Sequential decision making

Frederik Mallmann-Trenn
6CCS3AIN

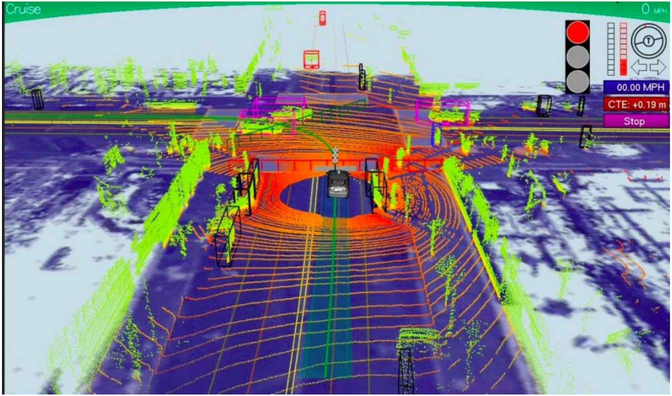
What to do?

(40 Acres and a Mule Filmworks/Universal Pictures)

What to do?

(mystorybook.com/books/42485)

What to do?



(Sebastian Thrun & Chris Urmson/Google)

Sequential decision making?

(mystorybook.com/books/42485)

- Ultimately, we are interested in sequential decision making
- One decision leads to another.
- Each decision depends on the ones before, and affects the ones after.

How to decide what to do

- Start simple.
- Single decision.
- Consider being offered a bet in which you pay £2 if an odd number is rolled on a die, and win £3 if an even number appears.
- Is this a good bet?

How to decide what to do

- Consider being offered a bet in which you pay £2 if an odd number is rolled on a die, and win £3 if an even number appears.
- Is this a good bet?
- To analyse this, we need the expected value of the bet.

How to decide what to do

- We do this in terms of a random variable, which we will call X.
- X can take two values:

  3 if the die rolls odd

  -2 if the die rolls even
- And we can also calculate the probability of these two values

  $P(X = 3) = 0.5$

  $P(X = -2) = 0.5$

How to decide what to do

- The expected value is then the weighted sum of the values, where the weights are the probabilities.
- Formally the expected value of X is defined by:

$$E[X] = \sum_k k \cdot P(X = k)$$

where the summation is over all values of k for which $P(X = k) \neq 0$.

How to decide what to do

- Here the expected value is:

$$E[X] = 3 \cdot 0.5 + (-2) \cdot 0.5$$

- Thus the expected value of X, $E[X]$, is £0.5, and we take this to be the value of the bet.

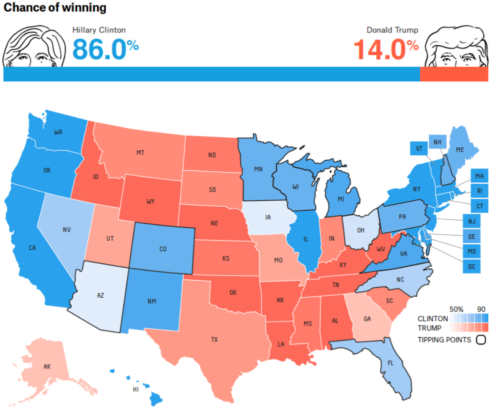
How to decide what to do

- Do you take the bet?
- Compare that $£0.5$ with not taking the bet.
- Not taking the bet has (expected) value £0

How to decide what to do

- £0.5 is not the value you will get.
- You can think of it as the long run average if you were offered the bet many times.
- Again, even after a large number of rounds you won't get that value (there will be some noise)

Sometimes the unlikely event can occur … it doesn’t mean the prediction was bad



Electoral votes

Hillary Clinton	338.9
Donald Trump	198.2
Evan McMullin	0.8
Gary Johnson	0.1

Popular vote

Hillary Clinton	49.6%
Donald Trump	43.3%
Gary Johnson	5.6%
Other	1.5%

(fivethirtyeight.com)

Example

- Pacman is at a T-junction
- Based on their knowledge, estimates that if they go Left:
  - Probability of 0.3 of getting a payoff of 10
  - Probability of 0.2 of getting a payoff of 1
  - With the remaining probability a payoff of -5
- What is the expected value of Left?

Example

- Pacman is at a T-junction
- Based on their knowledge, estimates that if they go Left:
  - Probability of 0.3 of getting a payoff of 10
  - Probability of 0.2 of getting a payoff of 1
  - With the remaining probability a payoff of -5
- What is the expected value of Left?
- $$\mathbb{E}[X] = 0.3 \cdot 10 + 0.2 \cdot 1 + (1 - 0.3 - 0.2) \cdot (-5) = 3 + 0.2 - 2.5 = 0.7$$

How to decide what to do

- Another bet: you get £1 if a 2 or a 3 is rolled, £5 if a six is rolled, and pay 3 otherwise.
- What's the expected value?

