Decision-Making

Paolo Turrini

Department of Computing, Imperial College London

Introduction to Artificial Intelligence
2nd Part

Outline

- Lotteries (and how to win them)
- Risky moves
- maybe "Time" but I very much doubt it

Lotteries

(and how to win them)

The main reference

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

Rewards

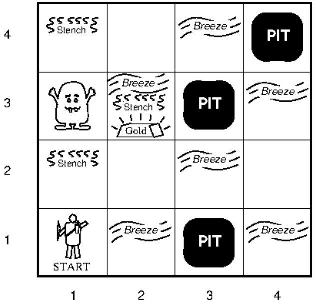
Sensors Breeze, Glitter, Smell

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



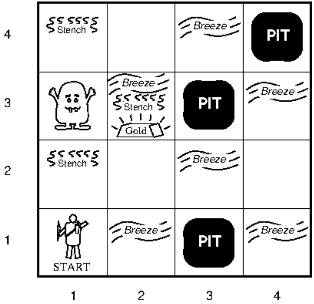
Rewards

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



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 squares in the Wumpus World.

State space

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

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

e.g., the squares in the Wumpus World.

- States can also take into account the inner state of the agent, e.g., the knowledge base KB;

State space

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

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

e.g., the squares in the Wumpus World.

- States can also take into account the inner state of the agent, e.g., the knowledge base KB;
- or the actions they have performed, e.g., climbing out of the cave with the gold.

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.

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.

(Paolo Turrini 2016)

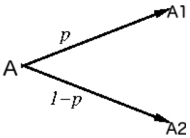
Lotteries

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 outcomes A_1 and A_2, and $p \in [0, 1]$

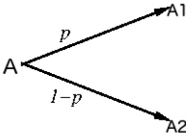
Lottery $A = [p, A_1;\ (1 - p), A_2]$



Lotteries

A lottery is a probability distribution over the set of states.
e.g., for outcomes A_1 and A_2, and $p \in [0, 1]$

Lottery $A = [p, A_1;\ (1 - p), A_2]$



L is the set of lotteries over S.

Simple Lotteries

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

Simple Lotteries

Observation: A state $s \in S$ can be seen as a lottery: where s is assigned probability 1 and all other states probability 0.

Simple Lotteries

Observation: A state $s \in S$ can be seen as a lottery: where s is assigned probability 1 and all other states probability 0.

e.g.,
$A = [1, A_1; 0, A_2; 0, A_3; \ldots]$

We get A_1 with probability 1, and the rest with probability 0.

Compound Lotteries

A lottery over the set of lotteries

Compound Lotteries

A lottery over the set of lotteries is itself a lottery.

Compound Lotteries

A lottery over the set of lotteries is itself a lottery.

$\mathbf{A} = [q_1, A; q_2, B; \ldots; q_n, C] =$

Compound Lotteries

A lottery over the set of lotteries is itself a lottery.

$$\mathbf{A} = [q_1, A; q_2, B; \ldots ; q_n, C] =$$
$$= [q_1, [p_1, A_1; p_2, A_2; \ldots p_n, A_n]; q_2, B; \ldots ; q_n, C] =$$

Compound Lotteries

A lottery over the set of lotteries is itself a lottery.

$$\mathbf{A} = [q_1, A; q_2, B; \ldots; q_n, C] =$$
$$= [q_1, [p_1, A_1; p_2, A_2; \ldots p_n, A_n]; q_2, B; \ldots; q_n, C] =$$
$$= [q_1 p_1, A_1; q_1 p_2, A_2; \ldots q_n p_n, A_n; q_2, B; \ldots; q_n, C] = \ldots$$

Compound Lotteries

A lottery over the set of lotteries is itself a lottery.

$$\mathbf{A} = [q_1, A; q_2, B; \ldots; q_n, C] =$$
$$= [q_1, [p_1, A_1; p_2, A_2; \ldots p_n, A_n]; q_2, B; \ldots; q_n, C] =$$
$$= [q_1 p_1, A_1; q_1 p_2, A_2; \ldots q_n p_n, A_n; q_2, B; \ldots; q_n, C] = \ldots$$

Compound lotteries can be reduced to simple lotteries

Expected Utility

Let $A = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be a lottery.

