Wrap-Up:
Unsupervised Learning, Summary, AMA

LING 571 — Deep Processing Methods in NLP

Shane Steinert-Threlkeld

Announcements

- Let Saiya know about free extension if you haven't used and want to retroactively apply it
- HW9: f-measures output in alphabetical order of the class labels

Pragmatics of the Week

Over three decades later, I walked up to a counter in Antalya Airport to tell a disbelieving airline employee that our flight would shortly be canceled because the tanks being reported in the streets of Istanbul meant that a coup attempt was under way.* *It*

**A previous version of this article misstated the amount of time between 1980 and 2016. It is over three decades, not two.*

https://www.theatlantic.com/ideas/archive/2020/12/trumps-farcical-inept-and-deadly-serious-coup-attempt/617309/#correction%202

Un-/Semi-supervised Learning in NLP

A Roadblock to Deep Processing

- Deep processing of natural language data helps with:
  - Information retrieval
  - QA
  - WSD
  - Conversational AI
  - …
- But….

Developing Deep Processing Systems

- Building a deep processing system requires lots of annotated data
  - For evaluation
  - For *training* an ML system

A roadblock

- The following are cheap:
  - Compute
  - Text [the web!]

- The following are expensive:
  - Human hours
    - Programmers
    - Data annotators

Main Idea

- Leverage the huge amounts of text to learn useful representations
- “Fine tune” on a much smaller amount of task-specific data
  - a.k.a. transfer learning

Can we leverage the cheap resources?

Yann LeCun

https://twitter.com/rgblong/status/916062474545319938?lang=en

Deep Contextualized Word Representations

Peters et. al (2018)

Deep Contextualized Word Representations

Peters et. al (2018)

- Prior vector-space embeddings have typically been derived:
  - Context-independent distributions (CBOW; e.g. GloVe)
  - CNNs over characters

Deep Contextualized Word Representations

Peters et. al (2018)

Deep Contextualized Word Representations

Peters et. al (2018)

- NAACL 2018 Best Paper Award

Deep Contextualized Word Representations

Peters et. al (2018)

- NAACL 2018 Best Paper Award
- **E**mbeddings from **L**anguage **Mo**dels (ELMo)
  - [aka the OG NLP Muppet]

Deep Contextualized Word Representations

Peters et. al (2018)

- NAACL 2018 Best Paper Award
- **E**mbeddings from **L**anguage **Mo**dels (ELMo)
  - [aka the OG NLP Muppet]
- Rather than treat embeddings as bag of words
  - Create embeddings by using sequential modeling (bi-LSTM)

Deep Contextualized Word Representations

Peters et. al (2018)

- Comparison to GloVe:

	Source	Nearest Neighbors
GloVe	play	playing, game, games, played, players, plays, player, Play, football, multiplayer
biLM	Chico Ruiz made a spectacular **play** on Alusik's grounder…	Kieffer, the only junior in the group, was commended for his ability to hit in the clutch, as well as his all-round excellent **play**.
	Olivia De Havilland signed to do a Broadway **play** for Garson…	…they were actors who had been handed fat roles in a successful **play**, and had talent enough to fill the roles competently, with nice understatement.

Deep Contextualized Word Representations

Peters et. al (2018)

- Intrinsic evaluation via WSD:

Model	F_1
WordNet 1st Sense Baseline	65.9
Raganato et al(2017a)	69.9
Iacobacci et al.(2016)	70.1
CoVe, First Layer	59.4
CoVe, Second Layer	64.7
biLM, First layer	67.4
biLM, Second layer	69.0

Deep Contextualized Word Representations

Peters et. al (2018)

- Used in place of other embeddings on multiple tasks:

Task	Previous SOTA		Our baseline	ELMo + baseline	Increase (absolute/ relative)
SQuAD	Liu et al. (2017)	84.4	81.1	85.8	4.7 / 24.9%
SNLI	Chen et al. (2017)	88.6	88.0	88.7 ± 0.17	0.7 / 5.8%
SRL	He et al. (2017)	81.7	81.4	84.6	3.2 / 17.2%
Coref	Lee et al. (2017)	67.2	67.2	70.4	3.2 / 9.8%
NER	Peters et al. (2017)	91.93 ± 0.19	90.15	92.22 ± 0.10	2.06 / 21%
SST-5	McCann et al. (2017)	53.7	51.4	54.7 ± 0.5	3.3 / 6.8%

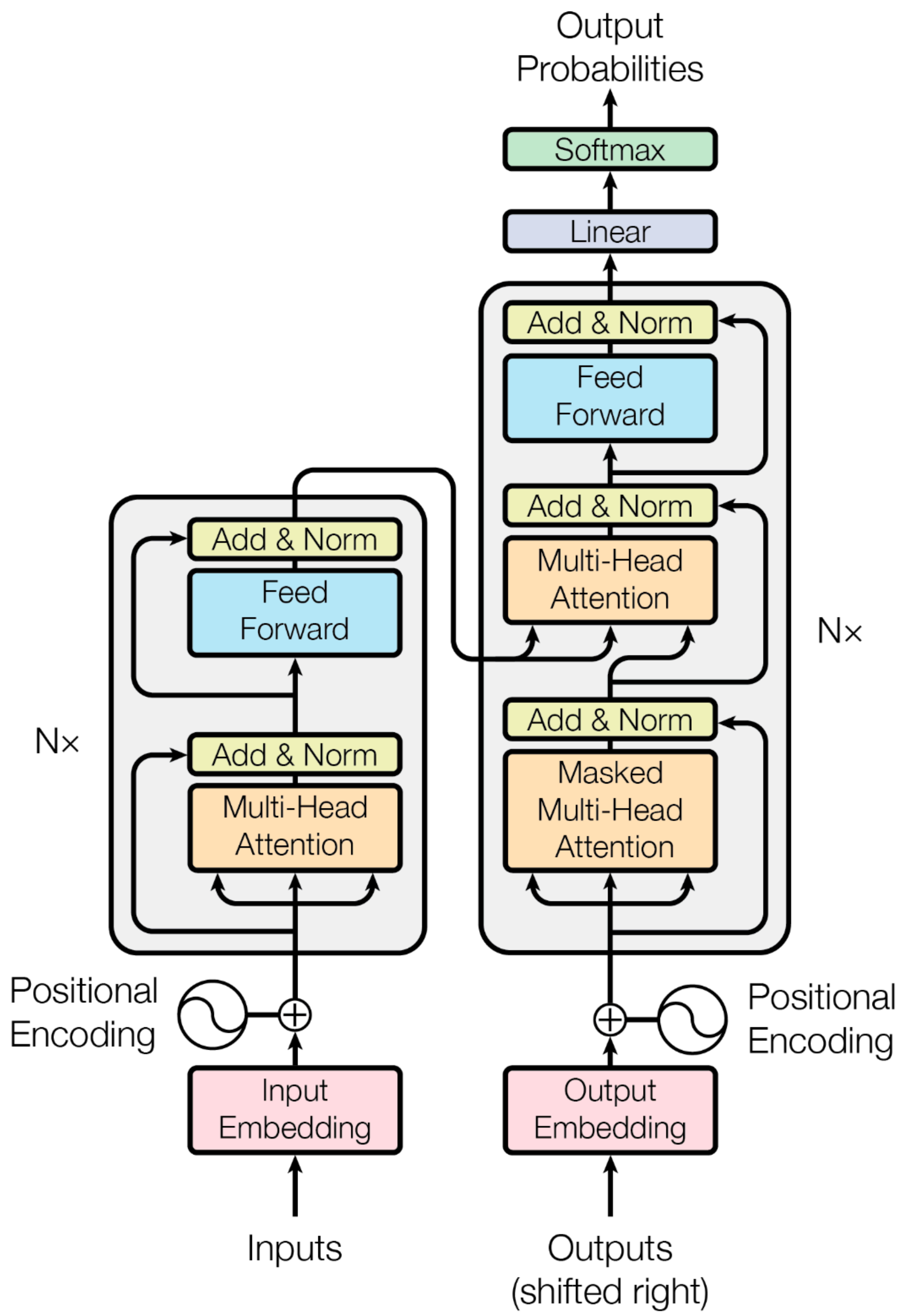
SQuAD = Stanford Question Answering Dataset
SNLI = Stanford Natural Language Inference Corpus
SST-5 = Stanford Sentiment Treebank

BERT

Bidirectional Encoder Representations from Transformers

Devlin et al 2018

Transformers [+ Encoder]

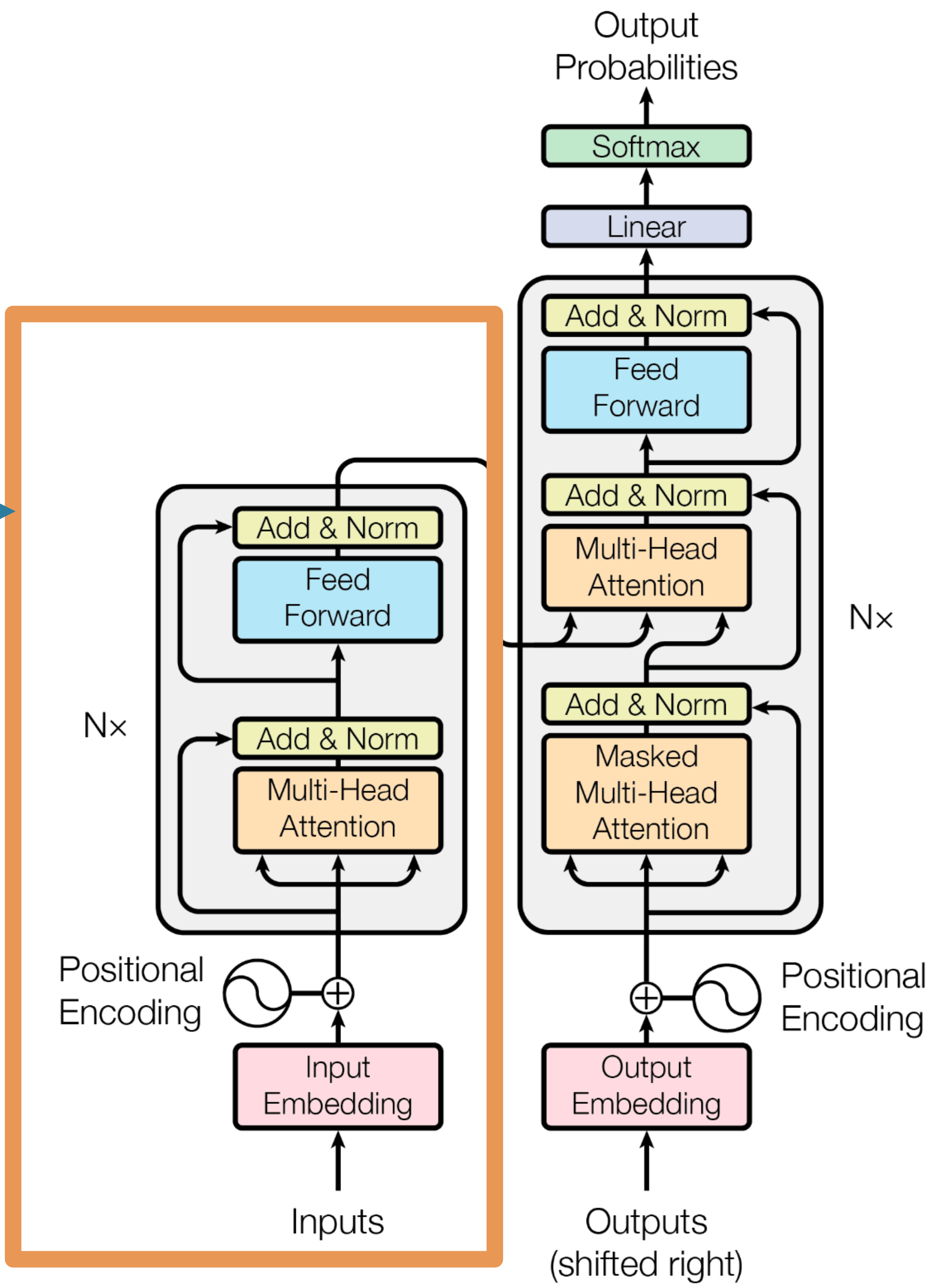


Vashwani et al 2017,
"[Attention is All You Need]"

[The Annotated Transformer]
[The Illustrated Transformer]

Transformers [+ Encoder]

Encoder



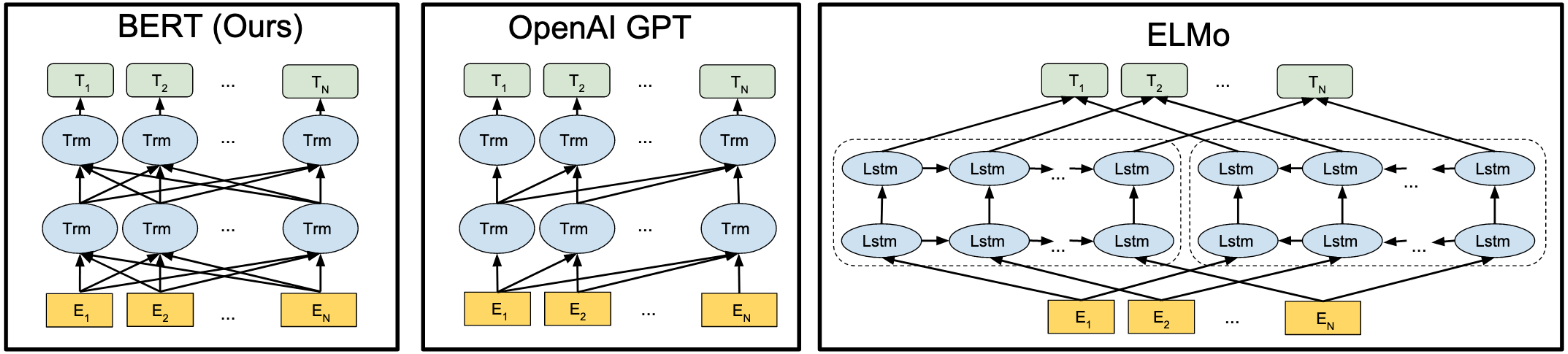
Vashwani et al 2017,
"Attention is All You Need"

The Annotated Transformer
The Illustrated Transformer

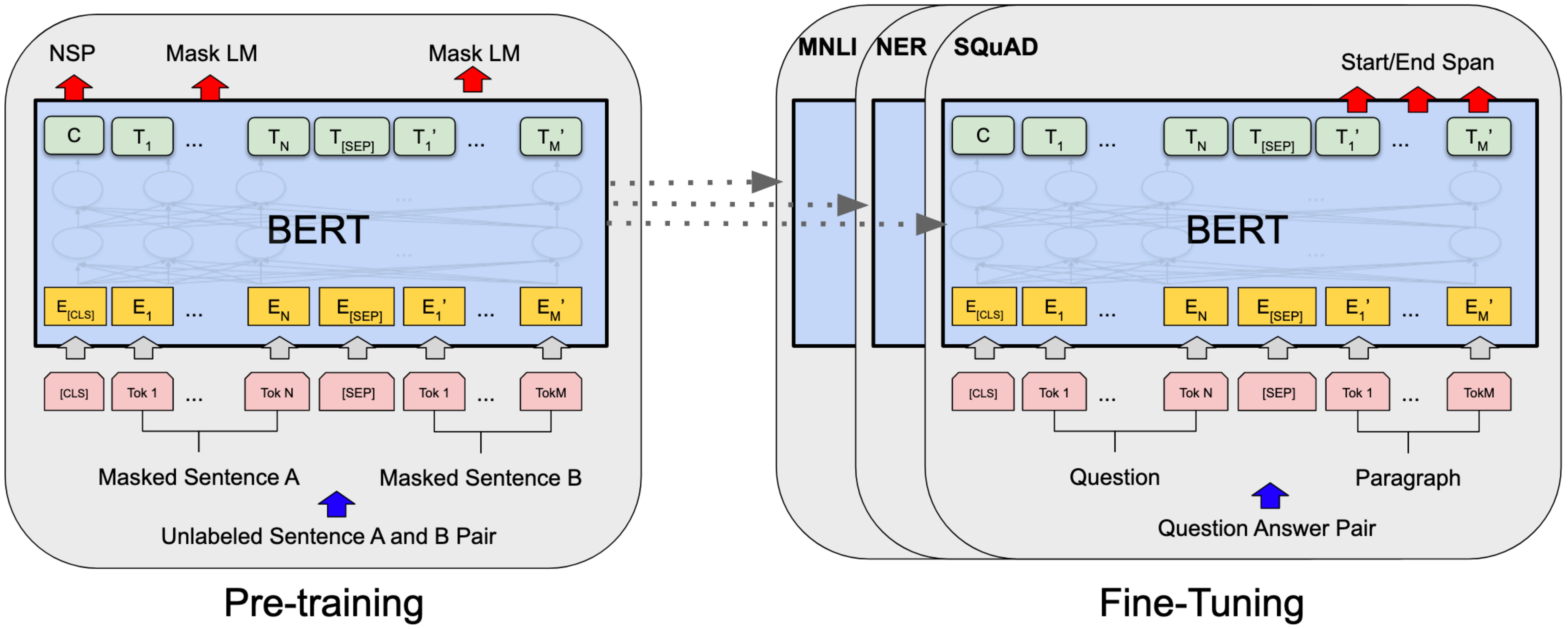
Bidirectional: Masked Language Modeling

- Main training task: *masked language modeling* (aka cloze task)
  - Raw text: “Seattle is the capital of Washington and is the home of UW.”
  - 15% of tokens are masked* (*some subtleties), e.g.:
  - Model input:
    - “Seattle is the [MASK] of Washington and [MASK] the home of UW.”
  - Task: predict the tokens in the [MASK] positions.
- [Also trained with Next Sentence Prediction: given two sentences, did the second follow the first in the text?]

