COMP 4446 / 5046
Lecture 5: Inference – Dynamic Programming

Jonathan K. Kummerfeld

Semester 1, 2025

THE UNIVERSITY OF SYDNEY

menti.com 6274 6616

Language Nerd

[Not to go all sentence fragment on you.]

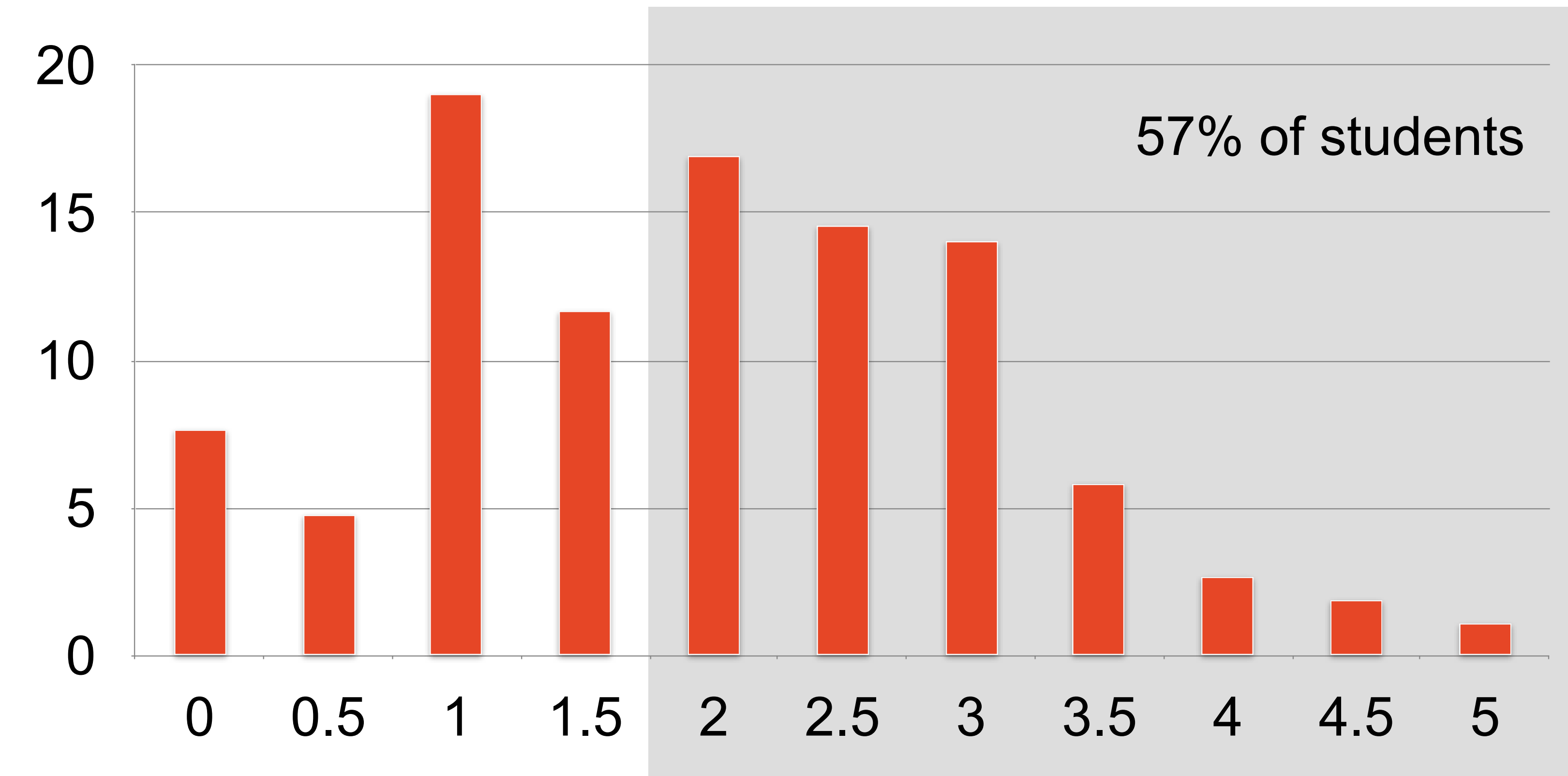
Source: https://xkcd.com/1443/

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

The quiz was representative of the exam and 43% of students scored under 2. What next?



Me: Working to give more support

You: Engage with all components of the unit

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Sequence Tagging

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We will consider various sequences and tags

way 1

B-PER	O	O
Joe	likes	trucks

2

Noun	Verb	Noun
Joe	likes	trucks

input

Sequences are a fixed, known, length
Tags are from a small-ish, fixed, set

should be small, otherwise, computational cost ↑

Independent predictions can be inconsistent

Possible independent predictions:

2 mistake

B-ORG	I-ORG	O	B-LOC	O	O	O	···

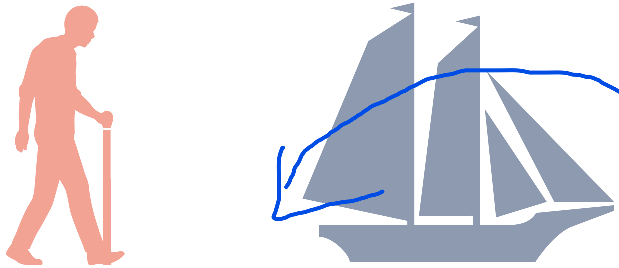
True labels:

B-ORG	I-ORG	I-ORG	I-ORG	O	O	O	···
The	University	of	Sydney	is	located	in	…

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

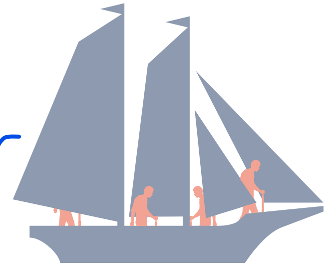
menti.com 6274 6616

Greedy predictions can make mistakes that become clear too late



Predictions:	Det	Adj.	Noun	Det	Noun

Error



True labels:	Det	Noun	Verb	Det	Noun
	The	old	man	the	boat

A greedy method can't go back and undo the choices Adj & Noun!

"The old man" originally good goes bad latter

menti.com 6274 6616

Considering each tag sequence separately is not feasible

17 tags in Universal Dependencies

$\text{Options} = |\text{tags}|^{|\text{words}|}$

For the last example,

17^5 = 1,419,857 options

For a 20 word sentence (average length in news):

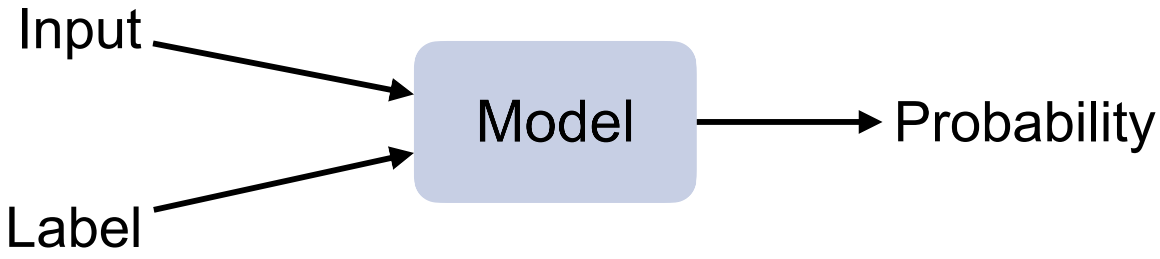
17^{20} = 4,064,231,406,647,572,522,401,601
(4.1 Septillion)

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

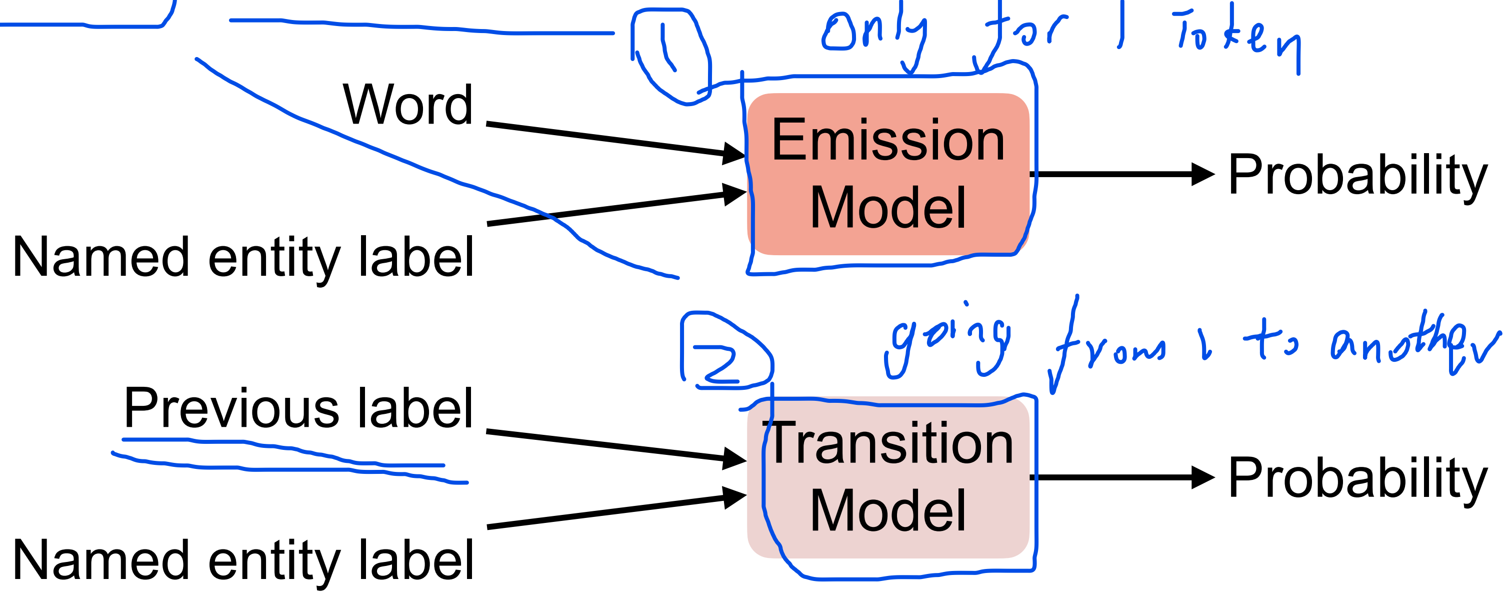
menti.com 6274 6616

For certain models we can find the highest scoring sequence efficiently



new

Our model needs to calculate the score in pieces of two types:



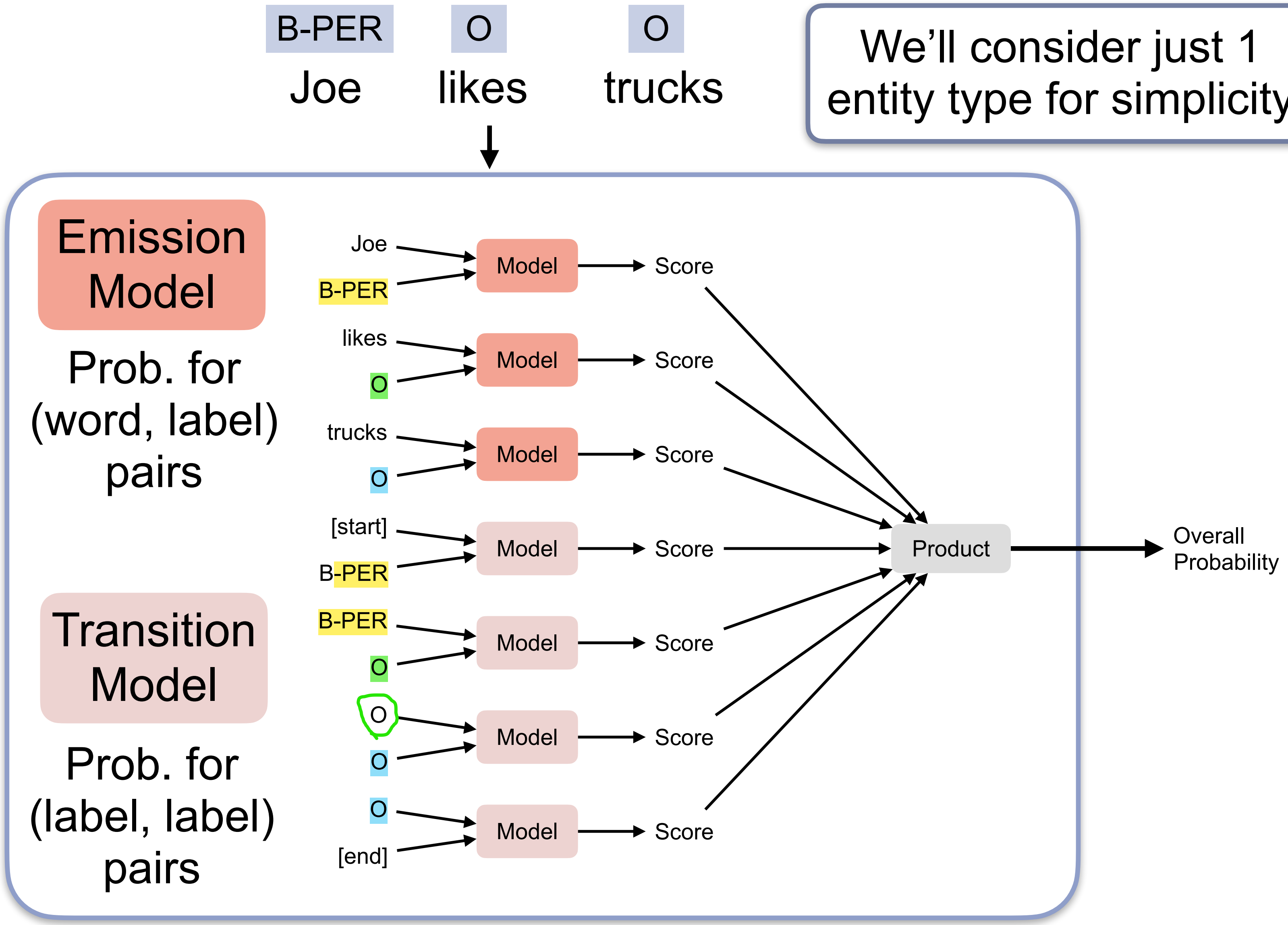
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We can still think of it as a model for whole sequences

We'll consider just 1 entity type for simplicity

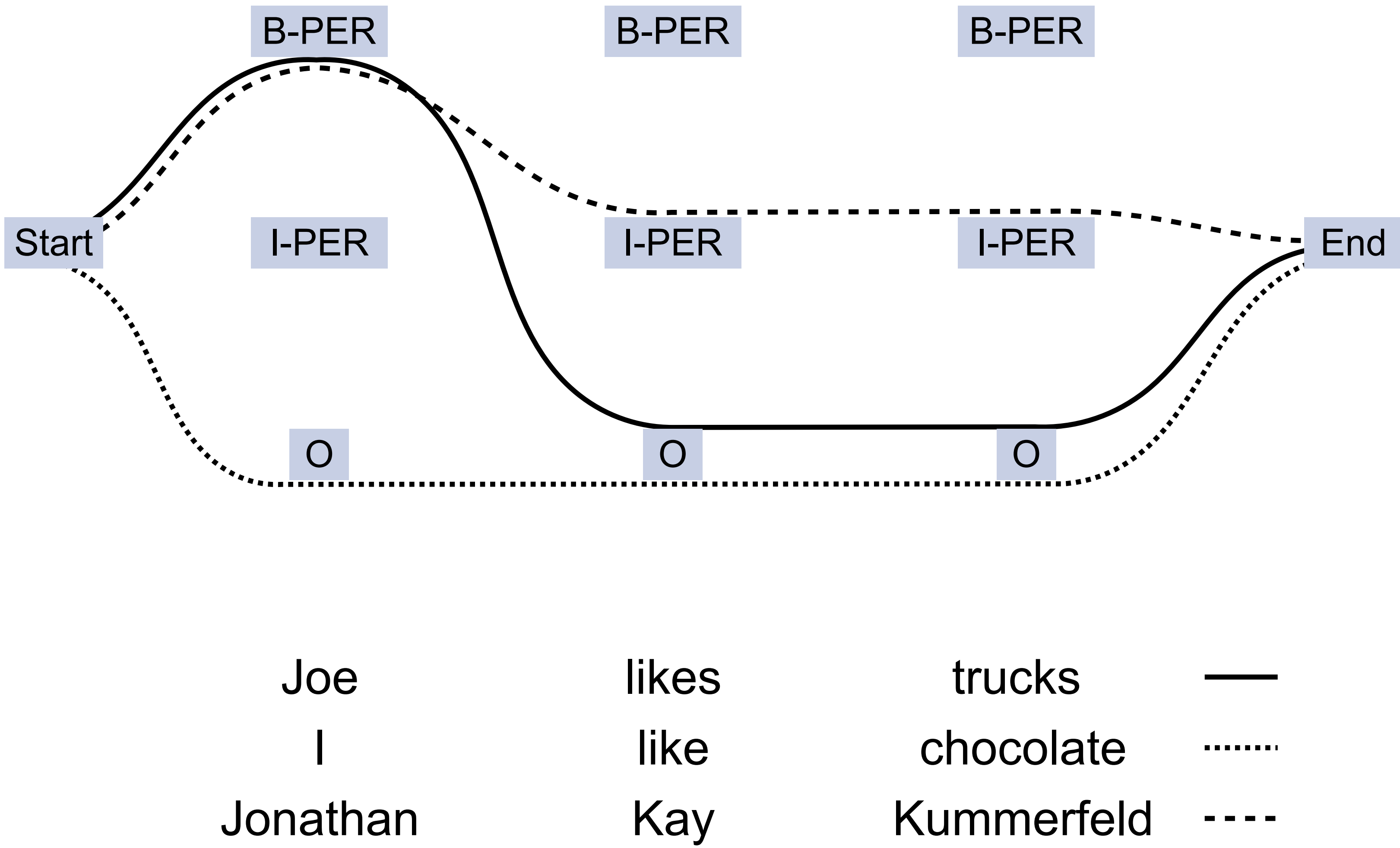


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

To understand the algorithm, we can draw paths through a grid of labels

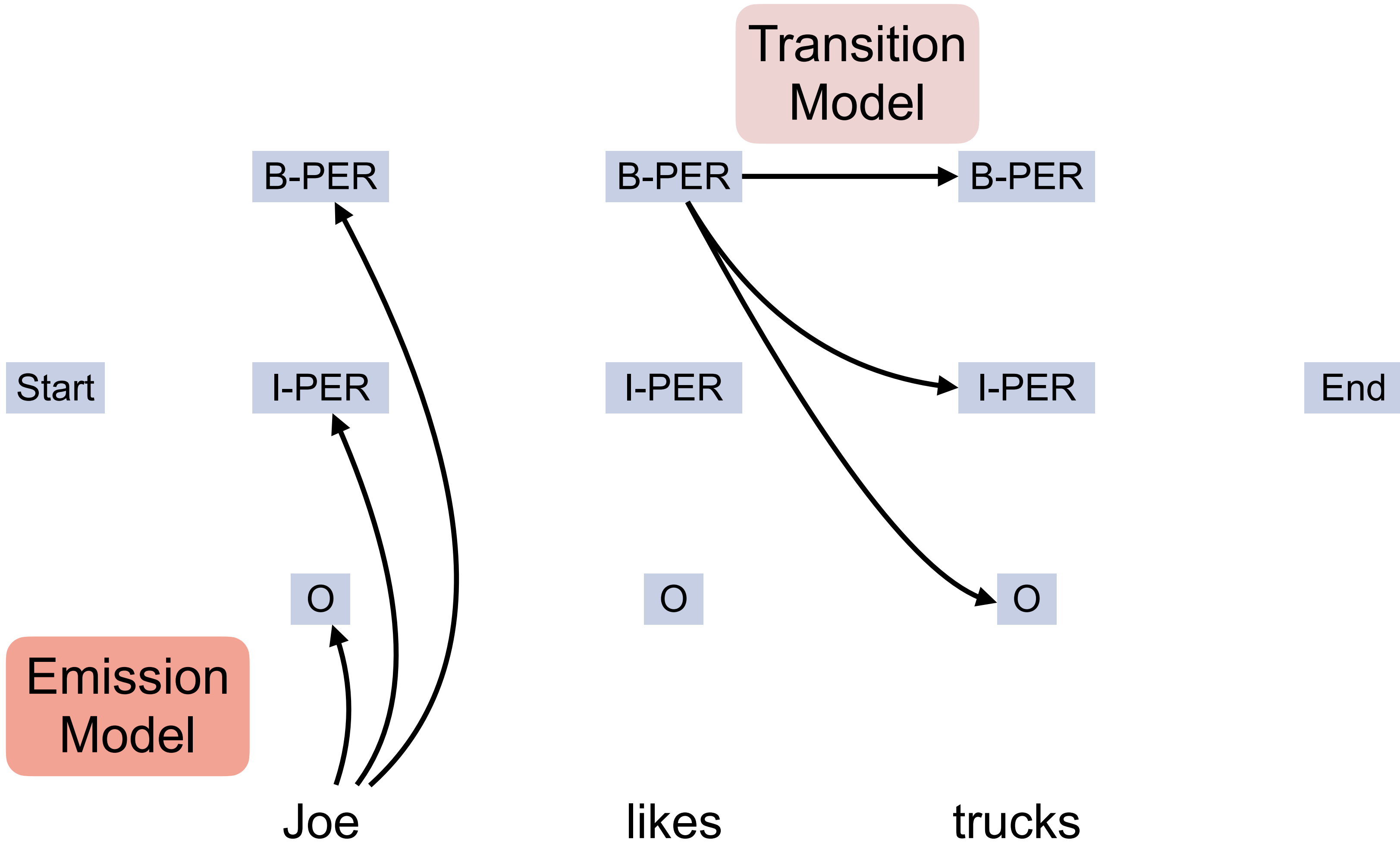


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

The two model parts from earlier correspond to parts of this grid



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

You can read it as output of transition model

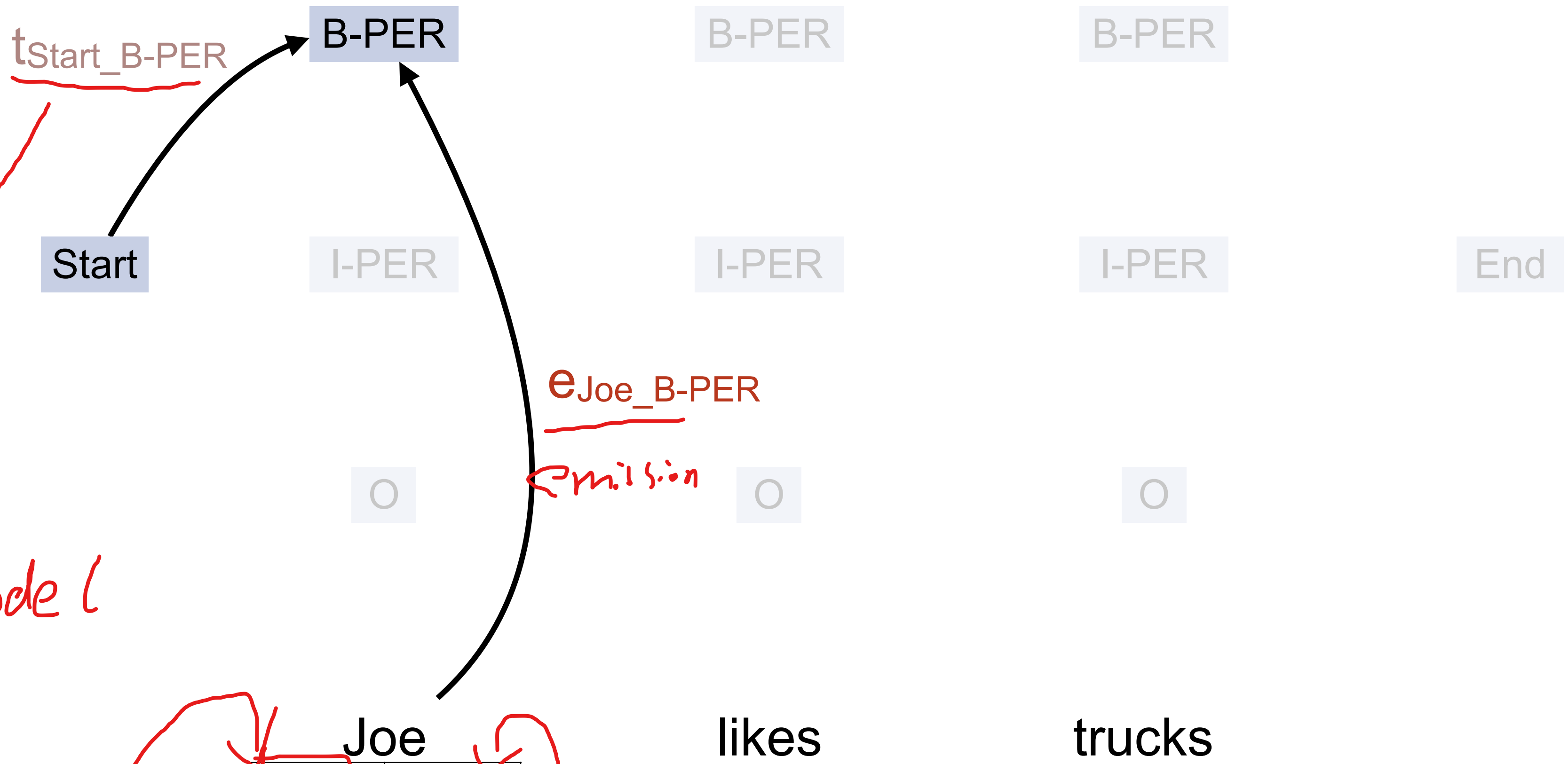
menti.com 6274 6616

We calculate intermediate scores, first for each possible start

emission transition

$$\boxed{prob_1} = e_{Joe\_B\text{-}PER} * t_{Start\_B\text{-}PER}$$

transition



Prob.

