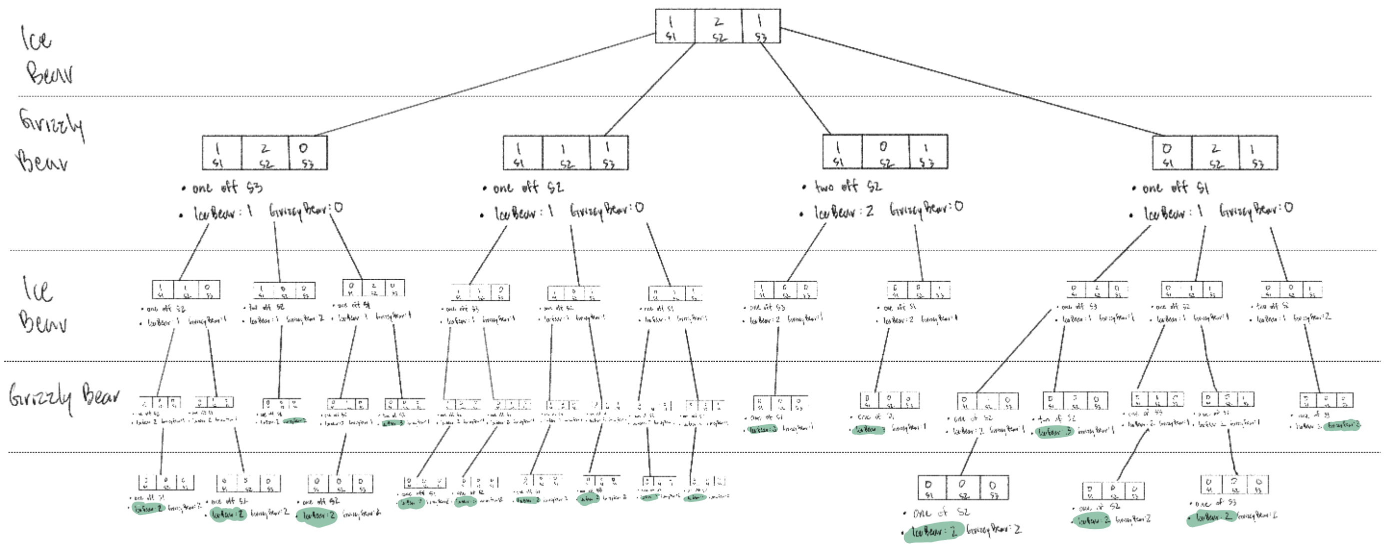
(1)

