



Why Language is Hard: Structure and Predictions

Introduction to Data Science Algorithms
Jordan Boyd-Graber and Michael Paul
STRUCTURED PREDICTION EXAMPLE

$$w_{START, VB} + w_{VB, answer} = 0.00 + 0.00 = 0.00$$

answer
$$_0$$
 the $_1$ question $_2$
$$\delta = VB / DET / PRO / NN$$
 (1)

$$w_{START, DET} + w_{DET, answer} = 0.00 + 0.00 = 0.00$$

answer₀ the₁ question₂

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ 0.00 \\ NN \end{pmatrix}$$
(1)

$$w_{START, PRO} + w_{PRO, answer} = 0.00 + 0.00 = 0.00$$

answer₀ the₁ question₂

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \end{pmatrix}$$
(1)

$$w_{START, NN} + w_{NN, answer} = 0.00 + 0.00 = 0.00$$

$$\delta = \begin{array}{c} VB \\ VB \\ DET \\ PRO \\ NN \end{array} \begin{pmatrix} 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ \end{array} \right) \tag{1}$$

$$\delta_0(VB) + w_{VB, VB} + w_{VB, the} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \begin{array}{c} \text{vB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ \end{pmatrix} \tag{1}$$

$$\delta_0(VB) + w_{VB, DET} + w_{DET, the} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \end{pmatrix}$$
 (1)

$$\delta_0(VB) + w_{VB, PRO} + w_{PRO, the} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \end{pmatrix}$$
 (1)

$$\delta_0(VB) + w_{VB, NN} + w_{NN, the} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \end{pmatrix}$$
 (1)

$$\delta_1(VB) + w_{VB, VB} + w_{VB, question} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \begin{array}{c} \text{vB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ \end{array} \right) \tag{1}$$

$$\delta_1(VB) + w_{VB, DET} + w_{DET, question} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \begin{array}{c} \text{vB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ \end{array} \right) \tag{1}$$

$$\delta_1(VB) + w_{VB, PRO} + w_{PRO, question} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0 \end{pmatrix}$$
 (1)

$$\delta_1(VB) + w_{VB, NN} + w_{NN, question} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} \tag{1}$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} \tag{1}$$

Backpointers

$$\beta = \begin{array}{ccc} & \text{the}_1 & \text{question}_2 \\ \text{VB} & VB & VB \\ \text{PRO} & VB & VB \\ \text{NN} & VB & VB \\ VB & VB & VB \end{array} \right) \tag{2}$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} \tag{1}$$

Backpointers

$$\beta = \begin{array}{ccc} & \text{the}_1 & \text{question}_2 \\ \text{VB} & & VB \\ \text{DET} & VB & VB \\ \text{PRO} & VB & VB \\ \text{NN} & VB & VB \end{array} \right) \tag{2}$$

Decoding Sentence 1

Scores

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix}$$
 (1)

Backpointers

$$\beta = \begin{array}{ccc} & \text{the}_1 & \text{question}_2 \\ \text{VB} & VB & VB \\ \text{DET} & VB & VB \\ \text{PRO} & VB & VB \\ \text{NN} & VB & VB \end{array} \right) \tag{2}$$

Reconstruction: VB VB VB

Prediction: VB VB VB

Prediction: VB VB VB

Prediction: VB VB VB

Gold Features

(DET, the) (DET, NN)

(VB, DET)

(NN, question)

Shared Features

(START, VB)

(VB, answer)

Predicted Features

(VB, the)

(VB, question)

(VB, VB)

Prediction: VB VB VB

Gold Features (DET, the) (DET, NN) (VB, DET) (NN, question)

Shared Features (START, VB) (VB, answer)

Predicted Features (VB, the) (VB, question) (VB, VB)

```
    New feature vector: (DET, NN): 1.00; (DET, the): 1.00; (NN, question): 1.00; (VB, DET): 1.00; (VB, VB): -2.00; (VB, question): -1.00; (VB, the): -1.00
```

Prediction: VB VB VB

Gold Features (DET, the) (DET, NN) (VB, DET) (NN, question)

Shared Features (START, VB) (VB, answer) Predicted Features (VB, the) (VB, question) (VB, VB)

```
    New feature vector: (DET, NN): 1.00; (DET, the): 1.00; (NN, question): 1.00; (VB, DET): 1.00; (VB, VB): -2.00; (VB, question): -1.00; (VB, the): -1.00
```

$$\delta = \frac{VB}{PRO} \begin{pmatrix} VB & \\ DET & \\ NN & \end{pmatrix}$$
 (3)

$$w_{START, VB} + w_{VB, question} = 0.00 + -1.00 = -1.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 \\ NN \end{pmatrix}$$
 question₀ the₁ answer₂
$$\frac{VB}{PRO} \begin{pmatrix} -1.00 \\ NN \end{pmatrix}$$
 (3)

wstart, DET + wDET, question =
$$0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 \\ 0.00 \\ NN \end{pmatrix}$$
 (3)

$$w_{START, PRO} + w_{PRO, question} = 0.00 + 0.00 = 0.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB & -1.00 & \\ DET & 0.00 & \\ NN & 0.00 & \\ \end{array} \right) \label{eq:delta_estimate_problem}$$

$$w_{START, NN} + w_{NN, question} = 0.00 + 1.00 = 1.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ DET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 \\ 0.00 \\ 0.00 \\ 1.00 \\ \end{pmatrix} \tag{3}$$

$$\delta_0(NN) + w_{NN, VB} + w_{VB, the} = 1.00 + 0.00 + -1.00 = 0.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ ODET \\ PRO \\ NN \\ \end{array} \begin{array}{c} -1.00 & 0.00 \\ 0.00 \\ 0.00 \\ 1.00 \\ \end{array}$$

$$\delta_0(NN) + w_{NN, DET} + w_{DET, the} = 1.00 + 0.00 + 1.00 = 2.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ OET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 0.00 \\ 0.00 & 2.00 \\ 0.00 \\ 1.00 \\ \end{array} \right) \tag{3}$$