$prob_1$	Start

Previous state

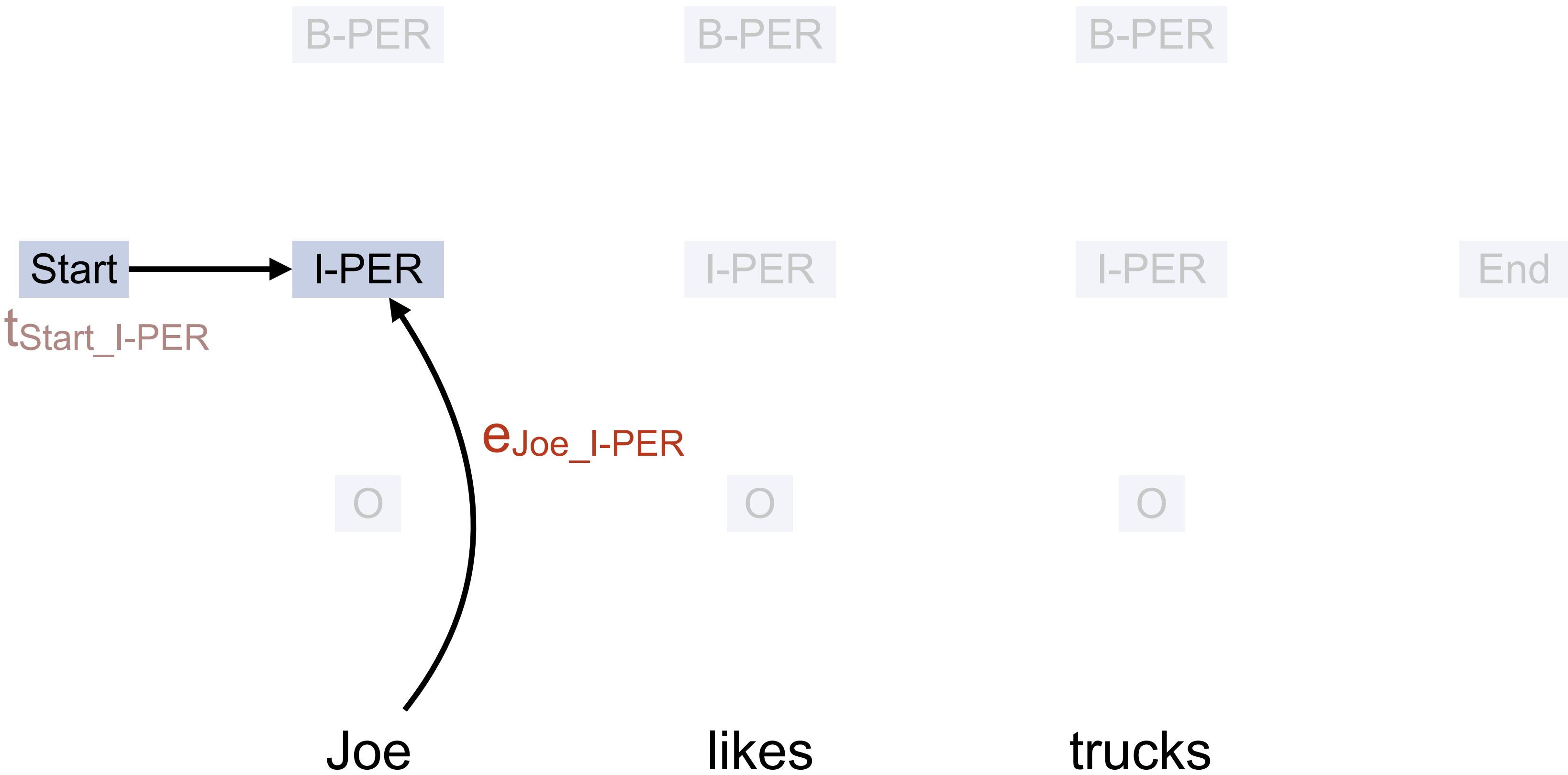
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, first for each possible start

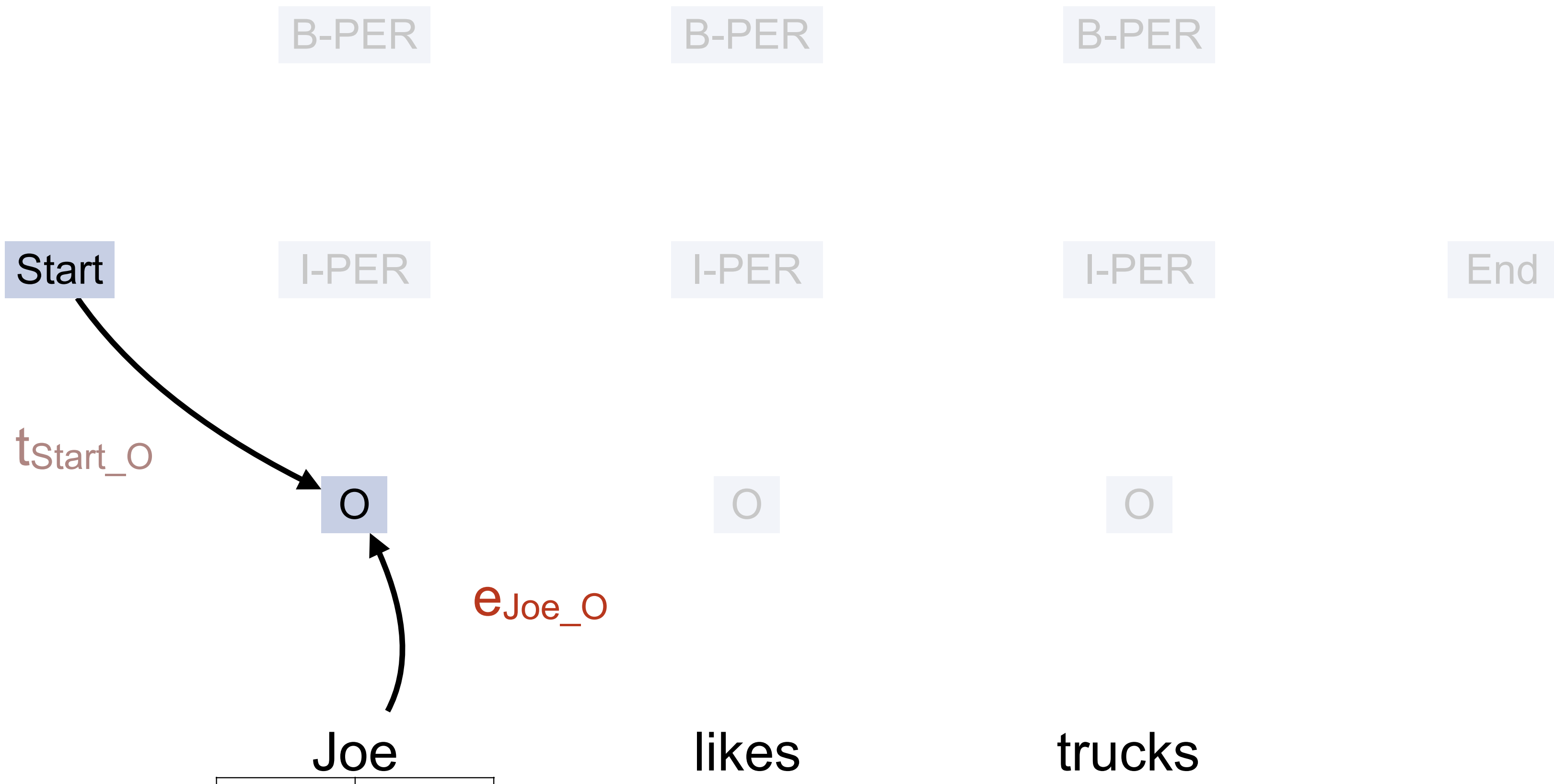
$$prob_2 = e_{Joe\_I\text{-}PER} * t_{Start\_I\text{-}PER}$$



$prob_1$	Start
$prob_2$	Start

We calculate intermediate scores, first for each possible start

$$prob_3 = e_{Joe\_O} * t_{Start\_O}$$



$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

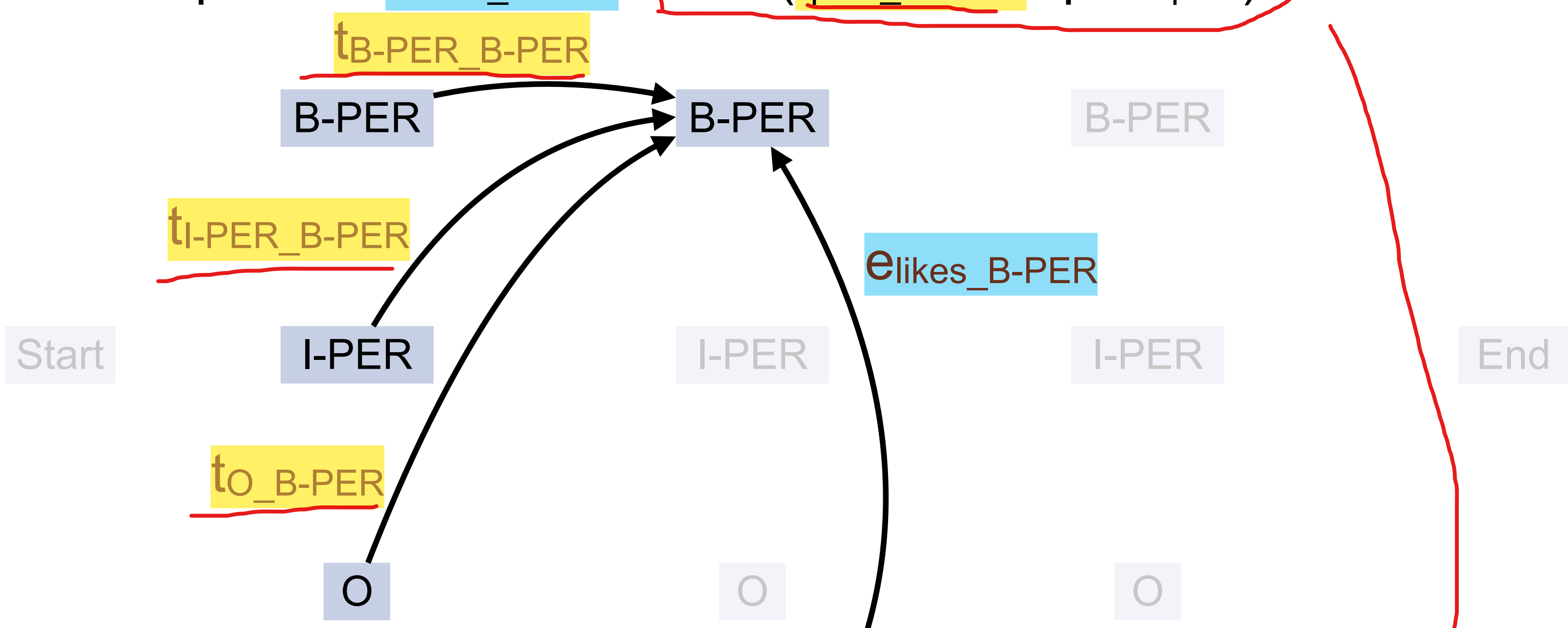
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, then continuations, noting the best

$$prob_4 = e_{likes\_B\text{-}PER} * \max(t_{prev\_B\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	

trucks

Example

O, B-PER is max

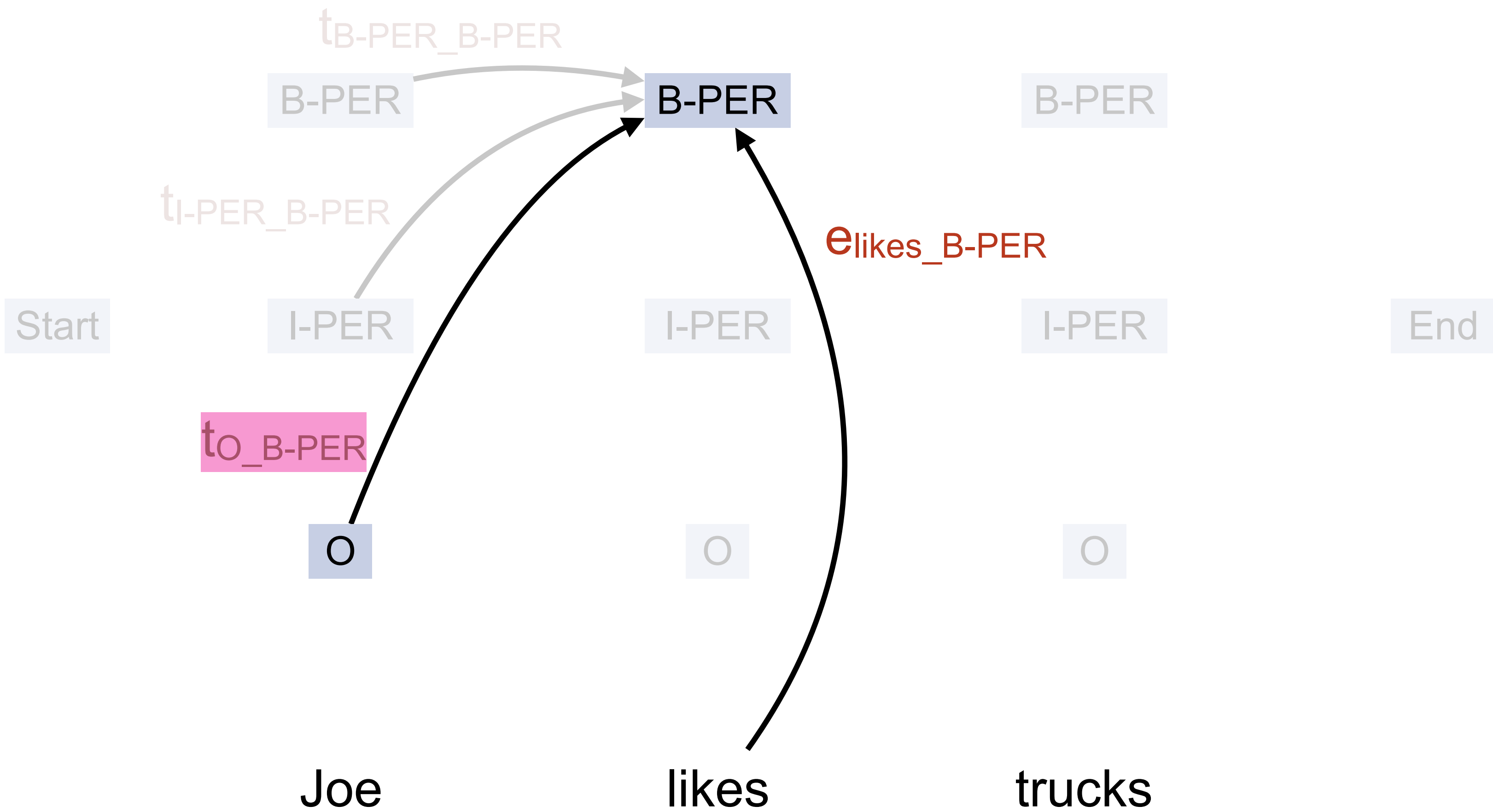
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, then continuations, noting the best

$$prob_4 = e_{likes\_B\text{-}PER} * \max(t_{prev\_B\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

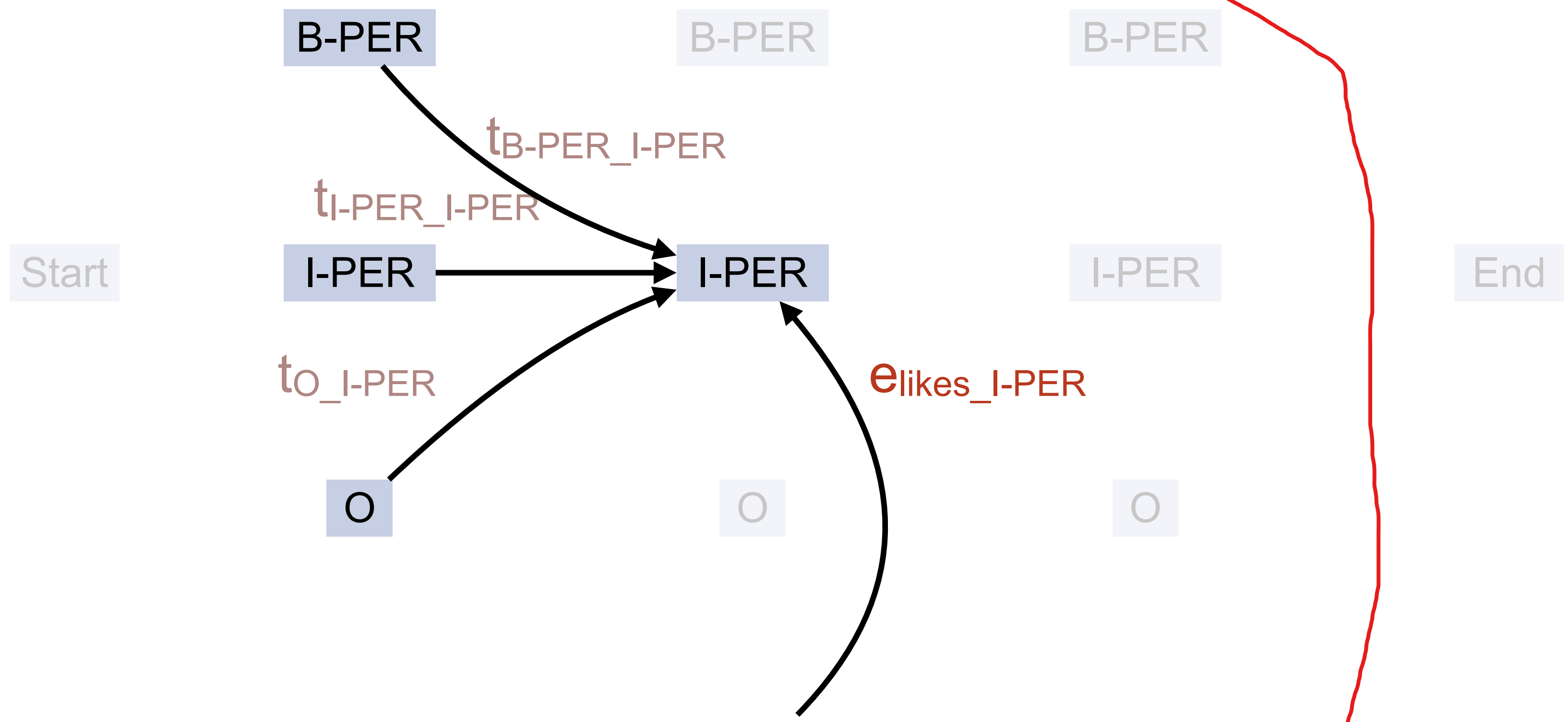
$prob_4$	O

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, then continuations, noting the best

$$\text{prob}_5 = e_{\text{likes\_I-PER}} * \max(t_{\text{prev\_I-PER}} * \text{prob}_{\text{prev}})$$



Joe

prob_1	Start
prob_2	Start
prob_3	Start

likes

prob_4	O
prob_5	

trucks

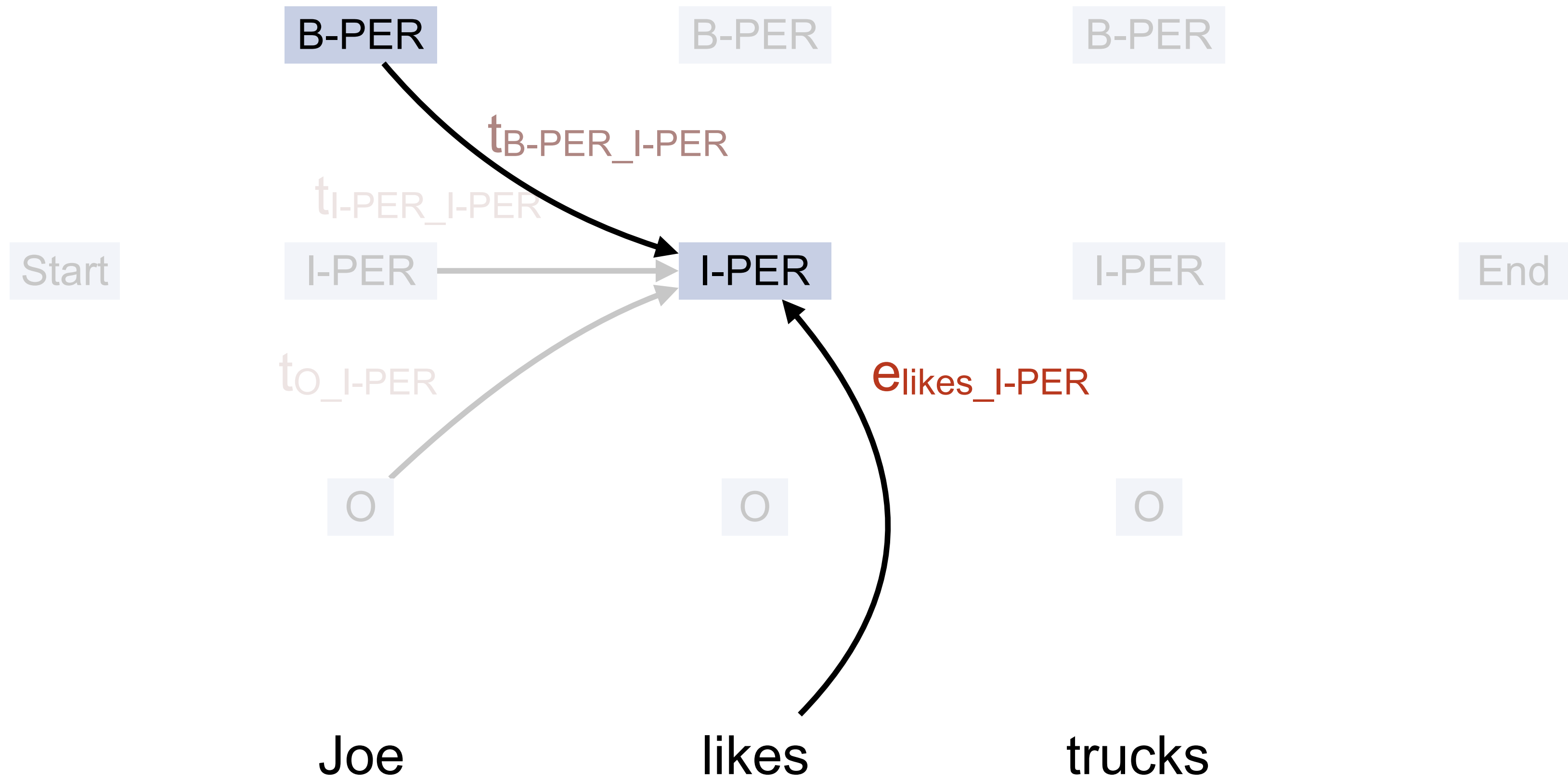
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, then continuations, noting the best

$$prob_5 = e_{likes\_I\text{-}PER} * \max(t_{prev\_I\text{-}PER} * prob_{prev})$$



$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

$prob_4$	O
$prob_5$	B-PER

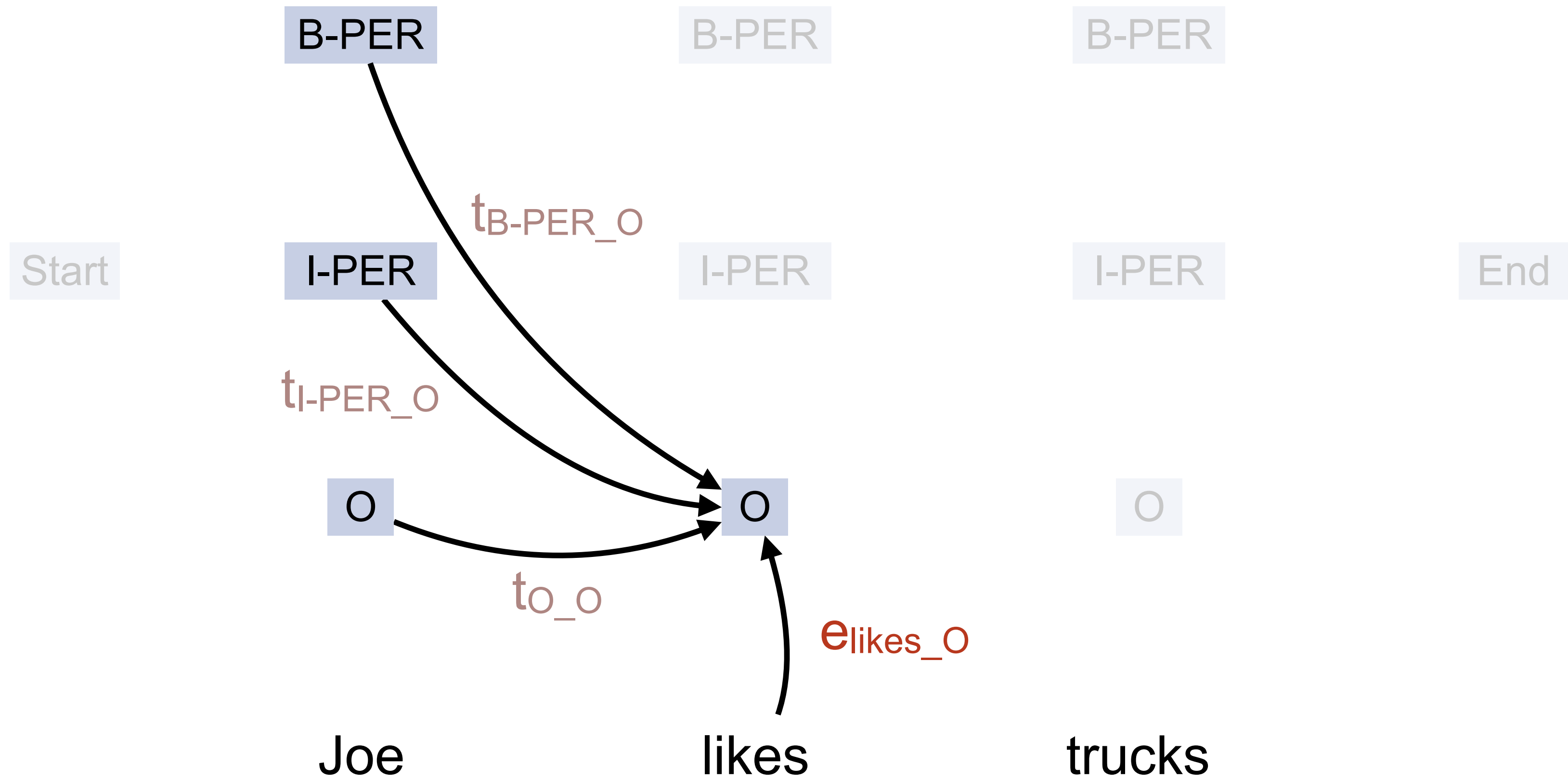
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, then continuations, noting the best

$$\text{prob}_6 = e_{\text{likes\_O}} * \max(t_{\text{prev\_O}} * \text{prob}_{\text{prev}})$$



prob_1	Start
prob_2	Start
prob_3	Start

prob_4	O
prob_5	B-PER
prob_6	

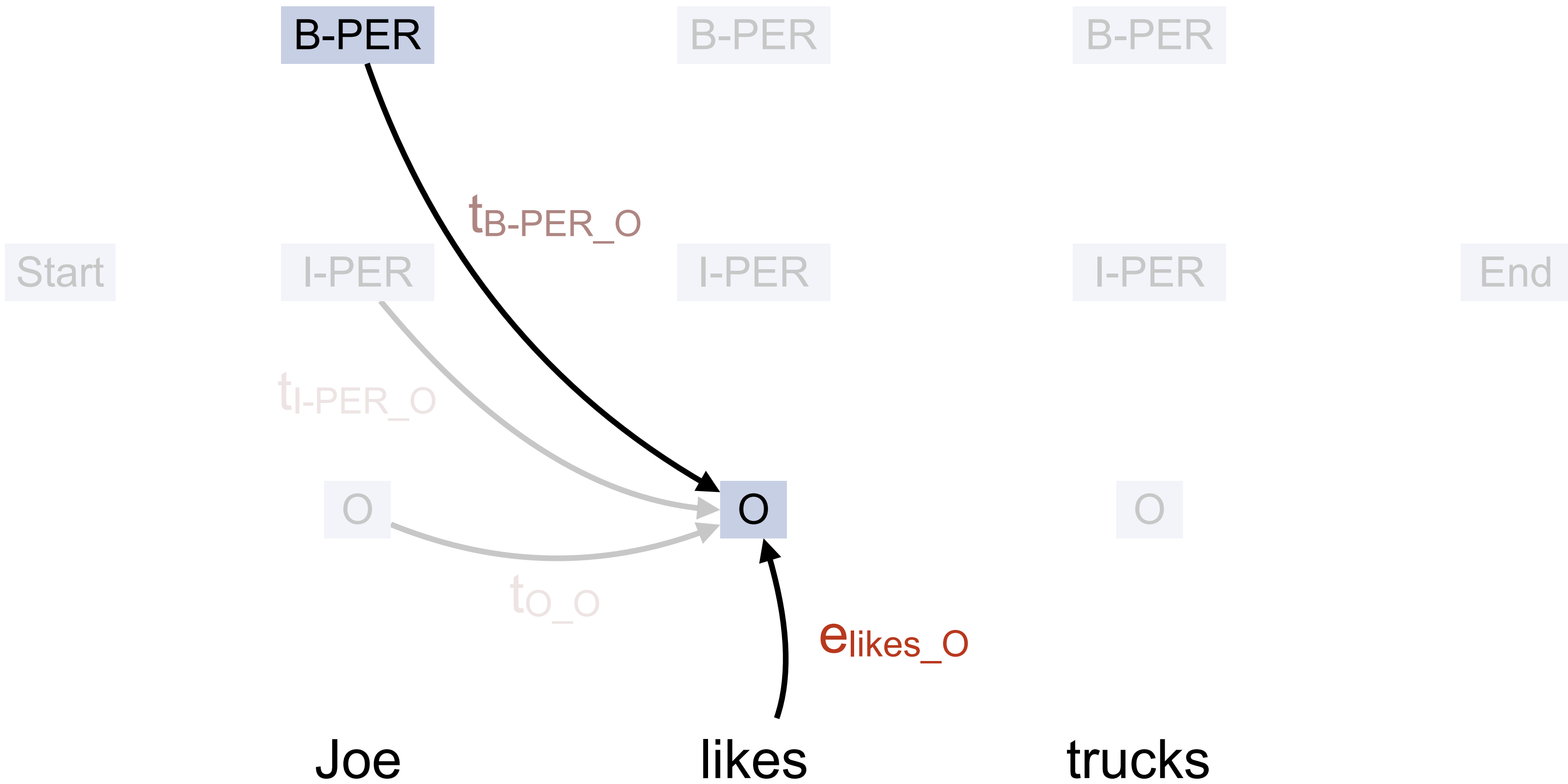
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We calculate intermediate scores, then continuations, noting the best

$$\text{prob}_6 = e_{\text{likes\_O}} * \max(t_{\text{prev\_O}} * \text{prob}_{\text{prev}})$$



prob_1	Start
prob_2	Start
prob_3	Start

prob_4	O
prob_5	B-PER
prob_6	B-PER

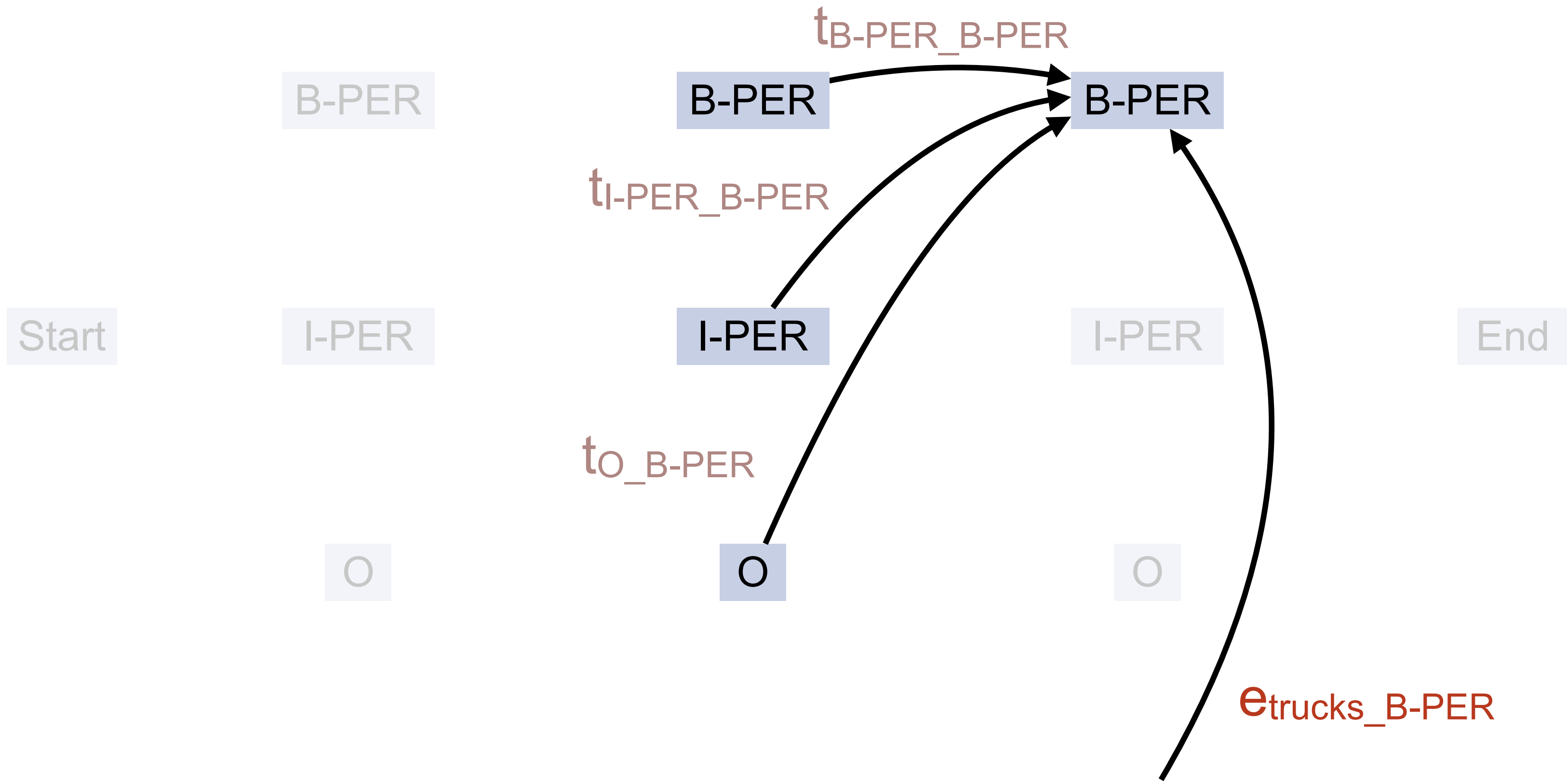
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Each step only looks back one step

$$prob_7 = e_{trucks\_B\text{-}PER} * \max(t_{prev\_B\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	