Turn

Possible States



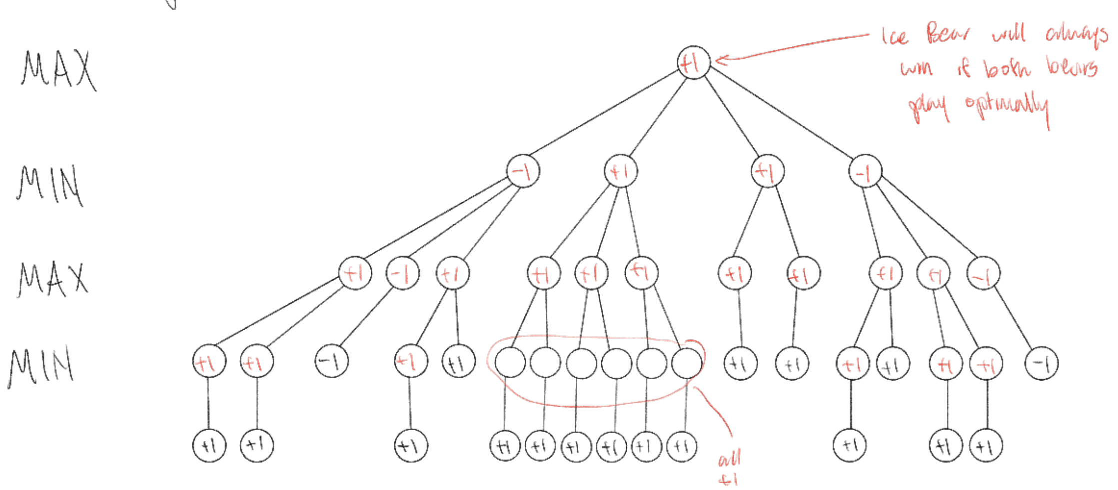
Ice Bear: 16

Grizzly Bear: 2

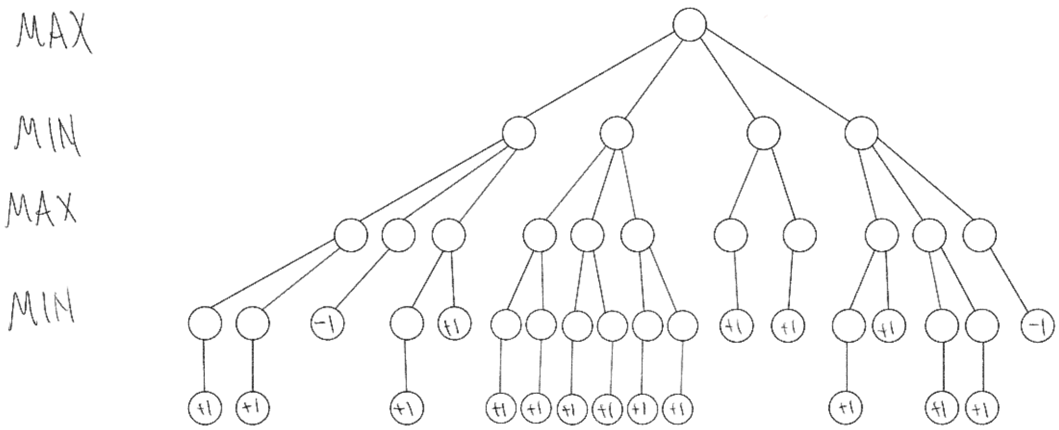
Total: 18 Leaf Nodes

MiniMax Algorithm

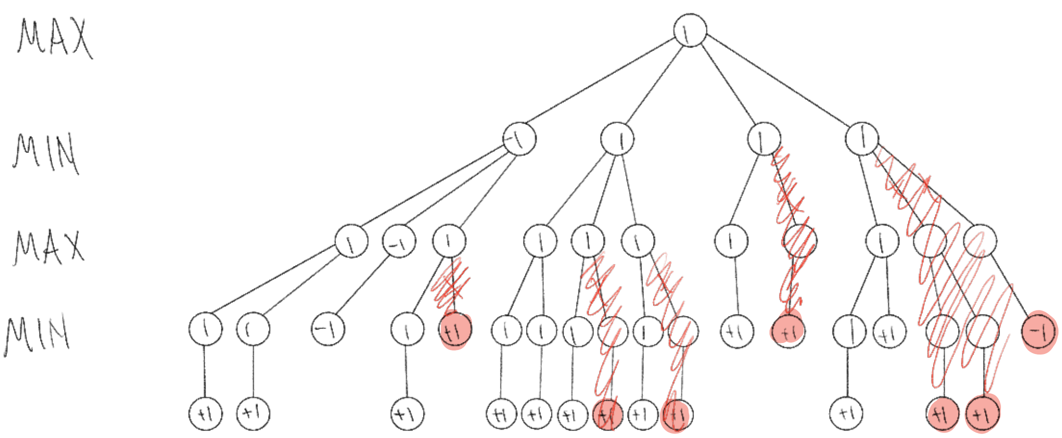
(3)



4

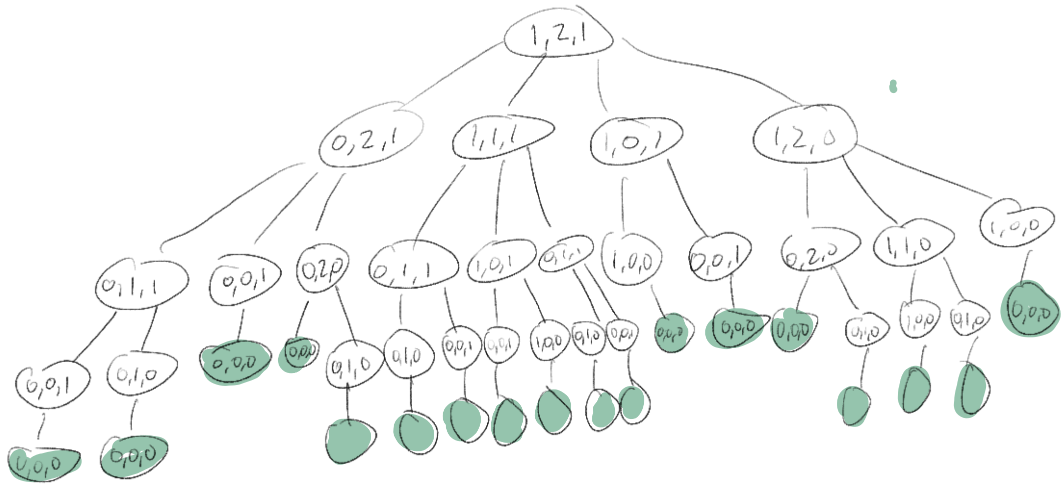


4



7 pruned
1 Grizzly bear
6 Ice bear

1) Checking 1



1,2,1
0,2,1
1,1,1
1,0,1
1,2,0
0,1,1
0,0,1
0,2,0
0,1,1
1,0,1
1,0,0
0,0,1
0,2,0
1,1,0
1,0,0
0,0,1
0,1,0
0,0,0
0,0,0
0,1,0
0,1,0
1,0,0
0,1,0
0,0,1
0,0,0
0,0,0
0,1,0
1,0,0
0,1,0
0,0,0
0,0,0
0,0,0

b) RL

0	V_1	V_2	-22
V_1	-18	-20	V_2
V_2	-20 S1	-18	V_1 S2
-22	V_2	V_1	0

North South East West

$$V_1 = \frac{1}{4}(-1+V_1) + \frac{1}{4}(-1-18) + \frac{1}{4}(-1+V_2) + \frac{1}{4}(-1)$$

$$V_2 = \frac{1}{4}(-1+V_2) + \frac{1}{4}(-1-20) + \frac{1}{4}(-1-22) + \frac{1}{4}(-1+V_1)$$

