Ethan Wong
5/26/22

CS 161 HW9

1. First we should find the attribute $\{A, B, C\}$ that minimizes the conditional entropy of D.

22

A) $ENT(A,d) = \sum_a Pr(a)\, ENT(D|A)$

- $ENT(D|a) = Pr(a) \cdot -[Pr(d|a) \cdot \log_2(Pr(d|a)) + Pr(\bar{d}|a) \cdot \log_2(Pr(\bar{d}|a))]$
$= \frac{11}{22} \cdot -[\frac{7}{11} \log_2(\frac{7}{11}) + \frac{4}{11} \log_2(\frac{4}{11})]$
$= 0.4728$

- $ENT(D|\bar{a}) = Pr(\bar{a}) \cdot -[Pr(d|\bar{a}) \cdot \log_2(Pr(d|\bar{a})) + Pr(\bar{d}|\bar{a}) \cdot \log_2(Pr(\bar{d}|\bar{a}))]$
$= \frac{11}{22} \cdot -[\frac{3}{11} \cdot \log_2(\frac{3}{11}) + \frac{8}{11} \log_2(\frac{8}{11})]$
$= 0.4226$

- $ENT(D|A) = 0.4728 + 0.4226 = 0.8954$

B) $ENT(D|B) = \sum_b Pr(b)\, ENT(D|b)$

- $ENT(D|b) = Pr(b) \cdot -[Pr(d|b) \cdot \log_2(Pr(d|b)) + Pr(\bar{d}|b) \cdot \log_2(Pr(\bar{d}|b))]$
$= \frac{14}{22} \cdot -[\frac{8}{14} \log_2(\frac{8}{14}) + \frac{6}{14} \log_2(\frac{6}{14})]$
$= 0.6270$

- $ENT(D|\bar{b}) = Pr(\bar{b}) \cdot -[Pr(d|\bar{b}) \cdot \log_2(Pr(d|\bar{b}) + Pr(\bar{d}|\bar{b}) \cdot \log_2(Pr(\bar{d}|\bar{b})]$
$= \frac{8}{22} \cdot -[\frac{2}{8} \log_2(\frac{2}{8}) + \frac{6}{8} \log_2(\frac{6}{8})]$
$= 0.2950$

- $ENT(D|B) = 0.6270 + 0.2950 = 0.922$

C) $ENT(D|C) = \sum_c Pr(c)\, ENT(D|c)$

- $ENT(D|c) = Pr(c) \cdot -[Pr(d|c) \cdot \log_2(Pr(d|c) + Pr(\bar{d}|c) \cdot \log_2(Pr(\bar{d}|c))]$
$= \frac{7}{22} \cdot -[\frac{4}{7} \log_2(\frac{4}{7}) + \frac{3}{7} \log(\frac{3}{7})]$
$= 0.3135$

- $ENT(D|\bar{c}) = Pr(\bar{c}) \cdot -[Pr(d|\bar{c}) \cdot \log_2(Pr(d|\bar{c})) + Pr(\bar{d}|\bar{c}) \cdot \log_2(Pr(\bar{d}|\bar{c}))]$
$= \frac{15}{22} \cdot -[\frac{6}{15} \log_2(\frac{6}{15}) + \frac{9}{15} \log_2(\frac{9}{15})]$
$= 0.6620$

- $ENT(D|C) = 0.3135 + 0.6620 = 0.9755$

Splitting on A yields the lowest conditional entropy and thus greatest information gain. The first split on the decision tree should be A.

$$\begin{cases} A=F: \{+x_5, -x_6, +x_7, -x_8\} \\ A=T: \{+x_1, +x_2, -x_3, -x_4\} \end{cases}$$