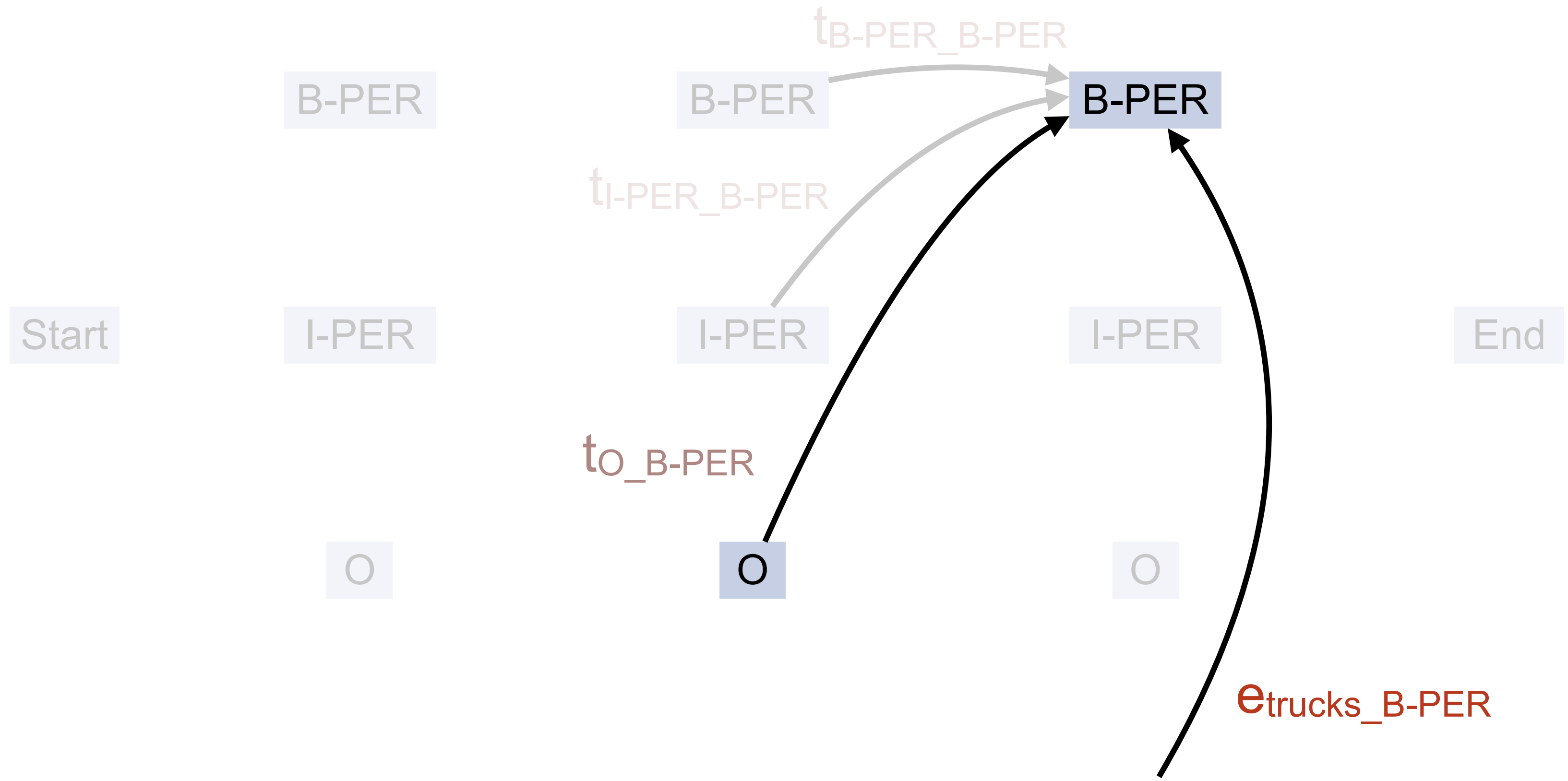
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Each step only looks back one step

$$prob_7 = e_{trucks\_B\text{-}PER} * \max(t_{prev\_B\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O

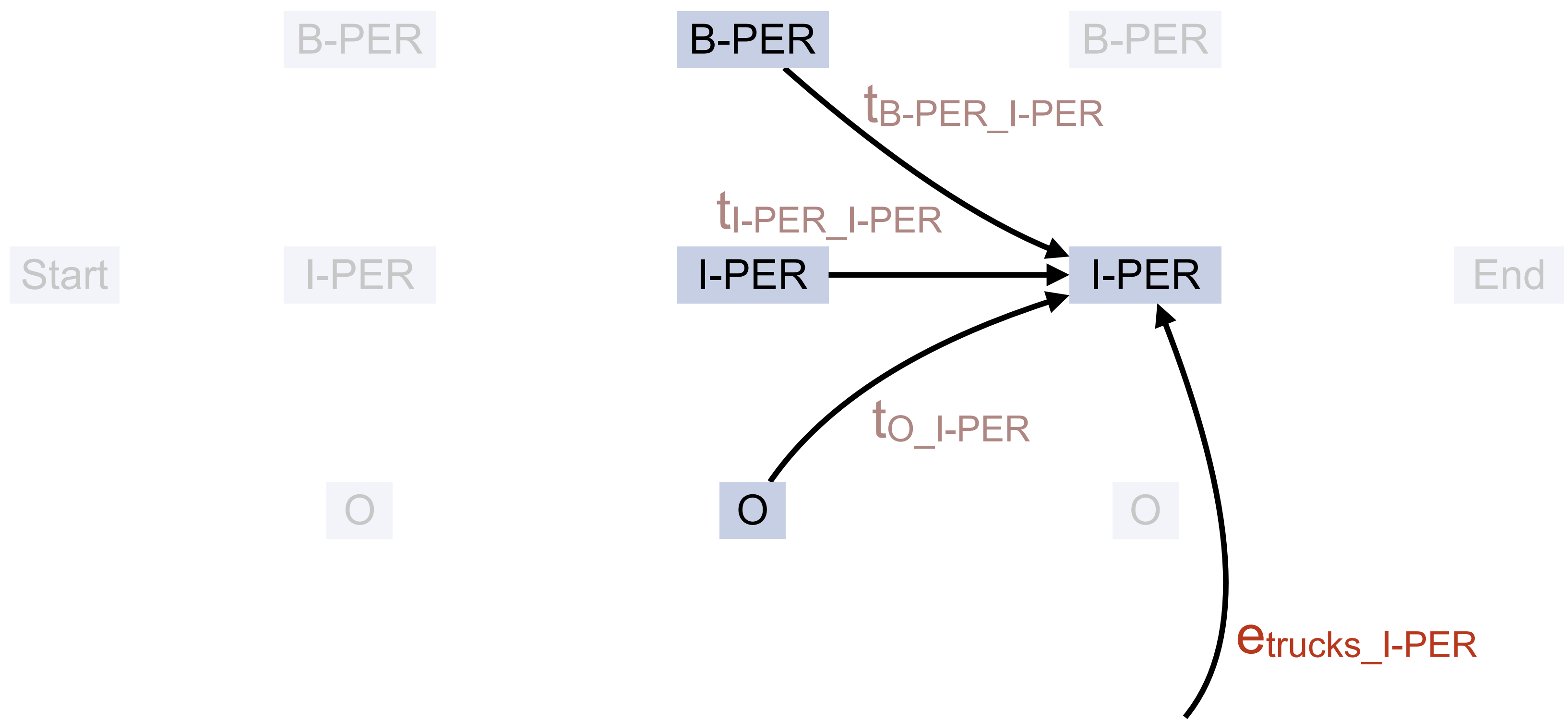
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Each step only looks back one step

$$prob_8 = e_{trucks\_I\text{-}PER} * \max(t_{prev\_I\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

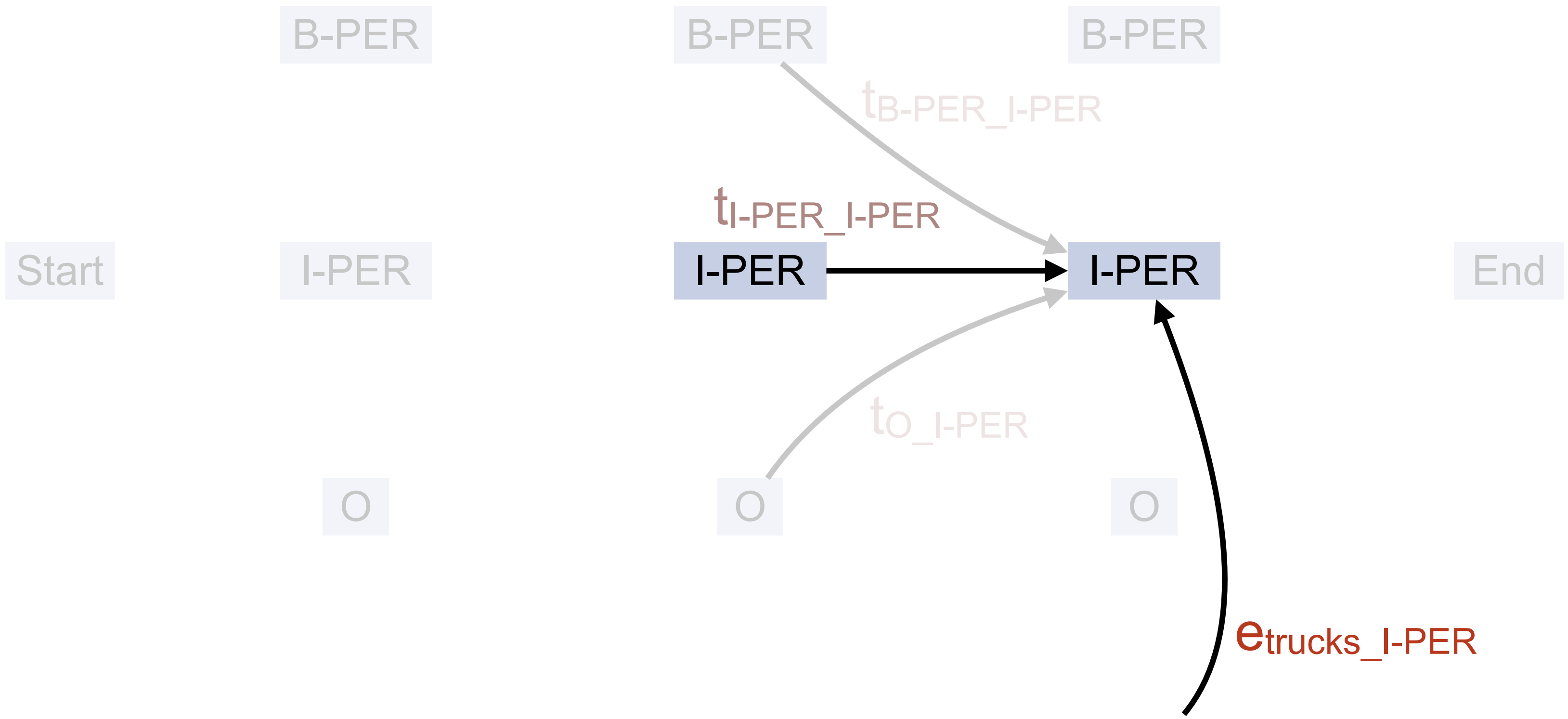
$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O
$prob_8$	

Each step only looks back one step

$$prob_8 = e_{trucks\_I\text{-}PER} * \max(t_{prev\_I\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O
$prob_8$	I-PER

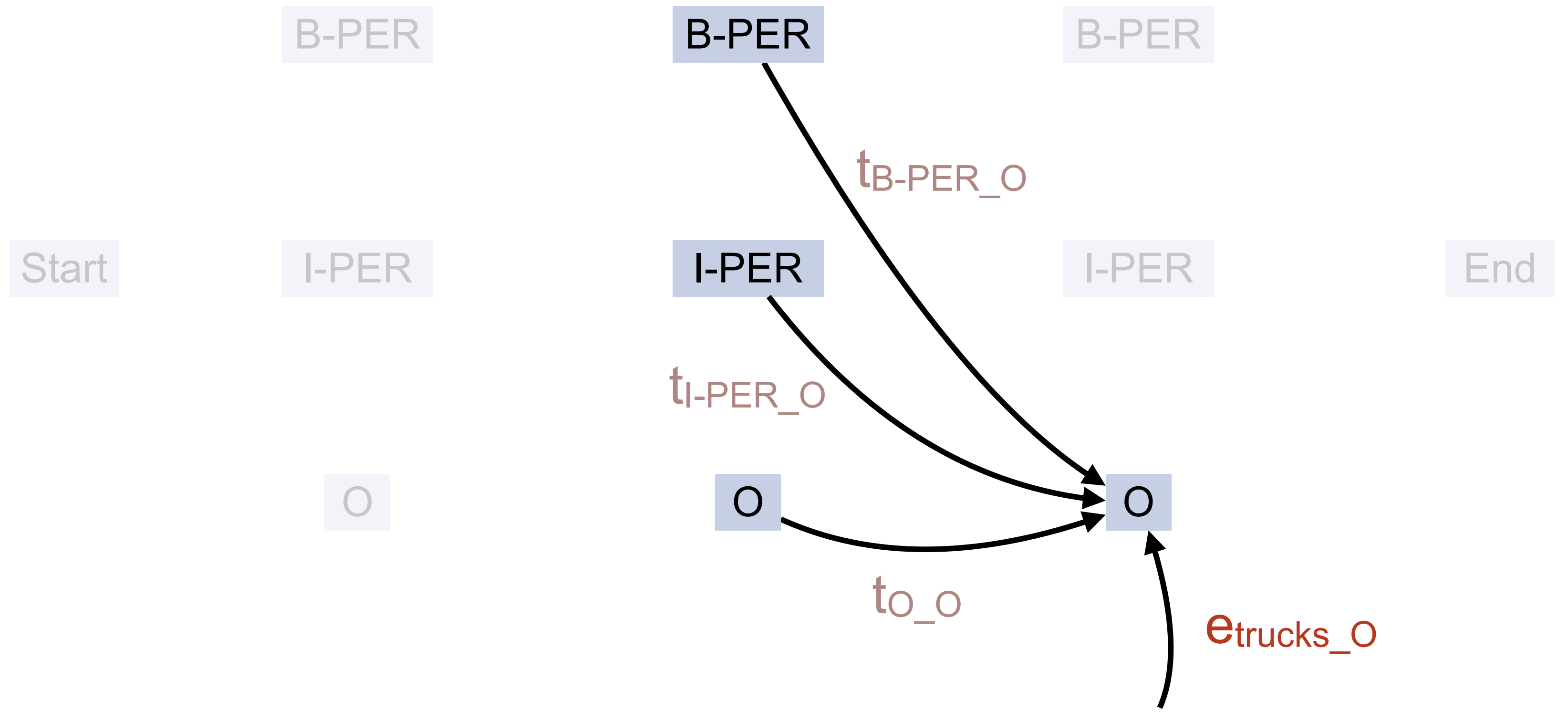
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Each step only looks back one step

$$prob_9 = e_{trucks\_I\text{-}PER} * \max(t_{prev\_I\text{-}PER} * prob_{prev})$$



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O
$prob_8$	I-PER
$prob_9$	

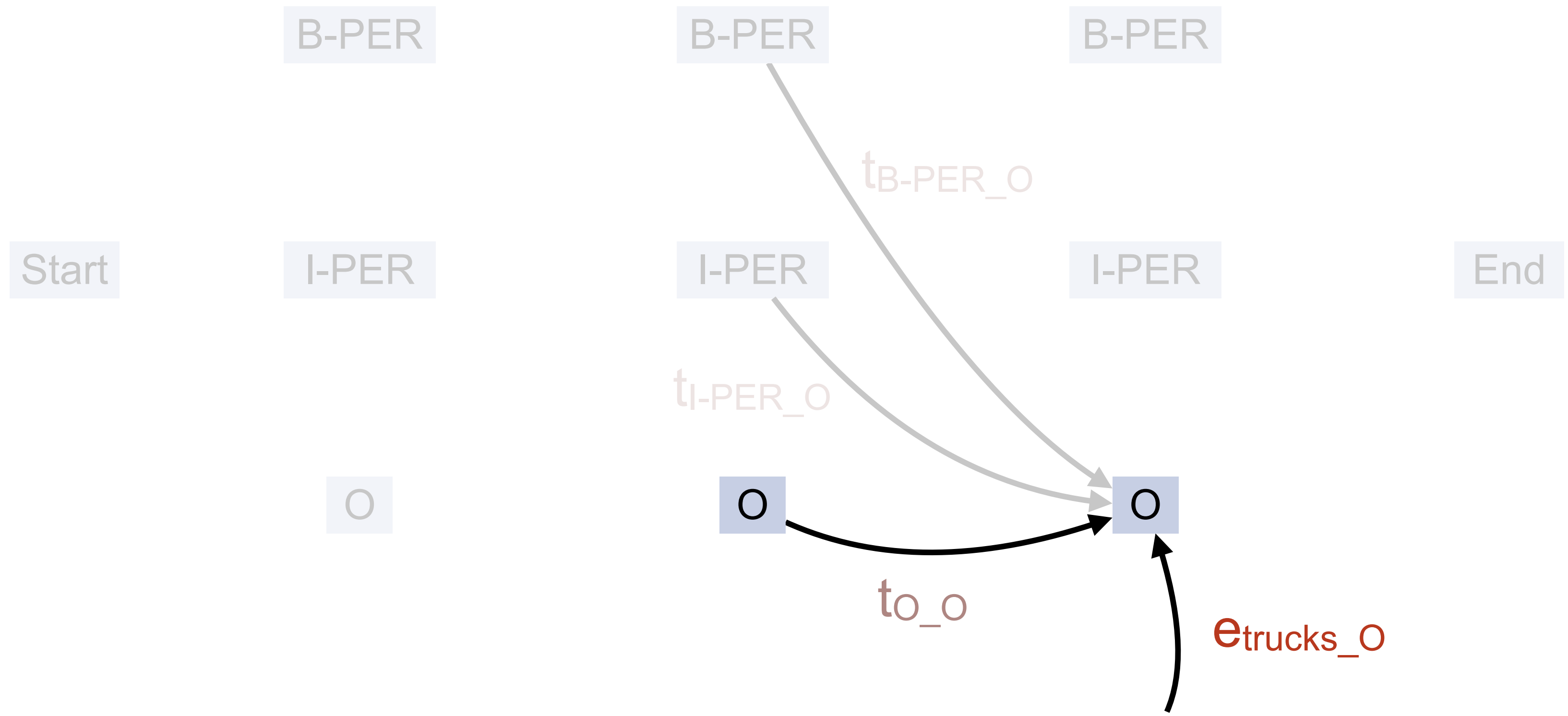
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Each step only looks back one step

$$prob_9 = e_{trucks\_I\text{-}PER} * \max(t_{prev\_I\text{-}PER} * prob_{prev})$$



Joe	
$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes	
$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks	
$prob_7$	O
$prob_8$	I-PER
$prob_9$	O

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Ending

Each step only looks back one step

$$\boxed{prob_{10} = \max(t_{prev\_End} * prob_{prev})}$$

Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O
$prob_8$	I-PER
$prob_9$	O

$prob_{10}$	

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Example

Each step only looks back one step

t_{O_End} $prob_9$

$$prob_{10} = \max(t_{prev\_End} * prob_{prev})$$

Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O
$prob_8$	I-PER
$prob_9$	O

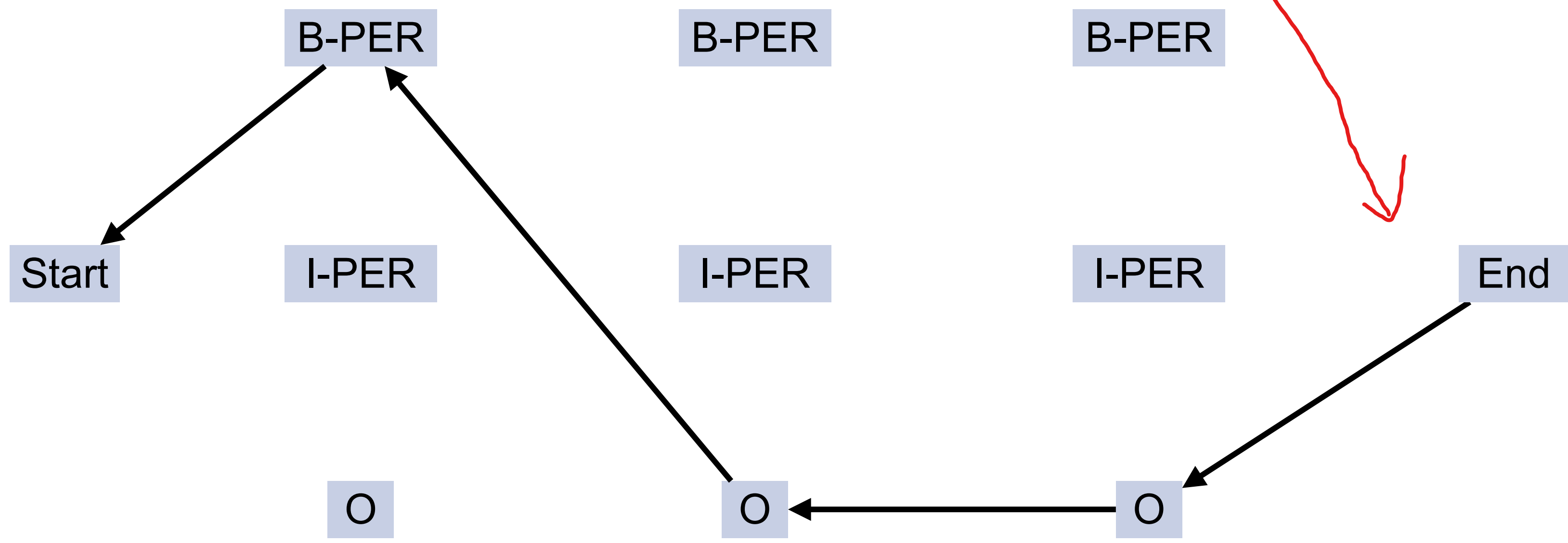
$prob_{10}$	O

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

To get the path, follow the references backwards to the start



Joe

$prob_1$	Start
$prob_2$	Start
$prob_3$	Start

likes

$prob_4$	O
$prob_5$	B-PER
$prob_6$	B-PER

trucks

$prob_7$	O
$prob_8$	I-PER
$prob_9$	O

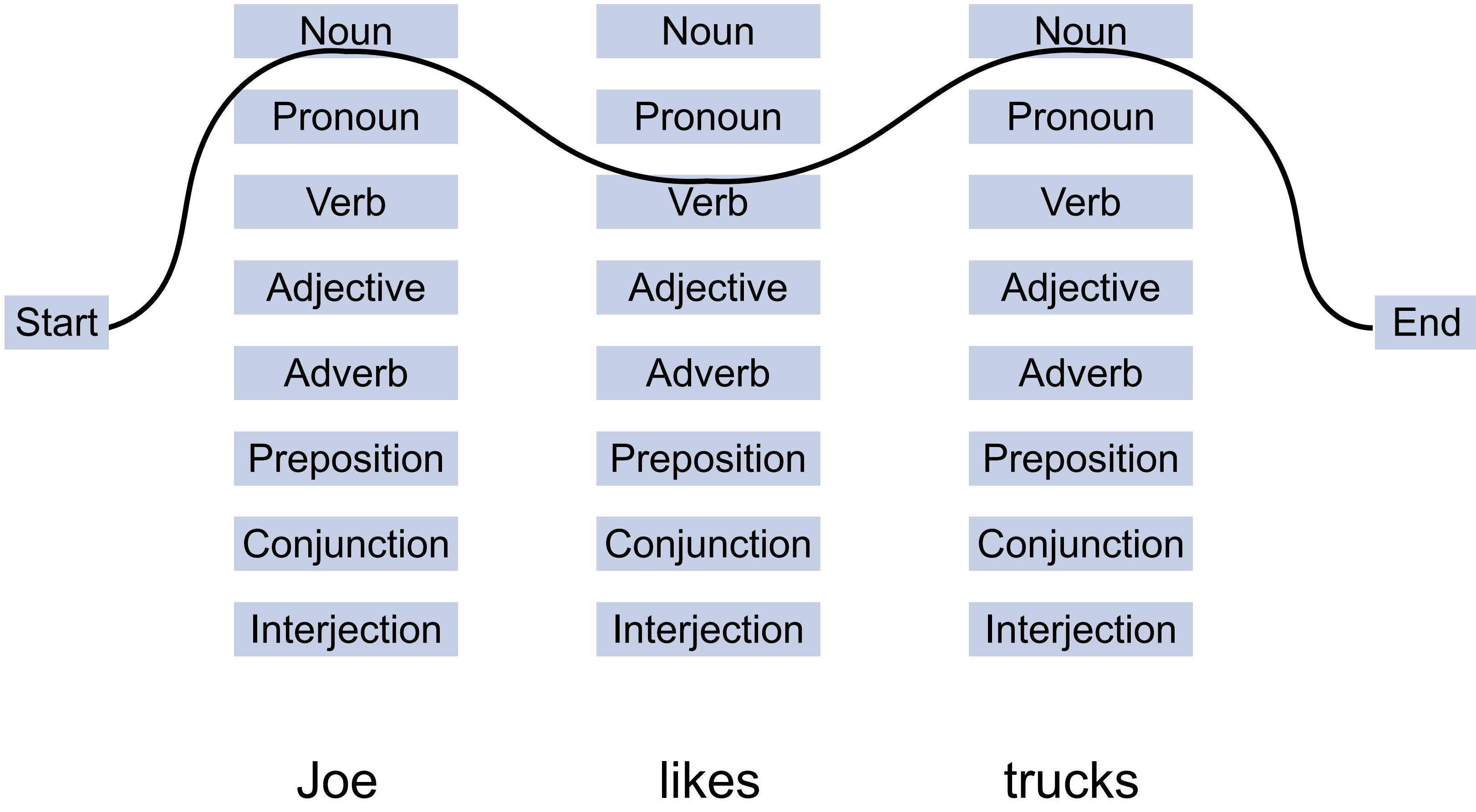
$prob_{10}$	O

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We can apply the same idea to POS tagging



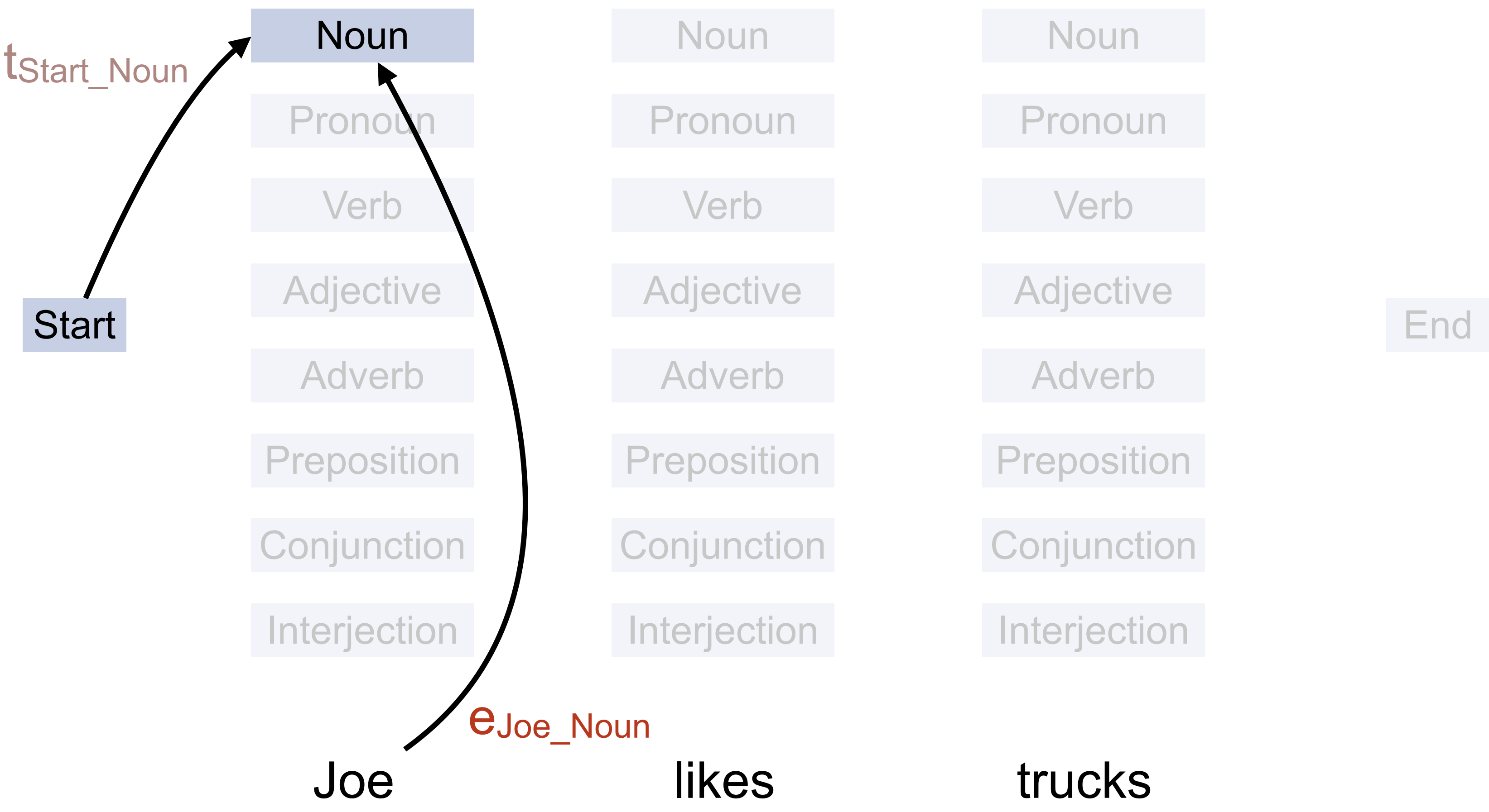
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

One more example

We can apply the same idea to POS tagging



Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Forward

Pseudocode for the algorithm to calculate scores and previous labels

```
Create storage for best option for each
(position, label) pair

For each position in input:
    For each label:
        Best score = -infinity
        Best previous = None
        For each previous_label:
            Score = previous_label's prob *
                    prob(label | input and
                    previous label)
            If score is better than best:
                Update best
        Save best score and best previous for
        this (position, label) pair
```

Called the Viterbi algorithm

(note, also need to account for start and end)