All the same for other states

$$V_1 = -\frac{1}{4} + \frac{1}{4}V_1 - \frac{1}{4} - \frac{18}{4} - \frac{1}{4} + \frac{1}{4}V_2 - \frac{1}{4}$$

$$V_1 = -\frac{4}{4} - \frac{18}{4} + \frac{1}{4}V_1 + \frac{1}{4}V_2$$

$$V_1 - \frac{1}{4}V_1 = -\frac{22}{4} + \frac{1}{4}V_2$$

$$\frac{4}{3}\cdot\frac{3}{4}V_1 = \left(-\frac{22}{\cancel{4}} + \frac{1}{\cancel{4}}V_2\right)\cdot\frac{\cancel{4}}{3}$$

$$V_1 = -\frac{22}{3} + \frac{1}{3}V_2$$

$$V_2 = \frac{1}{4}(-1+V_2) + \frac{1}{4}(-1-20) + \frac{1}{4}(-1-22) + \frac{1}{4}(-1+V_1)$$

$$V_2 = -\frac{1}{4} + \frac{1}{4}V_2 - \frac{21}{4} - \frac{23}{4} - \frac{1}{4} + \frac{1}{4}V_1$$

$$V_2 = -\frac{46}{4} + \frac{1}{4}V_2 + \frac{1}{4}V_1$$

$$V_2 - \frac{1}{4}V_2 = -\frac{46}{4} + \frac{1}{4}V_1$$

$$\frac{\cancel{4}}{\cancel{3}}\left(\frac{\cancel{3}}{\cancel{4}}V_2\right) = \left(-\frac{46}{\cancel{4}} + \frac{1}{\cancel{4}}V_1\right)\frac{4}{3}$$

$$V_2 = -\frac{46}{3} + \frac{1}{3}V_1$$

Solving For V_1 & V_2

$$V_1 = -\frac{22}{3} + \frac{1}{3}V_2 \quad \Big| \quad V_2 = -\frac{46}{3} + \frac{1}{3}V_1$$

$$V_1 = -\frac{22}{3} + \frac{1}{3}\left(-\frac{46}{3} + \frac{1}{3}V_1\right)$$

$$V_1 = -\frac{22}{3} - \frac{46}{9} + \frac{1}{9}V_1$$

$$\frac{9}{9}V_1 - \frac{1}{9}V_1 = \frac{-66-46}{9}$$

$$\frac{8}{\cancel{9}}V_1 = -\frac{112}{\cancel{9}}$$

$$V_1 = \frac{-112}{8} = -14$$ 6

$V_1 = -14$

$$V_2 = -\frac{46}{3} + \frac{1}{3}(-14)$$

$$V_2 = \frac{-46-14}{3} = -\frac{60}{3} = -20$$

$$V_2 = -20$$ 7

8, 9

0	-14	-20	-22
-14	-18	-20	-20
-20	-20 S1	-18	-14 S2
-22	-20	-14	0

8 $q_\pi(S1, SOUTH)$

9 $q_\pi(S2, SOUTH)$

$Q(s,a) = \sum_{s'} T(s,a,s')[R(s,a,s') + \cancel{\gamma} V(s')]$

$Q(s,a) = \sum_{s'} T(s,a,s')[R(s,a,s') + V(s')]$

$q_\pi(S2, South) = (-1 + 0) = -1$

$q_\pi(S1, South) = 1(-1-20) = -21$

(10)

Givens:

- $P(D_n | S_m)$
- $S_{rabbit} = 4 \cdot S_{cat}$
- $S_{dog} = 2 \cdot S_{cat}$

Find:

$$P(S_{rabbit} | D_{rabbit}) = \frac{P(D_{rabbit} | S_{rabbit})}{P(D_{rabbit})} P(S_{rabbit})$$

$$= \frac{(4/10)}{(3/10)} (4/7) = \frac{16}{21} = 0.\overline{761904}$$

Sample: c, r, r, r, r, d, d

$P(S_{cat}) = \frac{1}{7}$

$P(S_{rabbit}) = \frac{4}{7}$

$P(S_{dog}) = \frac{2}{7}$

$$P(D_{rabbit}) = P(D_{rabbit} | S_{dog}) \cdot P(S_{dog}) + P(D_{rabbit} | S_{cat}) \cdot P(S_{cat}) + P(D_{rabbit} | S_{rabbit}) \cdot P(S_{rabbit})$$

$$\left(\frac{2}{10}\right)\left(\frac{2}{7}\right) + \left(\frac{1}{10}\right)\left(\frac{1}{7}\right) + \left(\frac{4}{10}\right)\left(\frac{4}{7}\right) = \left(\frac{21}{70}\right) = \frac{3}{10} = 0.3 = P(D_{rabbit})$$

(11) Probability of success

- means what's the probability of showing a different animal

$P(S)$ = probability of success

S_{dog}

$P(S | S_{dog})$ = probability of success given a transformation of dog

$P(S | S_{dog}) = P(D_{cat} | S_{dog}) + P(D_{rabbit} | S_{dog})$

$P(S | S_{dog}) = 0.2 + 0.2 = 0.4$

S_{rabbit}

$P(S | S_{rabbit}) = P(D_{dog} | S_{rabbit}) + P(D_{cat} | S_{rabbit})$

$P(S | S_{rabbit}) = 0.3 + 0.3 = 0.6$

S_{cat}

$P(S | S_{cat}) = P(D_{dog} | S_{cat}) + P(D_{rabbit} | S_{cat})$

$P(S | S_{cat}) = 0.1 + 0.1 = 0.2$

Total Probability of $P(S)$

$$P(S) = P(S | S_{dog}) \cdot P(S_{dog}) + P(S | S_{cat}) \cdot P(S_{cat}) + P(S | S_{rabbit}) \cdot P(S_{rabbit})$$