Bidirectional



Fine Tuning



NSP
Mask LM
Mask LM
C
T1
...
TN
T[SEP]
T1'
...
TM'
BERT
E[CLS]
E1
...
EN
E[SEP]
E1'
...
EM'
[CLS]
Tok 1
...
Tok N
[SEP]
Tok 1
...
TokM
Masked Sentence A
Masked Sentence B
Unlabeled Sentence A and B Pair
Pre-training
MNLI
NER
SQuAD
Start/End Span
C
T1
...
TN
T[SEP]
T1'
...
TM'
BERT
E[CLS]
E1
...
EN
E[SEP]
E1'
...
EM'
[CLS]
Tok 1
...
Tok N
[SEP]
Tok 1
...
TokM
Question
Paragraph
Question Answer Pair
Fine-Tuning

Initial Results

System	MNLI-(m/mm)	QQP	QNLI	SST-2	CoLA	STS-B	MRPC	RTE	**Average**
	392k	363k	108k	67k	8.5k	5.7k	3.5k	2.5k	-
Pre-OpenAI SOTA	80.6/80.1	66.1	82.3	93.2	35.0	81.0	86.0	61.7	74.0
BiLSTM+ELMo+Attn	76.4/76.1	64.8	79.8	90.4	36.0	73.3	84.9	56.8	71.0
OpenAI GPT	82.1/81.4	70.3	87.4	91.3	45.4	80.0	82.3	56.0	75.1
$\text{BERT}_{\text{BASE}}$	84.6/83.4	71.2	90.5	93.5	52.1	85.8	88.9	66.4	79.6
$\text{BERT}_{\text{LARGE}}$	**86.7/85.9**	**72.1**	**92.7**	**94.9**	**60.5**	**86.5**	**89.3**	**70.1**	**82.1**

Major Application

The Keyword | Latest Stories | Product Updates | Company News

SEARCH

Understanding searches better than ever before

Pandu Nayak
Google Fellow and Vice President, Search

Published Oct 25, 2019

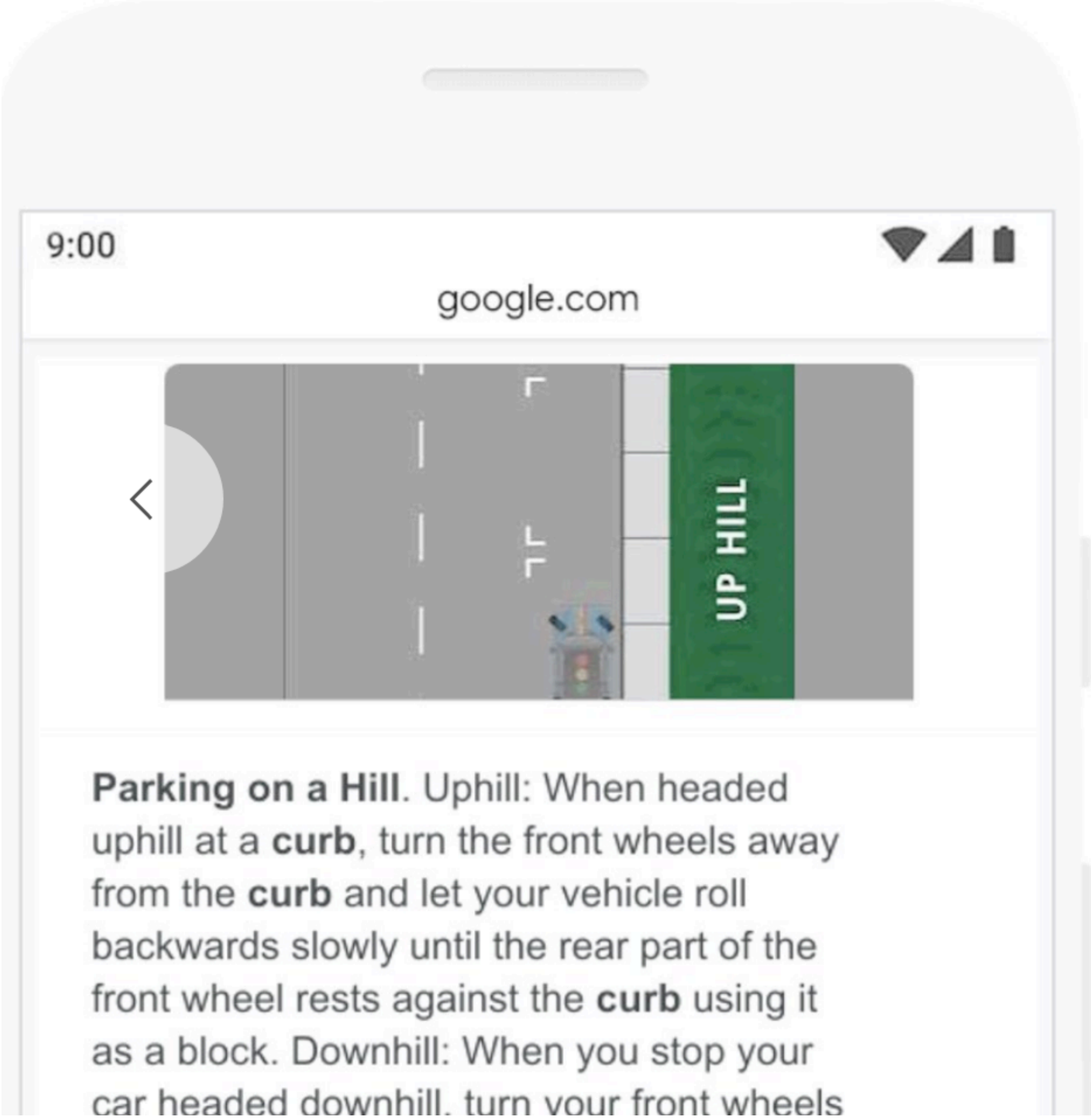
If there's one thing I've learned over the 15 years working on Google Search, it's that people's curiosity is endless. We see billions of searches every day, and 15 percent of those queries are ones we haven't seen before--so we've built ways to return results for queries we can't anticipate.

https://www.blog.google/products/search/search-language-understanding-bert/

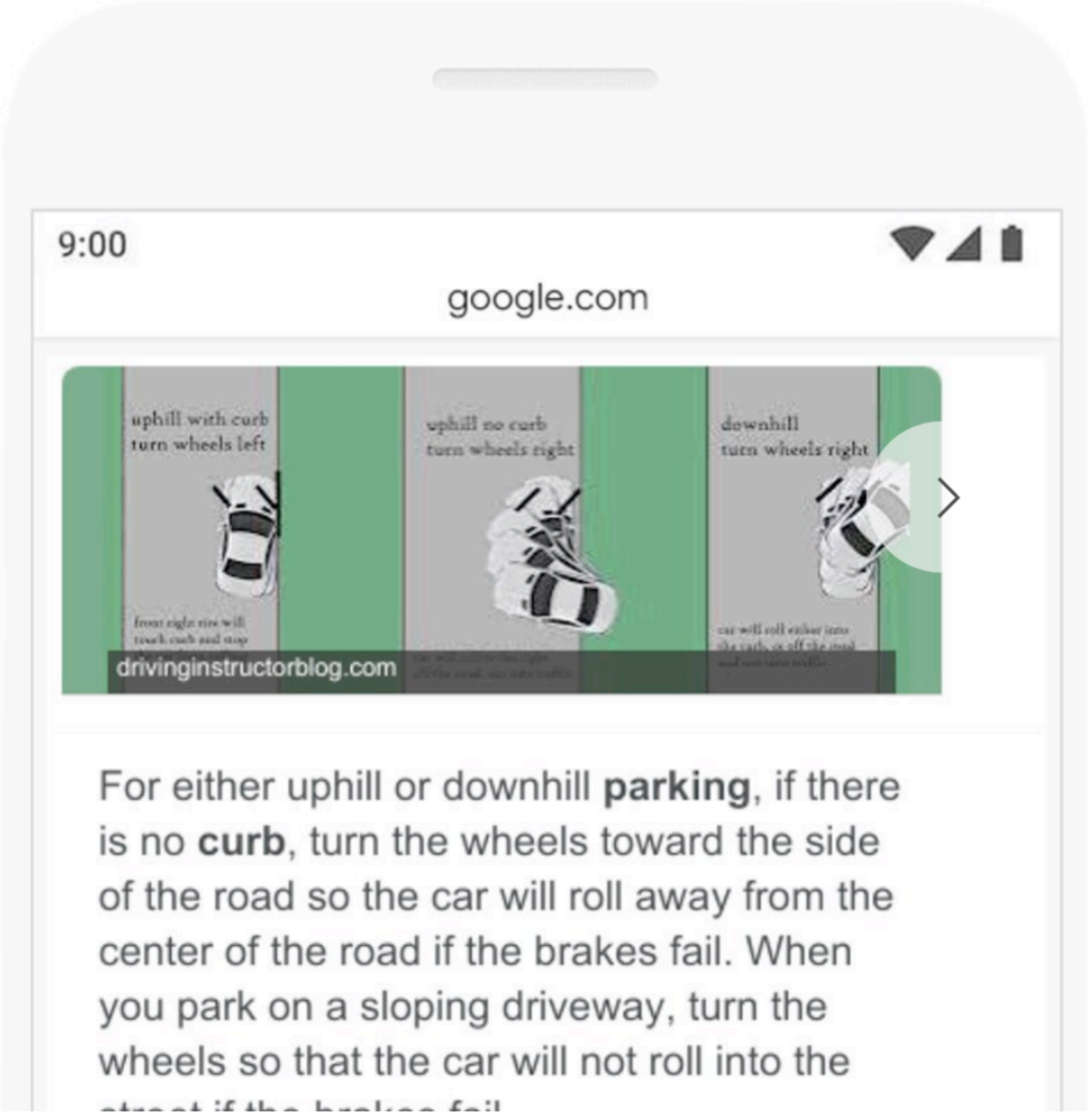
Major Application

parking on a hill with no curb

BEFORE



AFTER



Does BERT implicitly perform deep processing?

WHAT DO YOU LEARN FROM CONTEXT? PROBING FOR SENTENCE STRUCTURE IN CONTEXTUALIZED WORD REPRESENTATIONS

Ian Tenney,[*1] **Patrick Xia,**[2] **Berlin Chen,**[3] **Alex Wang,**[4] **Adam Poliak,**[2]
R. Thomas McCoy,[2] **Najoung Kim,**[2] **Benjamin Van Durme,**[2] **Samuel R. Bowman,**[4]
Dipanjan Das,[1] **and Ellie Pavlick**[1,5]

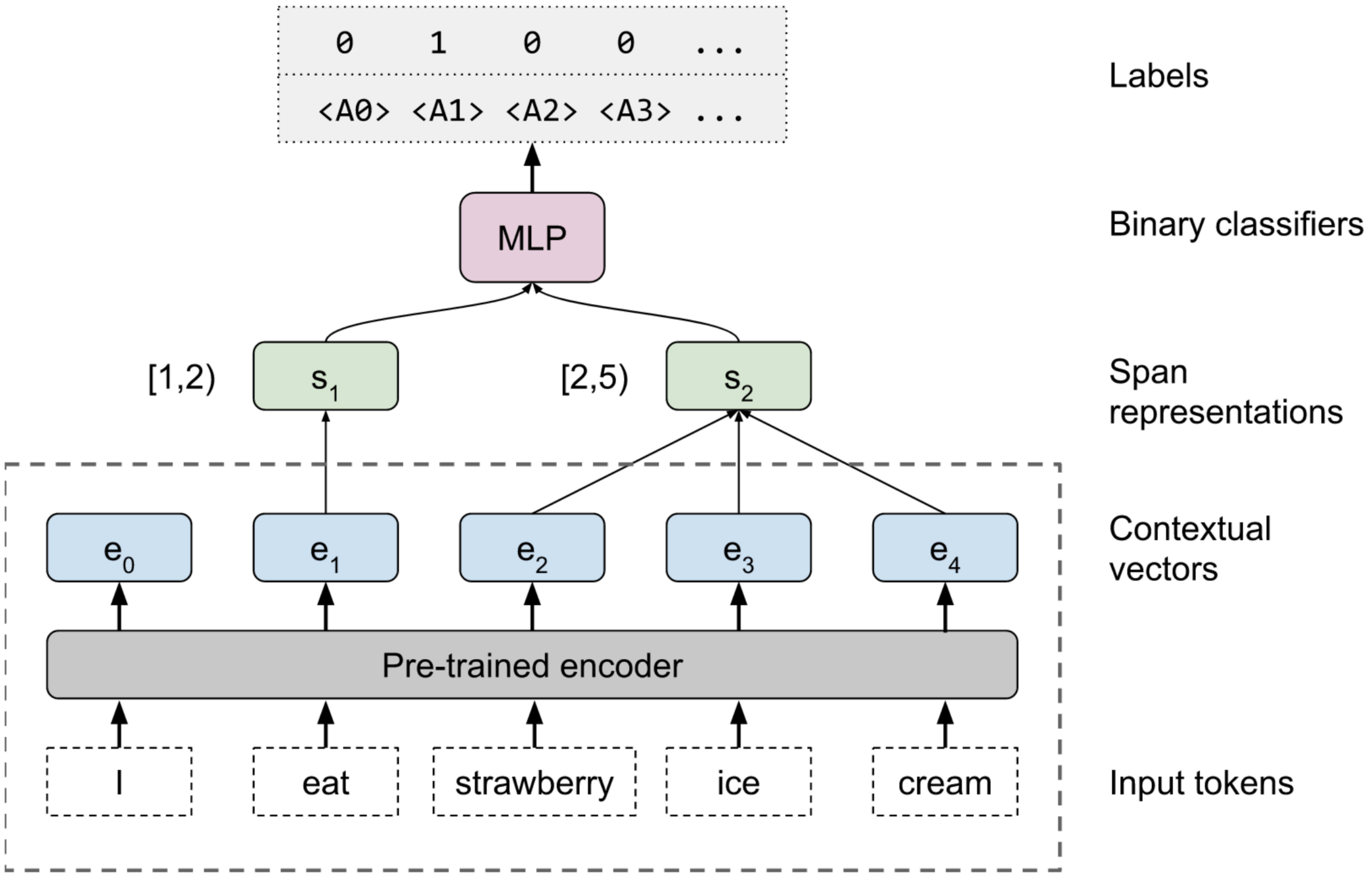
[1]Google AI Language, [2]Johns Hopkins University, [3]Swarthmore College,
[4]New York University, [5]Brown University

Tenney et al 2019

ABSTRACT

Contextualized representation models such as ELMo (Peters et al., 2018a) and BERT (Devlin et al., 2018) have recently achieved state-of-the-art results on a diverse array of downstream NLP tasks. Building on recent token-level probing work, we introduce a novel *edge probing* task design and construct a broad suite of sub-sentence tasks derived from the traditional structured NLP pipeline. We probe word-level contextual representations from four recent models and investigate how they encode sentence structure across a range of syntactic, semantic, local, and long-range phenomena. We find that existing models trained on language modeling and translation produce strong representations for syntactic phenomena, but only offer comparably small improvements on semantic tasks over a non-contextual baseline.