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Pseudocode for the algorithm to get the path

path back

```
Path = []
Current label = End
For position from end of sequence to start:
    Current label = Best for this position and
                    the current label
    Append current label to Path

Reverse path
```

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Time complexity is $\boxed{O(|words| * |labels|^2)}$

Joe, likes, trucks

|words| steps

B-PER
I-PER
O
likes

In each step, consider |labels|

O

For each option, consider |labels| previous options similar to Perceptron

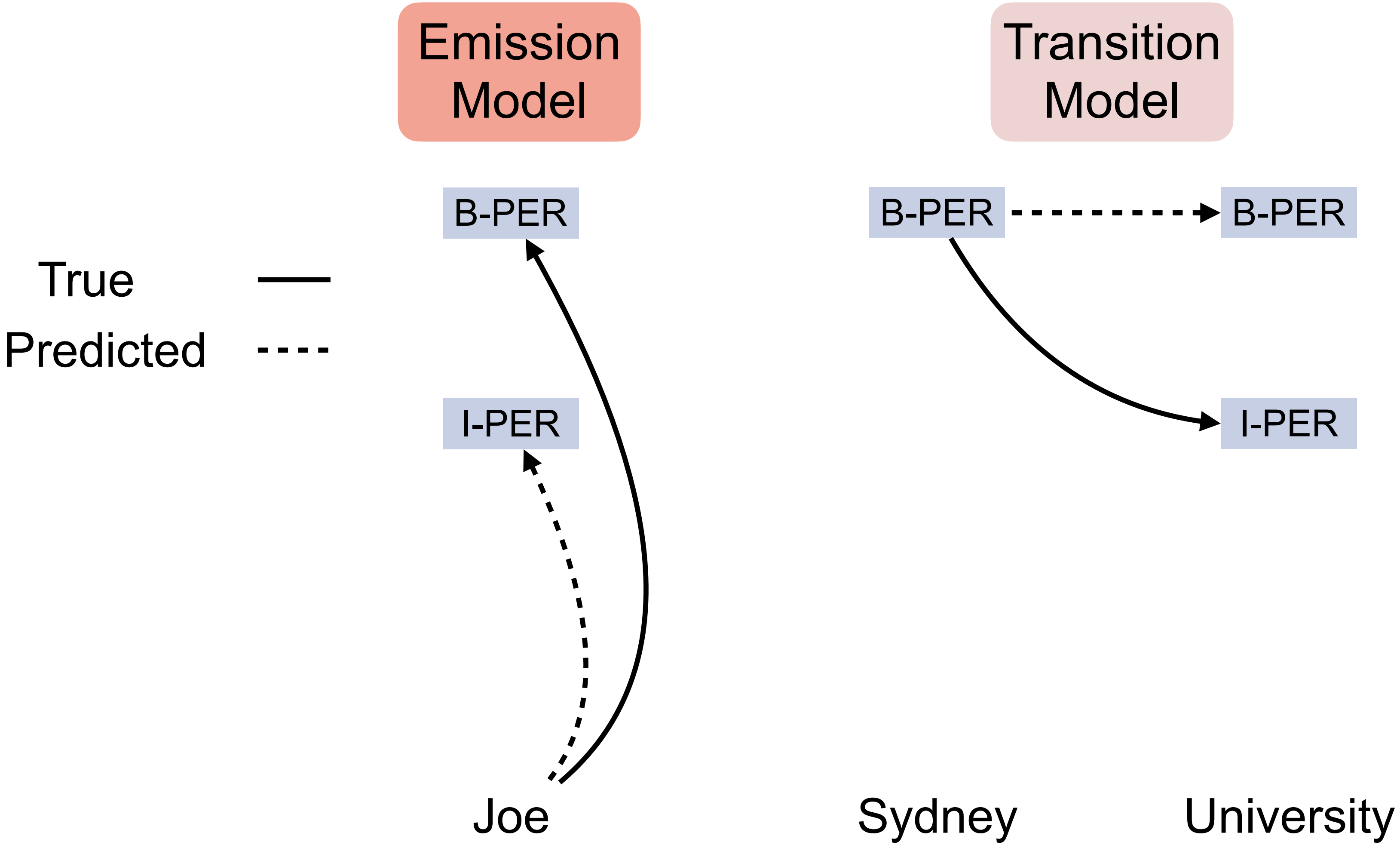
ex. backprob

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Learn by comparing the final output with the true answer



menti.com 6274 6616

The model can use all of the input, and the previous label

Emission Model

O

e_{likes_O}

Joe likes trucks

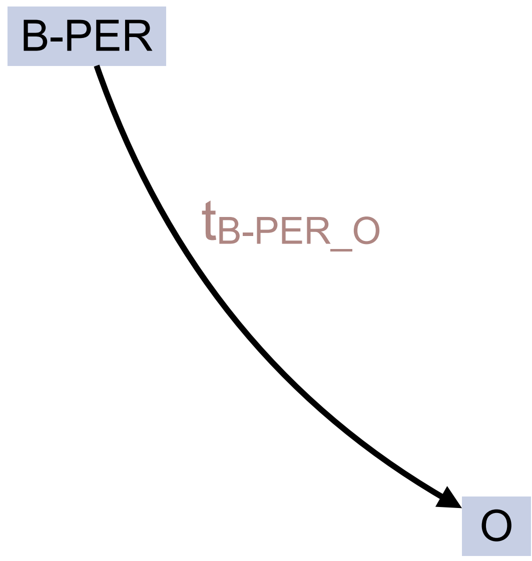
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

The model can use all of the input, and the previous label

Transition Model



Joe likes trucks

COMP 4446 / 5046
Lecture 5, 2025

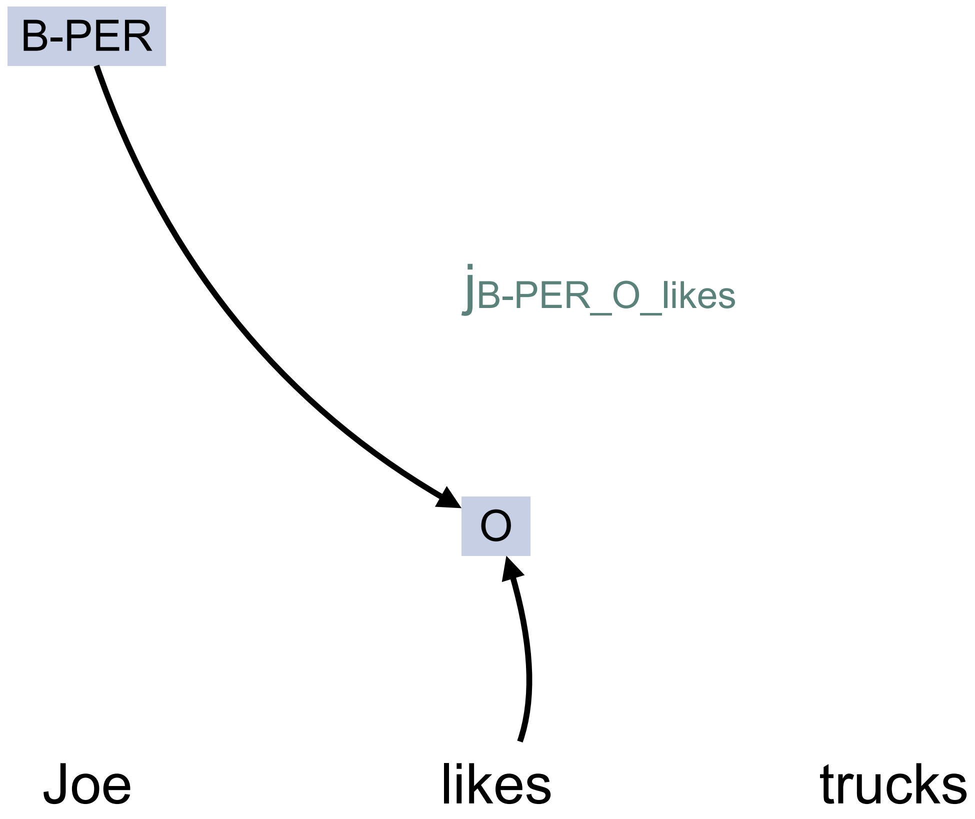
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We can also define a model that does both scores at once

Joint Model

$$\text{score}_{\text{current}} = \max(j_{\text{prev\_label\_word}} * \text{score}_{\text{prev}})$$



If we are willing to increase complexity, we can use more label context

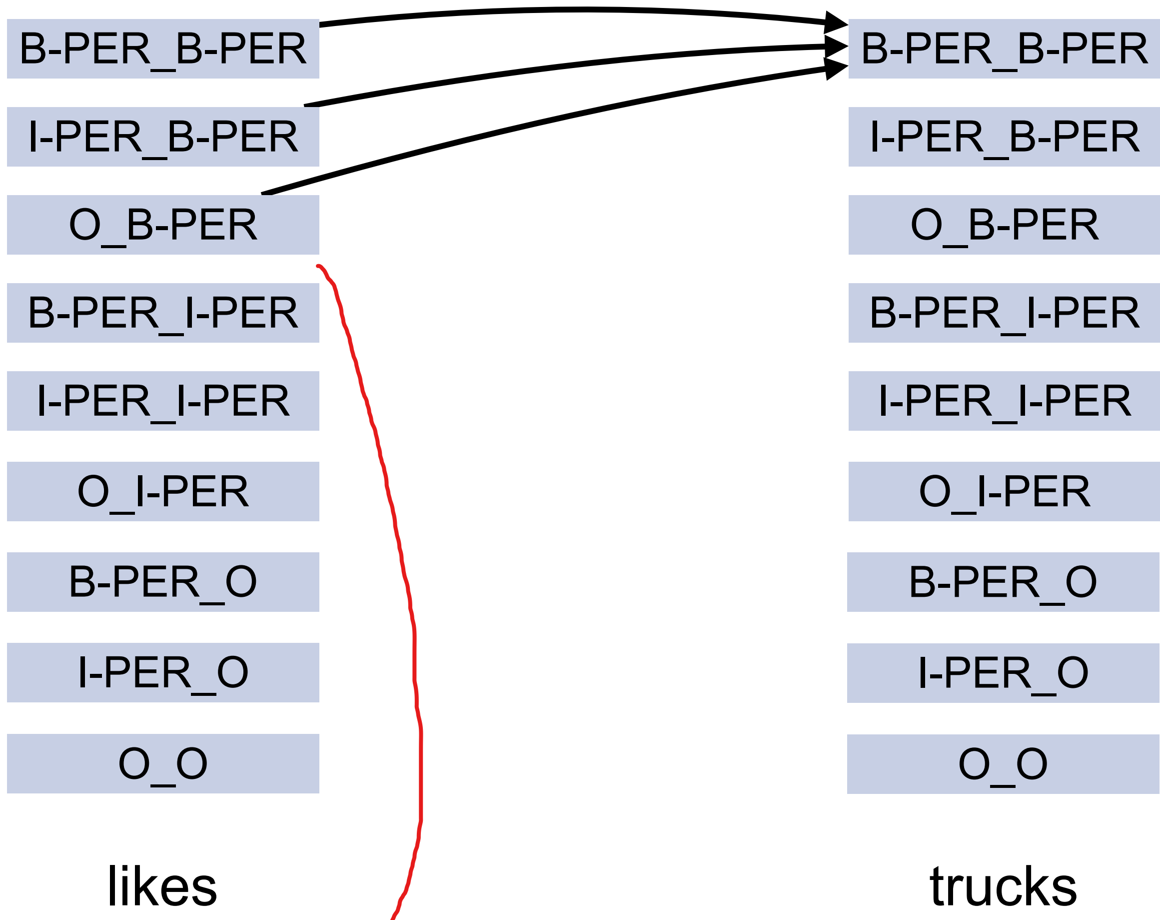
Previous_Current

	Joe	likes	trucks	
		B-PER_B-PER	B-PER_B-PER	
		I-PER_B-PER	I-PER_B-PER	
		O_B-PER	O_B-PER	
	Start_B-PER	B-PER_I-PER	B-PER_I-PER	
Start	Start_I-PER	I-PER_I-PER	I-PER_I-PER	End
	Start_O	O_I-PER	O_I-PER	
		B-PER_O	B-PER_O	
		I-PER_O	I-PER_O	
		O_O	O_O	

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

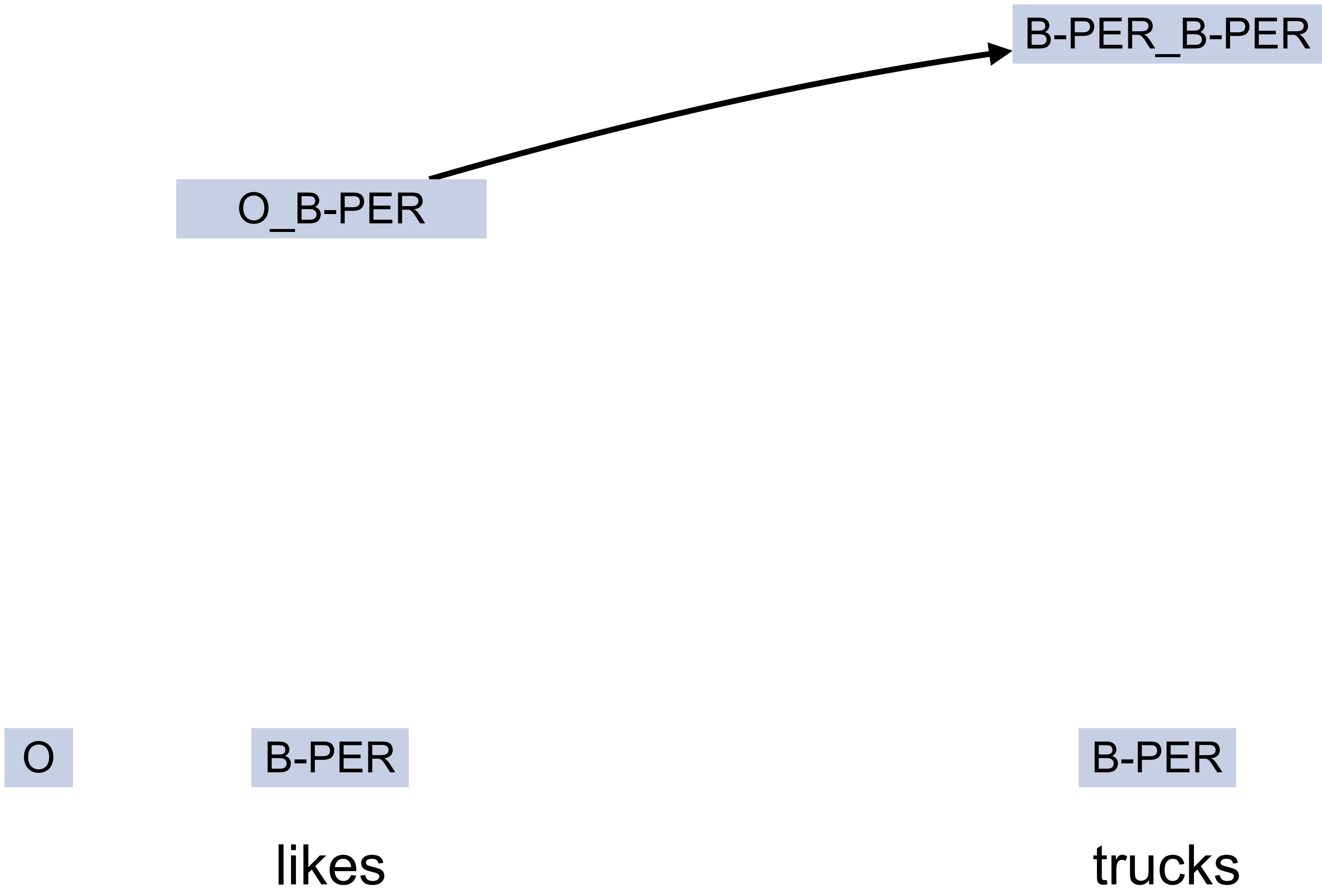
menti.com 6274 6616

If we are willing to increase complexity, we can use more label context



more complex, can aware
before token

If we are willing to increase complexity, we can use more label context



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Now, time complexity is $\boxed{O(|words| * |labels|^3)}$

difference

Joe, likes, trucks

|words| steps

B-PER_B-PER

I-PER_B-PER

O_B-PER

…

likes

In each step, consider |labels|*|labels|

O

For each option, consider |labels| previous options

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

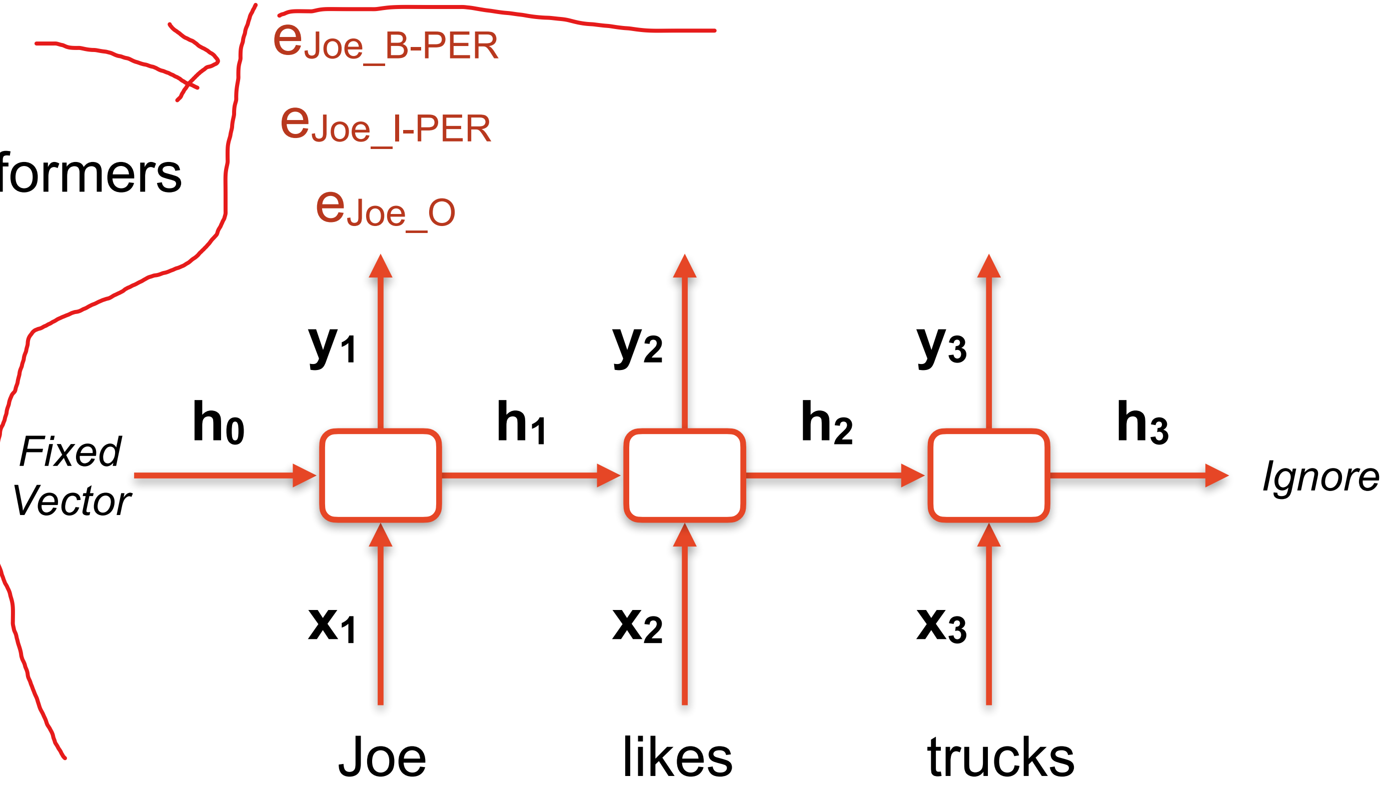
We can apply this method to a variety of models

Linear models (e.g., CRF)

Feedforward Networks

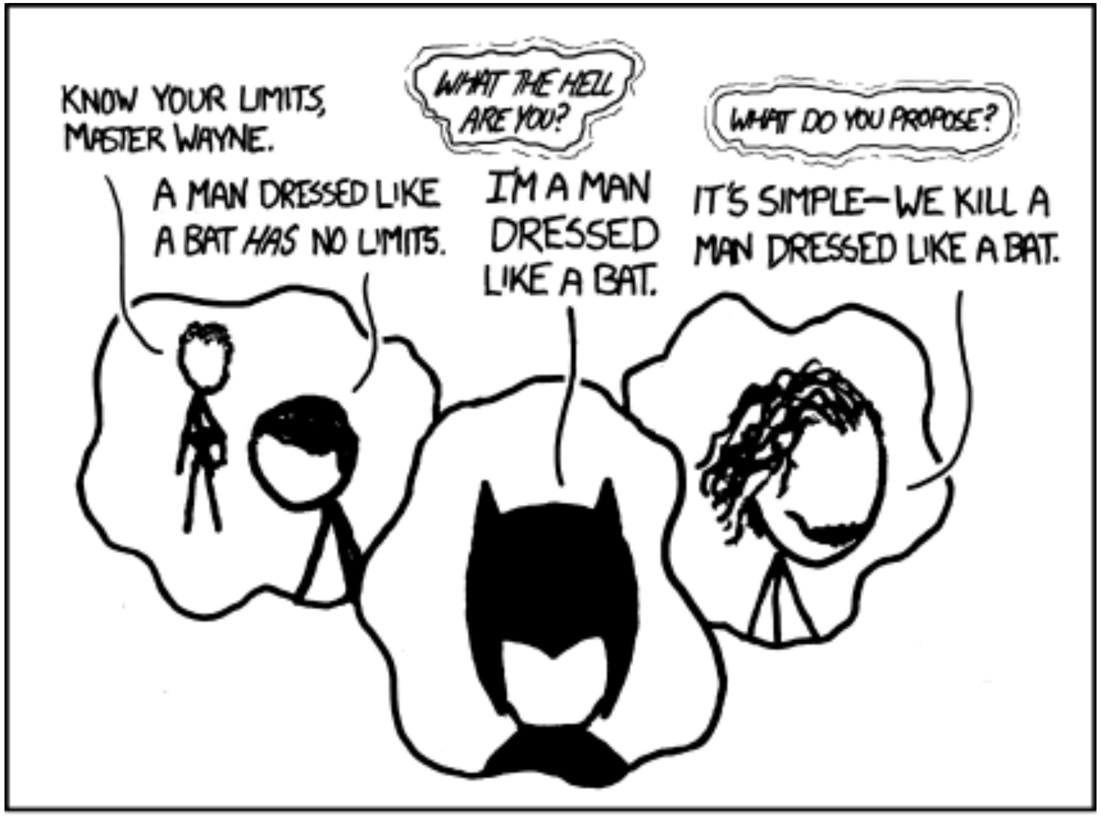
RNNs

Transformers



3 minute Break - stretch and visit Menti

menti.com
6274 6616



MY HOBBY:
WHENEVER ANYONE SAYS "BATMAN", I MENTALLY REPLACE IT WITH "A MAN DRESSED LIKE A BAT."

Batman

[I'm really worried Christopher Nolan will kill a man dressed like a bat in his next movie. (The man will be dressed like a bat, I mean. Christopher Nolan won't be, probably.)]

Source: https://xkcd.com/1004/

COMP 4446 / 5046
Lecture 5, 2025

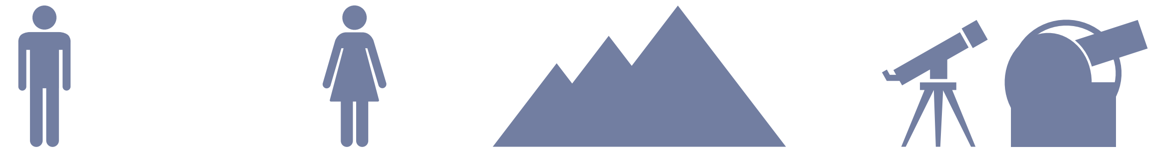
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Graph Parsing

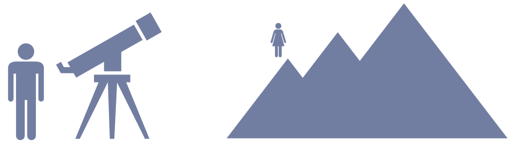
menti.com 6274 6616

Syntactic parsing can resolve ambiguities



I saw the person on the hill with a telescope

3 meanings



I saw the person on the hill with a telescope

I saw the person on the hill with a telescope

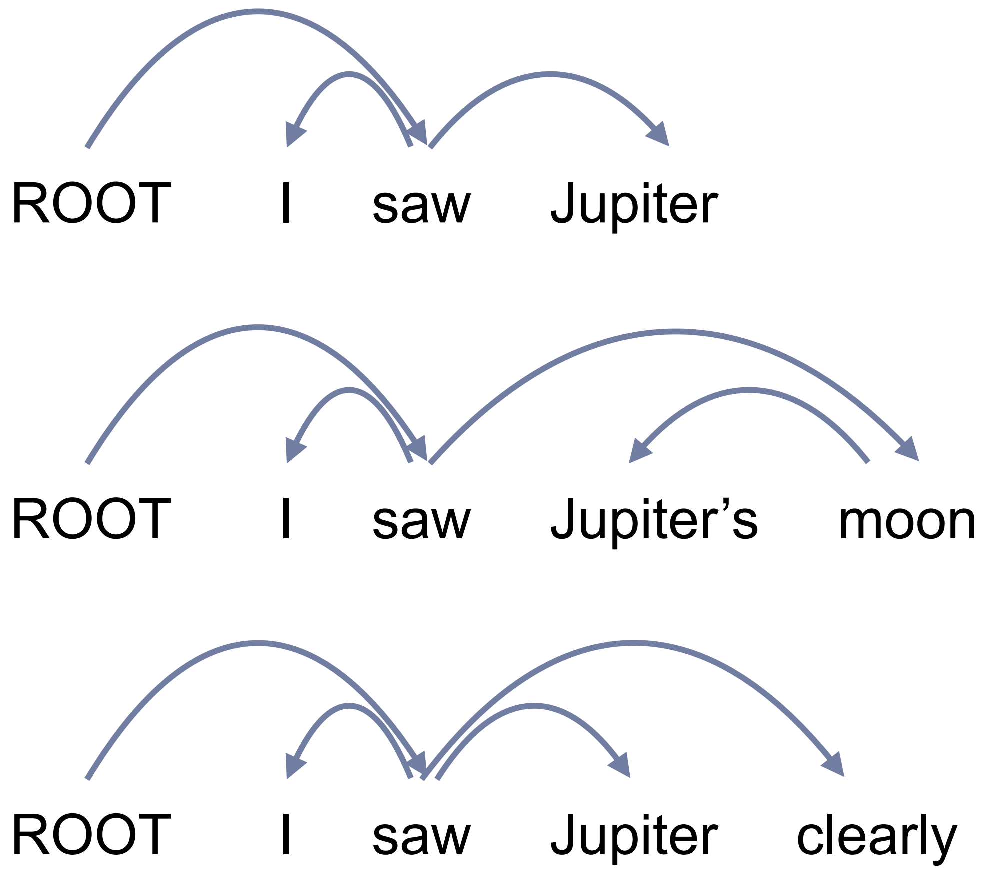
I saw the person on the hill with a telescope

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Arcs / Edges / Dependencies go from a head to a dependent

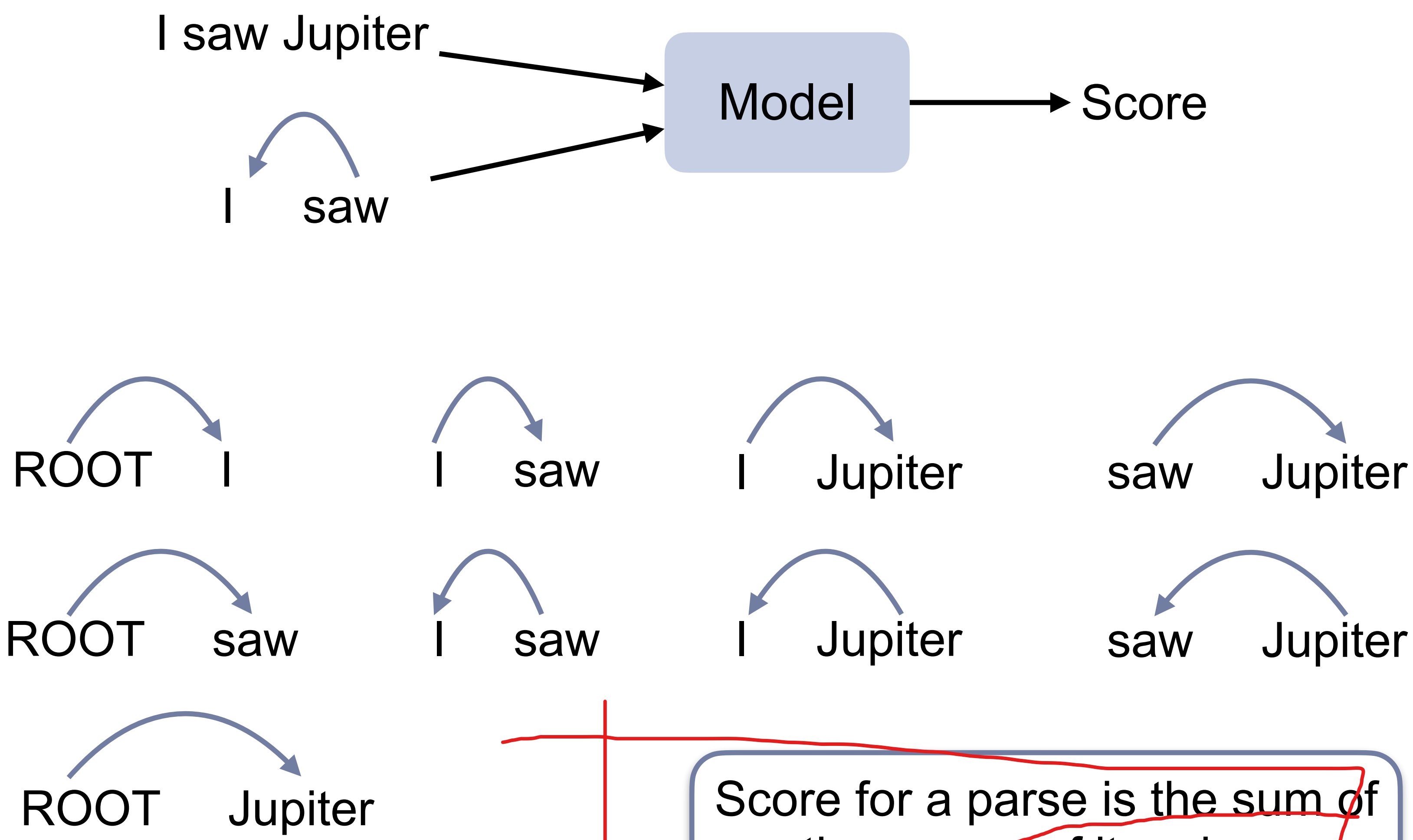


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

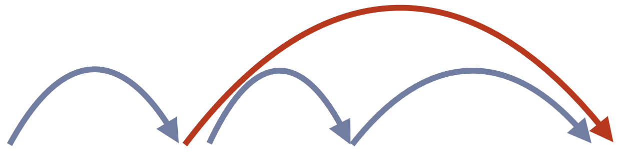
We'll use an arc model to score options



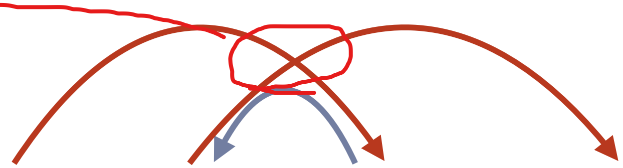
Score for a parse is the sum of the scores of its edges