D) Branch $A = T$ Examples: $\{x_1, x_2, x_3, x_4\}$

① Split on B

$ENT(D|B) = \sum_b Pr(B)\, ENT(D|b)$

• $ENT(D|b) = Pr(b) \cdot -[Pr(d|b) \cdot \log_2(Pr(d|b)) + Pr(\bar{d}|b) \cdot \log_2(Pr(\bar{d}|b))]$

$= \frac{7}{11} \cdot -[\frac{7}{7} \log_2(\frac{7}{7}) + \frac{0}{7} \log_2(\frac{0}{7})]$

$= 0$

• $ENT(D|\bar{b}) = Pr(\bar{b}) \cdot -[Pr(d|\bar{b}) \cdot \log_2(Pr(d|\bar{b})) + Pr(\bar{d}|\bar{b}) \cdot \log_2(Pr(\bar{d}|\bar{b}))]$

$= \frac{4}{11} \cdot -[\frac{0}{4} \log_2(\frac{0}{4}) + \frac{4}{4} \log_2(\frac{4}{4})]$

$= 0$

• $ENT(D|B) = 0 + 0 = 0$

This is the smallest possible conditional entropy, so we split on B.

E) Branch $A = F$ Examples $\{x_5, x_6, x_7, x_8\}$

① Split on C

$ENT(D|C) = \sum_c Pr(C)\, ENT(D|c)$

• $ENT(D|c) = Pr(c) \cdot -[Pr(d|c) \cdot \log_2(Pr(d|c)) + Pr(\bar{d}|c) \cdot \log_2(Pr(\bar{d}|c))]$

$= \frac{3}{11} \cdot -[\frac{3}{3} \log_2(\frac{3}{3}) + \frac{0}{3} \log_2(\frac{0}{3})]$

$= 0$

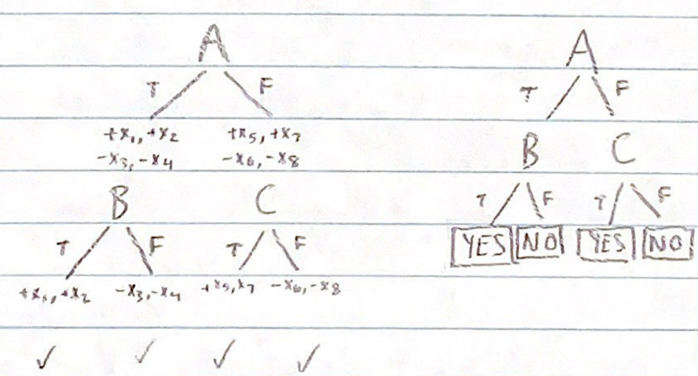
• $ENT(D|\bar{c}) = Pr(\bar{c}) \cdot -[Pr(d|\bar{c}) \cdot \log_2(Pr(d|\bar{c})) + Pr(\bar{d}|\bar{c}) \cdot \log_2(Pr(\bar{d}|\bar{c}))]$

$= \frac{8}{11} \cdot -[\frac{0}{8} \log_2(\frac{0}{8}) + \frac{8}{8} \log_2(\frac{8}{8})]$

$= 0$

• $ENT(D|C) = 0 + 0 = 0$

This is the smallest possible conditional entropy, so we split on C.



2. $(A \lor \neg B) \oplus (\neg C \lor D)$

Convert to CNF

$[(A \lor \neg B) \lor (\neg C \lor D)] \land [\neg(A \lor \neg B) \lor \neg(\neg C \lor D)]$

$[(A \lor \neg B) \lor (\neg C \lor D)] \land [(\neg A \land B) \lor (C \land \neg D)]$

$[A \lor \neg B \lor \neg C \lor D] \land \{[\neg A \lor (C \lor \neg D)] \land [B \lor (C \lor \neg D)]\}$

$(A \lor \neg B \lor \neg C \lor D) \land [(C \land \neg D) \lor \neg A] \land [(C \lor \neg D) \lor B]$

$(A \lor \neg B \lor \neg C \lor D) \land [(C \land \neg A) \land (\neg D \land \neg A)] \land [(C \land B) \land (\neg D \land B)]$