$$\delta_0(NN) + w_{NN, PRO} + w_{PRO, the} = 1.00 + 0.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ DET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 0.00 \\ 0.00 & 2.00 \\ 0.00 & 1.00 \\ 1.00 \\ \end{pmatrix} \tag{3}$$

$$\delta_0(DET) + w_{DET, NN} + w_{NN, the} = 0.00 + 1.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ OET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 0.00 \\ 0.00 & 2.00 \\ 0.00 & 1.00 \\ 1.00 & 1.00 \\ \end{array} \right) \tag{3}$$

$$\delta_1(DET) + w_{DET, VB} + w_{VB, answer} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ OET \\ PRO \\ NN \end{array} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 \\ 0.00 & 1.00 \\ 1.00 & 1.00 \end{pmatrix} \tag{3}$$

$$\delta_1(DET) + w_{DET, DET} + w_{DET, answer} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ DET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 & 2.00 \\ 0.00 & 1.00 \\ 1.00 & 1.00 \\ \end{pmatrix} \tag{3}$$

$$\delta_1(DET) + w_{DET, PRO} + w_{PRO, answer} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{question}_0 & \text{the}_1 & \text{answer}_2 \\ VB \\ DET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 & 2.00 \\ 0.00 & 1.00 & 2.00 \\ 1.00 & 1.00 \\ \end{array} \right) \tag{3}$$

$$\delta_1(DET) + w_{DET, NN} + w_{NN, answer} = 2.00 + 1.00 + 0.00 = 3.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 & 2.00 \\ 0.00 & 1.00 & 2.00 \\ 1.00 & 1.00 & 3.00 \end{pmatrix}$$
 (3)

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 & 2.00 \\ 0.00 & 1.00 & 2.00 \\ 1.00 & 1.00 & 3.00 \end{pmatrix}$$
 (3)

Backpointers

$$\beta = \begin{array}{c} \text{the}_1 & \text{answer}_2 \\ \text{VB} & NN & DET \\ \text{PRO} & NN & DET \\ \text{NN} & DET \\ \text{DET} & DET \end{array}$$

$$(4)$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 & 2.00 \\ 0.00 & 1.00 & 2.00 \\ 1.00 & 1.00 & 3.00 \end{pmatrix}$$
(3)

Backpointers

$$\beta = \begin{array}{c} \text{the}_1 & \text{answer}_2 \\ \text{VB} & NN & DET \\ \text{DET} & NN & DET \\ \text{NN} & DET \\ NN & DET \\ DET & DET \end{array}$$
 (4)

Scores

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 & 0.00 & 2.00 \\ 0.00 & 2.00 & 2.00 \\ 0.00 & 1.00 & 2.00 \\ 1.00 & 1.00 & 3.00 \end{pmatrix}$$
(3)

Backpointers

$$\beta = \begin{array}{c} \text{the}_1 & \text{answer}_2 \\ \text{VB} & NN & DET \\ \text{DET} & NN & DET \\ \text{PRO} & NN & DET \\ \text{NN} & DET \\ DET & DET \end{array}$$
 (4)

Reconstruction: NN DET NN

Prediction: NN DET NN

Prediction: NN DET NN

Prediction: NN DET NN

Gold Features

(VB, DET) (START, VB)

(VB, question)

Shared Features (DET, the) (DET, NN) (NN, answer)

Predicted Features

(START, NN)

(NN, question)

(NN, DET)

Prediction: NN DET NN

Gold Features (VB, DET) (START, VB) (VB, question)

Shared Features (DET, the) (DET, NN) (NN, answer) Predicted Features (START, NN) (NN, question) (NN, DET)

New feature vector: (DET, NN): 1.00; (DET, the): 1.00; (NN, DET): -1.00;
 (VB, DET): 2.00; (VB, VB): -2.00; (VB, the): -1.00; (START, NN): -1.00;
 (START, VB): 1.00

Prediction: NN DET NN

Gold Features (VB, DET) (START, VB) (VB, question)

Shared Features (DET, the) (DET, NN) (NN, answer) Predicted Features (START, NN) (NN, question) (NN, DET)

New feature vector: (DET, NN): 1.00; (DET, the): 1.00; (NN, DET): -1.00; (VB, DET): 2.00; (VB, VB): -2.00; (VB, the): -1.00; (START, NN): -1.00; (START, VB): 1.00

$$\delta = \frac{VB}{PRO} \begin{pmatrix} VB & \\ DET & \\ PRO & \\ NN \end{pmatrix}$$
 (5)

$$w_{START, VB} + w_{VB, you} = 1.00 + 0.00 = 1.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 1.00 \\ NN \end{pmatrix}$$
 you₀ demand₁ the₂ delay₃ (5)

$$w_{\text{START, DET}} + w_{\text{DET, you}} = 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} you_0 & demand_1 & the_2 & delay_3 \\ 1.00 & & & \\ 0.00 & & & \\ NN & & & \\ \end{pmatrix}$$
 (5)

$$w_{\text{START, PRO}} + w_{\text{PRO, you}} = 0.00 + 0.00 = 0.00$$

$$S = \frac{VB}{PRO} \begin{pmatrix} 1.00 \\ 0.00 \\ NN \end{pmatrix}$$
 (5)

$$w_{\text{START, NN}} + w_{\text{NN, you}} = -1.00 + 0.00 = -1.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ VB & 1.00 & \\ 0.00 & 0.00 & \\ NN & -1.00 & \\ \end{array} \right) \tag{5}$$

$$\delta_0(DET) + w_{DET, VB} + w_{VB, demand} = 0.00 + 0.00 + 0.00 = 0.00$$

$$\delta = \begin{array}{c} you_0 & demand_1 & the_2 & delay_3 \\ VB & 1.00 & 0.00 \\ OET & 0.00 & \\ PRO & 0.00 \\ NN & -1.00 & \\ \end{array} \right) \tag{5}$$

