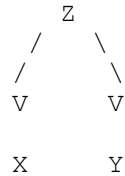


1a: MLE  $\operatorname{argmax}_h P(\text{data}|h)$  --- what makes the data most likely  
 MAP  $\operatorname{argmax}_h P(h|\text{data})$  --- what hypothesis does the data make most likely; requires prior on  $h$

1b:  $VC(H) = |X|$

1c (i):



1c(ii) 5 params

1c(iii) 7 params

1d : FALSE: estimate is optimistic because it might not be paying for training  
                   set noise

1e: TRUE

1f: TRUE (though not covered in 2003)

1g: FALSE: Both can hit local minima

1h: FALSE: MDPs have fully observed state whereas HMMs have observation symbols that are stochastically dependent on state

2a Initialize with if [Attribute value 1] then class

    Taking the next example...

        If it is correctly classified do nothing  
         Otherwise pace at the root one of the attributes with the corresponding classification

2b (Assuming lists are of depth  $d$ )

$$4^d * k! / (k - d)!$$

2c  $m \geq 1/\epsilon (\log |H| + \log (1/\delta))$  with  $\epsilon = 0.1$ ,  $\delta = 0.1$

2d  $|H| = 2^k$  and  $m \geq 10 (k \log 2 + \log 10)$

3a The centroids found by k-means can get pushed away from each other so the are further apart than the true means that generated the data.

3d One difference is that GMM elements may be long and thin

4a 0.8

4b 0.18

4c 0.44

4d 0.2 (Since  $P(\text{yell})$  is 0.2 no matter what the state)

4e HAAAA

5a: A and C

5b:

$$\begin{array}{ll} P(A \wedge B \mid C) = 1/8 & P(A \wedge B \mid \sim C) = 2/8 \\ P(\sim A \wedge B \mid C) = 4/8 & P(\sim A \wedge B \mid \sim C) = 1/8 \\ P(A \wedge \sim B \mid C) = 0 & P(A \wedge \sim B \mid \sim C) = 5/8 \\ P(\sim A \wedge \sim B \mid C) = 3/8 & P(\sim A \wedge \sim B \mid \sim C) = 0 \end{array}$$

$$P(C) = 1/2 \quad P(\sim C) = 1/2$$

5c:

$$P(C) = 1/2 \quad P(\sim C) = 1/2$$

$$P(A \mid C) = 1/8 \quad P(A \mid \sim C) = 7/8$$

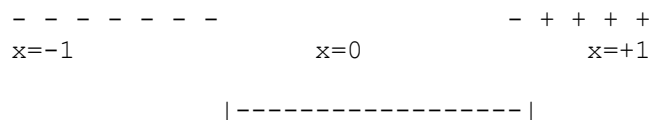
$P(B \mid C) = 5/8 \quad P(B \mid \sim C) = 3/8$  [  $P(\sim A \mid \dots)$  and  $P(\sim B \mid \dots)$  defined implicitly]

5d:  $C=1$

5e:  $C=1$

6b:  $w = (0, 2)$   $b = -5$

6c: It's possible



For small  $C$  would choose this margin

7a: 3

7b:  $30/7$

7c:  $18/4$

7d:  $18/4$

7e: Yes. TRAINING set error is zero because each point is closest to itself

7f: Same answer

8a: 0

8b:  $3/8$

8c:  $1/4$

8d:  $1/4$

8e:  $1/4$

8f: 0

9a: No arcs

9b: A ---> C <---- B

9c: Full connection

10,11: Am trying to find the figure that goes with this question!!