Requirements for our inference method

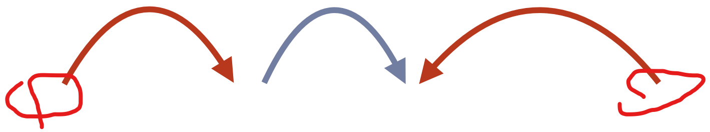
1. Produce a tree



2. No crossing arcs



3. A single root

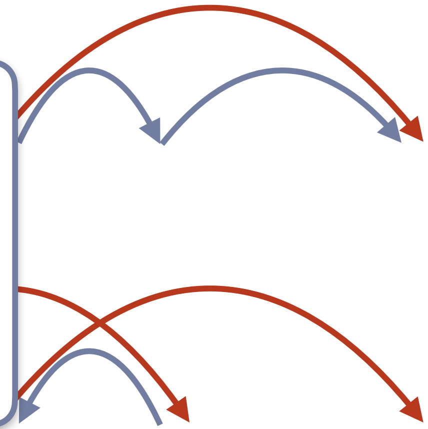


4. The tree has the highest score of all possible trees

Requirements for our inference method

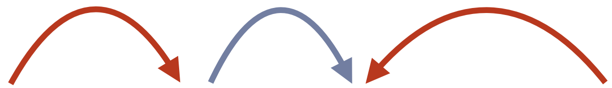
1. Produce a tree

In my research I invented an algorithm that overcomes these two constraints



2. No crossing arcs

3. A single root



4. The tree has the highest score of all possible trees

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Many inference methods - today we see dynamic programming

~~the room for my~~
in ~~my~~ class

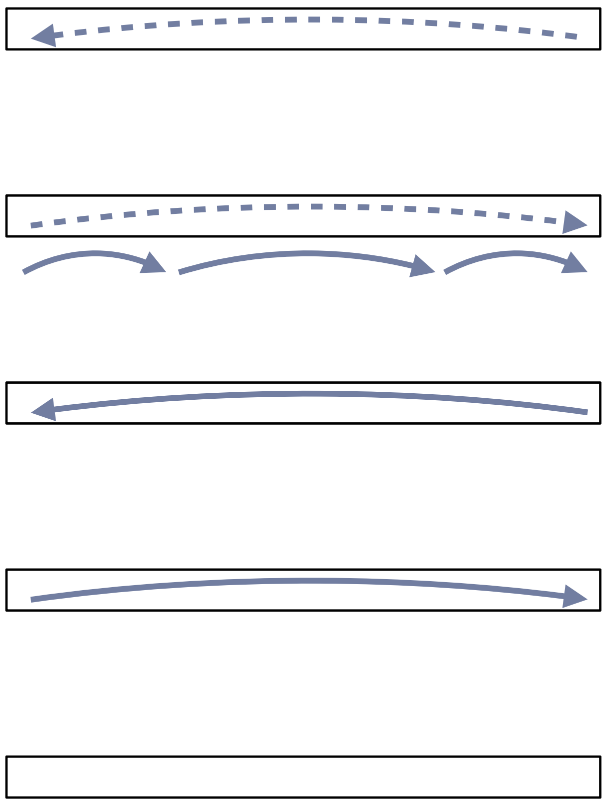
boxes meaning in rest pages

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Our inference algorithm is defined by **items** and rules

Item	Represents
	A structure with an **indirect path** from *class* to *in*
	A structure with an **indirect path** from *in* to *class*
	A structure with an **arc** from *class* to *in*
	A structure with an **arc** from *in* to *class*
	A structure with **no link** between *in* and *class*

in … class

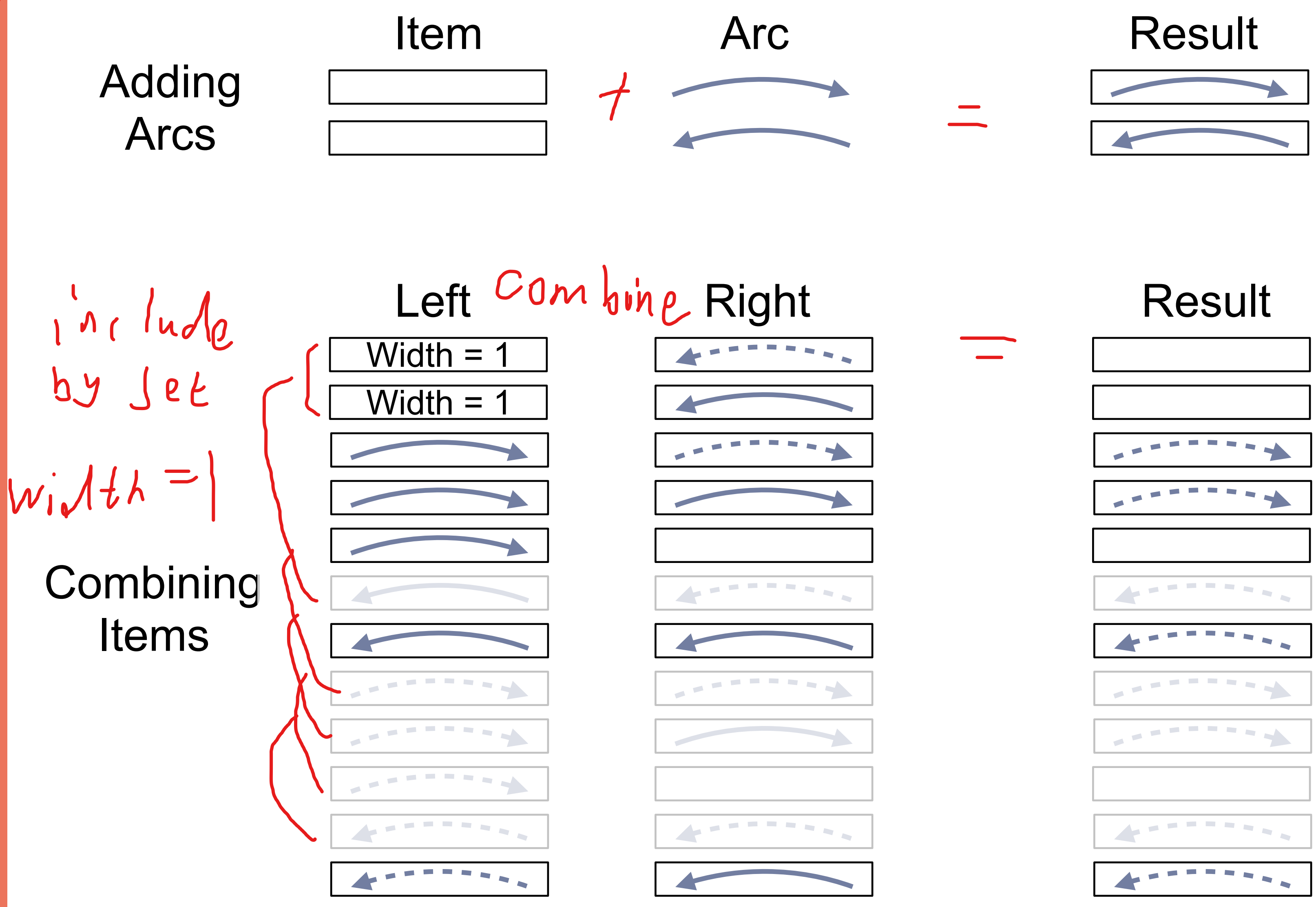
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

combine

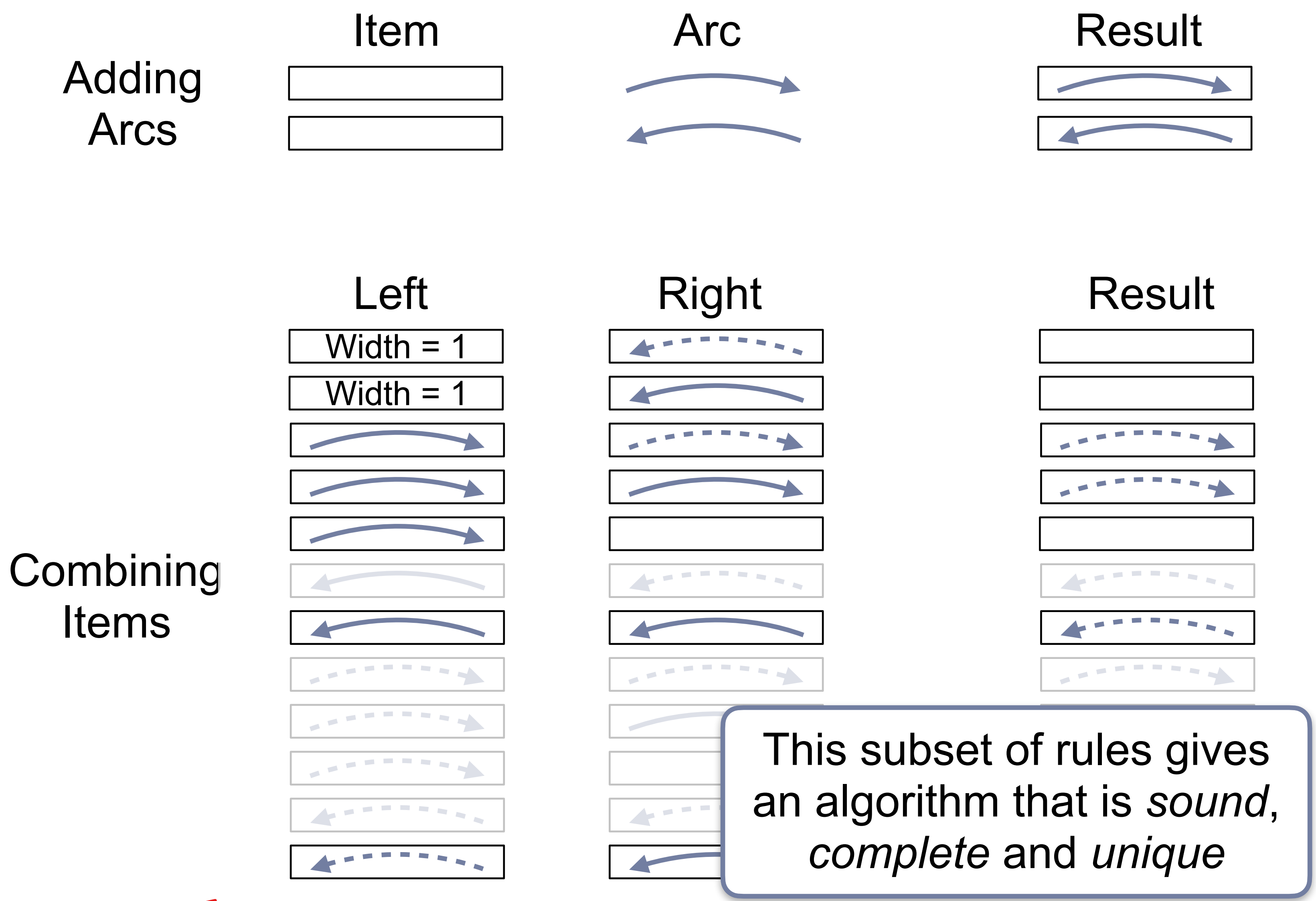
Our inference algorithm is defined by items and **rules**



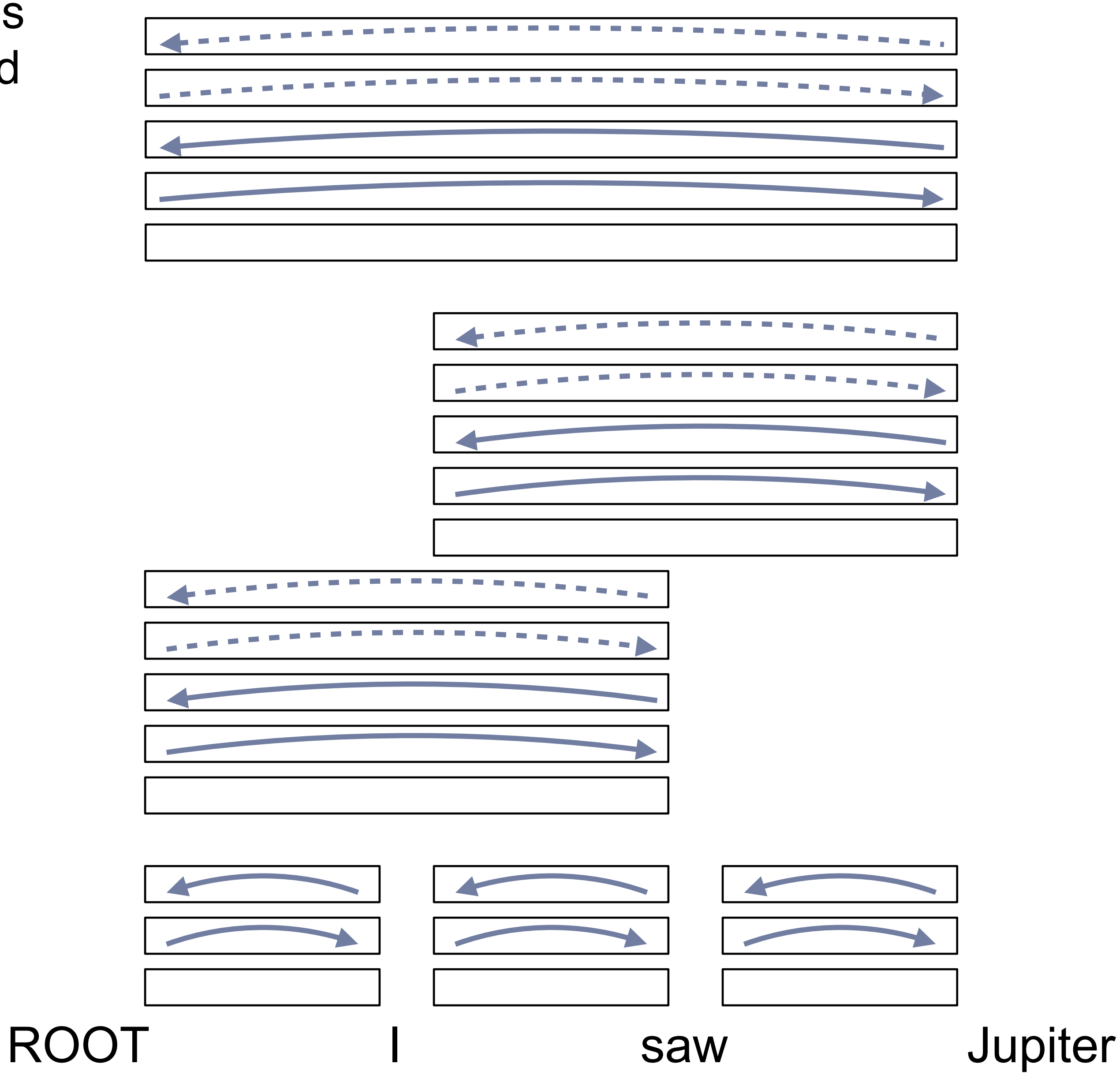
Example

menti.com 6274 6616

Our inference algorithm is defined by items and **rules**



We use the items and rules to build up the parse



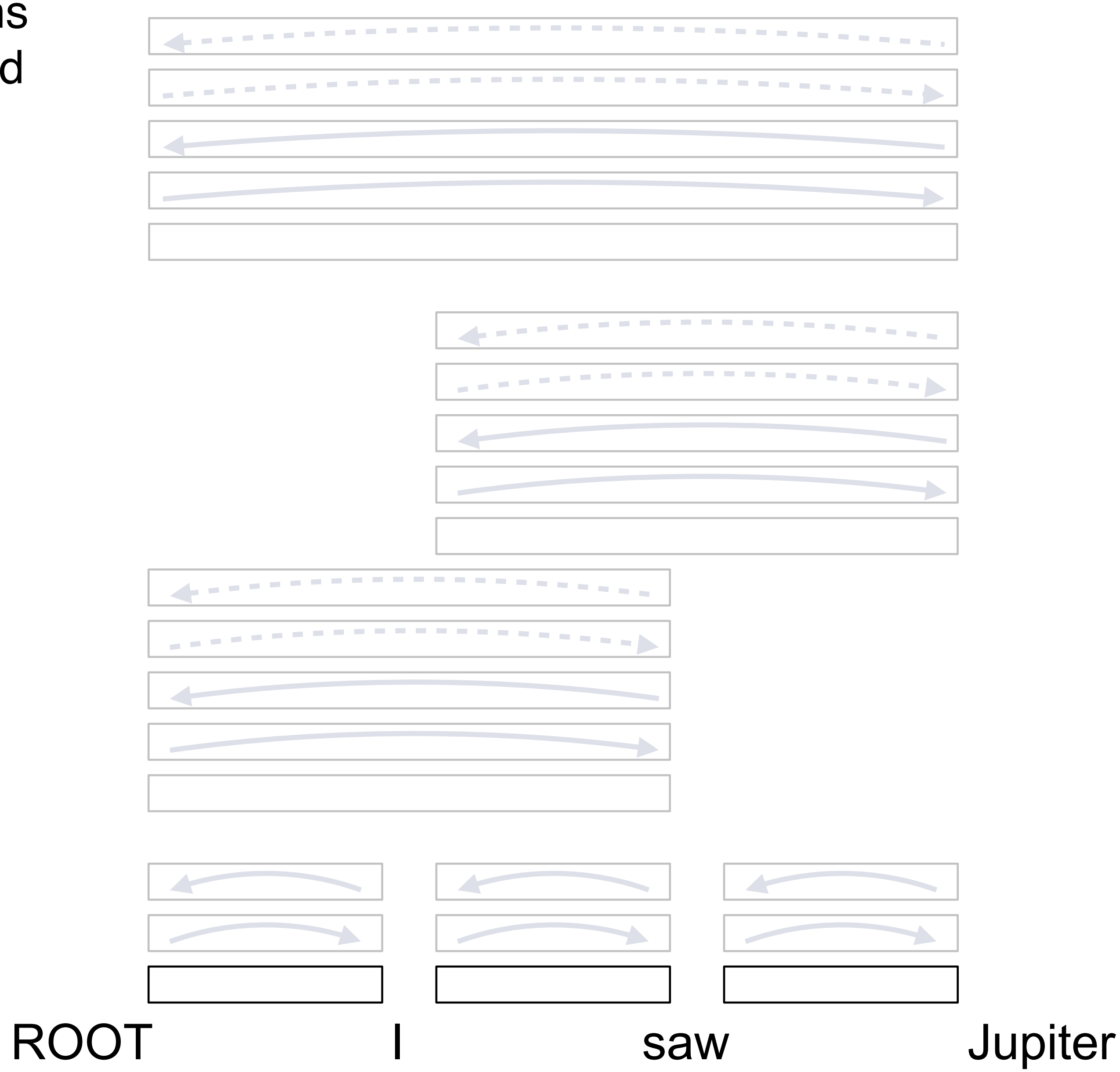
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Create starting items



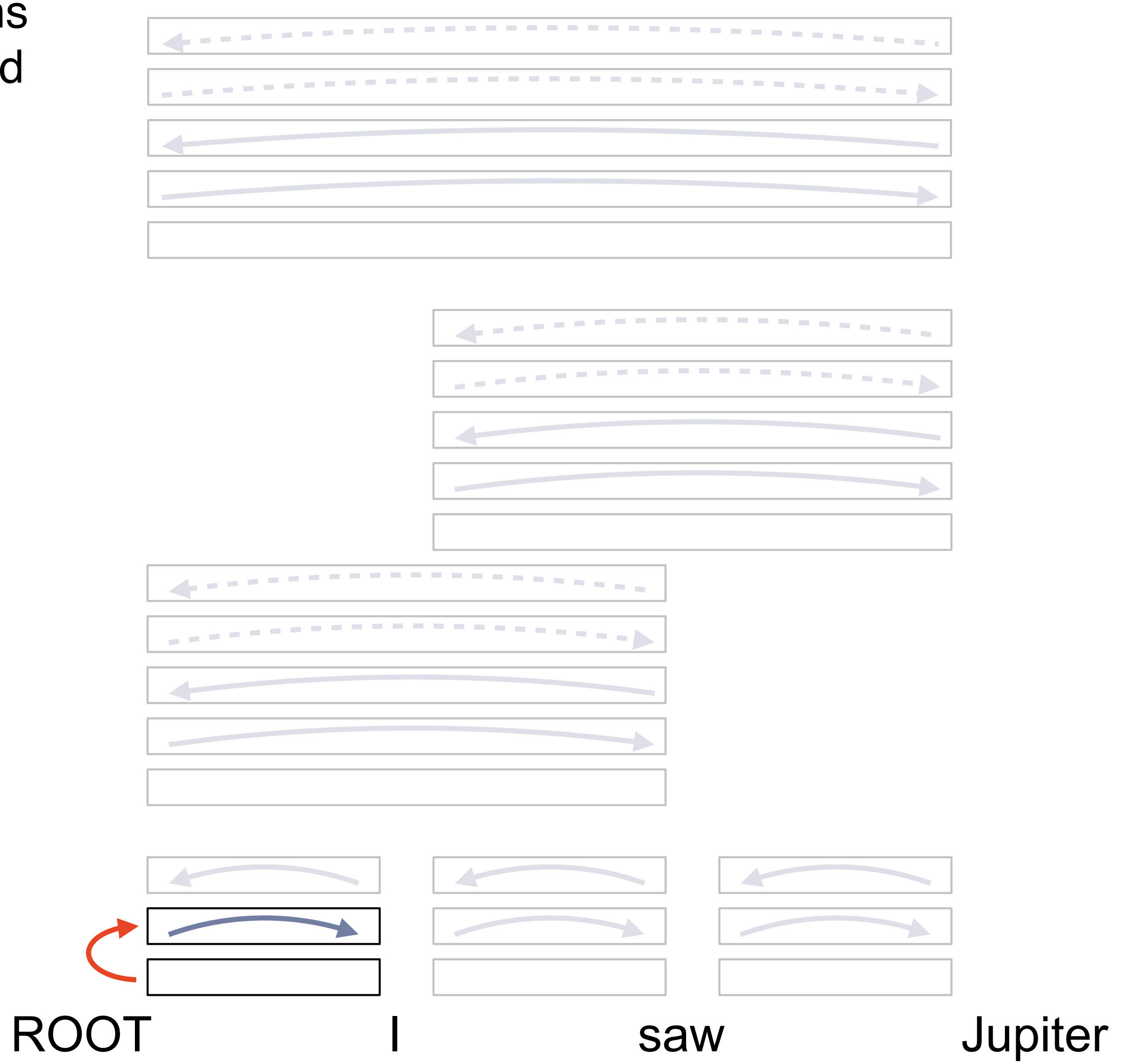
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

2!

