Risk and Decisions (II)

Paolo Turrini

Department of Computing, Imperial College London

Introduction to Artificial Intelligence

The lectures

- The agent and the world (**Knowledge Representation**)
  - Actions and knowledge
  - Inference
- Good decisions (**Risk and Decisions**)
  - Chance
  - Gains
- Good decisions in time (**Markov Decision Processes**)
  - Chance and gains in time
  - Patience
  - Finding the best strategy
- Learning from experience (**Reinforcement Learning**)
  - Finding a reasonable strategy

Today

- Utility and expected utility
- Risky moves

Risk and Decisions

Lotteries (and how to win them)

The book

Stuart Russell and Peter Norvig
Artificial Intelligence: a modern approach
Chapters 16-17

If you snooze you lose?

I set the alarm clock(s) to wake up on time for the lectures.

If you snooze you lose?

I set the alarm clock(s) to wake up on time for the lectures.

Let action S_t = snooze the alarm clock t times

If you snooze you lose?

I set the alarm clock(s) to wake up on time for the lectures.

Let action S_t = snooze the alarm clock t times

Will S_t get me there on time?

If you snooze you lose?

I set the alarm clock(s) to wake up on time for the lectures.

Let action S_t = snooze the alarm clock t times

Will S_t get me there on time?

Potential problems, e.g.:

1. planned engineering works
2. my phone dies
3. my mum forgets to call me (very unlikely)

If you snooze you lose?

Suppose I believe the following:

If you snooze you lose?

Suppose I believe the following:

$$
\begin{aligned}
P(S_0 \text{ gets me there on time} | \ldots) &= 0.99 \\
P(S_1 \text{ gets me there on time} | \ldots) &= 0.90 \\
P(S_3 \text{ gets me there on time} | \ldots) &= 0.6 \\
P(S_{10} \text{ gets me there on time} | \ldots) &= 0.1
\end{aligned}
$$

If you snooze you lose?

Suppose I believe the following:

$$
\begin{aligned}
P(S_0 \text{ gets me there on time} | \ldots) &= 0.99 \\
P(S_1 \text{ gets me there on time} | \ldots) &= 0.90 \\
P(S_3 \text{ gets me there on time} | \ldots) &= 0.6 \\
P(S_{10} \text{ gets me there on time} | \ldots) &= 0.1
\end{aligned}
$$

Which action should I choose?

If you snooze you lose?

Suppose I believe the following:

$$
\begin{aligned}
P(S_0 \text{ gets me there on time}|\ldots) &= 0.99\\
P(S_1 \text{ gets me there on time}|\ldots) &= 0.90\\
P(S_3 \text{ gets me there on time}|\ldots) &= 0.6\\
P(S_{10} \text{ gets me there on time}|\ldots) &= 0.1
\end{aligned}
$$

Which action should I choose?

IT DEPENDS

If you snooze you lose?

Suppose I believe the following:

$$
\begin{aligned}
P(S_0 \text{ gets me there on time} | \ldots) &= 0.99 \\
P(S_1 \text{ gets me there on time} | \ldots) &= 0.90 \\
P(S_3 \text{ gets me there on time} | \ldots) &= 0.6 \\
P(S_{10} \text{ gets me there on time} | \ldots) &= 0.1
\end{aligned}
$$

Which action should I choose?

IT DEPENDS on my preferences

If you snooze you lose?

Suppose I believe the following:

$$\begin{aligned} P(S_0 \text{ gets me there on time} | \ldots) &= 0.99 \\ P(S_1 \text{ gets me there on time} | \ldots) &= 0.90 \\ P(S_3 \text{ gets me there on time} | \ldots) &= 0.6 \\ P(S_{10} \text{ gets me there on time} | \ldots) &= 0.1 \end{aligned}$$

Which action should I choose?

IT DEPENDS on my preferences

e.g., missing class vs. sleeping

Chances + preferences

Utility theory is used to represent and reason with preferences

Chances + preferences

Utility theory is used to represent and reason with preferences

Decision theory = utility theory + probability theory

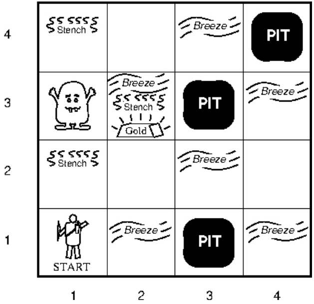
Rewards

Sensors Breeze, Glitter, Smell

Actuators Up, Down, Left, Right, Grab, Release, Shoot, Climb

Rewards 1000 escaping with gold, -1000 dying, -10 using arrow, -1 walking

Environment
- Squares adjacent to Wumpus are smelly
- Squares adjacent to pit are breezy
- Glitter iff gold is in the same square
- Shooting kills Wumpus if you are facing it
- Shooting uses up the only arrow
- Grabbing picks up gold if in same square
- Releasing drops the gold in same square



Rewards

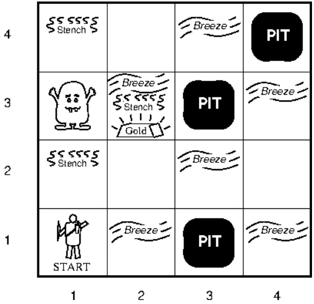
Sensors Breeze, Glitter, Smell

Actuators Up, Down, Left, Right, Grab, Release, Shoot, Climb

Rewards 1000 escaping with gold, -1000 dying, -10 using arrow, -1 walking

Environment

- Squares adjacent to Wumpus are smelly
- Squares adjacent to pit are breezy
- Glitter iff gold is in the same square
- Shooting kills Wumpus if you are facing it
- Shooting uses up the only arrow
- Grabbing picks up gold if in same square
- Releasing drops the gold in same square



