Agent-Based Systems

Paolo Turrini

www.dcs.warwick.ac.uk/~pturrini p.turrini@warwick.ac.uk

Risks and Decisions

Expected utility

Today

- Probabilities with value attached
  - How to compare options: goals vs utilities
  - Deciding means gambling
- Risky moves and rationality

The book

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

If you snooze you lose. Or do you?

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. Stagecoach busses run past me full of people
2. my phone and iPad die together
3. my mum forgets to call me

If you snooze you lose. Or do you?

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. Or do you?

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
Decision theory = utility theory + probability theory

Rewards

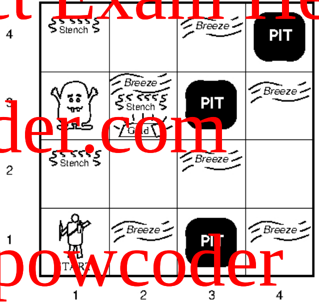
Sensors Breeze, Glitter, Smell

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

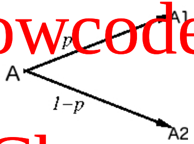
$$W = \{w_1, \dots, w_n\}$$

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

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

e.g., for states w_1 and w_2, and $p \in [0, 1]$



Lottery $L_1 = [p, w_1; (1-p), w_2]$

L is the set of lotteries over W.

Observation: A state $w \in W$ can be seen as a lottery :

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

e.g.,
$L_1 = [1, w_1; 0, w_2; \ldots 0, w_r]$

We get w_1 with probability 1, and the rest with probability 0.

Compound lotteries

Consider now the set L of lotteries over W.

Observation: A lottery over L is a lottery over W.

Compound lotteries

Consider now the set L of lotteries over W.

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

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

Compound lotteries

Consider now the set L of lotteries over W.

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

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_m, L_m]$$
$$= [q_1, [p_1, w_1; p_2, w_2; \ldots; p_n, w_n]; q_2, L_2; \ldots; q_m, L_m]$$

Compound lotteries

Consider now the set L of lotteries over W.

Observation: A lottery over $\mathbf{L}$ is a lottery over W:

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_m, L_m]$$
$$= [q_1, [p_1, w_1; p_2, w_2; \ldots; p_n, w_n]; q_2, L_2; \ldots; q_m, L_m]$$
$$= [q_1 p_1, w_1; q_1 p_2, w_2; \ldots; q_1 p_n, w_n; q_2, L_2; \ldots; q_m, L_m]$$

Compound lotteries

Consider now the set L of lotteries over W.

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

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \dots; q_m, L_m]$$
$$= [q_1, [p_1, w_1; p_2, w_2; \dots p_n, w_n]; q_2, L_2; \dots; q_m, L_m]$$
$$= [q_1 p_1, w_1; q_1 p_2, w_2; \dots q_1 p_n, w_n; q_2, L_2; \dots; q_m, L_m]$$
$$= [q_1, L_1; q_2 [r_1, w_1; r_2, w_2; \dots r_n, w_n]; \dots; q_m, L_m]$$

Compound lotteries

Consider now the set L of lotteries over W.

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

$$\mathbf{L}_1 = [q_1, L_1; q_2, L_2; \ldots; q_m, L_m]$$
$$= [q_1, [p_1, w_1; p_2, w_2; \ldots p_n, w_n]; q_2, L_2; \ldots; q_m, L_m]$$
$$= [q_1 p_1, w_1; q_1 p_2, w_2; \ldots q_1 p_n, w_n; q_2, L_2; \ldots; q_m, L_m]$$
$$= [q_1, L_1; q_2 [r_1, w_1; r_2, w_2; \ldots r_n, w_n]; \ldots; q_m, L_m]$$
$$= [q_1, L_1; [q_2 r_1, w_1; q_2 r_2, w_2; \ldots q_n r_n, w_n]; \ldots; q_m, L_m]$$

Compound lotteries

Consider now the set L of lotteries over W.

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

$$
\begin{aligned}
\mathbf{L}_1 &= [q_1, L_1; q_2, L_2; \dots; q_m, L_m] \\
&= [q_1, [p_1, w_1; p_2, w_2; \dots p_n, w_n]; q_2, L_2; \dots; q_m, L_m] \\
&= [q_1 p_1, w_1; q_1 p_2, w_2; \dots q_1 p_n, w_n; q_2, L_2; \dots; q_m, L_m] \\
&= [q_1, L_1; q_2[r_1, w_1; r_2, w_2; \dots r_n, w_n]; \dots; q_m, L_m] \\
&= [q_1, L_1; [q_2 r_1, w_1; q_2 r_2, w_2; \dots q_n r_n, w_n]; \dots; q_m, L_m] \\
&= [[(q_1 p_1 + q_2 r_1), w_1; (q_1 p_2 + q_2 r_2), w_2; \dots (q_1 p_n + q_2 r_n), w_n; \dots; q_m, L_m]
\end{aligned}
$$

Compound lotteries

Consider now the set L of lotteries over W.