Apply arc creation rules



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

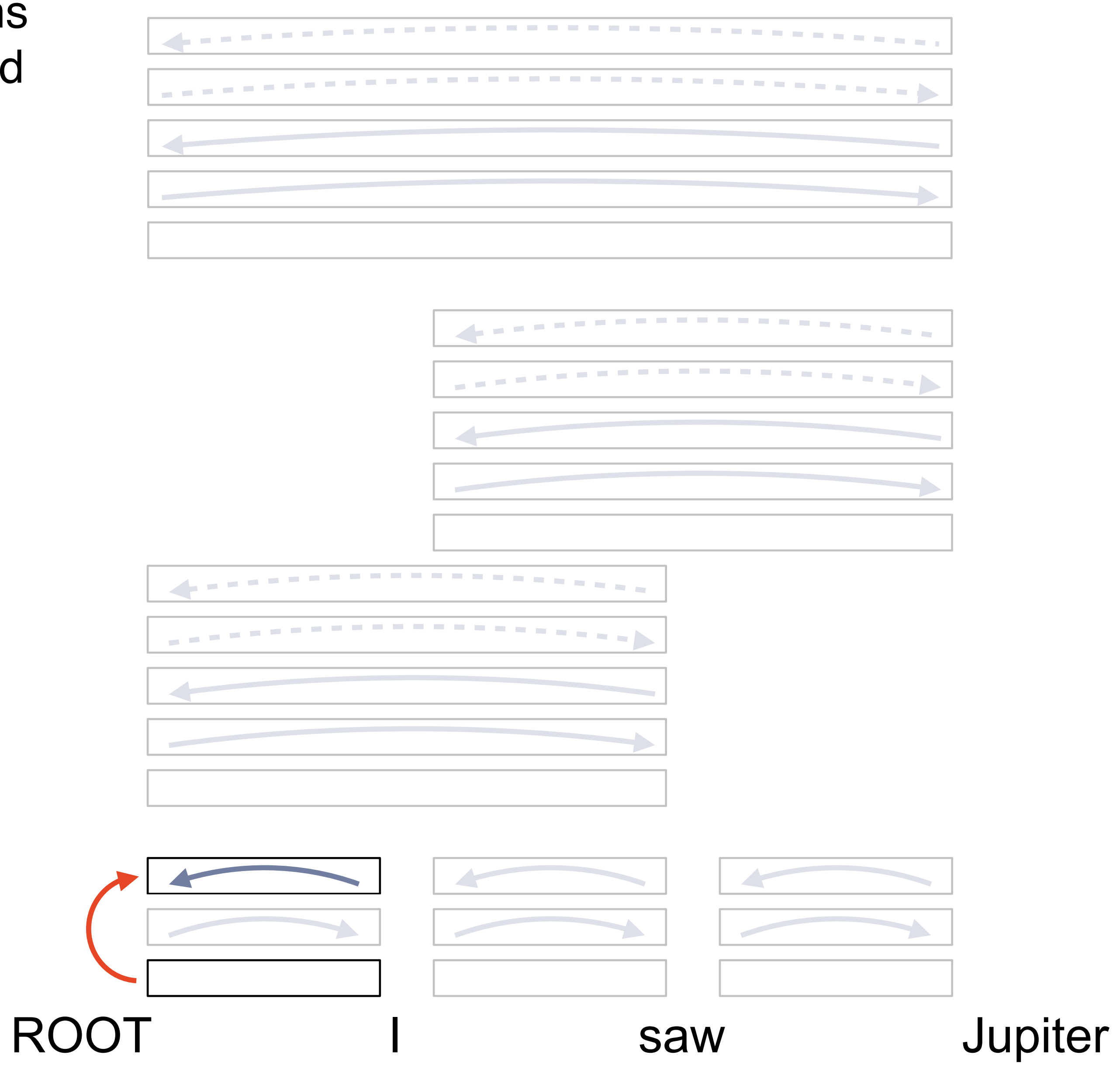
menti.com 6274 6616

We use the items and rules to build up the parse

Hiding

2.2

Apply arc creation rules



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

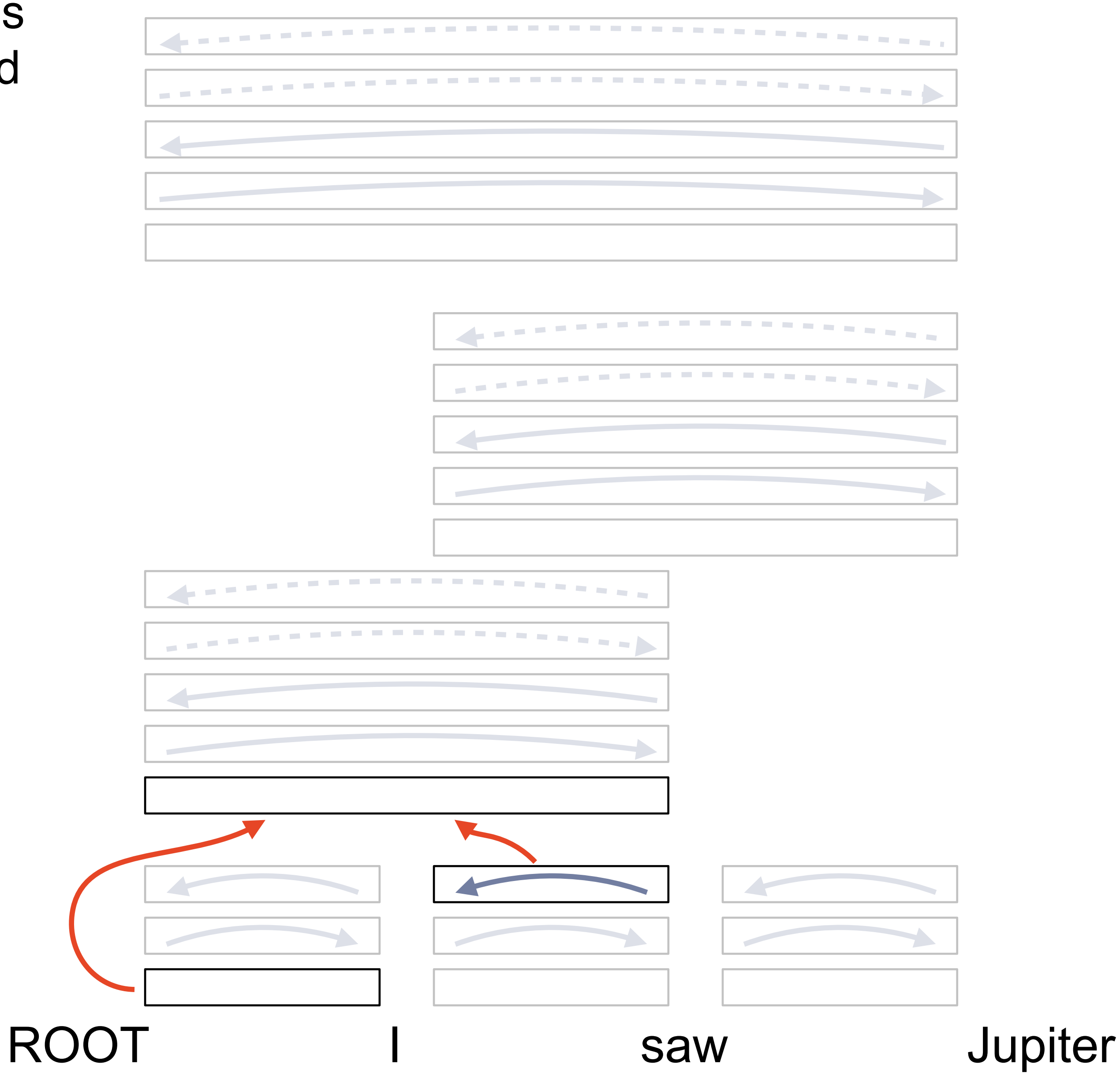
menti.com 6274 6616

We use the items and rules to build up the parse

Hiding

B

Apply combination rules



COMP 4446 / 5046
Lecture 5, 2025

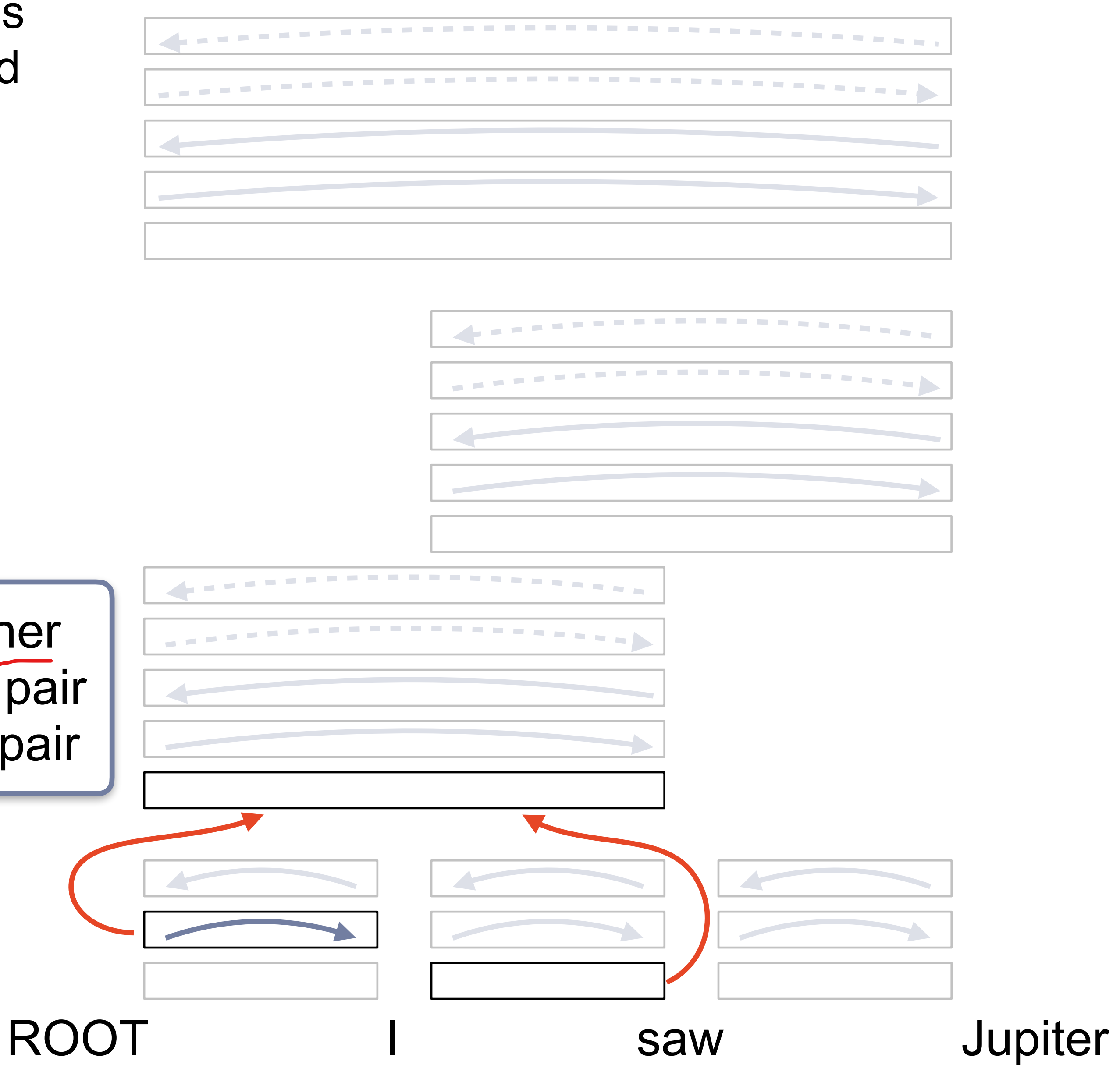
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Keep the higher scoring of this pair and the other pair

Apply combination rules

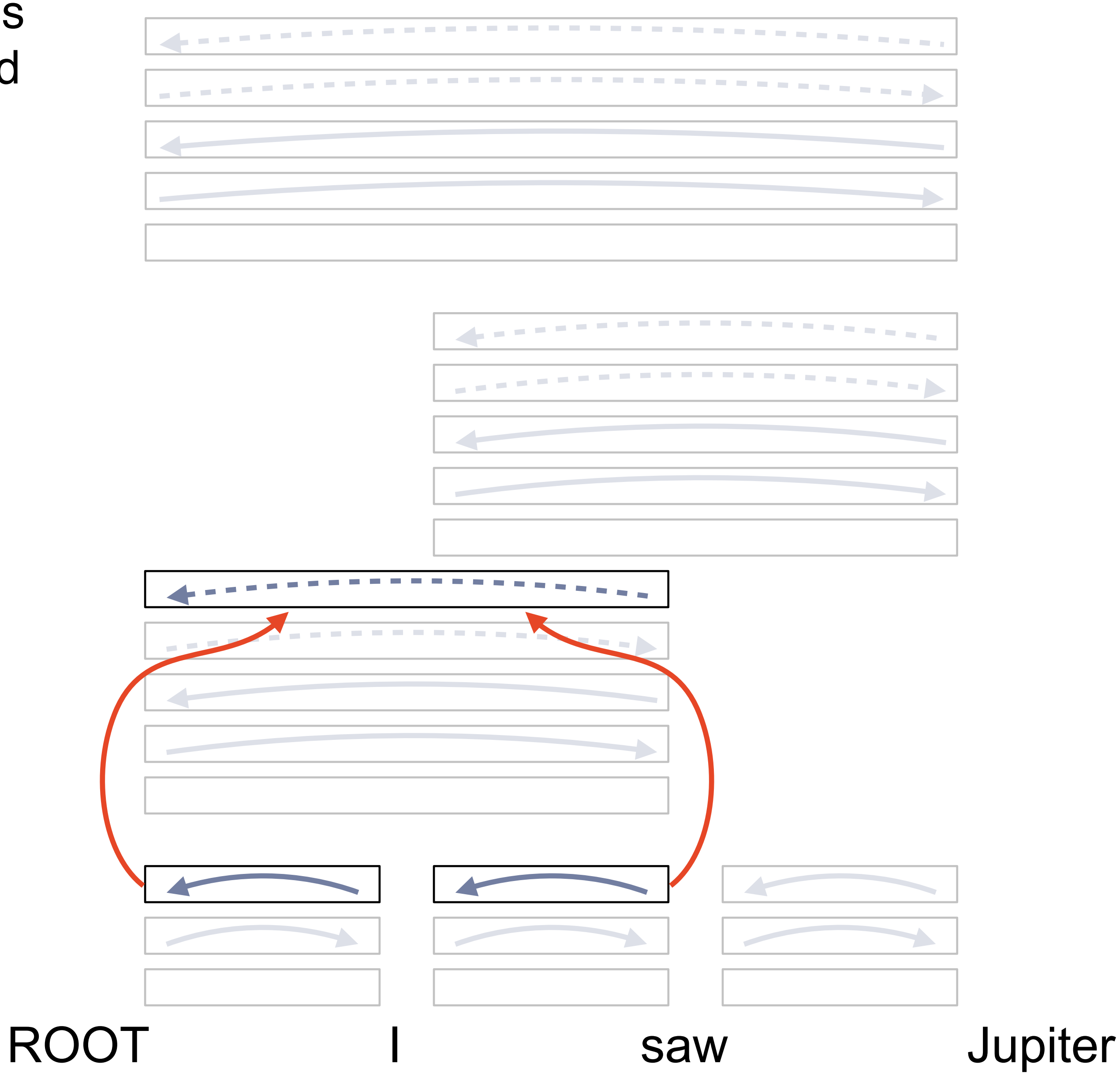


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse



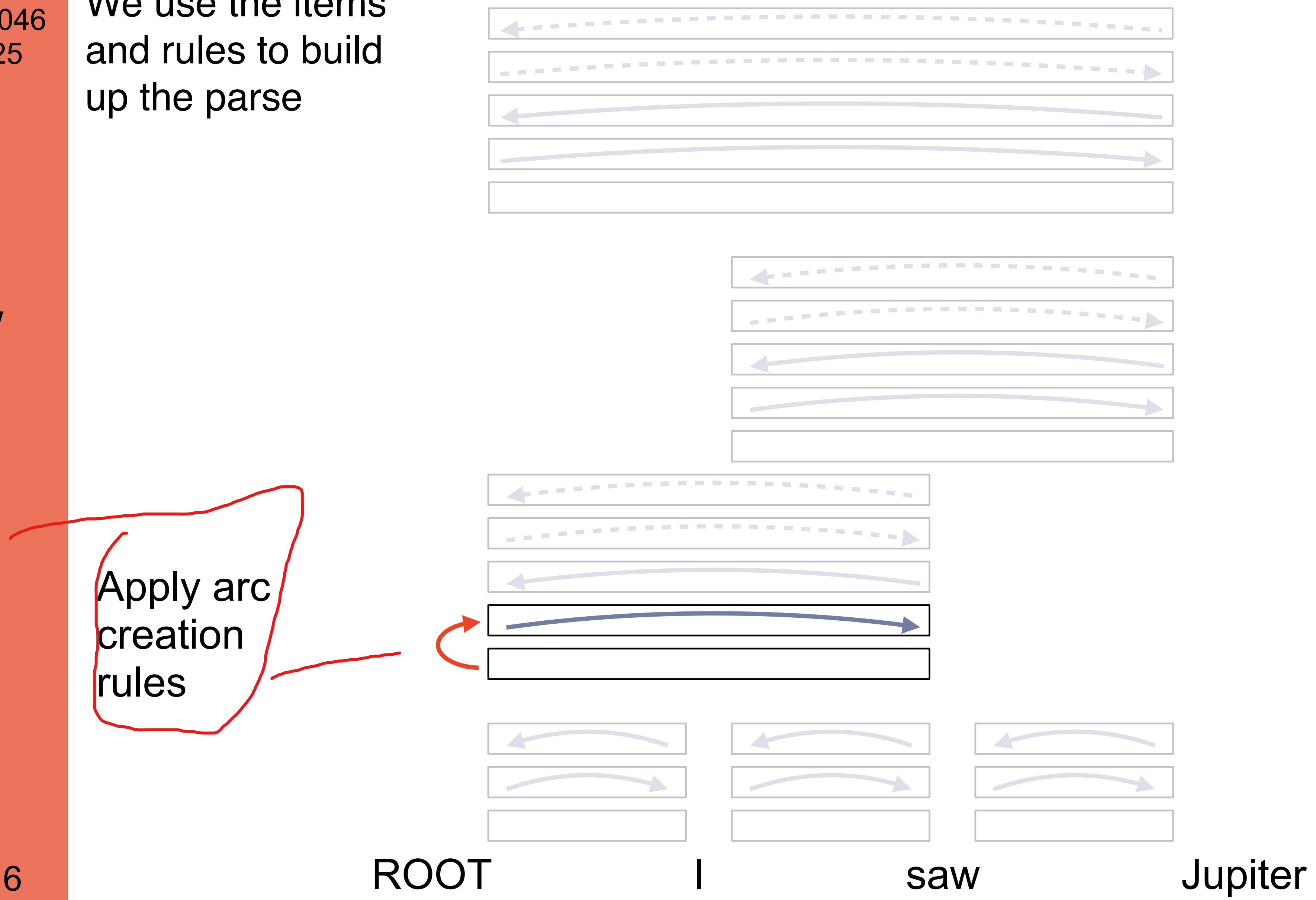
Apply combination rules

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Apply combination rules

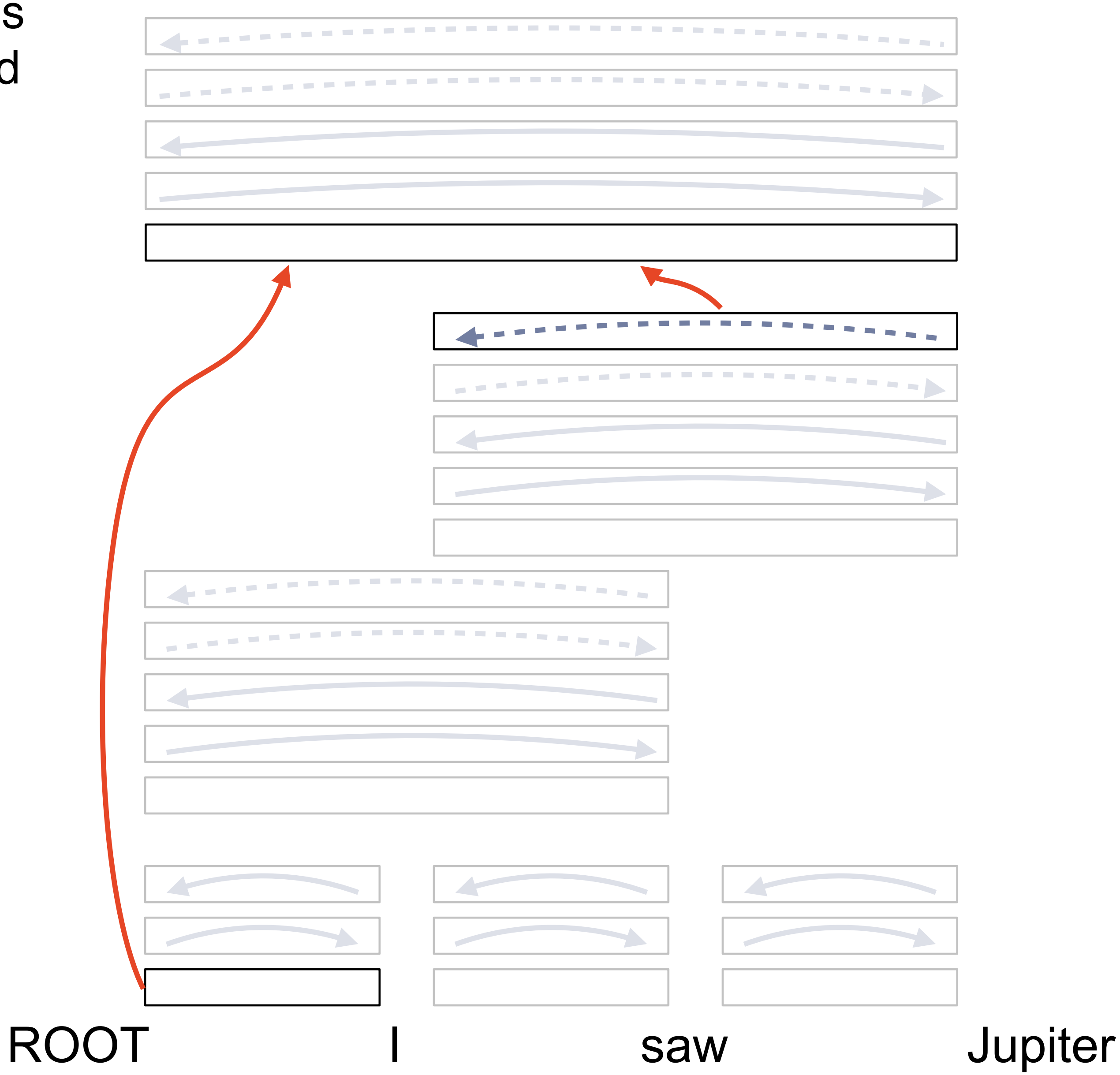
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Apply combination rules



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Apply combination rules

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Apply combination rules

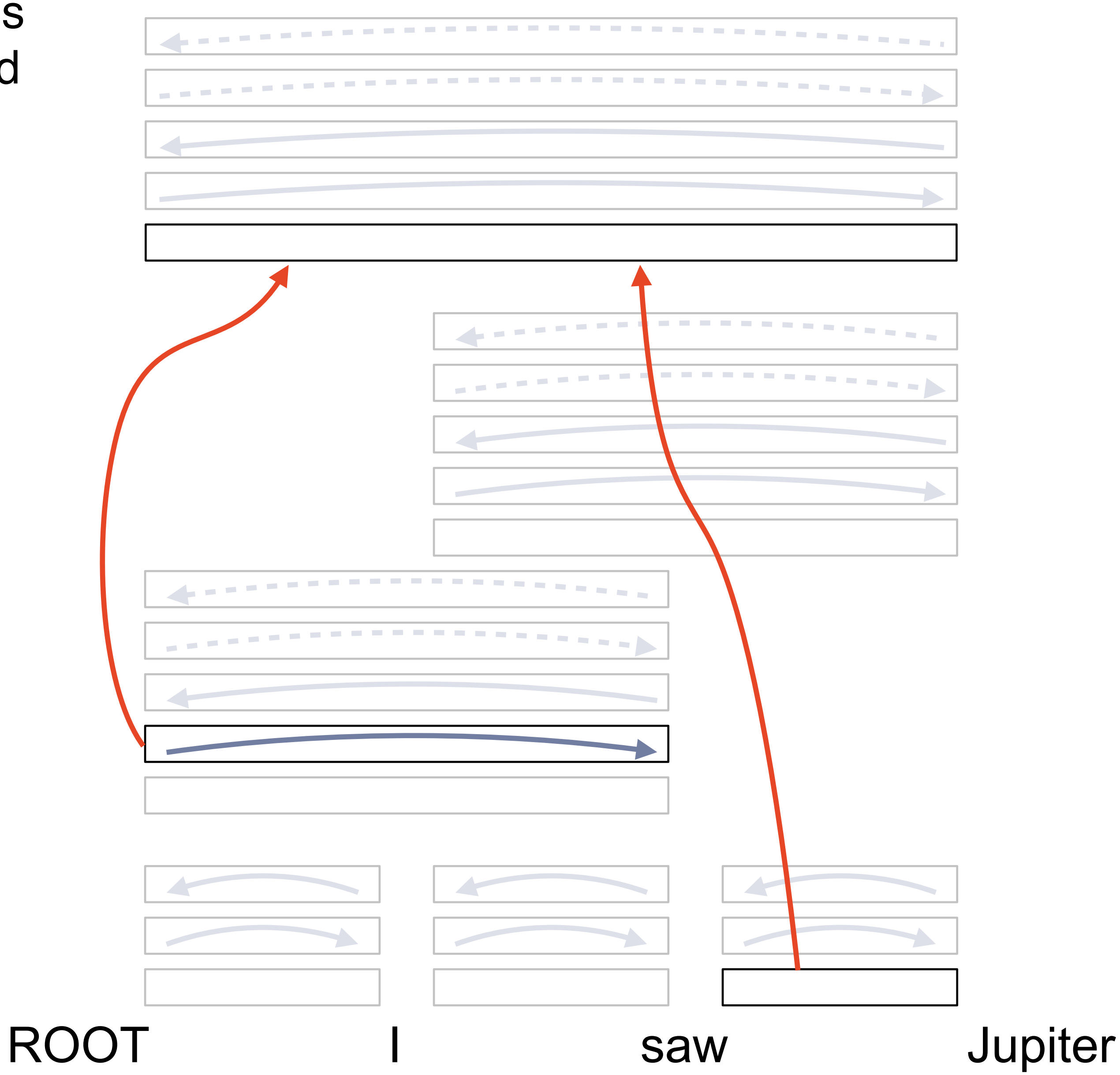
COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

We use the items and rules to build up the parse

Apply combination rules

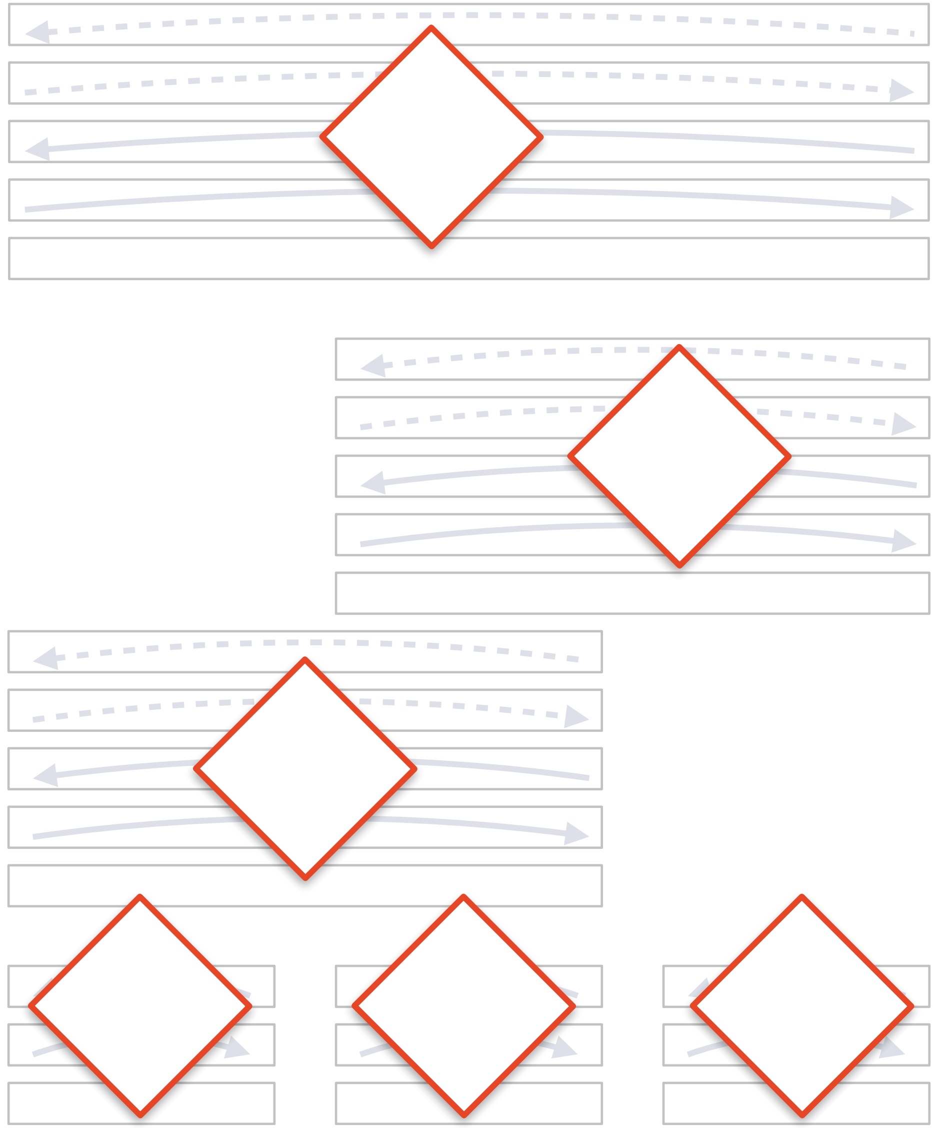


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

General idea: Consider each way to form each item, store the best and how you got it

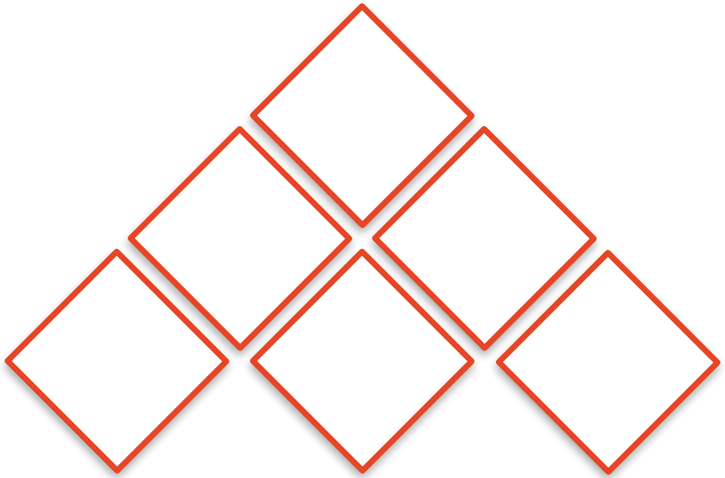


COMP 4446 / 5046
Lecture 5, 2025

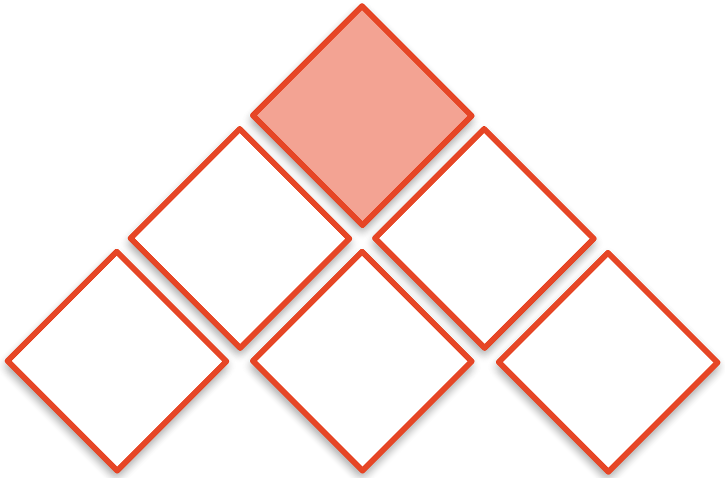
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

General idea: Consider each way to form each item, store the best and how you got it



General idea: Consider each way to form each item, store the best and how you got it

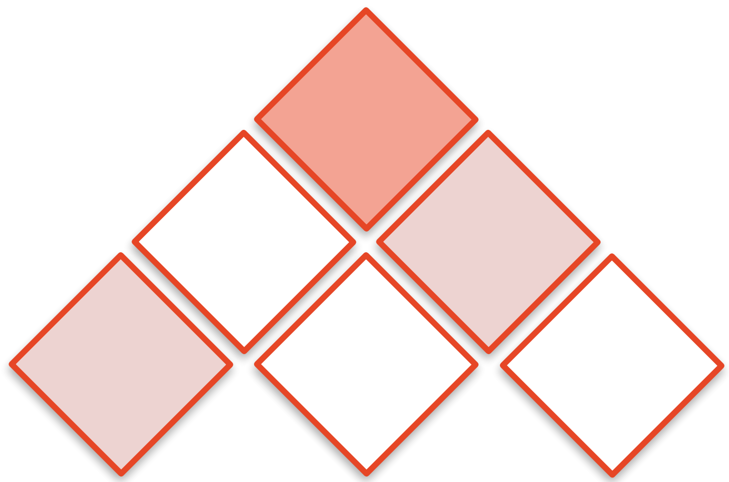


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

General idea: Consider each way to form each item, store the best and how you got it

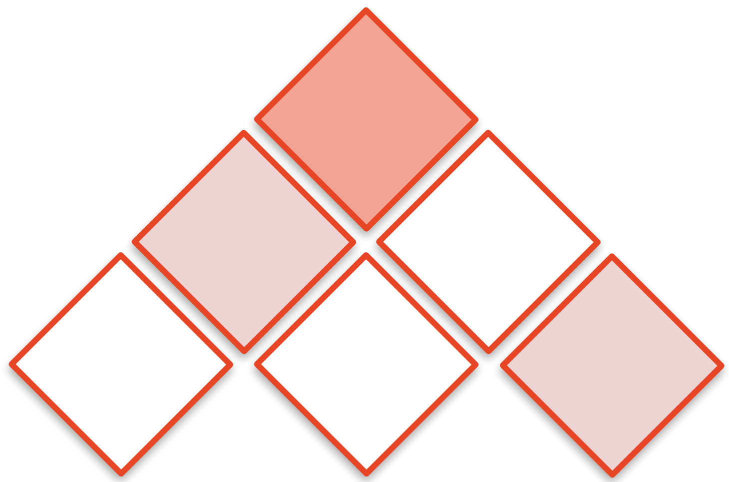


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

General idea: Consider each way to form each item, store the best and how you got it

