Dependency Parsing

Computational Linguistics: Jordan Boyd-Graber

University of Maryland

Examples

Processing Shift-Reduce Actions

I am the very model of a modern major general

Processing Shift-Reduce Actions

I am the very model of a modern major general

1. Shift
2. Left
3. Shift
4. Shift
5. Shift

Processing Shift-Reduce Actions

I am the very model of a modern major general

6. Left
7. Left
8. Shift
9. Shift
10. Shift
11. Shift
12. Shift

Processing Shift-Reduce Actions

I am the very model of a modern major general

13. Left
14. Left
15. Left
16. Right
17. Right
18. Right
19. Right
20. Shift

Stack

[root]

Buffer

[I, am, the, very, model, of, a, modern, major, general]

Edges

Next move: 1. Shift

Stack

[root , |]

Buffer

[am, the, very, model, of, a, modern, major, general]

Edges

Next move: 2. Left

Stack

[root]

Buffer

[am, the, very, model, of, a, modern, major, general]

Edges

, I ← am

Next move: 3. Shift

Stack

[root , am]

Buffer

[the, very, model, of, a, modern, major, general]

Edges

, I ← am

Next move: 4. Shift

Stack

[root , am , the]

Buffer

[very, model, of, a, modern, major, general]

Edges

, I ← am

Next move: 5. Shift

Stack

[root , am , the , very]

Buffer

[model, of, a, modern, major, general]

Edges

, I ← am

Next move: 6. Left

Stack

[root , am , the]

Buffer

[model, of, a, modern, major, general]

Edges

, I ← am

, very ← model

Next move: 7. Left

Stack

[root , am]

Buffer

[model, of, a, modern, major, general]

Edges

, I ← am

, very ← model

, the ← model

Next move: 8. Shift

Stack

[root , am , model]

Buffer

[of, a, modern, major, general]

Edges

, I ← am

, very ← model

, the ← model

Next move: 9. Shift

Stack

[root , am , model , of]

Buffer

[a, modern, major, general]

Edges

, I ← am

, very ← model

, the ← model

Next move: 10. Shift

Stack

[root , am , model , of , a]

Buffer

[modern, major, general]

Edges

, I ← am

, very ← model

, the ← model

Next move: 11. Shift

Stack

[root , am , model , of , a , modern]

Buffer

[major, general]

Edges

, I ← am
, very ← model
, the ← model

Next move: 12. Shift

Stack

[root , am , model , of , a , modern , major]

Buffer

[general]

Edges

, I ← am
, very ← model
, the ← model

Next move: 13. Left

Stack

[root , am , model , of , a , modern]

Buffer

[general]

Edges

, I ← am
, very ← model
, the ← model
, major ← general

Next move: 14. Left

Stack

[root , am , model , of , a]

Buffer

[general]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general

Next move: 15. Left

Stack

[root , am , model , of]

Buffer

[general]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general
, a ← general

Next move: 16. Right

Stack

[root , am , model]

Buffer

[of,]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general
, a ← general
, of → general

Next move: 17. Right

Stack

[root , am]

Buffer

[model,]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general
, a ← general
, of → general
, model → of

Next move: 18. Right

Stack

[root]

Buffer

[am]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general
, a ← general
, of → general
, model → of
, am → model

Next move: 19. Right

Stack

[ ]

Buffer

[root]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general
, a ← general
, of → general
, model → of
, am → model
, root → am

Next move: 20. Shift

Stack

[root]

Buffer

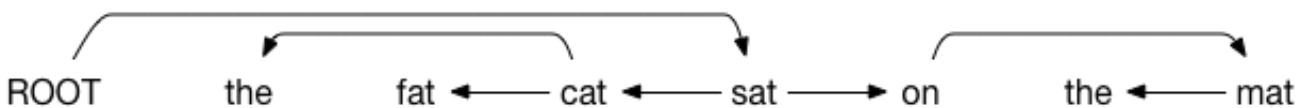
[]

Edges

, I ← am
, very ← model
, the ← model
, major ← general
, modern ← general
, a ← general
, of → general
, model → of
, am → model
, root → am

Transition Sequence Algorithm

- Start with root on stack, buffer with whole sentence
- If there's nothing on the stack, you must shift
- If the top of the stack is the child of the top of the buffer, then make a left edge
- If the top of the buffer is is a child of the top of the stack and the top of the buffer has no children that have yet to be added to the tree, then make a right



Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				
s				

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				
s				
l	7	mat	6	the

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				
s				
l	7	mat	6	the
r	5	on	7	mat

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				
s				
l	7	mat	6	the
r	5	on	7	mat
r	4	sat	5	on