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

$$
\begin{aligned}
\mathbf{L}_1 &= [q_1, L_1; q_2, L_2; \ldots; q_m, L_m] \\
&= [q_1, [p_1, w_1; p_2, w_2; \ldots p_n, w_n]; q_2, L_2; \ldots; q_m, L_m] \\
&= [q_1 p_1, w_1; q_1 p_2, w_2; \ldots q_1 p_n, w_n; q_2, L_2; \ldots; q_m, L_m] \\
&= [q_1, L_1; q_2[r_1, w_1; r_2, w_2; \ldots r_n, w_n]; \ldots; q_m, L_m] \\
&= [q_1, L_1; [q_2 r_1, w_1; q_2 r_2, w_2; \ldots q_n r_n, w_n]; \ldots; q_m, L_m] \\
&= [[(q_1 p_1 + q_2 r_1), w_1; (q_1 p_2 + q_2 r_2), w_2; \ldots (q_1 p_n + q_2 r_n), w_n]; \ldots; q_m, L_m] \\
&= \ldots
\end{aligned}
$$

Compound lotteries

Consider now the set L of lotteries over W.

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

$$
\begin{aligned}
\mathbf{L}_1 &= [q_1, L_1; q_2, L_2; \ldots; q_m, L_m] \\
&= [q_1, [p_1, w_1; p_2, w_2; \ldots p_n, w_n]; q_2, L_2; \ldots; q_m, L_m] \\
&= [q_1 p_1, w_1; q_1 p_2, w_2; \ldots q_1 p_n, w_n; q_2, L_2; \ldots; q_m, L_m] \\
&= [q_1, L_1; q_2[r_1, w_1; r_2, w_2; \ldots r_n, w_n]; \ldots; q_m, L_m] \\
&= [q_1, L_1; [q_2 r_1, w_1; q_2 r_2, w_2; \ldots q_n r_n, w_n]; \ldots; q_m, L_m] \\
&= [[(q_1 p_1 + q_2 r_1), w_1; (q_1 p_2 + q_2 r_2), w_2; \ldots (q_1 p_n + q_2 r_n), w_n]; \ldots; q_m, L_m] \\
&= \ldots
\end{aligned}
$$

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?

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.

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

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

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

'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, which is better than C.
But if you give me the right mix of A and C then this is 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

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

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

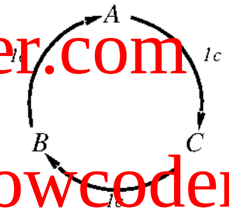
Violating the constraints leads to self-evident irrationality.

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 $\succeq$ satisfies the five previous principles if and only if there exists a real-valued function $u : L \to \mathbb{R}$ such that:

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, w_1; \ldots ; p_n, w_n]) = \sum_i p_i u(w_i)$

Representation Theorem

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

A preference relation $\succeq$ satisfies the five previous principles if and only if there exists a real-valued function $u : L \to \mathbb{R}$ such that:

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, w_1; \ldots; p_n, w_n]) = \sum_i p_i u(w_i)$

Proof idea:

Representation Theorem

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

A preference relation $\succeq$ satisfies the five previous principles if and only if there exists a real-valued function $u : L \to \mathbb{R}$ such that:

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, w_1; \ldots ; p_n, w_n]) = \sum_i p_i u(w_i)$

Proof idea:

[$\Leftarrow$] 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 $\succeq$ satisfies the five previous principles if and only if there exists a real-valued function $u : L \to \mathbb{R}$ such that:

- $u(A) \geq u(B) \quad \Leftrightarrow \quad A \succeq B$
- $u([p_1, w_1; \ldots ; p_n, w_n]) = \sum_i p_i u(w_i)$

Proof idea:

[$\Leftarrow$] 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

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

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

associating a real number to each state

A utility function is a function

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

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.

Let $L = [p_1, w_1; p_2, w_2; \ldots p_n, w_n]$ be a lottery.

The expected utility of L is

$$u(L) = \sum_{p_i, w_i} p_i \times u(w_i)$$

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

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

What would you do?

‘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. Not the same for everyone!

Humans and expected utility

Tverski and Kahneman's Prospect Theory:

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

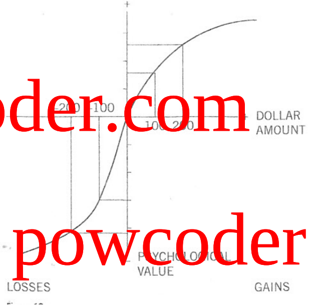


Figure: Typical empirical data

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.

