Tutorial 4

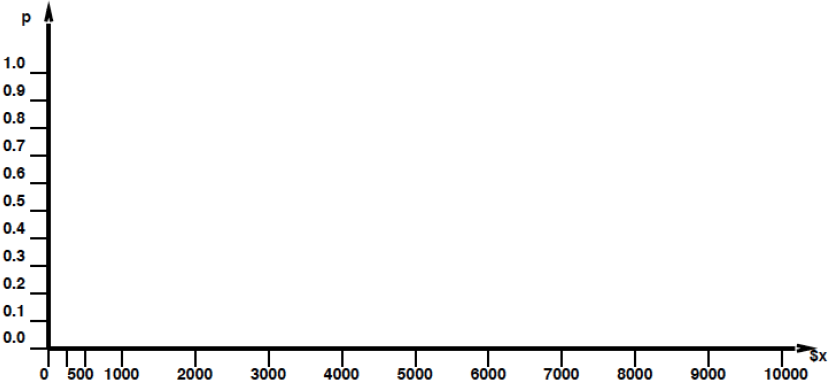
Rational Decisions

Pietro Pasotti

06 October 2016

Ex.1.: Drawing the utility curve

For each x, indicate at which value of p would you be indifferent between a sure prize x or a lottery with [p, M; (1-p), 0] for M=10,000?



Ex.2.: Buy the book or not?

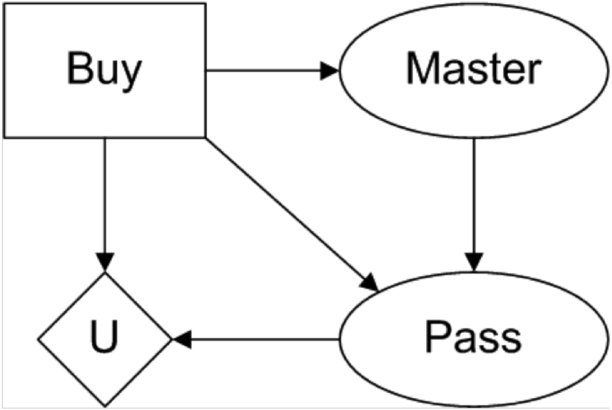
B - buy
M - master
P - pass
U - utility

a) Draw the decision network for the problem of Sam passing or not the AIT course (see lecture slide 28)

Decision node - rectangle
Chance node - oval
Utility node - diamond

Ex.2.: Buy the book or not?

a) Draw the decision network for the problem



Ex.2.: Buy the book or not?

b) Compute the expected utility of buying the book, and of not buying it (see lecture slide 38).

... how?

$$EU(\alpha|e) = max_a\big(\Sigma_{s'} P(s'|a,e) \cdot U(s')\big)$$

In our case, Sam's next state is one of two ways: either he passes the AIT exam (very good) or he doesn't (very bad).

The probability that he gets a pass depends on whether he has **mastered the book's contents** which in turn depends on whether he has **bought the book**.

The value of buying the book, then, is the expected value of *the best action* given that you buy it (denoted $EU(b)$), minus the expected value of *the best action* given that you don't (denoted $EU(\neg b)$).
However there is in fact no action, because we have no other choice than whether to buy the book or not!

Ex.2.: Buy the book or not?

$$EU(\alpha|e) = max_a\Big(\sum_{s'} P(s'|a,e) \cdot U(s')\Big)$$

$$EU(b|b) = max_a\Big(\sum_{s' \in \{p,\neg p\} \times \{b\}} P(s'|g,b) \cdot U(s')\Big)$$

– We introduce an action g ('go to exam') which has probability 1.

Ex.2.: Buy the book or not?

$$EU(b) = P(p|b) \cdot (U(p) + U(b))$$
$$+P(\neg p|b) \cdot (U(\neg p) + U(b))$$
$$EU(\neg b) = P(p|\neg b) \cdot (U(p) + U(\neg b))$$
$$+P(\neg p|\neg b) \cdot (U(\neg p) + U(\neg b))$$

We know from Sam:

$U(\neg b) = 0$

$U(b) = -100$

$U(p) = 2000$

$U(\neg p) = 0$

Ex.2.: Buy the book or not?