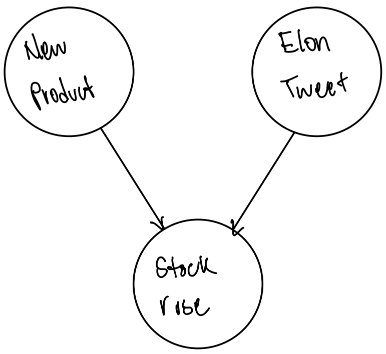
$$P(S) = \left(\frac{4}{10}\right)\left(\frac{2}{7}\right) + \left(\frac{2}{10}\right)\left(\frac{1}{7}\right) + \left(\frac{6}{10}\right)\left(\frac{4}{7}\right) = \frac{34}{70} = 0.485714$$

Bayes Nets

Alex, a student obsessed with tech stocks and memes, creates a Bayesian Network to predict whether NVIDIA's stock will rise tomorrow. The following is the Bayesian network that models the possibility of a stock price increase:

- NEW_PRODUCT: Will NVIDIA announce a new product or technology?
- ELON_TWEET: Will Elon Musk tweet about AI or graphics cards?
- STOCK_RISE: Will NVIDIA's stock price rise?

NVIDIA's stock won't rise if there's no new product announcement and Elon doesn't tweet about AI or graphics cards. If both a new product is announced and Elon tweets about AI or graphics cards, there's a 60% chance the stock will rise. Based on market analysis and Elon's Twitter habits, there's a 50% chance NVIDIA will announce a new product and an 70% chance Elon will tweet something relevant.



Posterior Probability

12 P(New Product)

+p	0.5
−p	0.5

P(Elon Tweet)

+e	0.7
−e	0.3

P(Stock rise)

+s	0.21
−s	0.79

$P(+s) = P(+s|+p,+e)\cdot P(+p,+e) + P(+s|+p,-e)\cdot P(+p,-e) + P(+s|-p,+e)\cdot P(-p,+e) + P(+s|-p,-e)\cdot P(-p,-e)$

$(0.6)\cdot(0.5\cdot 0.7) + (0)\cdot(0.5\cdot 0.3) + (0)\cdot(0.5\cdot 0.7) + (0)\cdot(0.5\cdot 0.3) = 0.21$

15

Conditional Probability

13 P(Stock Rise | New Product, Elon Tweet)

P: Product	E: Elon Tweet	S: Stock	P(S | P, E)
+p	+e	+s	0.6
+p	+e	−s	0.4
+p	−e	+s	0
+p	−e	−s	1.0
−p	+e	+s	0
−p	+e	−s	1.0
−p	−e	+s	0
−p	−e	−s	1.0

Given

$P(+s \mid +e, +a) = 0.6$

$P(-s \mid -e, -a) = 1.0$

16 Find $P(+p)$

$P(+s) = 0.28$

from previous calculations

$P(+s) = P(+s|+p,+e)\cdot P(+p,+e)$

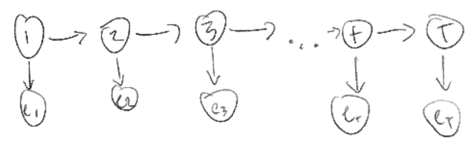
$P(+s) = P(+s|+p,+e)\cdot P(+p)\cdot P(+e)$

$0.28 = 0.6\cdot P(+p)\cdot 0.7$

$P(+p) = 0.66\bar{6}$

(17) Markov Assumption

- $P(X_t \mid X_{0:t-1}) = P(X_t \mid X_{t-1})$
- $P(E_t \mid X_{0:t}, E_{1:t-1}) = P(E_t \mid X_t)$
- $P(X_{0:t}, E_{1:t}) = P(X_0) \prod_{i=1}^{t} P(X_i \mid X_{i-1}) P(E_i \mid X_i)$



2. $P(X_t \mid e_{1:t}) = P(X_t \mid e_{1:T})$

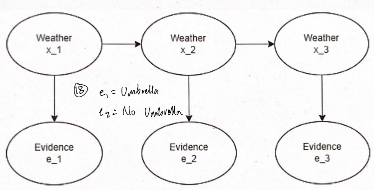
$\Updownarrow$

$P(X_t$

Hidden Markov Models

You want to predict the weather based on whether your friend Emily uses an umbrella. The hidden states represent the weather conditions: Sunny, Cloudy, and Rainy. The evidence (observation) is whether Emily uses an umbrella or not.

e_1 = umbrella e_2 = no umbrella



(18) e_1 = Umbrella
e_2 = No Umbrella

Initial Probabilities:

	Sunny	Cloudy	Rainy
$P(x_1)$	0.5	0.3	0.2

Transition Probabilities:

X_t	$P(X_{t+1} \mid X_t)$		
	Sunny	Cloudy	Rainy
Sunny	0.7	0.2	0.1
Cloudy	0.3	0.4	0.3
Rainy	0.2	0.3	0.5

Emission Probabilities:

X_t	$P(E_t \mid X_t)$	
	No Umbrella	Umbrella
Sunny	0.4	0.6
Cloudy	0.8	0.2
Rainy	0.1	0.9

a) $P(X_1 = \text{Sunny} \mid e_1) = 0.3$

$= P(X_1) \cdot P(e_t \mid X_t)$

$= P(X_1 = \text{Sunny}) \cdot P(E_t = \text{Umbrella} \mid X_t = \text{Sunny})$

$= 0.5 \cdot 0.6 = 0.3$

$= 0.3 / 0.54 = 0.555$

b) $P(X_1 = \text{Cloudy} \mid e_1) =$

$= P(X_1 = \text{Cloudy}) \cdot P(E_t = \text{umbrella} \mid X_t = \text{Cloudy})$

$= 0.3 \cdot 0.2 = 0.06$

$= 0.06 / 0.54 = 0.111$

c) $P(X_1 = \text{Rainy} \mid e_1)$:

$= P(X_1 = \text{Rainy}) \cdot P(E_t = \text{umbrella} \mid X_t = \text{Rainy})$

$= 0.2 \cdot 0.9 = 0.18$

$= 0.18 / 0.54 = 0.333$

Normalize:

$0.3 + 0.06 + 0.18 = 0.54$

d) $P(X_2 = \text{Sunny} \mid e_1, e_2)$

$= P(E_t = \text{No Umbrella} \mid X_t = \text{Sunny}) \left[\sum_{X_1} P(X_2 = \text{Sunny} \mid X_1) P(X_1 \mid e_1) \right]$

$= P(E_t = \text{No Umbrella} \mid X_t = \text{Sunny}) \left[(P(X_2 = \text{Sunny} \mid X_1 = \text{Sunny}) \cdot P(X_1 = \text{Sunny} \mid e_1)) + (P(X_2 = \text{Sunny} \mid X_1 = \text{Cloudy}) \cdot P(X_1 = \text{Cloudy} \mid e_1)) + (P(X_2 = \text{Sunny} \mid X_1 = \text{Rainy}) \cdot P(X_1 = \text{Rainy} \mid e_1)) \right]$

$= 0.4 [(0.7 \cdot 0.555) + (0.3 \cdot 0.111) + (0.2 \cdot 0.333)]$

$= 0.19536$ $= 0.19536 / 0.42513 = 0.45953$

e) $P(X_2 = \text{Cloudy} \mid e_1, e_2)$

$= P(E_t = \text{No Umbrella} \mid X_t = \text{Cloudy}) \left[(P(X_2 = \text{Cloudy} \mid X_1 = \text{Cloudy}) \cdot P(X_1 = \text{Cloudy} \mid e_1)) + (P(X_2 = \text{Cloudy} \mid X_1 = \text{Sunny}) \cdot P(X_1 = \text{Sunny} \mid e_1)) + (P(X_2 = \text{Cloudy} \mid X_1 = \text{Rainy}) \cdot P(X_1 = \text{Rainy} \mid e_1)) \right]$

$= 0.8 [(0.4 \times 0.111) + (0.2 \times 0.555) + (0.3 \times 0.333)]$

$= 0.20424$

$= 0.20424 / 0.42513 = 0.480417$

f) $P(X_2 = \text{Rainy} \mid e_1, e_2)$

$= P(E_t = \text{No Umbrella} \mid X_t = \text{Rainy}) \left[\sum_{X_1} P(X_2 = \text{Rainy} \mid X_1) \cdot P(X_1 \mid e_1) \right]$

$= 0.1 [(0.3 \times 0.111) + (0.1 \times 0.555) + (0.5 \times 0.333)]$

$= 0.02553$

$= 0.02553 / 0.42513 = 0.06005$

$\alpha = 0.02553 + 0.20424 + 0.19536 = 0.42513$

Book Formula:

$$P(X_{t+1} \mid e_{1:t+1}) = \underset{\text{normalization constant}}{\alpha} P(e_{t+1} \mid X_{t+1}) \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t \mid e_{1:t})$$

(19) Viterbi Algo from book

$m_{1:t+1} = P(e_{t+1} \mid X_{t+1}) \max_{x_t} P(X_{t+1} \mid x_t) \max_{x_{1:t-1}} P(x_{1:t-1}, x_t, e_{1:t})$

$[e_1 = U, e_2 = \text{No } U, e_3 = U]$ find $[X_1, X_2, X_3]$

X_1	$P(X_1, E_t = \text{Umbrella})$
Sunny	$P(X_1) \cdot P(e_t \mid X_t) = 0.5 \cdot 0.6 = 0.3$
Cloudy	$0.3 \cdot 0.2 = 0.06$
Rainy	$0.2 \cdot 0.9 = 0.18$

X_2	
Sunny	$P(E_t = \text{No U}, X_t = \text{Sunny}) \cdot \max_{1:t} P(X_{t+1}^{\text{Sunny}} \mid X_t) P(X_{1:t}, e_{1:t})$ $0.4 \max(\overset{0.21}{0.7 \cdot 0.3}, \overset{0.018}{0.3 \cdot 0.06}, \overset{0.036}{0.2 \cdot 0.18}) = 0.4 \cdot 0.21 = 0.084$
Cloudy	$0.8 \max(\overset{0.06}{0.2 \cdot 0.3}, \overset{0.024}{0.4 \cdot 0.06}, \overset{0.054}{0.18 \cdot 0.3}) = 0.8 \cdot 0.06 = 0.048$
Rainy	$0.1 \max(\overset{0.03}{0.1 \cdot 0.3}, \overset{0.018}{0.3 \cdot 0.06}, \overset{0.09}{0.5 \cdot 0.18}) = 0.1 \cdot 0.09 = 0.009$