Expected Utility

Let $A = [p_1, A_1; p_2, A_2; \dots p_n, A_n]$ be a lottery.
The expected utility of A is

$$u(A) = \sum_{p_i, A_i} p_i \times u(A_i)$$

Expected Utility

Let $A = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be a lottery.
The expected utility of A is

$$u(A) = \sum_{p_i, A_i} p_i \times u(A_i)$$

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

Expected Utility

Let $A = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be a lottery.
The expected utility of A is

$$u(A) = \sum_{p_i, A_i} p_i \times u(A_i)$$

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 $A = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be a lottery.
The expected utility of A is

$$u(A) = \sum_{p_i, A_i} p_i \times u(A_i)$$

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 $A = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be a lottery.
The expected utility of A is

$$u(A) = \sum_{p_i, A_i} p_i \times u(A_i)$$

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

Tverski and Kahneman's Prospect Theory:

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

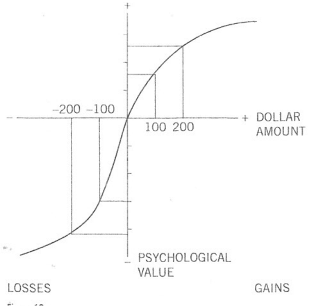


Figure: Typical empirical data

Humans and Expected Utility

Tverski and Kahneman's Prospect Theory:

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

Warning! controversial statement:

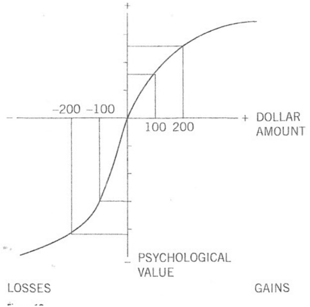


Figure: Typical empirical data

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.

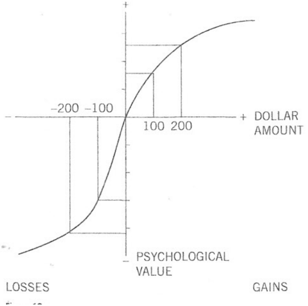


Figure: Typical empirical data

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.

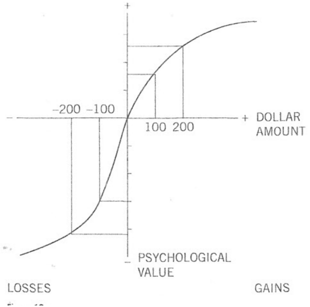


Figure: Typical empirical data

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 lotter 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 lotter 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 states and let $p, q \in [0, 1]$.

Rational preferences

Let A, B, C be three states and let $p, q \in [0, 1]$.
A preference relation $\succeq$ makes sense if it satisfies the following constraints

Rational preferences

Let A, B, C be three states and let $p, q \in [0, 1]$.
A preference relation $\succeq$ makes sense if it satisfies the following constraints

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

Rational preferences

Let A, B, C be three states and let $p, q \in [0, 1]$.
A preference relation $\succeq$ makes sense if it satisfies the following constraints

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

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

Rational preferences

Let A, B, C be three states and let $p, q \in [0, 1]$.
A preference relation $\succeq$ makes sense 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 $A \succ B \succ C \Rightarrow \exists p \; [p, A; \; 1 - p, C] \sim B$

Rational preferences

Let A, B, C be three states and let $p, q \in [0, 1]$.
A preference relation $\succeq$ makes sense 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 $A \sim B \Rightarrow [p, A;\ 1 - p, C] \sim [p, B; 1 - p, C]$

Rational preferences

Let A, B, C be three states and let $p, q \in [0, 1]$.
A preference relation $\succeq$ makes sense 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 $A \succ B \Rightarrow (p \geq q \Leftrightarrow [p, A;\ 1 - p, B] \succsim [q, A;\ 1 - q, B])$

Rational preferences contd.

Violating the constraints leads to self-evident irrationality.

Rational preferences contd.

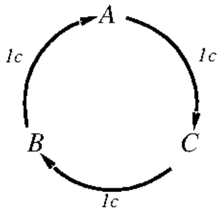
Violating the constraints leads to self-evident irrationality.

Take transitivity.

Rational preferences contd.

Violating the constraints leads to self-evident irrationality.

Take transitivity.

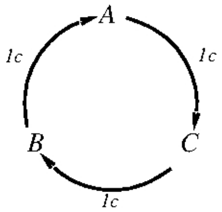


Rational preferences contd.

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



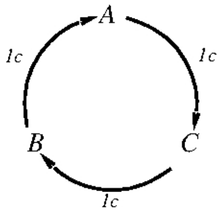
Rational preferences contd.

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 contd.

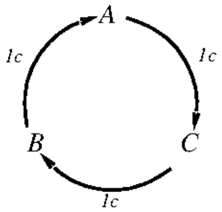
Violating the constraints leads to self-evident irrationality.