How to decide what to do

- Another bet: you get £1 if a 2 or a 3 is rolled, £5 if a six is rolled, and pay 3 otherwise.
- What's the expected value?
- $$\mathbb{E}[X] = \frac{2}{6} \cdot 1 + \frac{1}{6} \cdot 5 + \frac{3}{6} \cdot (-3) = -\frac{1}{3}$$

How to decide what to do

- What happens if you repeat this bet 10 times: you get £1 if a 2 or a 3 is rolled, £5 if a six is rolled, and pay 3 otherwise.
- What's the expected value now? (i.e., after all 10 games)

How to decide what to do

- Let $X_i, i \in \{1, 2, \ldots, 10\}$ and $X = \sum_{i=1}^{10} X_i$
- The expected value here is:

$$\mathbb{E}[X_i] = \frac{2}{6} \cdot 1 + \frac{1}{6} \cdot 5 + \frac{3}{6} \cdot (-3) = -\frac{1}{3}$$

- Thus, by linearity of expectation (i.e., $E[\alpha Y + Z] = \alpha E[Y] + E[Z]$, for all Y, Z and α),

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{10} X_i\right] = \sum_{i=1}^{10} E[X_i] = 10 \cdot E[X_i] = -10 \cdot \frac{1}{3} = -\frac{10}{3}$$

How an agent might decide what to do

- Consider an agent with a set of possible actions A.
- Each $a \in A$ has a set of possible outcomes s_a.
- Which action should the agent pick?

How an agent might decide what to do

- The action a^* which a rational agent should choose is that which maximises the agent's utility.
- In other words the agent should pick:

$$a^* = arg \max_{a \in A} u(s_a),$$

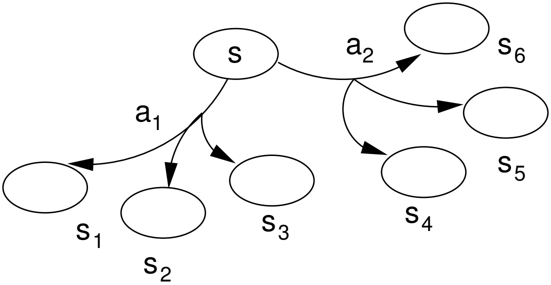
  - where s_a is the state obtained by choosing action a and
  - $u(s_a)$ is the utility of that state
- The problem is that in any realistic situation, the resulting state is probabilistic.
- Instead we have to calculate the expected utility of each action and make the choice on the basis of that.

How an agent might decide what to do

- In other words, for each action a with a set of outcomes s_a, the agent should calculate:

$$E[u(a)] = \sum_{s' \in s_a} u(s'). \Pr(s_a = s')$$

and pick the best. Here: decide between $E[u(a_1)]$ and $E[u(a_2)]$



How an agent might decide what to do

- That is it picks the action that has the greatest expected utility.
  - The right thing to do.

(40 Acres and a Mule Filmworks/Universal Pictures)

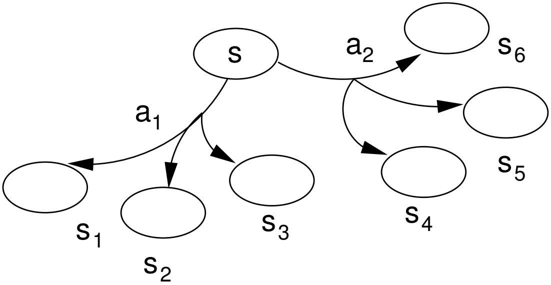
- Here “rational” means “rational in the sense of maximising expected utility”.

Example

- Pacman is at a T-junction
- Based on their knowledge, estimates that if they go Left:
  - Probability of 0.3 of getting a payoff of 10
  - Probability of 0.2 of getting a payoff of 1
  - Probability of 0.5 of getting a payoff of -5
- If they go Right:
  - Probability of 0.5 of getting a payoff of -5
  - Probability of 0.4 of getting a payoff of 3
  - Probability of 0.1 of getting a payoff of 15
- Should they choose Left or Right (MEU)?

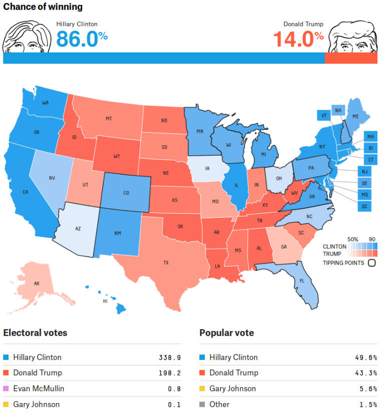
Stochastic

- Note that we are dealing with stochastic actions here.



- A given action has several possible outcomes.
- We don't know, in advance, which one will happen.

Stochastic



(fivethirtyeight.com)

- A lot like life.

Limitations of our notion of "rational"

- Consider the following game. Let's say your monthly income is m.
- W.p. $2/3$ I double your income every month
- With the remaining probability you have to give me your monthly income every month
- Expected value if playing $E[Income] = 2m\frac{2}{3} + 0\frac{1}{3} = \frac{4}{3}m$
- Expected value if not playing $E[Income] = m$.
- Would you play?