UNICAMP

Unicamp - Instituto de Computação

MC886 B / MO416 A

Introdução à Inteligência Artificial

Prof. Dr. Jacques Wainer

Exercício 5

Aluno:
George Gigilas Junior - 216741

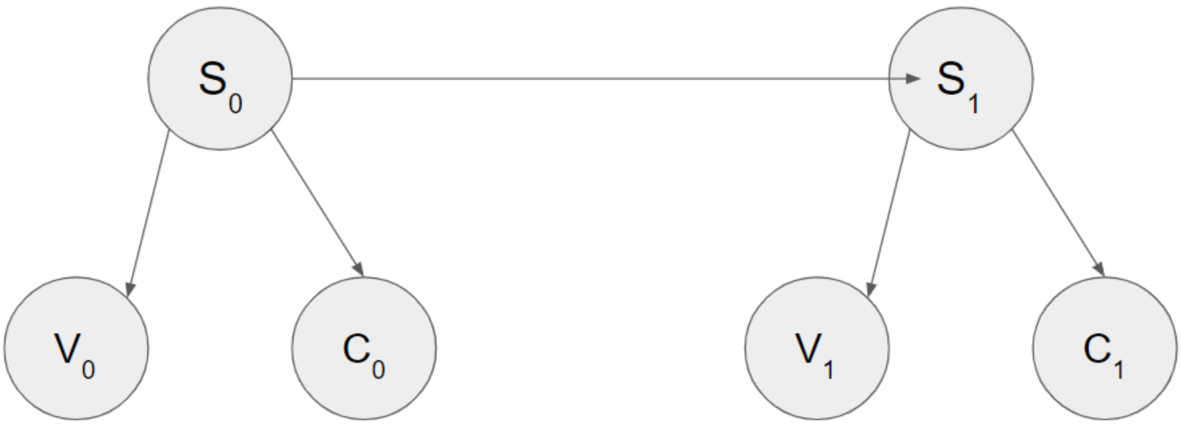
Introdução

Este exercício corresponde aos exercícios 15.13 e 15.14 do livro texto, cujos enunciados estão disponíveis juntamente com a atividade. Este arquivo corresponde às minhas soluções para os itens Q1, Q2, Q3, Q4 e Q5 do problema.

Solução

Item Q1 - formular a informação do problema como uma rede bayesiana que o professor poderia usar para filtragem ou previsão a partir de uma sequência de observações.

Para a representação, vou convencionar V como sendo a variável que indica olhos vermelhos, S como sendo a variável que indica sono suficiente e C como sendo a variável que indica se o aluno dormiu em classe. Segue a rede bayesiana:

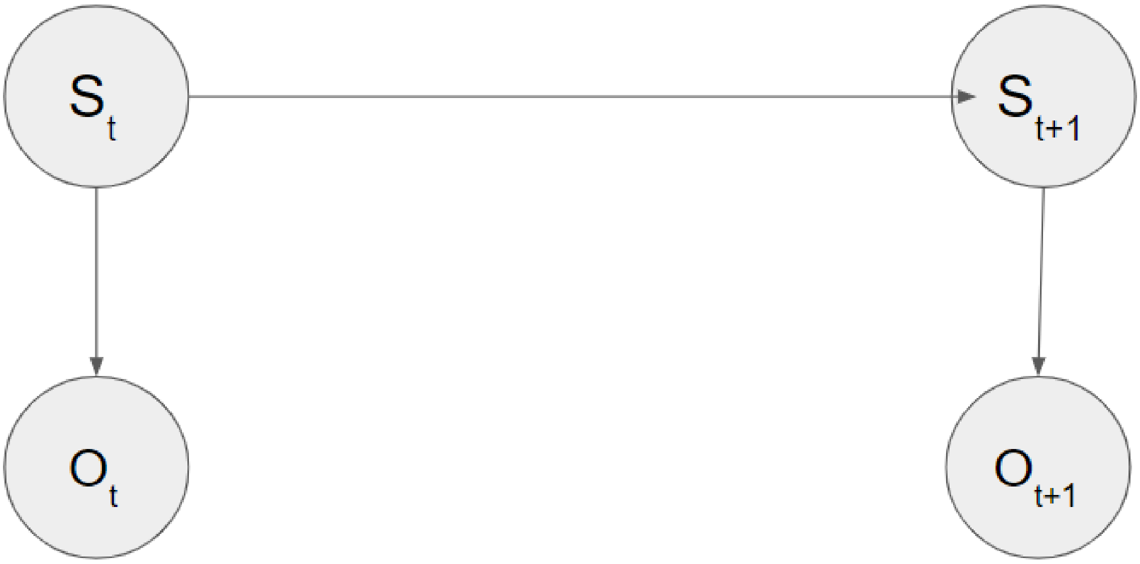


Para poder usá-la para filtragem ou previsão, é necessário fornecer $P(S_0 = true)$, $P(S_1=true|S_0)$, $P(V_0 = true|S_0)$ e $P(C_0 = true|S_0)$. Seguem esses valores:

- $P(S_0 = true) = 0,6$
- $P(S_1 = true| S_0 = true) = 0,8$
- $P(S_1 = true| S_0 = false) = 0,2$
- $P(V_0 = true| S_0 = true) = 0,2$
- $P(V_0 = true| S_0 = false) = 0,7$
- $P(C_0 = true| S_0 = true) = 0,1$
- $P(C_0 = true| S_0 = false) = 0,3$

Item Q2 - reformular como uma hidden Markov model com uma única variável de observação. Dê as tabelas de probabilidade completas para o modelo.

Para reformular como uma hidden Markov model (HMM), é necessário utilizar a técnica de compressão, porque as HMM possuem somente 1 variável de observação e 1 variável de estado. Para isso, vou convencionar O como sendo a variável de observação que combina V e C. A variável O pode assumir 4 valores (T = true e F = false): TT, se V = T e C = T; TF, se V = T e C = F; FT, se V = F e C = T; e FF, se V = F e C = F. Assim, temos a seguinte HMM:



Agora, precisamos construir as tabelas de prior para o estado, as tabelas de transição e o modelo de sensor. Seguem abaixo:

- Tabela de prior para o estado ($\mathbf{P}(S_0)$):

S_0	Probabilidade
T	0,6
F	0,4

- Tabela de transição ($\widehat{P}(S_{t+1} \mid S_t)$):

S_t	S_{t+1} = T	S_{t+1} = F
T	0,8	0,2
F	0,2	0,8

- Modelo de sensor ($\mathbf{P}(O_t \mid S_t)$):

S_t	O_t = TT	O_t = TF	O_t = FT	O_t = FF
T	0,2*0,1 = 0,02	0,2*0,9 = 0,18	0,8*0,1 = 0,08	0,8*0,9 = 0,72
F	0,7*0,3 = 0,21	0,7*0,7 = 0,49	0,3*0,3 = 0,09	0,3*0,7 = 0,21

Item Q3 - estimar o valor de estado (filtragem): calcular $P(EnoughSleep_t \mid e_{1:t})$, para t = 1, 2 e 3. Agora são convencionados: e1 = ¬V e ¬C, e2 = V e ¬C, e3 = V e C.

Para resolver essa questão, precisamos aplicar o algoritmo de forward, como visto em aula. Vou convencionar EnoughSleep como sendo S, e1 como sendo FF, e2 como sendo TF, e3 como sendo TT.

- Para t = 1:

  $\mathbf{P}(S_1 \mid e_{1:1}) = \mathbf{P}(S_1 \mid e_1) = \mathbf{P}(S_1 \mid O_1 = FF) = < \alpha_1(T), \alpha_1(F) > * \frac{1}{P(O1)}$

  $\alpha_1(T) = \pi_T * b_T(O_1)$
  - $b_T(O_1) = P(O_1 = FF \mid S_1 = T) = 0{,}72$
  - $\pi_T = P(S_1 = T) = P(S_1 = T \mid S_0 = T)*P(S_0 = T) + P(S_1 = T \mid S_0 = F)*P(S_0 = F)$
    $= 0{,}8 * 0{,}6 + 0{,}2 * 0{,}4 = 0{,}48 + 0{,}08 = 0{,}56$

  $\alpha_1(T) = 0{,}72 * 0{,}56 = 0{,}4032$

  $\alpha_1(F) = \pi_F * b_F(O_1)$
  - $b_F(O_1) = P(O_1 = FF \mid S_1 = F) = 0{,}21$
  - $\pi_F = P(S_1 = F) = P(S_1 = F \mid S_0 = T)*P(S_0 = T) + P(S_1 = F \mid S_0 = F)*P(S_0 = F)$
    $= 0{,}2 * 0{,}6 + 0{,}8 * 0{,}4 = 0{,}12 + 0{,}32 = 0{,}44$

  $\alpha_1(F) = 0{,}21 * 0{,}44 = 0{,}0924$

  $P(O_1) = P(O_1 \text{ e } S_1 = T) + P(O_1 \text{ e } S_1 = F) = \alpha_1(T) + \alpha_1(F) = 0{,}4032 + 0{,}0924 = 0{,}4956$

