Optimal Control and Dynamic Programming

TU/e

4SC000 Q2 2022-2023

Duarte Antunes

Lecture 1

Outline

- <u>Introduction to optimal control and applications</u>
- Dynamic programming algorithm

Optimal control

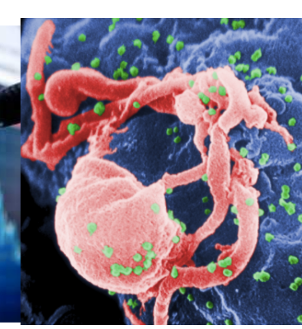
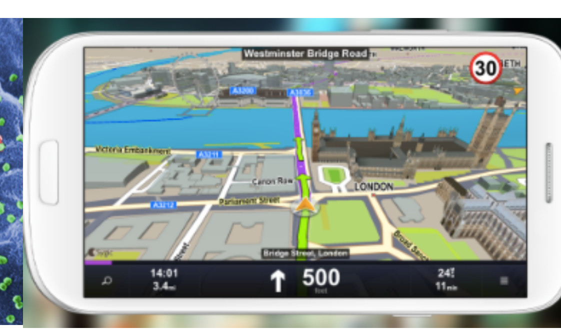
Optimality

- Useful design principle in many engineering contexts (optimize efficiency of a refrigerator, minimize the fuel consumption of a car, etc.).
- Nature is described by laws derived from optimality principles.
- We optimize every day when we make decisions (true?).

Optimal control

- Deals with problems in which optimal decisions or control actions are pursued over a time period in order to reach final and intermediate goals.
- Arises in the control of physical systems (e.g. mechanical, electrical, biological) and in many other contexts (e.g., economics, computer science, and game theory).

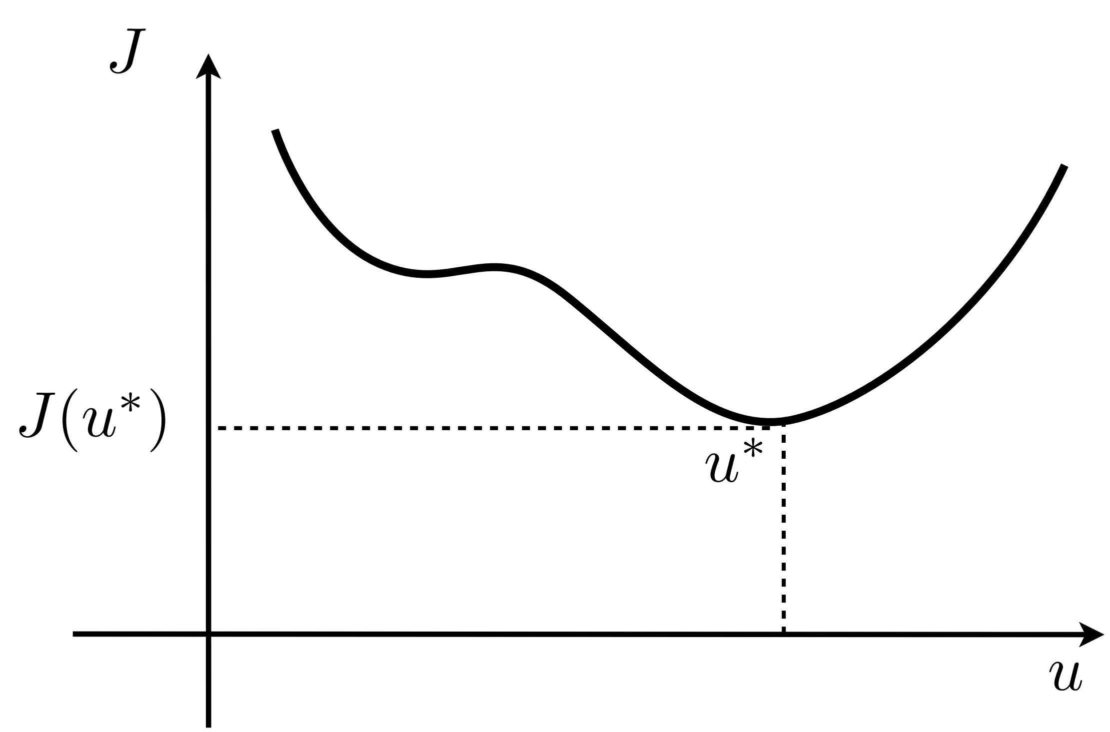


Optimal control vs static optimization

Static optimization

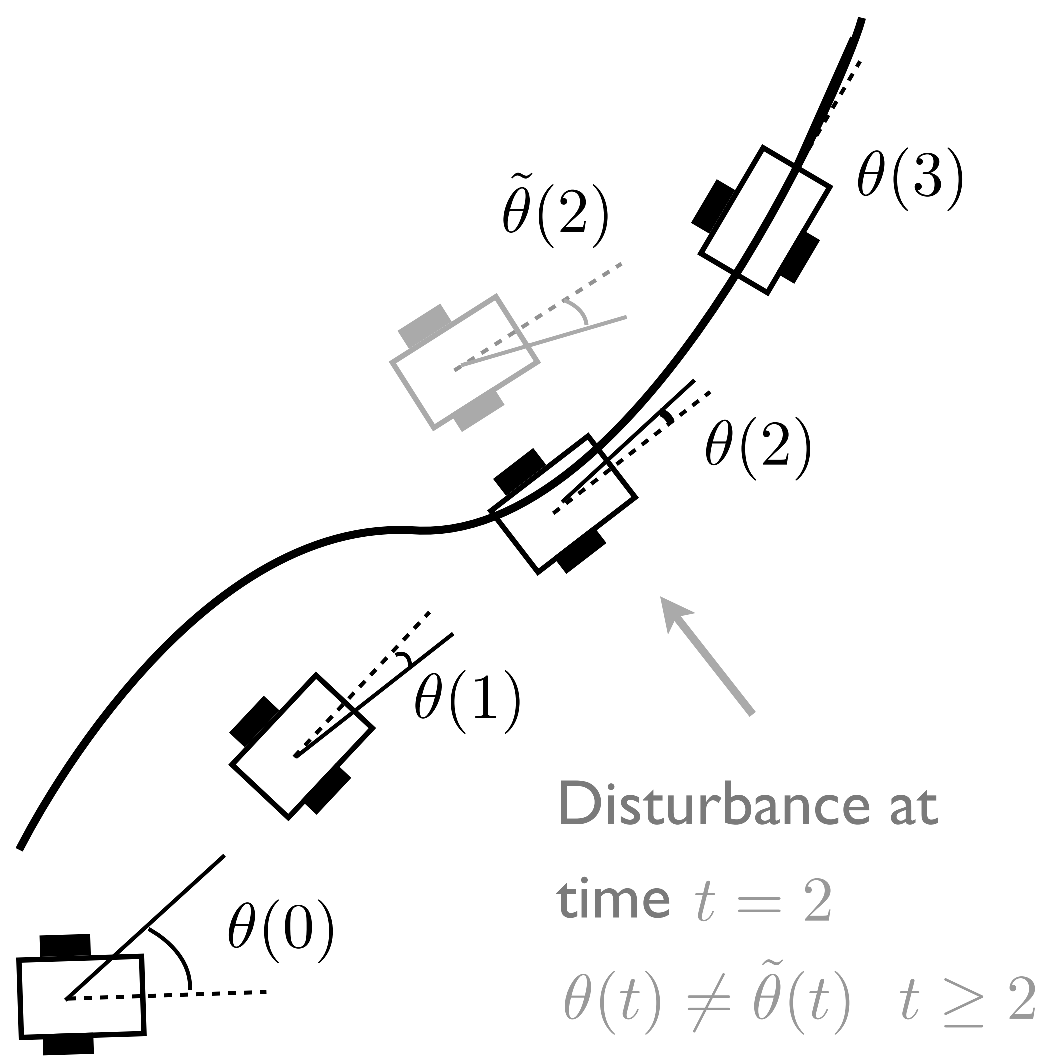
- Determine <u>one</u> optimal decision.
- Examples: decide on the price of a product, determine the slope of a straight line which best fits data, etc.



Optimal control vs static optimization

Optimal control

- Determine several optimal decisions over time.
- Decisions are functions of state, i.e., a control law to cope with disturbances.
- Examples: driving a car/bike in a race, positioning the tip of a robot arm in the presence of disturbances, playing chess, etc.



Optimal control formulation

Dynamic model

- Specifies the rules of the problem or the equations of the physical system.
- State: summarizes relevant information to make future decisions.
  - Control actions: influence the evolution of the state over time.
  - State evolution may be deterministic or stochastic (driven by disturbances).

Cost function

- Encapsulates the goals to be achieved in the problem.
- Typically additive over time and by convention should be minimized.

Goal: find a control policy which minimizes the cost

- Policy: set of functions mapping the state at each instant of time to an action.
- Related problem: compute an optimal path/trajectory consisting of optimal decisions over time for a given initial state.

Optimal control problems

Three classes of problems will be considered in the course

	time	state space
Discrete optimization problems	discrete	discrete (countable)
Stage decision problems	discrete	general
Continuous-time optimal control problems	continuous	general

Some applications are discussed next and more applications later. However, there are many others - see Appendix B.

Applications

Traditional process control

- controlling an inverted pendulum, mass-spring damper, double integrator, quadcopter, etc.

Aerospace

- minimum-fuel launch of a satellite, etc.

Operational research, management, finance

- inventory control, control of a queue, control of networks (data, traffic, etc.), etc.

Computer Science

- shortest path in graphs, scheduling, selection problems, among others.

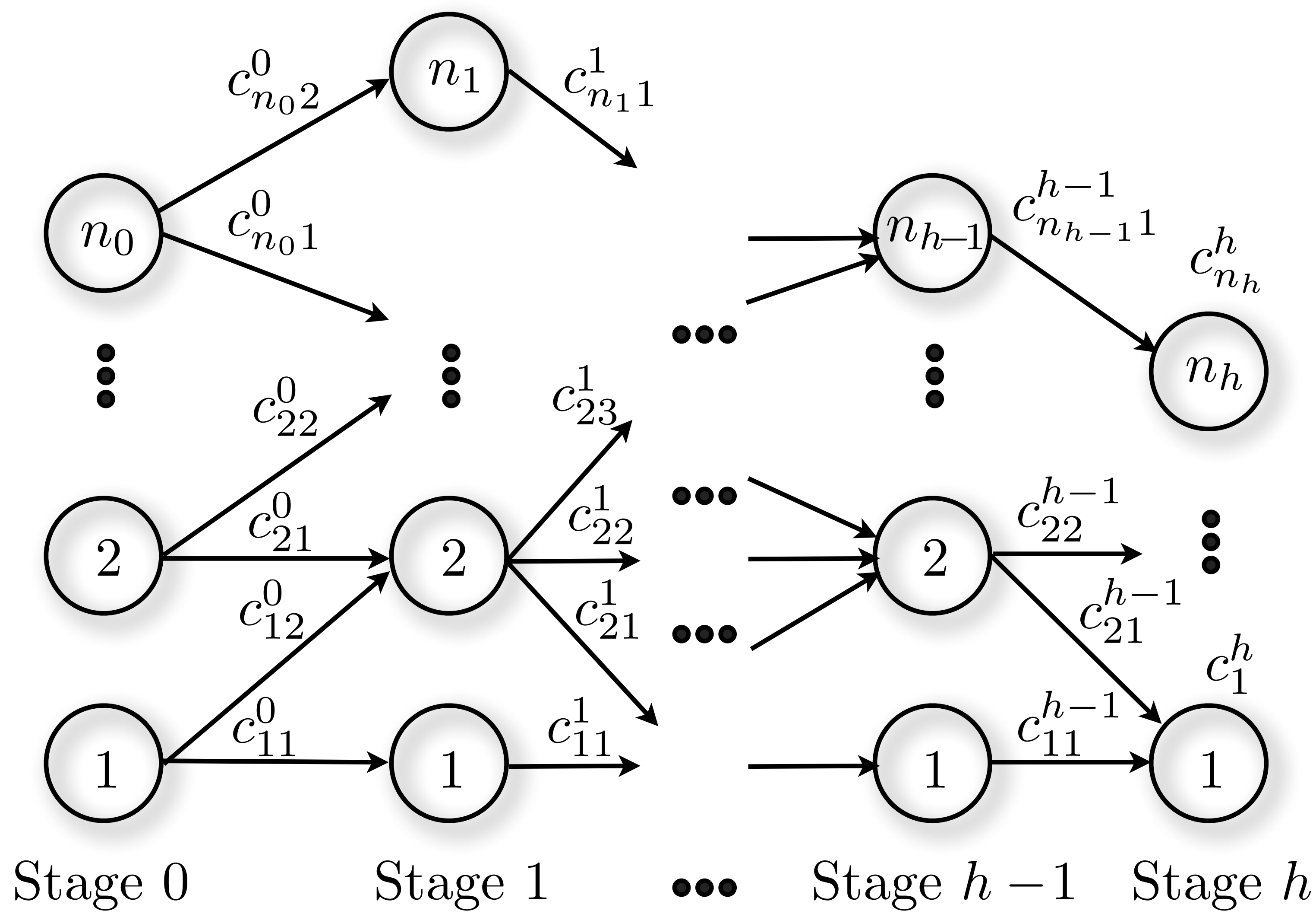
Other fields

- Computational biology, automotive, games, many others.

Next slides address some applications treated in the course, where we will consider also cases where uncertainty is present.

