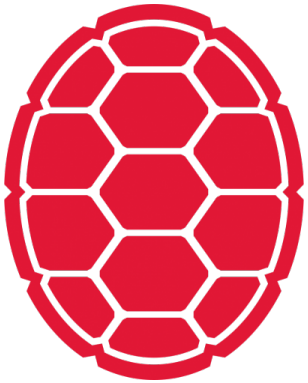

Hidden Markov Model

Natural Language Processing: Jordan Boyd-Graber

University of Maryland

EXAMPLE

(Estimated) HMM Parameters

$$\pi = \begin{bmatrix} 0.3 \\ 0.3 \\ 0.3 \\ 0.1 \end{bmatrix} \begin{matrix} \text{Det} \\ \text{Adj} \\ \text{N} \\ \text{V} \end{matrix} \qquad \theta = \begin{bmatrix} 0.1 & 0.4 & 0.45 & 0.05 \\ 0.1 & 0.3 & 0.5 & 0.1 \\ 0.05 & 0.05 & 0.1 & 0.8 \\ 0.3 & 0.2 & 0.3 & 0.2 \end{bmatrix} \begin{matrix} \text{Det} \\ \text{Adj} \\ \text{N} \\ \text{V} \end{matrix} \tag{1}$$

$$\beta = \begin{matrix} \\ \text{Det} \\ \text{Adj} \\ \text{N} \\ \text{V} \end{matrix} \begin{matrix} \begin{matrix} \text{the} & \text{old} & \text{man} & \text{blue} & \text{boat} & \text{a} & \text{an} \end{matrix} \\ \begin{pmatrix} 0.6 & 0.025 & 0.025 & 0.025 & 0.025 & 0.2 & 0.1 \\ 0.033 & 0.3 & 0.1 & 0.3 & 0.1 & 0.033 & 0.033 \\ 0.033 & 0.1 & 0.4 & 0.1 & 0.3 & 0.033 & 0.033 \\ 0.033 & 0.1 & 0.4 & 0.2 & 0.2 & 0.033 & 0.033 \end{pmatrix} \end{matrix} \tag{2}$$

Computing Score

What is the probability of the sequence “a/`Det` blue/`Adj` boat/`N`”?

Computing Score

What is the probability of the sequence "a/`Det` blue/`Adj` boat/`N`"?

$$\pi_d \beta_{d,the} \theta_{d,a} \beta_{a,blue} \theta_{a,n} \beta_{n,boat} = \tag{3}$$

$$0.3 * 0.6 * 0.4 * 0.3 * 0.5 * 0.1 = 0.00108 \tag{4}$$

Computing Score

What is the probability of the sequence "a/`Det` blue/`Adj` boat/`N`"?

$$\pi_d \beta_{d,the} \theta_{d,a} \beta_{a,blue} \theta_{a,n} \beta_{n,boat} = \tag{3}$$

$$0.3 * 0.6 * 0.4 * 0.3 * 0.5 * 0.1 = 0.00108 \tag{4}$$

Base Case

1. $\delta_1(a) =$

Base Case

1. $\delta_1(a) = -4.6$
2. $\delta_1(v) =$

Base Case

1. $\delta_1(a) = -4.6$
2. $\delta_1(v) = -5.7$
3. $\delta_1(d) =$

Base Case

1. $\delta_1(a) = -4.6$
2. $\delta_1(v) = -5.7$
3. $\delta_1(d) = -1.7$
4. $\delta_1(n) =$

Base Case

1. $\delta_1(a) = -4.6$
2. $\delta_1(v) = -5.7$
3. $\delta_1(d) = -1.7$
4. $\delta_1(n) = -4.6$

Position 2

1. $\delta_2(a) =$

Position 2

1. $\delta_2(a) = \max\left(\underbrace{-5.8}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-2.6}_{\textbf{d}}, \underbrace{-7.6}_{\text{n}}\right) + -1.2 = -2.6 + -1.2 = -3.8$

