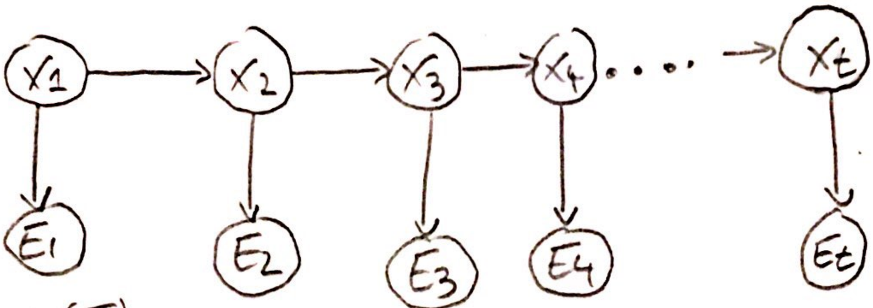
# COMP 424 A4

Nabil Choudhury
260622155

Q1) States: {High, Low} = X
Observations: {Buy, Sell, keep} = E

a)




Initial Prob($\pi$)

| X | P(X) |
|---|---|
| high | 0.6 |
| low | 0.4 |

Transition Prob. $P(X_{t+1} \mid X_t)$

| $X_t$ | $X_{t+1}$ high | low |
|---|---|---|
| high | 0.3 | 0.7 |
| low | 0.6 | 0.4 |

Emission Probabilities $P(E \mid X)$

| X | E: Buy | Sell | keep |
|---|---|---|---|
| high | 0.1 | 0.6 | 0.3 |
| low | 0.5 | 0.2 | 0.3 |

b) P({Buy, keep, Sell})

| | $E_1$ = Buy | $E_2$ = keep | $E_3$ = sell |
|---|---|---|---|
| X=high | 0.06 | 0.0414 | 0.020628 |
| X=low | 0.2 | 0.0366 | 0.008724 |

$\alpha_1(1) = P(E_1 = \text{Buy}, X = \text{high})$
$= P(\text{Buy} \mid X = \text{high}) P(X = \text{high})$
$= 0.1(0.6) = \underline{0.06}$

$\alpha_2(1) = P(\text{Buy} \mid X = \text{low}) P(X = \text{low})$
$= 0.5(0.6) = \underline{0.2}$

$\alpha_2(1) = 0.06(0.3)(0.3) + 0.2(0.6)(0.3)$
$= 0.0414$

$\alpha_2(2) = 0.06(0.7)(0.3) + 0.2(0.4)(0.3)$
$= 0.0366$

$\alpha_3(1) = 0.0414(0.3)(0.6) + 0.0366(0.6)(0.6) = \underline{0.020628}$

$\alpha_3(2) = 0.0414(0.7)(0.2) + 0.0366(0.4)(0.2) = \underline{0.008724}$

$\boxed{\text{Answer: } 0.020628 + 0.008724 = 0.029352}$

c) Filtering

$$P(X_3 = High \mid E_{1:3}) = \frac{P(X_3 = high, E_1 = buy, E_2 = keep, E_3 = sell)}{P(E_1 = buy, E_2 = keep, E_3 = sell)} = \frac{0.020628}{0.029352}$$

$$\boxed{= 0.703}$$

Prediction

$P(X_4 = high \mid E_{1:3})$

$$= P(X_4 = high \mid X_3 = high)\, P(X_3 = high \mid E_{1:3})$$

$$+$$

$$P(X_4 = high \mid X_3 = low)\, P(X_3 = low \mid E_{1:3})$$

$$= 0.3(0.703) + 0.6\left(\frac{0.008724}{0.029352}\right)$$

$$\boxed{= 0.389}$$

d) Low, high, high

Q2

Utilities

| | Accident | No. Acc |
|---|---|---|
| insured | -200 | -10 |
| not insured | -400 | 0 |

$P(S=\text{true}, D=\text{true}) = 0.02$
$P(S=\text{true}, D=\text{false}) = 0.18$
$P(S=\text{false}, D=\text{true}) = 0.08$
$P(S=\text{false}, D=\text{false}) = 0.72$

a) ⇒ Expected utility given no info.

$$P(A) = \sum_{sd} P(A=\text{true} \mid S=s, D=d)\, P(S=s, D=d)$$

$$= 0.5(0.02) + 0.3(0.18) + 0.2(0.08) + 0.1(0.72)$$

$P(A) = 0.152$
$P(\neg A) = 0.848$

if not insured ⇒ Utility $= 0.152 * 400$
$= \$60.8$

if insured ⇒ $0.152(200) + 0.848(10)$
$= \$38.88$

pay difference $\$60.8 - \$38.88$
$= \$21.92$

b)

insured

$0.3(200) + 0.7(10) = 67$

not insured

$0.3(400) = 120$

$120 - 67 = \$53$

c) not insured

$P(A|D) = P(A|D=\text{true}, S=\text{true})P(S=\text{true}) + P(A|D=\text{true}, S=\text{false})P(S=\text{false})$