Discrete Optimization Problems

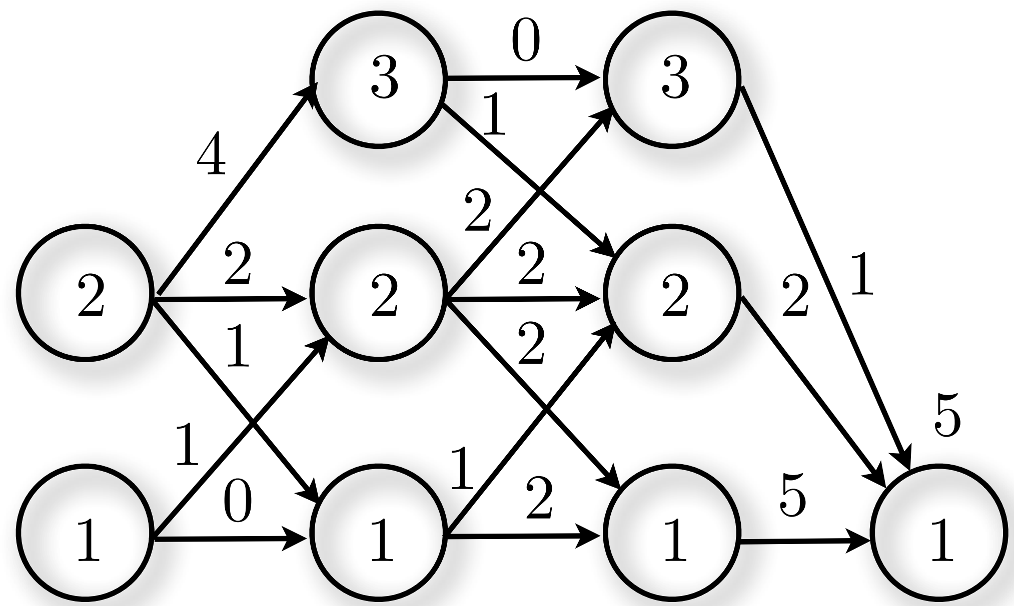
Specified by a <u>transition diagram</u> with $h-1$ decision stages



- <u>Dynamic model</u>: <u>circles</u> indicate states at each of h stages; <u>arrows</u> indicate actions for each state which lead to states at next stages.
- <u>Costs</u> $c^k_{i,j}$ are associated with actions j for each state i at each stage k; for the terminal stage h the costs c^h_i depend only on the state i.

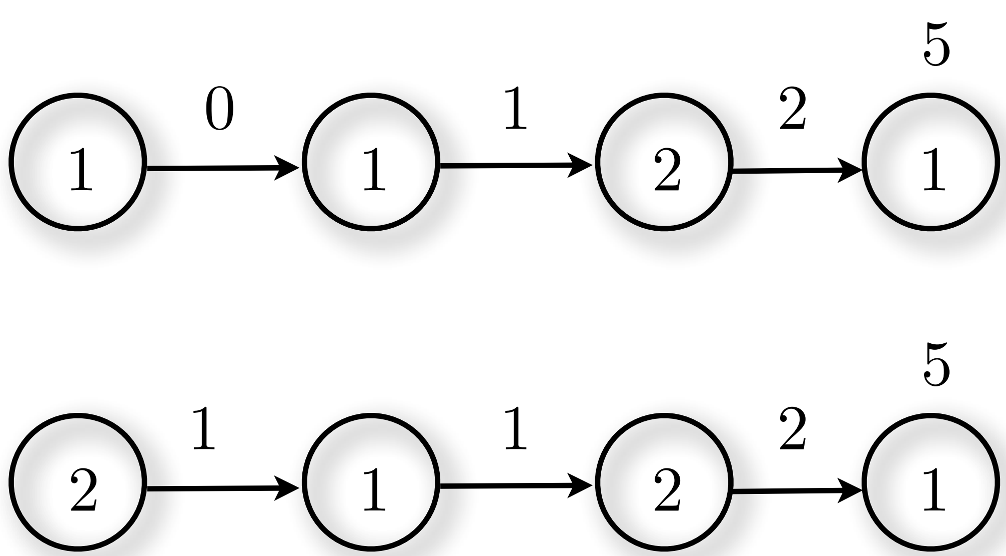
Discrete Optimization Problems

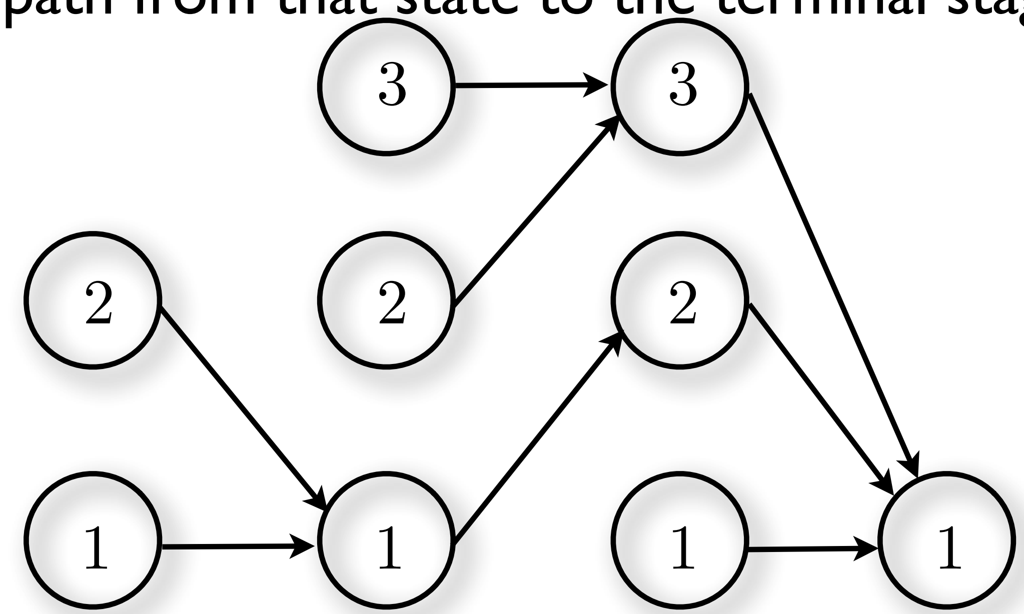
Challenges



(i) Determine an optimal path for a given initial state which minimizes the sum of costs incurred at every stage (including the terminal stage).

(ii) Determine an optimal policy specifying for each state the first decision of the optimal path from that state to the terminal stage.



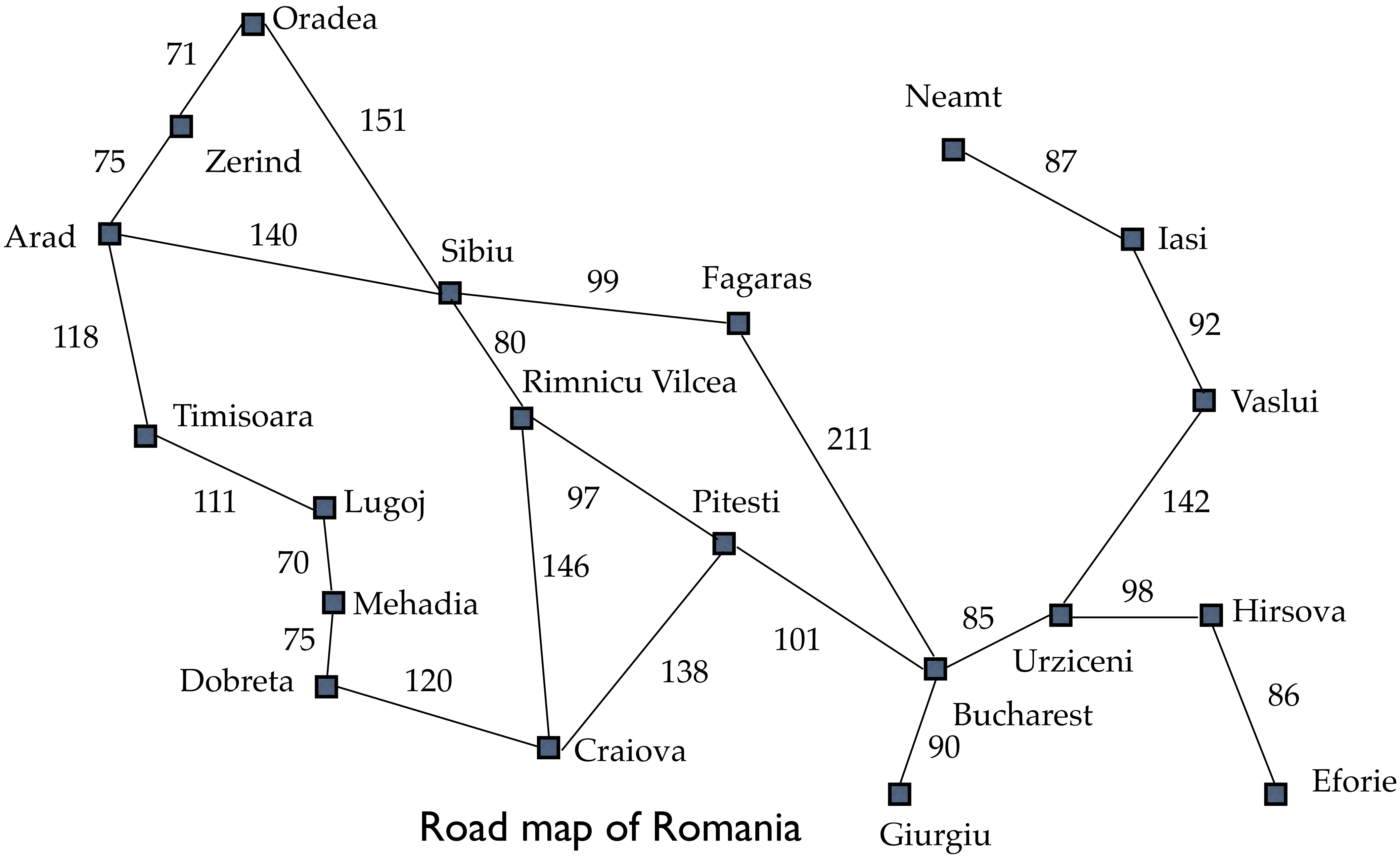


Inventory control

How to manage the supply of products in a shop? Overstock is prejudicial (physical space limitations, technological obsolescence, etc.) and under stock undermines sales.

Shortest paths in graphs

What is the shortest distance from Bucharest to Lugoj?

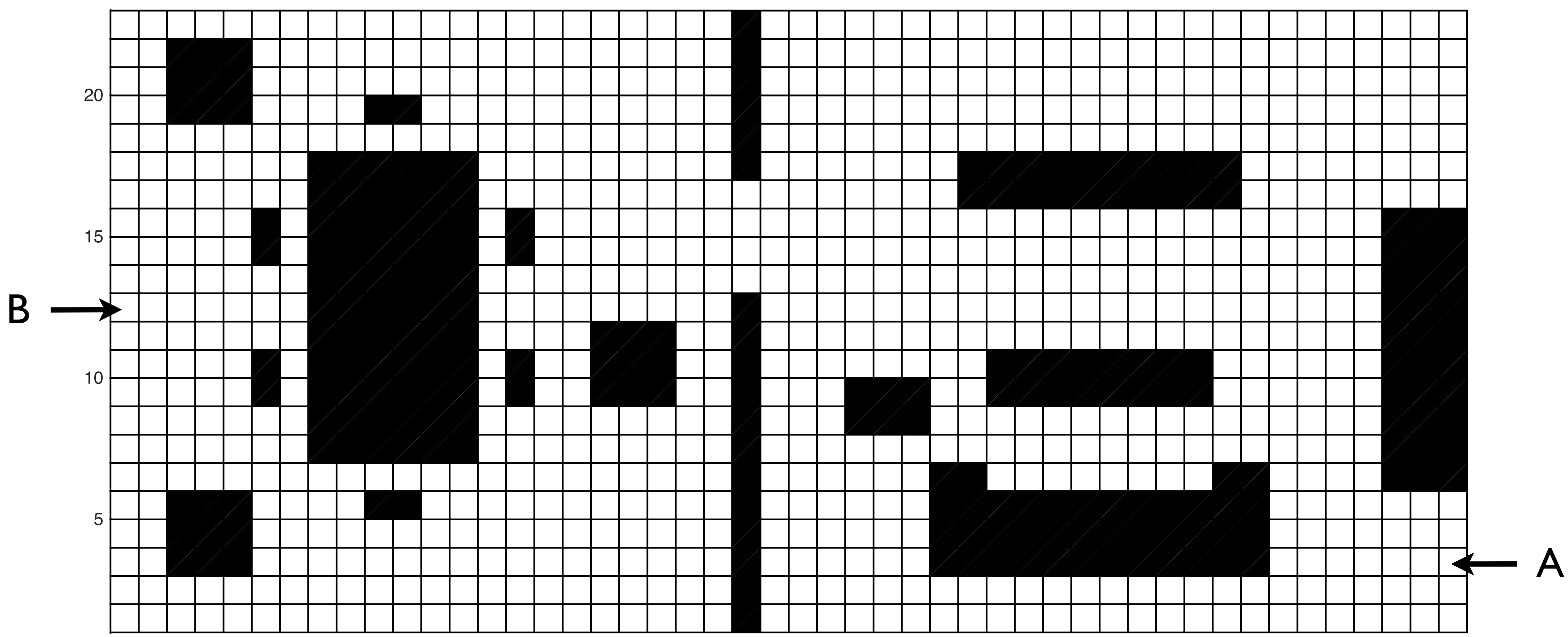


Road map of Romania

*Source: *Artificial intelligence: a modern approach*, Stuart J. Russel, Peter Novig, 3rd edition, 2016

Robot path planning

What is the shortest path for a robot to go from point A to B?



Games

How to make profit in expectation in a game such as blackjack?