2. $\delta_2(v) =$

Position 2

1. $\delta_2(a) = \max\left(\underbrace{-5.8}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-2.6}_{\textbf{d}}, \underbrace{-7.6}_{\text{n}}\right) + -1.2 = -2.6 + -1.2 = -3.8$
2. $\delta_2(v) = \max\left(\underbrace{-6.9}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-4.7}_{\textbf{d}}, \underbrace{-4.8}_{\text{n}}\right) + -2.3 = -4.7 + -2.3 = -7.0$
3. $\delta_2(d) =$

Position 2

1. $\delta_2(a) = \max\left(\underbrace{-5.8}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-2.6}_{\textbf{d}}, \underbrace{-7.6}_{\text{n}}\right) + -1.2 = -2.6 + -1.2 = -3.8$
2. $\delta_2(v) = \max\left(\underbrace{-6.9}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-4.7}_{\textbf{d}}, \underbrace{-4.8}_{\text{n}}\right) + -2.3 = -4.7 + -2.3 = -7.0$
3. $\delta_2(d) = \max\left(\underbrace{-6.9}_{\text{a}}, \underbrace{-6.9}_{\text{v}}, \underbrace{-4.0}_{\textbf{d}}, \underbrace{-7.6}_{\text{n}}\right) + -3.7 = -4.0 + -3.7 = -7.7$
4. $\delta_2(n) =$

Position 2

1. $\delta_2(a) = \max\left(\underbrace{-5.8}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-2.6}_{\textbf{d}}, \underbrace{-7.6}_{\text{n}}\right) + -1.2 = -2.6 + -1.2 = -3.8$
2. $\delta_2(v) = \max\left(\underbrace{-6.9}_{\text{a}}, \underbrace{-7.3}_{\text{v}}, \underbrace{-4.7}_{\textbf{d}}, \underbrace{-4.8}_{\text{n}}\right) + -2.3 = -4.7 + -2.3 = -7.0$
3. $\delta_2(d) = \max\left(\underbrace{-6.9}_{\text{a}}, \underbrace{-6.9}_{\text{v}}, \underbrace{-4.0}_{\textbf{d}}, \underbrace{-7.6}_{\text{n}}\right) + -3.7 = -4.0 + -3.7 = -7.7$
4. $\delta_2(n) = \max\left(\underbrace{-5.3}_{\text{a}}, \underbrace{-6.9}_{\text{v}}, \underbrace{-2.5}_{\textbf{d}}, \underbrace{-6.9}_{\text{n}}\right) + -1.9 = -2.5 + -1.9 = -4.4$

Position 3

1. $\delta_3(a) =$

Position 3

1. $\delta_3(a) = \max\left(\underbrace{-5.0}_{\mathbf{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-8.6}_{\text{d}}, \underbrace{-7.4}_{\text{n}}\right) + -2.3 = -5.0 + -2.3 = -7.3$
2. $\delta_3(v) =$

Position 3

1. $\delta_3(a) = \max\left(\underbrace{-5.0}_{\mathbf{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-8.6}_{\text{d}}, \underbrace{-7.4}_{\text{n}}\right) + -2.3 = -5.0 + -2.3 = -7.3$
2. $\delta_3(v) = \max\left(\underbrace{-6.1}_{\text{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-4.6}_{\mathbf{n}}\right) + -0.9 = -4.6 + -0.9 = -5.5$
3. $\delta_3(d) =$

Position 3

1. $\delta_3(a) = \max\left(\underbrace{-5.0}_{\mathbf{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-8.6}_{\text{d}}, \underbrace{-7.4}_{\text{n}}\right) + -2.3 = -5.0 + -2.3 = -7.3$
2. $\delta_3(v) = \max\left(\underbrace{-6.1}_{\text{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-4.6}_{\mathbf{n}}\right) + -0.9 = -4.6 + -0.9 = -5.5$
3. $\delta_3(d) = \max\left(\underbrace{-6.1}_{\mathbf{a}}, \underbrace{-8.2}_{\text{v}}, \underbrace{-10.0}_{\text{d}}, \underbrace{-7.4}_{\text{n}}\right) + -3.7 = -6.1 + -3.7 = -9.8$
4. $\delta_3(n) =$

