Hidden Markov Models

Natural Language Processing: Jordan Boyd-Graber
University of Colorado Boulder
LECTURE 20

Structured Prediction

- Thus far, we've assumed observations are iid (or at least exchangeable)
- Common tool in computational biology, information extraction, sequence modeling
- I'm going to use linguistic examples (more in 5832), but think about how it might be used in your favorite applications

Content Questions

Administrivia

- Have everything for last HW
- EC added by special request
- Keep working on projects

In class . . .

- Finding most likely sequence
- Garden pathing
  - The prime number few
  - The cotton clothing is made of grows in Mississippi

In class . . .

- Finding most likely sequence
- Garden pathing
  - The prime number few
  - The cotton clothing is made of grows in Mississippi
  - We'll see how Viterbi decoding can be confused by (and solve) part of speech ambiguities

Model parameters

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

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

In class . . .

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

(3)

(4)

(5)

(6)

In class . . .

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

$$\log(\pi_d \beta_{d,a} \theta_{d,a} \beta_{a,blue} \theta_{a,n} \beta_{n,boat}) = \tag{3}$$
$$\tag{4}$$
$$\tag{5}$$
$$\tag{6}$$

In class . . .

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

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

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

$$\tag{5}$$

$$\tag{6}$$

In class . . .

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

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

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

$$-.5 + -.7 + -.4 + -.5 + -0.3 + -0.5 = -3.0 \tag{5}$$

$$\tag{6}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} & & \\ & & \\ & & \\ & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$\pi_{\text{START, A}} + \beta_{\text{A, the}} = -0.52 + -1.48 = -2.00$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & & \\ & & \\ & & \\ & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$\pi_{\text{START, V}} + \beta_{\text{V, the}} = -1.00 + -1.48 = -2.48$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & & \\ -2.48 & & \\ & & \\ & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$\pi_{\text{START, D}} + \beta_{\text{D, the}} = -0.52 + -0.22 = -0.74$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & & \\ -2.48 & & \\ -0.74 & & \\ & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$\pi_{\text{START, N}} + \beta_{\text{N, the}} = -0.52 + -1.48 = -2.00$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & & \\ -2.48 & & \\ -0.74 & & \\ -2.00 & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_0(D) + \theta_{\text{D, A}} + \beta_{\text{A, old}} = -0.74 + -0.40 + -0.52 = -1.67$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & \\ -2.48 & & \\ -0.74 & & \\ -2.00 & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_0(D) + \theta_{\text{D, V}} + \beta_{\text{V, old}} = -0.74 + -1.30 + -1.00 = -3.05$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & \\ -2.48 & -3.05 & \\ -0.74 & & \\ -2.00 & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_0(D) + \theta_{\text{D, D}} + \beta_{\text{D, old}} = -0.74 + -1.00 + -1.60 = -3.35$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & \\ -2.48 & -3.05 & \\ -0.74 & -3.35 & \\ -2.00 & & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_0(D) + \theta_{\text{D, N}} + \beta_{\text{N, old}} = -0.74 + -0.35 + -1.00 = -2.09$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & \\ -2.48 & -3.05 & \\ -0.74 & -3.35 & \\ -2.00 & -2.09 & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_1(A) + \theta_{\text{A, A}} + \beta_{\text{A, man}} = -1.67 + -0.52 + -1.00 = -3.19$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & \\ -0.74 & -3.35 & \\ -2.00 & -2.09 & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_1(N) + \theta_{\text{N, V}} + \beta_{\text{V, man}} = -2.09 + -0.10 + -0.40 = -2.59$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & -2.59 \\ -0.74 & -3.35 & \\ -2.00 & -2.09 & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_1(A) + \theta_{\text{A, D}} + \beta_{\text{D, man}} = -1.67 + -1.00 + -1.60 = -4.27$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & -2.59 \\ -0.74 & -3.35 & -4.27 \\ -2.00 & -2.09 & \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

$$\delta_1(A) + \theta_{\text{A, N}} + \beta_{\text{N, man}} = -1.67 + -0.30 + -0.40 = -2.36$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & -2.59 \\ -0.74 & -3.35 & -4.27 \\ -2.00 & -2.09 & -2.36 \end{pmatrix} \end{array} \tag{7}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & -2.59 \\ -0.74 & -3.35 & -4.27 \\ -2.00 & -2.09 & -2.36 \end{pmatrix} \end{array} \tag{7}$$

