Structured Perceptron

Advanced Machine Learning for NLP

Jordan Boyd-Graber

HANDS-ON DEMO

Problem setup

- Restricted set of POS tags: adjective, preposition, verb, determiner, noun
- We first have sentence “time flies like an arrow” with true POS sequence N V P D N
- Features are $(z_i, z_{i+1}), (z_i, w_i)$
- What’s a maximum violation POS sequence?

Problem setup

- Restricted set of POS tags: adjective, preposition, verb, determiner, noun
- We first have sentence “time flies like an arrow” with true POS sequence N V P D N
- Features are $(z_i, z_{i+1}), (z_i, w_i)$
- What’s a maximum violation POS sequence?
- Can do on paper because search is tractable

Problem setup

- Restricted set of POS tags: adjective, preposition, verb, determiner, noun
- We first have sentence “time flies like an arrow” with true POS sequence N V P D N
- Features are $(z_i, z_{i+1}), (z_i, w_i)$
- What’s a maximum violation POS sequence?
- Can do on paper because search is tractable
- So we’re all on the same page, let’s all use A A A A A

- Correct answer: N V P D N
- Prediction: A A A A A

- Correct answer: N V P D N
- Prediction: A A A A A

- Correct answer: N V P D N
- Prediction: A A A A A

Gold Features

(P, D) (N, arrow)
(D, an) (N, time)
(V, P) (P, like)
(V, flies) (START, N)
(D, N) (N, V)

Shared Features

Predicted Features

(A, arrow) (A, A)
(A, an) (A, like)
(START, A) (A, flies)
(A, time)

- Correct answer: N V P D N
- Prediction: A A A A A

Gold Features

(P, D) (N, arrow)
(D, an) (N, time)
(V, P) (P, like)
(V, flies) (START, N)
(D, N) (N, V)

Shared Features

Predicted Features

(A, arrow) (A, A)
(A, an) (A, like)
(START, A) (A, flies)
(A, time)

- New feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00

- Correct answer: N V P D N
- Prediction: A A A A A

Gold Features

(P, D) (N, arrow)
(D, an) (N, time)
(V, P) (P, like)
(V, flies) (START, N)
(D, N) (N, V)

Shared Features

Predicted Features

(A, arrow) (A, A)
(A, an) (A, like)
(START, A) (A, flies)
(A, time)