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				
s				
l	7	mat	6	the
r	5	on	7	mat
r	4	sat	5	on
r	0	None	4	sat

Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
s				
s				
l	3	cat	2	fat
l	3	cat	1	the
s				
l	4	sat	3	cat
s				
s				
s				
l	7	mat	6	the
r	5	on	7	mat
r	4	sat	5	on
r	0	None	4	sat
s				

Eisner Parsing

Given the score function,

$$\lambda = \begin{array}{c} \\ \text{root} \\ \text{plastic} \\ \text{cup} \\ \text{holders} \end{array} \begin{array}{c} \begin{array}{cccc} \text{root} & \text{plastic} & \text{cup} & \text{holders} \end{array} \\ \begin{pmatrix} & 1.0 & 1.0 & 1.0 \\ -\infty & & -1.0 & -1.0 \\ -\infty & 2.0 & & -1.0 \\ -\infty & 0 & 4.0 & \end{pmatrix} \end{array}, \tag{1}$$

parse the sentence using the 4-dimensional table presented in the lecture. Where we use the rules:

1. $C[s][t][\rightarrow][\circ] = \max_{s \le q < t} C[s][q][\rightarrow][\cdot] + C[q+1][t][\leftarrow][\cdot] + \lambda_{(w_s, w_t)}$
2. $C[s][t][\leftarrow][\circ] = \max_{s \le q < t} C[s][q][\rightarrow][\cdot] + C[q+1][t][\leftarrow][\cdot] + \lambda_{(w_t, w_s)}$
3. $C[s][t][\rightarrow][\cdot] = \max_{s < q \le t} C[s][q][\rightarrow][\circ] + C[q][t][\rightarrow][\cdot]$
4. $C[s][t][\leftarrow][\cdot] = \max_{s \le q < t} C[s][q][\leftarrow][\cdot] + C[q][t][\leftarrow][\circ]$

to complete a 4-chart with entries for spans that start at s, end at t, and are either right $\rightarrow$ or left $\leftarrow$ directed and either incomplete $\circ$ or complete $\cdot$.

Span 1,2

1. $C[1,2,\leftarrow,\circ] =$
2. $C[1,2,\rightarrow,\circ] =$
3. $C[1,2,\leftarrow,\cdot] =$
4. $C[1,2,\rightarrow,\cdot] =$

Span 1,2

1. $C[1,2,\leftarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{2,1} = \lambda_{2,1} = -\infty$

2. $C[1,2,\rightarrow,\circ] =$

3. $C[1,2,\leftarrow,\cdot] =$

4. $C[1,2,\rightarrow,\cdot] =$

Span 1,2

1. $C[1,2,\leftarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{2,1} = \lambda_{2,1} = -\infty$

2. $C[1,2,\rightarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{1,2} = 1$

3. $C[1,2,\leftarrow,\cdot] =$

4. $C[1,2,\rightarrow,\cdot] =$

Span 1,2

1. $C[1,2,\leftarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{2,1} = \lambda_{2,1} = -\infty$

2. $C[1,2,\rightarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{1,2} = 1$

3. $C[1,2,\leftarrow,\cdot] = C[1,1,\leftarrow,\cdot] + C[1,2,\leftarrow,\circ] = C[1,2,\leftarrow,\circ] = \lambda_{2,1} = -\infty$

4. $C[1,2,\rightarrow,\cdot] =$

Span 1,2

1. $C[1,2,\leftarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{2,1} = \lambda_{2,1} = -\infty$

2. $C[1,2,\rightarrow,\circ] = C[1,1,\rightarrow,\cdot] + C[2,2,\leftarrow,\cdot] + \lambda_{1,2} = 1$

3. $C[1,2,\leftarrow,\cdot] = C[1,1,\leftarrow,\cdot] + C[1,2,\leftarrow,\circ] = C[1,2,\leftarrow,\circ] = \lambda_{2,1} = -\infty$

4. $C[1,2,\rightarrow,\cdot] = C[1,2,\rightarrow,\circ] + C[2,2,\rightarrow,\cdot] = C[1,2,\rightarrow,\circ] = \lambda_{1,2} = 1$

Span 2,3

1. $C[2,3,\leftarrow,\circ] =$
2. $C[2,3,\rightarrow,\circ] =$
3. $C[2,3,\leftarrow,\cdot] =$
4. $C[2,3,\rightarrow,\cdot] =$

Span 2,3

1. $C[2,3,\leftarrow,\circ] = \lambda_{3,2} = 2$
2. $C[2,3,\rightarrow,\circ] =$
3. $C[2,3,\leftarrow,\cdot] =$
4. $C[2,3,\rightarrow,\cdot] =$