$\underset{(1)}{(A \lor \neg B \lor \neg C \lor D)} \land \underset{(2)}{(C \land \neg A)} \land \underset{(3)}{(\neg D \land \neg A)} \land \underset{(4)}{(C \land B)} \land \underset{(5)}{(\neg D \land B)}$

(1)

A	B	C	D	$A \lor \neg B \lor \neg C \lor D$	$A - B - C + D$	
0	0	0	0	1	0	
0	0	0	1	1	1	
0	0	1	0	1	-1	
0	0	1	1	1	0	
0	1	0	0	1	-1	
0	1	0	1	1	0	
0	1	1	0	0	-2	☆
0	1	1	1	1	-1	
1	0	0	0	1	1	
1	0	0	1	1	2	
1	0	1	0	1	0	
1	0	1	1	1	1	
1	1	0	0	1	0	
1	1	0	1	1	1	
1	1	1	0	1	-1	
1	1	1	1	1	0	

$w_A = 1 \quad w_B = -1 \quad w_C = -1 \quad w_D = 1$

$t = -1.5$

$g = A - B - C + D \geq -1.5$

② $(C \vee \neg A)$

A	C	$C \vee \neg A$	g
0	0	1	0
0	1	1	1
1	0	0 ☆	-1
1	1	1	0

$W_A = -1$

$W_C = 1$

$t = -0.5$

$g = C - A \geq -0.5$

③ $(\neg D \vee \neg A)$

A	D	$\neg D \vee \neg A$	g
0	0	1	0
0	1	1	-1
1	0	1	-1
1	1	0 ☆	-2

$W_A = -1$

$W_D = -1$

$t = -1.5$

$g = -D - A \geq -1.5$

④ $(C \vee B)$

B	C	$C \vee B$	g
0	0	0 ☆	0
0	1	1	1
1	0	1	1
1	1	1	2

$W_B = 1$

$W_C = 1$

$t = 0.5$

$g = B + C \geq 0.5$

⑤ $(\neg D \vee B)$

B	D	$\neg D \vee B$	g
0	0	1	0
0	1	0 ☆	-1
1	0	1	1
1	1	1	0

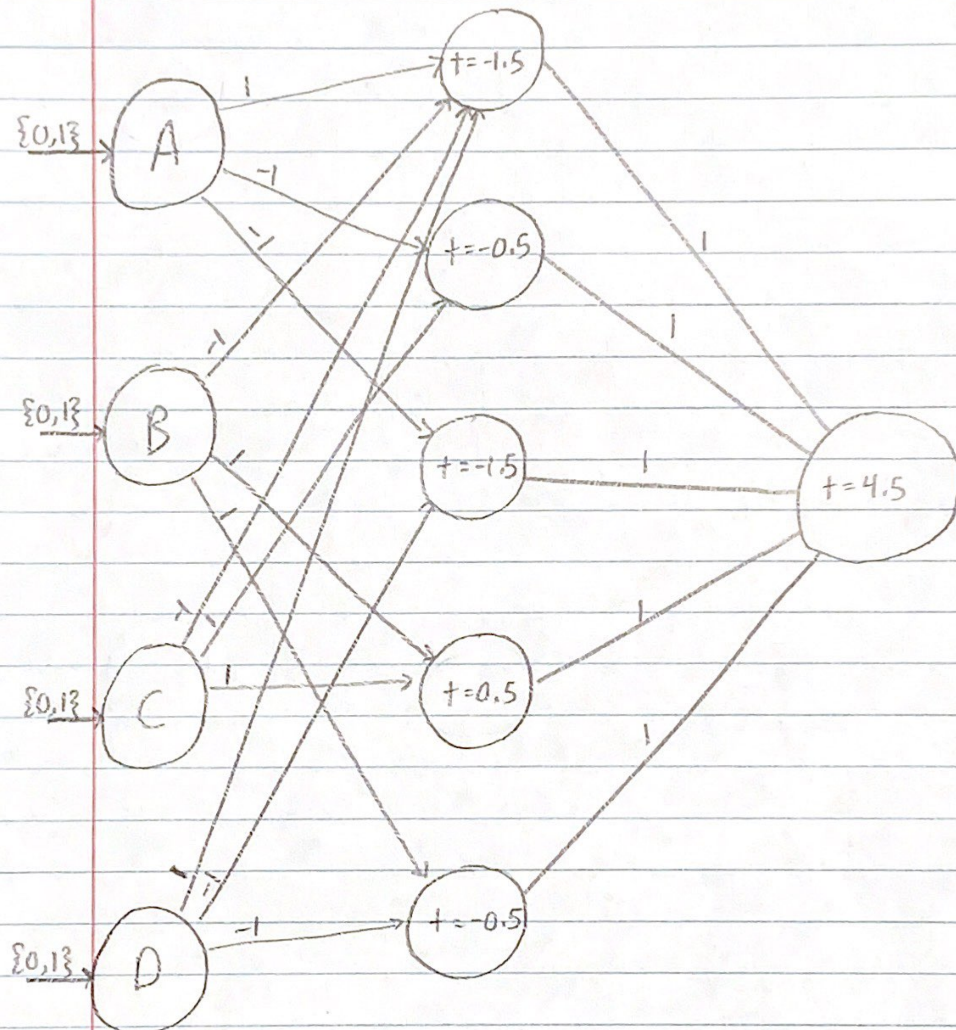
$W_B = 1$

$W_D = -1$

$t = -0.5$

$g = B - D \geq -0.5$

