MDP1.pdf

①

Modelling Probabilistic Dynamic Progr[illegible]

[illegible] we define:

t – stage

i – State at beginning at stage t

x – action (in our ~~[illegible]~~ previous egs, x – how many machies to produce etc).

Recurrence function

$$f_t(i) = \max_x \left\{ \begin{matrix} \text{Expected reward} \\ \text{during stage } t \\ \text{given } i, x \end{matrix} + \underbrace{[illegible]}_{\substack{\text{prob. state at beginning of} \\ \text{next stage is } j}} \; \underbrace{[illegible]}_{\substack{\text{max. in next stage when beginning} \\ \text{with state } j.}} \right\}$$

Oil Company Example

(Winston Chapter 21.4 example 5 Pg 1080)

Suno Oil $\boxed{D \text{ dollars to allocate}}$

$\boxed{\text{Site } 1}$ $\boxed{\text{Site } 2}$ - - - $\boxed{\text{Site } T}$

- Investing X dollars
- Prob. oil be found $q_t(x)$
- If oil is found gets r_t

Objective:

Maximise the expected value of all oil found on sites $1, 2, \cdots, T$.

Define:

$f_t(d)$ to be the max. expected value of the oil that can be found on sites $t, t+1, \ldots, T$ if d dollars are available to allocate to sites $t, t+1, \ldots, T$.

Assumption $q_T(x)$ non-decreasing

Stage T

$$f_T(d) = r_T \underbrace{q_T(d)}_{\text{if oil is found}} + \underbrace{(1-q_T(d))}_{\text{if oil is not found}}\, 0$$

reward reward

Stage $t < T$

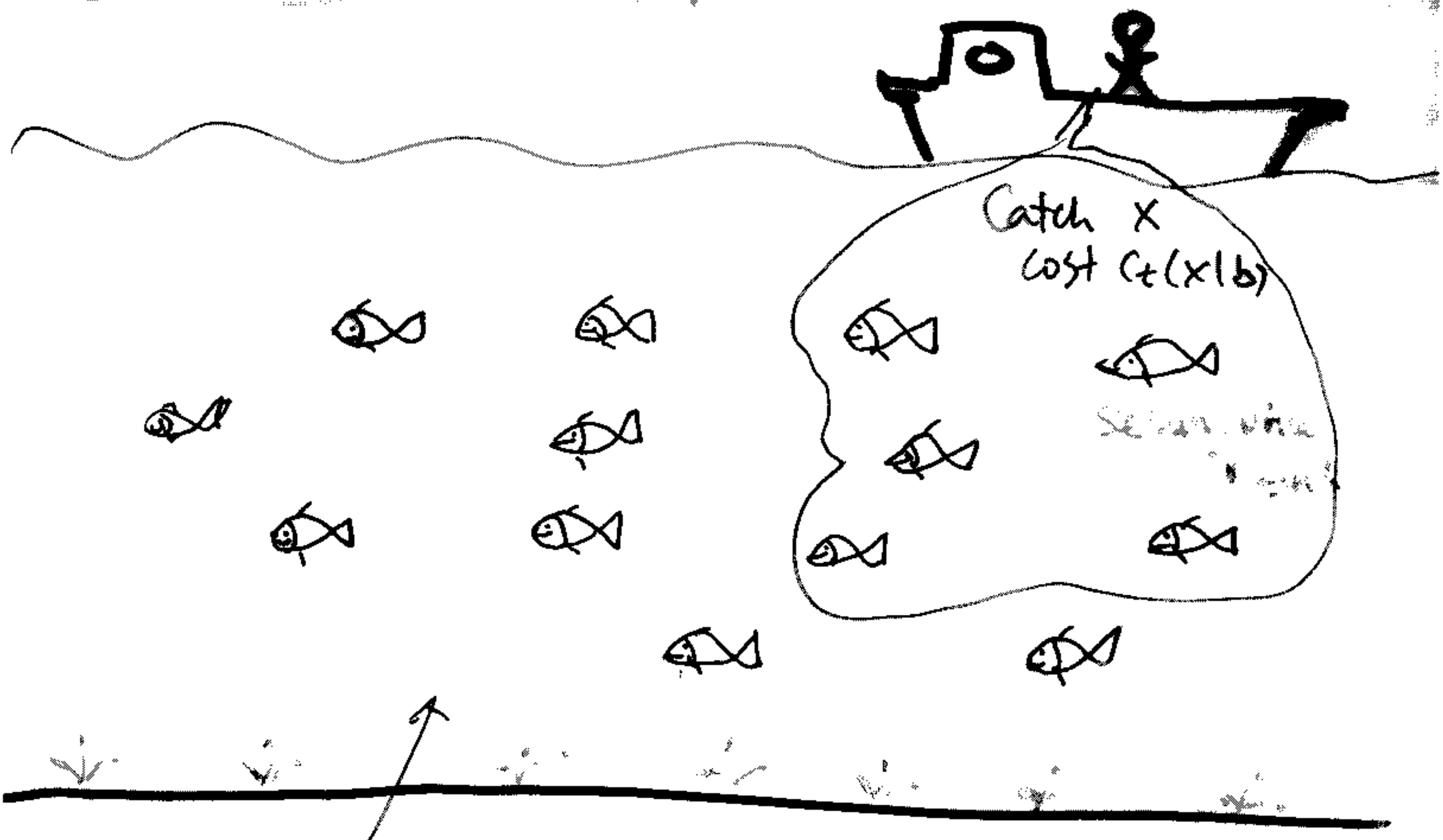
$$f_t(d) = \max_x \left\{ r_t\, q_t(x) + f_{t+1}(d-x) \right\}$$