Edge Probing Set-up



Results

	CoVe			ELMo			GPT		
	Lex.	**Full**	**Abs. Δ**	**Lex.**	**Full**	**Abs. Δ**	**Lex.**	**`cat`**	**`mix`**
Part-of-Speech	85.7	94.0	8.4	90.4	**96.7**	6.3	88.2	94.9	95.0
Constituents	56.1	81.6	25.4	69.1	**84.6**	15.4	65.1	81.3	**84.6**
Dependencies	75.0	83.6	8.6	80.4	**93.9**	13.6	77.7	92.1	**94.1**
Entities	88.4	90.3	1.9	92.0	**95.6**	3.5	88.6	92.9	92.5
SRL (all)	59.7	80.4	20.7	74.1	**90.1**	16.0	67.7	86.0	89.7
Core roles	*56.2*	*81.0*	*24.7*	*73.6*	***92.6***	*19.0*	*65.1*	*88.0*	*92.0*
Non-core roles	*67.7*	*78.8*	*11.1*	*75.4*	***84.1***	*8.8*	*73.9*	*81.3*	***84.1***
OntoNotes coref.	72.9	79.2	6.3	75.3	84.0	8.7	71.8	83.6	**86.3**
SPR1	73.7	77.1	3.4	80.1	**84.8**	4.7	79.2	83.5	83.1
SPR2	76.6	80.2	3.6	82.1	83.1	1.0	82.2	**83.8**	**83.5**
Winograd coref.	52.1	**54.3**	2.2	**54.3**	**53.5**	-0.8	51.7	52.6	**53.8**
Rel. (SemEval)	51.0	60.6	9.6	55.7	77.8	22.1	58.2	**81.3**	**81.0**
Macro Average	69.1	78.1	9.0	75.4	**84.4**	9.1	73.0	83.2	**84.4**

	BERT-base				BERT-large				
	F1 Score			Abs. Δ	F1 Score			Abs. Δ	
	Lex.	**`cat`**	**`mix`**	**ELMo**	**Lex.**	**`cat`**	**`mix`**	**(base)**	**ELMo**
Part-of-Speech	88.4	**97.0**	96.7	0.0	88.1	96.5	**96.9**	0.2	0.2
Constituents	68.4	83.7	86.7	2.1	69.0	80.1	**87.0**	0.4	2.5
Dependencies	80.1	93.0	95.1	1.1	80.2	91.5	**95.4**	0.3	1.4
Entities	90.9	96.1	96.2	0.6	91.8	96.2	**96.5**	0.3	0.9
SRL (all)	75.4	89.4	91.3	1.2	76.5	88.2	**92.3**	1.0	2.2
Core roles	*74.9*	*91.4*	*93.6*	*1.0*	*76.3*	*89.9*	***94.6***	*1.0*	*2.0*
Non-core roles	*76.4*	*84.7*	*85.9*	*1.8*	*76.9*	*84.1*	***86.9***	*1.0*	*2.8*
OntoNotes coref.	74.9	88.7	90.2	6.3	75.7	89.6	**91.4**	1.2	7.4
SPR1	79.2	84.7	**86.1**	1.3	79.6	85.1	**85.8**	-0.3	1.0
SPR2	81.7	83.0	**83.8**	0.7	81.6	83.2	**84.1**	0.3	1.0
Winograd coref.	54.3	53.6	54.9	1.4	53.0	53.8	**61.4**	6.5	7.8
Rel. (SemEval)	57.4	78.3	82.0	4.2	56.2	77.6	**82.4**	0.5	4.6
Macro Average	75.1	84.8	86.3	1.9	75.2	84.2	**87.3**	1.0	2.9

Conclusion

- “in general, contextualized embeddings improve over their non-contextualized counterparts largely on syntactic tasks (e.g. constituent labeling) in comparison to semantic tasks (e.g. coreference), suggesting that these embeddings encode syntax more so than higher-level semantics”

BERT Rediscovers the Classical NLP Pipeline

Ian Tenney[1] **Dipanjan Das**[1] **Ellie Pavlick**[1,2]
[1]Google Research [2]Brown University
{iftenney,dipanjand,epavlick}@google.com

Abstract

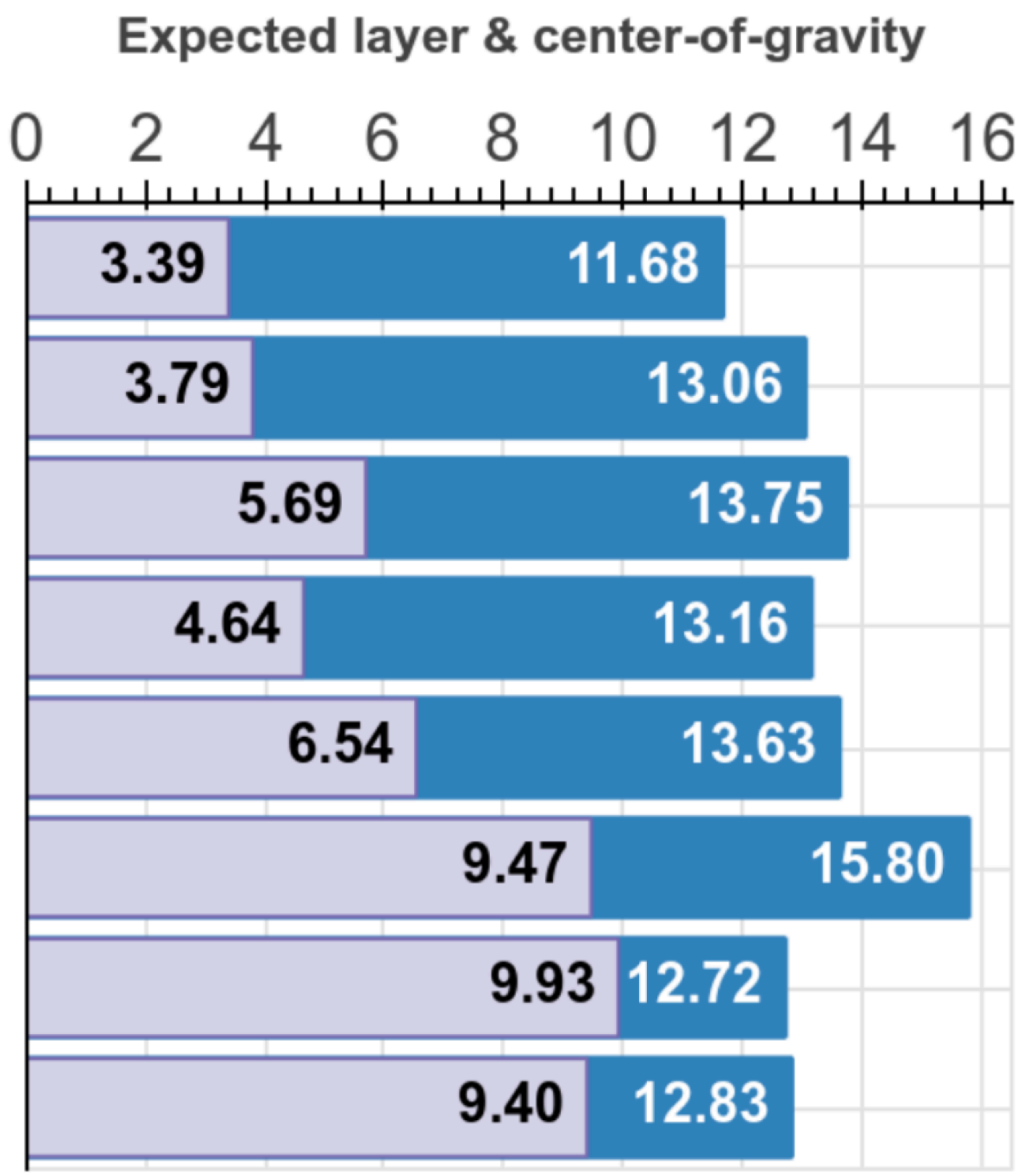
Pre-trained text encoders have rapidly advanced the state of the art on many NLP tasks. We focus on one such model, BERT, and aim to quantify where linguistic information is captured within the network. We find that the model represents the steps of the traditional NLP pipeline in an interpretable and localizable way, and that the regions responsible for each step appear in the expected sequence: POS tagging, parsing, NER, semantic roles, then coreference. Qualitative analysis reveals that the model can and often does adjust this pipeline dynamically, revising lower-level decisions on the basis of disambiguating information from higher-level representations.

of the network directly, to assess whether there exist localizable regions associated with distinct types of linguistic decisions. Such work has produced evidence that deep language models can encode a range of syntactic and semantic information (e.g. Shi et al., 2016; Belinkov, 2018; Tenney et al., 2019), and that more complex structures are represented hierarchically in the higher layers of the model (Peters et al., 2018b; Blevins et al., 2018).

We build on this latter line of work, focusing on the BERT model (Devlin et al., 2019), and use a suite of probing tasks (Tenney et al., 2019) derived from the traditional NLP pipeline to quantify where specific types of linguistic information are

Tenney et al 2019

	F1 Scores	
	ℓ=0	ℓ=24
POS	88.5	96.7
Consts.	73.6	87.0
Deps.	85.6	95.5
Entities	90.6	96.1
SRL	81.3	91.4
Coref.	80.5	91.9
SPR	77.7	83.7
Relations	60.7	84.2



A Structural Probe for Finding Syntax in Word Representations

John Hewitt
Stanford University
johnhew@stanford.edu

Christopher D. Manning
Stanford University
manning@stanford.edu

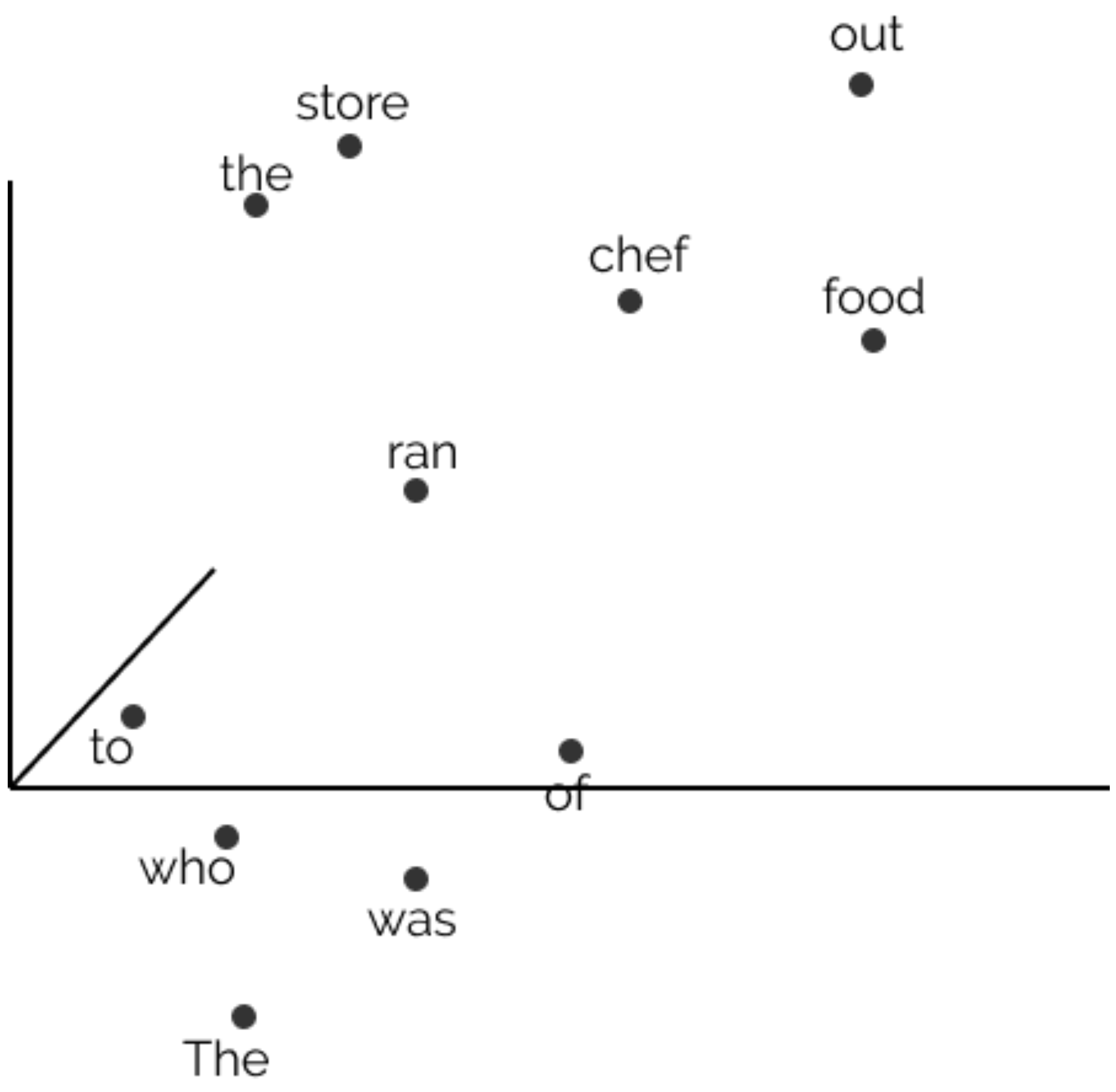
Abstract

Recent work has improved our ability to detect linguistic knowledge in word representations. However, current methods for detecting syntactic knowledge do not test whether syntax trees are represented in their entirety. In this work, we propose a *structural probe*, which evaluates whether syntax trees are embedded in a linear transformation of a neural network's word representation space. The probe identifies a linear transformation under which squared L2 distance encodes the distance between words in the parse tree, and one in which squared L2 norm encodes depth in the parse tree. Using our probe, we show

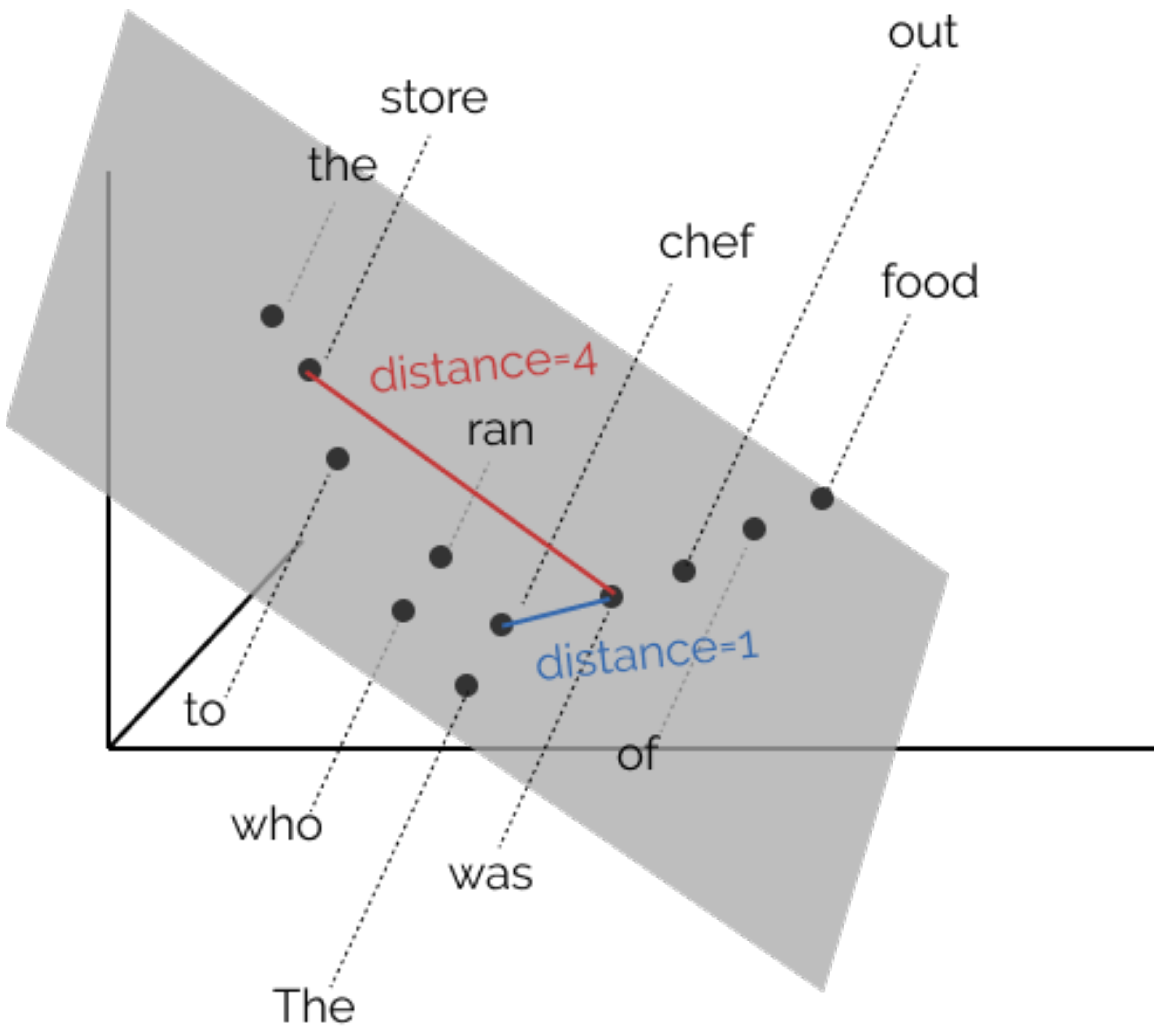
In this work, we propose a *structural probe*, a simple model which tests whether syntax trees are consistently embedded in a linear transformation of a neural network's word representation space. Tree structure is embedded if the transformed space has the property that squared L2 distance between two words' vectors corresponds to the number of edges between the words in the parse tree. To reconstruct edge directions, we hypothesize a linear transformation under which the squared L2 norm corresponds to the depth of the word in the parse tree. Our probe uses supervision to find the transformations under which these properties are best approximated for each model. If such transfor-

