



Department of Computer Science  
UNIVERSITY OF COLORADO **BOULDER**



## 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

$$\delta = \begin{matrix} A \\ P \\ D \\ N \end{matrix} V \begin{pmatrix} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{pmatrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} \begin{pmatrix} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ -1.00 & & & & \end{pmatrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} A \\ P \\ D \\ N \end{matrix} V \begin{matrix} \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \end{matrix} \begin{pmatrix} -1.00 \\ \textcolor{red}{0.00} \\ \\ \\ \end{pmatrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ D \\ N \end{matrix} & \begin{pmatrix} -1.00 & \textcolor{red}{0.00} & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ 0.00 & & & & \\ 1.00 & & & & \end{pmatrix} \end{matrix} \quad (1)$$



## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \left( \begin{array}{ccccc} -1.00 & 0.00 & & & \\ 0.00 & 1.00 & & & \\ 0.00 & \textcolor{red}{3.00} & & & \\ 0.00 & & & & \\ 1.00 & & & & \end{array} \right) \end{matrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

$$\delta = V \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ D \\ N \end{matrix} & \begin{pmatrix} -1.00 & 0.00 \\ 0.00 & 1.00 \\ 0.00 & 3.00 \\ 0.00 & \textcolor{red}{1.00} \\ 1.00 & \end{pmatrix} \end{matrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

$$\delta = V \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ D \\ N \end{matrix} & \begin{pmatrix} -1.00 & 0.00 \\ 0.00 & 1.00 \\ 0.00 & 3.00 \\ 0.00 & 1.00 \\ 1.00 & \textcolor{red}{1.00} \end{pmatrix} \end{matrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \left( \begin{array}{ccccc} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & \textcolor{red}{3.00} & & \\ 0.00 & 1.00 & & & \\ 1.00 & 1.00 & & & \end{array} \right) \end{matrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \left( \begin{array}{ccccc} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & \textcolor{red}{3.00} & & \\ 1.00 & 1.00 & & & \end{array} \right) \end{matrix} \quad (1)$$



## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \left( \begin{array}{ccccc} -1.00 & 0.00 & 2.00 & & \\ 0.00 & 1.00 & 5.00 & & \\ 0.00 & 3.00 & 3.00 & & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & \textcolor{red}{3.00} & & \end{array} \right) \end{matrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \left( \begin{array}{ccccc} -1.00 & 0.00 & 2.00 & 4.00 & \\ 0.00 & 1.00 & 5.00 & 5.00 & \\ 0.00 & 3.00 & 3.00 & 5.00 & \\ 0.00 & 1.00 & 3.00 & & \\ 1.00 & 1.00 & 3.00 & & \end{array} \right) \end{matrix} \quad (1)$$

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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



## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

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

## Decoding Sentence 2

---

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

- Scores

$$\delta = \begin{matrix} & \text{fruit}_0 & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ D \\ N \end{matrix} & \begin{pmatrix} -1.00 & 0.00 & 2.00 & 4.00 & 7.00 \\ 0.00 & 1.00 & 5.00 & 5.00 & 7.00 \\ 0.00 & 3.00 & 3.00 & 5.00 & 7.00 \\ 0.00 & 1.00 & 3.00 & 7.00 & 7.00 \\ 1.00 & 1.00 & 3.00 & 5.00 & 8.00 \end{pmatrix} \end{matrix} \quad (1)$$

## Decoding Sentence 2

- Scores

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

- Backpointers

$$\beta = \begin{matrix} & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \begin{pmatrix} N & V & P & D \\ N & V & P & D \\ N & V & P & D \\ P & V & P & D \\ N & V & P & D \end{pmatrix} \end{matrix} \quad (2)$$

## Decoding Sentence 2

- Scores

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

- Backpointers

$$\beta = \begin{matrix} & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ V \\ D \\ N \end{matrix} & \begin{pmatrix} N & V & P & D \\ N & V & P & D \\ N & V & P & D \\ P & V & P & D \\ N & V & P & D \end{pmatrix} \end{matrix} \quad (2)$$

## Decoding Sentence 2

---

- Scores

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

- Backpointers

$$\beta = V \begin{matrix} & \text{flies}_1 & \text{like}_2 & \text{an}_3 & \text{apple}_4 \\ \begin{matrix} A \\ P \\ D \\ N \end{matrix} & \begin{pmatrix} N & V & P & D \\ N & V & P & D \\ N & V & P & D \\ P & V & P & D \\ N & V & P & D \end{pmatrix} \end{matrix} \quad (2)$$

- 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

- 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



- 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)

- 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

- 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

## 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