X_3	$m_{1:t+1} = P(e_{t+1} \mid X_{t+1}) \max_{x_{1:t}} P(X_{t+1} \mid X_t) P(x_{1:t}, e_{1:t})$
Sunny	$P(e_1 = U \mid X_{t+1} = \text{Sunny})$ $0.6 \max(\underset{\text{Sunny}}{\overset{0.0588}{0.7 \cdot 0.084}}, \underset{\text{Cloudy}}{\overset{0.0144}{0.3 \cdot 0.048}}, \underset{\text{rainy}}{\overset{0.0018}{0.2 \cdot 0.009}}) = 0.6 \cdot 0.0588 = 0.03528$
Cloudy	$0.2 \max(\underset{\text{Sunny}}{\overset{0.0168}{0.2 \cdot 0.084}}, \underset{\text{Cloudy}}{\overset{0.0192}{0.4 \cdot 0.048}}, \underset{\text{rainy}}{\overset{0.0027}{0.3 \cdot 0.009}}) = 0.2 \cdot 0.0192 = 0.00384$
Rainy	$0.9 \max(\overset{0.0084}{0.1 \cdot 0.084}, \overset{0.0144}{0.3 \cdot 0.048}, \overset{0.0045}{0.5 \cdot 0.009})$ $= 0.9 \cdot 0.0144 = 0.01296$

(20) Joint Probability $P(x_1, x_2, x_3, e_1, e_2, e_3)$

$P(x_1 = S, e_1 = U, x_2 = S, e_2 = NO\ U, x_3 = S, e_3 = U)$

$= P(X_1 = Sunny) \cdot P(E_1 = U \mid X_1 = Sunny) \cdot P(X_2 = S, X_1 = S) \cdot P(E_2 = NO\ U \mid X_2 = S) \cdot P(X_3 = S \mid X_2 = S) \cdot P(E_3 = U \mid X_3 = S)$

$= 0.5 \cdot 0.6 \cdot 0.7 \cdot 0.4 \cdot 0.7 \cdot 0.6 = 0.03528$

(21)

$t = 2$

$x_{2,1} = Sunny$ $x_{2,2} = Cloudy$ $x_{2,3} = Cloudy$ $x_{2,4} = Rainy$

$e_2 = NO\ Umbrella$

Update : new information $e_t = e_2 =$ no umbrella

weight(particle) = $P(e_t = \text{no umbrella} \mid X_t = \text{particle})$

weight $(x_{2,1}) = P(e_t = \text{no umbrella} \mid X_t = sunny) = 0.4$

weight $(x_{2,2}) = P(\text{no } u \mid Cloudy) = 0.8$

weight $(x_{2,3}) = 0.8$

weight $(x_{2,4}) = 0.1$

Normalize

$0.4 + 0.8(2) + 0.1 = 2.1$

$w(x_{2,1}) = \frac{0.4}{2.1} = 0.190476$ Sunny

$w(x_{2,2}) = \frac{0.8}{2.1} = 0.380952$ Cloudy

$w(x_{2,3}) = \frac{0.8}{2.1} = 0.380952$

$w(x_{2,4}) = \frac{0.1}{2.1} = 0.047619$ rainy

(22)

Resample

Transition Probabilities:

X_t	$P(X_{t+1} \mid X_t)$		
	Sunny	**Cloudy**	**Rainy**
Sunny	0.7	0.2	0.1
Cloudy	0.3	0.4	0.3
Rainy	0.2	0.3	0.5

particles

$x_{2,1} = Sunny$

$P(X_{3,i} = Sunny \mid X_{2,1} = Sunny) = 0.7$

$P(X_3 = Cloudy \mid X_{2,1} = Sunny) = 0.2$

$P(X_3 = Rainy \mid X_{2,1} = Sunny) = 0.1$

$x_{2,2} = Cloudy$ $x_{2,3} = Cloudy$

$P(X_{3,i} = Sunny \mid X_{2,2} = Cloudy) = 0.3$

$P(X_{3,i} = Cloudy \mid X_{2,2} = Cloudy) = 0.4$

$P(X_{3,i} = Rainy \mid X_{2,2} = Cloudy) = 0.3$

$x_{2,4} = Rainy$

$P(X_{3,i} = Sunny \mid X_{2,4} = rainy) = 0.2$

$P(X_{3,i} = Cloudy \mid X_{2,4} = rainy) = 0.3$

$P(X_{3,i} = rainy \mid X_{2,4} = rainy) = 0.5$

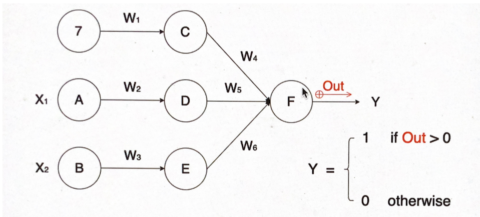
Weighted

$P(X_{3,i} = Sunny) = (0.190476 \cdot 0.7) + 2(0.380952 \cdot 0.3) + (0.047619 \cdot 0.2) = 0.371428$

$P(X_{3,i} = Cloudy) = (0.190476 \cdot 0.2) + 2(0.380952 \cdot 0.4) + (0.047619 \cdot 0.3) = 0.3571425$

$P(X_{3,i} = Rainy) = (0.190476 \cdot 0.1) + 2(0.380952 \cdot 0.3) + (0.047619 \cdot 0.5) = 0.2714283$

24

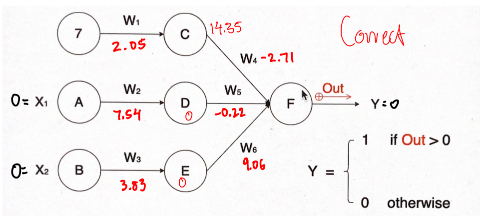


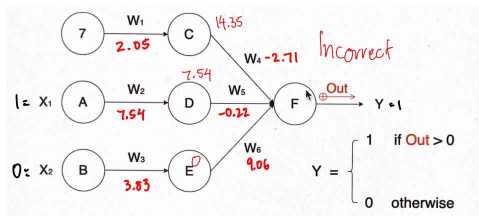
XOR Function

X_1	X_2	OUT
0	0	0
0	1	1
1	0	1
1	1	0

$W_1, W_2, W_3, W_4, W_5, W_6$

① [2.05, 7.54, 3.83, -2.71, -0.22, 9.06]