$$\delta_0(VB) + w_{VB, DET} + w_{DET, demand} = 1.00 + 2.00 + 0.00 = 3.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ VB & 1.00 & 0.00 \\ O.00 & 3.00 \\ O.00 & 0.00 \\ O.00 & -1.00 \end{array} \right) \tag{5}$$

$$\delta_0(VB) + w_{VB, PRO} + w_{PRO, demand} = 1.00 + 0.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ VB & 1.00 & 0.00 \\ DET & 0.00 & 3.00 \\ PRO & 0.00 & 1.00 \\ NN & -1.00 \end{array} \right) \tag{5}$$

$$\delta_0(VB) + w_{VB, NN} + w_{NN, demand} = 1.00 + 0.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ VB & 1.00 & 0.00 \\ O.00 & 3.00 \\ O.00 & 1.00 \\ NN & -1.00 & 1.00 \end{array} \right) \tag{5}$$

$$\delta_1(DET) + w_{DET, VB} + w_{VB, the} = 3.00 + 0.00 + -1.00 = 2.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 \\ 0.00 & 3.00 \\ 0.00 & 1.00 \\ -1.00 & 1.00 \\ \end{pmatrix} \tag{5}$$

$$\delta_1(DET) + w_{DET, DET} + w_{DET, the} = 3.00 + 0.00 + 1.00 = 4.00$$

$$\delta = \begin{array}{c} VB \\ VB \\ PRO \\ NN \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 \\ 0.00 & 3.00 & 4.00 \\ 0.00 & 1.00 \\ -1.00 & 1.00 \\ \end{pmatrix} \tag{5}$$

$$\delta_1(DET) + w_{DET, PRO} + w_{PRO, the} = 3.00 + 0.00 + 0.00 = 3.00$$

$$\delta = \begin{array}{c} VB \\ VB \\ PRO \\ NN \end{array} \left(\begin{array}{cccc} 1.00 & 0.00 & 2.00 \\ 0.00 & 3.00 & 4.00 \\ 0.00 & 1.00 & 3.00 \\ -1.00 & 1.00 \end{array} \right)$$
 (5)

$$\delta_1(DET) + w_{DET, NN} + w_{NN, the} = 3.00 + 1.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 \\ 0.00 & 3.00 & 4.00 \\ 0.00 & 1.00 & 3.00 \\ -1.00 & 1.00 & 4.00 \\ \end{pmatrix} \tag{5}$$

$$\delta_2(DET) + w_{DET, VB} + w_{VB, delay} = 4.00 + 0.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 \\ 0.00 & 1.00 & 3.00 \\ -1.00 & 1.00 & 4.00 \\ \end{pmatrix} \tag{5}$$

$$\delta_2(VB) + w_{VB, DET} + w_{DET, delay} = 2.00 + 2.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} VB \\ VB \\ PRO \\ NN \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 & 4.00 \\ 0.00 & 1.00 & 3.00 \\ -1.00 & 1.00 & 4.00 \\ \end{pmatrix} \tag{5}$$

$$\delta_2(DET) + w_{DET, PRO} + w_{PRO, delay} = 4.00 + 0.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 & 4.00 \\ 0.00 & 1.00 & 3.00 & 4.00 \\ -1.00 & 1.00 & 4.00 \\ \end{pmatrix} \tag{5}$$

$$\delta_2(DET) + w_{DET, NN} + w_{NN, delay} = 4.00 + 1.00 + 0.00 = 5.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 & 4.00 \\ 0.00 & 1.00 & 3.00 & 4.00 \\ -1.00 & 1.00 & 4.00 & 5.00 \\ \end{pmatrix} \tag{5}$$

Scores

$$\delta = \begin{array}{c} \text{you}_0 & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ VB & 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 & 4.00 \\ 0.00 & 1.00 & 3.00 & 4.00 \\ -1.00 & 1.00 & 4.00 & 5.00 \\ \end{array} \right) \tag{5}$$

Backpointers

$$\beta = \begin{array}{ccc} & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ \text{VB} & DET & DET & DET \\ \text{PRO} & VB & DET & VB \\ \text{NN} & VB & DET & DET \\ \end{array}$$
 (6)

Scores

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 & 4.00 \\ 0.00 & 1.00 & 3.00 & 4.00 \\ -1.00 & 1.00 & 4.00 & 5.00 \\ \end{pmatrix} \tag{5}$$

Backpointers

$$\beta = \begin{cases} VB & delay_3 \\ VB & DET & DET & DET \\ PRO & VB & DET & VB \\ NN & VB & DET & DET \\ VB & DET & DET \\ \end{pmatrix}$$
(6)

Scores

$$\delta = \begin{array}{c} \text{you}_0 & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ VB & 1.00 & 0.00 & 2.00 & 4.00 \\ 0.00 & 3.00 & 4.00 & 4.00 \\ 0.00 & 1.00 & 3.00 & 4.00 \\ -1.00 & 1.00 & 4.00 & 5.00 \\ \end{array} \right) \tag{5}$$

Backpointers

$$\beta = \begin{array}{ccc} & \text{demand}_1 & \text{the}_2 & \text{delay}_3 \\ \text{VB} & DET & DET & DET \\ \text{PRO} & VB & DET & VB \\ \text{NN} & VB & DET & DET \\ \end{array}$$
 (6)

Reconstruction: VB DET DET NN

Prediction: VB DET DET NN

Prediction: VB DET DET NN

Prediction: VB DET DET NN

Gold Features

(VB, demand) (PRO, you) (START, PRO)

(PRO, VB)

Shared Features

(DET, the) (DET, NN)

(VB, DET)

(NN, delay)

Predicted Features

(DET, DET)

(START, VB)

(DET, demand)

(VB, you)

Prediction: VB DET DET NN

Gold Features

(VB, demand) (PRO, you) (START, PRO) (PRO, VB)