State space

The universe in which the agent moves is a finite set of states

$$S = \{s_1, \ldots, s_n\}$$

State space

The universe in which the agent moves is a finite set of states

$$S = \{s_1, \ldots, s_n\}$$

e.g., the possible grid configurations in the Wumpus World

State space

The universe in which the agent moves is a finite set of states

$$S = \{s_1, \dots, s_n\}$$

e.g., the possible grid configurations in the Wumpus World

- States can also contain a description of:

State space

The universe in which the agent moves is a finite set of states

$$S = \{s_1, \ldots, s_n\}$$

e.g., the possible grid configurations in the Wumpus World

- States can also contain a description of:
  - the inner state of the agent, e.g., the knowledge base KB
  - relevant changes happened
  - the history of the game so far

State space

The universe in which the agent moves is a finite set of states

$$S = \{s_1, \ldots, s_n\}$$

e.g., the possible grid configurations in the Wumpus World

- States can also contain a description of:
  - the inner state of the agent, e.g., the knowledge base KB
  - relevant changes happened
  - the history of the game so far
- The set of states is our sample space

Lotteries

A lottery is a probability distribution over the set of states.

Lotteries

A lottery is a probability distribution over the set of states.
e.g., for states s_1 and s_2, and $p \in [0, 1]$

Lottery $L_1 = [p, s_1;\ (1 - p), s_2]$

Lotteries

A lottery is a probability distribution over the set of states.
e.g., for states s_1 and s_2, and $p \in [0, 1]$

Lottery $L_1 = [p, s_1;\ (1 - p), s_2]$

L is the set of lotteries over S.

States and lotteries

Observation: A state $s \in S$ can be seen as a lottery

States and lotteries

Observation: A state $s \in S$ can be seen as a lottery:

- s is assigned probability 1
- all other states probability 0

States and lotteries

Observation: A state $s \in S$ can be seen as a lottery:

- s is assigned probability 1
- all other states probability 0

e.g.,
$L_1 = [1, s_1; 0, s_2; \ldots 0, s_n]$

We get s_1 with probability 1, and the rest with probability 0.

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_n, L_n]$

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1, [p_1, s_1; p_2, s_2; \ldots p_n, s_n]; q_2, L_2; \ldots; q_n, L_n]$$

Compound lotteries

Consider now the set L of lotteries over S.

Observation: A lottery over L is a lottery over S:

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1, [p_1, s_1; p_2, s_2; \ldots p_n, s_n]; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1 p_1, s_1; q_1 p_2, s_2; \ldots q_n p_n, s_n; q_2, L_2; \ldots; q_n, L_n]$$

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1, [p_1, s_1; p_2, s_2; \ldots p_n, s_n]; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1 p_1, s_1; q_1 p_2, s_2; \ldots q_n p_n, s_n; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1, L_1; q_2[r_1, s_1; r_2, s_2; \ldots r_n, s_n]; \ldots; q_n, L_n]$$

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$$
\begin{aligned}
\mathbf{L}_1 &= [q_1, L_1; q_2, L_2; \ldots; q_n, L_n] \\
&= [q_1, [p_1, s_1; p_2, s_2; \ldots p_n, s_n]; q_2, L_2; \ldots; q_n, L_n] \\
&= [q_1 p_1, s_1; q_1 p_2, s_2; \ldots q_n p_n, s_n; q_2, L_2; \ldots; q_n, L_n] \\
&= [q_1, L_1; q_2[r_1, s_1; r_2, s_2; \ldots r_n, s_n]; \ldots; q_n, L_n] \\
&= [q_1, L_1; [q_2 r_1, s_1; q_2 r_2, s_2; \ldots q_n r_n, s_n]; \ldots; q_n, L_n]
\end{aligned}
$$

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1, [p_1, s_1; p_2, s_2; \ldots p_n, s_n]; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1 p_1, s_1; q_1 p_2, s_2; \ldots q_n p_n, s_n; q_2, L_2; \ldots; q_n, L_n]$$
$$= [q_1, L_1; q_2[r_1, s_1; r_2, s_2; \ldots r_n, s_n]; \ldots; q_n, L_n]$$
$$= [q_1, L_1; [q_2 r_1, s_1; q_2 r_2, s_2; \ldots q_n r_n, s_n]; \ldots; q_n, L_n]$$
$$= [[(q_1 p_1 + q_2 r_1), s_1; (q_1 p_2 + q_2 r_2), s_2; \ldots (q_1 p_n + q_2 r_n), s_n]; \ldots; q_n, L_n]$$

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$$
\begin{aligned}
\mathbf{L}_1 &= [q_1, L_1; q_2, L_2; \ldots ; q_n, L_n] \\
&= [q_1, [p_1, s_1; p_2, s_2; \ldots p_n, s_n]; q_2, L_2; \ldots ; q_n, L_n] \\
&= [q_1 p_1, s_1; q_1 p_2, s_2; \ldots q_n p_n, s_n; q_2, L_2; \ldots ; q_n, L_n] \\
&= [q_1, L_1; q_2[r_1, s_1; r_2, s_2; \ldots r_n, s_n]; \ldots ; q_n, L_n] \\
&= [q_1, L_1; [q_2 r_1, s_1; q_2 r_2, s_2; \ldots q_n r_n, s_n]; \ldots ; q_n, L_n] \\
&= [[(q_1 p_1 + q_2 r_1), s_1; (q_1 p_2 + q_2 r_2), s_2; \ldots (q_1 p_n + q_2 r_n), s_n]; \ldots ; q_n, L_n] \\
&= \ldots
\end{aligned}
$$

Compound lotteries

Consider now the set L of lotteries over S.
Observation: A lottery over L is a lottery over S:

$$\begin{aligned}
&\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \dots; q_n, L_n] \\
&= [q_1, [p_1, s_1; p_2, s_2; \dots p_n, s_n]; q_2, L_2; \dots; q_n, L_n] \\
&= [q_1 p_1, s_1; q_1 p_2, s_2; \dots q_n p_n, s_n; q_2, L_2; \dots; q_n, L_n] \\
&= [q_1, L_1; q_2[r_1, s_1; r_2, s_2; \dots r_n, s_n]; \dots; q_n, L_n] \\
&= [q_1, L_1; [q_2 r_1, s_1; q_2 r_2, s_2; \dots q_n r_n, s_n]; \dots; q_n, L_n] \\
&= [[(q_1 p_1 + q_2 r_1), s_1; (q_1 p_2 + q_2 r_2), s_2; \dots (q_1 p_n + q_2 r_n), s_n]; \dots; q_n, L_n] \\
&= \dots
\end{aligned}$$

Compound lotteries can be reduced to simple lotteries

Comparing lotteries: the plan

Rewards are defined only on some states, and not on others.