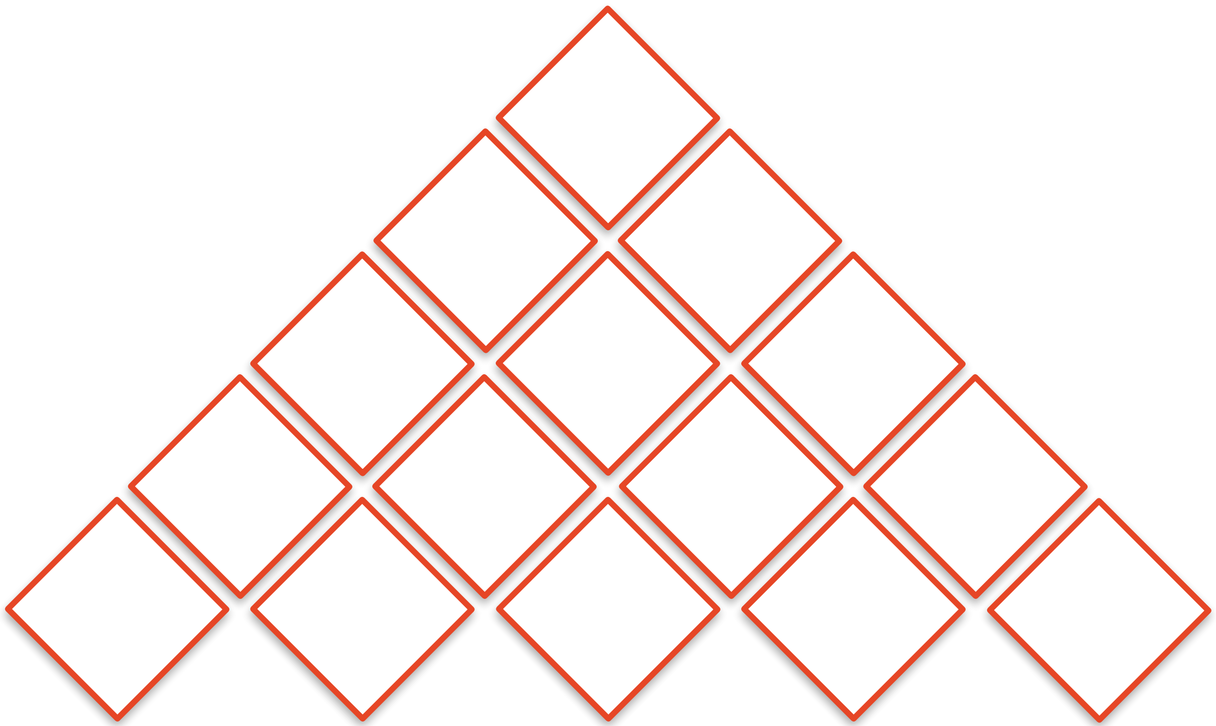


COMP 4446 / 5046
Lecture 5, 2025

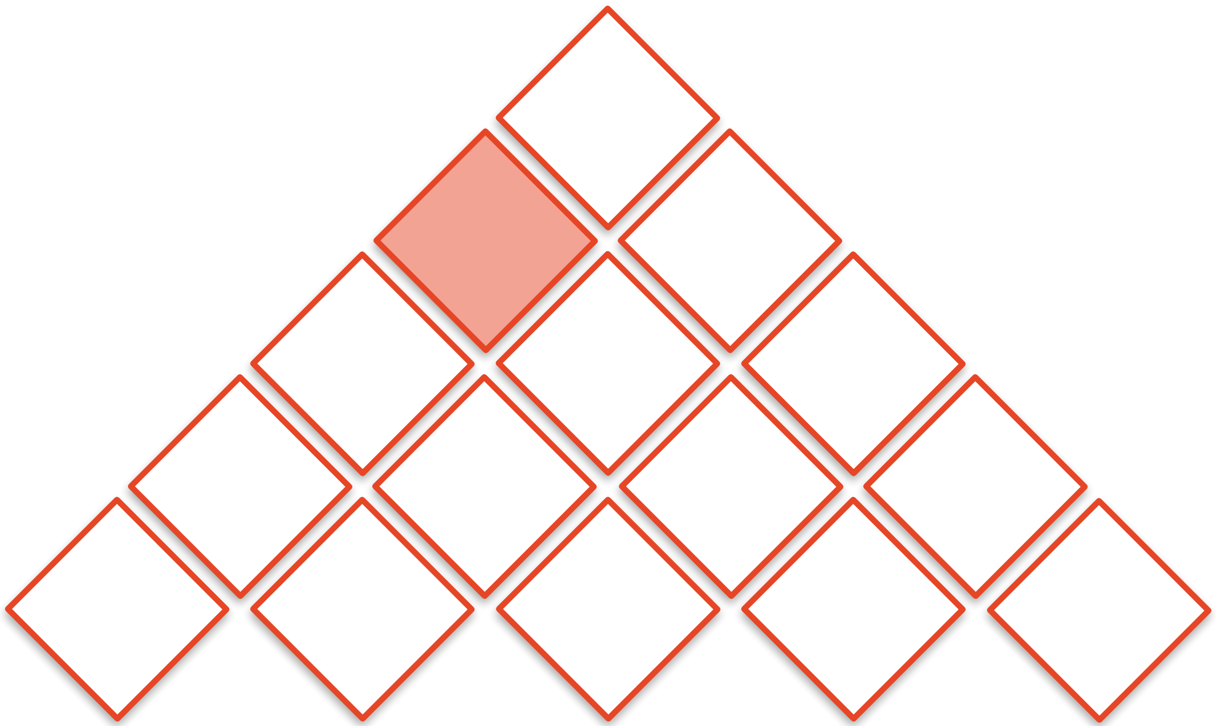
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

General idea: Consider each way to form each item, store the best and how you got it



General idea: Consider each way to form each item, store the best and how you got it

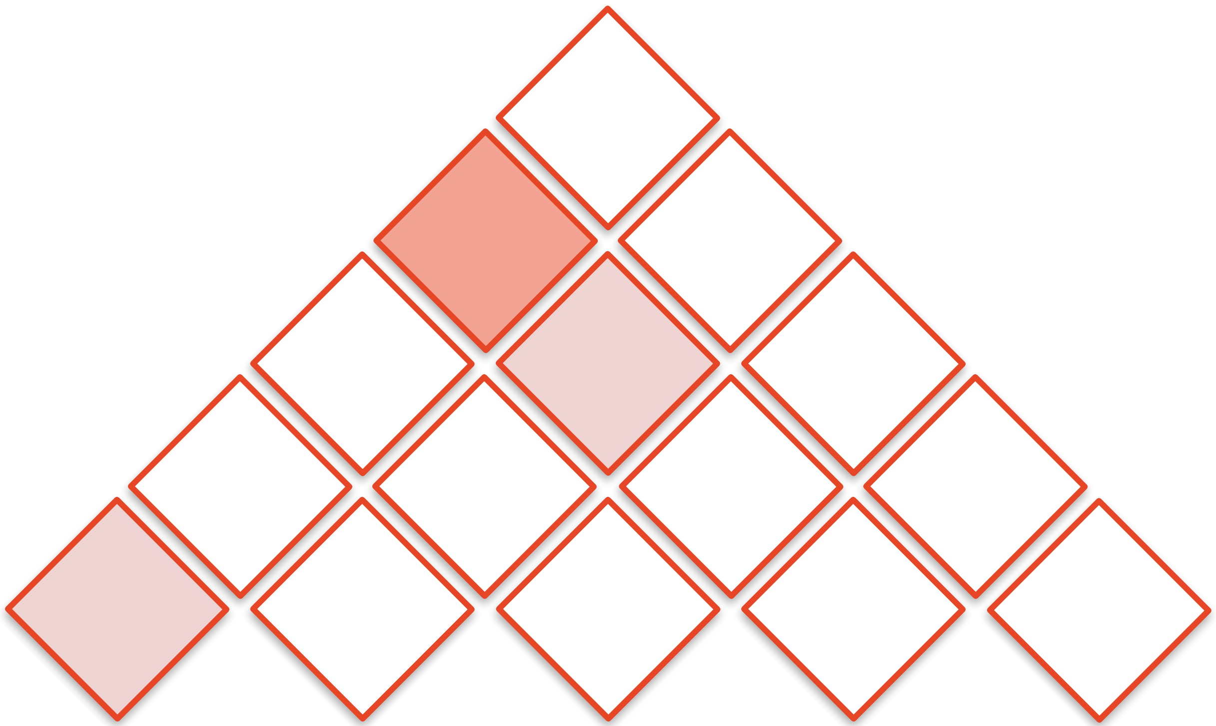


COMP 4446 / 5046
Lecture 5, 2025

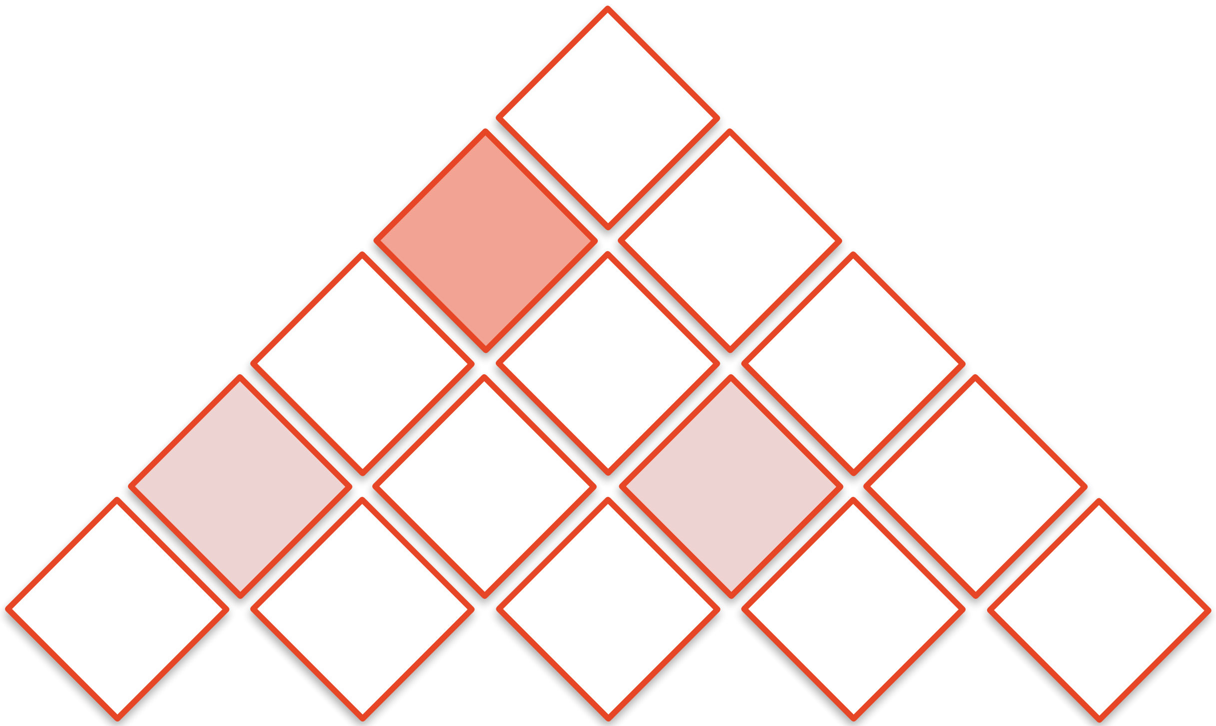
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

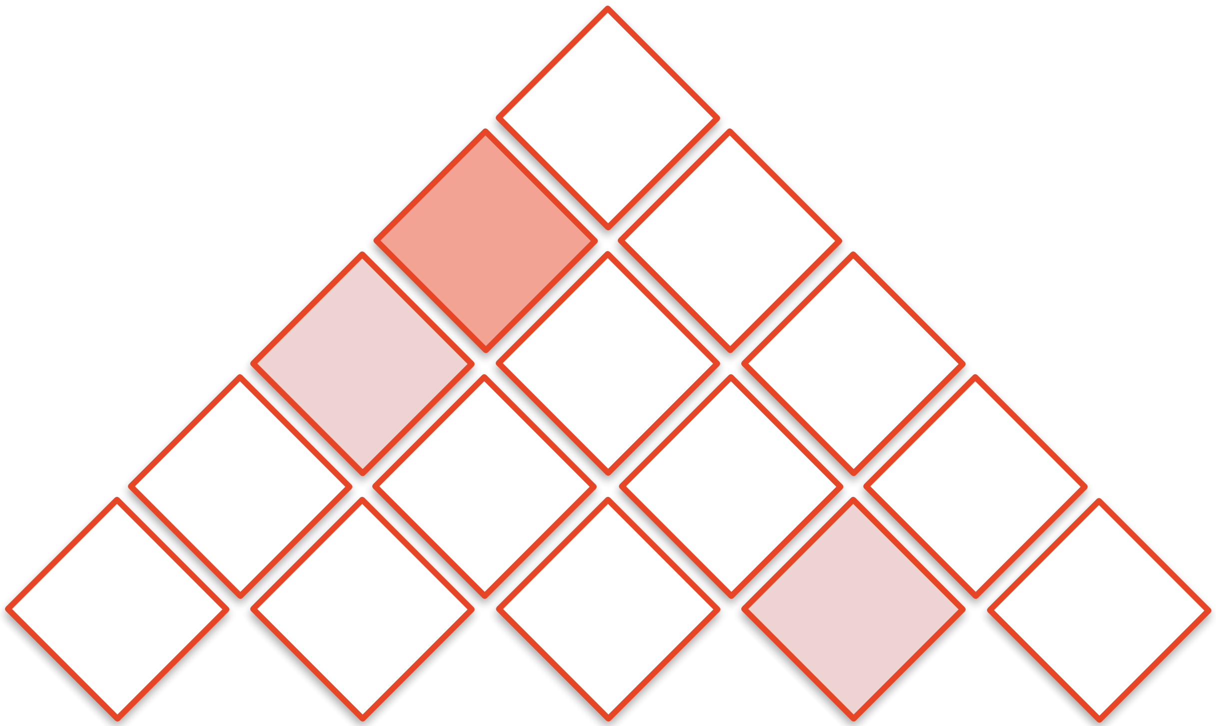
General idea: Consider each way to form each item, store the best and how you got it



General idea: Consider each way to form each item, store the best and how you got it

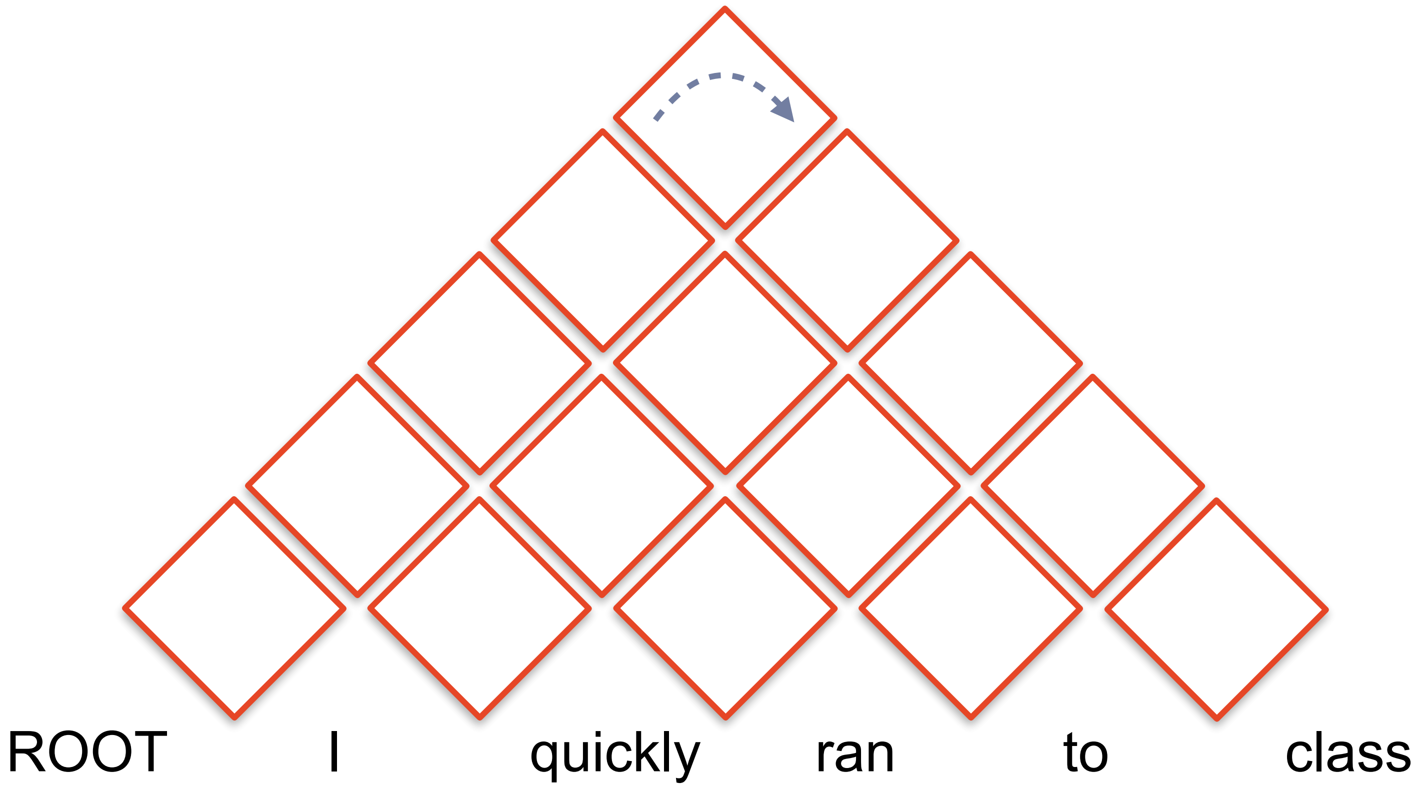


General idea: Consider each way to form each item, store the best and how you got it



menti.com 6274 6616

At the end, start at the top and work downwards to get the parse

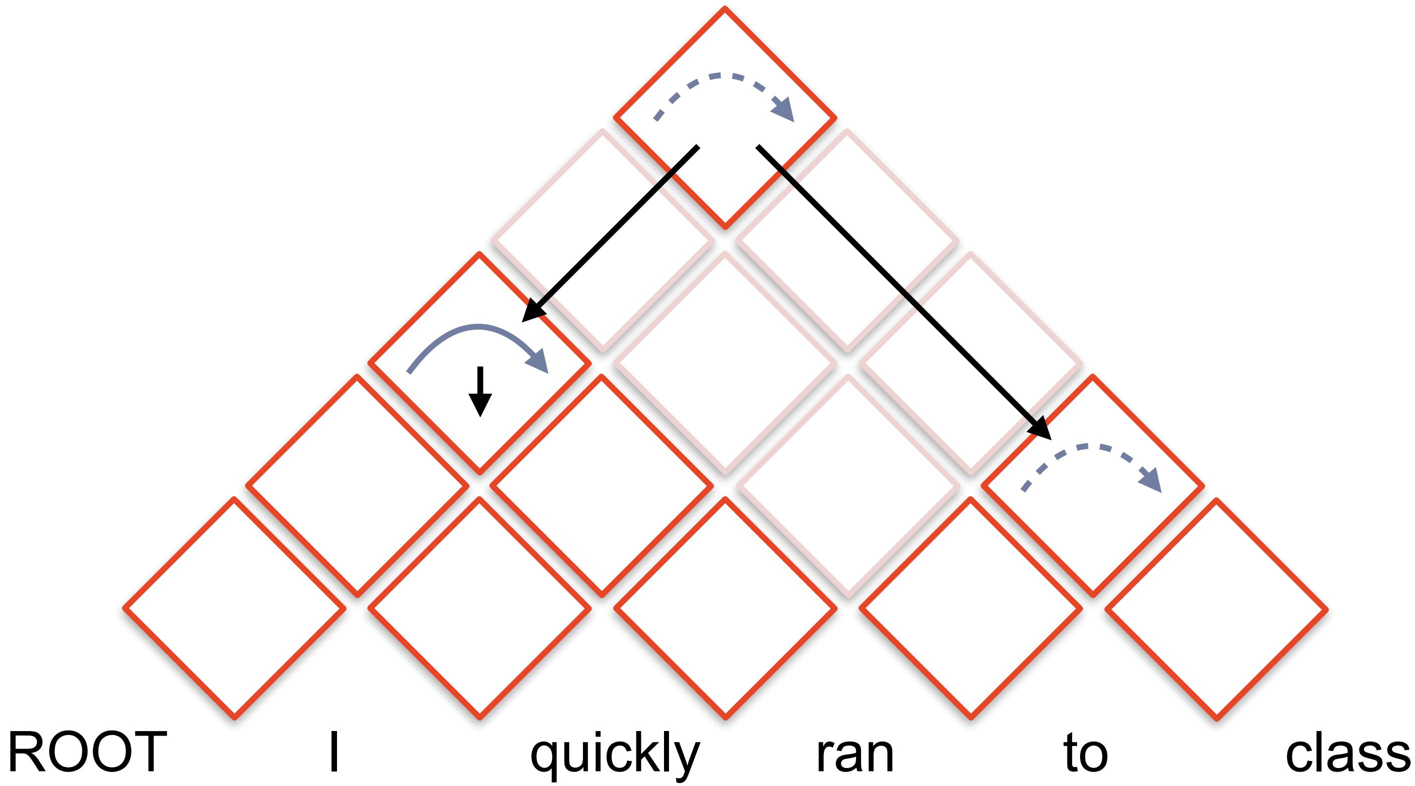


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

At the end, start at the top and work downwards to get the parse

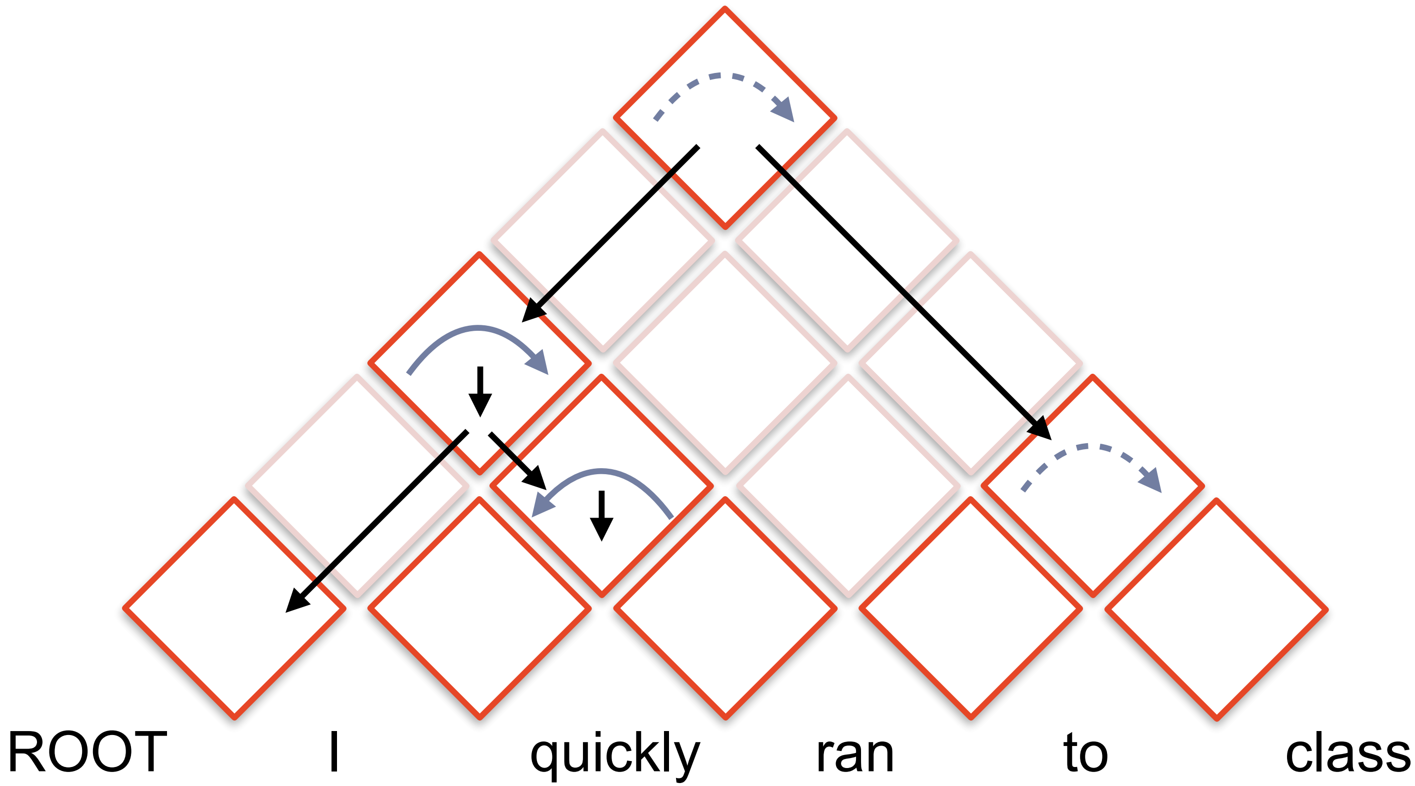


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

At the end, start at the top and work downwards to get the parse

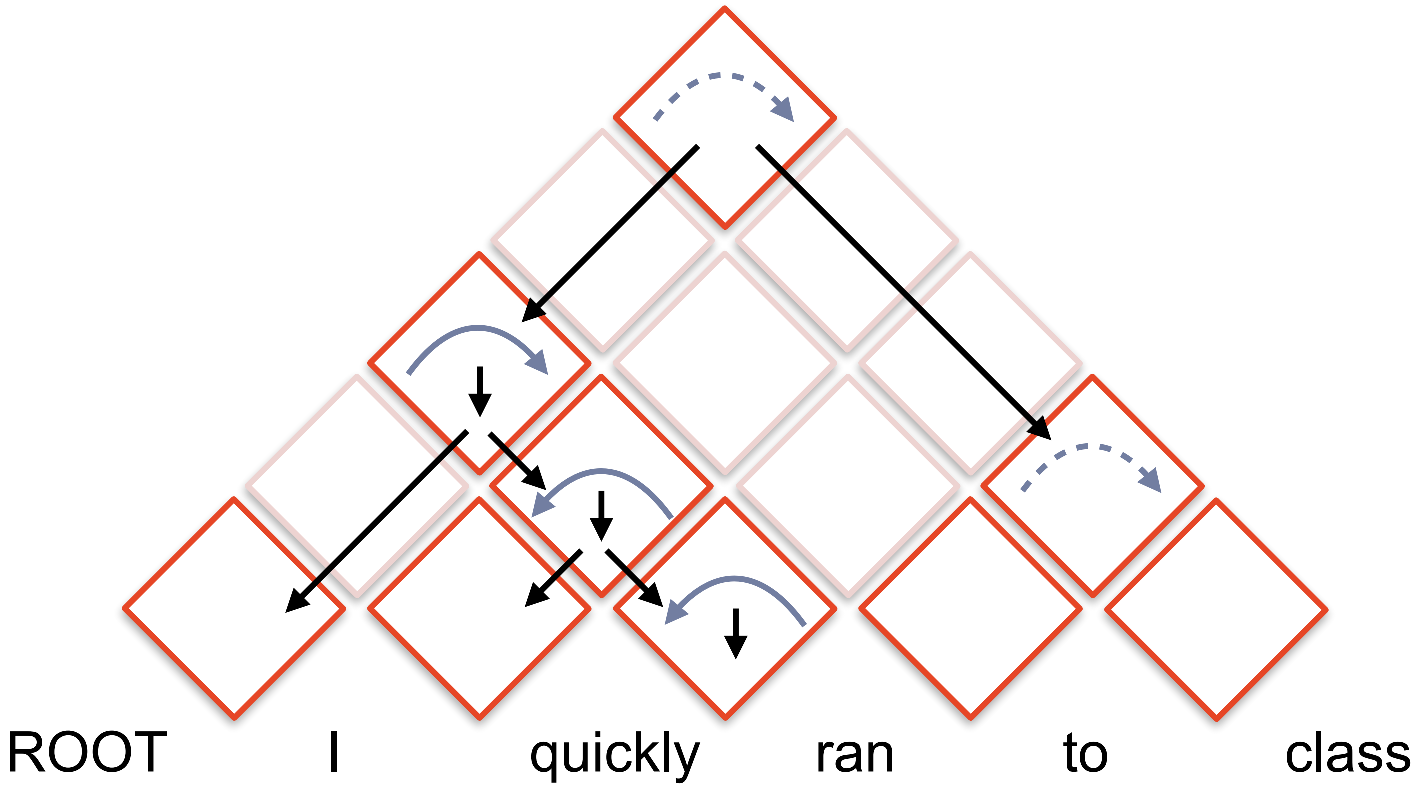


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

At the end, start at the top and work downwards to get the parse

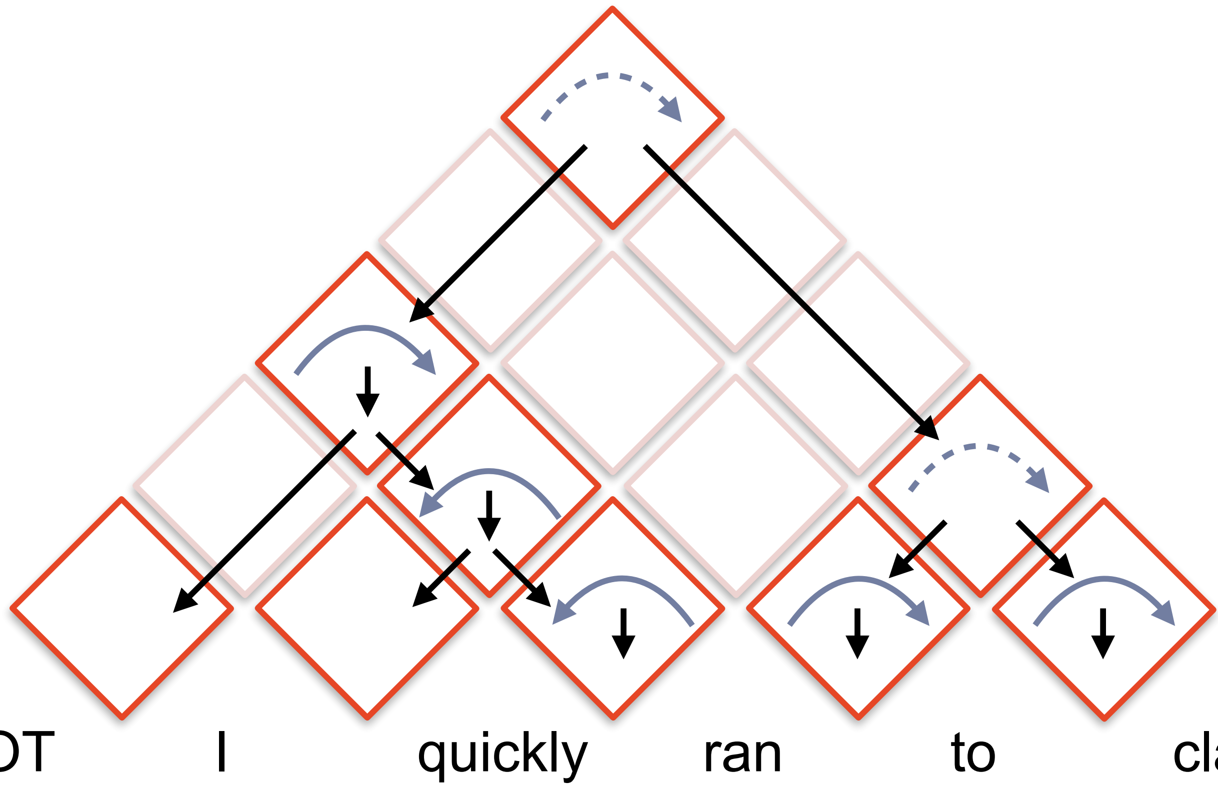


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

At the end, start at the top and work downwards to get the parse

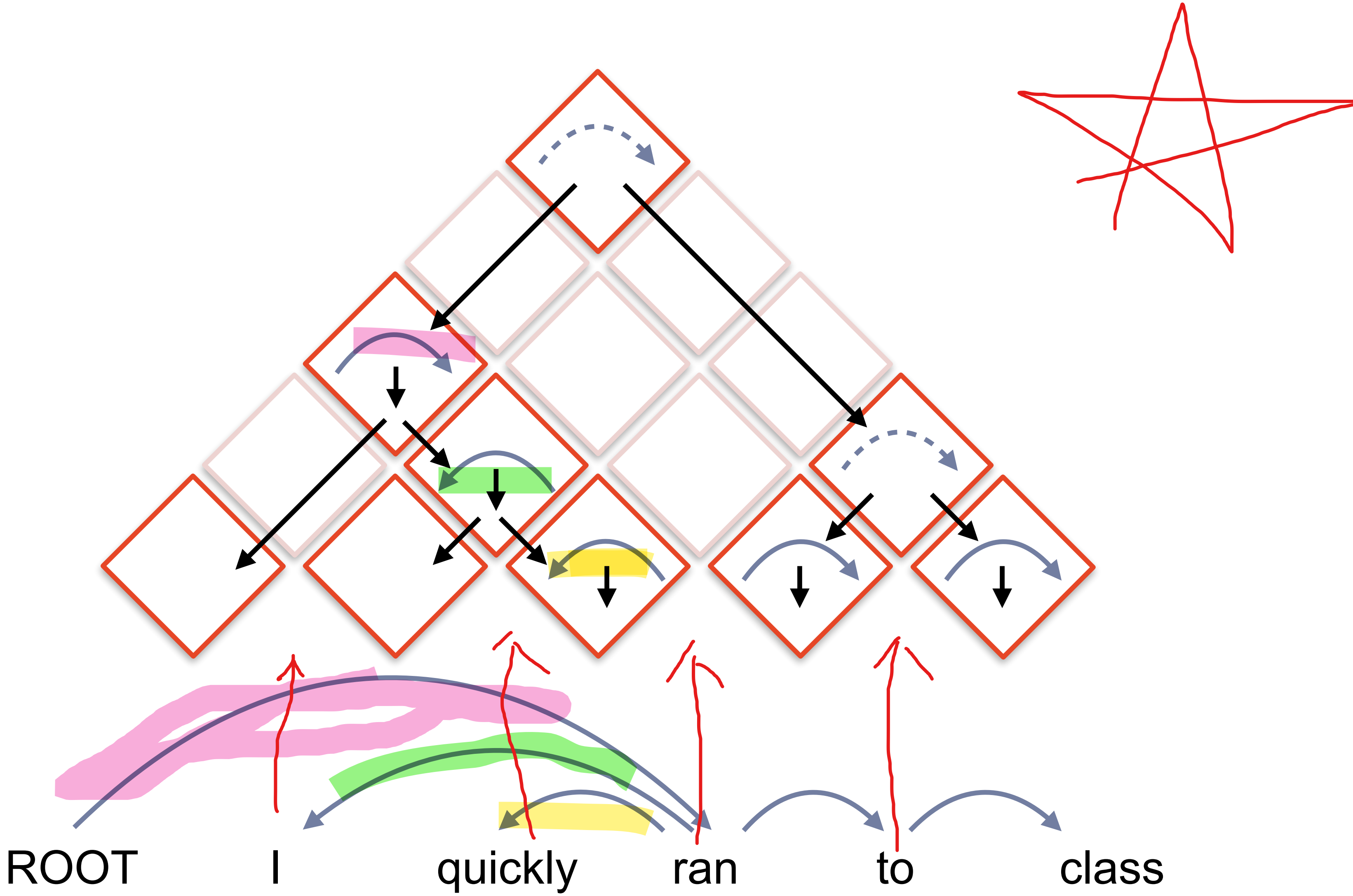


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

At the end, start at the top and work downwards to get the parse

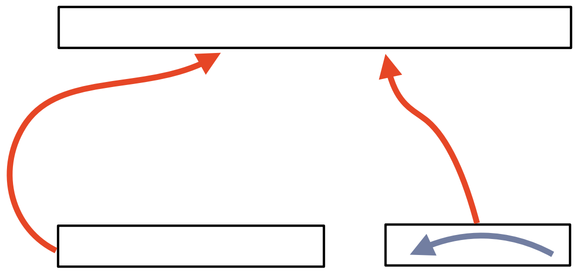


Algorithm in pseudocode

Called the CKY algorithm
..or CYK algorithm

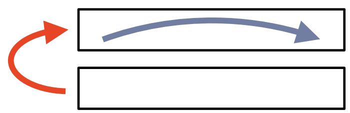
```
Create storage for best options for each span
and item type

Insert initial items
For each possible span length
    For each possible span start
        For each possible item type
            Best = None
            For each possible split point
                For each rule
                    Get the score for this
                    combination and update best
            Record the best
        Get the scores for adding an arc to the
        empty items and record them
```