Shared Features

(DET, the) (DET, NN) (VB, DET) (NN, delay)

Predicted Features

(DET, DET) (START, VB) (DET, demand) (VB, you)

New feature vector: (DET, DET): -1.00; (DET, NN): 1.00;
(DET, demand): -1.00; (DET, the): 1.00; (NN, DET): -1.00;
(PRO, VB): 1.00; (PRO, you): 1.00; (VB, DET): 2.00; (VB, VB): -2.00;
(VB, demand): 1.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00

Prediction: VB DET DET NN

Gold Features (VB, demand)

(PRO, you) (START, PRO)

(PRO, VB)

Shared Features

(DET, the) (DET, NN) (VB, DET) (NN, delay)

Predicted Features

(DET, DET) (START, VB) (DET, demand) (VB, you)

New feature vector: (DET, DET): -1.00; (DET, NN): 1.00; (DET, demand): -1.00; (DET, the): 1.00; (NN, DET): -1.00; (PRO, VB): 1.00; (PRO, you): 1.00; (VB, DET): 2.00; (VB, VB): -2.00; (VB, demand): 1.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00

$$\delta = \frac{VB}{PRO} \begin{pmatrix} VB & \\ DET & \\ PRO & \\ NN \end{pmatrix}$$
 (7)

$$w_{START, VB} + w_{VB, you} = 0.00 + -1.00 = -1.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} -1.00 \\ NN \end{pmatrix}$$
 the demand $\frac{1}{2}$ demand $\frac{1}{2}$ (7)

$$w_{\text{START, DET}} + w_{\text{DET, you}} = 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} v_{100} & v_{100} & v_{100} \\ -1.00 & v_{100} \\ 0.00 & v_{100} \\ 0.00 & v_{100} \end{pmatrix}$$
(7)

$$w_{\text{START, PRO}} + w_{\text{PRO, you}} = 1.00 + 1.00 = 2.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB & -1.00 & \\ DET & 0.00 & \\ NN & 2.00 & \\ \end{array} \right) \tag{7}$$

$$w_{\text{START, NN}} + w_{\text{NN, you}} = -1.00 + 0.00 = -1.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB & -1.00 & \\ \delta = \begin{array}{c} DET \\ PRO \\ NN & -1.00 \end{array} \right) \tag{7}$$

$$\delta_0(PRO) + w_{PRO, VB} + w_{VB, delay} = 2.00 + 1.00 + 0.00 = 3.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ \delta = \begin{array}{c} \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & \textbf{3.00} \\ 0.00 \\ 2.00 \\ -1.00 \\ \end{array} \right) \tag{7}$$

$$\delta_0(PRO) + w_{PRO, DET} + w_{DET, delay} = 2.00 + 0.00 + 0.00 = 2.00$$
• Scores

$$\delta = \begin{array}{cccc} & \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB & -1.00 & 3.00 \\ 0.00 & 2.00 \\ PRO & 2.00 \\ NN & -1.00 \end{array} \right)$$

$$\delta_0(PRO) + w_{PRO, PRO} + w_{PRO, delay} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ OET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 3.00 \\ 0.00 & 2.00 \\ 2.00 & 2.00 \\ -1.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_0(PRO) + w_{PRO, NN} + w_{NN, delay} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ OET \\ PRO \\ NN \\ \end{array} \begin{pmatrix} -1.00 & 3.00 \\ 0.00 & 2.00 \\ 2.00 & 2.00 \\ -1.00 & 2.00 \\ \end{array} \right) \tag{7}$$

$$\delta_1(PRO) + w_{PRO, VB} + w_{VB, the} = 2.00 + 1.00 + -1.00 = 2.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ \delta = \begin{array}{c} \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 \\ 0.00 & 2.00 \\ 2.00 & 2.00 \\ -1.00 & 2.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_1(VB) + w_{VB, DET} + w_{DET, the} = 3.00 + 2.00 + 1.00 = 6.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 \\ 0.00 & 2.00 & 6.00 \\ 2.00 & 2.00 \\ -1.00 & 2.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_1(VB) + w_{VB, PRO} + w_{PRO, the} = 3.00 + 0.00 + 0.00 = 3.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 \\ 0.00 & 2.00 & 6.00 \\ 2.00 & 2.00 & 3.00 \\ -1.00 & 2.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_1(VB) + w_{VB, NN} + w_{NN, the} = 3.00 + 0.00 + 0.00 = 3.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ \delta = \begin{array}{c} \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 \\ 0.00 & 2.00 & 6.00 \\ 2.00 & 2.00 & 3.00 \\ -1.00 & 2.00 & 3.00 \\ \end{array} \right) \tag{7}$$

$$\delta_2(DET) + w_{DET, VB} + w_{VB, demand} = 6.00 + 0.00 + 1.00 = 7.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ \delta = \begin{array}{c} \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 \\ 2.00 & 2.00 & 3.00 \\ -1.00 & 2.00 & 3.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_2(DET) + w_{DET, DET} + w_{DET, demand} = 6.00 + -1.00 + -1.00 = 4.00$$

$$\delta = \begin{array}{c} VB \\ DET \\ PRO \\ NN \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 & 4.00 \\ 2.00 & 2.00 & 3.00 \\ -1.00 & 2.00 & 3.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_2(DET) + w_{DET, PRO} + w_{PRO, demand} = 6.00 + 0.00 + 0.00 = 6.00$$

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ OET \\ PRO \\ NN \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 & 4.00 \\ 2.00 & 2.00 & 3.00 & 6.00 \\ -1.00 & 2.00 & 3.00 \\ \end{pmatrix} \tag{7}$$

$$\delta_2(DET) + w_{DET, NN} + w_{NN, demand} = 6.00 + 1.00 + 0.00 = 7.00$$

$$\delta = \begin{array}{c} \text{VB} \\ \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 & 4.00 \\ 2.00 & 2.00 & 3.00 & 6.00 \\ -1.00 & 2.00 & 3.00 & 7.00 \\ \end{pmatrix} \tag{7}$$