Position 3

1. $\delta_3(a) = \max\left(\underbrace{-5.0}_{\mathbf{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-8.6}_{\text{d}}, \underbrace{-7.4}_{\text{n}}\right) + -2.3 = -5.0 + -2.3 = -7.3$
2. $\delta_3(v) = \max\left(\underbrace{-6.1}_{\text{a}}, \underbrace{-8.6}_{\text{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-4.6}_{\mathbf{n}}\right) + -0.9 = -4.6 + -0.9 = -5.5$
3. $\delta_3(d) = \max\left(\underbrace{-6.1}_{\mathbf{a}}, \underbrace{-8.2}_{\text{v}}, \underbrace{-10.0}_{\text{d}}, \underbrace{-7.4}_{\text{n}}\right) + -3.7 = -6.1 + -3.7 = -9.8$
4. $\delta_3(n) = \max\left(\underbrace{-4.5}_{\mathbf{a}}, \underbrace{-8.2}_{\text{v}}, \underbrace{-8.5}_{\text{d}}, \underbrace{-6.7}_{\text{n}}\right) + -0.9 = -4.5 + -0.9 = -5.4$

Position 4

1. $\delta_4(a) =$

Position 4

1. $\delta_4(a) = \max\left(\underbrace{-8.5}_{\text{a}}, \underbrace{-7.2}_{\textbf{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-8.4}_{\text{n}}\right) + -3.4 = -7.2 + -3.4 = -10.6$
2. $\delta_4(v) =$

Position 4

1. $\delta_4(a) = \max\left(\underbrace{-8.5}_{\text{a}}, \underbrace{-7.2}_{\textbf{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-8.4}_{\text{n}}\right) + -3.4 = -7.2 + -3.4 = -10.6$
2. $\delta_4(v) = \max\left(\underbrace{-9.6}_{\text{a}}, \underbrace{-7.2}_{\text{v}}, \underbrace{-12.8}_{\text{d}}, \underbrace{-5.7}_{\textbf{n}}\right) + -3.4 = -5.7 + -3.4 = -9.1$
3. $\delta_4(d) =$

Position 4

1. $\delta_4(a) = \max\left(\underbrace{-8.5}_{\text{a}}, \underbrace{-7.2}_{\textbf{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-8.4}_{\text{n}}\right) + -3.4 = -7.2 + -3.4 = -10.6$
2. $\delta_4(v) = \max\left(\underbrace{-9.6}_{\text{a}}, \underbrace{-7.2}_{\text{v}}, \underbrace{-12.8}_{\text{d}}, \underbrace{-5.7}_{\textbf{n}}\right) + -3.4 = -5.7 + -3.4 = -9.1$
3. $\delta_4(d) = \max\left(\underbrace{-9.6}_{\text{a}}, \underbrace{-6.8}_{\textbf{v}}, \underbrace{-12.1}_{\text{d}}, \underbrace{-8.4}_{\text{n}}\right) + -0.5 = -6.8 + -0.5 = -7.3$
4. $\delta_4(n) =$

Position 4

1. $\delta_4(a) = \max\left(\underbrace{-8.5}_{\text{a}}, \underbrace{-7.2}_{\textbf{v}}, \underbrace{-10.7}_{\text{d}}, \underbrace{-8.4}_{\text{n}}\right) + -3.4 = -7.2 + -3.4 = -10.6$
2. $\delta_4(v) = \max\left(\underbrace{-9.6}_{\text{a}}, \underbrace{-7.2}_{\text{v}}, \underbrace{-12.8}_{\text{d}}, \underbrace{-5.7}_{\textbf{n}}\right) + -3.4 = -5.7 + -3.4 = -9.1$
3. $\delta_4(d) = \max\left(\underbrace{-9.6}_{\text{a}}, \underbrace{-6.8}_{\textbf{v}}, \underbrace{-12.1}_{\text{d}}, \underbrace{-8.4}_{\text{n}}\right) + -0.5 = -6.8 + -0.5 = -7.3$
4. $\delta_4(n) = \max\left(\underbrace{-8.0}_{\text{a}}, \underbrace{-6.8}_{\textbf{v}}, \underbrace{-10.6}_{\text{d}}, \underbrace{-7.7}_{\text{n}}\right) + -3.4 = -6.8 + -3.4 = -10.2$