Hewitt and Manning 2019
blog post

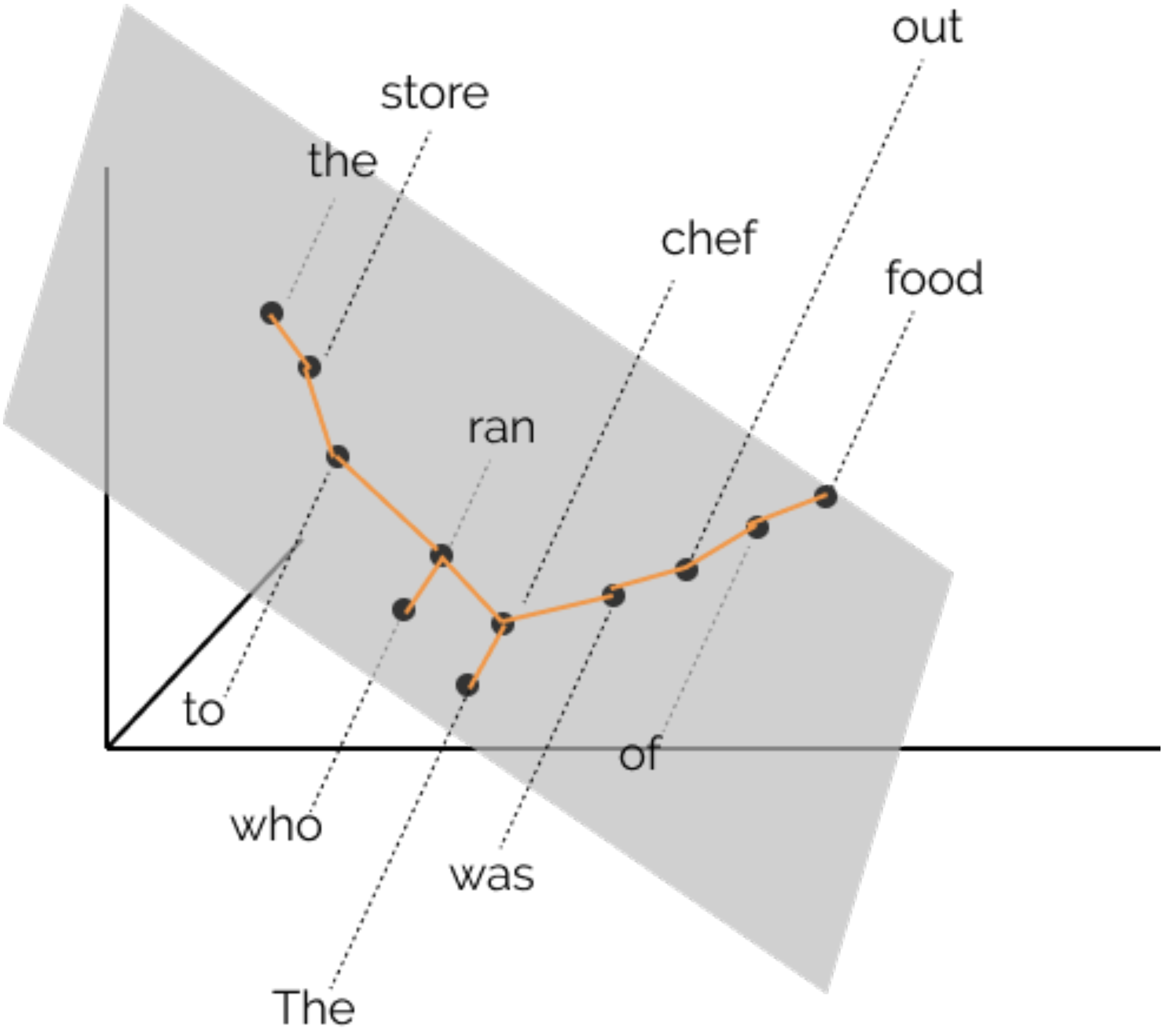
“The chef who ran to the store was out of food.”



“The chef who ran to the store was out of food.”



“The chef who ran to the store was out of food.”

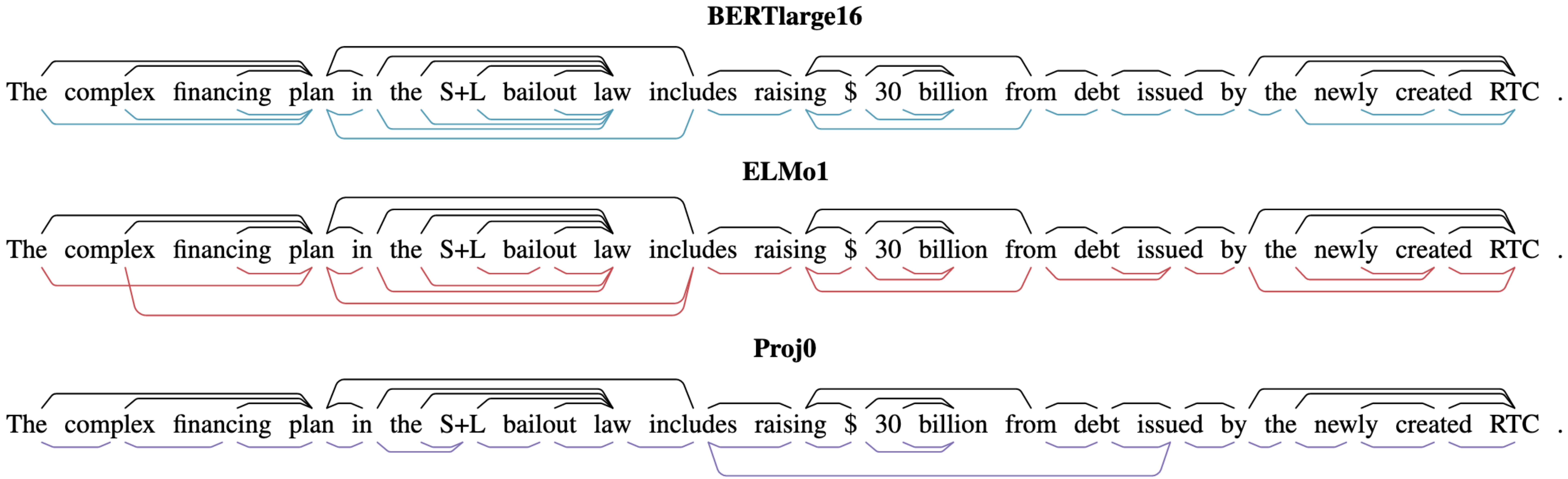


Results

Method	Distance		Depth	
	UUAS	DSpr.	Root%	NSpr.
LINEAR	48.9	0.58	2.9	0.27
ELMO0	26.8	0.44	54.3	0.56
DECAY0	51.7	0.61	54.3	0.56
PROJ0	59.8	0.73	64.4	0.75
ELMO1	77.0	0.83	86.5	0.87
BERTBASE7	79.8	0.85	88.0	0.87
BERTLARGE15	**82.5**	0.86	89.4	0.88
BERTLARGE16	81.7	**0.87**	**90.1**	**0.89**

[SOTA: directed UAS >97%]

Examples



Black = gold parse.
Model parses: Maximum Spanning Tree from distances in transformed space.

Limitations of Large LMs

Right for the Wrong Reasons: Diagnosing Syntactic Heuristics in Natural Language Inference

R. Thomas McCoy,[1] Ellie Pavlick,[2] & Tal Linzen[1]
[1]Department of Cognitive Science, Johns Hopkins University
[2]Department of Computer Science, Brown University
tom.mccoy@jhu.edu, ellie_pavlick@brown.edu, tal.linzen@jhu.edu

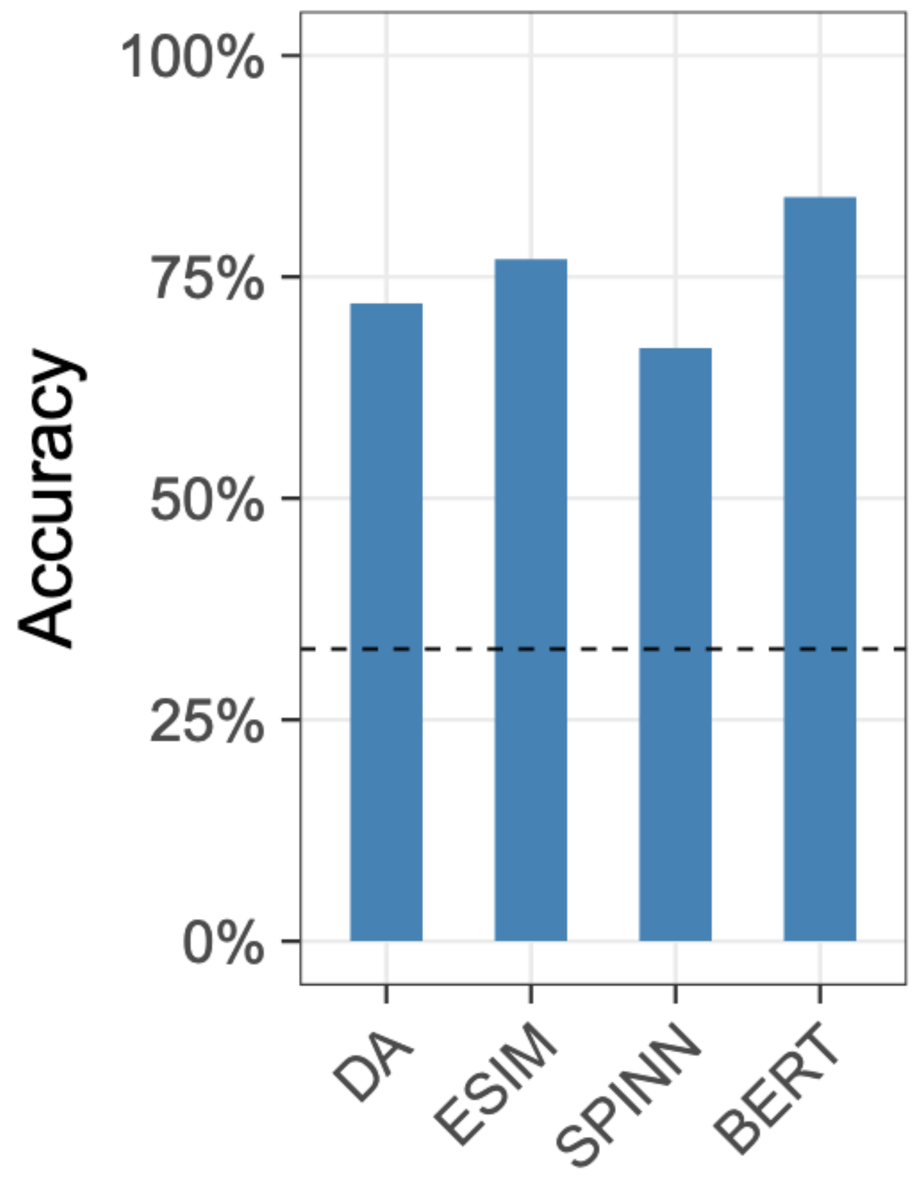
McCoy et al 2019

Main Idea

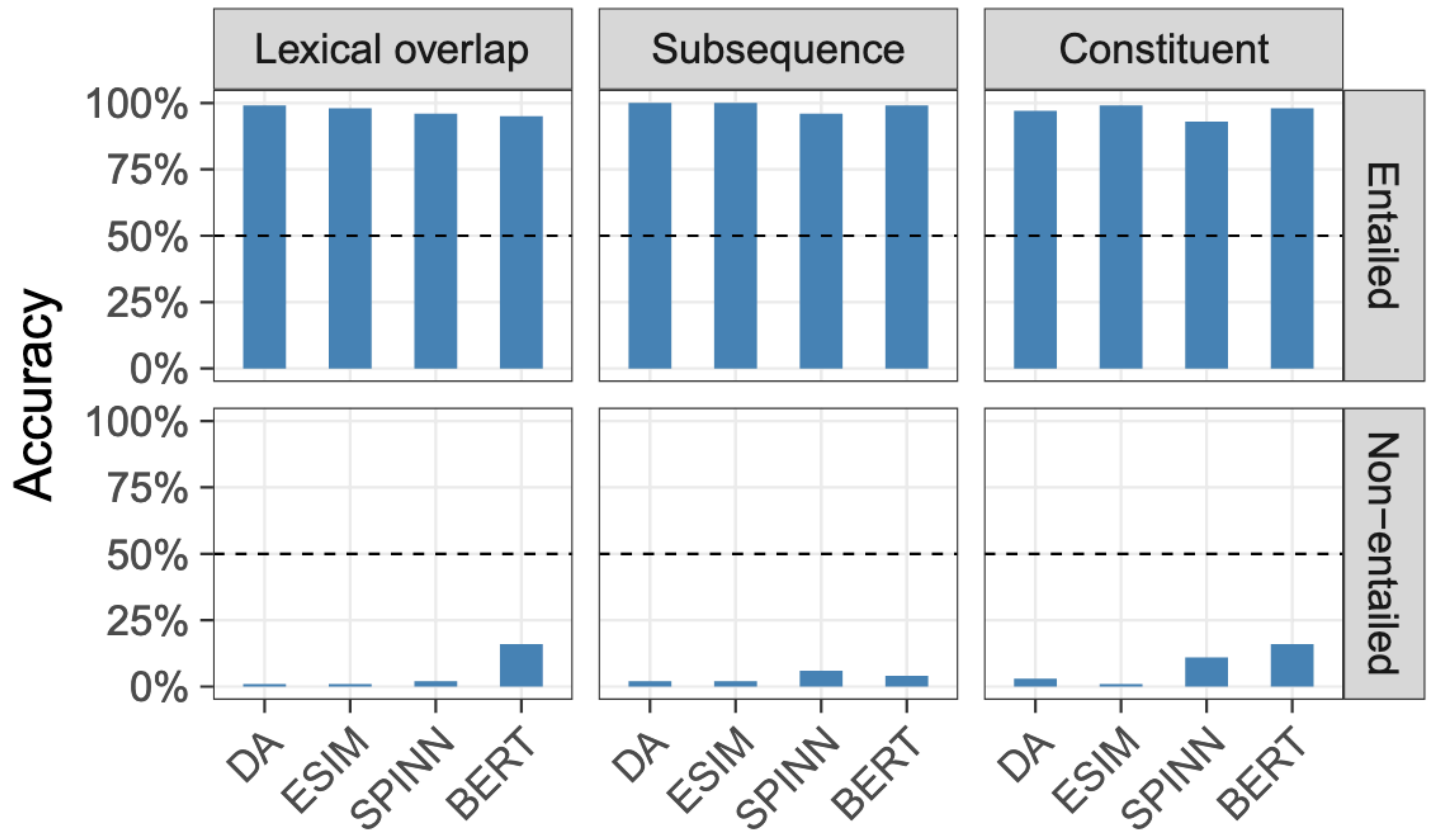
- BERT et al do really well on natural language understanding tasks like NLI (natural language inference)
- Do they do so “for the right reasons”?
- In other words:
  - Or does solving the existing datasets mean they’ve solved the task?
  - Or can success reflect other features than deep language understanding?