- Backpointers

$$\tau = \begin{array}{c} \\ \text{A} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cc} \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} D & A \\ D & N \\ D & A \\ D & A \end{pmatrix} \end{array} \tag{8}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & -2.59 \\ -0.74 & -3.35 & -4.27 \\ -2.00 & -2.09 & -2.36 \end{pmatrix} \end{array} \tag{7}$$

- Backpointers

$$\tau = \begin{array}{c} \\ \text{A} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cc} \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} D & A \\ D & N \\ D & A \\ D & A \end{pmatrix} \end{array} \tag{8}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccc} \text{the}_0 & \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 \\ -2.48 & -3.05 & -2.59 \\ -0.74 & -3.35 & -4.27 \\ -2.00 & -2.09 & -2.36 \end{pmatrix} \end{array} \tag{7}$$

- Backpointers

$$\tau = \begin{array}{c} \\ \text{A} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cc} \text{old}_1 & \text{man}_2 \end{array} \\ \begin{pmatrix} D & A \\ D & N \\ D & A \\ D & A \end{pmatrix} \end{array} \tag{8}$$

- Reconstruction: D A N

Decoding Sentence 3

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\pi_{\text{START, A}} + \beta_{\text{A, the}} = -0.52 + -1.48 = -2.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\pi_{\text{START, V}} + \beta_{\text{V, the}} = -1.00 + -1.48 = -2.48$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & & & & \\ -2.48 & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\pi_{\text{START, D}} + \beta_{\text{D, the}} = -0.52 + -0.22 = -0.74$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & & & & \\ -2.48 & & & & \\ -0.74 & & & & \\ & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\pi_{\text{START, N}} + \beta_{\text{N, the}} = -0.52 + -1.48 = -2.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & & & & \\ -2.48 & & & & \\ -0.74 & & & & \\ -2.00 & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_0(D) + \theta_{\text{D, A}} + \beta_{\text{A, old}} = -0.74 + -0.40 + -0.52 = -1.67$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & & & \\ -2.48 & & & & \\ -0.74 & & & & \\ -2.00 & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_0(D) + \theta_{\text{D, V}} + \beta_{\text{V, old}} = -0.74 + -1.30 + -1.00 = -3.05$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & & & \\ -2.48 & -3.05 & & & \\ -0.74 & & & & \\ -2.00 & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_0(D) + \theta_{\text{D, D}} + \beta_{\text{D, old}} = -0.74 + -1.00 + -1.60 = -3.35$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & & & \\ -2.48 & -3.05 & & & \\ -0.74 & -3.35 & & & \\ -2.00 & & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_0(D) + \theta_{\text{D, N}} + \beta_{\text{N, old}} = -0.74 + -0.35 + -1.00 = -2.09$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & & & \\ -2.48 & -3.05 & & & \\ -0.74 & -3.35 & & & \\ -2.00 & -2.09 & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_1(A) + \theta_{\text{A, A}} + \beta_{\text{A, man}} = -1.67 + -0.52 + -1.00 = -3.19$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & & \\ -2.48 & -3.05 & & & \\ -0.74 & -3.35 & & & \\ -2.00 & -2.09 & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_1(N) + \theta_{\text{N, V}} + \beta_{\text{V, man}} = -2.09 + -0.10 + -0.40 = -2.59$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & & \\ -2.48 & -3.05 & -2.59 & & \\ -0.74 & -3.35 & & & \\ -2.00 & -2.09 & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_1(A) + \theta_{\text{A, D}} + \beta_{\text{D, man}} = -1.67 + -1.00 + -1.60 = -4.27$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & & \\ -2.48 & -3.05 & -2.59 & & \\ -0.74 & -3.35 & -4.27 & & \\ -2.00 & -2.09 & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_1(A) + \theta_{\text{A, N}} + \beta_{\text{N, man}} = -1.67 + -0.30 + -0.40 = -2.36$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 & \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & & & \\ -2.48 & -3.05 & -2.59 & & & \\ -0.74 & -3.35 & -4.27 & & & \\ -2.00 & -2.09 & -2.36 & & & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_2(V) + \theta_{\text{V, A}} + \beta_{\text{A, the}} = -2.59 + -0.70 + -1.48 = -4.77$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & \\ -2.48 & -3.05 & -2.59 & & \\ -0.74 & -3.35 & -4.27 & & \\ -2.00 & -2.09 & -2.36 & & \end{pmatrix} \end{array} \quad (9)$$