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 $\succsim$ makes sense if and only if there exists a real-valued function u such that:

Representation Theorem

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

A preference relation $\succsim$ makes sense if and only if there exists a real-valued function u such that:

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

Representation Theorem

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

A preference relation $\succsim$ makes sense if and only if there exists a real-valued function u such that:

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

Representation Theorem

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

A preference relation $\succsim$ makes sense if and only if there exists a real-valued function u such that:

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

$[\Leftarrow]$

Representation Theorem

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

A preference relation $\succsim$ makes sense if and only if there exists a real-valued function u such that:

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

[$\Leftarrow$] 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 $\succsim$ makes sense if and only if there exists a real-valued function u such that:

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

[⇐] 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 $\succsim$ makes sense if and only if there exists a real-valued function u such that:

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

[$\Leftarrow$] 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 functions.

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 utility function!

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 factor 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 factor 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

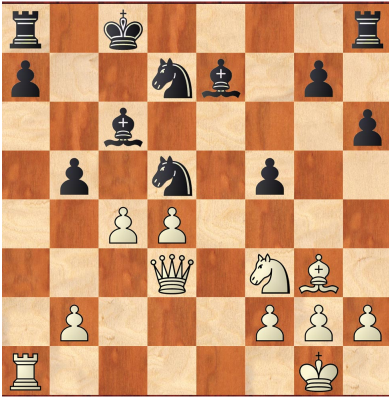


Figure: Deep Blue- Kasparov 1996, Final Game. Material favours Black but the position is hopeless

Multicriteria decision-making

How can we handle utility functions of many variables $X_1 \ldots X_n$?

Multicriteria decision-making

How can we handle utility functions of many variables $X_1 \ldots X_n$?
e.g., what is
$U(\text{king safety}, \text{material advantage}, \text{control of the centre})$?

Multicriteria decision-making

How can we handle utility functions of many variables $X_1 \dots X_n$?
e.g., what is
$U(\text{king safety}, \text{material advantage}, \text{control of the centre})$?

- We need to find ways to compare bundles of factors, but might be difficult in general (strict dominance, stochastic dominance).

Multicriteria decision-making

How can we handle utility functions of many variables $X_1 \dots X_n$?
e.g., what is
$U(\text{king safety}, \text{material advantage}, \text{control of the centre})$?

- We need to find ways to compare bundles of factors, but might be difficult in general (strict dominance, stochastic dominance).
- Search methods to avoid multicriteria altogether: Monte Carlo Tree Search generates random endgames.

Multicriteria decision-making

How can we handle utility functions of many variables $X_1 \dots X_n$?
e.g., what is
$U(\text{king safety}, \text{material advantage}, \text{control of the centre})$?

- We need to find ways to compare bundles of factors, but might be difficult in general (strict dominance, stochastic dominance).
- Search methods to avoid multicriteria 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.

Rationality and expected utility

Robert J. Aumann
Nobel Prize Winner
Economics

"A person's behavior is rational if it is in his best interests, given his information"

Rationality and expected utility

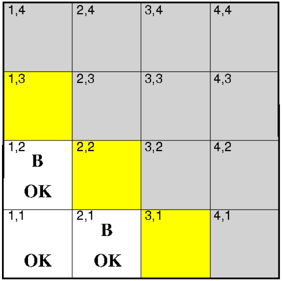
Robert J. Aumann
Nobel Prize Winner

Economics

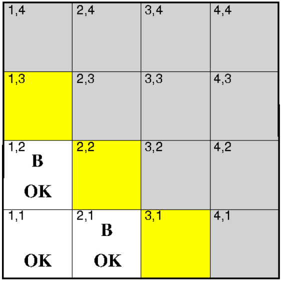
"A person's behavior is rational if it is in his best interests, given his information"

Choose an action that maximises the expected utility

Beliefs and Expected Utility



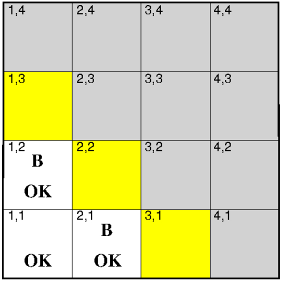
Beliefs and Expected Utility



Rewards:

- -1000 for dying
- 0 any other square

Beliefs and Expected Utility

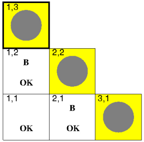


Rewards:

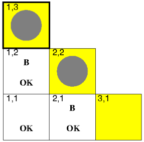
- -1000 for dying
- 0 any other square

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

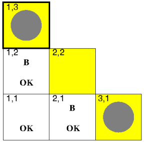
Using conditional independence contd.