How do we choose between lotteries?

Comparing lotteries: the plan

Rewards are defined only on some states, and not on others.

How do we choose between lotteries?

Here is the plan:

- First we introduce a comparison relation between lotteries
- Then some intuitive properties this relation ought to have
- Then prove that it can be reduced to numbers.
  Notice: I said numbers, I haven't said money.

Comparing lotteries: the plan

Rewards are defined only on some states, and not on others.

How do we choose between lotteries?

Here is the plan:

- First we introduce a comparison relation between lotteries
- Then some intuitive properties this relation ought to have
- Then prove that it can be reduced to numbers.
  Notice: I said numbers, I haven't said money.

When we don't have numbers, we can often make them up.

Preferences

A preference relation is a relation $\succeq \subseteq L \times L$ over the set of lotteries.

Preferences

A preference relation is a relation $\succeq \subseteq L \times L$ over the set of lotteries.

- $A \succeq B$ means that lottery A is weakly preferred to lottery B.

Preferences

A preference relation is a relation $\succeq \subseteq L \times L$ over the set of lotteries.

- $A \succeq B$ means that lottery A is weakly preferred to lottery B.
- $A \succ B = (A \succeq B \text{ and not } B \succeq A)$ means that lottery A is strictly preferred to lottery B.

Preferences

A preference relation is a relation $\succeq \subseteq L \times L$ over the set of lotteries.

- $A \succeq B$ means that lottery A is weakly preferred to lottery B.
- $A \succ B = (A \succeq B \text{ and not } B \succeq A)$ means that lottery A is strictly preferred to lottery B.
- $A \sim B = (A \succeq B \text{ and } B \succeq A)$ means that lottery A the same as lottery B value-wise (indifference).

Rational preferences

Let A, B, C be three lotteries and let $p, q \in [0, 1]$.

Rational preferences

Let A, B, C be three lotteries and let $p, q \in [0, 1]$.
We call a preference relation $\succeq$ reasonable if it satisfies the following constraints:

Rational preferences

Let A, B, C be three lotteries and let $p, q \in [0, 1]$.
We call a preference relation $\succeq$ reasonable if it satisfies the following constraints:

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

Transitivity $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$

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

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

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

Orderability

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

'Either A over B, or B over A, or I don't care.'

Transitivity

$$(A \succ B) \land (B \succ C) \Rightarrow (A \succ C)$$

'If A is better than B, and B better than C, then A is better than C.'

Continuity

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

'A is better than B , that is better than C. But if you give me the right mix of A and C then this would be the same as B.'

Substituability

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

'If I'm indifferent to A and B,
then I also don't care of how likely they are.'

Monotonicity

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

'If I like A more than B,
then I'd rather have a bit more of A than a bit more of B.'

Rational preferences

Violating the constraints leads to self-evident irrationality.

Rational preferences

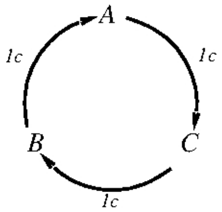
Violating the constraints leads to self-evident irrationality.

Take transitivity.

Rational preferences

Violating the constraints leads to self-evident irrationality.

Take transitivity.

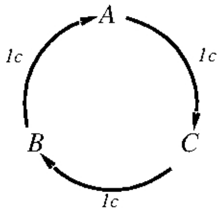


Rational preferences

Violating the constraints leads to self-evident irrationality.

Take transitivity.

If $B \succ C$, then an agent who has C would pay (say) 1 cent to get B



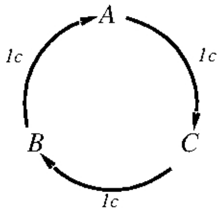
Rational preferences

Violating the constraints leads to self-evident irrationality.

Take transitivity.

If $B \succ C$, then an agent who has C would pay (say) 1 cent to get B

If $A \succ B$, then an agent who has B would pay (say) 1 cent to get A



Rational preferences

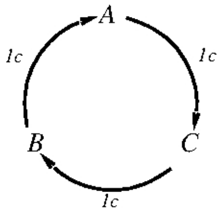
Violating the constraints leads to self-evident irrationality.

Take transitivity.

If $B \succ C$, then an agent who has C would pay (say) 1 cent to get B

If $A \succ B$, then an agent who has B would pay (say) 1 cent to get A

If $C \succ A$, then an agent who has A would pay (say) 1 cent to get C



Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$

Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i \ p_i u(S_i)$

Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i p_i u(S_i)$

$[\Leftarrow]$

Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i \ p_i u(S_i)$

[$\Leftarrow$] By contraposition. E.g., pick transitivity and show that if the relation is not transitive there is no way of associating numbers to outcomes.

Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i \ p_i u(S_i)$

[$\Leftarrow$] By contraposition. E.g., pick transitivity and show that if the relation is not transitive there is no way of associating numbers to outcomes.

[$\Rightarrow$]

Representation Theorem

Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944)

A preference relation $\succeq$ is reasonable *if and only if there exists a real-valued function* $u : L \to \mathbb{R}$ *such that:*

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \sum_i p_i u(S_i)$

[$\Leftarrow$] By contraposition. E.g., pick transitivity and show that if the relation is not transitive there is no way of associating numbers to outcomes.
[$\Rightarrow$] We use the axioms to show that there are infinitely many functions that satisfy them, but they are all "equivalent" to a unique real-valued utility function.