Position 5

1. $\delta_5(a) =$

Position 5

1. $\delta_5(a) = \max\left(\underbrace{-11.8}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-8.2}_{\textbf{d}}, \underbrace{-13.2}_{\text{n}}\right) + -2.3 = -8.2 + -2.3 = -10.5$
2. $\delta_5(v) =$

Position 5

1. $\delta_5(a) = \max\left(\underbrace{-11.8}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-8.2}_{\textbf{d}}, \underbrace{-13.2}_{\text{n}}\right) + -2.3 = -8.2 + -2.3 = -10.5$
2. $\delta_5(v) = \max\left(\underbrace{-12.9}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-10.3}_{\textbf{d}}, \underbrace{-10.4}_{\text{n}}\right) + -1.6 = -10.3 + -1.6 = -11.9$
3. $\delta_5(d) =$

Position 5

1. $\delta_5(a) = \max\left(\underbrace{-11.8}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-8.2}_{\textbf{d}}, \underbrace{-13.2}_{\text{n}}\right) + -2.3 = -8.2 + -2.3 = -10.5$
2. $\delta_5(v) = \max\left(\underbrace{-12.9}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-10.3}_{\textbf{d}}, \underbrace{-10.4}_{\text{n}}\right) + -1.6 = -10.3 + -1.6 = -11.9$
3. $\delta_5(d) = \max\left(\underbrace{-12.9}_{\text{a}}, \underbrace{-10.3}_{\text{v}}, \underbrace{-9.6}_{\textbf{d}}, \underbrace{-13.2}_{\text{n}}\right) + -3.7 = -9.6 + -3.7 = -13.3$
4. $\delta_5(n) =$

Position 5

1. $\delta_5(a) = \max\left(\underbrace{-11.8}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-8.2}_{\textbf{d}}, \underbrace{-13.2}_{\text{n}}\right) + -2.3 = -8.2 + -2.3 = -10.5$
2. $\delta_5(v) = \max\left(\underbrace{-12.9}_{\text{a}}, \underbrace{-10.7}_{\text{v}}, \underbrace{-10.3}_{\textbf{d}}, \underbrace{-10.4}_{\text{n}}\right) + -1.6 = -10.3 + -1.6 = -11.9$
3. $\delta_5(d) = \max\left(\underbrace{-12.9}_{\text{a}}, \underbrace{-10.3}_{\text{v}}, \underbrace{-9.6}_{\textbf{d}}, \underbrace{-13.2}_{\text{n}}\right) + -3.7 = -9.6 + -3.7 = -13.3$
4. $\delta_5(n) = \max\left(\underbrace{-11.3}_{\text{a}}, \underbrace{-10.3}_{\text{v}}, \underbrace{-8.1}_{\textbf{d}}, \underbrace{-12.5}_{\text{n}}\right) + -1.2 = -8.1 + -1.2 = -9.3$

Reconstruction

1. What is the most likely pos sequence for “the old man”?

Reconstruction

1. What is the most likely pos sequence for “the old man”?
   For “the old man”, the reconstruction starts with the best part of speech at Position 3, which is noun (-5.4), which has an adjective back pointer, which as a back pointer to determiner. The overall sequence is “The/det old/adj man/n”.
2. What is the most likely pos sequence for “the old man the boat”?

Reconstruction

1. What is the most likely pos sequence for “the old man”?
   For “the old man”, the reconstruction starts with the best part of speech at Position 3, which is noun (-5.4), which has an adjective back pointer, which as a back pointer to determiner. The overall sequence is “The/det old/adj man/n”.
2. What is the most likely pos sequence for “the old man the boat”?

For “the old man the boats”, the reconstruction starts with the best part of speech at Position 5, which is a noun (-9.3), which leads to the sequence “The/det old/n man/v the/det boats/n”.