Heuristic	Premise	Hypothesis	Label
Lexical overlap heuristic	The banker near the judge saw the actor.	The banker saw the actor.	E
	The lawyer was advised by the actor.	The actor advised the lawyer.	E
	The doctors visited the lawyer.	The lawyer visited the doctors.	N
	The judge by the actor stopped the banker.	The banker stopped the actor.	N
Subsequence heuristic	The artist and the student called the judge.	The student called the judge.	E
	Angry tourists helped the lawyer.	Tourists helped the lawyer.	E
	The judges heard the actors resigned.	The judges heard the actors.	N
	The senator near the lawyer danced.	The lawyer danced.	N
Constituent heuristic	Before the actor slept, the senator ran.	The actor slept.	E
	The lawyer knew that the judges shouted.	The judges shouted.	E
	If the actor slept, the judge saw the artist.	The actor slept.	N
	The lawyers resigned, or the artist slept.	The artist slept.	N

Results



(a)

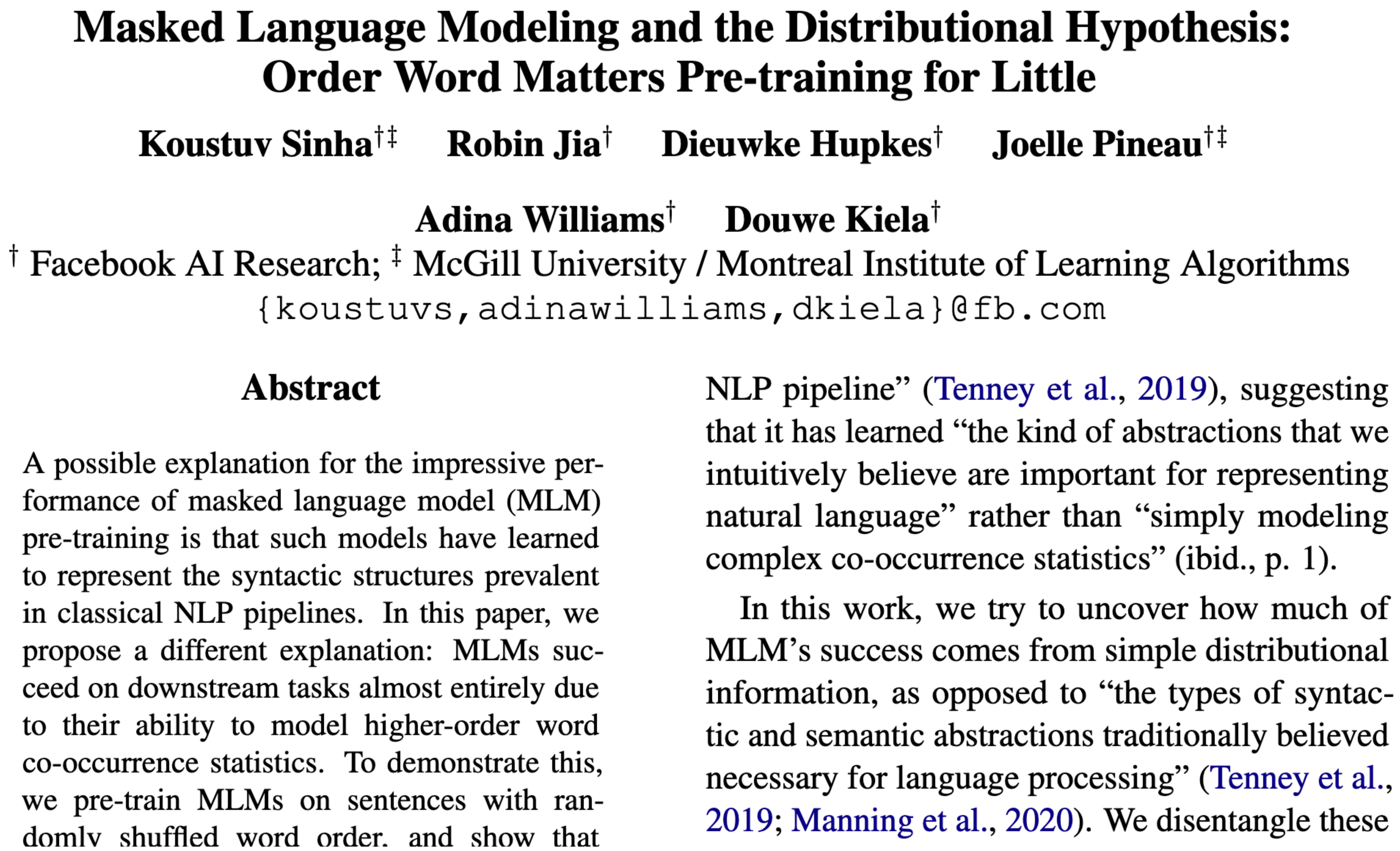


(b)

(performance improves if fine-tuned on this challenge set)

Word Order in the Large LM Era

- ‘Early’ demo that neural bag-of-words works well: “Deep Unordered Composition Rivals Syntactic Methods for Text Classification” —2015
- Large (M)LM success is not due to word order (paper):



Masked Language Modeling and the Distributional Hypothesis: Order Word Matters Pre-training for Little

Koustuv Sinha[†‡] Robin Jia[†] Dieuwke Hupkes[†] Joelle Pineau[†‡]

Adina Williams[†] Douwe Kiela[†]

[†] Facebook AI Research; [‡] McGill University / Montreal Institute of Learning Algorithms

{koustuvs,adinawilliams,dkiela}@fb.com

Abstract

A possible explanation for the impressive performance of masked language model (MLM) pre-training is that such models have learned to represent the syntactic structures prevalent in classical NLP pipelines. In this paper, we propose a different explanation: MLMs succeed on downstream tasks almost entirely due to their ability to model higher-order word co-occurrence statistics. To demonstrate this, we pre-train MLMs on sentences with randomly shuffled word order, and show that

NLP pipeline” (Tenney et al., 2019), suggesting that it has learned “the kind of abstractions that we intuitively believe are important for representing natural language” rather than “simply modeling complex co-occurrence statistics” (ibid., p. 1).

In this work, we try to uncover how much of MLM’s success comes from simple distributional information, as opposed to “the types of syntactic and semantic abstractions traditionally believed necessary for language processing” (Tenney et al., 2019; Manning et al., 2020). We disentangle these

Word Order in the Large LM Era

- ‘Early’ demo that neural bag-of-words works well: “Deep Unordered Composition Rivals Syntactic Methods for Text Classification” —2015
- Large (M)LM success is not due to word order (paper):

Masked Language Modeling and the Distributional Hypothesis: Order Word Matters Pre-training for Little

Koustuv Sinha†‡ **Robin Jia**† **Dieuwke Hupkes**† **Joelle Pineau**†‡
Adina Williams† **Douwe Kiela**†
† Facebook AI Research; ‡ McGill University / Montreal Institute of Learning Algorithms
{koustuvs,adinawilliams,dkiela}@fb.com

Abstract

A possible explanation for the impressive performance of masked language model (MLM) pre-training is that such models have learned to represent the syntactic structures prevalent in classical NLP pipelines. In this paper, we propose a different explanation: MLMs succeed on downstream tasks almost entirely due to their ability to model higher-order word co-occurrence statistics. To demonstrate this, we pre-train MLMs on sentences with randomly shuffled word order, and show that

NLP pipeline” (Tenney et al., 2019), suggesting that it has learned “the kind of abstractions that we intuitively believe are important for representing natural language” rather than “simply modeling complex co-occurrence statistics” (ibid., p. 1).

In this work, we try to uncover how much of MLM’s success comes from simple distributional information, as opposed to “the types of syntactic and semantic abstractions traditionally believed necessary for language processing” (Tenney et al., 2019; Manning et al., 2020). We disentangle these

Costs of LMs

Costs of LMs

- Currently something of an ‘arms race’ between e.g. Google, Facebook, OpenAI, MS, Baidu

Costs of LMs

- Currently something of an ‘arms race’ between e.g. Google, Facebook, OpenAI, MS, Baidu
- Hugely expensive
  - Carbon emissions
  - Monetarily
    - Inequitable access

Costs of LMs

- Currently something of an 'arms race' between e.g. Google, Facebook, OpenAI, MS, Baidu
- Hugely expensive
  - Carbon emissions
  - Monetarily
    - Inequitable access

Energy and Policy Considerations for Deep Learning in NLP

Emma Strubell **Ananya Ganesh** **Andrew McCallum**
College of Information and Computer Sciences
University of Massachusetts Amherst
{strubell, aganesh, mccallum}@cs.umass.edu

Abstract

Recent progress in hardware and methodology for training neural networks has ushered in a new generation of large networks trained on abundant data. These models have obtained notable gains in accuracy across many NLP tasks. However, these accuracy improvements depend on the availability of exceptionally large computational resources that necessitate similarly substantial energy consumption. As a result these models are costly to train and develop, both financially, due to the cost of hardware and electricity or cloud compute time, and environmentally, due to the carbon footprint required to fuel modern tensor

Consumption	CO_2e (lbs)
Air travel, 1 person, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experiments	78,468
Transformer (big)	192
w/ neural arch. search	626,155

Table 1: Estimated CO_2 emissions from training common NLP models, compared to familiar consumption.[1]

Costs of LMs

- Currently something of an ‘arms race’ between e.g. Google, Facebook, OpenAI, MS, Baidu
- Hugely expensive
  - Carbon emissions
  - Monetarily
    - Inequitable access

Costs of LMs

- Currently something of an ‘arms race’ between e.g. Google, Facebook, OpenAI, MS, Baidu
- Hugely expensive
  - Carbon emissions
  - Monetarily
    - Inequitable access

Green AI

Roy Schwartz*♢ Jesse Dodge*♢♣ Noah A. Smith♢♡ Oren Etzioni♢

♢Allen Institute for AI, Seattle, Washington, USA
♣ Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
♡ University of Washington, Seattle, Washington, USA

July 2019

Abstract

The computations required for deep learning research have been doubling every few months, resulting in an estimated 300,000x increase from 2012 to 2018 [2]. These computations have a surprisingly large carbon footprint [40]. Ironically, deep learning was inspired by the human brain, which is remarkably energy efficient. Moreover, the financial cost of the computations can make it difficult for academics, students, and researchers, in particular those from emerging economies, to engage in deep learning research.

This position paper advocates a practical solution by making **efficiency** an evaluation criterion for research alongside accuracy and related measures. In addition, we propose reporting the financial cost or “price tag” of developing, training, and running models to provide baselines for the investigation of increasingly efficient methods. Our goal is to make AI both greener and more inclusive—enabling any inspired undergraduate with a laptop to write high-quality research papers. Green AI is an emerging focus at the Allen Institute for AI.

“Deep” Understanding?

Climbing towards NLU:
On Meaning, Form, and Understanding in the Age of Data

Emily M. Bender
University of Washington
Department of Linguistics
ebender@uw.edu

Alexander Koller
Saarland University
Dept. of Language Science and Technology
koller@coli.uni-saarland.de

Abstract

The success of the large neural language models on many NLP tasks is exciting. However, we find that these successes sometimes lead to hype in which these models are being described as “understanding” language or capturing “meaning”. In this position paper, we argue that a system trained only on form has *a priori* no way to learn meaning. In keeping with the ACL 2020 theme of “Taking Stock of Where We’ve Been and Where We’re Going”, we argue that a clear understanding of the dis-

the structure and use of language and the ability to ground it in the world. While large neural LMs may well end up being important components of an eventual full-scale solution to human-analogous NLU, they are not nearly-there solutions to this grand challenge. We argue in this paper that genuine progress in our field — climbing the right hill, not just the hill on whose slope we currently sit — depends on maintaining clarity around big picture notions such as *meaning* and *understanding* in task design and reporting of experimental results.

https://www.aclweb.org/anthology/2020.acl-main.463/

L’Affaire Gebru

- Bender, Gebru, and others’ “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? ”
  - Environmental + financial costs
  - Research opportunity costs
  - Datasets so large they are impossible to audit
- Initial media coverage (now many others):
  - https://www.nytimes.com/2020/12/03/technology/google-researcher-timnit-gebru.html
  - https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/
- Gebru’s new initiative: Distributed AI Research (DAIR)

Google Researcher Says She Was Fired Over Paper Highlighting Bias in A.I.

Timnit Gebru, one of the few Black women in her field, had voiced exasperation over the company’s response to efforts to increase minority hiring.

MIT Technology Review

Artificial intelligence / Machine learning

We read the paper that forced Timnit Gebru out of Google. Here’s what it says.

Summary

- Pre-trained large LMs are very powerful
- Transfer learning from them often leads to very strong performance on NLP tasks
- Why?
  - Some evidence of *some* internal deep processing (esp. syntax)
  - Very clever exploitation of spurious correlations in the data
- Drawbacks:
  - Costs
  - Limited understanding
  - Inscrutability

From LMs to Chatbots

GPT Assistant training pipeline

Stage	Pretraining	Supervised Finetuning	Reward Modeling	Reinforcement Learning
Dataset	**Raw internet** text trillions of words low-quality, large quantity	**Demonstrations** Ideal Assistant responses, ~10-100K (prompt, response) written by contractors low quantity, high quality	**Comparisons** 100K –1M comparisons written by contractors low quantity, high quality	**Prompts** ~10K-100K prompts written by contractors low quantity, high quality
Algorithm	**Language modeling** predict the next token	**Language modeling** predict the next token	**Binary classification** predict rewards consistent w preferences	**Reinforcement Learning** generate tokens that maximize the reward
		init from	init from	init from SFT use RM
Model	Base model	SFT model	RM model	RL model
Notes	1000s of GPUs months of training ex: GPT, LLaMA, PaLM can deploy this model	1-100 GPUs days of training ex: Vicuna-13B can deploy this model	1-100 GPUs days of training	1-100 GPUs days of training ex: ChatGPT, Claude can deploy this model

Source: Andrej Karpathy @ MS Build

AMA / General Discussion

AMA Questions, Models

- In the context of question answering algorithms, what are common or promising approaches for building systems that can answer questions and provide some explanation of how the machine arrived at its answer? I'm thinking of both cases where we want some kind of reasoning chain (x because of y) and other cases where we wan the machine to not only provide an answer, but also provide some source or set of sources ideally tagged to each corresponding claim in the answer. What are the different elements needed to implement something like this?
  - Rigorously assessing natural language explanations: https://arxiv.org/abs/2309.10312v1
  - Retrieval augmented generation (much follow up now): https://arxiv.org/abs/2005.11401

AMA Questions, Stochastic Parrots

- How much of what the authors raise are issues particular to LLM's vs industries at large? The environmental cost for instance doesn't seem particular to LLM training, but true for just about any energy intensive industry.

- As far as documenting the data sets, this seems reasonable, but what exactly documentation means might be unclear. What is the criteria exactly?
  - Data Statements: https://techpolicylab.uw.edu/data-statements/
  - Data Sheets: https://arxiv.org/abs/1803.09010
  - Dataset Nutrition Label: https://datanutrition.org/labels/

AMA Questions: Moving Forward

- How to self-study and practice the application of PyTorch and Tensorflow on NLP tasks during the holidays (to prepare for internship interviews)?
- Opportunities within UW/compling to engage in topics related to this class.
- Differences between computational linguistics and NLP in general and with regard to roles in industry.