- New feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$w_{\text{START, A}} + w_{\text{A, fruit}} = -1.00 + 0.00 = -1.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$w_{\text{START, P}} + w_{\text{P, fruit}} = 0.00 + 0.00 = 0.00$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & & & & \\ 0.00 & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$w_{\text{START, V}} + w_{\text{V, fruit}} = 0.00 + 0.00 = 0.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ & & & & \\ & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$w_{\text{START, D}} + w_{\text{D, fruit}} = 0.00 + 0.00 = 0.00$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$w_{\text{START, N}} + w_{\text{N, fruit}} = 1.00 + 0.00 = 1.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{c} \text{fruit}_0 \end{array} \\ \begin{pmatrix} -1.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 1.00 \end{pmatrix} \end{array} \quad \begin{array}{c} \text{flies}_1 \\ \\ \\ \\ \\ \\ \end{array} \quad \begin{array}{c} \text{like}_2 \\ \\ \\ \\ \\ \\ \end{array} \quad \begin{array}{c} \text{an}_3 \\ \\ \\ \\ \\ \\ \end{array} \quad \begin{array}{c} \text{apple}_4 \\ \\ \\ \\ \\ \\ \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_0(N) + w_{\text{N, A}} + w_{\text{A, flies}} = 1.00 + 0.00 + -1.00 = 0.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ 1.00 & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_0(N) + w_{\text{N, P}} + w_{\text{P, flies}} = 1.00 + 0.00 + 0.00 = 1.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & & & \\ 0.00 & 1.00 & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ 1.00 & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_0(N) + w_{\text{N, V}} + w_{\text{V, flies}} = 1.00 + 1.00 + 1.00 = 3.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & & & \\ 0.00 & 1.00 & & & \\ 0.00 & 3.00 & & & \\ 0.00 & & & & \\ 1.00 & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_0(P) + w_{\text{P, D}} + w_{\text{D, flies}} = 0.00 + 1.00 + 0.00 = 1.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & & & \\ 0.00 & 1.00 & & & \\ 0.00 & 3.00 & & & \\ 0.00 & 1.00 & & & \\ 1.00 & & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_0(N) + w_{\text{N, N}} + w_{\text{N, flies}} = 1.00 + 0.00 + 0.00 = 1.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & & & \\ 0.00 & 1.00 & & & \\ 0.00 & 3.00 & & & \\ 0.00 & 1.00 & & & \\ 1.00 & 1.00 & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_1(V) + w_{\text{V, A}} + w_{\text{A, like}} = 3.00 + 0.00 + -1.00 = 2.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & & & \\ 0.00 & 3.00 & & & \\ 0.00 & 1.00 & & & \\ 1.00 & 1.00 & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_1(V) + w_{V,\,P} + w_{P,\,\text{like}} = 3.00 + 1.00 + 1.00 = 5.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & & & \\ 0.00 & 1.00 & & & \\ 1.00 & 1.00 & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_1(V) + w_{V,\ V} + w_{V,\ \text{like}} = 3.00 + 0.00 + 0.00 = 3.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & & & \\ 1.00 & 1.00 & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_1(V) + w_{\text{V, D}} + w_{\text{D, like}} = 3.00 + 0.00 + 0.00 = 3.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_1(V) + w_{\text{V, N}} + w_{\text{N, like}} = 3.00 + 0.00 + 0.00 = 3.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & 3.00 & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_2(P) + w_{\text{P, A}} + w_{\text{A, an}} = 5.00 + 0.00 + -1.00 = 4.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & 3.00 & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_2(P) + w_{\text{P, P}} + w_{\text{P, an}} = 5.00 + 0.00 + 0.00 = 5.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & \\ 0.00 & 1.00 & 5.00 & 5.00 & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & 3.00 & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_2(P) + w_{\text{P, V}} + w_{\text{V, an}} = 5.00 + 0.00 + 0.00 = 5.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & \\ 0.00 & 1.00 & 5.00 & 5.00 & \\ 0.00 & 3.00 & 3.00 & 5.00 & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & 3.00 & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_2(P) + w_{\text{P, D}} + w_{\text{D, an}} = 5.00 + 1.00 + 1.00 = 7.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & \\ 0.00 & 1.00 & 5.00 & 5.00 & \\ 0.00 & 3.00 & 3.00 & 5.00 & \\ 0.00 & 1.00 & 3.00 & 7.00 & \\ 1.00 & 1.00 & 3.00 & & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_2(P) + w_{\text{P, N}} + w_{\text{N, an}} = 5.00 + 0.00 + 0.00 = 5.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{cccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 \quad \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 1.00 & 5.00 & 5.00 \\ 0.00 & 3.00 & 3.00 & 5.00 \\ 0.00 & 1.00 & 3.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_3(D) + w_{\text{D, A}} + w_{\text{A, apple}} = 7.00 + 0.00 + 0.00 = 7.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & \\ 0.00 & 3.00 & 3.00 & 5.00 & \\ 0.00 & 1.00 & 3.00 & 7.00 & \\ 1.00 & 1.00 & 3.00 & 5.00 & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_3(D) + w_{\text{D, P}} + w_{\text{P, apple}} = 7.00 + 0.00 + 0.00 = 7.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & \\ 0.00 & 1.00 & 3.00 & 7.00 & \\ 1.00 & 1.00 & 3.00 & 5.00 & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_3(D) + w_{\text{D, V}} + w_{\text{V, apple}} = 7.00 + 0.00 + 0.00 = 7.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & \\ 1.00 & 1.00 & 3.00 & 5.00 & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_3(D) + w_{\text{D, D}} + w_{\text{D, apple}} = 7.00 + 0.00 + 0.00 = 7.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 & \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

$$\delta_3(D) + w_{\text{D, N}} + w_{\text{N, apple}} = 7.00 + 1.00 + 0.00 = 8.00$$

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 & 8.00 \end{pmatrix} \end{array} \tag{1}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 & 8.00 \end{pmatrix} \end{array} \tag{1}$$