Representation Theorem

Michael Maschler, Eilon Solan and Shmiel Zamir
Game Theory (Ch. 2)
Cambridge University Press, 2013.

Representation Theorem

Michael Maschler, Eilon Solan and Shmiel Zamir
Game Theory (Ch. 2)
Cambridge University Press, 2013.

The main message: Give me any order on outcomes that makes sense and I can turn it into a real-valued function!

Utility functions

A utility function is a function

$$u : S \to \mathbb{R}$$

associating a real number to each state.

Utility functions

A utility function is a function

$$u : S \to \mathbb{R}$$

associating a real number to each state.

Important:
Utility functions are not the same as money. Utility functions are a representation of happiness, goal satisfaction, fulfilment and the like. They are just a mathematical tool to represent a comparison between outcomes. So altruism, unselfishness, and so fort **can** be modelled using utility functions.

Expected utility

Let $L_1 = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$ be a lottery.

Expected utility

Let $L_1 = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$ be a lottery.
The expected utility of A is

$$u(L_1) = \sum_{p_i, s_i} p_i \times u(s_i)$$

Expected utility

Let $L_1 = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$ be a lottery.
The expected utility of A is

$$u(L_1) = \sum_{p_i, s_i} p_i \times u(s_i)$$

e.g., rolling a fair six-sided dice, I win $27k$ if 6 comes out, lose $3k$ otherwise.

Expected utility

Let $L_1 = [p_1, s_1; p_2, s_2; \dots p_n, s_n]$ be a lottery.
The expected utility of A is

$$u(L_1) = \sum_{p_i, s_i} p_i \times u(s_i)$$

e.g., rolling a fair six-sided dice, I win $27k$ if 6 comes out, lose $3k$ otherwise. The expected utility is

Expected utility

Let $L_1 = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$ be a lottery.
The expected utility of A is

$$u(L_1) = \sum_{p_i, s_i} p_i \times u(s_i)$$

e.g., rolling a fair six-sided dice, I win $27k$ if 6 comes out, lose $3k$ otherwise. The expected utility is $= \frac{1}{6}27k - \frac{5}{6}3k$

Expected utility

Let $L_1 = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$ be a lottery.
The expected utility of A is

$$u(L_1) = \sum_{p_i, s_i} p_i \times u(s_i)$$

e.g., rolling a fair six-sided dice, I win $27k$ if 6 comes out, lose $3k$ otherwise. The expected utility is $= \frac{1}{6}27k - \frac{5}{6}3k = 2k.$

Humans and expected utility

'rolling a fair six-sided dice,
you win $27k$ if 6 comes out,
lose $3k$ otherwise'

Humans and expected utility

'rolling a fair six-sided dice,
you win $27k$ if 6 comes out,
lose $3k$ otherwise'

What would you do?

Humans and expected utility

'rolling a fair six-sided dice,
you win $27k$ if 6 comes out,
lose $3k$ otherwise'

What would you do?

Let's change the setup a little bit...

Humans and expected utility

'rolling a fair six-sided dice,
you win $27k$ if 6 comes out,
lose $3k$ otherwise'

What would you do?

Let's change the setup a little bit...

Modifying utilities and probabilities we can find the indifference point, passed which we change our mind.

Humans and expected utility

'rolling a fair six-sided dice,
you win $27k$ if 6 comes out,
lose $3k$ otherwise'

What would you do?

Let's change the setup a little bit...

Modifying utilities and probabilities we can find the indifference point, passed which we change our mind. Not the same for everyone!

Humans and expected utility

Tverski and Kahneman's Prospect Theory:

- Humans have complex utility estimates
- Risk aversion, satisfaction level

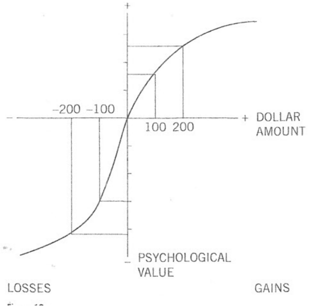


Figure: Typical empirical data

Humans and expected utility

Tverski and Kahneman's Prospect Theory:

- Humans have complex utility estimates
- Risk aversion, satisfaction level

Warning! controversial statement:

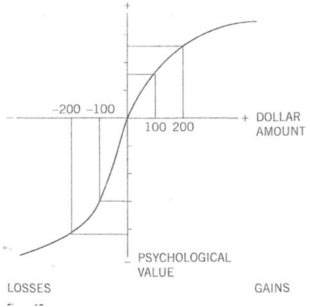


Figure: Typical empirical data

Humans and expected utility

Tverski and Kahneman's Prospect Theory:

- Humans have complex utility estimates
- Risk aversion, satisfaction level

Warning! controversial statement:

PT does not refute the principle
of maximization of expected utility.

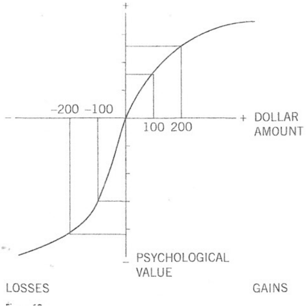


Figure: Typical empirical data

Humans and expected utility

Tverski and Kahneman's Prospect Theory:

- Humans have complex utility estimates
- Risk aversion, satisfaction level

Warning! controversial statement:

PT does not refute the principle of maximization of expected utility.

We can incorporate risk aversion and satisfaction as properties of outcomes.

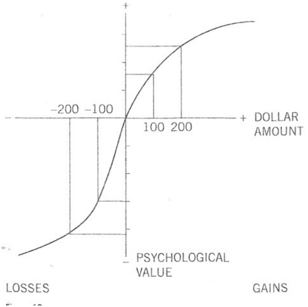


Figure: Typical empirical data

Multicriteria decision-making

- Certain outcomes seem difficult to compare:

Multicriteria decision-making

- Certain outcomes seem difficult to compare:
    - what factors are more important?