  Portanto, $\mathbf{P}(S_1 \mid e_1) = < \frac{0{,}4032}{0{,}4956}, \frac{0{,}0924}{0{,}4956} > = <0{,}8136; 0{,}1864>$

- Para t = 2:

  $\mathbf{P}(S_2 \mid e_{1:2}) = < \alpha_2(T), \alpha_2(F) > * \frac{1}{P(O)}$

  $\alpha_2(T) = [\sum_{j=1}^{N} \alpha 1(j) \ * \ ajT] \ * \ bT(O2)$

- $a_{jT} = P(S_2 = T \mid S_1 = j)$
  - j = F: $a_{FT} = P(S_2 = T \mid S_1 = F) = 0{,}2$
  - j = T: $a_{TT} = P(S_2 = T \mid S_1 = T) = 0{,}8$

- $b_T(O_2) = P(O_2 = TF \mid S_2 = T) = 0{,}18$

$\alpha_2(T) = [\alpha 1(F) * a_{FT} + \alpha 1(T) * a_{TT}] * b_T(O_2) = [0{,}0924 * 0{,}2 + 0{,}4032 * 0{,}8] * 0{,}18 = 0{,}0613872$

$\alpha_2(F) = [\sum_{j=1}^{N} \alpha 1(j) \; * \; ajF] \; * \; bF(O2)]$

- $a_{jF} = P(S_2 = F \mid S_1 = j)$
  - j = F: $a_{FF} = P(S_2 = F \mid S_1 = F) = 0{,}8$
  - j = T: $a_{TF} = P(S_2 = F \mid S_1 = T) = 0{,}2$

- $b_F(O_2) = P(O_2 = TF \mid S_2 = F) = 0{,}49$

$\alpha_2(F) = [\alpha 1(F) * a_{FF} + \alpha 1(T) * a_{TF}] * b_F(O_2) = [0{,}0924 * 0{,}8 + 0{,}4032 * 0{,}2] * 0{,}49 = 0{,}0757344$

$P(O_2) = \alpha_2(T) + \alpha_2(F) = 0{,}0613872 + 0{,}0757344 = 0{,}1371216$

Portanto, $\mathbf{P}(S_2 \mid e_{1:2}) = < \frac{0{,}0613872}{0{,}1371216}, \frac{0{,}0757344}{0{,}1371216} > = <0{,}4477; 0{,}5523>$

- Para t = 3:

$\mathbf{P}(S_3 \mid e_{1:3}) = < \alpha_3(T), \alpha_3(F) > * \frac{1}{P(O)}$

$\alpha_3(T) = [\sum_{j=1}^{N} \alpha 2(j) \; * \; ajT] \; * \; bT(O3)$

- $a_{jT} = P(S_3 = T \mid S_2 = j)$
  - j = F: $a_{FT} = P(S_3 = T \mid S_2 = F) = 0{,}2$
  - j = T: $a_{TT} = P(S_3 = T \mid S_2 = T) = 0{,}8$

- $b_T(O_3) = P(O_3 = TT \mid S_3 = T) = 0{,}02$

$\alpha_3(T) = [\alpha 2(F) * a_{FT} + \alpha 2(T) * a_{TT}] * b_T(O_3) = [0{,}0757344 * 0{,}2 + 0{,}0613872 * 0{,}8] * 0{,}02 = 0{,}00128513$

$\alpha_3(F) = [\sum_{j=1}^{N} \alpha 2(j) \; * \; ajF] \; * \; bF(O3)]$

- $a_{jF} = P(S_3 = F \mid S_2 = j)$
  - j = F: $a_{FF} = P(S_3 = F \mid S_2 = F) = 0{,}8$

- ◦ j = T: $a_{TF} = P(S_3 = F \mid S_2 = T) = 0,2$
- $b_F(O_3) = P(O_3 = TT \mid S_2 = F) = 0,21$

$\alpha_3(F) = [\alpha 2(F) * a_{FF} + \alpha 2(T) * a_{TF}] * b_F(O_3) = [0,0757344 * 0,8 + 0,0613872 * 0,2] * 0,21 = 0,01530164$

$P(O_3) = \alpha_3(T) + \alpha_3(F) = 0,00128513 + 0,01530164 = 0,01658677$

Portanto, $\mathbf{P}(S_3 \mid e_{1:3}) = < \frac{0,00128513}{0,01658677}, \frac{0,01530164}{0,01658677} > = <0,0775; 0,9225>$

Item Q4 - suavização: calcular $P(EnoughSleep_t \mid e_{1:3})$, para t = 1, 2 e 3.