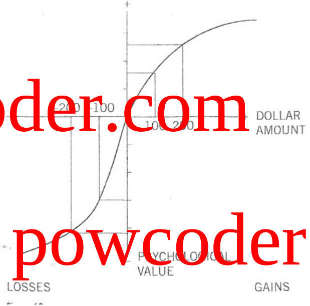


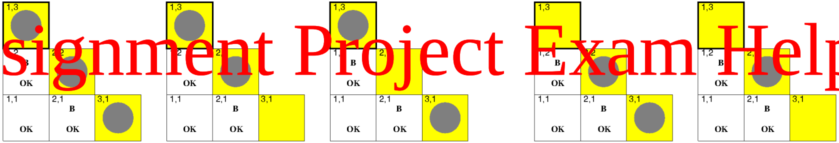
Figure: Typical empirical data

Beliefs and expected utility

Rewards:

- -1000 for dying
- 0 any other square

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



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

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

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

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

$u(1,3) =$

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

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

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:

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

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

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:

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

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

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

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

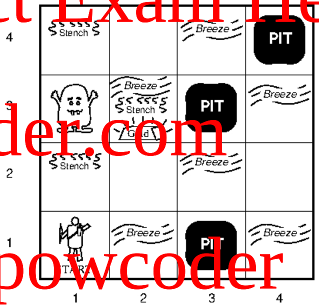
Actuators

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



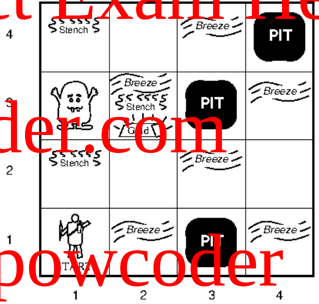
Actuators

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



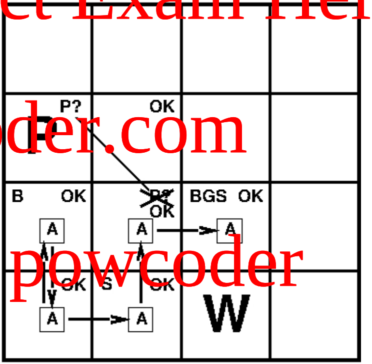
Deterministic actions

Actions in the Wumpus World are **deterministic**

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

The probability of successfully executing action $(2,2)$ at $[2,3]$:

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



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!

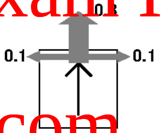
Stochastic actions

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

$$(w, a) = [p_1, w_1; p_2, w_2; \ldots p_n, w_n]$$

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

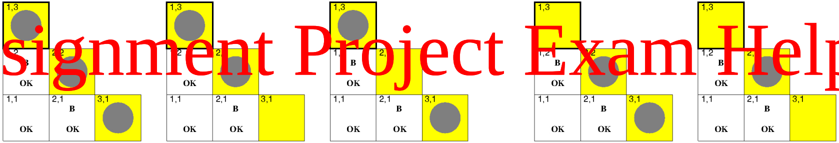
- Goes to $[2, 2]$ with probability 0.5
- Goes to $[3, 1]$ with probability 0.3
- Goes back to $[1, 1]$ with probability 0.1
- Bumps their head to the wall and stays in $[2, 1]$ with prob. 0.1
- Goes to any other square with probability 0



Rewards:

- -1000 for dying
- 0 any other square

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



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

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

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

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

Then the utility of such action is given by:

$$u(w, a) = \sum_{p_i, L_i} p_i u(L_i) = \sum_{p_i} p_i \sum_{q_i, w_i} q_i u(w_i)$$

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

It is a lottery of lotteries!

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

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

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

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

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

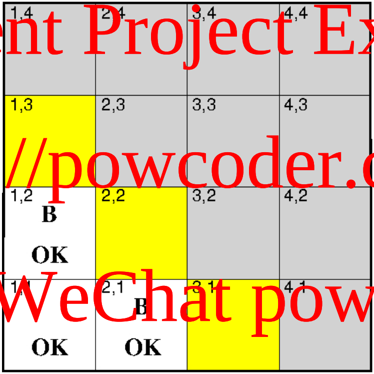
Going to $[2,2]$ is still the irrational choice, but not as bad.
The rational choice is either going to $[1,3]$ or $[3,1]$

- A purely knowledge-based agent has nothing better to do than choosing at random. Which means $\frac{2}{3}u(1,3) + \frac{1}{3}u(2,2)$.

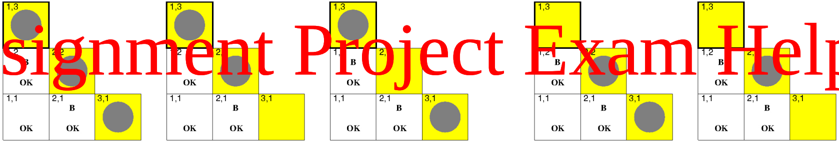
- A belief-based agent can improve the payoff using probabilistic reasoning and going for $u(1,3)$.

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

New probability model



Assume pits can be in a square with probability 0.01



0.0001 0.0099 0.0099 0.0001 0.0099

Obviously, we can use exactly the same reasoning!

- 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.

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

- Game Theory = Decision Theory with many agents