Multicriteria decision-making

- Certain outcomes seem difficult to compare:
    - what factors are more important?
    - have we considered all the relevant ones?

Multicriteria decision-making

- Certain outcomes seem difficult to compare:
  - what factors are more important?
  - have we considered all the relevant ones?
  - do factors interfere with one another?

Multicriteria decision-making

- Certain outcomes seem difficult to compare:
    - what factors are more important?
    - have we considered all the relevant ones?
    - do factors interfere with one another?
- In other situations the utility function may be updated because of new incoming information (e.g., evaluating non-terminal positions in a long extensive game like Chess or Go)

Multicriteria decision-making

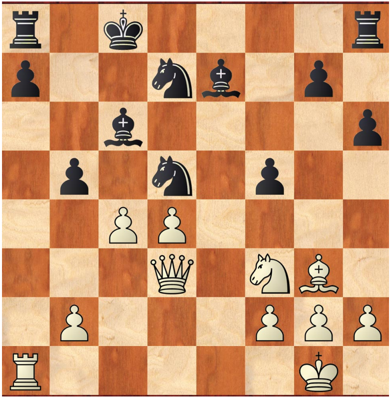


Figure: Deep Blue vs. Kasparov 1996, Final Game. Garry Kasparov (Black) to move: material favours him but the position is hopeless.

Multicriteria decision-making

How do we handle multiple many variables?

Multicriteria decision-making

How do we handle multiple many variables?
e.g., what is
$u(\text{king safety}, \text{material advantage}, \text{control of the centre})$?

Multicriteria decision-making

How do we handle multiple many variables?

e.g., what is

$u(\text{king safety}, \text{material advantage}, \text{control of the centre})$?

how do we compare lotteries over sets of states?

and sequences of states?

Multicriteria decision-making

How do we handle multiple many variables?
e.g., what is
u(king safety, material advantage, control of the centre)?
how do we compare lotteries over sets of states?
and sequences of states?

- We need to find ways to compare bundles of factors, might be difficult in general

Multicriteria decision-making

How do we handle multiple many variables?
e.g., what is
$u(\text{king safety}, \text{material advantage}, \text{control of the centre})$?
how do we compare lotteries over sets of states?
and sequences of states?

- We need to find ways to compare bundles of factors, might be difficult in general
- Search methods to avoid multi-criteria altogether: Monte Carlo Tree Search generates random endgames.

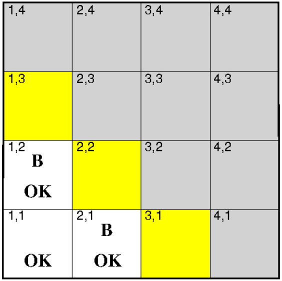
Multicriteria decision-making

How do we handle multiple many variables?
e.g., what is
$u(\text{king safety}, \text{material advantage}, \text{control of the centre})$?
how do we compare lotteries over sets of states?
and sequences of states?

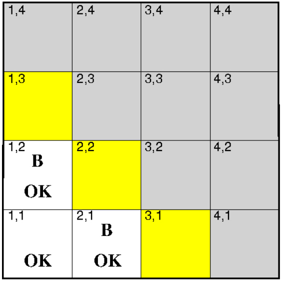
- We need to find ways to compare bundles of factors, might be difficult in general
- Search methods to avoid multi-criteria altogether: Monte Carlo Tree Search generates random endgames.

We assume there is a way of assigning a utility function to bundles of factors and therefore compare them.

Beliefs and expected utility



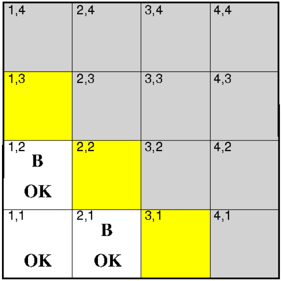
Beliefs and expected utility



Rewards:

- -1000 for dying
- 0 any other square

Beliefs and expected utility

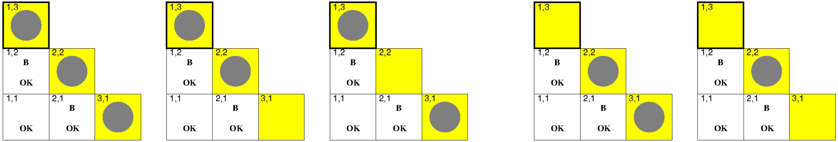


Rewards:

- -1000 for dying
- 0 any other square

What's the expected utility of going to $[3, 1], [2, 2], [1, 3]$?

Using conditional independence contd.



0.2 x 0.2 = 0.04 | 0.2 x 0.8 = 0.16 | 0.8 x 0.2 = 0.16 | 0.2 x 0.2 = 0.04 | 0.2 x 0.8 = 0.16

$$
\begin{aligned}
\mathbf{P}(P_{1,3}|known, b) &= \alpha' \langle 0.2(0.04+0.16+0.16),\ 0.8(0.04+0.16) \rangle \\
&\approx \langle 0.31, 0.69 \rangle
\end{aligned}
$$

$$
\mathbf{P}(P_{2,2}|known, b) \approx \langle 0.86, 0.14 \rangle
$$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$u(1,3) =$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$u(1,3) = u[0.31, -1000; 0.69, 0]$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

$$u(3,1) = u(1,3)$$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

$$u(3,1) = u(1,3)$$

$$u(2,2) =$$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

$$u(3,1) = u(1,3)$$

$$u(2,2) = u[0.86, -1000; 0.14, 0]$$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

$$u(3,1) = u(1,3)$$

$$u(2,2) = u[0.86, -1000; 0.14, 0] = -860$$

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

$$u(3,1) = u(1,3)$$

$$u(2,2) = u[0.86, -1000; 0.14, 0] = -860$$

Clearly going to $[2,2]$ from either $[1,2]$ or $[2,1]$ is irrational.