Para resolver essa questão, precisamos aplicar o algoritmo de backward, como visto em aula, juntamente com os resultados de forward calculados no item anterior. Novamente, vou convencionar EnoughSleep como sendo S, e1 como sendo FF, e2 como sendo TF, e3 como sendo TT.

- Para t = 1:

  $\mathbf{P}(S_1 \mid e_{1:3}) = < P(S_1 = T \mid O_1, O_2, O_3), P(S_1 = F \mid O_1, O_2, O_3) >$
  $= <\alpha 1(T) * \beta_1(T), \alpha 1(F) * \beta_1(F) > * \frac{1}{P(O)}$

  Note que $\alpha 1(T)$ e $\alpha 1(F)$ já foram calculados no item anterior.

  $\beta_1(T) = \sum_{j=1}^{N} aTj \ * \ bj(O2) \ * \ \beta 2(j)$

  - $\beta_2(T) = \sum_{j=1}^{N} aTj \ * \ bj(O3) \ * \ \beta 3(j)$
    - ◦ $\beta 3(T) = 1$ (pois é o último valor de t)
    - ◦ j = T:
      - $a_{TT} = P(S_3 = T \mid S_2 = T) = 0,8$
      - $b_T(O_3) = P(O_3 = TT \mid S_3 = T) = 0,02$
    - ◦ j = F:
      - $a_{TF} = P(S_3 = F \mid S_2 = T) = 0,2$
      - $b_F(O_3) = P(O_3 = TT \mid S_3 = F) = 0,21$

    $\beta_2(T) = 0,8*0,02*1 + 0,2*0,21*1 = 0,058$
  - $\beta_2(F) = \sum_{j=1}^{N} aFj \ * \ bj(O3) \ * \ \beta 3(j)$
    - ◦ $\beta 3(F) = 1$ (pois é o último valor de t)
    - ◦ j = T:
      - $a_{FT} = P(S_3 = T \mid S_2 = F) = 0,2$
      - $b_T(O_3) = P(O_3 = TT \mid S_3 = T) = 0,02$
    - ◦ j = F:

- $a_{FF} = P(S_3 = F \mid S_2 = F) = 0,8$
- $b_F(O_3) = P(O_3 = TT \mid S_3 = F) = 0,21$

$\beta_2(F) = 0,2*0,02*1 + 0,8*0,21*1 = 0,172$

- $a_{Tj} = P(S_3 = j \mid S_2 = T)$
  - j = T: $a_{TT} = P(S_3 = T \mid S_2 = T) = 0,8$
  - j = F: $a_{TF} = P(S_3 = F \mid S_2 = T) = 0,2$
- $b_j(O_2) = P(O_2 = TF \mid S_2 = j)$
  - j = T: $b_T(O_2) = P(O_2 = TF \mid S_2 = T) = 0,18$
  - j = F: $b_F(O_2) = P(O_2 = TF \mid S_2 = F) = 0,49$

$\beta_1(T) = a_{TT} * b_T(O_2) * \beta_2(T) + a_{TF} * b_F(O_2) * \beta_2(F) = 0,8*0,18*0,058 + 0,2*0,49*0,172 = 0,025208$

$\beta_1(F) = \sum_{j=1}^{N} aFj \;*\; bj(O2) \;*\; \beta 2(j)$

- $\beta_2(T) = 0,058$
- $\beta_2(F) = 0,172$
- $a_{Fj} = P(S_3 = j \mid S_2 = F)$
  - j = T: $a_{FT} = P(S_3 = T \mid S_2 = F) = 0,2$
  - j = F: $a_{FF} = P(S_3 = F \mid S_2 = F) = 0,8$
- $b_j(O_2) = P(O_2 = TF \mid S_2 = j)$
  - j = T: $b_T(O_2) = P(O_2 = TF \mid S_2 = T) = 0,18$
  - j = F: $b_F(O_2) = P(O_2 = TF \mid S_2 = F) = 0,49$

$\beta_1(F) = a_{FT} * b_T(O_2) * \beta_2(T) + a_{FF} * b_F(O_2) * \beta_2(F) = 0,2*0,18*0,058 + 0,8*0,49*0,172 = 0,069512$

$P(O) = \beta_1(T)*\alpha_1(T) + \beta_1(F)+\alpha_1(F) = 0,4032 * 0,025208 + 0,0924 * 0,069512 = 0,0165867744$