Course Recap / Highlights

Wrapping Up

Deep Processing

- Building of deep linguistic structures for NLP
  - Syntax
  - Semantics
  - Pragmatics

- Used and useful in many applications, e.g.
  - IR/QA/search
  - Conversational AI

Syntax

- Constituency Parsing
  - (P)CFGs
  - Grammar induction
- Dependency Parsing
  - Transition vs. MST based parsers

CKY Parsing Example

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]		
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]		
		Det [2,3]	**NP** [2,4]		
			Noun, Nom [3,4]		
				Prep [4,5]	

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

0 *I* 1 *prefer* 2 *a* 3 *flight* 4 *on* 5 *TWA* 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]		
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]		
		Det [2,3]	NP [2,4]		
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]		
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]		
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]		
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]		
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]		
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	
		Det [2,3]	NP [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathscr{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]		
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* | *include* | *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* | *she* | *me*
NP → *TWA* | *Houston*
NP → *Det Nominal*

Nominal → *book* | *flight* | *meal* | *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* | *include* | *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]		
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I prefer a flight on TWA

0 1 2 3 4 5 6

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	
		Det [2,3]	NP [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* | *include* | *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* | *she* | *me*
NP → *TWA* | *Houston*
NP → *Det Nominal*

Nominal → *book* | *flight* | *meal* | *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* | *include* | *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar
S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP
NP → I / she / me
NP → TWA / Houston
NP → Det Nominal
Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP
VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP
PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon
Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar
S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP
NP → I / she / me
NP → TWA / Houston
NP → Det Nominal
Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP
VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP
PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	

Lexicon
Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathscr{L}_1$ Grammar
S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP
NP → I / she / me
NP → TWA / Houston
NP → Det Nominal
Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP
VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP
PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	
					NNP, NP [5,6]

Lexicon
Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

0 I 1 prefer 2 a 3 flight 4 on 5 TWA 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* | *include* | *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* | *she* | *me*
NP → *TWA* | *Houston*
NP → *Det Nominal*

Nominal → *book* | *flight* | *meal* | *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* | *include* | *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	
		Det [2,3]	NP [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathscr{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	**NP** [2,6]
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	
		Det [2,3]	**NP** [2,4]	[2,5]	**NP** [2,6]
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* | *include* | *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* | *she* | *me*
NP → *TWA* | *Houston*
NP → *Det Nominal*

Nominal → *book* | *flight* | *meal* | *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* | *include* | *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I prefer a flight on TWA

0 1 2 3 4 5 6

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	[1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* / *include* / *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* / *she* / *me*
NP → *TWA* / *Houston*
NP → *Det Nominal*

Nominal → *book* / *flight* / *meal* / *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* / *include* / *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I *prefer* *a* *flight* *on* *TWA*

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar
$S \rightarrow NP\ VP$
$S \rightarrow X1\ VP$
$X1 \rightarrow Aux\ NP$
$S \rightarrow book \mid include \mid prefer$
$S \rightarrow Verb\ NP$
$S \rightarrow X2\ PP$
$S \rightarrow Verb\ PP$
$S \rightarrow VP\ PP$
$NP \rightarrow I \mid she \mid me$
$NP \rightarrow TWA \mid Houston$
$NP \rightarrow Det\ Nominal$
$Nominal \rightarrow book \mid flight \mid meal \mid money$
$Nominal \rightarrow Nominal\ Noun$
$Nominal \rightarrow Nominal\ PP$
$VP \rightarrow book \mid include \mid prefer$
$VP \rightarrow Verb\ NP$
$VP \rightarrow X2\ PP$
$X2 \rightarrow Verb\ NP$
$VP \rightarrow Verb\ PP$
$VP \rightarrow VP\ PP$
$PP \rightarrow Preposition\ NP$

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2 [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon
$Det \rightarrow that \mid this \mid a$
$Noun \rightarrow book \mid flight \mid meal \mid money$
$Pronoun \rightarrow I \mid she \mid me$
$Proper\text{-}Noun \rightarrow Houston \mid TWA$
$Aux \rightarrow does$
$Preposition \rightarrow from \mid to \mid on \mid near \mid through$
$Verb \rightarrow book \mid include \mid prefer$

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* | *include* | *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* | *she* | *me*
NP → *TWA* | *Houston*
NP → *Det Nominal*

Nominal → *book* | *flight* | *meal* | *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* | *include* | *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I *prefer* *a* *flight* *on* *TWA*

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	**VP, X2, S** [1,6]
		Det [2,3]	**NP** [2,4]	[2,5]	**NP** [2,6]
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* / *include* / *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* / *she* / *me*
NP → *TWA* / *Houston*
NP → *Det Nominal*

Nominal → *book* / *flight* / *meal* / *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* / *include* / *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I *prefer* *a* *flight* *on* *TWA*

0 1 2 3 4 5 6

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	**VP, X2, S** [1,6]
		Det [2,3]	**NP** [2,4]	[2,5]	**NP** [2,6]
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

$\mathscr{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → *NP VP*
S → *X1 VP*
X1 → *Aux NP*
S → *book* | *include* | *prefer*
S → *Verb NP*
S → *X2 PP*
S → *Verb PP*
S → *VP PP*

NP → *I* | *she* | *me*
NP → *TWA* | *Houston*
NP → *Det Nominal*

Nominal → *book* | *flight* | *meal* | *money*
Nominal → *Nominal Noun*
Nominal → *Nominal PP*

VP → *book* | *include* | *prefer*
VP → *Verb NP*
VP → *X2 PP*
X2 → *Verb NP*
VP → *Verb PP*
VP → *VP PP*

PP → *Preposition NP*

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → *that* | *this* | *a*
Noun → *book* | *flight* | *meal* | *money*
Pronoun → *I* | *she* | *me*
Proper-Noun → *Houston* | *TWA*
Aux → *does*
Preposition → *from* | *to* | *on* | *near* | *through*
Verb → *book* | *include* | *prefer*

I *prefer* *a* *flight* *on* *TWA*

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	[0,6]
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	**VP, X2, S** [1,6]
		Det [2,3]	**NP** [2,4]	[2,5]	**NP** [2,6]
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

$\mathscr{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	S [0,6]
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	S [0,6]
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathscr{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	**S** [0,2]	[0,3]	**S** [0,4]	[0,5]	**S** [0,6]
	Verb, VP, S [1,2]	[1,3]	**VP, X2, S** [1,4]	[1,5]	**VP, X2, S** [1,6]
		Det [2,3]	**NP** [2,4]	[2,5]	**NP** [2,6]
			Noun, Nom [3,4]	[3,5]	**Nom** [3,6]
				Prep [4,5]	**PP** [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book | include | prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I | she | me
NP → TWA | Houston
NP → Det Nominal

Nominal → book | flight | meal | money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book | include | prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	S [0,6]
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

Lexicon

Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

NP, Pronoun [0,1]	S [0,2]	[0,3]	S [0,4]	[0,5]	S [0,6]
	Verb, VP, S [1,2]	[1,3]	VP, X2, S [1,4]	[1,5]	VP, X2, S [1,6]
		Det [2,3]	NP [2,4]	[2,5]	NP [2,6]
			Noun, Nom [3,4]	[3,5]	Nom [3,6]
				Prep [4,5]	PP [4,6]
					NNP, NP [5,6]

$\mathcal{L}_1$ Grammar

S → NP VP
S → X1 VP
X1 → Aux NP
S → book / include / prefer
S → Verb NP
S → X2 PP
S → Verb PP
S → VP PP

NP → I / she / me
NP → TWA / Houston
NP → Det Nominal

Nominal → book / flight / meal / money
Nominal → Nominal Noun
Nominal → Nominal PP

VP → book / include / prefer
VP → Verb NP
VP → X2 PP
X2 → Verb NP
VP → Verb PP
VP → VP PP

PP → Preposition NP

Lexicon

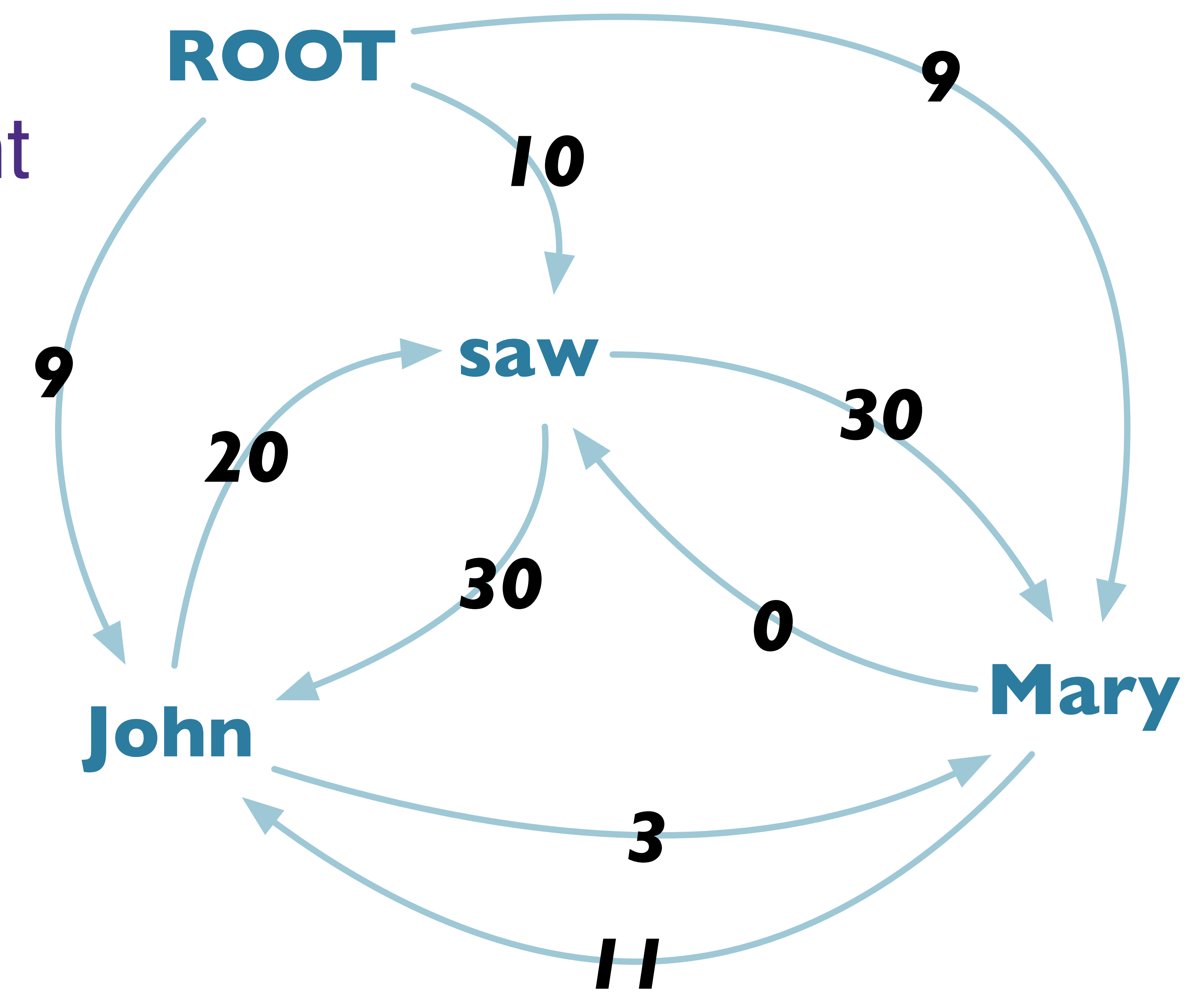
Det → that | this | a
Noun → book | flight | meal | money
Pronoun → I | she | me
Proper-Noun → Houston | TWA
Aux → does
Preposition → from | to | on | near | through
Verb → book | include | prefer

I prefer a flight on TWA

0 1 2 3 4 5 6

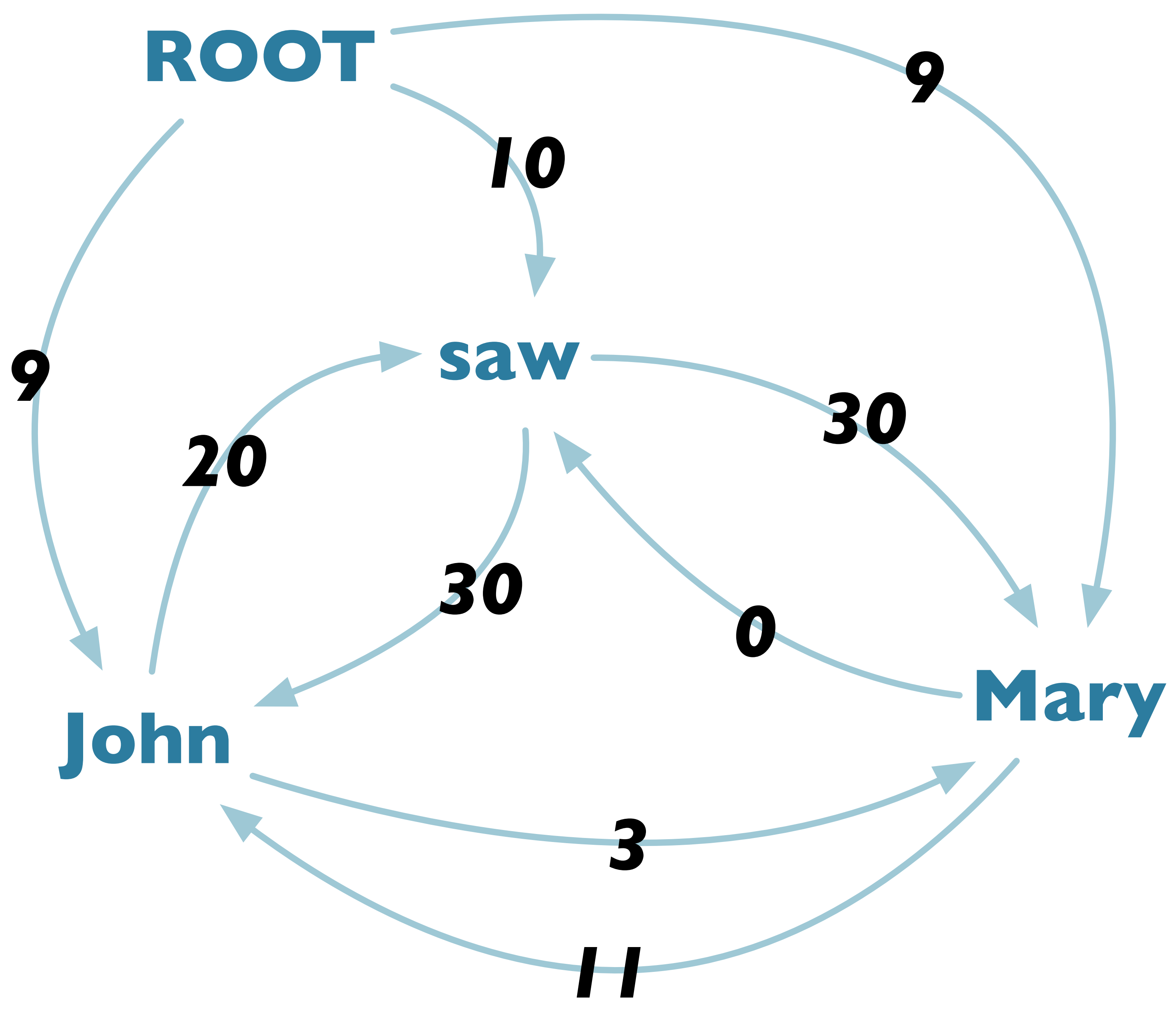
Maximum Spanning Tree

- McDonald et al, 2005 use variant of Chu-Liu-Edmonds algorithm for MST (CLE)
- Sketch of algorithm:
  - For each node, greedily select incoming arc with max weight
  - If the resulting set of arcs forms a tree, this is the MST.
  - If not, there must be a cycle.
    - "Contract" the cycle: Treat it as a single vertex
    - Recalculate weights into/out of the new vertex
    - Recursively do MST algorithm on resulting graph
- Running time: naïve: O(n3); Tarjan: O(n2)
  - Applicable to non-projective graphs



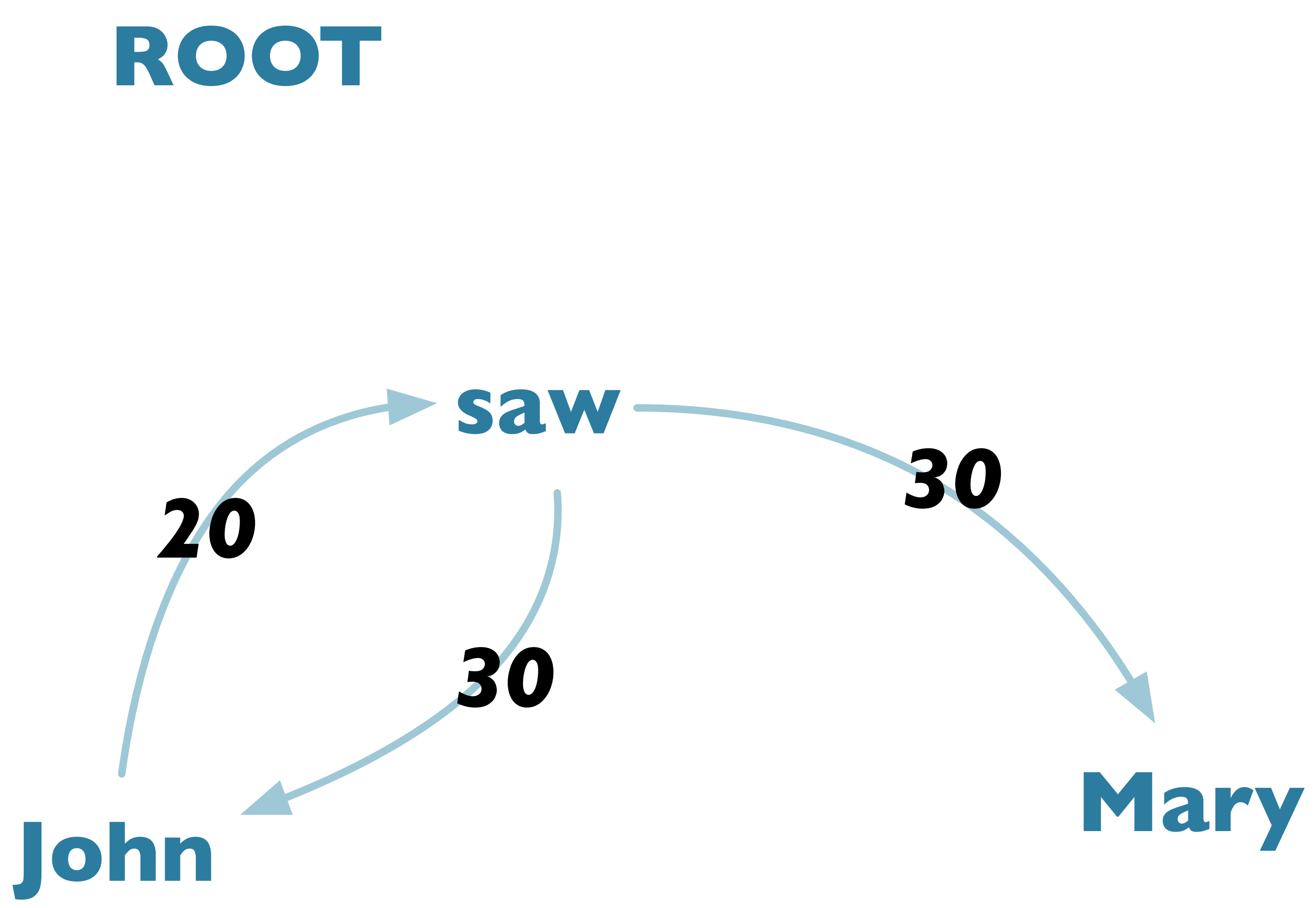
Step 1 & 2

- Find, for each word, the highest scoring incoming edge.