Decoding Sentence 4

Scores

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ DET \\ PRO \\ NN \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 & 4.00 \\ 2.00 & 2.00 & 3.00 & 6.00 \\ -1.00 & 2.00 & 3.00 & 7.00 \\ \end{pmatrix} \tag{7}$$

Backpointers

$$\beta = \begin{array}{cccc} & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ \text{VB} & PRO & PRO & DET \\ \text{PRO} & PRO & VB & DET \\ \text{PRO} & PRO & VB & DET \\ \text{NN} & PRO & VB & DET \\ \end{array} \right) \tag{8}$$

Decoding Sentence 4

Scores

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ \delta = \begin{array}{c} \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 & 4.00 \\ 2.00 & 2.00 & 3.00 & 6.00 \\ -1.00 & 2.00 & 3.00 & 7.00 \\ \end{pmatrix} \tag{7}$$

Backpointers

$$\beta = \begin{array}{cccc} & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ \text{VB} & \begin{array}{cccc} PRO & PRO & DET \\ PRO & PRO & VB & DET \\ PRO & VB & DET \\ PRO & VB & DET \\ \end{array} \right) \end{subarray}$$

Decoding Sentence 4

Scores

$$\delta = \begin{array}{c} \text{you}_0 & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ VB \\ \delta = \begin{array}{c} \text{VB} \\ \text{DET} \\ \text{PRO} \\ \text{NN} \end{array} \begin{pmatrix} -1.00 & 3.00 & 2.00 & 7.00 \\ 0.00 & 2.00 & 6.00 & 4.00 \\ 2.00 & 2.00 & 3.00 & 6.00 \\ -1.00 & 2.00 & 3.00 & 7.00 \\ \end{pmatrix} \tag{7}$$

Backpointers

$$\beta = \begin{array}{cccc} & \text{delay}_1 & \text{the}_2 & \text{demand}_3 \\ \text{VB} & PRO & PRO & DET \\ \text{PRO} & PRO & VB & DET \\ \text{PRO} & PRO & VB & DET \\ \text{NN} & PRO & VB & DET \\ \end{array} \right) \tag{8}$$

Reconstruction: PRO VB DET VB

Prediction: PRO VB DET VB

Prediction: PRO VB DET VB

Prediction: PRO VB DET VB

Gold Features (DET, NN) (NN, demand) Shared Features
(VB, delay)
(DET, the) (VB, DET)
(PRO, you)
(START, PRO)
(PRO, VB)

Predicted Features (DET, VB) (VB, demand)

Prediction: PRO VB DET VB

Gold Features (DET, NN)

(NN, demand)

Shared Features

(VB, delay) (DET, the) (VB, DET) (PRO, you) (START, PRO) (PRO, VB)

Predicted Features

(DET, VB) (VB, demand)

New feature vector: (DET, DET): -1.00; (DET, NN): 2.00;
(DET, VB): -1.00; (DET, demand): -1.00; (DET, the): 1.00; (NN, DET):
-1.00; (NN, demand): 1.00; (PRO, VB): 1.00; (PRO, you): 1.00; (VB, DET): 2.00; (VB, VB): -2.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00

- Correct answer: PRO VB DET NN
- Prediction: PRO VB DET VB

Gold Features (DET, NN) (NN, demand)

Shared Features (VB, delay) (DET, the) (VB, DET) (PRO, you) (START, PRO) (PRO, VB)

Predicted Features (DET, VB) (VB, demand)

```
New feature vector: (DET, DET): -1.00; (DET, NN): 2.00;
(DET, VB): -1.00; (DET, demand): -1.00; (DET, the): 1.00; (NN, DET):
-1.00; (NN, demand): 1.00; (PRO, VB): 1.00; (PRO, you): 1.00; (VB, DET): 2.00; (VB, VB): -2.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00
```

$$w_{START, VB} + w_{VB, what} = 0.00 + 0.00 = 0.00$$

$$S = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ NN \end{pmatrix}$$
 what₀ silence₁ can₂ show₃
$$0.00 \\ 0.00$$

$$w_{START, DET} + w_{DET, what} = 0.00 + 0.00 = 0.00$$

$$\mathcal{S} = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ 0.00 \\ NN \end{pmatrix}$$
 what₀ silence₁ can₂ show₃
$$0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00$$
 (9)

$$w_{START, PRO} + w_{PRO, what} = 1.00 + 0.00 = 1.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ 0.00 \\ 1.00 \\ NN \end{pmatrix}$$
 (9

$$w_{START, NN} + w_{NN, what} = -1.00 + 0.00 = -1.00$$

$$\delta = \begin{array}{c} \text{what}_0 \quad \text{silence}_1 \quad \text{can}_2 \quad \text{show}_3 \\ 0.00 \\ 0.00 \\ \text{PRO} \\ \text{NN} \end{array}$$

$$\delta_0(PRO) + w_{PRO, VB} + w_{VB, silence} = 1.00 + 1.00 + 0.00 = 2.00$$
• Scores

$$\delta = \begin{array}{c} \text{what}_0 \quad \text{silence}_1 \quad \text{can}_2 \quad \text{show}_3 \\ VB \\ 0.00 \quad 2.00 \\ 0.00 \\ 0.00 \\ 1.00 \\ NN \\ -1.00 \end{array} \right) \tag{9}$$

$$\delta_0(VB) + w_{VB, DET} + w_{DET, silence} = 0.00 + 2.00 + 0.00 = 2.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 2.00 \\ 0.00 & 2.00 \\ 1.00 \\ NN \end{pmatrix}$$
(9)

$$\delta_0(PRO) + w_{PRO, PRO} + w_{PRO, silence} = 1.00 + 0.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ VB & 0.00 & 2.00 & \\ DET & 0.00 & 2.00 & \\ PRO & 1.00 & 1.00 & \\ NN & -1.00 & \\ \end{array} \right) \tag{9}$$

