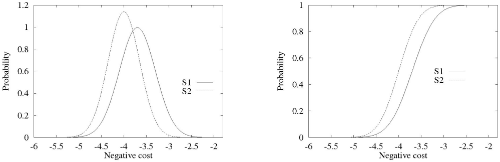
Típicamente se definen atributos tales que para cada uno $U$ sea monotónico

Dominancia Estricta: la opción $B$ domina estrictamente a la opción $A$ si y solo si

$\forall\, i \;\; X_i(B) \geq X_i(A)$ (y por lo tanto $U(B) \geq U(A)$)




La dominancia estricta se ve poco en la práctica




La distribución $p_1$ domina estocásticamente a la distribución $p_2$ si y solo si

$$\forall\, t \quad \int_{-\infty}^{t} p_1(x)dx \leq \int_{-\infty}^{t} p_2(t)dt$$

Si $U$ es monotona en $x$, entonces $A_1$ con distribución de salida $p_1$ domina estocásticamente a $A_2$ con distribución de salida $p_2$:

$$\int_{-\infty}^{\infty} p_1(x)U(x)dx \geq \int_{-\infty}^{\infty} p_2(x)U(x)dx$$

Caso Multi-atributo: dominancia estocástica sobre todos los atributos $\implies$ optimal

# Dominancia estocástica (cont.)

La dominancia estocástica muchas veces puede se determinada sin la distribución exacta usando razonamiento *cualitativo*

Ejm.Los costos de construcción se incrementan con la distancia de la ciudad
$S_1$ estás más cerca de la ciudad que $S_2$
$\implies S_1$ domina estocásticamente $S_2$ en costo

Ejm. Los daños se incrementan con la velocidad del choque.
Se pueden anotar las redes de creencias con información sobre dominancia estocástica:
$X \xrightarrow{+} Y$ ($X$ influencia positivamente $Y$) significa que:
Para cada valor $\mathbf{z}$ de los otros padres de $Y$ $\mathbf{Z}$
$\forall x_1, x_2 \;\; x_1 \geq x_2 \implies \mathbf{P}(Y|x_1, \mathbf{z})$ domina estocásticamente a $\mathbf{P}(Y|x_2, \mathbf{z})$

# Estructura de Preferencias: Determinística

$X_1$ y $X_2$ son idependiente preferencialmente de $X_3$ si y solo si las preferencias entre $x_1, x_2, x_3$ y $x_1', x_2', x_3$ no dependen de $x_3$

Ejm. $Noise, Cost, Safety$ :

20,000 suffer, $4.6 billion, 0.06 deaths/mpm vs.

70,000 suffer, $4.2 billion, 0.06 deaths/mpm

*Teorema* (Leontief, 1947): Si cada par de atributos es P.I. de sus complementos, entonces cada subconjunto de atributos es P.I. de sus complementos: P.I. mútua.

*Teorema* (Debreu, 1960): P.I. mútua $\implies \exists$ función de valor aditiva:

$$V(S) = \sum_i V_i(X_i(S))$$

Luego, asumir mat*n* como una función con 1 atributo, suele ser una buena aproximación.

# Estructura de Preferencias: Estocástica

Necesitamos considerar las preferencias sobre las loterías:
**X** es independiente-en-utilidad de **Y** si y solo si
la preferencias sobre las loterías en **X** no dependen de **y.ß**
U.I. Mútua: Cada subconjunto es U.I. de su complemento
$\implies \exists$ fución de utilidad multiplicativa:

$$U = k_1 U_1 + k_2 U_2 + k_3 U_3$$
$$+ k_1 k_2 U_1 U_2 + k_2 k_3 U_2 U_3 + k_3 k_1 U_3 U_1$$
$$+ k_1 k_2 k_3 U_1 U_2 U_3$$

Procedimientos de rutina y paquetes de software para generar pruebas de preferencias para identificar varias familias canónicas de funciones de utilidad.

# Valor de la información

Idea: Calcular el valor de adquirir cada posible pieza de evidencia
Se puede hacer **directamente a partir de una red de decisión**
Ejemplo: Dos yacimientos $A$ y $B$, exactamente uno tiene petróleo, con un valor de $k$

Probabilidades a priori 0,5 cada uno
El precio actual de cada yacimiento es $k/2$
Un "consultor ofrece un estudio preciso sobre $A$. Cuánto se debe pagar por este estudio?

Solución: Calcular el valor esperado de la información
= valor esperado de la mejor acción dada la información
El estudio dice: "hay petróleo en A" o "no hay petróleo en A", *prob. 0.5 each* (dado!)
$= [0{,}5 \times$ valor de "comprar A" dado "petróleo en A"
$+ 0{,}5 \times$ valor de "comprar B" dado "no hay petróleo en A"]
$= (0{,}5 \times k/2) + (0{,}5 \times k/2) = k/2$

# Fórmula General

Evidencia Actual $E$, mejor acción actual $\alpha$
Posibles efectos de lasacciones $S_i$, nuevas evidencias potenciales $E_j$

$$EU(\alpha|E) = \max_a \sum\nolimits_i U(S_i)\ P(S_i|E, a)$$

Supongamos que sabemos $E_j = e_{jk}$, entonces escogeríamos $\alpha_{e_{jk}}$ s.t.

$$EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) = \max_a \sum\nolimits_i U(S_i)\ P(S_i|E, a, E_j = e_{jk})$$

$E_j$ es una variable aleatória cuyo valor *actual* es desconocido
$\Longrightarrow$ se debe calcular la ganancia esperada sobre todos sus valores posibles:

$$VPI_E(E_j) = \left(\sum\nolimits_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$$

(VPI = *v*alue of perfect information)