As portrayed in the movie '*21*' the MIT blackjack team had an answer to this problem (using optimal control?). The same movie illustrates the principle to achieve this, using a famous game show problem https://www.youtube.com/watch?v=Q5nCtgcL4jU

Stage decision problems

Dynamic model $$x_{k+1} = f_k(x_k, u_k) \qquad k \in \{0, \dots, h-1\}$$

Cost function $$\sum_{k=0}^{h-1} g_k(x_k, u_k) + g_h(x_h)$$

Goals

(i) find policy $\pi = \{\mu_0, \dots, \mu_{h-1}\}$ $\quad u_k = \mu_k(x_k)$

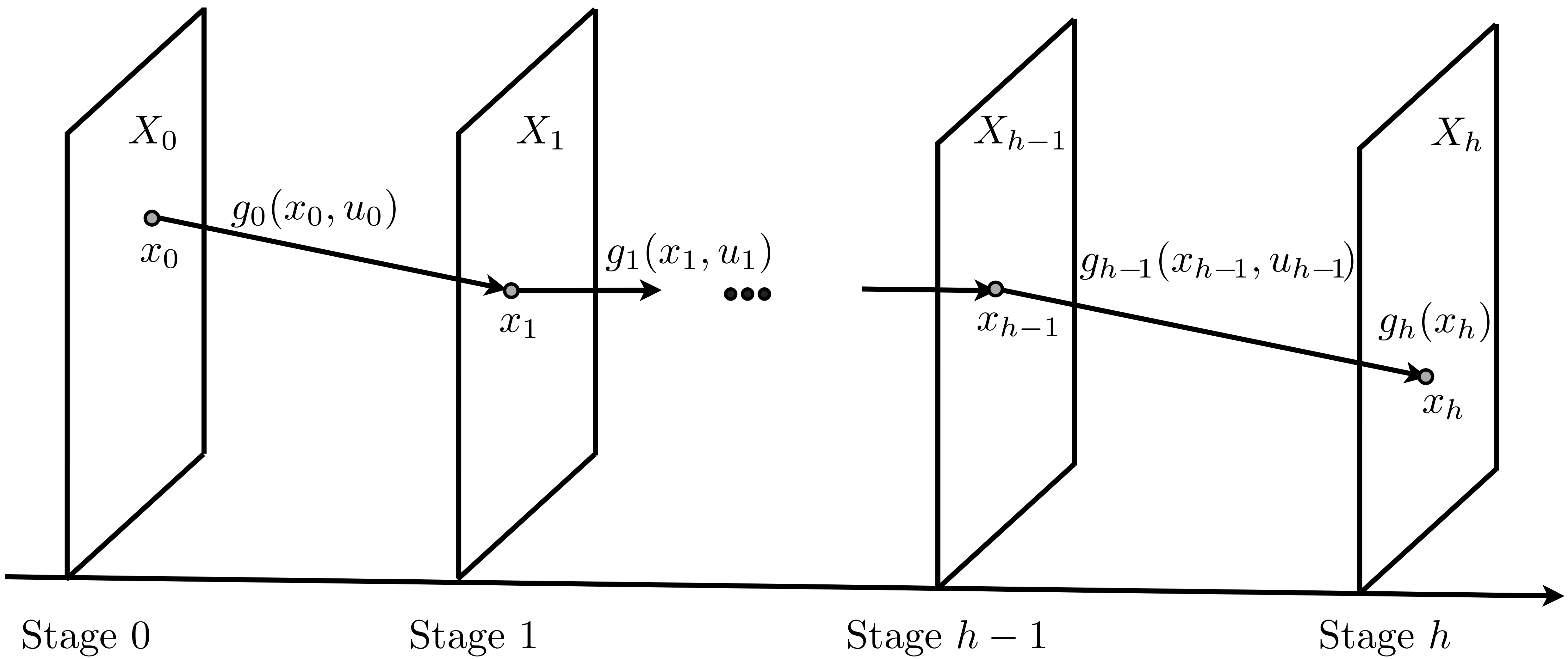
that leads to the minimum cost for <u>every</u> initial condition.

(ii) find path $\{(x_0, u_0), (x_1, u_1), \dots, (x_{h-1}, u_{h-1})\}$

that leads to the minimum cost for a <u>given</u> initial condition.

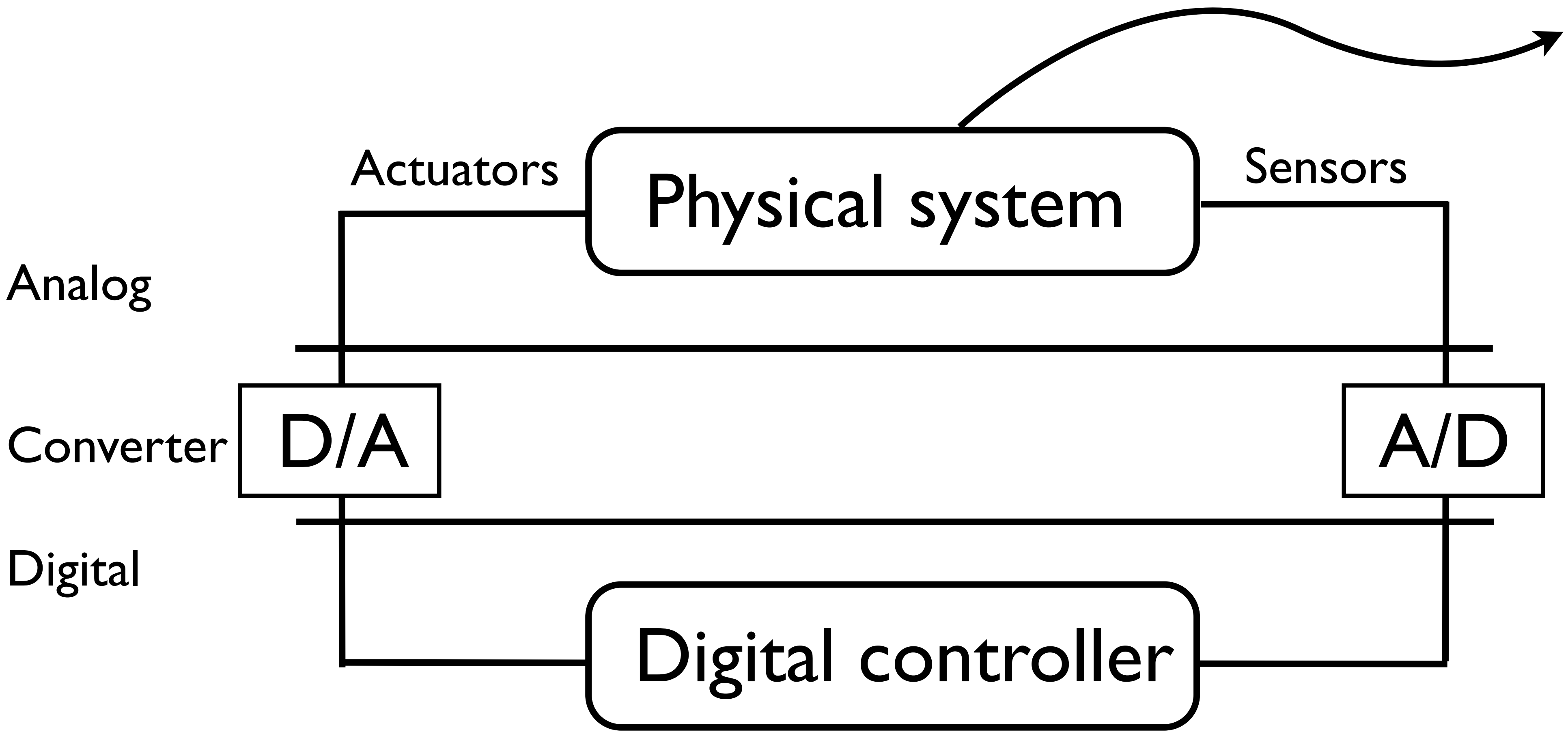
Stage decision problems

Generalization of discrete optimization problem considering general state and input spaces, e.g., $\mathbb{R}^n$



Digital control

Prime application: how to design a digital controller for a physical system?



or

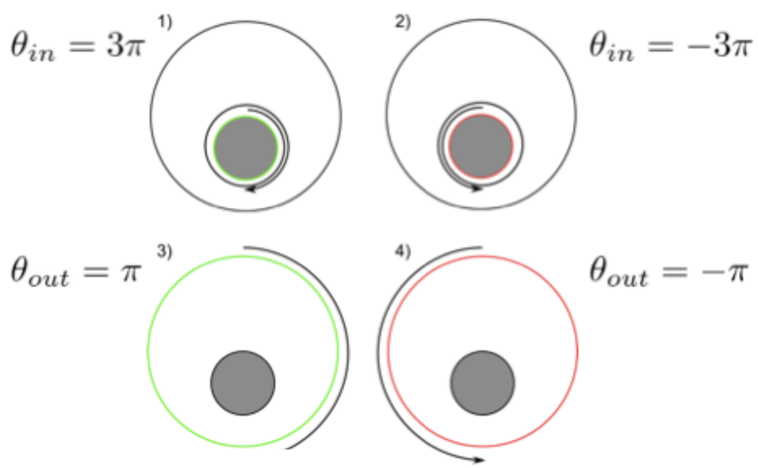
or

or ....

Several variants: full state is available or only an output, system can have disturbances or not, etc.

Mixing

How to mix two fluids in minimum time?



Camera

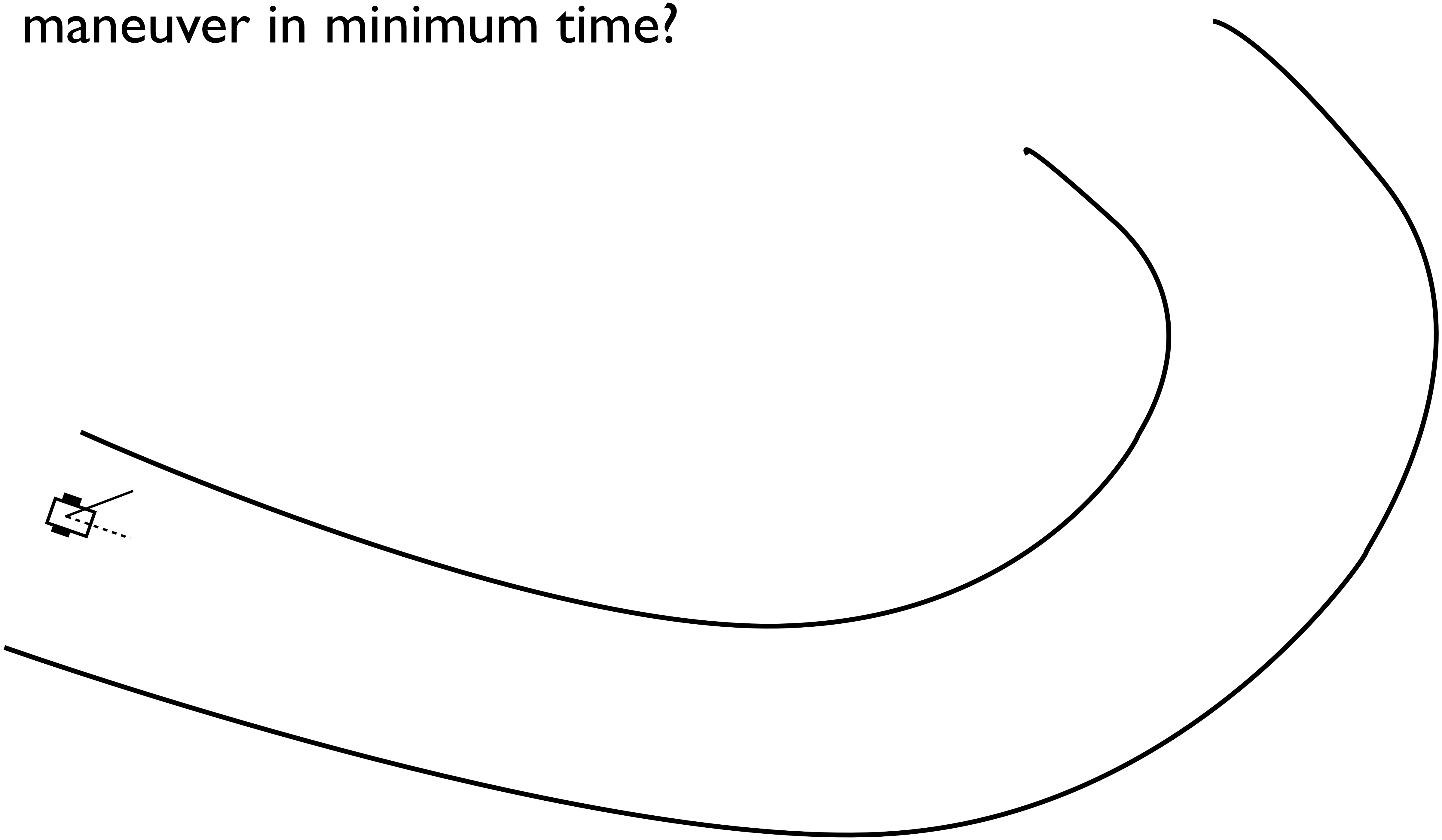
images

Actuation: 4 possible rotations decided once every h seconds

Control law

Digital control of a unicycle robot

Given a unicycle robot with constraints on speed and rotation rate and controlled digitally, how to perform a maneuver in minimum time?



Continuous-time optimal control problems

Dynamic model

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(0) = x_0, \quad t \in [0, h]$$

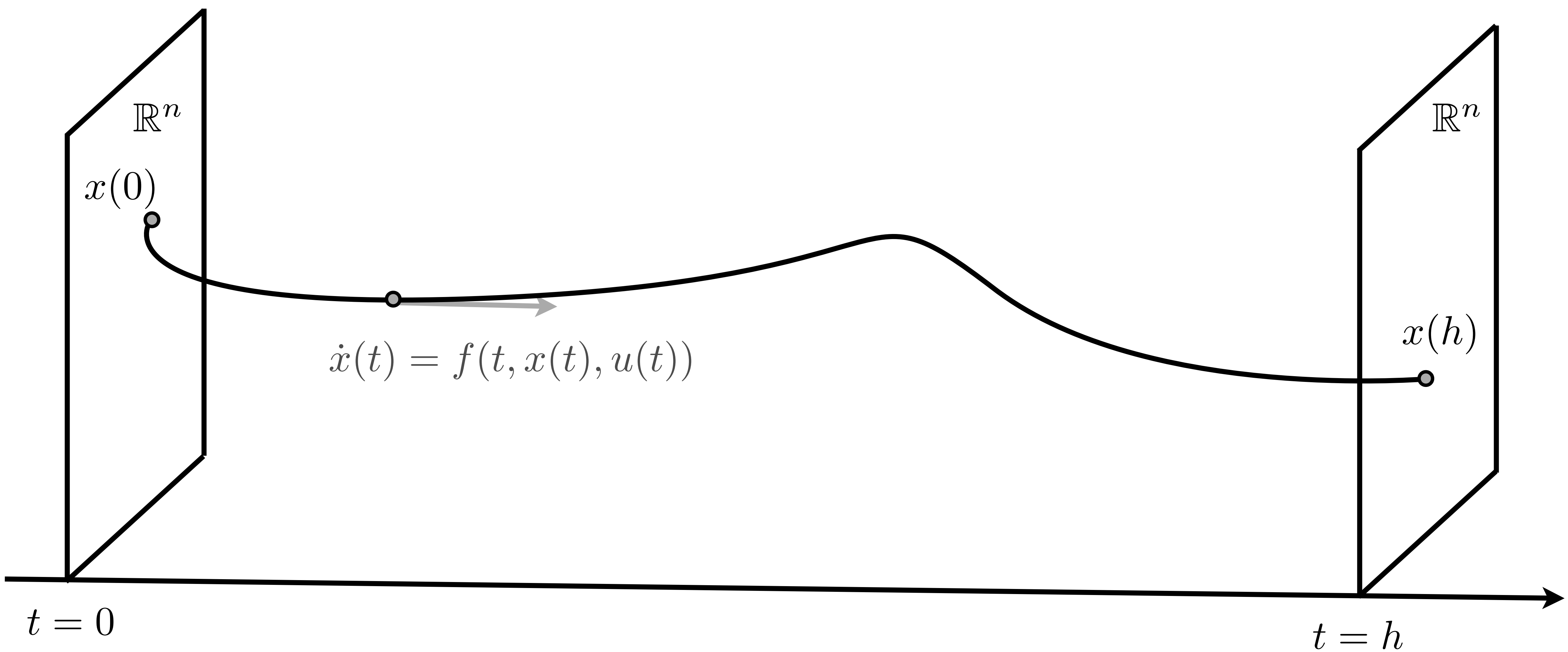
Cost function

$$\int_0^h g(t, x(t), u(t))dt + g_h(x(h))$$

Goals

- Find a <u>feedback</u> policy $u(t) = \mu(t, x(t))$ minimizing the cost function for <u>every</u> initial condition.
- Find a control input $u(t), \quad t \in [0, h]$, minimizing the cost function for a <u>given</u> initial condition.