Decoding Sentence 3

$$\delta_2(N) + \theta_{\text{N, V}} + \beta_{\text{V, the}} = -2.36 + -0.10 + -1.48 = -3.94$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 \end{array} \quad \text{boat}_4 \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 \\ -2.48 & -3.05 & -2.59 & -3.94 \\ -0.74 & -3.35 & -4.27 & \\ -2.00 & -2.09 & -2.36 & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_2(V) + \theta_{\text{V, D}} + \beta_{\text{D, the}} = -2.59 + -0.52 + -0.22 = -3.33$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 \end{array} \quad \text{boat}_4 \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 \\ -2.48 & -3.05 & -2.59 & -3.94 \\ -0.74 & -3.35 & -4.27 & -3.33 \\ -2.00 & -2.09 & -2.36 & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_2(V) + \theta_{\text{V, N}} + \beta_{\text{N, the}} = -2.59 + -0.52 + -1.48 = -4.59$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 \end{array} \quad \text{boat}_4 \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 \\ -2.48 & -3.05 & -2.59 & -3.94 \\ -0.74 & -3.35 & -4.27 & -3.33 \\ -2.00 & -2.09 & -2.36 & -4.59 \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_3(D) + \theta_{\text{D, A}} + \beta_{\text{A, boat}} = -3.33 + -0.40 + -1.00 = -4.73$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & \\ -0.74 & -3.35 & -4.27 & -3.33 & \\ -2.00 & -2.09 & -2.36 & -4.59 & \end{pmatrix} \end{array} \quad (9)$$

Decoding Sentence 3

$$\delta_3(D) + \theta_{\text{D, V}} + \beta_{\text{V, boat}} = -3.33 + -1.30 + -0.70 = -5.33$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & -5.33 \\ -0.74 & -3.35 & -4.27 & -3.33 & \\ -2.00 & -2.09 & -2.36 & -4.59 & \end{pmatrix} \end{array} \tag{9}$$

Decoding Sentence 3

$$\delta_3(D) + \theta_{\text{D, D}} + \beta_{\text{D, boat}} = -3.33 + -1.00 + -1.60 = -5.93$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & -5.33 \\ -0.74 & -3.35 & -4.27 & -3.33 & -5.93 \\ -2.00 & -2.09 & -2.36 & -4.59 & \end{pmatrix} \end{array} \quad (9)$$

Decoding Sentence 3

$$\delta_3(D) + \theta_{\text{D, N}} + \beta_{\text{N, boat}} = -3.33 + -0.35 + -0.52 = -4.20$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & -5.33 \\ -0.74 & -3.35 & -4.27 & -3.33 & -5.93 \\ -2.00 & -2.09 & -2.36 & -4.59 & -4.20 \end{pmatrix} \end{array} \quad (9)$$

Decoding Sentence 3

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & -5.33 \\ -0.74 & -3.35 & -4.27 & -3.33 & -5.93 \\ -2.00 & -2.09 & -2.36 & -4.59 & -4.20 \end{pmatrix} \end{array} \tag{9}$$

- Backpointers

$$\tau = \begin{array}{c} \\ \text{A} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cccc} \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} D & A & V & D \\ D & N & N & D \\ D & A & V & D \\ D & A & V & D \end{pmatrix} \end{array} \tag{10}$$

Decoding Sentence 3

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & -5.33 \\ -0.74 & -3.35 & -4.27 & -3.33 & -5.93 \\ -2.00 & -2.09 & -2.36 & -4.59 & -4.20 \end{pmatrix} \end{array} \tag{9}$$

- Backpointers

$$\tau = \begin{array}{c} \\ \text{A} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cccc} \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} D & A & V & D \\ D & N & N & D \\ D & A & V & D \\ D & A & V & D \end{pmatrix} \end{array} \tag{10}$$

Decoding Sentence 3

- Scores

$$\delta = \begin{array}{c} \\ A \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{the}_0 & \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} -2.00 & -1.67 & -3.19 & -4.77 & -4.73 \\ -2.48 & -3.05 & -2.59 & -3.94 & -5.33 \\ -0.74 & -3.35 & -4.27 & -3.33 & -5.93 \\ -2.00 & -2.09 & -2.36 & -4.59 & -4.20 \end{pmatrix} \end{array} \tag{9}$$

- Backpointers

$$\tau = \begin{array}{c} \\ \text{A} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cccc} \text{old}_1 & \text{man}_2 & \text{the}_3 & \text{boat}_4 \end{array} \\ \begin{pmatrix} D & A & V & D \\ D & N & N & D \\ D & A & V & D \\ D & A & V & D \end{pmatrix} \end{array} \tag{10}$$

- Reconstruction: D N V D N