Span 2,3

1. $C[2,3,\leftarrow,\circ] = \lambda_{3,2} = 2$

2. $C[2,3,\rightarrow,\circ] = \lambda_{2,3} = -1$

3. $C[2,3,\leftarrow,\cdot] =$

4. $C[2,3,\rightarrow,\cdot] =$

Span 2,3

1. $C[2,3,\leftarrow,\circ] = \lambda_{3,2} = 2$

2. $C[2,3,\rightarrow,\circ] = \lambda_{2,3} = -1$

3. $C[2,3,\leftarrow,\cdot] = C[2,3,\leftarrow,\circ] = 2$

4. $C[2,3,\rightarrow,\cdot] =$

Span 2,3

1. $C[2,3,\leftarrow,\circ] = \lambda_{3,2} = 2$

2. $C[2,3,\rightarrow,\circ] = \lambda_{2,3} = -1$

3. $C[2,3,\leftarrow,\cdot] = C[2,3,\leftarrow,\circ] = 2$

4. $C[2,3,\rightarrow,\cdot] = C[2,3,\rightarrow,\circ] = -1$

Span 3,4

1. $C[3,4,\leftarrow,\circ] = \lambda_{4,3} = 4$

2. $C[3,4,\rightarrow,\circ] = \lambda_{3,4} = -1$

3. $C[3,4,\leftarrow,\cdot] = C[3,4,\leftarrow,\circ] = 4$

4. $C[3,4,\rightarrow,\cdot] = C[3,4,\rightarrow,\circ] = -1$

Span 1,3

1.

$$\begin{aligned} C[1,3,\leftarrow,\circ] = \max(&C[1,1,\rightarrow,\cdot] + C[2,3,\leftarrow,\cdot], \\ &C[1,2,\rightarrow,\cdot] + C[3,3,\leftarrow,\cdot]) + \lambda_{3,1} \\ =&\lambda_{3,1} + \lambda_{3,2} = -\infty \end{aligned}$$

Span 1,3

1.

$$C[1,3,\leftarrow,\circ] = -\infty$$

Span 1,3

1.

$$C[1,3,\leftarrow,\circ] = -\infty$$

2.

$$\begin{aligned} C[1,3,\rightarrow,\circ] &= \max(C[2,3,\leftarrow,\cdot], C[1,2,\rightarrow,\cdot]) + \lambda_{1,3} \\ &= \lambda_{3,2} + \lambda_{1,3} = 3 \end{aligned}$$

Span 1,3

1.

$$C[1,3,\leftarrow,\circ] = -\infty$$

2.

$$C[1,3,\rightarrow,\circ] = 3$$

Span 1,3

1.

$$C[1,3,\leftarrow,\circ] = -\infty$$

2.

$$C[1,3,\rightarrow,\circ] = 3$$

3.

$$\begin{aligned} C[1,3,\leftarrow,\cdot] &= \max(C[1,1,\leftarrow,\cdot] + C[1,3,\leftarrow,\circ], C[1,2,\leftarrow,\cdot] + C[2,3,\rightarrow,\circ] \\ &= \max(0-\infty, -\infty+2) = -\infty \end{aligned}$$

Span 1,3

1.

$$C[1,3,\leftarrow,\circ] = -\infty$$

2.

$$C[1,3,\rightarrow,\circ] = 3$$

3.

$$C[1,3,\leftarrow,\cdot] - \infty$$

4.

$$\begin{aligned} C[1,3,\rightarrow,\cdot] = \max(&C[1,2,\rightarrow,\circ] + C[2,3,\rightarrow,\cdot], \\ &\underline{C[1,3,\rightarrow,\circ] + C[3,3,\rightarrow,\cdot]}) \\ &= \max(\lambda_{1,2} + \lambda_{2,3}, \underline{\lambda_{1,3} + \lambda_{3,2}}) = \max(0, \underline{3}) = 3 \end{aligned}$$

Span 2,4

1.

$$
\begin{aligned}
C[2,4,\leftarrow,\circ] &= \max(C[2,2,\rightarrow,\cdot] + C[3,4,\leftarrow,\cdot], \\
&\quad C[2,3,\rightarrow,\cdot] + C[4,4,\leftarrow,\cdot]) + \lambda_{4,2} \\
&= \max(\underline{C[3,4,\leftarrow,\cdot]}, C[2,3,\rightarrow,\cdot]) + \lambda_{4,2} \\
&= \max(\underline{\lambda_{4,3}}, \lambda_{2,3}) + \lambda_{4,2} = 4 + 0 = 4
\end{aligned}
$$

