Sequential Decisions

Mark van der Wilk

Department of Computing
Imperial College London

@markvanderwilk
m.vdwilk@imperial.ac.uk

February 13, 2023

Sequential Decision Making

- Last time: Principle for taking an action based on beliefs.

Sequential Decision Making

- Last time: Principle for taking an action based on beliefs.
- This time: How do we plan for decisions made after each other?

Sequential Decision Making

- Last time: Principle for taking an action based on beliefs.
- This time: How do we plan for decisions made after each other?
- Key idea: At each point, apply MEU, given all the knowledge you have at that point

Sequential Decision Making

- Last time: Principle for taking an action based on beliefs.
- This time: How do we plan for decisions made after each other?
- Key idea: At each point, apply MEU, given all the knowledge you have at that point
- Consequence: After making an earlier decision, you can **learn**. This information has value.

Sequential Decision Making

- Last time: Principle for taking an action based on beliefs.
- This time: How do we plan for decisions made after each other?
- Key idea: At each point, apply MEU, given all the knowledge you have at that point
- Consequence: After making an earlier decision, you can **learn**. This information has value.
- This is key reason why things become complicated, but principle is no more complex!

Example: Commuting I

Paula has an exam at 9am today, and it is 8:35am. She will not be admitted if she is late, so her loss function is

$$L(t, a) = \begin{cases} 0 & \text{if } t \leqslant 25 \text{ min} \\ 100 & \text{if } t > 25 \text{ min} \end{cases}. \tag{1}$$

Paula decides to take a taxi. The taxi driver knows of routes A and B. Route A can either be clear and fast ($Q = 0$) or jammed and slow ($Q = 1$). The time for route B follows a Gaussian.

$$t_{\mathrm{A}} = \begin{cases} 10 \text{ min} & \text{if } Q = 0 \\ 50 \text{ min} & \text{if } Q = 1 \end{cases}, \qquad P(Q) = \begin{cases} 0.5 & \text{for } Q = 0 \\ 0.5 & \text{for } Q = 1 \end{cases}, \tag{2}$$

$$p(t_{\mathrm{B}}) = \mathcal{N}\left(t_{\mathrm{B}}; 25, 5^2\right). \tag{3}$$

Example: Commuting II

The driver suggests the following strategy:

- Drive the first 2.5 minutes of route A to see whether it is clear (i.e. to observe Q).
- At that point, decide whether to keep going with route A, or to drive back 2.5 minutes and take route B.

Question: Should Paula:

- Ask the driver to take route A?
- Ask the driver to take route B?
- Ask the driver to continue with his strategy?

Decision Trees

- Arrange sequence of observations and decisions into a **tree**
- Helps us to compute utility based on decisions
- Helps us to find sequence of optimal decisions

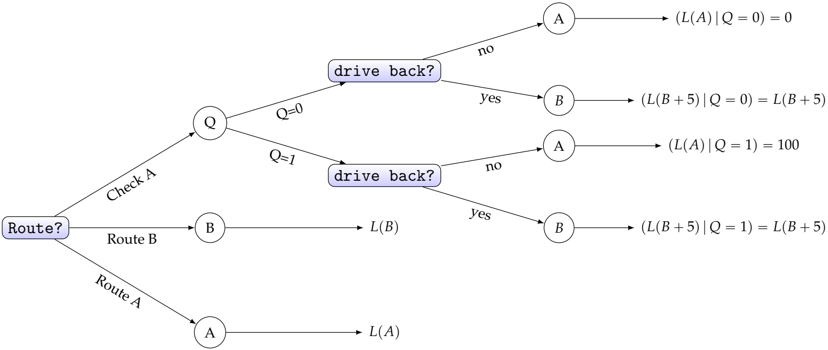
- Rectangle: Decision nodes.
- Circle: Observation of a random variable.
- Leaf nodes: Outcome of utility.
- Nodes are functions of the outcome of parent RVs.

Finding optimal decisions

Simple process:

- Start at leaf nodes and move towards root.
- RV nodes: Calculate expectation over utility values coming in.
- Decision nodes: Choose action corresponding to child with maximum utility. Pass on this utility.

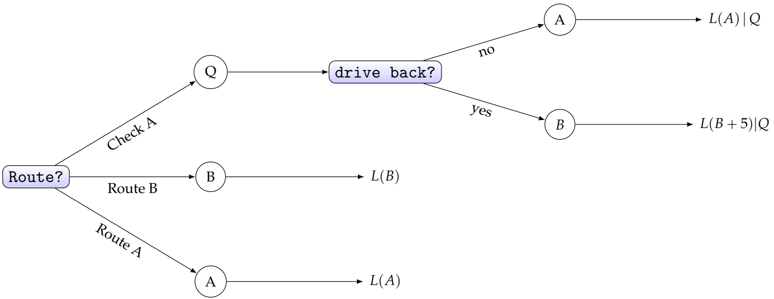
Example: Commuting



- We can split up outcomes of RVs explicitly
- This allows us to find explicit numerical values at all points.
- Note: $(L(B+5) \mid Q = 0) = L(B+5)$, since $B \perp\!\!\!\perp Q$.

Board: solution.

Example: Commuting



- Can make graph smaller by not writing out all outcomes of RVs.
- Decisions become *functions* of parent outcomes.

Board: solution

Exploration-Exploitation Trade-Off

- We saw that “Check A” was the best option.
- Spending effort to gather information, can help us to make better decisions later.
- Uncertainty determines how much exploration can gain.
- (Would exploration be worth it if $P(Q = 1) = 0.99$?)
- ⏩ Exploration-exploitation trade-off.

Conclusion

- Sequential decision making is the same as earlier.
- Can deal with complexity using decision trees.
- Exploration-exploitation trade-off.

Recommended reference: MacKay [1] chapter 36.
Further reading: Russell [2] chapter 17.
There's a whole theory of MDPs out there!

References I

[1] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[2] S. J. Russell. *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010.