Beliefs and expected utility

The expected utility $u(1,3)$ of the action $(1,3)$ of going to $[1,3]$ from an explored adjacent square is:

$$u(1,3) = u[0.31, -1000; 0.69, 0] = -310$$

$$u(3,1) = u(1,3)$$

$$u(2,2) = u[0.86, -1000; 0.14, 0] = -860$$

Clearly going to $[2,2]$ from either $[1,2]$ or $[2,1]$ is irrational.
Either going to $[1,3]$ or $[3,1]$ is the rational choice.

Risky moves

Actuators

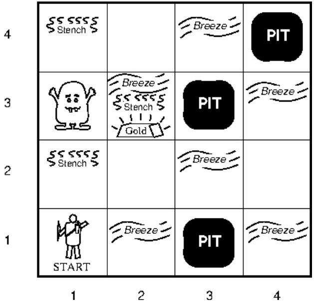
Sensors Breeze, Glitter, Smell

Actuators Turn L/R, Go, Grab, Release, Shoot, Climb

Rewards 1000 escaping with gold, -1000 dying, -10 using arrow, -1 walking

Environment

- Squares adjacent to Wumpus are smelly
- Squares adjacent to pit are breezy
- Glitter iff gold is in the same square
- Shooting kills Wumpus if you are facing it
- Shooting uses up the only arrow
- Grabbing picks up gold if in same square
- Releasing drops the gold in same square

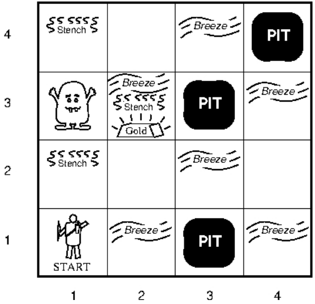


Actuators

Sensors	Breeze, Glitter, Smell
Actuators	Turn L/R, Go, Grab, Release, Shoot, Climb
Rewards	1000 escaping with gold, -1000 dying, -10 using arrow, -1 walking

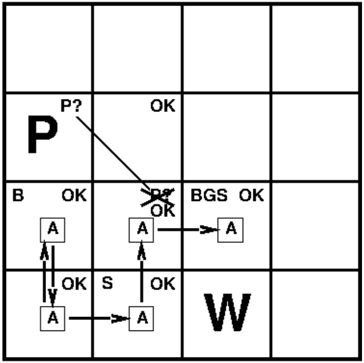
Environment

- Squares adjacent to Wumpus are smelly
- Squares adjacent to pit are breezy
- Glitter iff gold is in the same square
- Shooting kills Wumpus if you are facing it
- Shooting uses up the only arrow
- Grabbing picks up gold if in same square
- Releasing drops the gold in same square



Deterministic actions

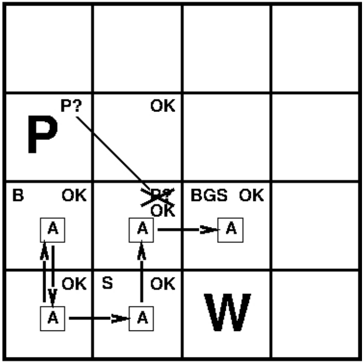
Actions in the Wumpus World are **deterministic**



Deterministic actions

Actions in the Wumpus World are **deterministic**

If I want to go from [2, 3] to [2, 2] I just go.

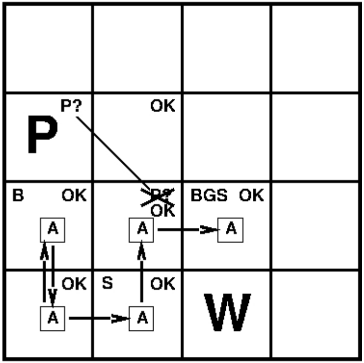


Deterministic actions

Actions in the Wumpus World are **deterministic**

If I want to go from $[2,3]$ to $[2,2]$ I just go.

$P([2,2] \mid [2,3],(2,2))$

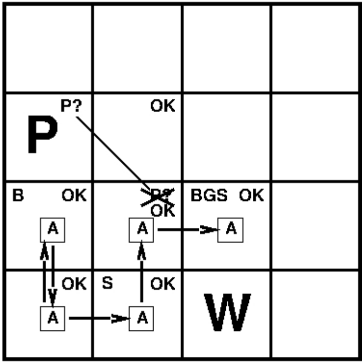


Deterministic actions

Actions in the Wumpus World are **deterministic**

If I want to go from [2, 3] to [2, 2] I just go.

$P([2,2] \mid [2,3],(2,2)) = 1$



Stochastic actions

Stochastic actions 'simulate' lack of control. The agent can try to go to the intended direction but much can work against:

- The environment
- The opponents
- The agent themselves!

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \dots p_n, s_n]$$

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$$

e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \dots p_n, s_n]$$

e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

- Goes to $[2, 2]$ with probability 0.5

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$$

e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

- Goes to $[2, 2]$ with probability 0.5
- Goes to $[3, 1]$ with probability 0.3

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \ldots p_n, s_n]$$

e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

- Goes to $[2, 2]$ with probability 0.5
- Goes to $[3, 1]$ with probability 0.3
- Goes back to $[1, 1]$ with probability 0.1

Stochastic actions

The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \dots p_n, s_n]$$

e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

- Goes to $[2, 2]$ with probability 0.5
- Goes to $[3, 1]$ with probability 0.3
- Goes back to $[1, 1]$ with probability 0.1
- Bumps their head to the wall and stays in $[2, 1]$ with prob. 0.1

Stochastic actions

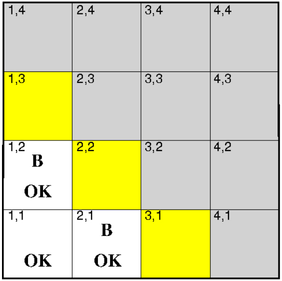
The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, s_1; p_2, s_2; \dots p_n, s_n]$$

