1. Controllable dynamical systems in discrete time

1.1 A _discrete-time controllable dynamical system_ with _state-space_ S and _action-space_ A is a map

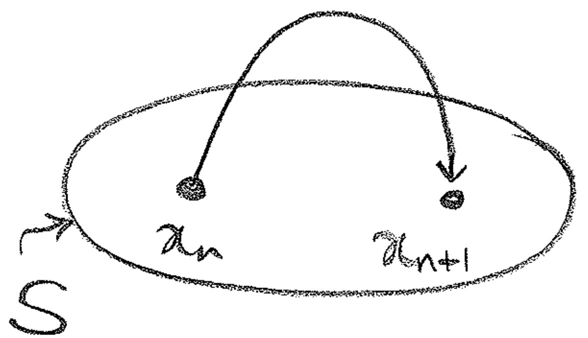
$$f : \mathbb{Z}^+ \times S \times A \longrightarrow S.$$

A _control_ is a map

$$u : \mathbb{Z}^+ \longrightarrow A.$$

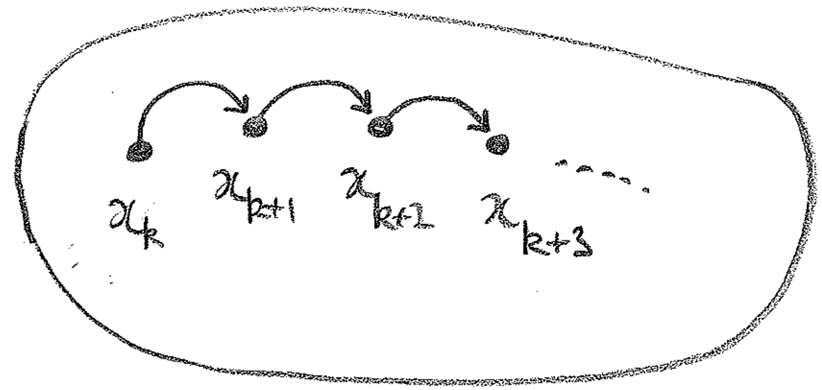
The _controlled sequence_ for control u starting from (k, x) is given by

$$x_k = x, \qquad x_{n+1} = f(n, x_n, u_n), \qquad n \geq k.$$



$$x_{n+1} = f(n, x_n, u_n)$$

plant equation



- The case where

$$f : S \times A \longrightarrow S$$

is called <u>time-homogeneous</u>.

- Sometimes there is a different set of actions A_x available in each state x. Then we replace $S \times A$ in the preceding by $\bigcup_{x \in S} \{x\} \times A_x$. This does not change much in the theory.

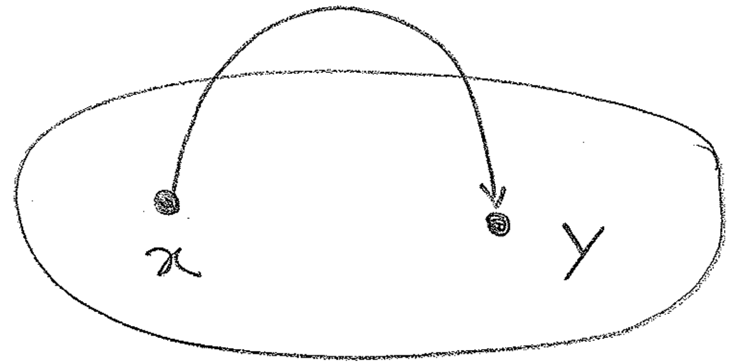
1.2 A discrete-time stochastic controllable dynamical system with state-space S and action-space A is a map

$$P : \mathbb{Z}^+ \times S \times A \longrightarrow \mathrm{Prob}(S).$$

The set of all probability measures on S.

We shall assume that S is countable for now, in which case

$$\mathrm{Prob}(S) = \{ (p_x : x \in S) : p_x \geq 0, \ \sum_x p_x = 1 \}.$$



Given that we are in state x at time n and choose action a, then we go to Y at time $n+1$, where

$$Y = y \quad \text{with probability} \quad P(n, x, a)_y .$$

We write, for a function $F : \mathbb{Z}^+ \times S \longrightarrow \mathbb{R}^+$,

$$PF(n, x, a) = \mathbb{E}\big(F(n+1, Y)\big) = \sum_y P(n, x, a)_y \, F(n+1, y) .$$

In the <u>time-homogeneous</u> case

$$P : S \times A \longrightarrow \mathrm{Prob}(S)$$

and, for a function $F : S \to \mathbb{R}^+$, we write

$$PF(x, a) = \mathbb{E}(F(Y)) = \sum_y P(x, a)_y \, F(y).$$

A <u>Markov control</u> is a map

$$u : \mathbb{Z}^+ \times S \longrightarrow A.$$

We call a map $u : S \longrightarrow A$ a <u>stationary Markov control</u>.