We need to compute:

$P(p|b) =$
$P(p|\neg b) =$
$P(\neg p|b) =$
$P(\neg p|\neg b) =$

Ex.2.: Buy the book or not?

b) Compute the expected utility of buying the book, and of not buying it.

We need to compute:

$$P(p|b) = P(m|b) \cdot P(p|b, m) + P(\neg m|b) \cdot P(p|b, \neg m)$$
$$P(p|\neg b) = P(m|\neg b) \cdot P(p|\neg b, m) + P(\neg m|\neg b) \cdot P(p|\neg b, \neg m)$$
$$P(\neg p|b) = 1 - P(p|b)$$
$$P(\neg p|\neg b) = 1 - P(p|\neg b)$$

Ex.2.: Buy the book or not?

b) Compute the expected utility of buying the book, and of not buying it.

We need to compute:

$P(p|b) = P(m|b) \cdot P(p|b, m) + P(\neg m|b) \cdot P(p|b, \neg m)$

$P(p|b) = 0.9 \cdot 0.9 + 0.1 \cdot 0.5 = 0.86$

$P(p|\neg b) = P(m|\neg b) \cdot P(p|\neg b, m) + P(\neg m|\neg b) \cdot P(p|\neg b, \neg m)$

$P(p|\neg b) = 0.7 \cdot 0.8 + 0.3 \cdot 0.3 = 0.65$

$P(\neg p|b) = 1 - P(p|b) = 0.14$

$P(\neg p|\neg b) = 1 - P(p|\neg b) = 0.35$

Ex.2.: Buy the book or not?

b) Compute the expected utility of buying the book, and of not buying it.

Putting it all together:

$P(p \mid b) = 0.86$	$U(\neg b) = 0$
$P(p \mid \neg b) = 0.65$	$U(b) = -100$
$P(\neg p \mid b) = 0.14$	$U(p) = 2000$
$P(\neg p \mid \neg b) = 0.35$	$U(\neg p) = 0$

$$EU(b) = P(p \mid b) \cdot (U(p) + U(b)) + P(\neg p \mid b) \cdot (U(\neg p) + U(b))$$
$$EU(b) = 0.86 \cdot (2000 - 100) + 0.14 \cdot (0 - 100) = 1620$$
$$EU(\neg b) = P(p \mid \neg b) \cdot (U(p) + U(\neg b)) + P(\neg p \mid \neg b) \cdot (U(\neg p) + U(\neg b))$$
$$EU(\neg b) = 0.65 \cdot (2000 + 0) + 0.35(0 + 0) = 1300$$

Ex.2.: Buy the book or not?

c) What should Sam do?

$EU(buying) = 1620$
$EU(\neg buying) = 1300$

Ex.2.: Buy the book or not?

c) What should Sam do?

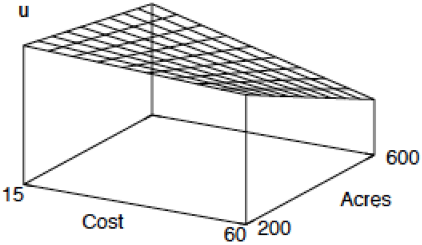
$EU(buying) = 1620$
$EU(\neg buying) = 1300$

$\Rightarrow$ Sam should buy the book.

Ex.3.: Utrecht power generation station

Draw the decision network for the problem.

Ex.3.: Utrecht power generation station



$u(Cost, Acres)$

$u(60, 600) = 0$, $u(15, 200) = 1 \Rightarrow$ worst and best options
Model the utility function $u(Cost, Acres)$ as an additive function.
Conditional utility functions u_C and u_A values are given.
Cost is three times as important as Acres lost.

Calculate the resulting utilities for the following Cost-Acre pairs:
$u(50, 300)$, $u(30, 400)$, $u(60, 200)$, $u(15, 600)$, $u(15, 200)$.

Ex.3.: Utrecht power generation station

$u(Cost, Acres)$
$u(60, 600) = 0$, $u(15, 200) = 1$
Additive utility function: $u(X, Y) = k_X \cdot u_X(X) + k_Y \cdot u_Y(Y)$, where $u_X(X)$ is a conditional utility function for X and k_X, k_Y are weights or scaling constants.
Conditional utility functions u_C and u_A values are given.
Cost is three times as important as Acres lost.