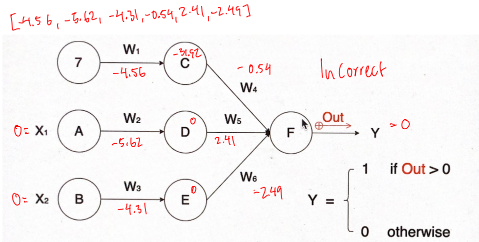
F = -2.71(14.35) + 0 + 0 = -38.8885

Y = 0

F = -2.71(14.35) + -0.22(7.54) = -40.5473

Y = 0

② [-4.56, -5.62, -4.31, -0.54, 2.41, -2.49]

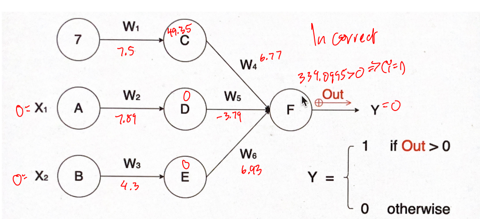


C = 7·(-4.56) = -31.92

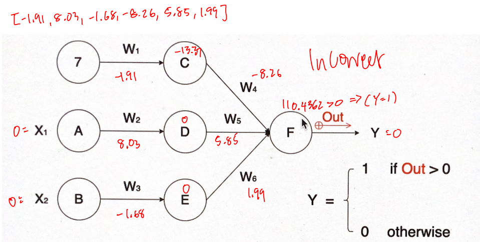
F = (-31.92)(-0.54) + 0 + 0 = 17.2368

17.2368 > 0 , Y = 1

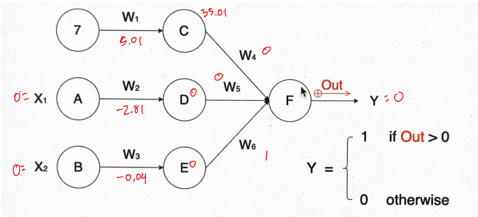
③ [7.5, 7.89, 4.3, 6.77, -3.79, 6.93]

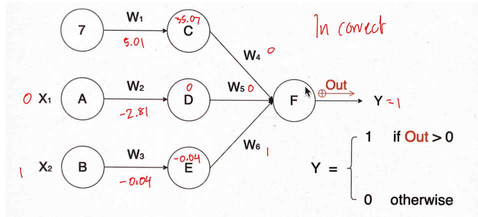


④ [-1.91, 8.03, -1.68, -8.26, 5.85, 1.99]



5 [5.01, -2.81, -0.04, 0, 0, 1]





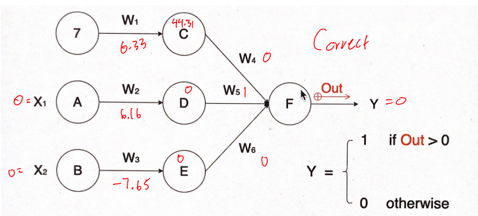
$F = (35.07)(0) + (0)(-2.81) + (0)(-0.04) = 0$

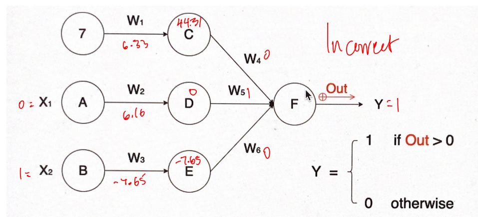
$0 > 0 \Rightarrow (Y = 0)$

$F = \cancel{(35.07)(0)} + \cancel{(0)(-2.81)} + (-0.04)(1) = -0.04$

$-0.04 > 0 \Rightarrow (Y = 0)$

6 [6.33, 6.16, -7.65, 0, 1, 0]



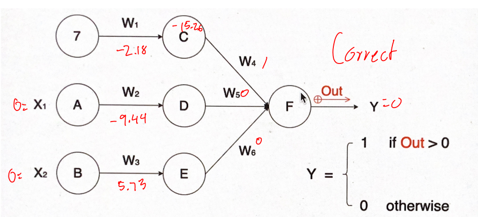


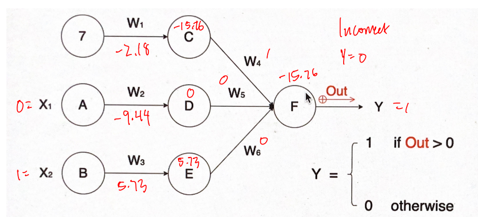
$F = (0)(44.31) + (0)(1) + (0)(-7.65) = 0$

$F = (44.31)(0) + (0)(1) + (0)(-7.65) = 0$

$0 > 0 \Rightarrow (Y = 0)$

7 [-2.18, -9.44, 5.73, 1, 0, 0]