for $0 \le x \le d$

Working backward until $\boxed{f_1(D)}$

Fishing Problem

(Winston 21-4 Ex 6 p1081)



Remaining bass: Reproduction rate / year

a multiple of d with probibility $q(d)$
(e.g. 30% chance at rate of $d=2$
$\Rightarrow q(2) = 0.3$).

The problem
- Fish caught <u>once</u> per year.
- <u>Ten</u> year span

- Determine catching strategy that <u>maximises profit</u> over next ten years
- 10000 bass in lake in beginning of year 1.

Define:

$f_t(b)$ – max. expected net profit than can be earning during years $t, t+1, \ldots, 10$ if the lake contains b bass at the beginning of year t.

$$f_{10}(b) = \max_{x} \{ x P_{10} - C_{10}(x \mid b) \}$$

cost

action: no. of bass to catch

$0 \le x \le b$

Recursive formula for $t < 10$

Use this to work backwards until $f_1(10{,}000)$

$$f_t(b) = \max_{x} \Big\{ x P_t - C_t(x \mid b) + \sum_{d} q(d) f_{t+1}(d(b-x)) \Big\}$$

rate of reproduct^{n}

probability that rate of reproduction is d

$$\boxed{0 \le x \le b}$$

Department Store Cash Management Example

[illegible] the parameter [illegible]

(Winston Ex 8 Chap 12.4, p1082)

- Daily demand of cash at store:
  Random variable D with $P(D=d) = p(d)$

$\boxed{\text{9:00am}}$ \$ Cash left from yesterday

+ \$ Withdrawl ($K each transact#) $\boxed{\text{or}}$ − \$ Deposit ($K each transaction)

− \# demand

= $\boxed{???}$

$\boxed{\text{5pm}}$

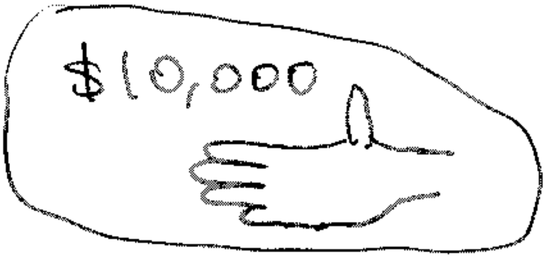
If $\boxed{???}$ is negative

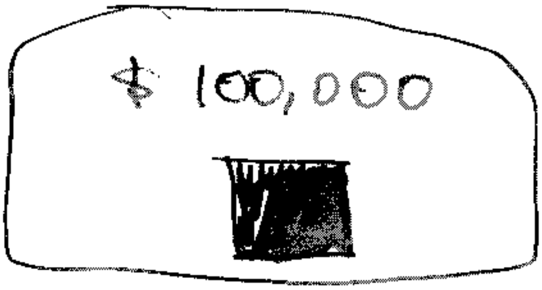
⇒ Shortage cost \$ S / dollar

If $\boxed{???}$ is positive

⇒ penalty cost (due to loss in interest) \$ i / dollar

At beginning of Day 1,





Planning period: 30 days

Objective: minimise the expected cost of filling the store's cash needs.

阶段 — Days

状态 — (c, b) at beginning of day $\$C$ in hand and $\$b$ in bank.

Define

$f_t(c, b)$ - min expected cost incurred by the store during days $t, t+1, \ldots, 30$, given that a beginning of day t, the store has c dollars in hand and b dollars in bank.

$$\boxed{\text{Total Cost} = \text{Cost of transact}^{n} + \text{Shortage Cost} + \text{penalty cost}}$$

Cost of transaction

$$K \delta(x), \quad \text{for } \delta(x) = \begin{cases} 1, & \text{if } x \neq 0 \\ 0, & \text{o.w.} \end{cases}$$

K: Cost per transaction

x: amt deposit / withdraw from bank

Shortage cost

$$\sum_{\underbrace{d \geq c+x}_{\text{if demand} \geq \text{supply}}} \underbrace{p(d)}_{\text{prob demand is } d} \underbrace{(d-c-x)^{+}}_{\text{excess in demand}} \underbrace{s}_{\text{shortage cost per dollar}}$$

Penalty cost

$$\sum_{\underbrace{d \leq c+x}_{\text{if supply} \geq \text{demand}}} \underbrace{p(d)}_{\text{prob. demand is } d} \underbrace{(c+x-d)^{+}}_{\text{excess in supply}} \underbrace{i}_{\text{loss in interest per dollar}}$$

At Stage 30

$$f_{30}(c,b) = \min_{x} \Big\{ K\delta(x) + \sum_{d \le c+x} p(d)(c+x-d)^{+} i$$

$$+ \sum_{d \ge c+x} p(d)\,(d-c-x)^{+} s \Big\}$$

$x \le b$ ← cannot withdraw more than b dollars ∵ that's all in bank

$-x \le c$ ← cannot deposit more than c dollars ∵ that's all in store.