Continuous-time optimal control problems



Most applications in control systems: motion control, aerospace, etc.

Minimum energy control

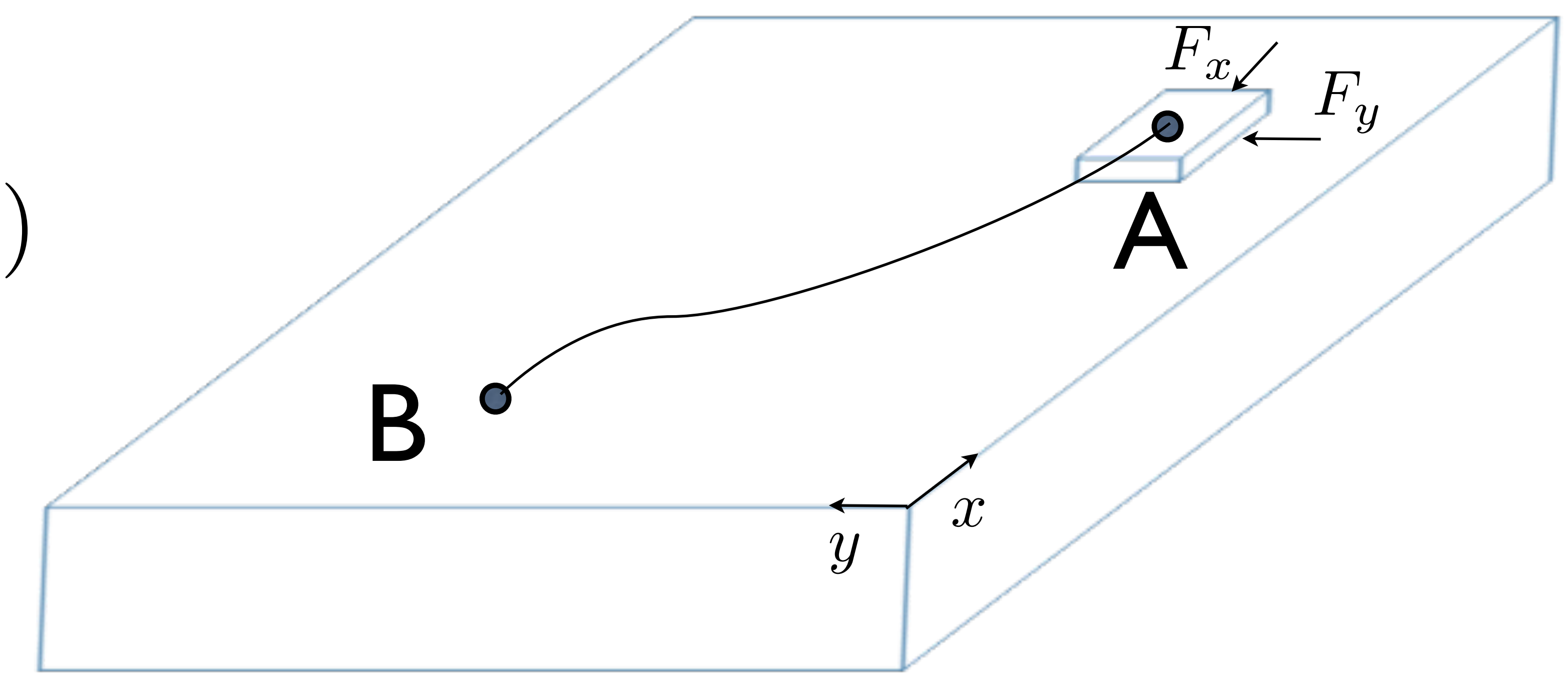
How to move a mass, described by a linear model, from point A to point B with minimum energy?

$$\min_u \int_0^T g(u(t))dt$$

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$x(0) = x_0$$

$$x(T) = x_{\text{desired}}$$



Minimum time control

How to move a quadcopter, described by a linear model, from one hovering position to another one in minimum time?

x_0 x_T

$$\min T$$

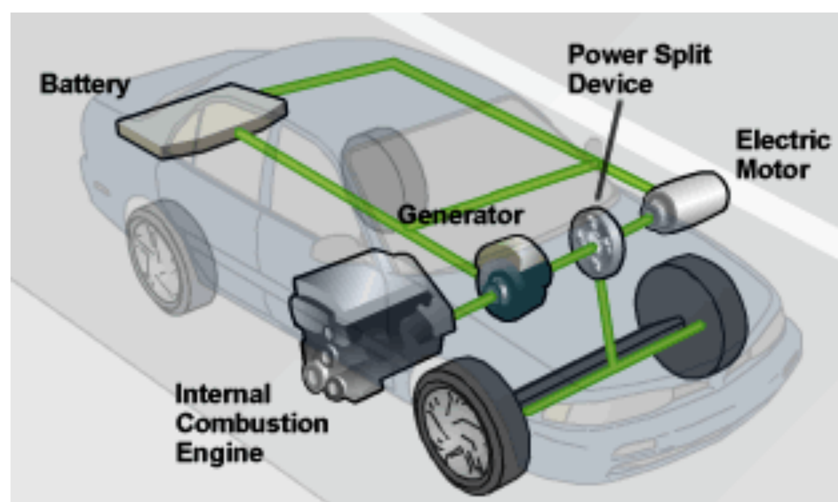
$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$x(0) = x_0$$

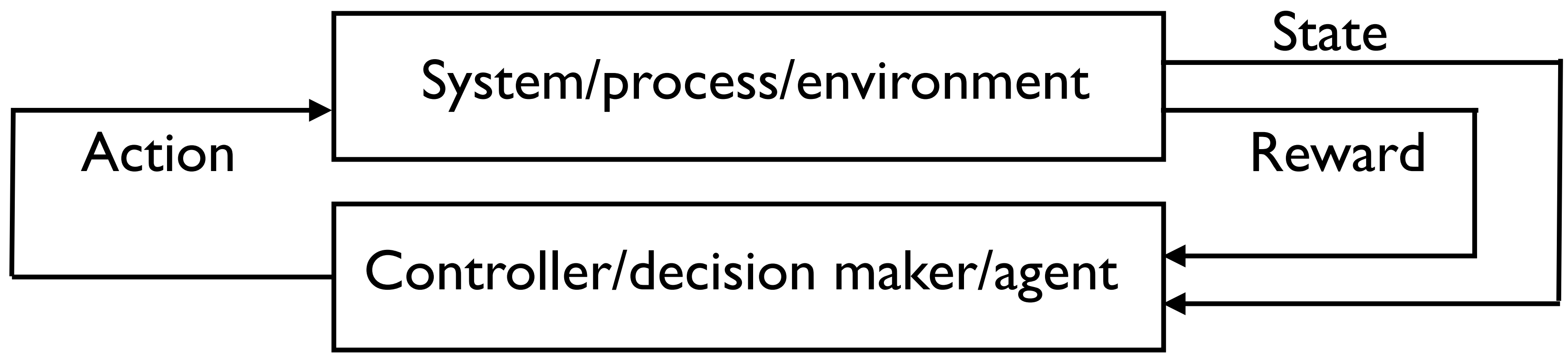
$$x(T) = x_{\text{desired}}$$

Energy management of EHV

Hybrid electric vehicles (HEV) have a battery where energy can be stored (e.g. during braking). Given a drive cycle, how to design the power slip between the battery and the internal combustion engine to minimize fuel consumption?



Reinforcement learning formulation



Dynamic model

- None! Only the real system or a simulator which can be ran is available
- Still, for analysis, a model taking the form $x_{k+1} = f(x_k, u_k, w_k)$ is subsumed.

disturbances

Reward/Cost function

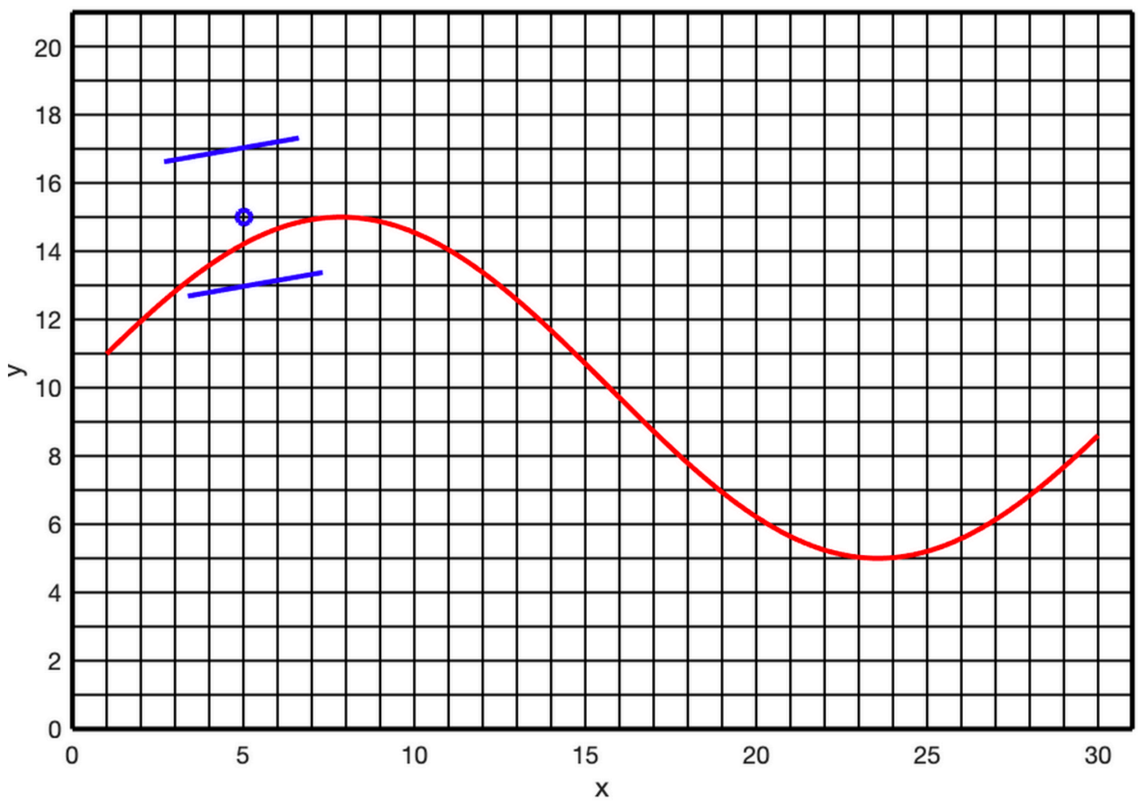
- Encapsulates the goals to be achieved in the problem. Obtained through the agent's interaction with the environment/simulator.
- <u>Additive over time</u> $\sum_{k=0}^{h-1} g_k(x_k, u_k) + g_h(x_h)$. Typically in the literature it <u>should be maximised</u>. However, we will consider that it <u>should be minimized</u>.

Goal

- Find a control policy (function from states to actions) which minimizes the cost based on data (trial and error)

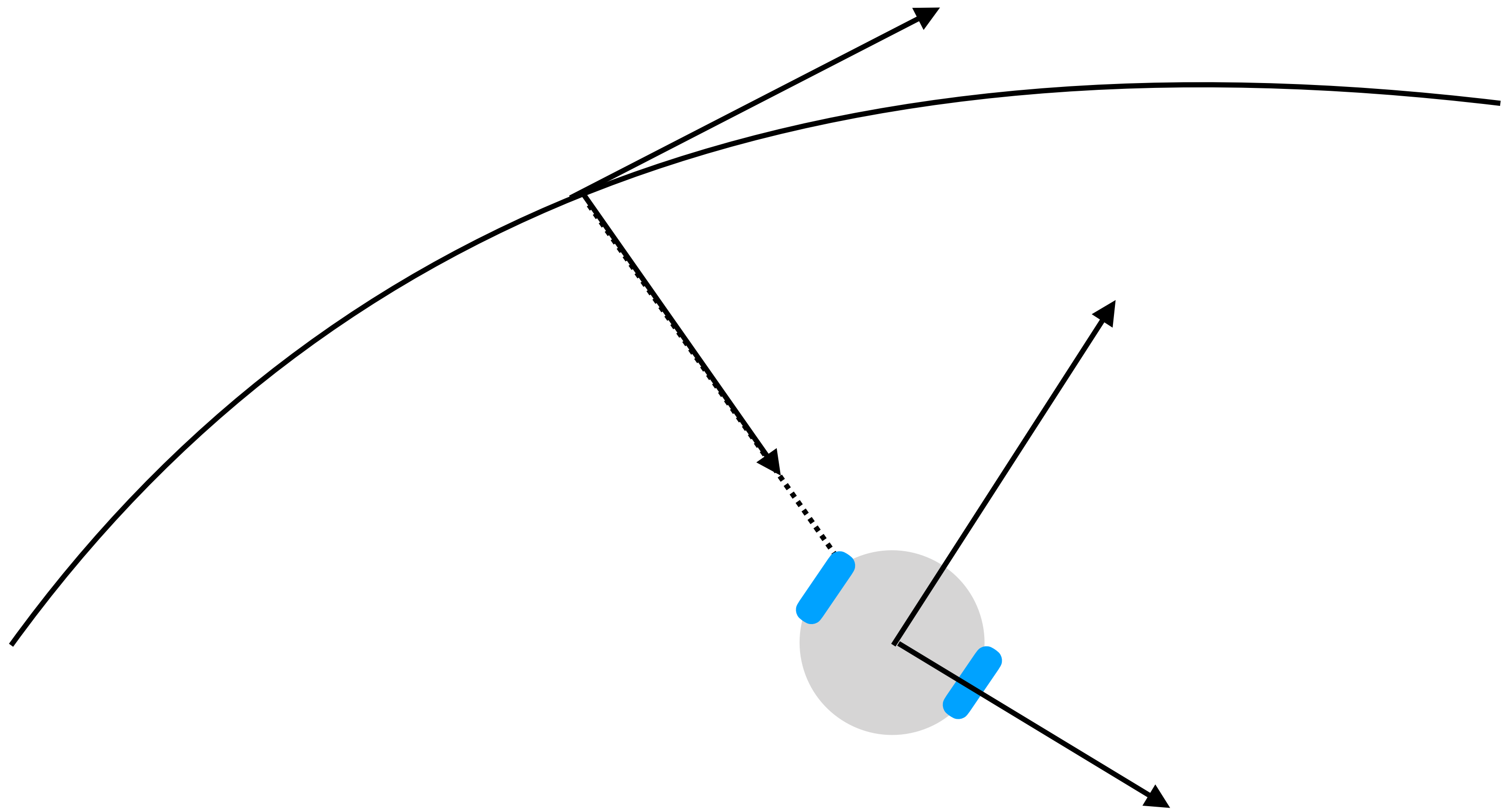
Buzz wire challenge

Learn fastest control strategies for the buzz wire challenge (guide a metal loop along a wire without touching the wire)