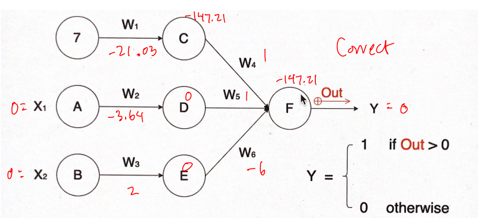


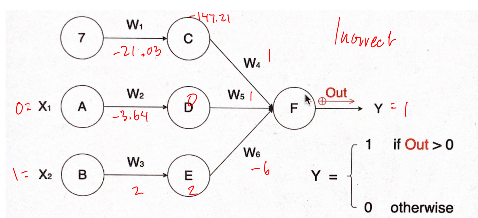


$F = -15.26$

$Y = 0$

8

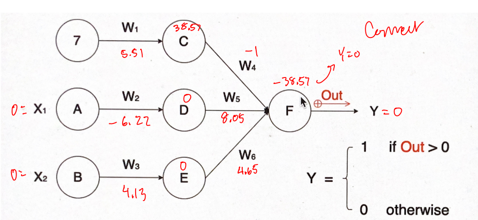


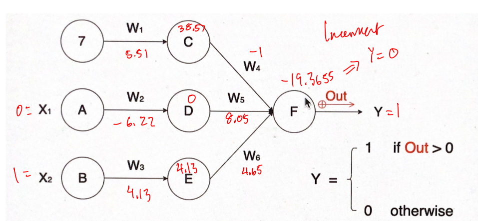


$F = -147.21 + (2)(-6) = -159.21$

$Y = 0$

9





$F = -1(38.57) + (4.13)(4.65) = -19.3655$

25

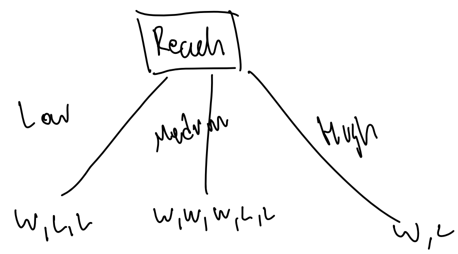
Reach	Skills	Experience	Outcome
High	Low	Low	Loss
Low	Medium	High	Loss
Medium	Low	High	Loss
Medium	Medium	High	Win
High	High	Medium	Win
Low	Medium	Medium	Win
Medium	High	Low	Loss
Low	Medium	Low	Loss
Medium	Low	High	Win
Medium	Medium	Low	Win

entropy $= -\sum_i (p_i)\log_2(p_i)$

5 wins, 5 losses, 10 total

$$-\left(\frac{5}{10}\right)\log_2\left(\frac{5}{10}\right) - \left(\frac{5}{10}\right)\log_2\left(\frac{5}{10}\right) = 1 \text{ bit}$$

26



Reach	Skills	Experience	Outcome
High	Low	Low	Loss
Low	Medium	High	Loss
Medium	Low	High	Loss
Medium	Medium	High	Win
High	High	Medium	Win
Low	Medium	Medium	Win
Medium	High	Low	Loss
Low	Medium	Low	Loss
Medium	Low	High	Win
Medium	Medium	Low	Win

Calculation

Low entropy $= -\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right) = 0.918296$ bits

Medium entropy: $-\frac{3}{5}\log_2\left(\frac{3}{5}\right) - \frac{2}{5}\log_2\left(\frac{2}{5}\right) = 0.970951$ bits

High entropy $= -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$ bits

Weighting

$$\frac{3}{10}(0.918296) + \frac{5}{10}(0.970951) + \frac{2}{10}(1) = 0.960964$$

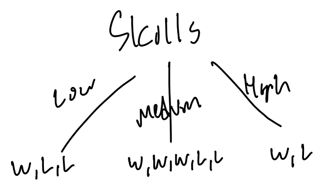
27 Information Gained

Original Entropy − weighted entropy after splitting at "Reach"

$1 - 0.960964 = 0.039036$ bits

(28) Skills

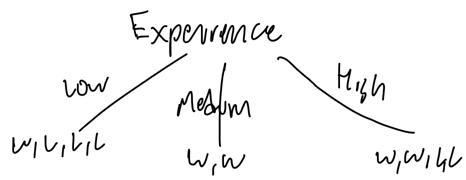
Reach	Skills	Experience	Outcome
High	Low	Low	Loss
Low	Medium	High	Loss
Medium	Low	High	Loss
Medium	Medium	High	Win
High	High	Medium	Win
Low	Medium	Medium	Win
Medium	High	Low	Loss
Low	Medium	Low	Loss
Medium	Low	High	Win
Medium	Medium	Low	Win



Same as Reach

(30) Experience

Reach	Skills	Experience	Outcome
High	Low	Low	Loss
Low	Medium	High	Loss
Medium	Low	High	Loss
Medium	Medium	High	Win
High	High	Medium	Win
Low	Medium	Medium	Win
Medium	High	Low	Loss
Low	Medium	Low	Loss
Medium	Low	High	Win
Medium	Medium	Low	Win



Entropy Calculation

Low: $-\frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{3}{4}\log_2\left(\frac{3}{4}\right) = 0.811278$

Medium: $-\frac{2}{2}\log_2\left(\frac{2}{2}\right) = 0$

High: $-\frac{2}{4}\log_2\left(\frac{2}{4}\right) - \frac{2}{4}\log_2\left(\frac{2}{4}\right) = 1$

Weighted

$$\frac{4}{10}(0.811278) + \cancel{\frac{2}{10}(0)} + \frac{4}{10}(1) = 0.724511$$

(31) Information Gained

$1 - 0.724511 = 0.275489$

(32) Experience has highest Gain at 0.275489

(33) Reach and Skills have equal number of gain at 0.039 bits thus are equally "bad"

(34) 28×28 input layer : 784

80 hidden nodes : 80

10 output layer : 10

$(784 \times 80) + (80 \times 10) = 63520$

Biases:

$80 + 10 = 90$

Total : $90 + 63520 = 63610$