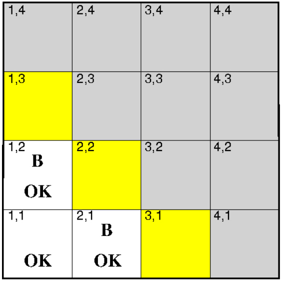
e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

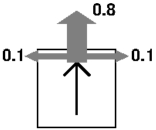
- Goes to $[2, 2]$ with probability 0.5
- Goes to $[3, 1]$ with probability 0.3
- Goes back to $[1, 1]$ with probability 0.1
- Bumps their head to the wall and stays in $[2, 1]$ with prob. 0.1
- Goes to any other square with probability 0

Beliefs, expected utility and stochastic actions

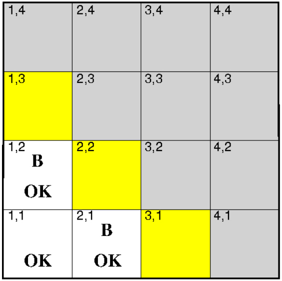


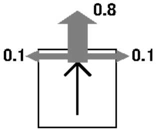
Beliefs, expected utility and stochastic actions





Beliefs, expected utility and stochastic actions

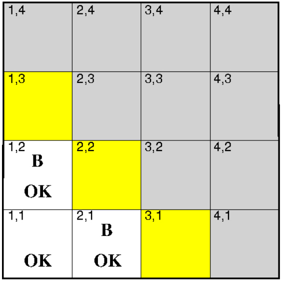


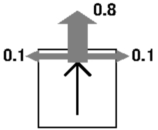


Rewards:

- -1000 for dying
- 0 any other square

Beliefs, expected utility and stochastic actions



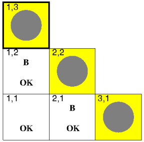


Rewards:

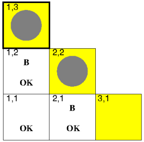
- -1000 for dying
- 0 any other square

What's the expected utility of going to $[3, 1]$, $[2, 2]$, $[1, 3]$?

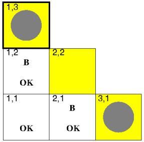
Beliefs, expected utility and stochastic actions



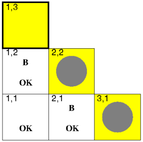
0.2 x 0.2 = 0.04



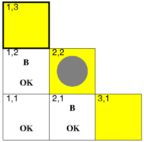
0.2 x 0.8 = 0.16



0.8 x 0.2 = 0.16



0.2 x 0.2 = 0.04



0.2 x 0.8 = 0.16

$$\begin{aligned}\mathbf{P}(P_{1,3}|known, b) &= \alpha' \langle 0.2(0.04+0.16+0.16),\ 0.8(0.04+0.16)\rangle \\ &\approx \langle 0.31, 0.69\rangle\end{aligned}$$

$$\mathbf{P}(P_{2,2}|known, b) \approx \langle 0.86, 0.14\rangle$$

Beliefs, expected utility and stochastic actions

Let $(s, a) = [p_1, L_1; p_2, L_2; \dots p_n, L_n]$ be the result of performing action a in state s

Beliefs, expected utility and stochastic actions

Let $(s, a) = [p_1, L_1; p_2, L_2; \ldots p_n, L_n]$ be the result of performing action a in state s, where each L_i is of the form $[q_1, s_{1i}; q_2, s_{2i}, \ldots, q_n, s_{ni}]$.

Beliefs, expected utility and stochastic actions

Let $(s, a) = [p_1, L_1; p_2, L_2; \ldots p_n, L_n]$ be the result of performing action a in state s, where each L_i is of the form $[q_1, s_{1i}; q_2, s_{2i}, \ldots, q_n, s_{ni}]$.
Then the utility of such action is given by:

$$u(s, a) = \sum_{p_i, L_i} p_i u(L_i) = \sum_{p_i} p_i \sum_{q_i, s_i} q_i u(s_i)$$

Beliefs, expected utility and stochastic actions

Let $(s, a) = [p_1, L_1; p_2, L_2; \ldots p_n, L_n]$ be the result of performing action a in state s, where each L_i is of the form $[q_1, s_{1i}; q_2, s_{2i}, \ldots, q_n, s_{ni}]$.
Then the utility of such action is given by:

$$u(s, a) = \sum_{p_i, L_i} p_i u(L_i) = \sum_{p_i} p_i \sum_{q_i, s_i} q_i u(s_i)$$

The expected utility of each outcome times the probability of reaching it.

Beliefs, expected utility and stochastic actions

Let $(s, a) = [p_1, L_1; p_2, L_2; \ldots p_n, L_n]$ be the result of performing action a in state s, where each L_i is of the form $[q_1, s_{1i}; q_2, s_{2i}, \ldots, q_n, s_{ni}]$.
Then the utility of such action is given by:

$$u(s, a) = \sum_{p_i, L_i} p_i u(L_i) = \sum_{p_i} p_i \sum_{q_i, s_i} q_i u(s_i)$$

The expected utility of each outcome times the probability of reaching it.

It is a lottery of lotteries!