- Backpointers

$$\beta = \begin{array}{c} \\ \text{A} \\ \text{P} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cccc} \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} N & V & P & D \\ N & V & P & D \\ N & V & P & D \\ P & V & P & D \\ N & V & P & D \end{pmatrix} \end{array} \tag{2}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 & 8.00 \end{pmatrix} \end{array} \tag{1}$$

- Backpointers

$$\beta = \begin{array}{c} \\ \text{A} \\ \text{P} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cccc} \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} N & V & P & D \\ N & V & P & D \\ N & V & P & D \\ P & V & P & D \\ N & V & P & D \end{pmatrix} \end{array} \tag{2}$$

Decoding Sentence 2

- Scores

$$\delta = \begin{array}{c} \\ A \\ P \\ V \\ D \\ N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 & 8.00 \end{pmatrix} \end{array} \tag{1}$$

- Backpointers

$$\beta = \begin{array}{c} \\ \text{A} \\ \text{P} \\ \text{V} \\ \text{D} \\ \text{N} \end{array} \begin{array}{c} \begin{array}{cccc} \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{array} \\ \begin{pmatrix} N & V & P & D \\ N & V & P & D \\ N & V & P & D \\ P & V & P & D \\ N & V & P & D \end{pmatrix} \end{array} \tag{2}$$

- Reconstruction: N V P D N

- Old feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00
- Correct answer: A N V D N
- Prediction: N V P D N; (P, D): 0.00

- Old feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00
- Correct answer: A N V D N
- Prediction: N V P D N; (P, D): 0.00

- Old feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00
- Correct answer: A N V D N
- Prediction: N V P D N

Gold Features

(V, D) (A, N) (A, fruit) (V, like) (START, A) (N, flies)

; (P, D): 0.00

Shared Features

(D, an) (N, V) (N, apple) (D, N)

Predicted Features

(P, D) (V, P) (P, like) (V, flies) (START, N) (N, fruit)

- Old feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00
- Correct answer: A N V D N
- Prediction: N V P D N

Gold Features

(V, D) (A, N) (A, fruit) (V, like) (START, A) (N, flies)

Shared Features

(D, an) (N, V) (N, apple) (D, N)

Predicted Features

(P, D) (V, P) (P, like) (V, flies) (START, N) (N, fruit)

- New feature vector: (A, A): -4.00; (A, N): 1.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, fruit): 1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, flies): 1.00; (N, fruit): -1.00; (N, time): 1.00; (V, D): 1.00; (V, like): 1.00; (P, D): 0.00

- Old feature vector: (A, A): -4.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, time): 1.00; (P, D): 1.00; (P, like): 1.00; (V, P): 1.00; (V, flies): 1.00; (START, A): -1.00; (START, N): 1.00
- Correct answer: A N V D N
- Prediction: N V P D N

Gold Features

(V, D) (A, N) (A, fruit) (V, like) (START, A) (N, flies)

Shared Features

(D, an) (N, V) (N, apple) (D, N)

Predicted Features

(P, D) (V, P) (P, like) (V, flies) (START, N) (N, fruit)

- New feature vector: (A, A): -4.00; (A, N): 1.00; (A, an): -1.00; (A, arrow): -1.00; (A, flies): -1.00; (A, fruit): 1.00; (A, like): -1.00; (A, time): -1.00; (D, N): 1.00; (D, an): 1.00; (N, V): 1.00; (N, arrow): 1.00; (N, flies): 1.00; (N, fruit): -1.00; (N, time): 1.00; (V, D): 1.00; (V, like): 1.00; (P, D): 0.00

Wrapup

- Not just for POS tagging: parsing, machine translation
- Hard to overstate how important features $\vec{\Phi}$ are
- Next time: can we get algorithm to find features for us?

Wrapup

- Not just for POS tagging: parsing, machine translation
- Hard to overstate how important features $\vec{\Phi}$ are
- Next time: can we get algorithm to find features for us?
- Project ideas:
  - Deep learning of features
  - Applying perceptron to your favorite problem, designing great features
  - Efficient data structures for finding max violation