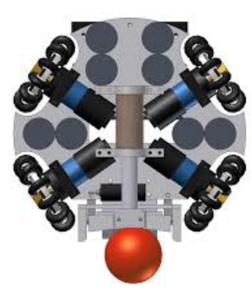


Optimal path following for any robot

Learn a control strategy to follow a path independently of the robot



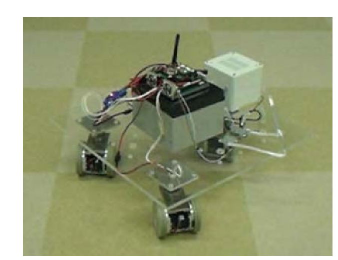
Car-like robot



Omni directional

Two wheeled robot

Kludge robot



2 caster &1 steering wheel

Outline

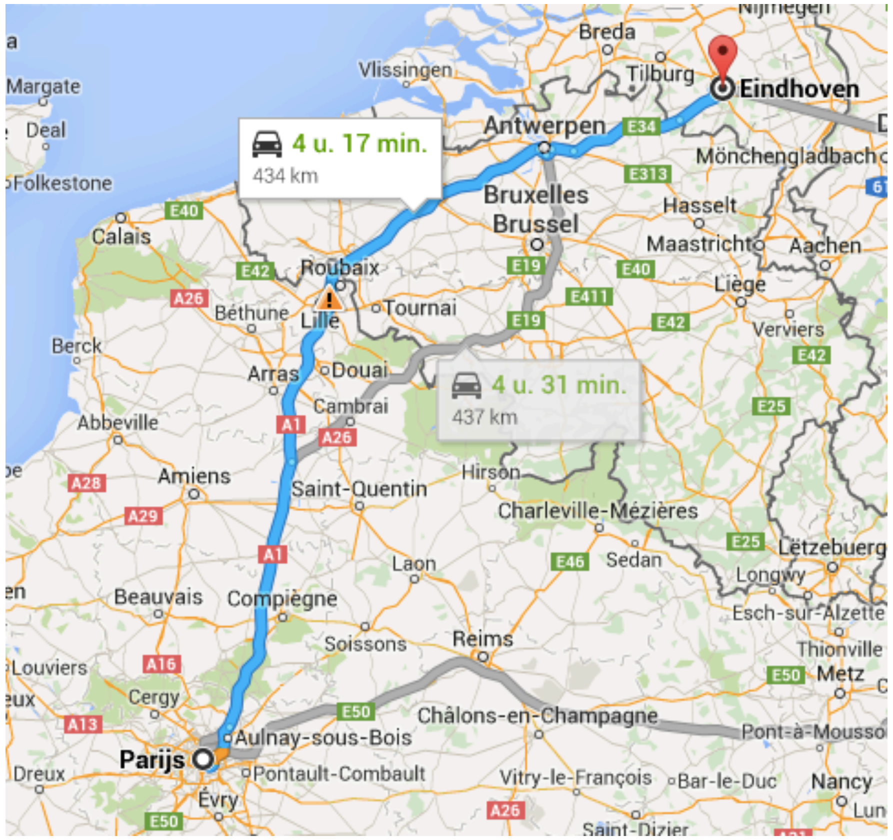
- Introduction to optimal control and applications
- <u>Dynamic programming algorithm</u>

Dynamic progamming

- Dynamic programming is an approach to solve optimal control problems.
- It allows to find functions mapping states into actions. These functions we call <u>policies</u> or <u>control laws</u>.
- One can use these functions to control a system in the presence of <u>disturbances</u>.
- It also allows to compute optimal paths/trajectories, although, as we shall see, other methods might be more efficient.

The principle of optimality

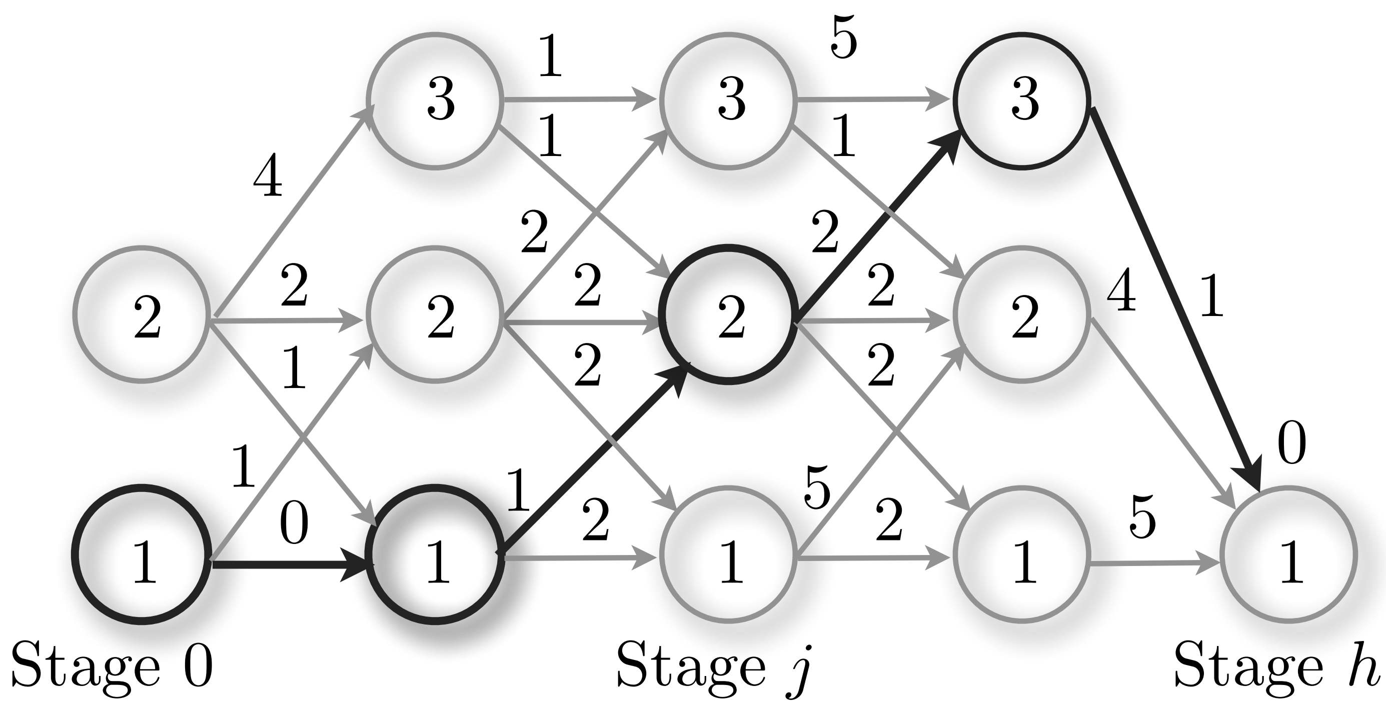
Example: shortest route from Eindhoven to Paris passes through Antwerp. Then the piece of the route from Antwerp to Paris is the shortest route between the two cities.

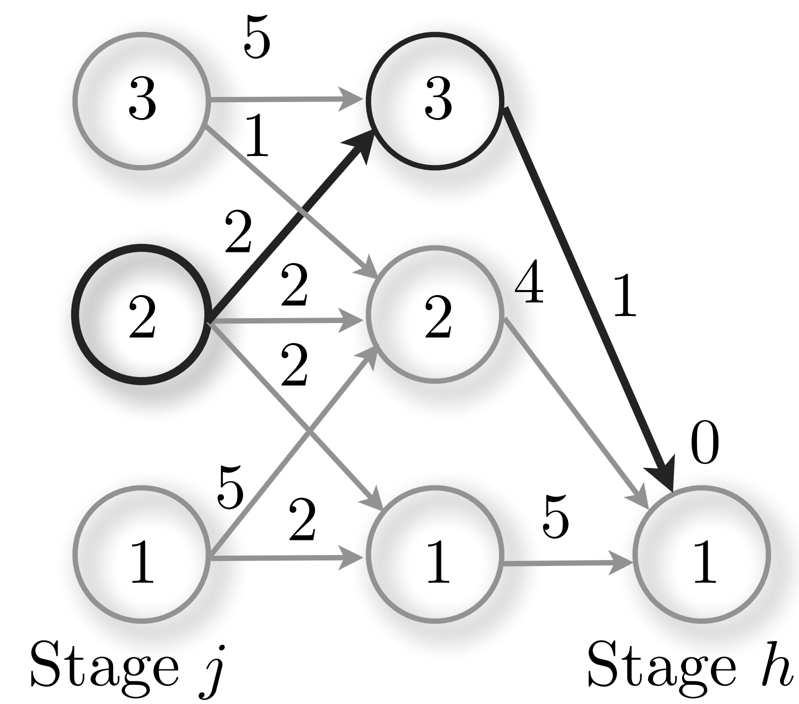


The principle of optimality

The tail of an optimal path is also optimal

- Given an optimal path for a discrete optimization problem from stage 0 to stage h consider the state x_j at stage j belonging to the optimal path.
- Then the decisions along the optimal path from stages j to h are also optimal for the discrete optimization problem with initial stage j, initial state x_j and final stage h.

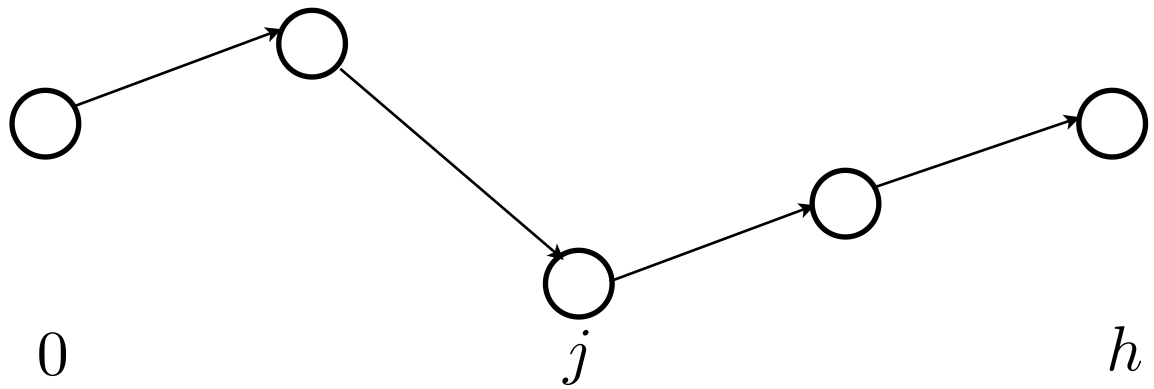




- The principle of optimality also applies to stage decision problems and continuous-time optimal control problems and is the basis of the <u>dynamic programming algorithm.</u>

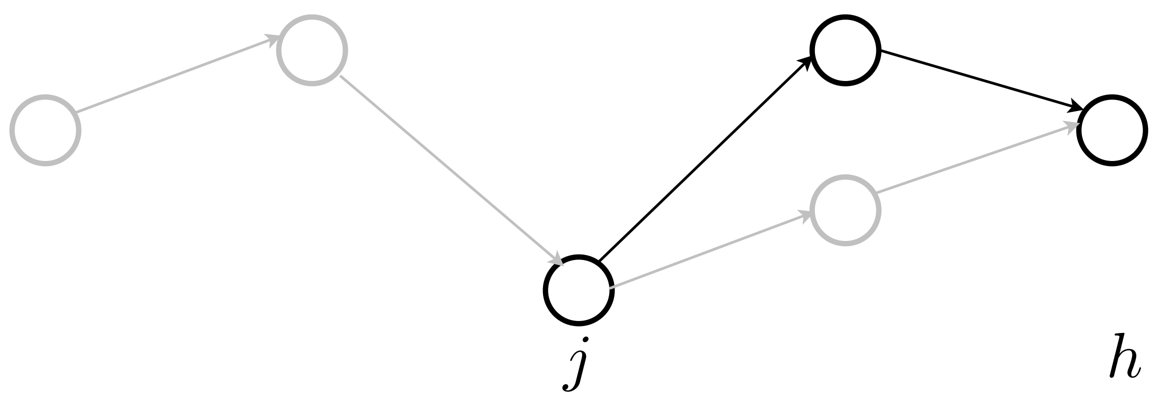
Proof of the principle of optimality

1. Consider an optimal path from stage 0 to stage h.



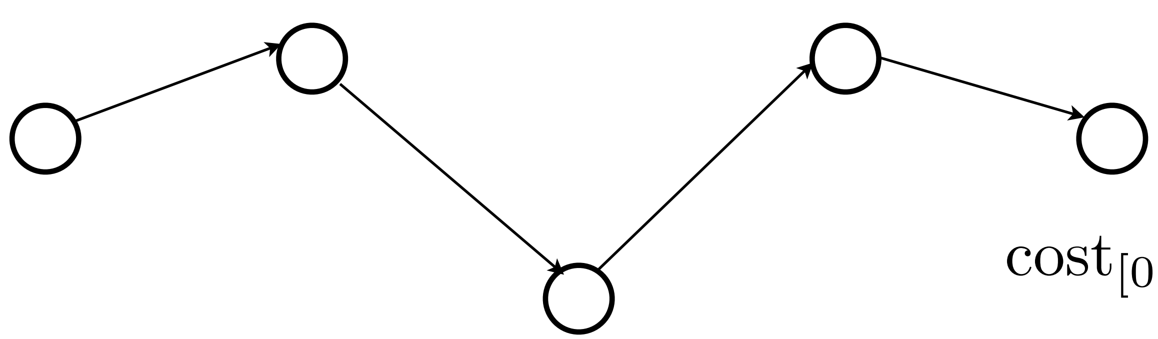
$$\mathrm{cost} = \mathrm{cost}_{[0,j)} + \mathrm{cost}_{[j,h]}$$

2. Suppose that the piece of the path from stage j to stage h is not optimal.



$$\underline{\mathrm{cost}}_{[j,h]} < \mathrm{cost}_{[j,h]}$$

3. Then there exists a path with smaller cost from stage 0 to stage h - contradiction!



$$\mathrm{cost}_{[0,j)} + \underline{\mathrm{cost}}_{[j,h]} < \mathrm{cost}_{[0,j)} + \mathrm{cost}_{[j,h]}$$

Dynamic programming algorithm

Main idea

- Find first the optimal policy and paths from stage j to h, and then use these to compute the optimal policy and paths from stage $j-1$ to stage h (principle of optimality).