Define

$$P^u(n) = \left(p^u_{xy}(n) : x, y \in S \right), \qquad p^u_{xy}(n) = P(n, x, u_n(x))_y .$$

Then $\left(P^u(n) : n \in \mathbb{Z}^+ \right)$ is a time-dependent stochastic matrix. Time-dependence disappears for a time-homogeneous system with a stationary Markov control.

Say that $(X_n)_{n \geq k}$ is a <u>controlled process</u> of the system P, using control u, and starting from (k, x), if

$$\mathbb{P}(X_k = x_k, X_{k+1} = x_{k+1}, \ldots, X_n = x_n)$$

$$= \delta_{x x_k} P(k, x_k, u_k(x_k))_{x_{k+1}} \cdots P(n-1, x_{n-1}, u_{n-1}(x_{n-1}))_{x_n}$$

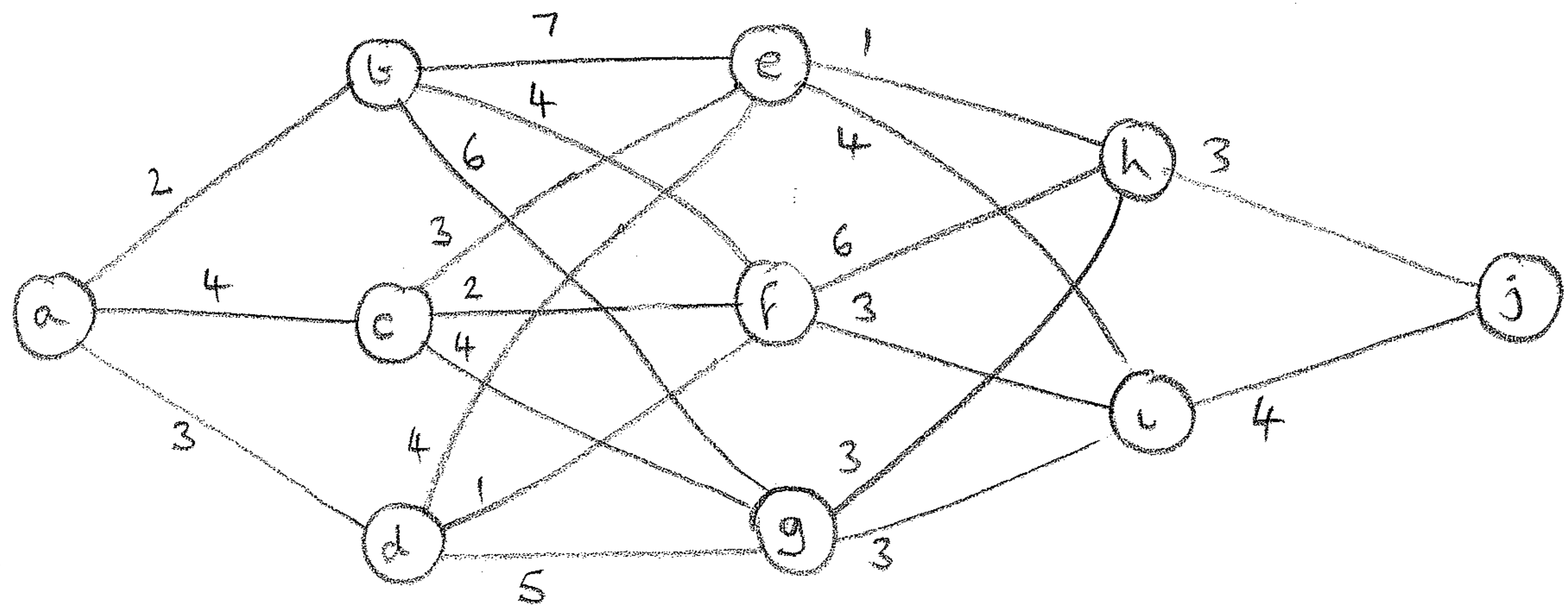
for all $n \geq k$ and all $x_k, \ldots, x_n \in S$.

This controlled process is a <u>time-inhomogeneous Markov chain</u>.

For a time-homogeneous system P and a stationary Markov control u, and for $k = 0$,

$$(X_n)_{n \geq 0} \sim \text{Markov}(\delta_x, P^u), \quad P^u_{xy} = P(x, u(x))_y$$

1.3 Example – finding the shortest path



What is the shortest path from a to j?

Putting the problem in the general framework ...