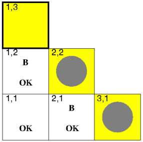
0.2 x 0.2 = 0.04



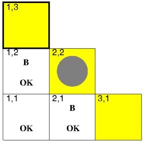
0.2 x 0.8 = 0.16



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

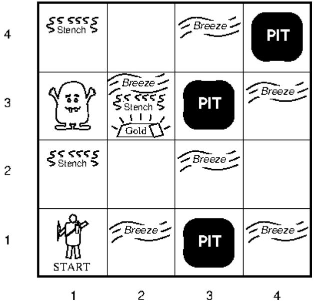
Sensors Breeze, Glitter, Smell

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

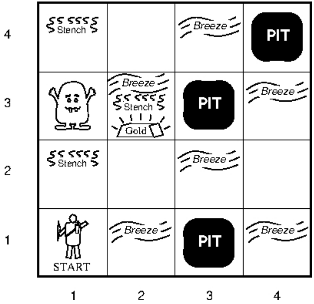
Sensors Breeze, Glitter, Smell

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

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

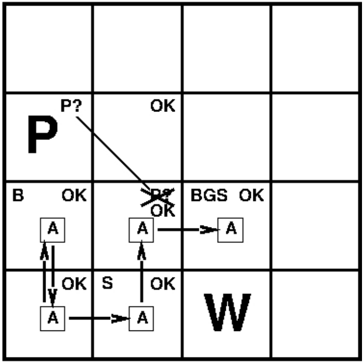
Environment

- 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

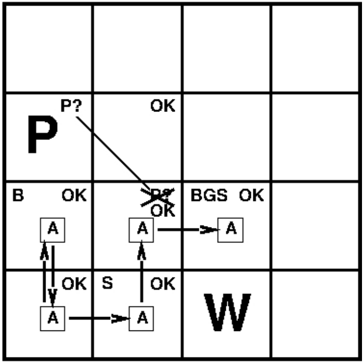
Actions in the Wumpus World are **deterministic**



Deterministic actions

Actions in the Wumpus World are **deterministic**

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

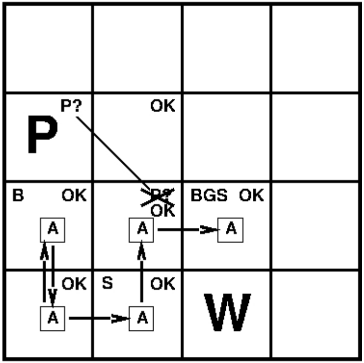


Deterministic actions

Actions in the Wumpus World are **deterministic**

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

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

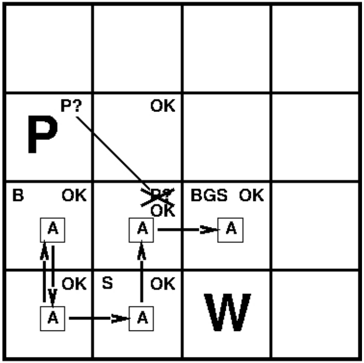


Deterministic actions

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

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.

$$(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$$

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.

$$(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$$

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.

$$(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$$

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.

$$(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$$

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.

$$(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$$

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.

$$(s, a) = [p_1, A_1; p_2, A_2; \dots p_n, A_n]$$

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 his head on the wall and stays in $[2, 1]$ with prob. 0.1

Stochastic actions

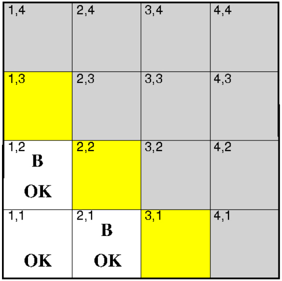
The result of performing a in state s is a lottery over S, i.e., probability distribution over the set of all possible states.

$$(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$$

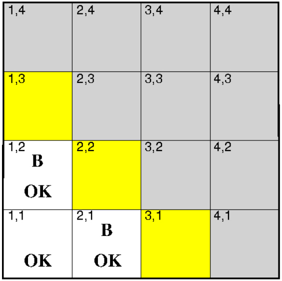
e.g., the agent decides to go from $[2, 1]$ to $[2, 2]$ but:

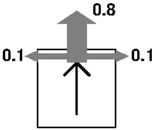
- Goes to $[2, 2]$ with probability 0.5
- Goes to $[3, 1]$ with probability 0.3
- Goes back to $[1, 1]$ with probability 0.1
- Bumps his head on the wall and stays in $[2, 1]$ with prob. 0.1
- Goes to any other square with probability 0