The dyn. prog. algorithm for discrete optimization problems:

(1) Start at the <u>final decision stage</u> and denote the terminal cost by cost-to-go at stage h,

$$J_h(i) = c_i^h$$

(2) For every state i at stage $k = h-1$ compute the optimal action j as follows

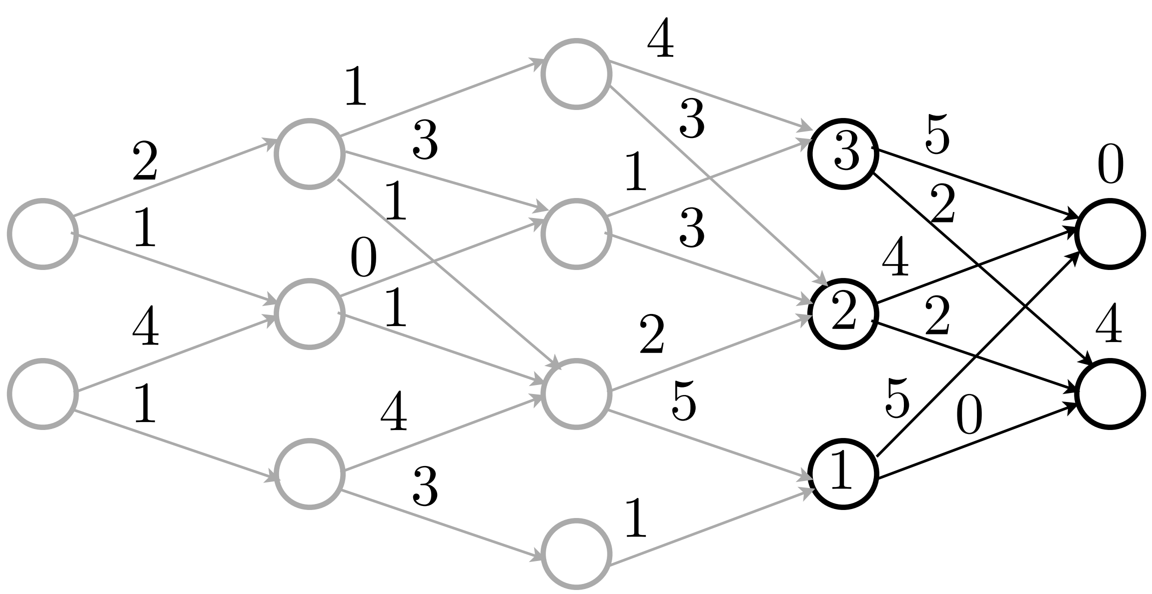
$$\min_{j \in \text{actions/arrows}} c_{ij}^k + J_{k+1}(\text{state at stage k+1 when } j \text{ is picked})$$

Denote the minimum value by cost-to-go at stage k by $J_k(i)$.

(3) Repeat (2) for stages $k \in \{h-2, h-3, \ldots, 1, 0\}$ moving backwards.

Then, the function which maps each state to the action obtained in (2) is an optimal policy.

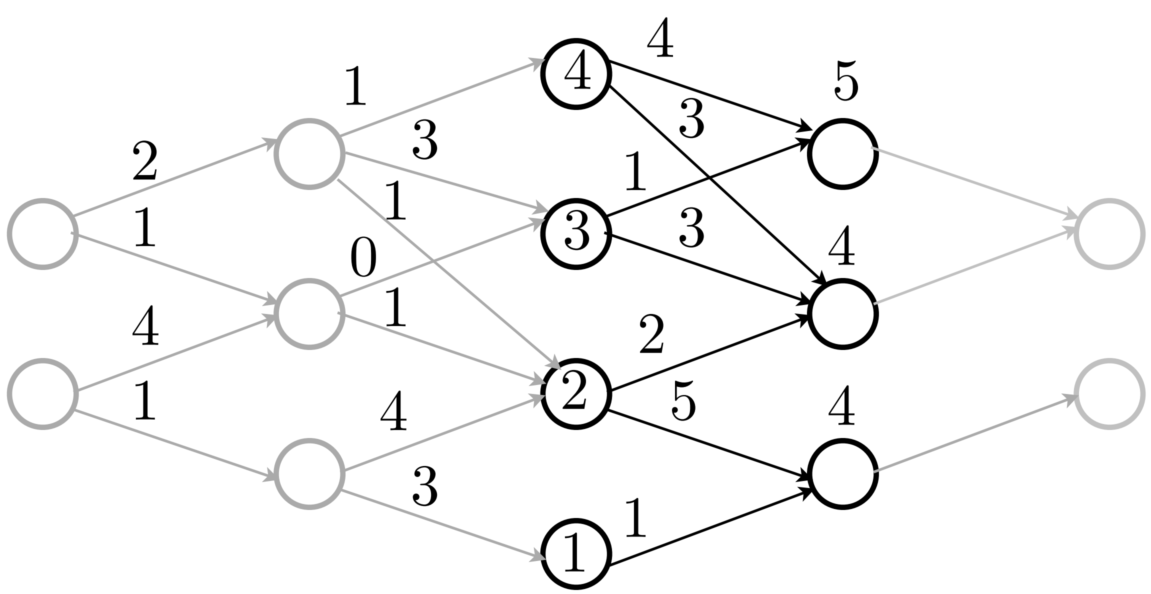
Example



Iteration 1 - Stage 3

State	Cost-to-go
1	$\min\{0+4, 5+0\} = 4$
2	$\min\{4+0, 2+4\} = 4$
3	$\min\{5+0, 2+4\} = 5$

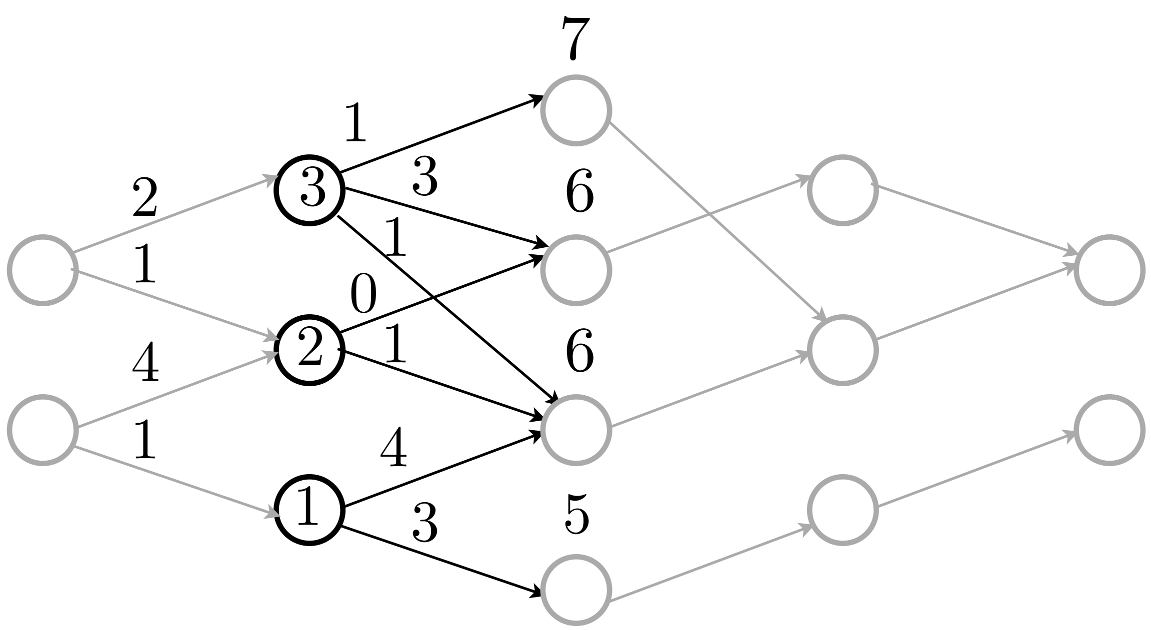
Example



Iteration 2 - Stage 2

State	Cost-to-go
1	$1 + 4 = 5$
2	$\min\{2 + 4, 5 + 4\} = 6$
3	$\min\{1 + 5, 3 + 4\} = 6$
4	$\min\{4 + 5, 3 + 4\} = 7$

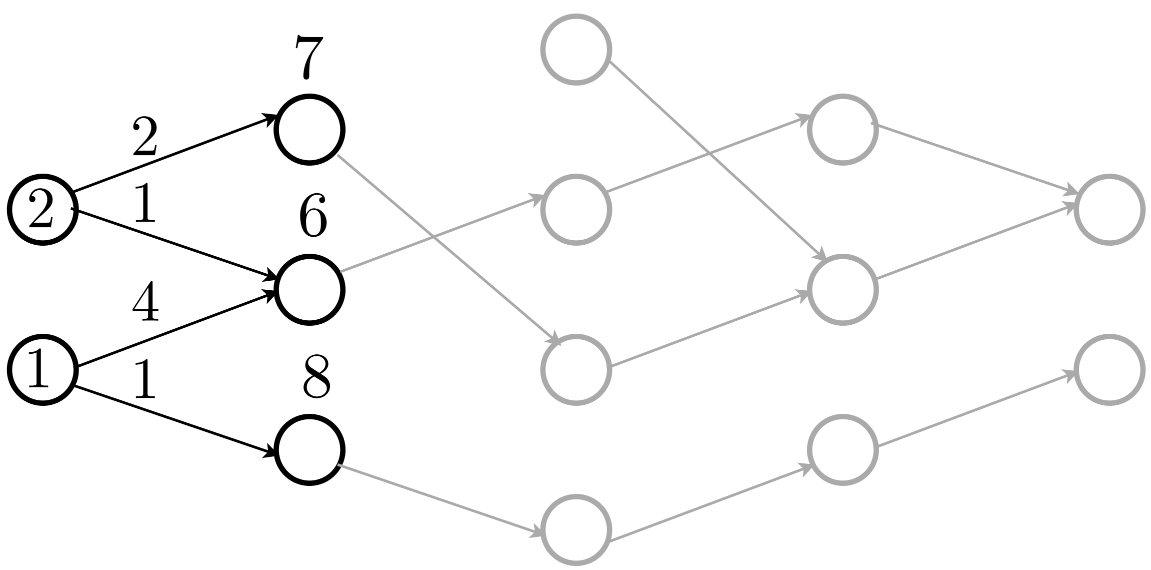
Example



Iteration 3 - Stage 1

State	Cost-to-go
1	$\min\{4+6, 3+5\} = 8$
2	$\min\{0+6, 1+6\} = 6$
3	$\min\{1+7, 3+6, 1+6\} = 7$