$$\delta_0(DET) + w_{DET, NN} + w_{NN, silence} = 0.00 + 2.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ VB & 0.00 & 2.00 & \\ 0.00 & 2.00 & \\ 1.00 & 1.00 & \\ NN & -1.00 & 2.00 & \\ \end{array} \right) \tag{9}$$

$$\delta_1(PRO) + w_{PRO, VB} + w_{VB, can} = 1.00 + 1.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ VB & 0.00 & 2.00 & 2.00 \\ O.00 & 2.00 & 0.00 \\ PRO & 1.00 & 1.00 \\ NN & -1.00 & 2.00 \end{array} \right) \tag{9}$$

$$\delta_1(VB) + w_{VB, DET} + w_{DET, can} = 2.00 + 2.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ VB & 0.00 & 2.00 & 2.00 \\ O.00 & 2.00 & 4.00 \\ PRO & 1.00 & 1.00 \\ NN & -1.00 & 2.00 \\ \end{array} \right) \tag{9}$$

$$\delta_1(VB) + w_{VB, PRO} + w_{PRO, can} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{c|cccc} & \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ \hline VB & 0.00 & 2.00 & 2.00 \\ \hline DET & 0.00 & 2.00 & 4.00 \\ PRO & 1.00 & 1.00 & 2.00 \\ \hline NN & -1.00 & 2.00 \\ \hline \end{array} \right) \qquad (9)$$

$$\delta_1(DET) + w_{DET, NN} + w_{NN, can} = 2.00 + 2.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} \text{what}_0 \quad \text{silence}_1 \quad \text{can}_2 \quad \text{show}_3 \\ VB \\ O.00 \quad 2.00 \quad 2.00 \\ O.00 \quad 2.00 \quad 4.00 \\ 1.00 \quad 1.00 \quad 2.00 \\ -1.00 \quad 2.00 \quad 4.00 \\ \end{array} \right) \tag{9}$$

$$\delta_2(NN) + w_{NN, VB} + w_{VB, show} = 4.00 + 0.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c|cccc} & \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ \hline VB & 0.00 & 2.00 & 2.00 & 4.00 \\ \hline DET & 0.00 & 2.00 & 4.00 \\ \hline PRO & 1.00 & 1.00 & 2.00 \\ \hline NN & -1.00 & 2.00 & 4.00 \\ \hline \end{array} \right) \tag{9}$$

$$\delta_2(VB) + w_{VB, DET} + w_{DET, show} = 2.00 + 2.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c|cccc} & \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ \hline VB & 0.00 & 2.00 & 2.00 & 4.00 \\ \hline DET & 0.00 & 2.00 & 4.00 & 4.00 \\ \hline PRO & 1.00 & 1.00 & 2.00 \\ \hline NN & -1.00 & 2.00 & 4.00 \\ \hline \end{array} \right) \tag{9}$$

$$\delta_2(DET) + w_{DET, PRO} + w_{PRO, show} = 4.00 + 0.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c} \text{what}_0 & \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ VB \\ 0.00 & 2.00 & 2.00 & 4.00 \\ 0.00 & 2.00 & 4.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 4.00 \\ NN & -1.00 & 2.00 & 4.00 \\ \end{array} \right) \tag{9}$$

$$\delta_2(DET) + w_{DET, NN} + w_{NN, show} = 4.00 + 2.00 + 0.00 = 6.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 2.00 & 2.00 & 4.00 \\ 0.00 & 2.00 & 4.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 4.00 \\ -1.00 & 2.00 & 4.00 & 6.00 \end{pmatrix} \tag{9}$$

Scores

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 2.00 & 2.00 & 4.00 \\ 0.00 & 2.00 & 4.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 4.00 \\ -1.00 & 2.00 & 4.00 & 6.00 \end{pmatrix}$$
(9)

Backpointers

$$\beta = \begin{cases} \text{Silence}_1 & \text{can}_2 & \text{show}_3 \\ \text{VB} & PRO & PRO & NN \\ \text{VB} & VB & VB \\ PRO & VB & DET \\ NN & PRO & PRO & PRO & PRO \\ \end{pmatrix}$$
(10)

Scores

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 2.00 & 2.00 & 4.00 \\ 0.00 & 2.00 & 4.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 4.00 \\ -1.00 & 2.00 & 4.00 & 6.00 \end{pmatrix}$$
(9)

Backpointers

$$\beta = \begin{array}{ccc} \text{silence}_1 & \text{can}_2 & \text{show}_3 \\ \text{VB} & \begin{array}{ccc} PRO & PRO & NN \\ VB & VB & VB \\ PRO & VB & DET \\ NN & DET & DET & DET \end{array} \right) \tag{10}$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 2.00 & 2.00 & 4.00 \\ 0.00 & 2.00 & 4.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 4.00 \\ NN & -1.00 & 2.00 & 4.00 & 6.00 \end{pmatrix}$$
 (9)

Backpointers

silence₁ can₂ show₃

$$\beta = \begin{array}{cccc} VB & PRO & PRO & NN \\ DET & VB & VB & VB \\ PRO & VB & DET \\ NN & DET & DET & DET \end{array}$$
(10)

Reconstruction: PRO VB DET NN

Prediction: PRO VB DET NN

Prediction: PRO VB DET NN

Prediction: PRO VB DET NN

Gold Features

(VB, show) (VB, can) (PRO, NN) (NN, silence) (NN, VB) (VB, VB) Shared Features (START, PRO)

(PRO, what)

Predicted Features
(DET, can)
(NN, show)
(VB, silence)
(DET, NN) (VB, DET)
(PRO, VB)

- Correct answer: PRO NN VB VB
- Prediction: PRO VB DET NN

Gold Features

(VB, show) (VB, can) (PRO, NN) (NN, silence) (NN, VB) (VB, VB) Shared Features (START, PRO) (PRO, what)

