Istanbul Technical University- Fall 2017
Machine Learning
Homework4

Aydin Ayanzadeh
student number: 504161503
ayanzadeh17@itu.edu.tr
a.ayanzadeh@gmail.com

December 28, 2017

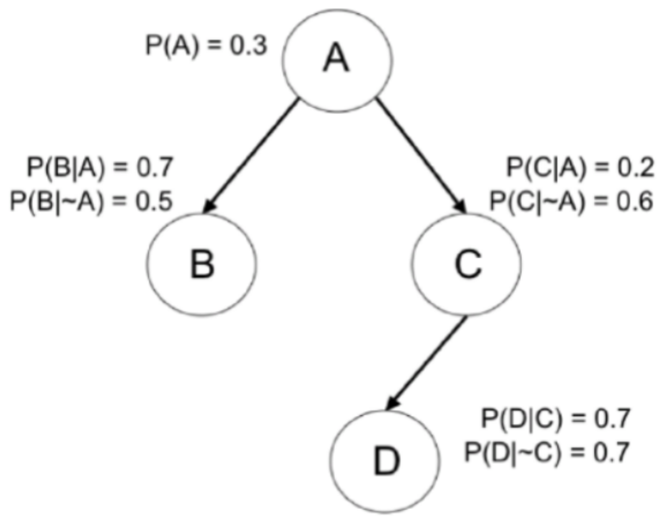

Figure 1: Bayesian network(Graphical Model)

Q1)

$a)p(A,B,C,D) = P(D|C)P(C|A)P(B|A)P(A) = 0.7*0.2*0.7*0.3 = 0.0294$

$b)p(A|B) = \dfrac{P(B,A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|\ A)P(\ A)}$

$= \frac{0.7*0.3}{0.7*0.3+0.5*0.7} = 0.375$

$c)P(C|B) = \dfrac{P(C,B)}{P(B)} = \frac{P(B)P(C)}{P(B)} = P(C)$
$= P(C|A)*P(A) + P(C|\sim A)*P(\sim A) = 0.2*0.3 + 0.6*0.7 = 0.48$

Q2

a) $\{O_1 = a, O_2 = a, O_3 = b\}$, we calculate $P(O|A,B,\Pi)$.

Initialization steps:

$$\alpha_{t=1}(S_1) = \pi_{S_1} b_{S_1}(a) \quad = 0.9*0.1 = 0.09 \tag{1}$$

$$\alpha_{t=1}(S_2) = \pi_{S_2} b_{S_2}(a) = 0.1*0.9 = 0.09 \tag{2}$$

Recursion Steps:

$$\alpha_{t=2}(S_1) = [\Sigma_{i=1}^{2}\alpha_{t=1}(i)a_{i1}]b_{S_1}(a) = (0.09 * 0.8 + 0.09 * 0.2) * 0.1 = 0.009 \tag{3}$$

$$\alpha_{t=2}(S_2) = [\Sigma_{i=1}^{2}\alpha_{t=1}(i)a_{i2}]b_{S_2}(a) = (0.09 * 0.2 + 0.09 * 0.8) * 0.9 = 0.081 \tag{4}$$

$$\alpha_{t=3}(S_1) = [\Sigma_{i=1}^{2}\alpha_{t=2}(i)a_{i1}]b_{S_1}(b) = [0.09 * 0.8 + 0.81 * 0.2]0.9 = 0.2106 \tag{5}$$

$$\alpha_{t=3}(S_2) = [\Sigma_{i=1}^{2}\alpha_{t=2}(i)a_{i2}]b_{S_2}(b) = (0.09 * 0.2 + 0.81 * 0.8) * 0.1 = 0.0666 \tag{6}$$

$$p(O|A, B, \pi) = \Sigma_{i=1}^{2}\alpha_T(i) = \alpha_{t=3}(S_1) + \alpha_{t=3}(S_2) = 0.0164 + 0.065 = 0.2772 \tag{7}$$

Q2 b) Viterbi Algorithm for state sequences O = (a, a, b).