$u(50, 300) =$
$u(30, 400) =$
$u(60, 200) =$
$u(15, 600) =$
$u(15, 200) =$

Ex.3.: Utrecht power generation station

$u(Cost, Acres)$
$u(60, 600) = 0$, $u(15, 200) = 1$
$u(X, Y) = k_X \cdot u_X(X) + k_Y \cdot u_Y(Y)$
Conditional utility functions u_C and u_A values are given.
Cost is three times as important as Acres lost.
$k_X = 0.75; k_Y = 0.25$

$u(50, 300) = 0.75 \cdot u_C(50) + 0.25 \cdot u_A(300) =$
$0.75 \cdot 0.2 + 0.25 \cdot 0.8 = 0.35$
$u(30, 400) = 0.75 \cdot 0.5 + 0.25 \cdot 0.5 = 0.5$
$u(60, 200) = 0.75 \cdot 0.0 + 0.25 \cdot 1 = 0.25$
$u(15, 600) = 0.75 \cdot 1.0 + 0.25 \cdot 0.0 = 0.75$
$u(15, 200) = 0.75 \cdot 1.0 + 0.25 \cdot 1.0 = 1.0$

Ex.4.: Immediate reward vs long-term punishment

Start state: Up or Down,
All other states: Right.
Assuming a discounted reward function, for what values of the discount γ should the agent choose Up and for which Down?
Compute the utility of each action as a function of γ.

+50	-1	-1	-1	...	-1	-1	-1	$\boxed{-1}$
Start				...				
-50	+1	+1	+1	...	+1	+1	+1	$\boxed{+1}$

Ex.4.: Immediate reward vs long-term punishment

Compute the utility of each action as a function of γ.

Discounted rewards:
$U([s_0, s_1, s_2, ...]) = R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + ...$

$U(Up) =$
$U(Down) =$

+50	-1	-1	-1	...	-1	-1	-1	$\boxed{-1}$
Start				...				
-50	+1	+1	+1	...	+1	+1	+1	$\boxed{+1}$

Ex.4.: Immediate reward vs long-term punishment

Discounted rewards:

$$U([s_0, s_1, s_2, \ldots]) = R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \ldots$$

$$U(Up) = R(Start) + \gamma R(Up) + \gamma^2 R(Right) + \ldots + \gamma^{101} R(Right)$$
$$U(Down) = R(Start) + \gamma R(Down) + \gamma^2 R(Right) + \ldots + \gamma^{101} R(Right)$$

+50	-1	-1	-1	...	-1	-1	-1	$\boxed{-1}$
Start				...				
-50	+1	+1	+1	...	+1	+1	+1	$\boxed{+1}$

Ex.4.: Immediate reward vs long-term punishment

Discounted rewards:
$U([s_0, s_1, s_2, ...]) = R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + ...$

$U(Up) = 0 + 50\gamma + \sum_{i=2}^{101} \gamma^i(-1)$
$U(Down) = 0 - 50\gamma + \sum_{i=2}^{101} \gamma^i$

+50	-1	-1	-1	. . .	-1	-1	-1	-1
Start				. . .				
-50	+1	+1	+1	. . .	+1	+1	+1	+1

Ex.4.: Immediate reward vs long-term punishment

Discounted rewards:
$U([s_0, s_1, s_2, \ldots]) = R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \ldots$

$U(Up) = 50\gamma - \sum_{i=2}^{101} \gamma^i$
$U(Down) = -50\gamma + \sum_{i=2}^{101} \gamma^i$

At what value of γ is there indifference between Up or Down?

+50	-1	-1	-1	. . .	-1	-1	-1	$\boxed{-1}$
Start				. . .				
-50	+1	+1	+1	. . .	+1	+1	+1	$\boxed{+1}$

Ex.4.: Immediate reward vs long-term punishment

Discounted rewards:
$U([s_0, s_1, s_2, \ldots]) = R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \ldots$

$U(Up) = 50\gamma - \sum_{i=2}^{101} \gamma^i$
$U(Down) = -50\gamma + \sum_{i=2}^{101} \gamma^i$

We must solve
$U(Up) = U(Down)$

Ex.4.: Immediate reward vs long-term punishment

At around $\gamma \approx 0.9844$ is the indifference point.
Higher than this, we want to go Down to avoid long-term expensive consequences.
Lower than this, we want to go Up to get the immediate benefit.
Hence, γ describes the preference of rewards in time.