$\Rightarrow \boxed{-c \le x \le b}$

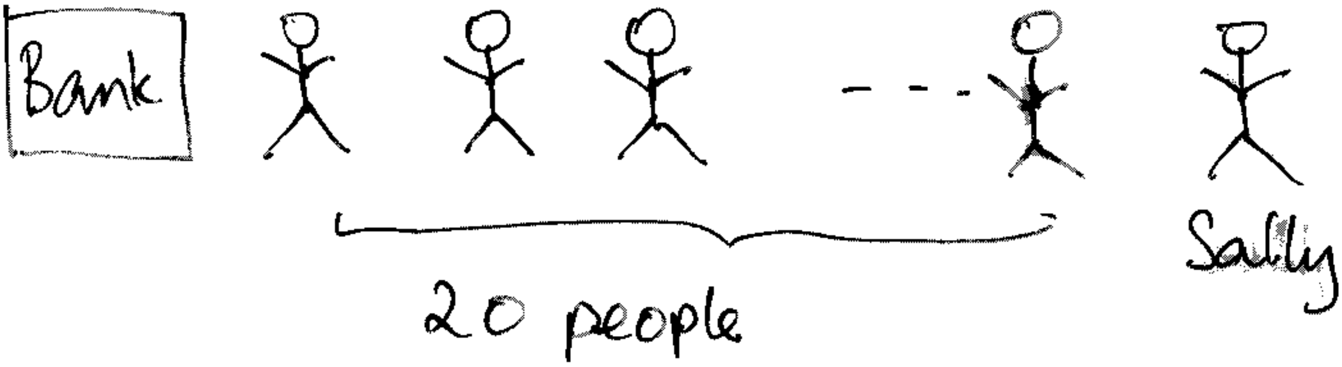
Recursive Formula

$$f_t(c,b) = \max_{-c \le x \le b} \Big\{ K\delta(x) + \sum_{d \le c+x} p(d)(c+x-d)^{+} i$$

$$+ \sum_{d \ge c+x} p(d)(d-c-x)^{+} s + \sum_{d} p(d)\, \underline{f_{t+1}(c+x-d, b-x)}$$

weighed by prob. demand = d

Working backwards until $\boxed{f_1(10{,}000,\ 100{,}000)}$



- Sally has a 30-minute break
- Reward: r if by end of break, she gets money
- Cost: c (Sally hates waiting)
- Rate of service (per minute):
  If n people in the line, prob x people being served or quit

$$\boxed{p(x \mid n)}$$

Objective Maximise expected reward:

$$\boxed{\text{reward} - \text{waiting costs}}$$

: No. of minutes left
(Sally has a 30-mins break)

: No. of people ahead of her in queue.

$\boxed{\text{Define:}}$ $f_t(n)$ – maximum expected net reward that Sally can receive from time t to the end of her lunch break if at time t, n people are ahead of her.

$\boxed{\text{Let}}$ $t = 0$ now
$t = 30$ end of planning period

Last stage $t = 29$ – beginning of the last minute of the problem.

At last stage, $t=29$

$$f_{29}(n) = \max \begin{cases} 0 & \text{(Leave)} \\ \underset{\text{reward}}{r}\, \underbrace{p(n|n)}_{\substack{\text{prob. everyone} \\ \text{in front of Sal disappears.}}} - \underset{\text{waiting cost}}{c} & \text{(Stay)} \end{cases}$$

For $t<29$

$$f_t(n) = \max \begin{cases} 0 & \text{(Leave)} \\ \underbrace{r\,p(n|n)}_{\substack{\text{expected reward if} \\ \text{she gets served}}} - \underset{\text{waiting cost}}{c} + \underbrace{\sum_{k<n} p(k|n)\, f_{t+1}(n-k)}_{\substack{\text{if she didn't get served, what} \\ \text{happens in next stage.}}} \end{cases}$$

Working this way until $f_0(20)$.

$$\underset{\substack{\text{Summing over all} \\ k<n.}}{\sum_{k<n}} \left(\underbrace{p(k|n)}_{\substack{\text{prob. } k \\ \text{people getting} \\ \text{served or quit.}}} \; f_{t+1}\underset{\substack{n-k \text{ people} \\ \text{in front of Sal for} \\ \text{next stage.}}}{(n-k)} \right)$$