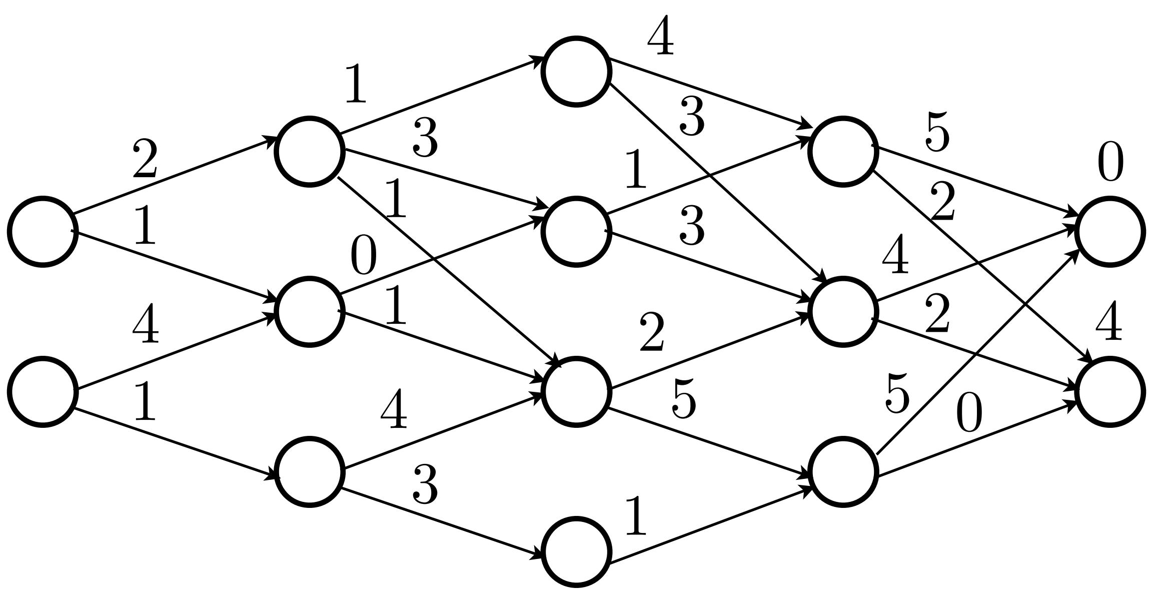
Example



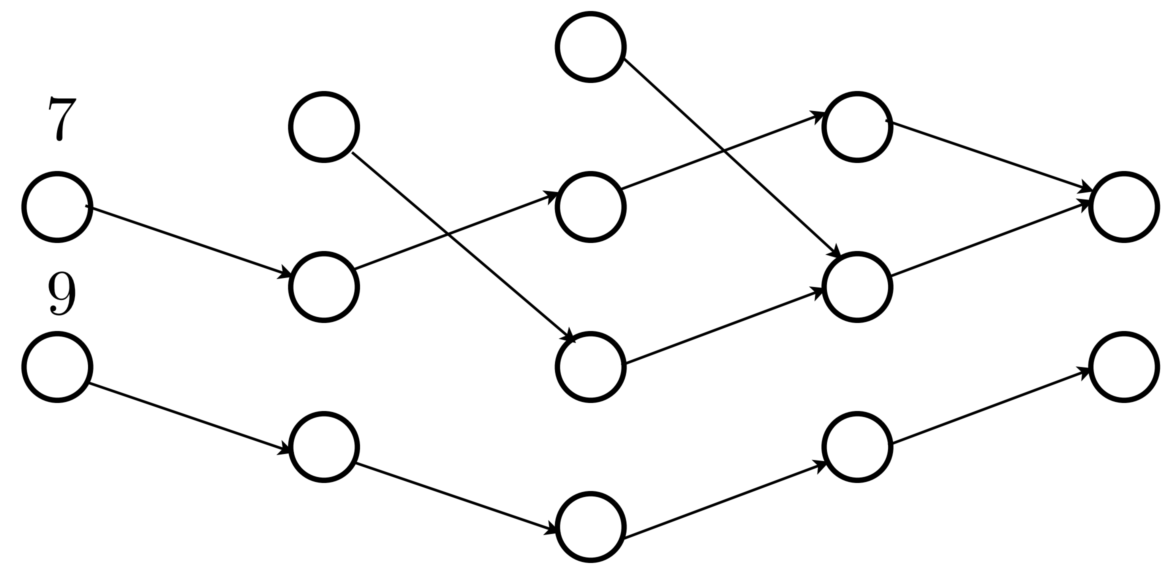
Iteration 4 - Stage 0

State	Cost-to-go
1	$\min\{1+8, 4+6\} = 9$
2	$\min\{2+7, 1+6\} = 7$

Optimal policy and optimal paths



Initial transition diagram



Optimal policy

Optimal policy

LiveScript/Lecture1/dp.mlx

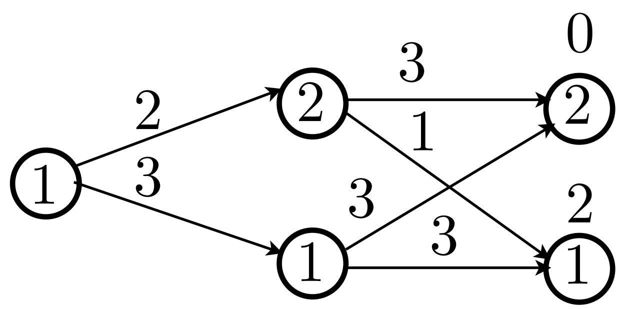
- While running the DP algorithm, for each state at each stage a decision is made to compute the cost-to-go. That decision is precisely the decision specified by the optimal policy.

Optimal path

- For a given initial state, follow the arrows leading to the final stage. This is the optimal path.
- The cost-to-go at stage 0 of that initial state coincides with the cost of the optimal path.

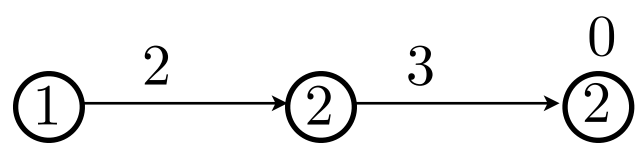
Non-uniqueness

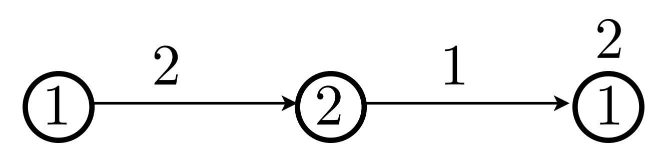
The optimal policy and the optimal paths may not be unique



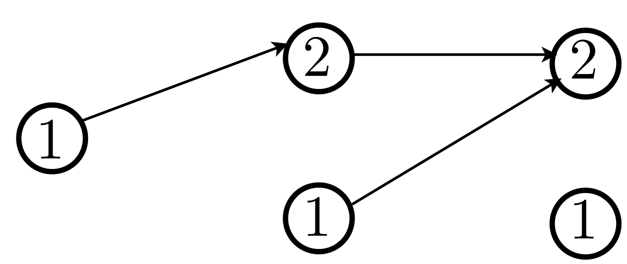
Stage 1, state 2 - both decisions have the same cost 3.

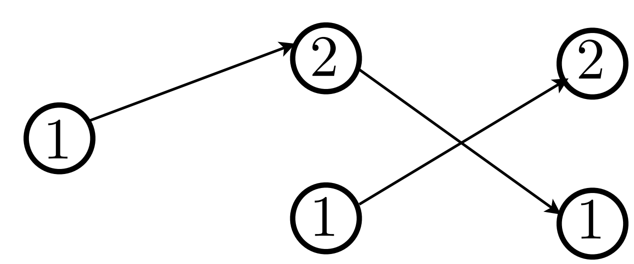
Two optimal paths





Two optimal policies





If more than one option has the same cost while running the dynamic programming algorithm, simply pick one of the options. In the end, <u>one</u> optimal policy is obtained (while several may be optimal).

Inventory control

Controlling the supply of one product

- Dynamic model $x_{k+1} = \max\{x_k + u_k - d_k, 0\}$

  $u_k \in \{0, 1, \ldots, N - x_k\}$

- Cost $\sum_{k=0}^{h-1} g_k(x_k, u_k) + g_h(x_h)$

$$g_k(x_k, u_k) = (c_1(x_k) + cu_k + c_{\mathrm{tr}}\|u_k\|_0) - p\min\{d_k, x_k + u_k\}$$

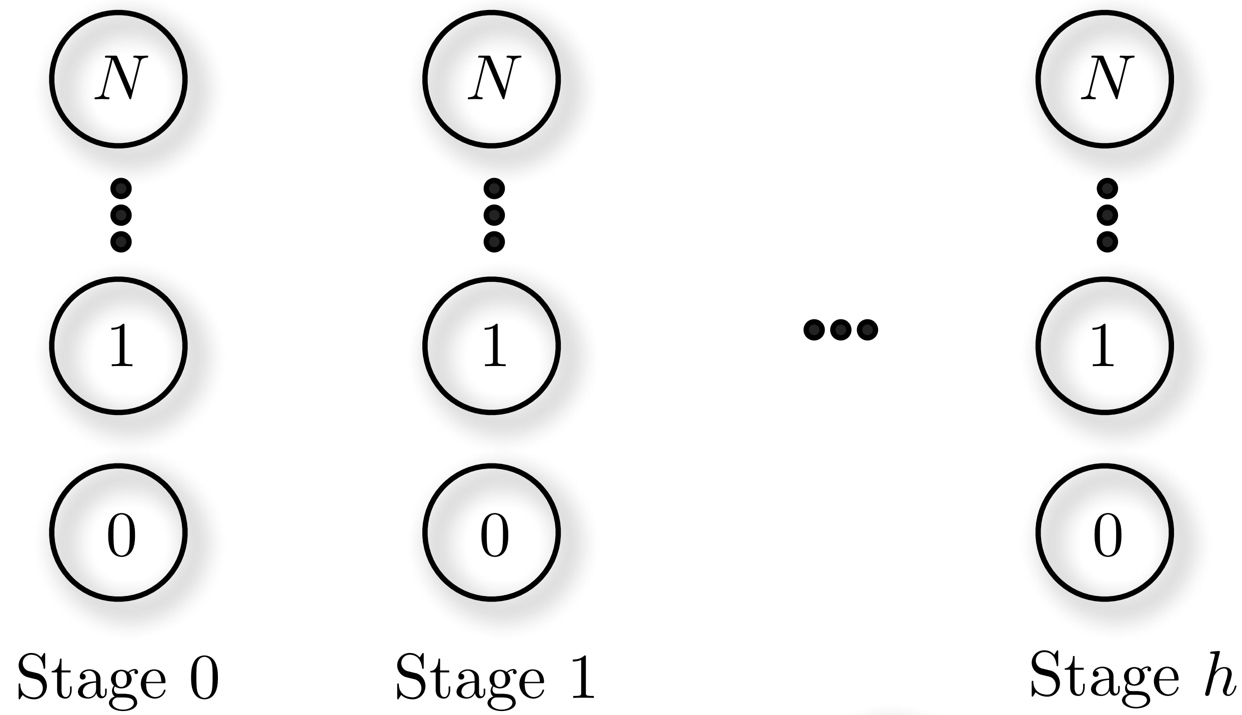
x_k number of items
N capacity
u_k supply
d_k demand
g_h terminal cost
c_1 storage cost
p selling price
c purchase price
c_{tr} transportation price

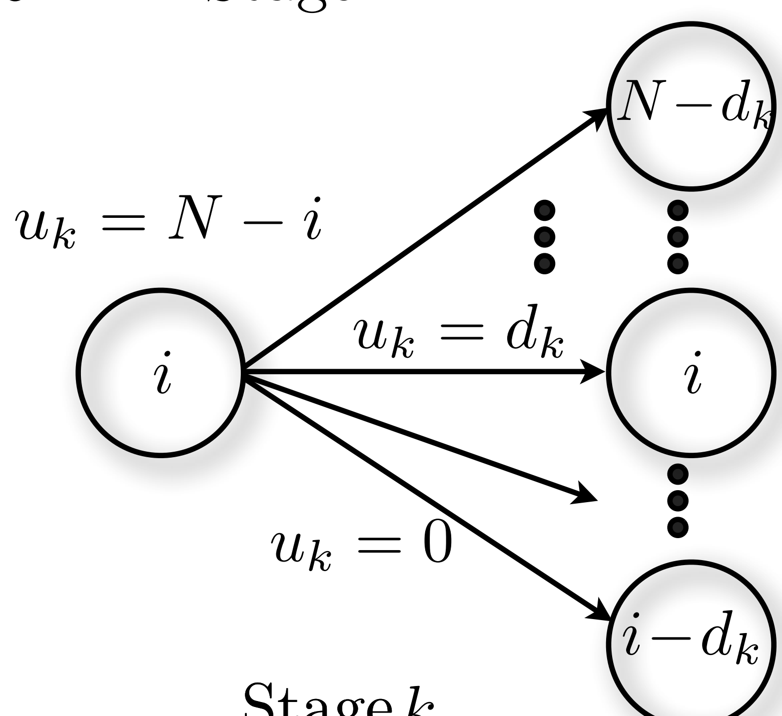
$$\|u_k\|_0 = \begin{cases} 0 \text{ if } u_k = 0 \\ 1 \text{ if } u_k \neq 0 \end{cases}$$

Formulation

Transition diagram

- circles at each stage indicate number of items, supplies determine transitions





Inventory control

$$x_{k+1} = \max\{x_k + u_k - d_k, 0\}$$

$$u_k \in \{0, 1, \ldots, N - x_k\}$$

x_k number of items

u_k supply

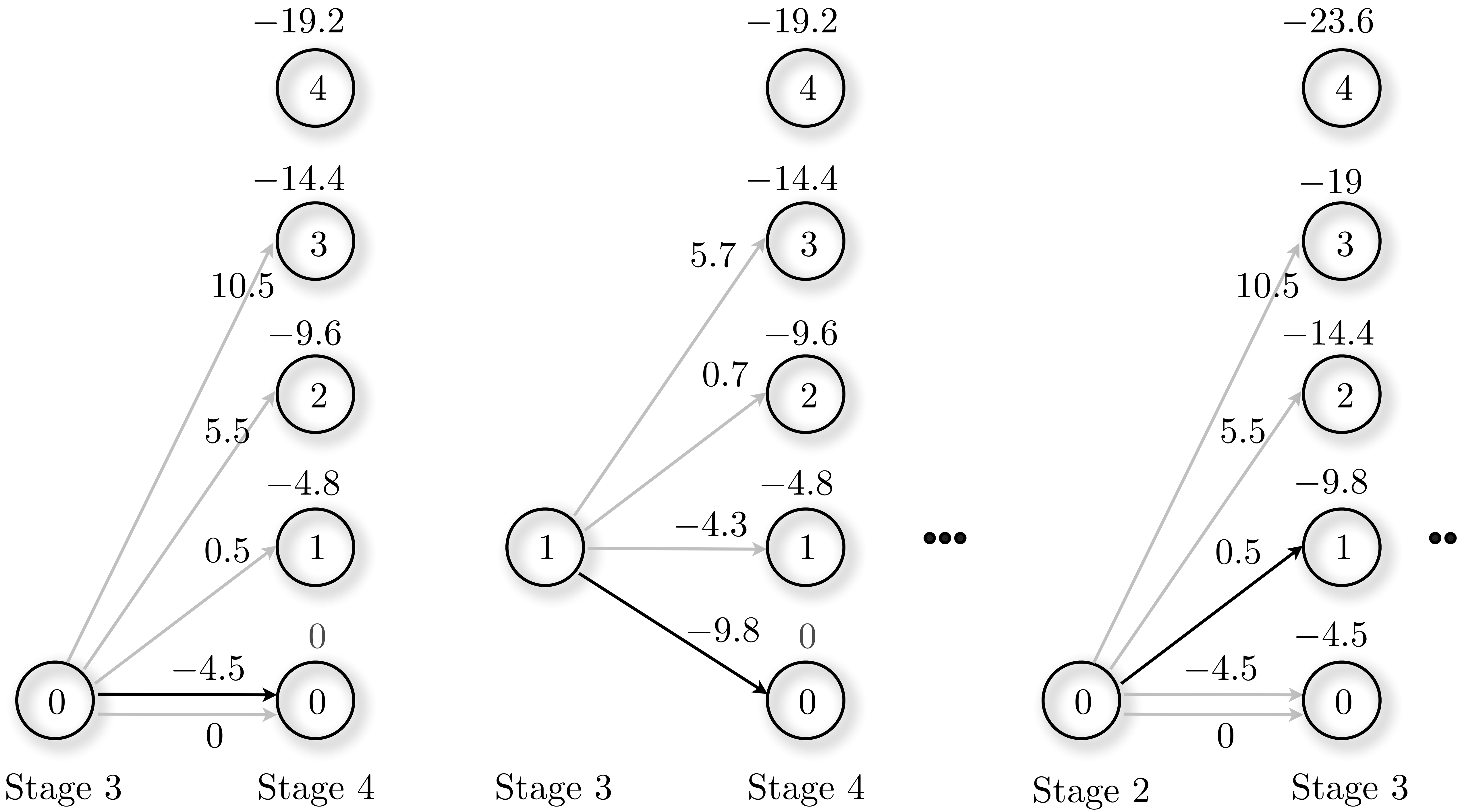
$$\sum_{k=0}^{h-1} g_k(x_k, u_k) + g_h(x_h)$$

$$g_k(x_k, u_k) = c_1(x_k) + cu_k + c_{\text{tr}}\|u_k\|_0 - p\min\{d_k, x_k + u_k\}$$