Initialization Steps:

$$\begin{aligned} \delta_{t=1}(1) &= \pi_1 b_1(a) = 0.9 * 0.1 = 0.09 \\ \delta_{t=1}(2) &= \pi_2 b_2(a) = 0.1 * 0.9 = 0.09 \\ \\ \psi_{t=1}(1) &= 0 \\ \psi_{t=1}(2) &= 0 \end{aligned} \tag{8}$$

Recursion Steps:

$$\begin{aligned} \delta_{t=2}(1) &= \max_i \delta_{t=1}(i)a_{i1}.b_1(a) = max(\delta_{t=1}(1)a_{11}.b_1(a), \delta_{t=1}(2)a_{21}.b_1(a)) \\ &= max(0.09 * 0.8 * 0.1, 0.09 * 0.2 * 0.1) = max(0.0072, 0.0018) = 0.0072 \\ \\ \delta_{t=2}(2) &= \max_i \delta_{t=1}(i)a_{i2}.b_2(a) = max(\delta_{t=1}(1)a_{12}.b_2(a), \delta_{t=1}(2)a_{22}.b_2(a)) \\ &= max(0.09 * 0.2 * 0.9, 0.09 * 0.8 * 0.9) = max(0.0162, 0.0648) = 0.0648 \\ \\ \delta_{t=3}(1) &= \max_i \delta_{t=2}(i)a_{i1}.b_1(b) = max(\delta_{t=2}(1)a_{11}.b_1(b), \delta_{t=2}(2)a_{21}.b_1(b)) \\ &= max(0.0072 * 0.8 * 0.9, 0.0648 * 0.2 * 0.9) = max(0.005184, 0.011664) = 0.011664 \\ \\ \delta_{t=3}(2) &= \max_i \delta_{t=2}(i)a_{i2}.b_2(b) = max(\delta_{t=2}(1)a_{12}.b_2(b), \delta_{t=2}(2)a_{22}.b_2(b)) \\ &= max(0.0072 * 0.2 * 0.1, 0.0648 * 0.8 * 0.1) = max(0.000144, 0.005184) = 0.005184 \end{aligned} \tag{9}$$

$$
\begin{aligned}
\psi_{t=2}(1) &= \underset{i}{argmax}(\delta_{t=1}(i)a_{i1}) = argmax(\delta_{t=1}(1)a_{11}, \delta_{t=1}(2)a_{21}) \\
&= argmax(0.09 * 0.8, 0.9 * 0.2) = argmax(0.072, 0.018) = S_1 \\
\\
\psi_{t=2}(2) &= \underset{i}{argmax}(\delta_{t=1}(i)a_{i2}) = argmax(\delta_{t=1}(1)a_{12}, \delta_{t=1}(2)a_{22}) \\
&= argmax(0.09 * 0.2, 0.09 * 0.8) = argmax(0.018, 0.072) = S_2 \\
\\
\psi_{t=3}(1) &= \underset{i}{argmax}(\delta_{t=2}(i)a_{i1}) = argmax(\delta_{t=2}(1)a_{11}, \delta_{t=2}(2)a_{21}) \\
&= argmax(0.0072 * 0.8, 0.0648 * 0.2) = argmax(0.00576, 0.01296) = S_2 \\
\\
\psi_{t=3}(2) &= \underset{i}{argmax}(\delta_{t=2}(i)a_{i2}) = argmax(\delta_{t=2}(1)a_{12}, \delta_{t=2}(2)a_{22}) \\
&= argmax(0.0072 * 0.2, 0.0648 * 0.8) = argmax(0.00144, 0.05184) \\
&= S_2
\end{aligned} \tag{10}
$$

Termination Step:

$$p^* = \max_i(\delta_{t=3}(i)) = max(0.011664, 0.005184) = 0.011664$$

$$q^*_{t=3} = \underset{i}{argmax}(\delta_{t=3}(i)) = argmax(0.011664, 0.005184) = S_1 \tag{11}$$

Path:

$$q^*_{t=2} = \psi_{t=3}(q^*_{t=3}) = \psi_{t=3}(S_1) = S_2$$

$$q^*_{t=1} = \psi_{t=2}(q^*_{t=2}) = \psi_{t=2}(S_2) = S_2 \tag{12}$$

basee on the Viterbi algorithm, after using this algorithm the sequence state given O whith the most probability of state sequence is $Q = S_2S_2S_1$.