Beliefs, expected utility and stochastic actions

$u(1,3) =$

Beliefs, expected utility and stochastic actions

$$u(1,3) = 0.8 \times u[0.31, -1000; 0.69, 0] + 0.1 \times u[1, 0] +$$
$$+ 0.1 \times u[0.86, -1000; 0.14, 0]$$

Beliefs, expected utility and stochastic actions

$$u(1,3) = 0.8 \times u[0.31, -1000; 0.69, 0] + 0.1 \times u[1, 0] +$$
$$+ 0.1 \times u[0.86, -1000; 0.14, 0] = 0.8 \times -310 + 0.1 \times -860 =$$

Beliefs, expected utility and stochastic actions

$u(1,3) = 0.8 \times u[0.31, -1000; 0.69, 0] + 0.1 \times u[1, 0] +$
$+ 0.1 \times u[0.86, -1000; 0.14, 0] = 0.8 \times -310 + 0.1 \times -860 =$
$-248 - 86$

Beliefs, expected utility and stochastic actions

$$u(1,3) = 0.8 \times u[0.31, -1000; 0.69, 0] + 0.1 \times u[1, 0] +$$
$$+0.1 \times u[0.86, -1000; 0.14, 0] = 0.8 \times -310 + 0.1 \times -860 =$$
$$-248 - 86 = -334$$

Beliefs, expected utility and stochastic actions

$$u(1,3) = 0.8 \times u[0.31, -1000; 0.69, 0] + 0.1 \times u[1, 0] +$$
$$+0.1 \times u[0.86, -1000; 0.14, 0] = 0.8 \times -310 + 0.1 \times -860 =$$
$$-248 - 86 = -334$$

Beliefs, expected utility and stochastic actions

$u(1,3) = 0.8 \times u[0.31, -1000; 0.69, 0] + 0.1 \times u[1, 0] +$
$+0.1 \times u[0.86, -1000; 0.14, 0] = 0.8 \times -310 + 0.1 \times -860 =$
$-248 - 86 = -334$

We can can get to $[2, 2]$ from two directions, but by symmetry it's the same.

Beliefs, expected utility and stochastic actions

$u(2,2) =$

Beliefs, expected utility and stochastic actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0] +$
$+0.1 \times u[1, 0]$

Beliefs, expected utility and stochastic actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0]+$
$+0.1 \times u[1, 0] = 0.8 \times -860 + 0.1 \times -310 =$

Beliefs, expected utility and stochastic actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0] +$
$+ 0.1 \times u[1, 0] = 0.8 \times -860 + 0.1 \times -310 = \; - 688 - 31$

Beliefs, expected utility and stochastic actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0]+$
$+0.1 \times u[1, 0] = 0.8 \times -860 + 0.1 \times -310 = -688 - 31 = -719$

Beliefs, expected utility and stochastic actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0] +$
$+0.1 \times u[1, 0] = 0.8 \times -860 + 0.1 \times -310 = -688 - 31 = -719$

$u(1,3) = u(3,1)$ (because of symmetry)

Beliefs, expected utility and stochastic actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0] +$
$+0.1 \times u[1,0] = 0.8 \times -860 + 0.1 \times -310 = -688 - 31 = -719$

$u(1,3) = u(3,1)$ (because of symmetry)

Going to $[2,2]$ is still the irrational choice, but not as bad.
The rational choice is either going to $[1,3]$ or $[3,1]$.

Beliefs versus knowledge

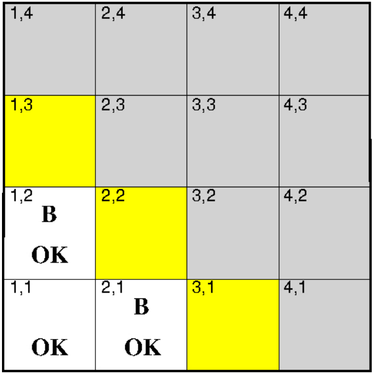
- A purely knowledge-based agent has nothing better to do than choosing at random. Which means $\frac{2}{3}u(1,3) + \frac{1}{3}u(2,2)$.
- A belief-based agent can improve the payoff using probabilistic reasoning and going for $u(1,3)$.

Beliefs versus knowledge

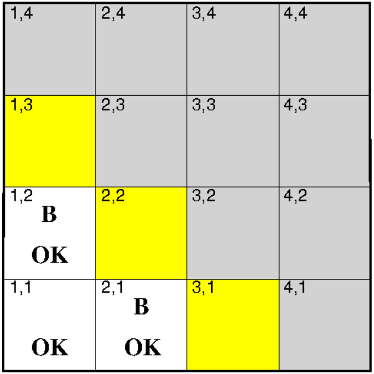
- A purely knowledge-based agent has nothing better to do than choosing at random. Which means $\frac{2}{3}u(1,3) + \frac{1}{3}u(2,2)$.
- A belief-based agent can improve the payoff using probabilistic reasoning and going for $u(1,3)$.

Obviously, the more chaotic the decision system the less the impact of reward difference.

New probability model

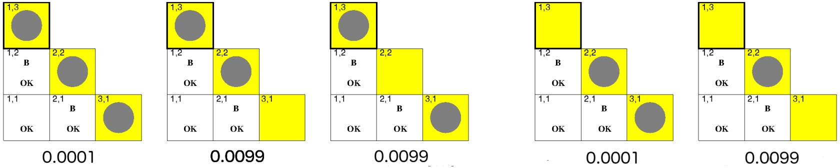


New probability model



Assume pits can be in a square with probability 0.01

The fringe



Obviously, we can use exactly the same reasoning!

Beliefs, expected utility and stochastic actions

- With deterministic agents, the chance of death is 0.9902 when trying to go to $[2, 2]$.
- With deterministic agents, it tends to 1 with the probability of pit in a square tending to 0;
- The more deterministic the agent, the higher the chance of death.
- Because the way rewards are defined, the expected utility follows the same pattern.

Beliefs, expected utility and stochastic actions

- With deterministic agents, the chance of death is 0.9902 when trying to go to $[2, 2]$.
- With deterministic agents, it tends to 1 with the probability of pit in a square tending to 0;
- The more deterministic the agent, the higher the chance of death.
- Because the way rewards are defined, the expected utility follows the same pattern.

Again, belief-based agents, perform much better than knowledge-based ones

Today's class

- Utility, lotteries and preferences
- Maximisation of expected utility
- Stochastic actions
- Knowledge-based versus belief-based agents

Coming next

- Time
- Risky plans
- What's the best "strategy" to follow?
- Estimating future gains: how patient should we be?

Coming next

- Time
- Risky plans
- What's the best "strategy" to follow?
- Estimating future gains: how patient should we be?