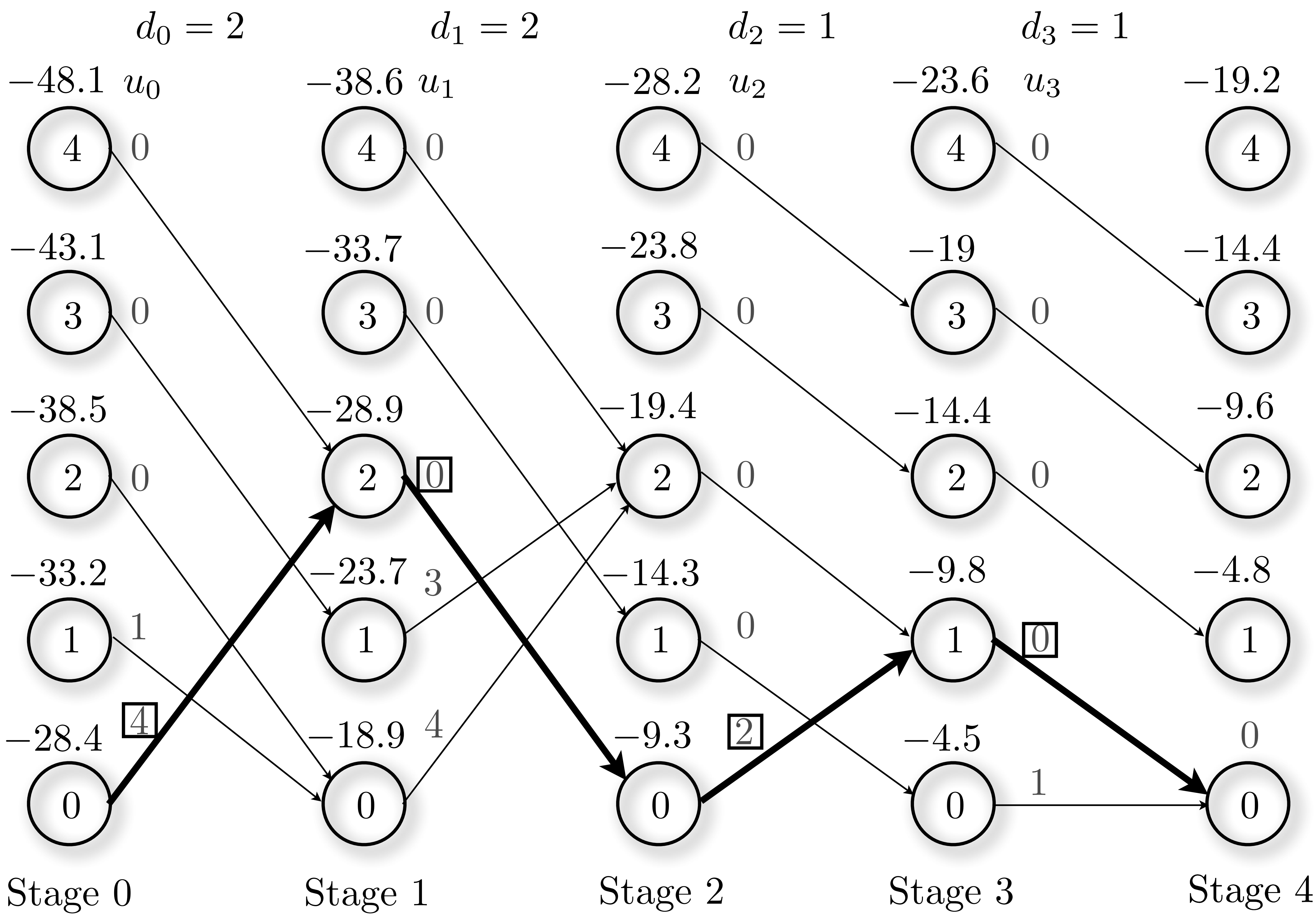
number of stages	$h = 4$
capacity	$N = 4$
demand	$d_0 = d_1 = 2,\ d_2 = d_3 = 1$
selling price	$p = 10$
purchase price	$c = 5$
transportation price	$c_{\text{tr}} = 0.5$
storage cost	$c_1(i) = 0.2i,\ i \in \{0, \ldots, N\}$
terminal cost	$g_4(i) = -r_{i+1},\ i \in \{0, \ldots, 4\},\ \ r = [0\ 4.8\ 9.6\ 14.4\ 19.2]$

What are the optimal supplies for a zero initial inventory?

Some iterations



Final policy and optimal path



Cost for a zero initial inventory -28.4

LiveScript/Lecture1/inventorycontrol.mlx

Richard E. Bellman

Historical note

- Dynamic programming was proposed in the 1940s by Richard E. Bellman

'I was intrigued by dynamic programming. It was clear to me that there was a good deal of good analysis there. Furthermore, I could see many applications. It was a clear choice. I could either be a traditional intellectual, or a modern intellectual using the results of my research for the problems of contemporary society.'
R. E. Bellman (1920-1984)

Concluding remarks

Summary

- Optimal control: determine several optimal decisions over time and as a function of the state.
- Optimal control problems: three classes (discrete optimization, stage-decision, continuous-time control), many applications.
- Dynamic programming: optimal decisions computed from the end to the initial stage.

After this lecture, you should be able to:

- Apply the dynamic programming algorithm.
- Solve deterministic inventory control problems.

Additional live scripts

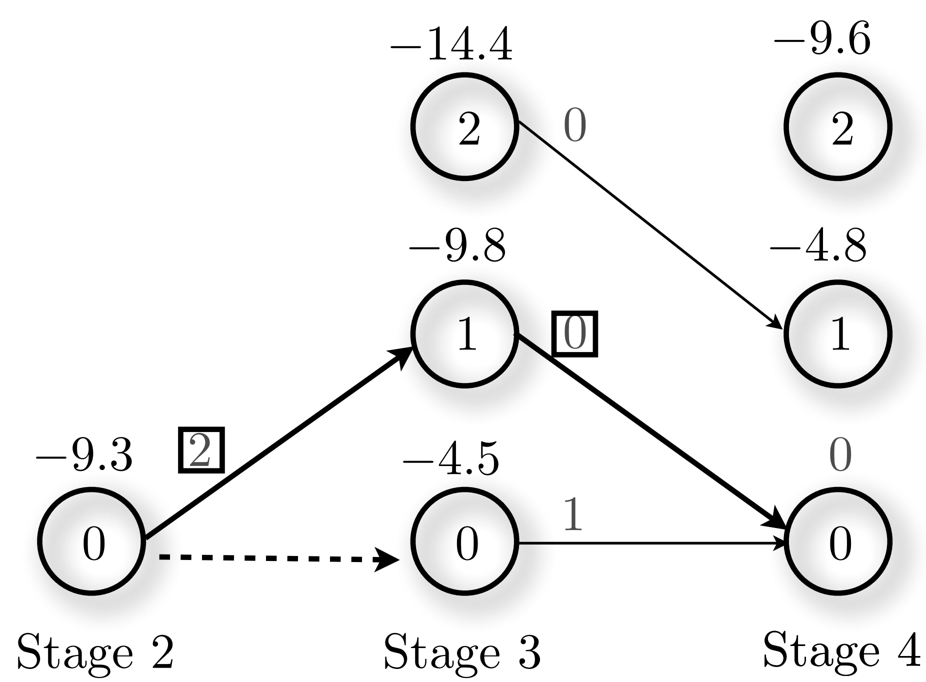
- Solving state estimation problems for a hidden Markov chain with the dynamic programming algorithm (Viterbi algorithm) LiveScript/Lecture1/viterbi.mlx

Appendix A

Inventory control with uncertainty

Coping with disturbances

In the context of the example of slides 40-44



What if the demand is $d_2 = 2$ instead of the expected $d_2 = 1$?

- The state at stage 3 is then $x_3 = 0$ instead of $x_3 = 1$.
- Open loop: blindly pick $u_3 = 0$ as initially planned.
- Closed-loop (using DP policy): pick $u_3 = 1$.

Open loop vs closed loop

Costs

- Open loop $g_0(0,4) + g_1(2,0) + g_2(0,2) + g_3(0,\boxed{0}) + g_4(0)$

$$0.5 \quad - \quad 19.6 \quad + \quad -9.5 \quad + \quad 0 \quad + \quad 0 = \quad -28.6$$

- Closed loop $g_0(0,4) + g_1(2,0) + g_2(0,2) + g_3(0,\boxed{1}) + g_4(0)$

$$0.5 \quad - \quad 19.6 \quad + \quad -9.5 \quad + \quad -4.5 \quad + \quad 0 = \quad -33.1$$

Expected cost if $\mathrm{Prob}[d_2 = 1] = 0.5,\ \mathrm{Prob}[d_2 = 2] = 0.5$

- Open loop $0.5 \times (-28.4) + 0.5 \times (-28.6) = -28.5$

- Closed loop $0.5 \times (-28.4) + 0.5 \times (-33.1) = -30.75$

Appendix B

References and applications

References

Textbooks

[1] D. Bertsekas, *Dynamic Programming and Optimal Control,* 3rd edition, Athena Scientific, Vol I and II, 2005

[2] A. E. Bryson and Y. C. Ho, *Applied Optimal Control, CRC Press, 1975*

[3] D. Liberzon, *Calculus of Variations and optimal control*, Princeton University Press, 2012,

[4] M. Athans and P.L. Falb. *Optimal control*. McGraw Hill, New York, 1966. Reprint by Dover in 2006.

[5] B. D. O. Anderson and J. B. Moore. *Optimal control: Linear quadratic methods*. Prentice Hall, New Jersey, 1990. Reprinted by Dover in 2007

[6] F. L. Lewis, Draguna Vrabie, Vassilis L. Syrmos, *Optimal control,* 3rd edition, John Wiley & Sons, 2012

[7] D. E. Kirk, *Optimal control theory: an introduction,* Dover books on electrical engineering, 2004

[8] R. Bellman. *Dynamic programming*. Princeton University, 1957.

[9] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, New Jersey, 1996.

[10] G. Chen, G. Chen, S-H, Hsu. *Linear Stochastic Control Systems*, CRC Press Book, 1995.

[11] M. H. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall, 1977.

[12] P. Whittle, *Optimal control: basics and beyond.* John Wiley and Sons Ltd, 1996.

[13] D. Bertsekas and S. E. Shreve, *Stochastic optimal control: the discrete-time case,* Athena Scientific, 1996

Numerical methods

[14] John Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*, SIAM, 2010

References

Applications

Aerospace

[A1] A. E. Bryson, JR.. *Applications of optimal control theory in aerospace engineering* Journal of Spacecraft and Rockets, Vol. 4, No. 5 (1967), pp. 545-553

[A2] J. M. Longuski, J. J. Guzmán, J.E. Prussing, *Optimal control with aerospace applications, Springer, 2014*

Biomedicine and sequential alignment of DNA

[A3] G .W. Swan. *Applications of optimal control theory in biomedicine*. Marcel Dekker, New York, 1984.

[A4] Eddy, S. R., *What is dynamic programming?*, Nature Biotechnology, 22, 909–910 (2004)

[A5] M. G. Neubert. *Marine reserves and optimal harvesting.* Ecology Letters, 6:843-849, 2003.

Power Systems

[A6] G. S. Christensen, M. E. El-Hawary, and S. A. Soliman. *Optimal control applications in electric power systems*. Plenum press, New York, 1987.

References

Applications

Operational research and inventory control

[A7] M. H. Davis, *Markov models and optimization*, Chapman and Hall/CRC, 1993.

[A8] A. Bensoussan, *Dynamic programming and inventory control*, IOS press, 2011.

Finance and economics

[A9] S.P. Sethi and G.L. Thompson. *Optimal control theory: applications to management science and economics*, Springer, New York, 2nd edition, 2005.

Computer science and scheduling problems

[A10] Art Lew and Holger Mauch, *Dynamic Programming: A Computational Tool*, Springer Verlag, 2007

Automotive

[A11] Bram de Jager, Thijs van Keulen, John Kessels *Optimal Control of Hybrid Vehicles*, Springer, 2013

References

History

[H1] A. E. Bryson Jr. *Optimal control - 1950-1985*. IEEE Control Systems Magazine, 1996.

[H2] Héctor J. Sussmann and Jan C. Willems, *300 Years Of Optimal Control: From The Brachystochrone To The Maximum Principle* (1997)

[H3] Richard Ernest Bellman *Eye of the Hurricane: An Autobiography, 1984*

[H4] Stuart Dreyfus (2002). *Richard Bellman on the Birth of Dynamic Programming*. In: Operations Research. Vol. 50, No. 1, Jan–Feb 2002, pp. 48–51

Seminal papers

[S1] Anonymous, "*Letter sent to Charles Montague, President of the Royal Society, where two mathematical problems proposed by the celebrated Johann Bernoulli are solved*" . Acta Eruditorum Lipsi", (1697) 223.

[S2] Bellman, Richard. *The theory of dynamic programming*. Bull. Amer. Math. Soc. 60 (1954), no. 6, 503--515

[S3] R. E. Kalman. *Contributions to the theory of optimal control*. Bol. Soc. Mat. Mexicana, 5:102-119, 1960. Reprint in Control Theory: Twenty-Five Seminal Papers (1931-1981), T. Basar, editor, IEEE Press, New York, 2001, pages 149-166.

[S4] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. The Mathematical Theory of Optimal Processes. Interscience, New York, 1962.

[S5] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis. *State-space solutions to standard H2 and Hinf control problems.* IEEE Trans. Automat. Control 34:831-847, 1989