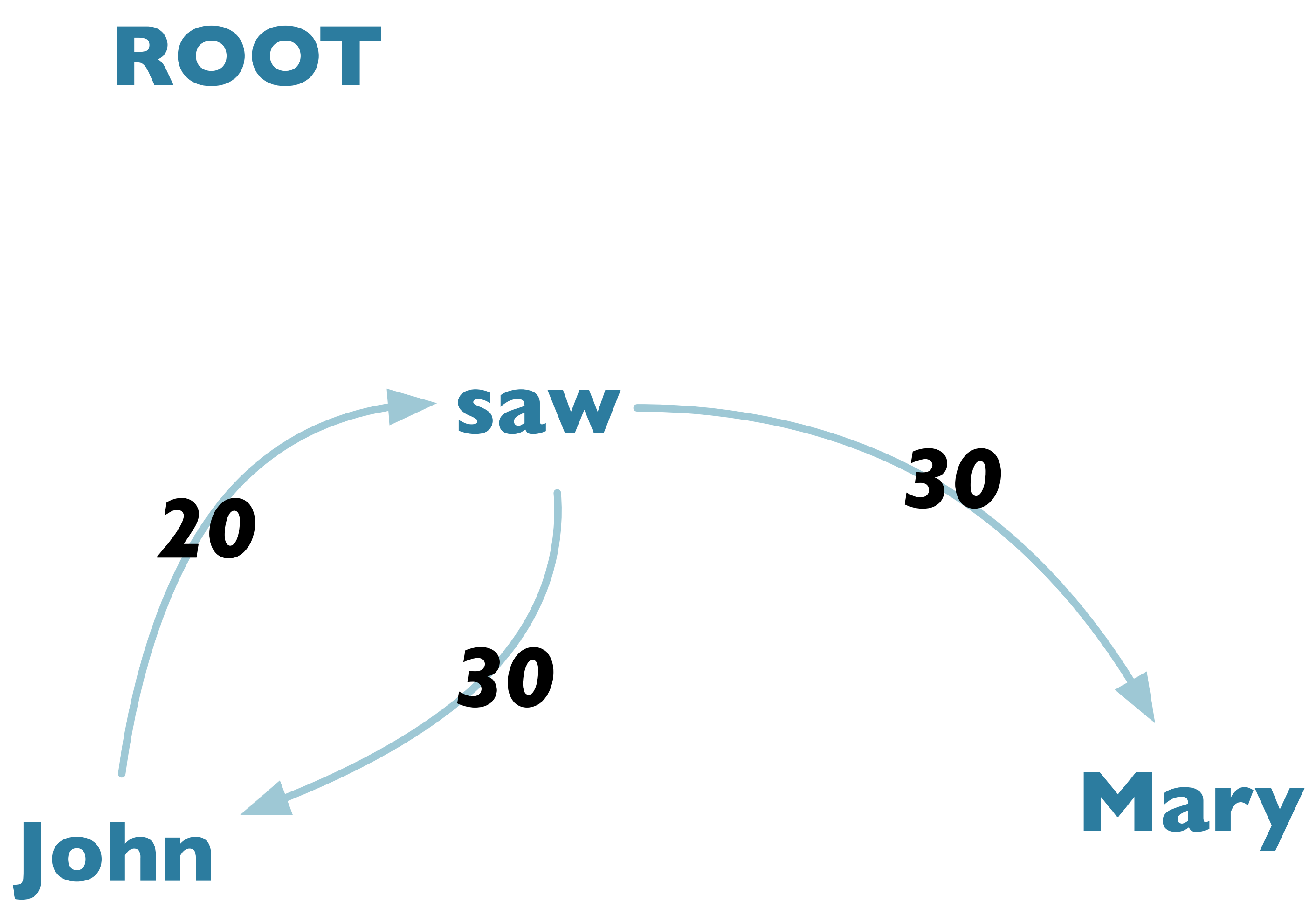
Step 1 & 2

- Find, for each word, the highest scoring incoming edge.



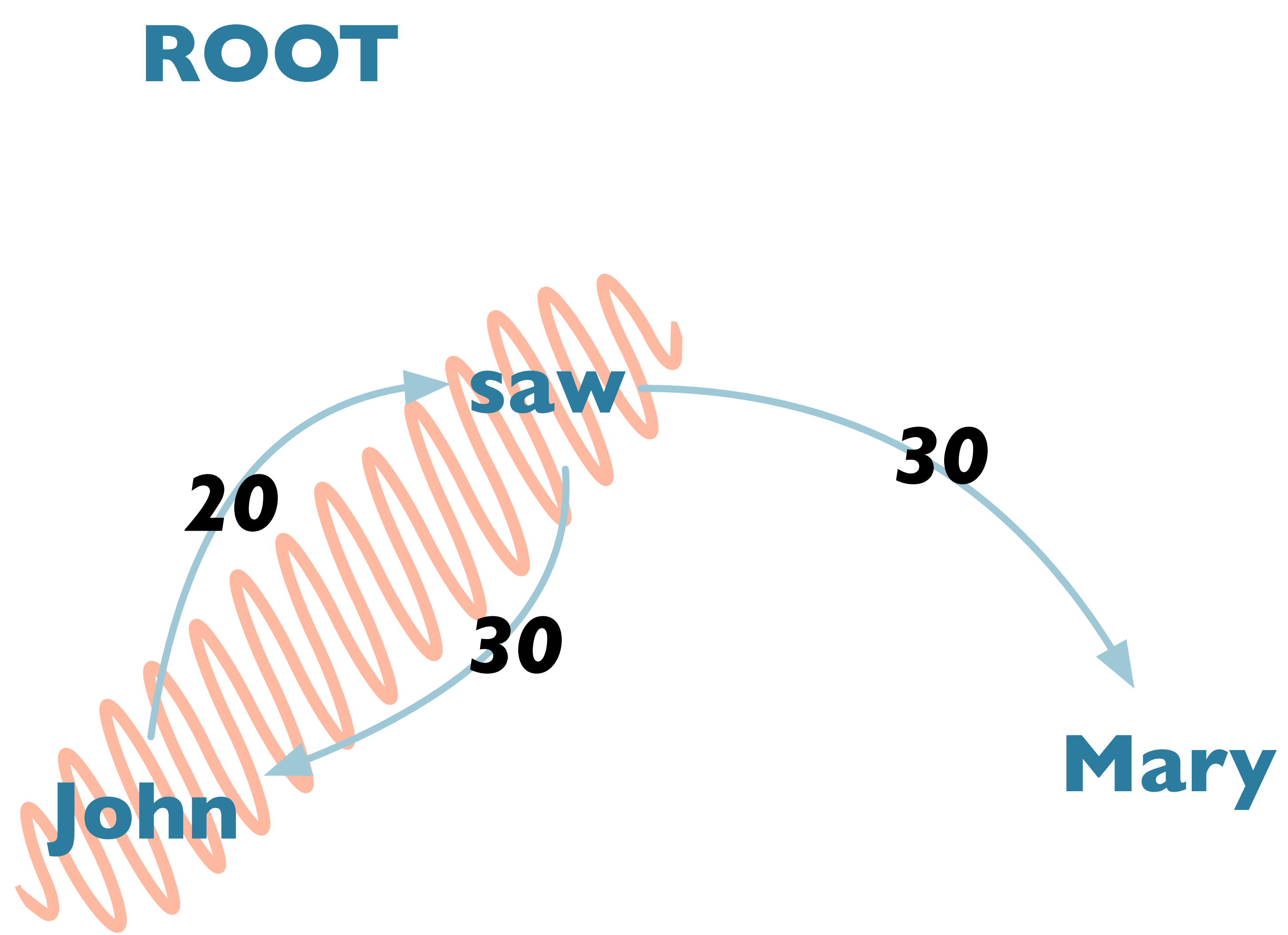
Step 1 & 2

- Find, for each word, the highest scoring incoming edge.
- Is it a tree?



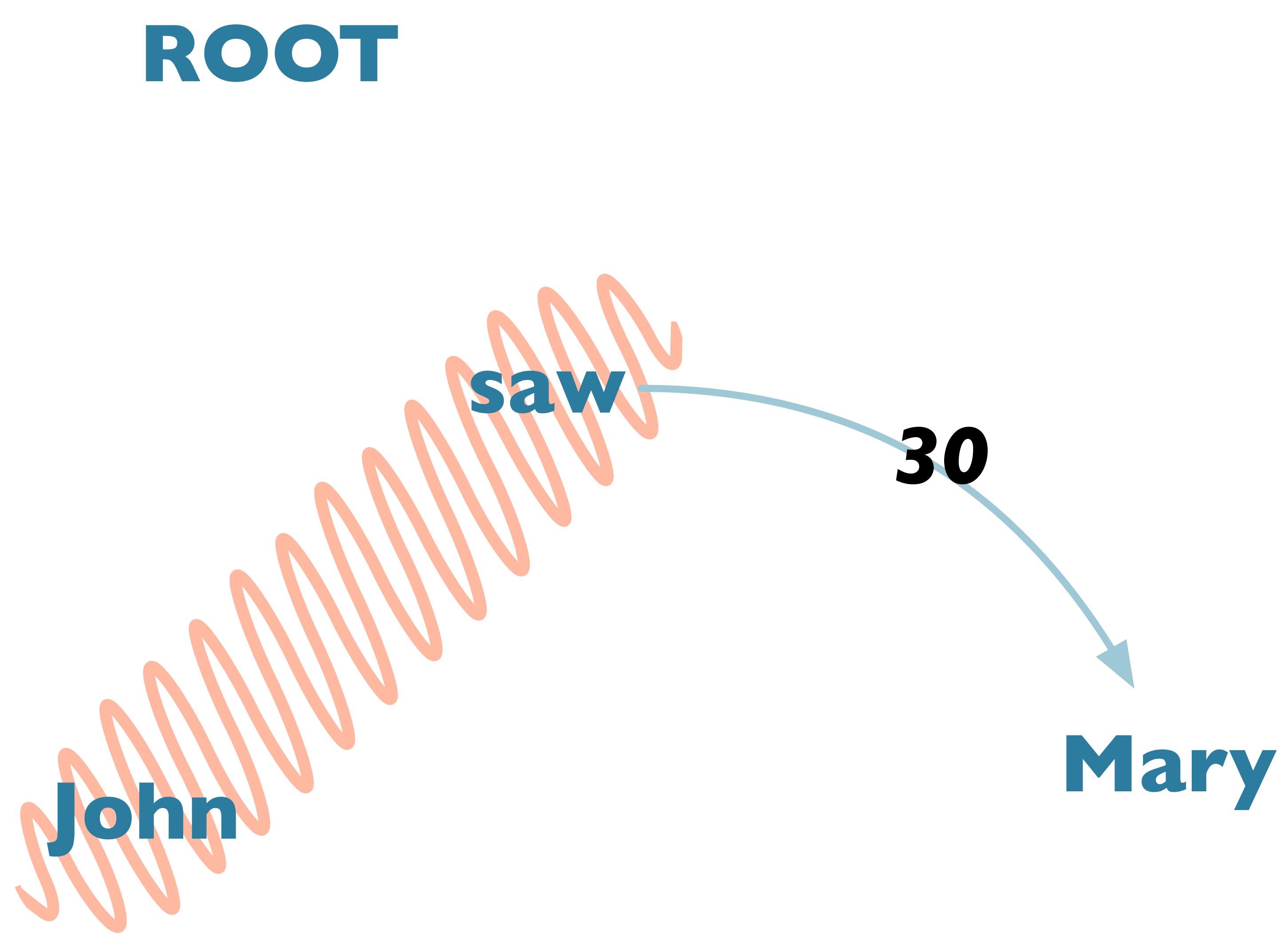
Step 1 & 2

- Find, for each word, the highest scoring incoming edge.
- Is it a tree?
  - No, there's a cycle.



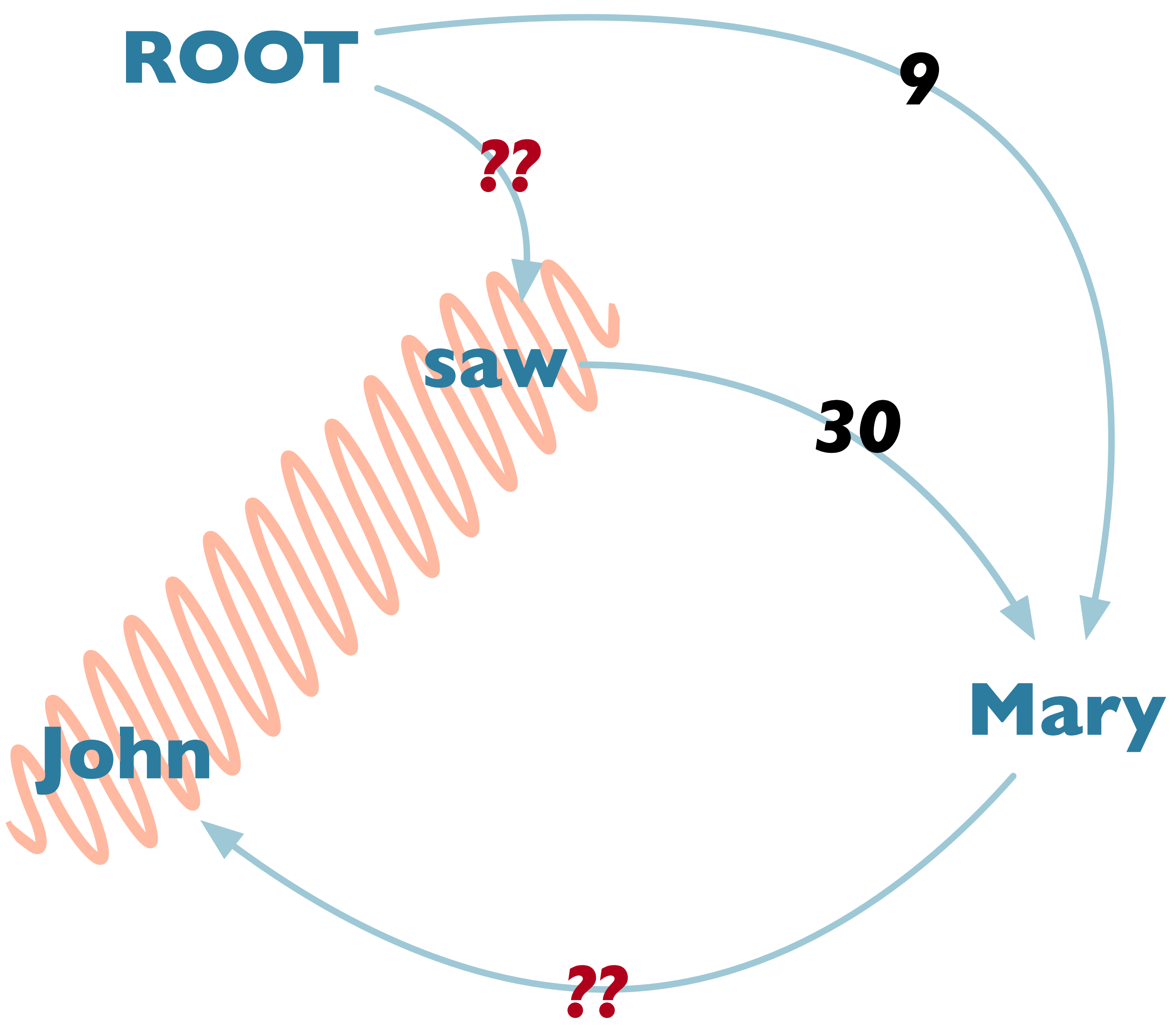
Step 1 & 2

- Find, for each word, the highest scoring incoming edge.
- Is it a tree?
  - No, there's a cycle.
- Collapse the cycle