$= 0.5(0.2) + 0.2(0.8) = 0.26$

$0.26(400) = 104$

insured

$0.26(200) + 0.74(10) = 59.4$

$\boxed{= \$44.6}$

d) From part a) cost of insurance is 38.88

$0.848(x) + 200(0.8)(0.152) + 400(0.2)(0.152) = 38.88$

$x = \$2.83$

Q3

a) $4^6 = 4096$ (4 actions, 6 states)

b) $V^0(s1) = 0 + 0.9(0.8(0) + 0.2(0)) = \underline{0}$

$V^0(s2) = 0 + 0.9(0.8(+20) + 0.2(0)) = \underline{14.4}$

$V^0(s3) = 20 + 0.9(1(20)) = \underline{38}$

$V^0(s4) = 10 + 0.9(0.8(0) + 0.2(+10)) = \underline{11.8}$

$V^0(s5) = 0 + 0.9(0.8(-10) + 0.2(0)) = \underline{-7.2}$

$V^0(s6) = -10 + 0.9(1(-10)) = \underline{-19}$

c) $\pi^1(s1) =$ Right

$\left.\begin{array}{l} 0 + 0.9(0.8(14.4)) \\ 0 + 0.9(0.8(11.8)) \end{array}\right\}$ S1 Right

$\pi^1(s2) =$ Right

$\left.\begin{array}{l} 14.4 + 0.9(0.8(38)) \\ 14.4 + 0.9(0.8(-7.2)) \end{array}\right\}$ S2 right

$\pi^1(s3) =$ Down

$\pi^1(s4) =$ Down

$\pi^1(s5) =$ UP

$\pi^1(s6) =$ UP

d) using policy for <u>C</u>

$v^1(s1) = 0 + 0.9(0.8(14.4) + 0.2(0)) = 10.368$

$v^1(s2) = 10 + 0.9(0.8(38) + 0.2(14.4)) = 29.952$

$v^1(s3) = 20 + 0.9(1(38)) = 54.2$

$v^1(s4) = 10 + 0.9(1(11.8)) = 20.62$

$v^1(s5) = 0 + 0.9(0.8(14.4) + 0.2(-7.2)) = 9.072$

$v^1(s6) = -10 + 0.9(0.8(38) + 0.2(-19)) = 13.94$

<u>One more iteration</u>

$\pi^2(s1) = \text{Right}$

$\pi^2(s2) = \text{Right}$

$\pi^2(s3) = \text{Down}$

$\pi^2(s4) = \text{Down}$

$\pi^2(s5) = \text{UP}$

$\pi^2(s6) = \text{UP}$

$v^2(s1) = 0 + 0.9(0.8(29.952) + 0.2(10.368)) = 23.431$

$v^2(s2) = 0 + 0.9(0.8(54.2) + 0.2(29.952)) = 44.14$

$v^2(s3) = 20 + 0.9(1(54.2)) = 68.78$

$v^2(s4) = 10 + 0.9(1(20.62)) = 28.558$

$v^2(s5) = 0 + 0.9(0.8(29.952) + 0.2(9.072))$
$= 23.1984$

$v^2(s6) = -10 + 0.9(0.8(54.2) + 0.2(13.94))$

$= 31.5332$

↑ stop when policy does not change.

These are the optimal values $V^*$

e) No. We did two iterations where the policy did not change, and both value functions were optimal. These functions were different and hence not unige.

f) As shown in part d) The optimal policy is (breaking ties alphabetically):

$\pi^*(s1) =$ Right

$\pi^*(s2) =$ Right

$\pi^*(s3) =$ down

$\pi^*(s4) =$ down

$\pi^*(s5) =$ up

$\pi^*(s6) =$ up

g) Yes, or else the policy would have changed in the two iterations shown in part d).

h) Changing S6 to −100 would not change the policy as the optimal policy moves up from S6.

$Q = \{1, 1, 3, 1, 1\}$

$A = \{1, 2, 3, 4, 5\}$

Q4) a)

$\boxed{Q_0 = \{1, 1, 3, 1, 1\}}$

$\underline{t=1}$

$$Q_1(A=3) = Q_0(A=3) + (r_{n+1} - Q_0(A=3))/1$$

$$= 3 + \frac{(2-3)}{1}$$

$$= 2$$

$\boxed{Q_1 = \{1, 1, 2, 1, 1\}}$

$\underline{t=2}$

$$Q_2(A=5) = 1 + \frac{0-1}{1}$$

$$= 0$$

$\boxed{Q_2 = \{1, 1, 2, 1, 0\}}$

$\underline{t=3}$

$$Q_3(A=3) = 2 + \frac{1-2}{2}$$

$$= 1.5$$

$\boxed{Q_3 = \{1, 1, 1.5, 1, 0\}}$

$\underline{t=4}$

$$Q_4(A=1) = 1 + \frac{0-1}{1}$$

$$= 0$$

$\boxed{Q_4 = \{0, 1, 1.5, 1, 0\}}$

$\underline{t=5}$

$$Q_5(A=3) = 1.5 + \frac{0-1.5}{3}$$

$$= 1$$

$\boxed{Q_5 = \{0, 1, 1, 1, 0\}}$

Scanned with CamScanner

b) $t=1$: cannot be concluded

$t=2$: can be concluded as best move is $A=3$, and we chose $A=5$

$t=3$: cannot be concluded since $A=3$ is best and we chose it

$t=4$: can be concluded as $A=1$ is not the best move.

$t=5$: cannot be concluded since we chose best move.