Portanto, $\mathbf{P}(S_1 \mid e_{1:3}) = <\frac{0,4032 * 0,025208}{0,0165867744}, \frac{0,0924 * 0,069512}{0,0165867744}> = <0,6128, 0,3872>$

- Para t = 2:

$\mathbf{P}(S_2 \mid e_{1:3}) = < P(S_2 = T \mid O_1, O_2, O_3), P(S_2 = F \mid O_1, O_2, O_3) >$

$= <\alpha 2(T)*\beta_2(T), \alpha 2(F)*\beta_2(F) > * \frac{1}{P(O)}$

Note que $\alpha 2(T)$ e $\alpha 2(F)$ já foram calculados no item anterior.

- $\beta_2(T) = 0,058$
- $\beta_2(F) = 0,172$

$P(O) = \beta_2(T)*\alpha_2(T) + \beta_2(F)+\alpha_2(F) = 0,058*0,0613872 + 0,172*0,0757344 = 0,0165867744$

Portanto, $\mathbf{P}(S_2 \mid e_{1:3}) = <\frac{0,058*0,0613872}{0,0165867744}, \frac{0,172*0,0757344}{0,0165867744}> = <0,2147, 0,7853>$

- Para t = 3:
  $\mathbf{P}(S_3 \mid e_{1:3})$ = < $P(S_3 = T \mid O_1, O_2, O_3)$, $P(S_3 = F \mid O_1, O_2, O_3)$ >
  = <$\alpha 3(T)$*β_3(T), $\alpha 3(F)$*β_3(F) > * $\frac{1}{P(O)}$
  Note que $\alpha 3(T)$ e $\alpha 3(F)$ já foram calculados no item anterior.

  - $\beta_3(T) = 1$
  - $\beta_3(F) = 1$

  $P(O) = \beta_3(T)*\alpha_3(T) + \beta_3(F)+\alpha_3(F) = 1*0{,}00128513 + 1*0{,}01530164 = 0{,}01658677$

  Portanto, $\mathbf{P}(S_3 \mid e_{1:3})$ = < $\frac{0{,}00128513}{0{,}01658677}$, $\frac{0{,}01530164}{0{,}01658677}$> = <0,0775; 0,9225>

Item Q5 - comparar as probabilidades filtrada e suavizada para t = 1 e 3.

Quanto à probabilidade filtrada, ela leva em conta apenas o que foi observado no presente (e_t) e os estados anteriores (S_{t-1}, S_{t-2}, etc.). Já a probabilidade suavizada, leva em conta, também, os estados futuros (S_{t+1}, S_{t+2}, …, S_T) e o que foi observado no futuro (e_{t+1}, …, e_T).

- Para t = 1:
  - Filtrada: $\mathbf{P}(S_1 \mid e_1)$ = <0,8136; 0,1864>
  - Suavizada: $\mathbf{P}(S_1 \mid e_{1:3})$ = <0,6128, 0,3872>
  - Comparação: como a probabilidade filtrada leva em conta apenas o estado anterior S_0 e e_1, ela possui pouca variação com relação ao que é visto no modelo de sensor (é um dos primeiros estados). Assim, ela afirma com mais certeza que o aluno teve um sono suficiente no estado S_1. Já a probabilidade filtrada, inclui os estados futuros S_2 e S_3, além de e_2 e e_3, tentando suavizar a diferença entre S_1, S_2 e S_3. Porém, a probabilidade acaba sendo bastante afetada e ela não afirma, com a mesma certeza de antes, que o aluno teve sono suficiente em S_1, o que faz com que haja uma dúvida maior. Isso ocorre porque e_2 e e_3 começam a apresentar sintomas de sono, propagando isso para os estados anteriores, pela característica da suavização.

- Para t = 3:
  - Filtrada: $\mathbf{P}(S_3 \mid e_{1:3})$ = <0,0775; 0,9225>
  - Suavizada: $\mathbf{P}(S_3 \mid e_{1:3})$ = <0,0775; 0,9225>
  - Comparação: como não tem mais nenhum estado ou observação depois de t=3, a suavização não surte nenhum efeito na probabilidade $\mathbf{P}(S_3 \mid e_{1:3})$. Isso não significa que essa probabilidade está suavizada para todos os valores de t, significa apenas que, com os dados que temos, não tem como suavizá-la. Quanto ao valor obtido, a probabilidade de o aluno não ter tido sono suficiente é muito alta, indicando bastante confiança quanto ao que foi calculado.