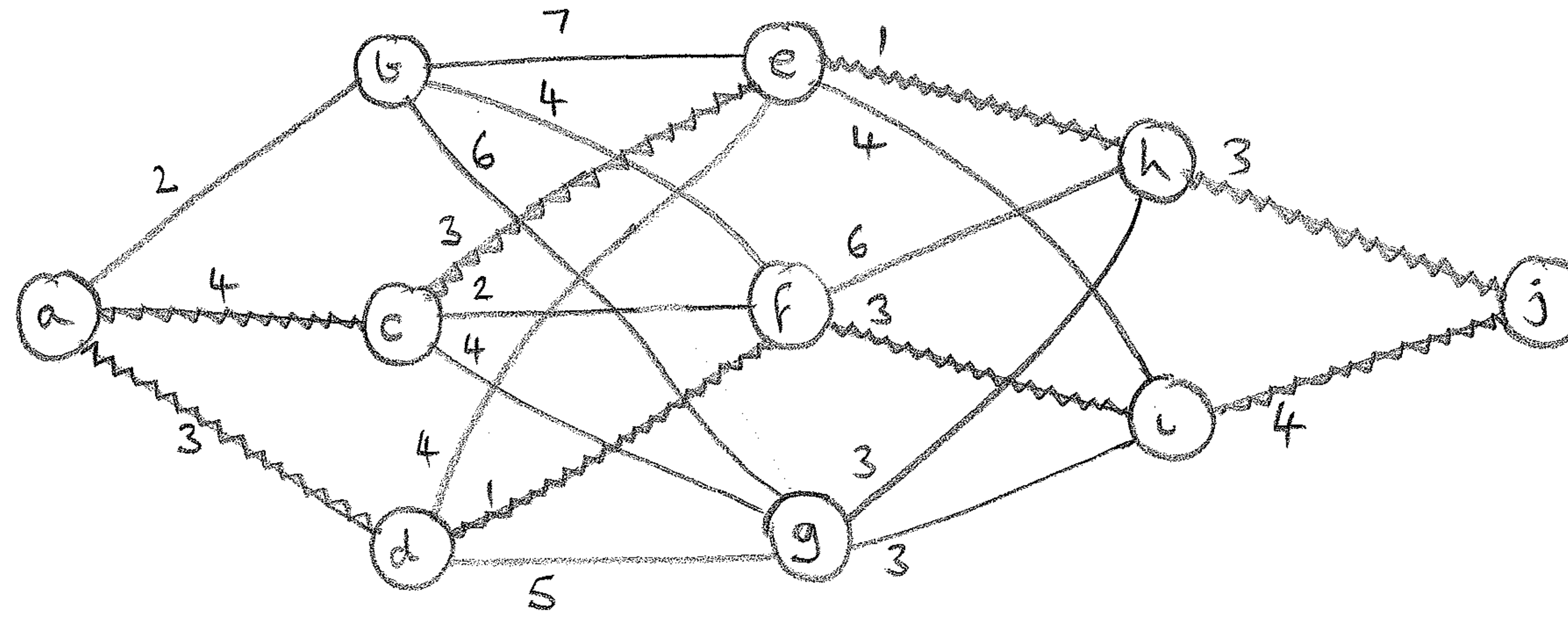
(not very helpful is solving the problem, but good exercise in understanding the framework)

$S = \{a, b, \ldots, j\}$, A_x = set of neighbours of x

We have to minimize $\sum_{n=0}^{\tau-1} c(x_n, x_{n+1})$

over controlled sequences $(x_n)_{n \geq 0}$ with $x_0 = a$, $x_\tau = j$,

where $c(x, a)$ = distance from x to a (directly)



General method : work backwards

1.4 Example - impatient gambling

You have £1 and a limitless overdraft facility. You can place bets on the outcome of a fair coin toss of any size you wish. If you predict the outcome correctly you win double your stake, otherwise you lose it. You wish to increase your wealth to £10 as quickly as possible. Find a strategy which achieves this aim in two bets on average. Show this cannot be improved upon.

In the general framework ...

$$S = \mathbb{Z}, \qquad A = \mathbb{Z}^+$$

$$P(x,a)_{x+a} = P(x,a)_{x-a} = \frac{1}{2}$$

We will restrict to stationary Markov strategies

$$u : S \longrightarrow A.$$

The controlled process $(X_n)_{n \geq 0}$ is then a Markov chain, starting from 1, with transition matrix

$$P^u = (p^u_{xy} : x, y \in S), \qquad p^u_{x,x+u(x)} = p^u_{x,x-u(x)} = \frac{1}{2}$$

The problem is to

$$\underline{\text{minimize}} \quad \mathbb{E}^u(T) \qquad \underline{\text{over}} \quad u,$$

where

$$T = \inf\{n \geq 0 : X_n \geq 10\}$$

the first time we hit 10.