Predicted Features

(DET, can) (NN, show) (VB, silence) (DET, NN) (VB, DET) (PRO, VB)

New feature vector: (DET, DET): -1.00; (DET, NN): 1.00; (DET, VB): -1.00; (DET, can): -1.00; (DET, demand): -1.00; (DET, the): 1.00; (NN, DET): -1.00; (NN, VB): 1.00; (NN, demand): 1.00; (NN, show): -1.00; (NN, silence): 1.00; (PRO, NN): 1.00; (PRO, you): 1.00; (VB, DET): 1.00; (VB, VB): -1.00; (VB, can): 1.00; (VB, show): 1.00; (VB, silence): -1.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00

- Correct answer: PRO NN VB VB
- Prediction: PRO VB DET NN

Gold Features

(VB, show) (VB, can) (PRO, NN) (NN, silence) (NN, VB) (VB, VB) Shared Features (START, PRO) (PRO, what)

Predicted Features

(DET, can) (NN, show) (VB, silence) (DET, NN) (VB, DET) (PRO, VB)

New feature vector: (DET, DET): -1.00; (DET, NN): 1.00; (DET, VB): -1.00; (DET, can): -1.00; (DET, demand): -1.00; (DET, the): 1.00; (NN, DET): -1.00; (NN, VB): 1.00; (NN, demand): 1.00; (NN, show): -1.00; (NN, silence): 1.00; (PRO, NN): 1.00; (PRO, you): 1.00; (VB, DET): 1.00; (VB, VB): -1.00; (VB, can): 1.00; (VB, show): 1.00; (VB, silence): -1.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00

$$w_{START, VB} + w_{VB, what} = 0.00 + 0.00 = 0.00$$

$$w_{START, DET} + w_{DET, what} = 0.00 + 0.00 = 0.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ 0.00 \\ NN \end{pmatrix}$$
 what₀ show₁ can₂ silence₃
$$0.00 \\ 0.00 \\ 0.11$$

$$w_{START, PRO} + w_{PRO, what} = 1.00 + 0.00 = 1.00$$

$$\tilde{D} = \frac{VB}{PRO} \begin{pmatrix} 0.00 \\ 0.00 \\ 1.00 \\ NN \end{pmatrix}$$
 what₀ show₁ can₂ silence₃
$$\frac{VB}{0.00} \begin{pmatrix} 0.00 \\ 0.00 \\ 1.00 \\ 0.00 \end{pmatrix}$$
 (11)

$$w_{START, NN} + w_{NN, what} = -1.00 + 0.00 = -1.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & \\ O.00 & 0.00 & \\ PRO & 1.00 & \\ NN & -1.00 & \\ \end{array} \right) \tag{11}$$

$$\delta_0(PRO) + w_{PRO, VB} + w_{VB, show} = 1.00 + 0.00 + 1.00 = 2.00$$

$$\delta = \frac{VB}{PRO} \begin{pmatrix} 0.00 & 2.00 \\ 0.00 & 0.00 \\ 1.00 & 0.00 \\ -1.00 \end{pmatrix}$$
 (11)

$$\delta_0(VB) + w_{VB, DET} + w_{DET, show} = 0.00 + 1.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 \\ O.00 & 1.00 \\ PRO & 1.00 \\ NN & -1.00 \end{array} \right) \tag{11}$$

$$\delta_0(PRO) + w_{PRO, PRO} + w_{PRO, show} = 1.00 + 0.00 + 0.00 = 1.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 & \\ DET & 0.00 & 1.00 & \\ PRO & 1.00 & 1.00 & \\ NN & -1.00 & \\ \end{array} \right) \tag{11}$$

$$\delta_0(PRO) + w_{PRO, NN} + w_{NN, show} = 1.00 + 1.00 + -1.00 = 1.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 \\ O.00 & 1.00 \\ PRO & 1.00 & 1.00 \\ NN & -1.00 & 1.00 \\ \end{array} \right) \tag{11}$$

$$\delta_1(NN) + w_{NN, VB} + w_{VB, can} = 1.00 + 1.00 + 1.00 = 3.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB \\ \delta = \begin{array}{c} O.00 & 2.00 & 3.00 \\ O.00 & 1.00 & \\ PRO \\ NN & -1.00 & 1.00 \\ \end{array} \right) \tag{11}$$

$$\delta_1(VB) + w_{VB, DET} + w_{DET, can} = 2.00 + 1.00 + -1.00 = 2.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB \\ \delta = \begin{array}{c} O.00 & 2.00 & 3.00 \\ 0.00 & 1.00 & 2.00 \\ PRO \\ NN & -1.00 & 1.00 \\ \end{array} \right) \tag{11}$$

$$\delta_1(VB) + w_{VB, PRO} + w_{PRO, can} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{ccccc} & \text{what}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 & 3.00 \\ \hline DET & 0.00 & 1.00 & 2.00 \\ PRO & 1.00 & 1.00 & 2.00 \\ NN & -1.00 & 1.00 & \end{array} \right) \tag{11}$$

$$\delta_1(VB) + w_{VB, NN} + w_{NN, can} = 2.00 + 0.00 + 0.00 = 2.00$$

$$\delta = \begin{array}{ccccc} & \text{what}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 & 3.00 \\ O.00 & 1.00 & 2.00 \\ PRO & 1.00 & 1.00 & 2.00 \\ NN & -1.00 & 1.00 & 2.00 \\ \end{array} \right) \tag{11}$$

$$\delta_2(NN) + w_{NN, VB} + w_{VB, silence} = 2.00 + 1.00 + -1.00 = 2.00$$

$$\delta = \begin{array}{c|ccccc} & \text{what}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 & 3.00 & 2.00 \\ \hline DET & 0.00 & 1.00 & 2.00 \\ PRO & 1.00 & 1.00 & 2.00 \\ NN & -1.00 & 1.00 & 2.00 \\ \end{array} \right) \tag{11}$$

