Dynamic Programming & Reinforcement Learning

Tutorial session 1

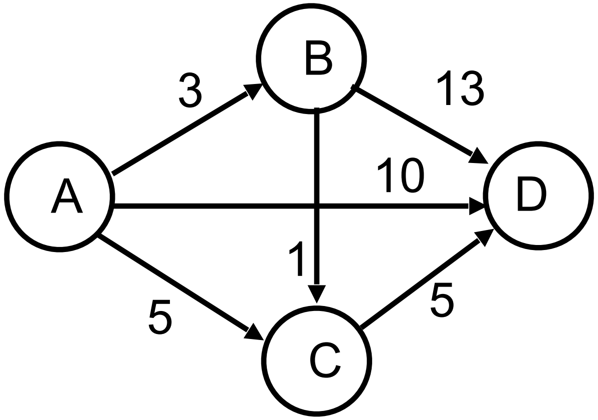
Exercise 1

Compute by hand the discounted ($\beta=3/4$) and long-run average rewards for the following infinite series:

- 1,1,1,1,1,1,1,….
- 2,0,2,0,2,0,2,….
- 100 1's, then ∞ many 0's
- 100 0's, then ∞ many 1's
- Give a reason why you prefer discounting over averaging
- Give a sufficient condition on the rewards under which discounted series converge
- Give an example of a diverging series

Exercise 2

- Find the shortest path from A to D in the directed graph below using dynamic programming (by hand)



- After how many iterations are you certain to have found the optimal solution? Give 2 termination conditions

Exercise 3

- Look up on the internet how Dijkstra’s algorithm works and apply it to the example of Exercise 2
- Construct a simple graph for which Dijkstra does not work and backward recursion does (hint: you are allowed to use negative arc lengths)

Exercise 4

- Determine by enumeration (i.e., try all combinations) the shortest path in the first example (the drawing of the city) of the transparencies of lecture 1
- Check the correctness by verifying the backward recursion
- Implement the problem in some suitable tool or language (R/Excel/…) and solve it

Exercise 5

- Consider an inventory problem with 10 items and immediate replenishments
- Let N be the maximum stock level for each item
- What is the number of states?
- You have a computer with 200GB memory available for computations
- How big can N be such that you can still execute a backward recursion algorithm?
  - Hint: argue first why it suffices to store only 2 vectors the size of $\mathcal{X}$ in memory

Exercise 6

- Estimate the total number of positions of chess pieces on the chess board (feasible & unfeasible)
- Estimate the total number of feasible positions
- Compare your estimation with information you find on the web

Exercise 7

- Consider a knapsack problem with $W=10$, $T=4$, $w=(5,4,3,3)$ and $v=(3,2,2,2)$
- Solve the problem by:

a) dynamic programming

b) a decision tree in which you consider all possibilities

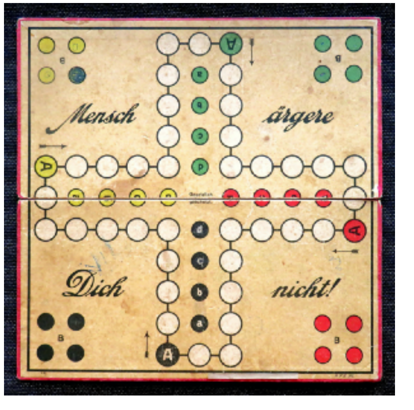
c) What is the complexity (≈ number of calculations) of both methods as a function of W and T?

Exercise 8

- Prove the following property of knapsack problems: $V_t(x) \leq V_t(y)$ for all t and $x \leq y$
- Hint: use induction on t starting from T

Exercise 9

- A variant of the Ludo board game (Dutch: Mens-erger-je-niet) has the following rules. A token advances with the role of a die, when a player roll a 6 he/she can roll again until the outcome is less than 6. What is the expected number of squares that the token advances in a full turn?



German board (wikipedia)

Exercise 10

- Two common examples of discrete and continuous distributions are the Poisson and exponential distribution:

$$N \sim \text{Poisson}(\lambda) \Leftrightarrow P(N = n) = \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}$$

$$X \sim \exp(\mu) \Leftrightarrow P(X \leq t) = 1 - \mathrm{e}^{-\mu t}$$

- What are their expectations EN and EX?

(Hint: for the exponential distribution first derive the density and then use partial integration; or use a formula for the expectation that uses the tail of F)