



# 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 , I ]

## 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

, l ← am

Next move: 3. Shift



## Stack

[root , am ]

## Buffer

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

## Edges

, I  $\leftarrow$  am

Next move: 4. Shift

## Stack

[root , am , the ]

## Buffer

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

## Edges

, l ← 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  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

Next move: 8. Shift

## Stack

[root , am , model ]

## Buffer

[of, a, modern, major, general]

## Edges

, I  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

Next move: 9. Shift

## Stack

[root , am , model , of ]

## Buffer

[a, modern, major, general]

## Edges

, I  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

Next move: 10. Shift

## Stack

[root , am , model , of , a ]

## Buffer

[modern, major, general]

## Edges

, I  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

Next move: 11. Shift



## Stack

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

## Buffer

[major, general]

## Edges

, I  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

Next move: 12. Shift

## Stack

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

## Buffer

[general]

## Edges

, I  $\leftarrow$  am  
, very  $\leftarrow$  model  
, the  $\leftarrow$  model

Next move: 13. Left

## Stack

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

## Buffer

[general]

## Edges

, I  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

, major  $\leftarrow$  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  $\leftarrow$  am

, very  $\leftarrow$  model

, the  $\leftarrow$  model

, major  $\leftarrow$  general

, modern  $\leftarrow$  general

, a  $\leftarrow$  general

Next move: 16. Right

## Stack

[root , am , model ]

## Buffer

[of, ]

## Edges

, I  $\leftarrow$  am  
, very  $\leftarrow$  model  
, the  $\leftarrow$  model  
, major  $\leftarrow$  general  
, modern  $\leftarrow$  general  
, a  $\leftarrow$  general  
, of  $\rightarrow$  general

Next move: 17. Right

## Stack

[root , am ]

## Buffer

[model, ]

## Edges

, I  $\leftarrow$  am  
, very  $\leftarrow$  model  
, the  $\leftarrow$  model  
, major  $\leftarrow$  general  
, modern  $\leftarrow$  general  
, a  $\leftarrow$  general  
, of  $\rightarrow$  general  
, **model**  $\rightarrow$  **of**

Next move: 18. Right

## Stack

[root     ]

## Buffer

[am]

## Edges

, I  $\leftarrow$  am  
, very  $\leftarrow$  model  
, the  $\leftarrow$  model  
, major  $\leftarrow$  general  
, modern  $\leftarrow$  general  
, a  $\leftarrow$  general  
, of  $\rightarrow$  general  
, model  $\rightarrow$  of  
, **am  $\rightarrow$  model**

Next move: 19. Right



## Stack

[     ]

## Buffer

[root]

## Edges

, I  $\leftarrow$  am  
, very  $\leftarrow$  model  
, the  $\leftarrow$  model  
, major  $\leftarrow$  general  
, modern  $\leftarrow$  general  
, a  $\leftarrow$  general  
, of  $\rightarrow$  general  
, model  $\rightarrow$  of  
, am  $\rightarrow$  model  
, **root**  $\rightarrow$  **am**

Next move: 20. Shift

## Stack

[root     ]

## Buffer

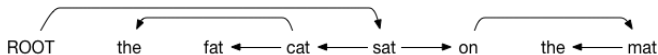
[]

## Edges

, I  $\leftarrow$  am  
, very  $\leftarrow$  model  
, the  $\leftarrow$  model  
, major  $\leftarrow$  general  
, modern  $\leftarrow$  general  
, a  $\leftarrow$  general  
, of  $\rightarrow$  general  
, model  $\rightarrow$  of  
, am  $\rightarrow$  model  
, root  $\rightarrow$  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				
	3	cat	2	fat
	3	cat	1	the

## Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
S				
S				
I	3	cat	2	fat
I	3	cat	1	the
S				



## Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
S				
S				
I	3	cat	2	fat
I	3	cat	1	the
S				
I	4	sat	3	cat

## Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
S				
S				
I	3	cat	2	fat
I	3	cat	1	the
S				
I	4	sat	3	cat
S				

## Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
S				
S				
I	3	cat	2	fat
I	3	cat	1	the
S				
I	4	sat	3	cat
S				
S				

## Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
S				
S				
I	3	cat	2	fat
I	3	cat	1	the
S				
I	4	sat	3	cat
S				
S				
S				

## Parse to Transition Sequence

Action	Head Index	Head Word	Dep Index	Dep Word
S				
S				
I	3	cat	2	fat
I	3	cat	1	the
S				
I	4	sat	3	cat
S				
S				
S				
I	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				

## Span 1,2

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

## 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.

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

## 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.

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

## 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.

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

2.

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

## Span 2,4

1.

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

2.

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

3.

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

## Span 2,4

1.

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

2.

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

3.

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

4.

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

## Span 1,4

1.

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

## Span 1,4

1.

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

2.

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



## Span 1,4

1.

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

2.

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

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

## Span 1,4

1.

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

2.

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

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

4.

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

## Reconstruction