$$\delta_2(VB) + w_{VB, DET} + w_{DET, silence} = 3.00 + 1.00 + 0.00 = 4.00$$

$$\delta = \begin{array}{c|ccccccc} & \text{what}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB & 0.00 & 2.00 & 3.00 & 2.00 \\ \hline DET & 0.00 & 1.00 & 2.00 & 4.00 \\ PRO & 1.00 & 1.00 & 2.00 \\ \hline NN & -1.00 & 1.00 & 2.00 \\ \end{array} \right) \tag{11}$$

$$\delta_2(VB) + w_{VB, PRO} + w_{PRO, silence} = 3.00 + 0.00 + 0.00 = 3.00$$

$$\delta = \begin{array}{c} \text{What}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ VB \\ \delta = \begin{array}{c} O.00 & 2.00 & 3.00 & 2.00 \\ 0.00 & 1.00 & 2.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 3.00 \\ -1.00 & 1.00 & 2.00 & \end{array} \right) \tag{11}$$

$$\delta_2(VB) + w_{VB, NN} + w_{NN, silence} = 3.00 + 0.00 + 1.00 = 4.00$$

$$\delta = \begin{array}{c|ccccc} & \text{what}_0 & \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ \hline VB & 0.00 & 2.00 & 3.00 & 2.00 \\ \hline DET & 0.00 & 1.00 & 2.00 & 4.00 \\ \hline PRO & 1.00 & 1.00 & 2.00 & 3.00 \\ \hline NN & -1.00 & 1.00 & 2.00 & 4.00 \\ \hline \end{array} \right) \end{tabular}$$

Scores

$$\delta = \begin{array}{c} VB \\ DET \\ PRO \\ NN \end{array} \begin{pmatrix} 0.00 & 2.00 & 3.00 & 2.00 \\ 0.00 & 1.00 & 2.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 3.00 \\ -1.00 & 1.00 & 2.00 & 4.00 \\ \end{pmatrix} \tag{11}$$

Backpointers

$$\beta = \begin{cases} NB & Show_1 & Can_2 & Silence_3 \\ NB & PRO & NN & NN \\ NN & VB & VB & VB \\ PRO & VB & VB \\ PRO & VB & VB \\ PRO & VB & VB \\ \end{cases}$$
(12)

Scores

$$\delta = \begin{array}{c} VB \\ DET \\ PRO \\ NN \end{array} \begin{pmatrix} 0.00 & 2.00 & 3.00 & 2.00 \\ 0.00 & 1.00 & 2.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 3.00 \\ -1.00 & 1.00 & 2.00 & 4.00 \\ \end{pmatrix} \tag{11}$$

Backpointers

$$\beta = \begin{array}{cccc} \text{show}_1 & \text{can}_2 & \text{silence}_3 \\ \text{VB} & PRO & NN & NN \\ \text{VB} & VB & VB \\ \text{PRO} & VB & VB \\ \text{PRO} & VB & VB \\ \text{PRO} & VB & VB \\ \end{array} \right) \tag{12}$$

Scores

$$\delta = \begin{array}{c} VB \\ DET \\ PRO \\ NN \end{array} \begin{pmatrix} 0.00 & 2.00 & 3.00 & 2.00 \\ 0.00 & 1.00 & 2.00 & 4.00 \\ 1.00 & 1.00 & 2.00 & 3.00 \\ -1.00 & 1.00 & 2.00 & 4.00 \\ \end{pmatrix} \tag{11}$$

Backpointers

$$\beta = \frac{\text{VB}}{\text{PRO}} \begin{pmatrix} PRO & NN & NN \\ VB & VB & VB \\ PRO & PRO & VB & VB \\ PRO & VB & VB \end{pmatrix}$$

$$(12)$$

Reconstruction: PRO NN VB DET

Prediction: PRO NN VB DET

Prediction: PRO NN VB DET

Prediction: PRO NN VB DET

Gold Features

(VB, silence)

(VB, VB)

Shared Features

(NN, show) (VB, can) (PRO, NN) (NN, VB)

(START, PRO)

(PRO, what)

Predicted Features

(DET, silence)

(VR DET)

- Correct answer: PRO NN VB VB
- Prediction: PRO NN VB DET

Gold Features

(VB, silence) (VB, VB)

Shared Features

(NN, show) (VB, can) (PRO, NN) (NN, VB) (START, PRO) (PRO, what)

Predicted Features

(DET, silence) (VB, DET)

```
New feature vector: (DET, DET): -1.00; (DET, NN): 1.00; (DET, VB): -1.00; (DET, can): -1.00; (DET, demand): -1.00; (DET, silence): -1.00; (DET, the): 1.00; (NN, DET): -1.00; (NN, VB): 1.00; (NN, demand): 1.00; (NN, show): -1.00; (NN, silence): 1.00; (PRO, NN): 1.00; (PRO, you): 1.00; (VB, can): 1.00; (VB, show): 1.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00
```

- Correct answer: PRO NN VB VB
- Prediction: PRO NN VB DET

Gold Features

(VB, silence) (VB, VB)

Shared Features

(NN, show) (VB, can) (PRO, NN) (NN, VB) (START, PRO) (PRO, what)

Predicted Features

(DET, silence) (VB, DET)

```
New feature vector: (DET, DET): -1.00; (DET, NN): 1.00; (DET, VB): -1.00; (DET, can): -1.00; (DET, demand): -1.00; (DET, silence): -1.00; (DET, the): 1.00; (NN, DET): -1.00; (NN, VB): 1.00; (NN, demand): 1.00; (NN, show): -1.00; (NN, silence): 1.00; (PRO, NN): 1.00; (PRO, you): 1.00; (VB, can): 1.00; (VB, show): 1.00; (VB, the): -1.00; (VB, you): -1.00; (START, NN): -1.00; (START, PRO): 1.00
```