Beliefs, Expected Utility and Stochastic Actions

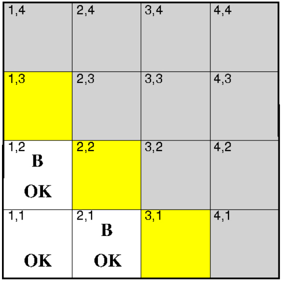


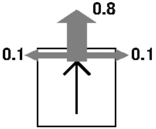
Beliefs, Expected Utility and Stochastic Actions





Beliefs, Expected Utility and Stochastic Actions

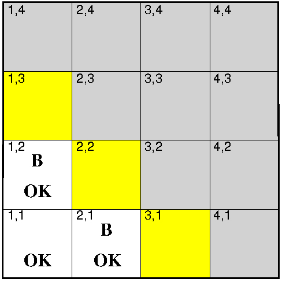


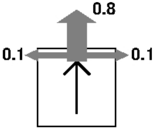


Rewards:

- -1000 for dying
- 0 any other square

Beliefs, Expected Utility and Stochastic Actions



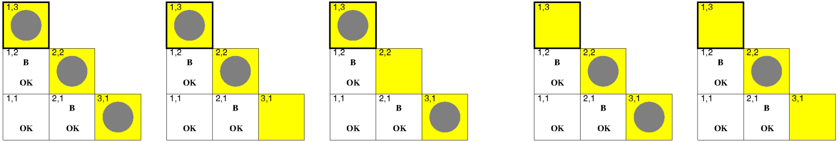


Rewards:

- -1000 for dying
- 0 any other square

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

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

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

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

Beliefs, Expected Utility and Stochastic Actions

Let $(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be the result of performing action a in state s

Beliefs, Expected Utility and Stochastic Actions

Let $(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be the result of performing action a in state s, where each A_i is of the form $[q_1, A_{1i}; q_2, A_{2i}, \ldots, q_n, A_{ni}]$.

Beliefs, Expected Utility and Stochastic Actions

Let $(s, a) = [p_1, A_1; p_2, A_2; \dots p_n, A_n]$ be the result of performing action a in state s, where each A_i is of the form $[q_1, A_{1i}; q_2, A_{2i}, \dots, q_n, A_{ni}]$.
Then the utility of such action is given be:

$$u(s, a) = \sum_{p_i, A_i} p_i \times u(A_i)$$

Beliefs, Expected Utility and Stochastic Actions

Let $(s, a) = [p_1, A_1; p_2, A_2; \ldots p_n, A_n]$ be the result of performing action a in state s, where each A_i is of the form $[q_1, A_{1i}; q_2, A_{2i}, \ldots, q_n, A_{ni}]$.
Then the utility of such action is given be:

$$u(s, a) = \sum_{p_i, A_i} p_i \times u(A_i)$$

The expected utility of each outcome, assuming we have reached it, times the probability of actually reaching it.

Beliefs, Expected Utility and Stochastic Actions

Let $(s, a) = [p_1, A_1; p_2, A_2; \dots p_n, A_n]$ be the result of performing action a in state s, where each A_i is of the form $[q_1, A_{1i}; q_2, A_{2i}, \dots, q_n, A_{ni}]$.
Then the utility of such action is given be:

$$u(s, a) = \sum_{p_i, A_i} p_i \times u(A_i)$$

The expected utility of each outcome, assuming we have reached it, times the probability of actually 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

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0]+$
$+0.1 \times u[1,0] = 0.8 \times -860 + 0.1 \times -310 = -688 - 31 = -729$

Beliefs, Expected Utility and Stochastic Actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0]+$
$+0.1 \times u[1, 0] = 0.8 \times -860 + 0.1 \times -310 = -688 - 31 = -729$

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

Beliefs, Expected Utility and Stochastic Actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0]+$
$+0.1 \times u[1,0] = 0.8 \times -860 + 0.1 \times -310 = \ -688 - 31 = -729$

$u(1,3) = u(3,1)$ (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, Expected Utility and Stochastic Actions

$u(2,2) =$
$0.8 \times u[0.86, -1000; 0.14, 0] + 0.1 \times u[0.31, -1000; 0.69, 0] +$
$+0.1 \times u[1,0] = 0.8 \times -860 + 0.1 \times -310 = -688 - 31 = -729$

$u(1,3) = u(3,1)$ (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 $[3,1]$.

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

Summary

- Utility, lotteries and preferences
- Maximisation of expected utility
- Stochastic actions

What's next

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

What's next

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