Time complexity is $O(|rules_{comb}||words|^3 + |rules_{arc}||words|^2)$



Combinations:

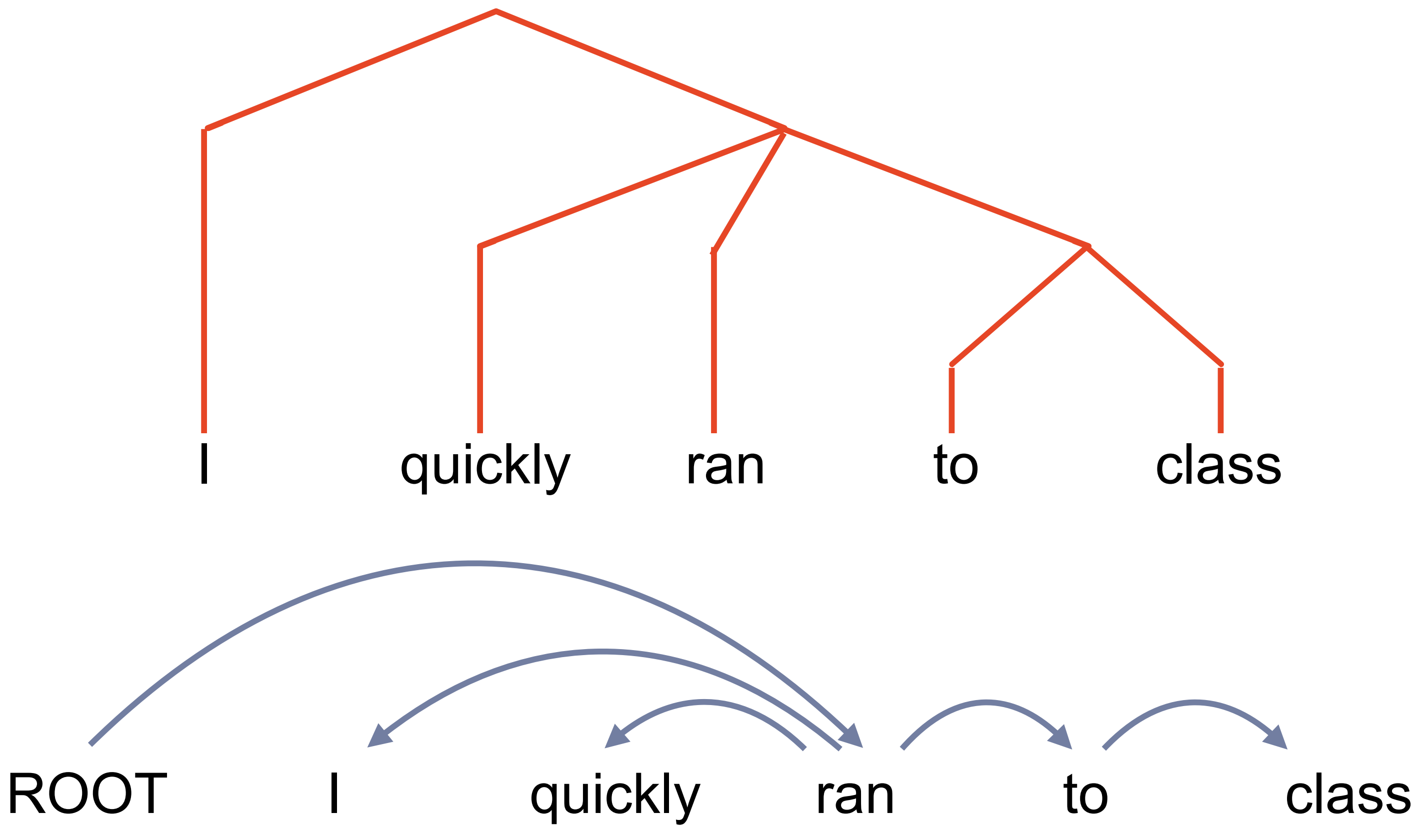
Each start point	$O(\|words\|)$
Each length	$O(\|words\|)$
Each split point	$O(\|words\|)$
Each rule	$O(\|rules_{comb}\|)$



Arc creation:

Each start point	$O(\|words\|)$
Each length	$O(\|words\|)$
Each rule	$O(\|rules_{arc}\|)$

Dependencies are one representation, constituency parses are another

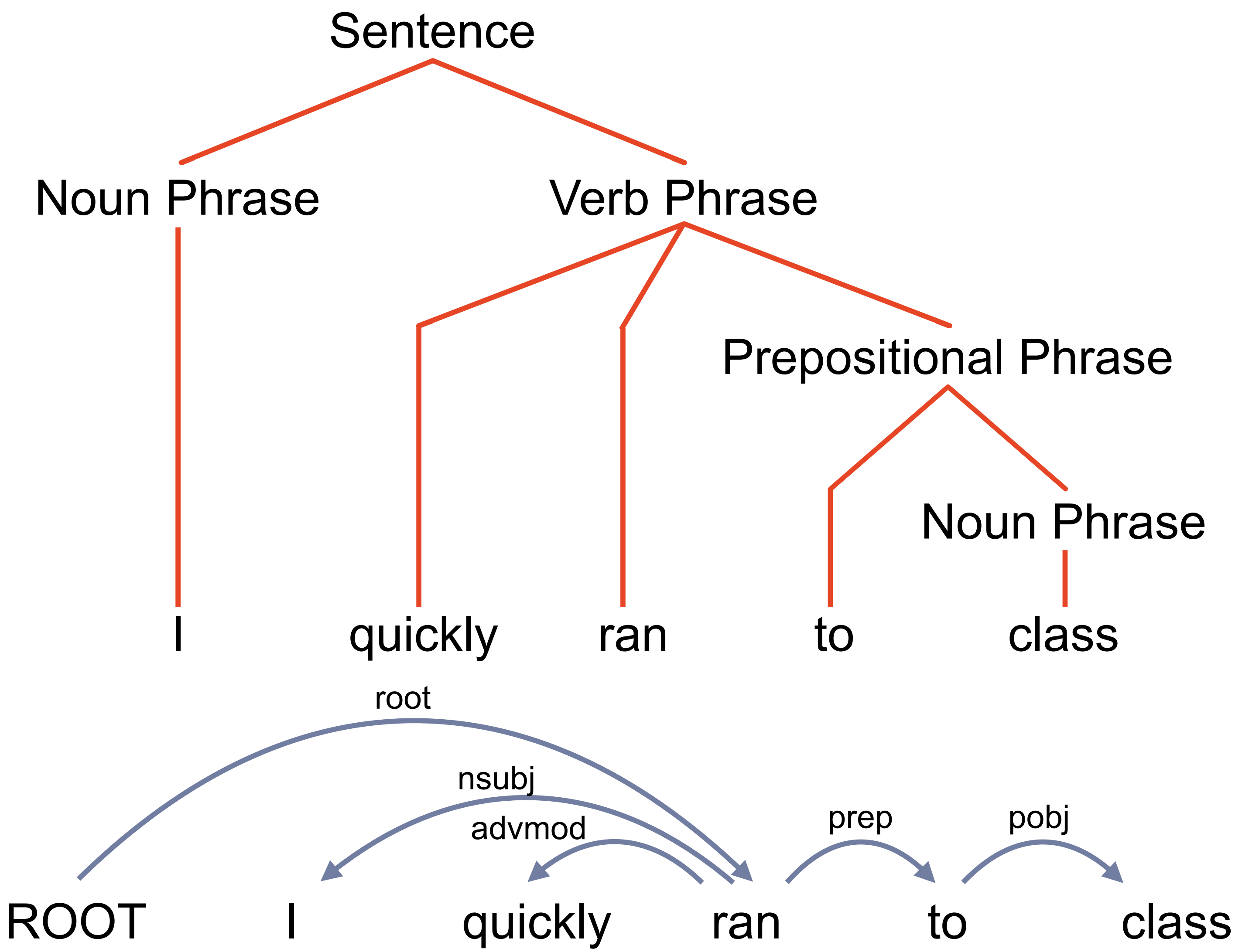


COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Dependencies are one representation, constituency parses are another



COMP 4446 / 5046
Lecture 5, 2025

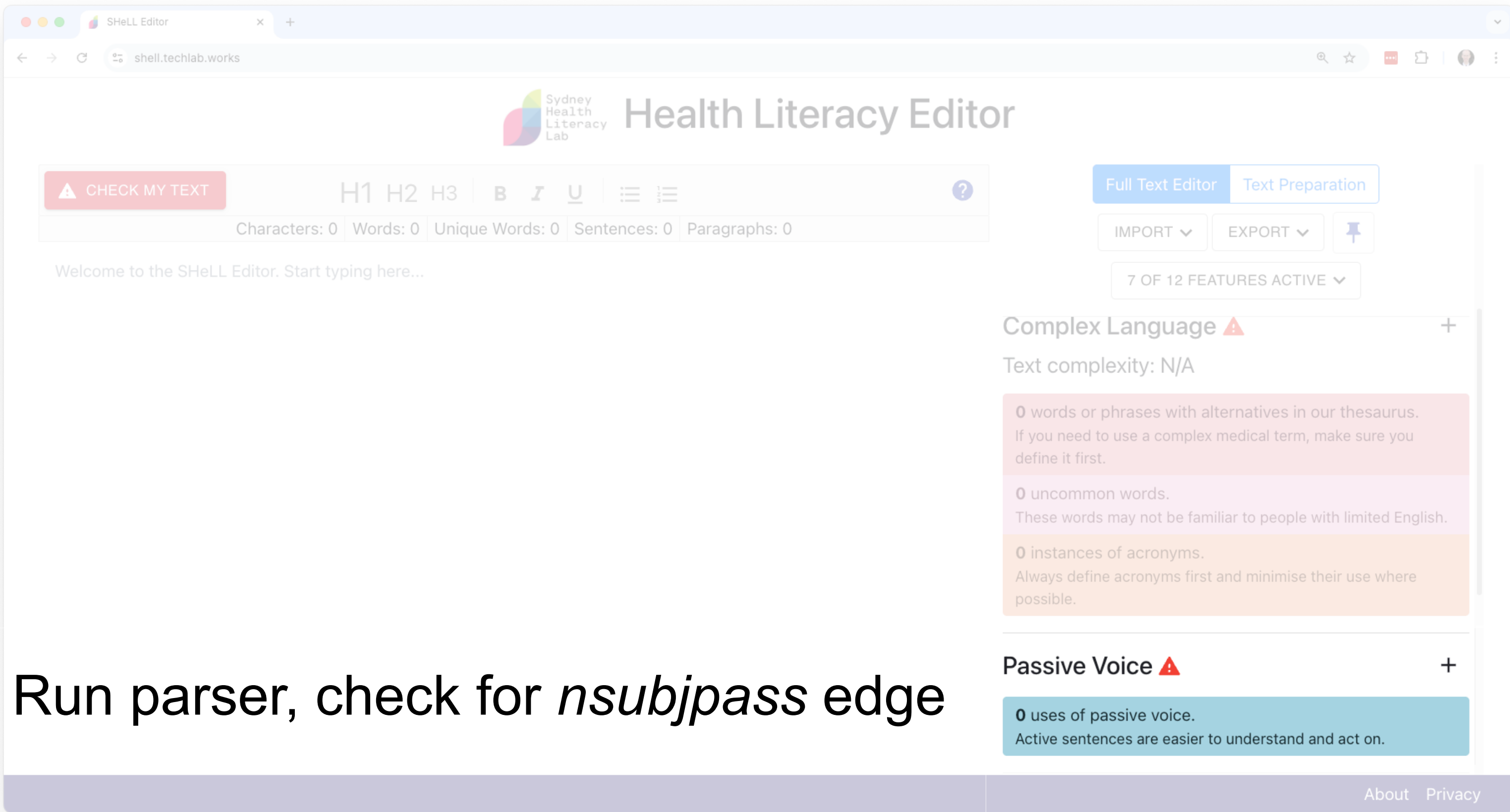
Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Why is this still relevant in the age of LLMs?

Usage

1. Occasionally, this is what you need



Run parser, check for *nsubjpass* edge

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Why is this still relevant in the age of LLMs?

2. You can use parsing as part of a system

We propose a new representation and algorithm for a class of graph structures that is flexible enough to cover almost all treebank structures, while still admitting efficient learning and inference

Why is this still relevant in the age of LLMs?

2. You can use parsing as part of a system

We propose a new representation and algorithm for a class of graph structures that is flexible enough to cover almost all treebank structures, while still admitting efficient learning and inference.

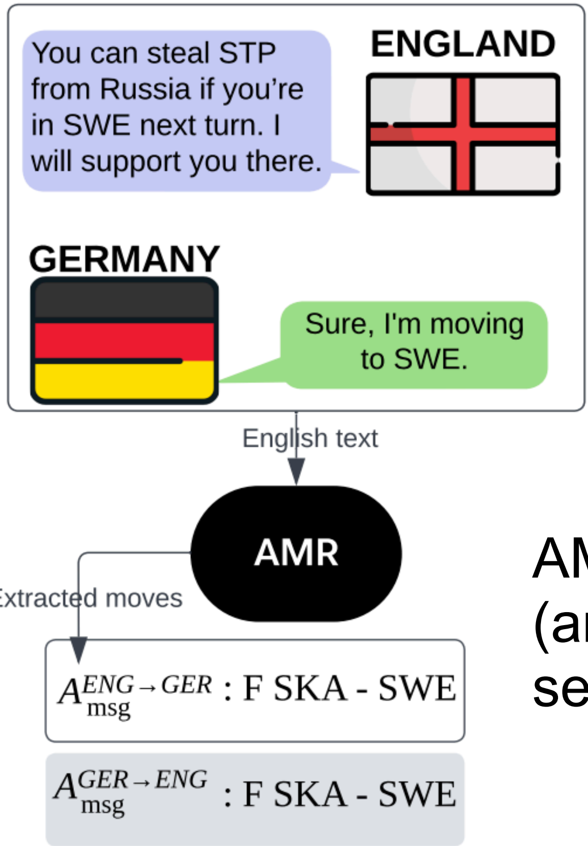
Limit parts of text to query LLM with based on the parse, e.g., don't cut in the middle of an arc

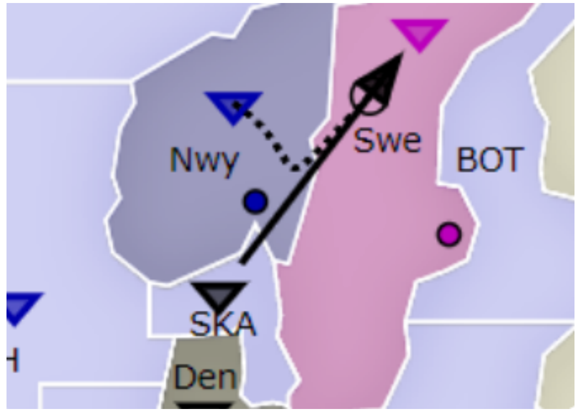
menti.com 6274 6616

menti.com 6274 6616

Why is this still relevant in the age of LLMs?

3. LLMs are still bad at some things



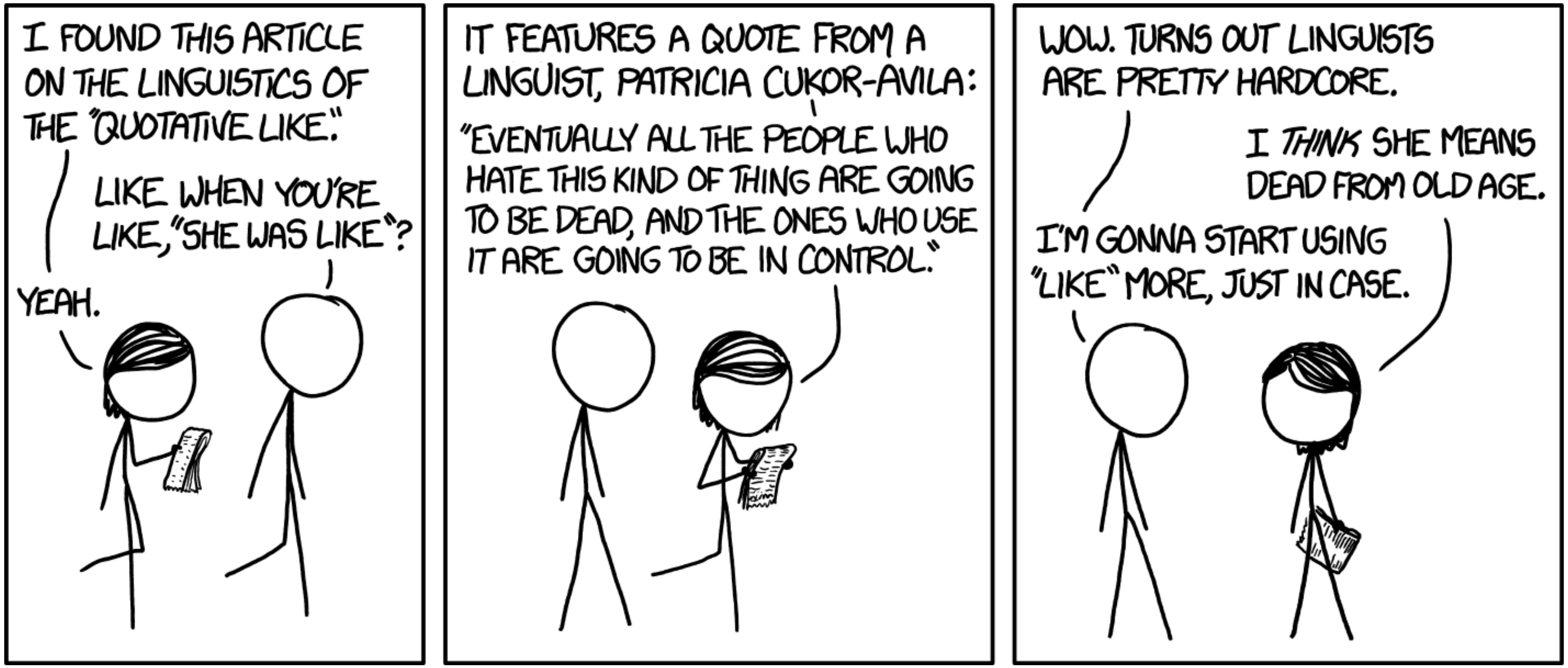


AMR = Abstract Meaning Representation (another type of parsing, focused on semantics rather than syntax)

3 minute Break - stretch and visit Menti

menti.com
6274 6616

Quotative Like



[God was like, "Let there be light," and there was light.]

Source: https://xkcd.com/1483/

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Coreference

References are when entities are mentioned in text

<ins>Victoria Chen</ins>, CFO of <ins>Megabucks Banking</ins>, saw <ins>her pay</ins> jump to $2.3 million, as the 38-year-old became the company's president. It is widely known that she came to Megabucks from rival <ins>Lotsabucks</ins>.

Example text is from the J+M textbook, chapter 23

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Coreference is identifying which mentions refer to the same entity

Victoria Chen

Megabucks Banking

Lotsabucks

her pay

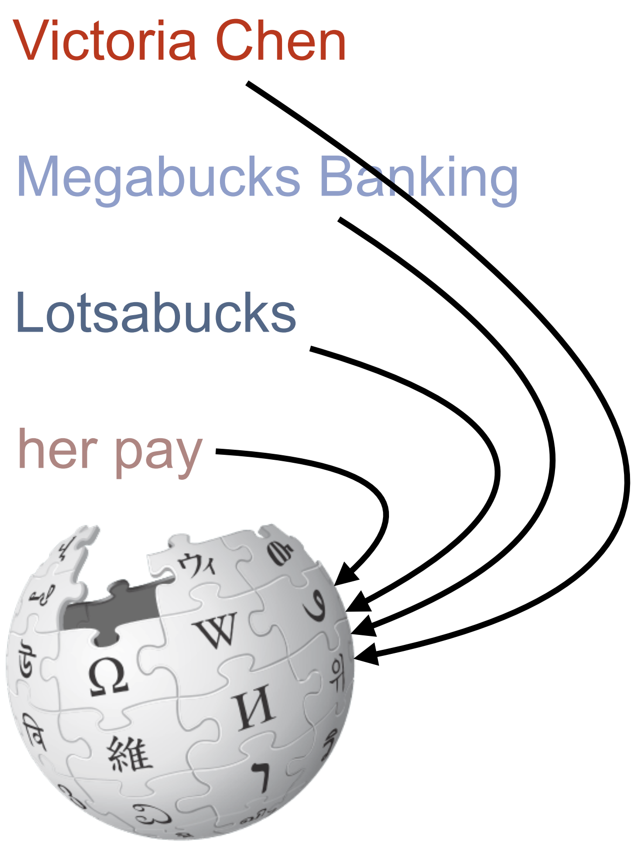
Victoria Chen, CFO of Megabucks Banking, saw her pay jump to $2.3 million, as the 38-year-old became the company's president. It is widely known that she came to Megabucks from rival Lotsabucks.

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Entity Linking / Wikification is a related task that links each entity to an external knowledge base



Victoria Chen, CFO of Megabucks Banking, saw her pay jump to $2.3 million, as the 38-year-old became the company’s president. It is widely known that she came to Megabucks from rival Lotsabucks.

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

The search space is huge!

Possible references:
|words| * |words|

Possible sets:
All ways to cluster
all references

Victoria Chen, CFO of Megabucks Banking, saw her pay jump to $2.3 million, as the 38-year-old became the company's president. It is widely known that she came to Megabucks from rival Lotsabucks.

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

First idea for simplification: do mention detection and filtering first

Possible references:
|words| * |words|

Victoria
Victoria Chen
Victoria Chen ,
Victoria Chen , CFO
Victoria Chen , CFO of
Victoria Chen , CFO of Megabucks
Victoria Chen , CFO of Megabucks Banking
Victoria Chen , CFO of Megabucks Banking ,
Victoria Chen , CFO of Megabucks Banking , saw
Victoria Chen , CFO of Megabucks Banking , saw her

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

First idea for simplification: do mention detection and filtering first

Possible references:
|words| * |words|

Victoria
Victoria Chen
~~Victoria Chen ,~~
Victoria Chen , CFO
~~Victoria Chen , CFO of~~
~~Victoria Chen , CFO of Megabucks~~
~~Victoria Chen , CFO of Megabucks Banking~~
~~Victoria Chen , CFO of Megabucks Banking ,~~
~~Victoria Chen , CFO of Megabucks Banking , saw~~
~~Victoria Chen , CFO of Megabucks Banking , saw her~~

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Second idea for simplification: Link pairs of mentions

Just find the best previous mention for this mention to be coreferent with

Victoria Chen, CFO of Megabucks Banking, saw her pay jump to $2.3 million, as the 38-year-old became the company's president. It is widely known that she came to Megabucks from rival Lotsabucks.

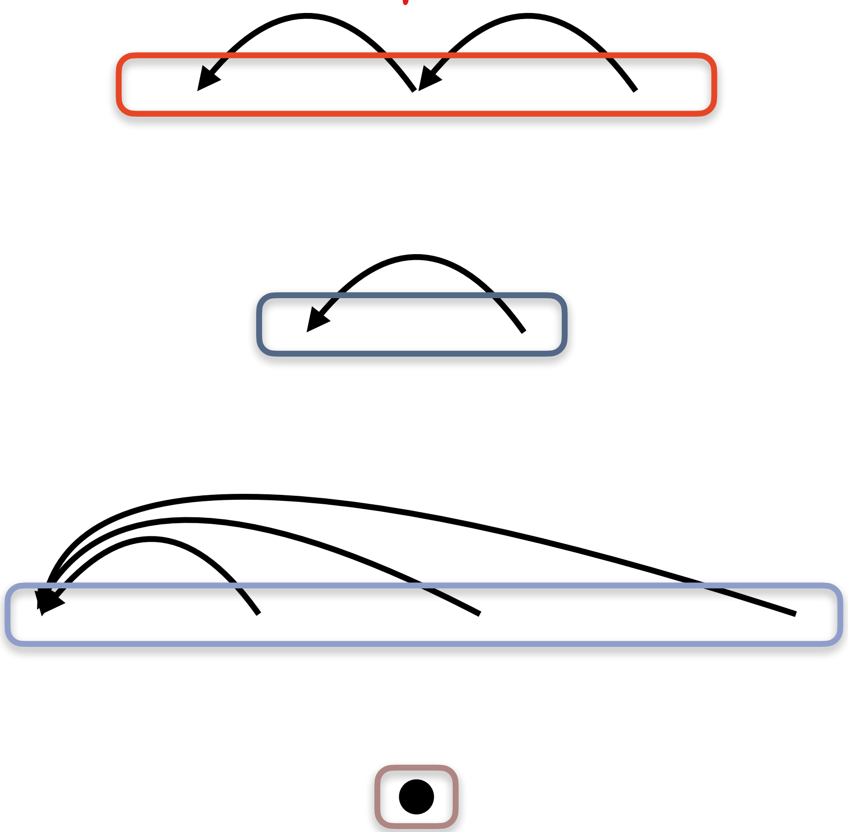
Second idea for simplification: Link pairs of mentions

Just find the best previous mention for this mention to be coreferent with

Victoria Chen, CFO of Megabucks Banking, saw her pay jump to $2.3 million, as the 38-year-old became the company's president. It is widely known that she came to Megabucks from rival Lotsabucks.

Third idea for simplification: Find the transitive closure

Cluster are separate i.e. no arrow toward other cluster



COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

But what about fancy inference algorithms?

why we not always use llm

Sometimes the simple idea ***does*** work!

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Workshop Preview

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Workshop preview - Keras and Tensorflow

Pre-Work	Confirm you can run the notebook on your own machine / a web service! Read the initial sections
In Workshop	LSTM variations GPT-2 text generation

COMP 4446 / 5046
Lecture 5, 2025

Sequence Tagging
Graph Parsing
Coreference
Workshop Preview

menti.com 6274 6616

Muddy Card

Open shortly, closes at 7:05pm

https://saipll.shinyapps.io/student-interface/

If you do not wish to participate in the study, use the Ed form instead

Go to Ed → Lessons → Muddy Cards Lecture 5