- We converted to CNF → we want every clause to be true. We will thus set our threshold at 4.5, and we only meet this threshold when every clause is true (outputs 1). All inputs are given a weight of 1.



* Weights are written above each of the lines.

3. $\theta_a = 0.9$ $\theta_{b|a} = 0.2$ $\theta_{\bar{b}|\bar{a}} = 0.5$ $\theta_{c|a} = 0.7$ $\theta_{c|\bar{a}} = 0.4$

$e_1 = \bar{a}, c$ $e_2 = \bar{a}, \bar{c}$ $e_3 = \bar{a}$

a) ① $\lambda_a = 0$ $\lambda_{\bar{a}} = 1$

$\lambda_c = 1$ $\lambda_{\bar{c}} = 0$

☆ $p^*(\bar{b}) = [(\theta_{\bar{c}|\bar{a}} \cdot \lambda_{\bar{c}}) + (\theta_{c|\bar{a}} \cdot \lambda_c)] \cdot [\lambda_{\bar{a}} \cdot \theta_{\bar{a}}] \cdot \theta_{\bar{b}|\bar{a}} +$

$[(\lambda_{\bar{c}} \cdot \theta_{\bar{c}|a}) + (\lambda_c \cdot \theta_{c|a})] \cdot [\lambda_a \cdot \theta_a] \cdot \theta_{\bar{b}|a}$

$p^*(\bar{b}) = [(0.6 \cdot 0) + (0.4 \cdot 1)] \cdot [1 \cdot 0.1] \cdot 0.5 +$

$[(0 \cdot 0.3) + (1 \cdot 0.7)] \cdot [0 \cdot 0.9] \cdot 0.8 = 0.02$

☆ $p^*(b) = [(\theta_{\bar{c}|\bar{a}} \cdot \lambda_{\bar{c}}) + (\theta_{c|\bar{a}} \cdot \lambda_c)] \cdot [\lambda_{\bar{a}} \cdot \theta_{\bar{a}}] \cdot \theta_{b|\bar{a}} +$

$[(\theta_{c|a} \cdot \lambda_c) + (\lambda_{\bar{c}} \cdot \theta_{\bar{c}|a})] \cdot [\lambda_a \cdot \theta_a] \cdot \theta_{b|a}$