Span 2,4

1.

$$C[2,4,\leftarrow,\circ] = \max(\underline{\lambda_{4,3}}, \lambda_{2,3}) + \lambda_{4,2} = 4 + 0 = 4$$

2.

$$\begin{aligned} C[2,4,\rightarrow,\circ] &= \max(C[2,2,\rightarrow,\cdot] + C[3,4,\leftarrow,\cdot], \\ &\qquad C[2,3,\rightarrow,\cdot] + C[4,4,\leftarrow,\cdot]) + \lambda_{2,4} \\ &= \max(\underline{\lambda_{4,3}}, \lambda_{2,3}) + \lambda_{2,4} = 3 \end{aligned}$$

Span 2,4

1.

$$C[2,4,\leftarrow,\circ] = \max(\underline{\lambda_{4,3}}, \lambda_{2,3}) + \lambda_{4,2} = 4 + 0 = 4$$

2.

$$C[2,4,\rightarrow,\circ] = 3$$

3.

$$\begin{aligned} C[2,4,\leftarrow,\cdot] &= \max(C[2,4,\rightarrow,\circ], C[2,3,\leftarrow,\cdot] + C[3,4,\leftarrow,\circ]) \\ &= \max(\lambda_{4,2} + \lambda_{4,3}, \\ &\qquad \underline{\lambda_{3,2} + \lambda_{4,3}}) = 6 \end{aligned}$$

Span 2,4

1.

$$C[2,4,\leftarrow,\circ] = \max(\underline{\lambda_{4,3}}, \lambda_{2,3}) + \lambda_{4,2} = 4 + 0 = 4$$

2.

$$C[2,4,\rightarrow,\circ] = 3$$

3.

$$C[2,4,\leftarrow,\cdot] = 6$$

4.

$$C[2,4,\rightarrow,\cdot] = \max(C[2,3,\rightarrow,\circ] + C[3,4,\rightarrow,\cdot], C[2,4,\rightarrow,\circ]) =$$
$$\max(\lambda_{2,3} + \lambda_{3,4}, \underline{\lambda_{2,4} + \lambda_{4,3}}) = 3$$

Span 1,4

1.

$$C[1,4,\leftarrow,\circ] = \lambda_{4,1} + \cdots = -\infty$$

Span 1,4

1.

$$C[1,4,\leftarrow,\circ] = \lambda_{4,1} + \cdots = -\infty$$

2.

$$\begin{aligned} C[1,4,\rightarrow,\circ] = \max(&C[2,4,\leftarrow,\cdot], \\ &C[1,2,\rightarrow,\cdot] + C[3,4,\leftarrow,\cdot], \\ &C[1,3,\rightarrow,\cdot]) + \lambda_{1,4} \\ = \max(&\lambda_{3,2} + \lambda_{4,3}, \underline{\lambda_{1,2} + \lambda_{4,3}}, \lambda_{1,2} + \lambda_{4,3}, \lambda_{1,3} + \lambda_{3,2}) \\ = &6 + 1 = 7 \end{aligned}$$

Span 1,4

1.

$$C[1,4,\leftarrow,\circ] = \lambda_{4,1} + \cdots = -\infty$$

2.

$$C[1,4,\rightarrow,\circ] = 7$$

3. $C[1,4,\leftarrow,\cdot] = \max(0 + -\infty, -\infty + \ldots, -\infty + \ldots) = -\infty$

Span 1,4

1.

$$C[1,4,\leftarrow,\circ] = \lambda_{4,1} + \cdots = -\infty$$

2.

$$C[1,4,\rightarrow,\circ] = 7$$

3. $C[1,4,\leftarrow,\cdot] = \max(0 + -\infty, -\infty + \ldots, -\infty + \ldots) = -\infty$

4.

$$\begin{aligned} C[1,4,\rightarrow,\cdot] = \max(&C[1,2,\rightarrow,\circ] + C[2,4,\rightarrow,\cdot], \\ &C[1,3,\rightarrow,\circ] + C[3,4,\rightarrow,\cdot], \\ &C[1,4,\rightarrow,\circ]) \\ = \max(&\lambda_{1,2} + \lambda_{2,4} + \lambda_{4,3} + \lambda_{1,3} + \lambda_{3,2} + \lambda_{3,4}, \\ &\underline{\lambda_{1,4} + \lambda_{4,3} + \lambda_{3,2}}) \\ = \max(&1+3, 3-1, \underline{1+4+2}) = 7 \end{aligned}$$

Reconstruction