$p^*(b) = [(0.6 \cdot 0) + (0.4 \cdot 1)] \cdot [1 \cdot 0.1] \cdot 0.5 +$ $p^*(\bar{b}) = 0.02$

$[(0.7 \cdot 1) + (0 \cdot 0.3)] \cdot [0 \cdot 0.9] \cdot 0.2 = 0.02$ $p^*(b) = 0.02$

② $\lambda_a = 0$ $\lambda_{\bar{a}} = 1$

$\lambda_c = 0$ $\lambda_{\bar{c}} = 1$

$p^*(\bar{b}) = [(0.6 \cdot 1) + (0.4 \cdot 0)] \cdot [1 \cdot 0.1] \cdot 0.5 +$

$[(1 \cdot 0.3) + (0 \cdot 0.7)] \cdot [0 \cdot 0.9] \cdot 0.8 = 0.03$ $p^*(b) = 0.03$

$p^*(b) = [(0.6 \cdot 1) + (0.4 \cdot 0)] \cdot [1 \cdot 0.1] \cdot 0.5 +$ $p^*(\bar{b}) = 0.03$

$[(0.7 \cdot 0) + (1 \cdot 0.3)] \cdot [0 \cdot 0.9] \cdot 0.2 = 0.03$

③ $\lambda_a = 0$ $\lambda_{\bar{a}} = 1$

$\lambda_c = 1$ $\lambda_{\bar{c}} = 1$

$p^*(\bar{b}) = [(0.6 \cdot 1) + (0.4 \cdot 1)] \cdot [1 \cdot 0.1] \cdot 0.5 +$

$[(1 \cdot 0.3) + (1 \cdot 0.7)] \cdot [0 \cdot 0.9] \cdot 0.8 = 0.05$

$p^*(b) = [(0.6 \cdot 1) + (0.4 \cdot 1) \cdot [1 \cdot 0.1] \cdot 0.5 +$ $p^*(\bar{b}) = 0.05$

$[(0.7 \cdot 1) + (1 \cdot 0.3)] \cdot [0 \cdot 0.9] \cdot 0.2 = 0.05$ $p^*(b) = 0.05$

b) The outputs of the circuit represent the probabilities

$P(b, e_1)$ $P(\bar{b}, e_1)$ $P(b, e_2)$ $P(\bar{b}, e_2)$ $P(b, e_3)$ $P(\bar{b}, e_3)$

These are the probabilities of query Q happening simultaneously with the various given pieces of evidence.

c) ① $Pr(\bar{b}|e_1) = \frac{Pr(\bar{b}, e_1)}{Pr(e_1)} = \frac{0.02}{Pr(e_1)} = \frac{0.02}{0.04} = \frac{1}{2}$

$Pr(e_1) = Pr(b, e_1) + Pr(\bar{b}, e_1)$

$Pr(e_1) = 0.02 + 0.02 = 0.04$

$\boxed{Pr(\bar{b}|e_1) = 0.5}$

② $Pr(\bar{b}|e_2) = \frac{Pr(\bar{b}, e_2)}{Pr(e_2)} = \frac{0.03}{0.06} = \frac{1}{2}$

$Pr(e_2) = Pr(b, e_2) + Pr(\bar{b}, e_2)$

$= 0.03 + 0.03 = 0.06$

$\boxed{Pr(\bar{b}|e_2) = 0.5}$

③ $Pr(\bar{b}|e_3) = \frac{Pr(\bar{b}, e_3)}{Pr(e_3)} = \frac{0.05}{0.10} = \frac{1}{2}$

$Pr(e_3) = Pr(b, e_3) + Pr(\bar{b}, e_3)$

$= 0.05 + 0.05 = 0.10$

$\boxed{Pr(\bar{b}|e_3) = 0.5}$