Step 1 & 2

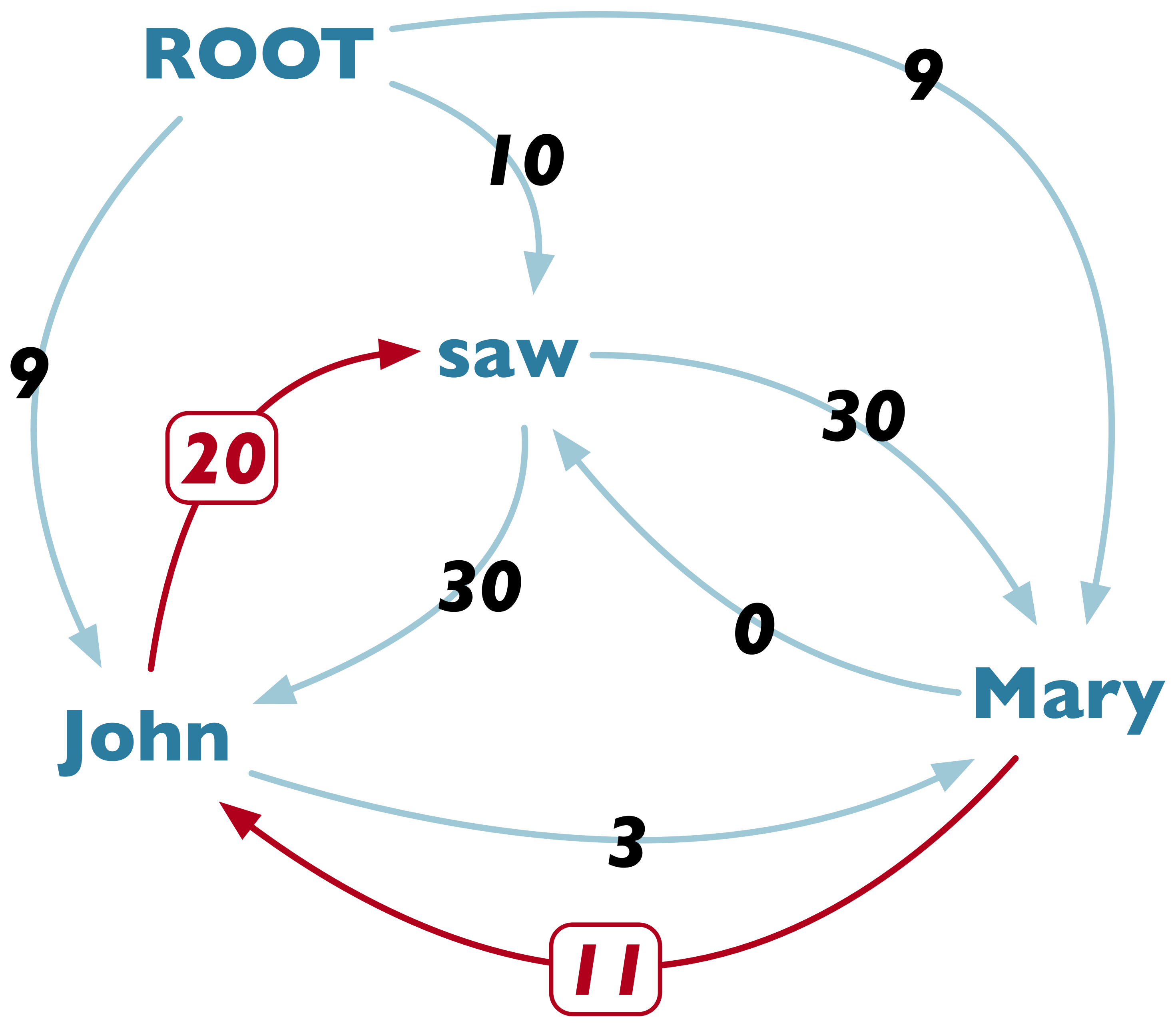
- Find, for each word, the highest scoring incoming edge.
- Is it a tree?
  - No, there's a cycle.
- Collapse the cycle
- And re-examine the edges again



Calculating Weights for Collapsed Vertex

- Since there's a cycle:
  - Contract cycle & reweight
  - John+saw as single vertex
  - Calculate weights in & out as:
- Recurse

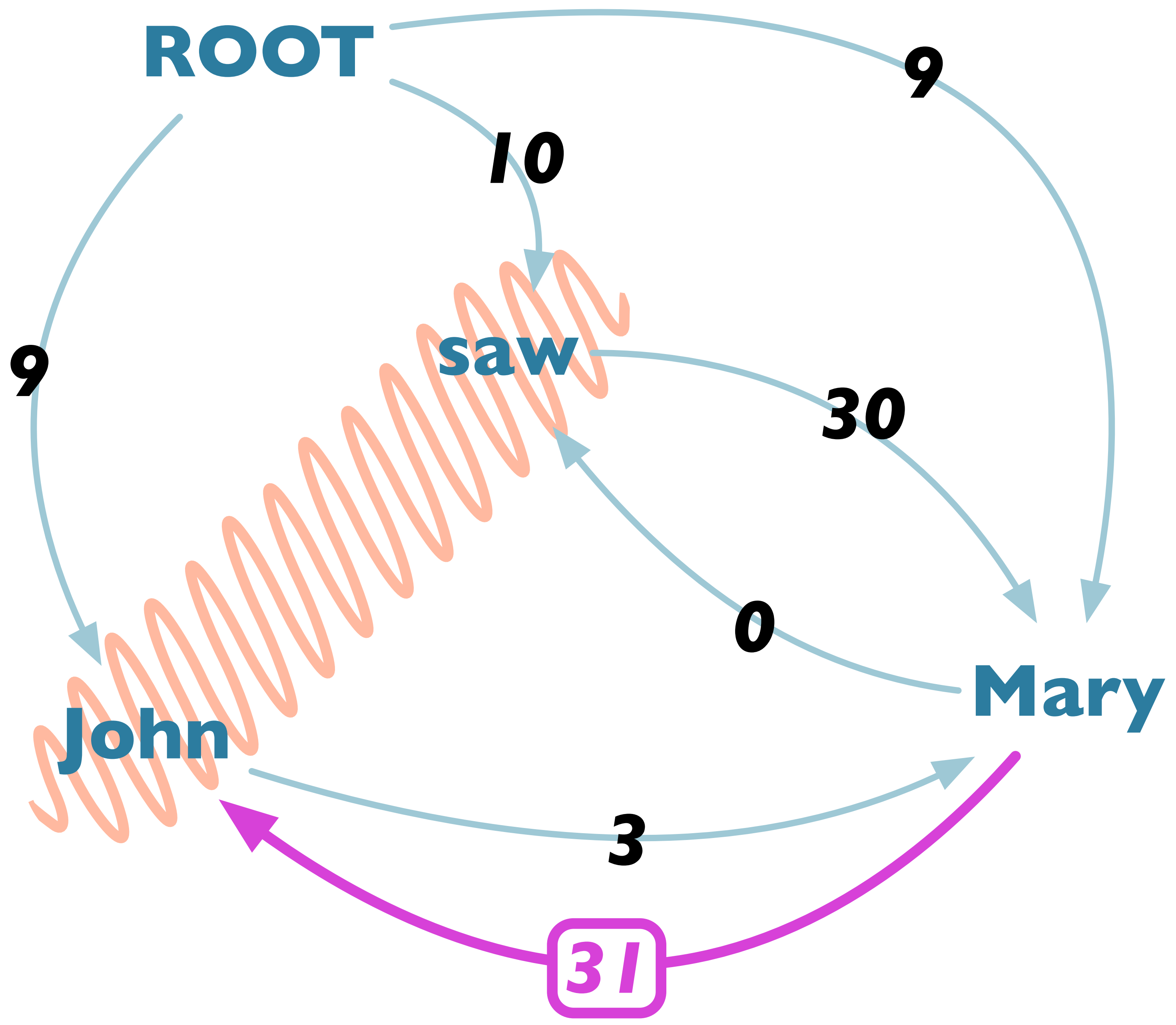
s(Mary, C) 11 + 20 = 31



Calculating Weights for Collapsed Vertex

- Since there's a cycle:
  - Contract cycle & reweight
  - John+saw as single vertex
  - Calculate weights in & out as:
- Recurse

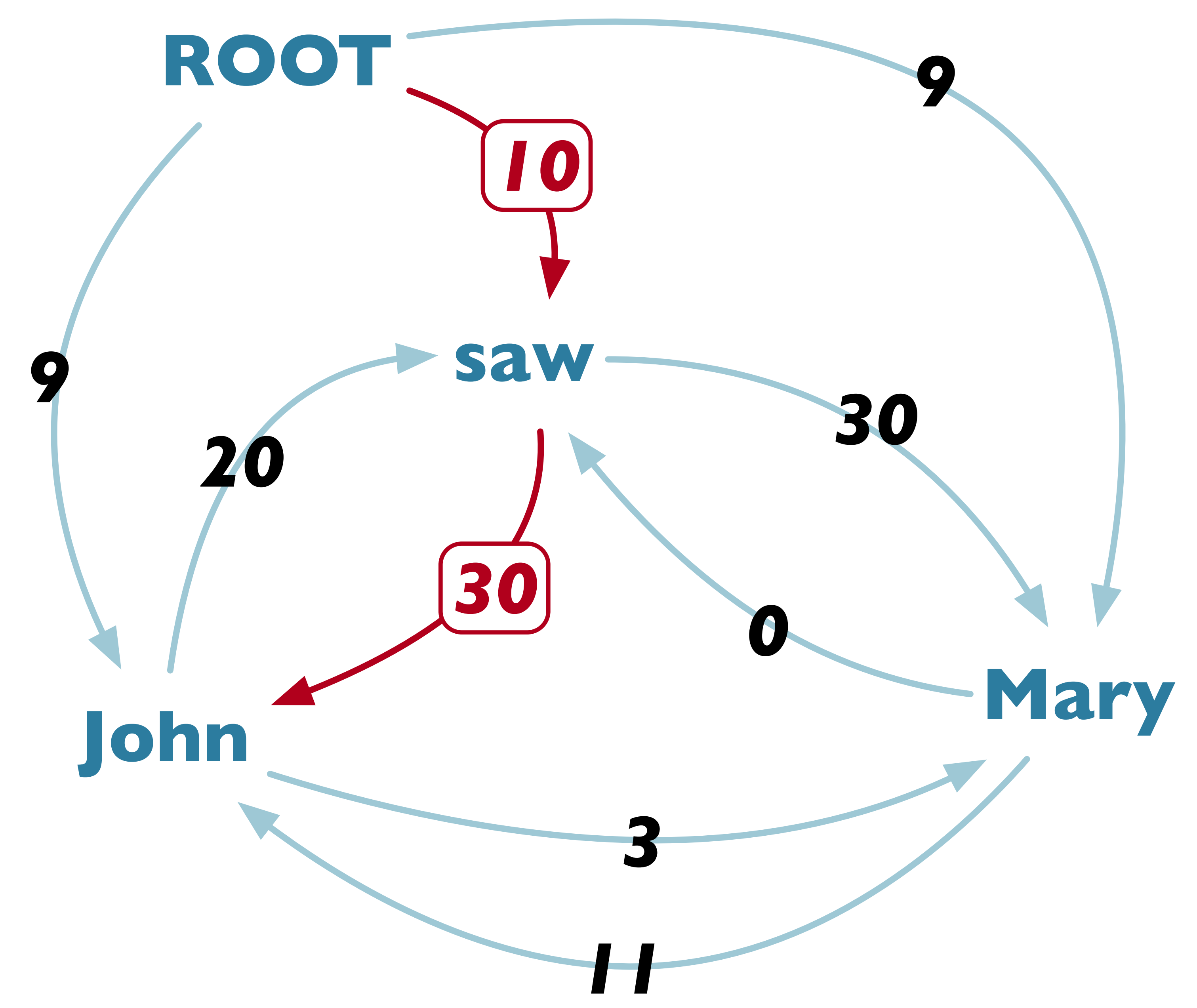
s(Mary, C) 11 + 20 = 31



Calculating Weights for Collapsed Vertex

- Since there's a cycle:
  - Contract cycle & reweight
  - John+saw as single vertex
  - Calculate weights in & out as:
- Recurse

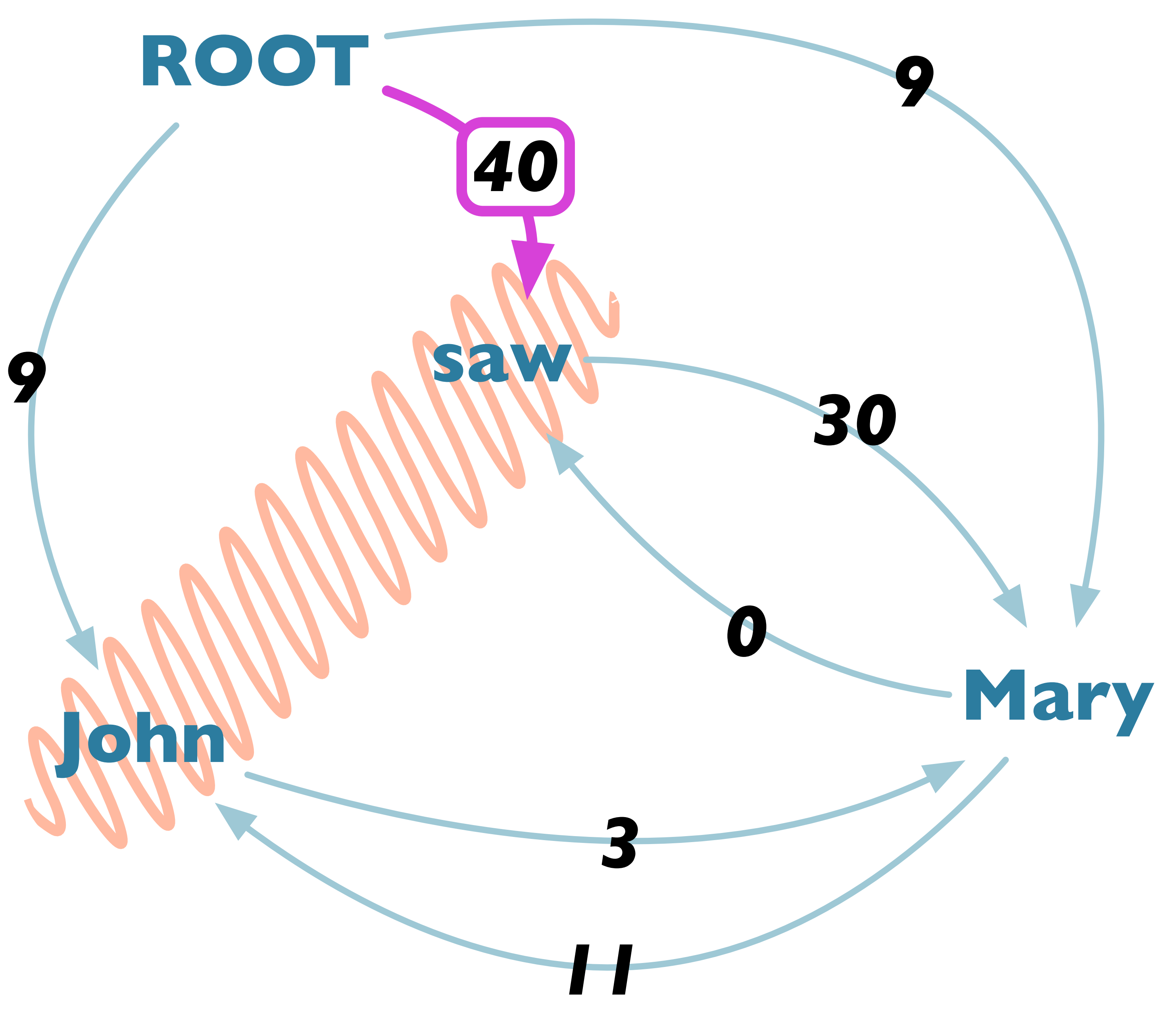
s(ROOT, C) 10 + 30 = 40



Calculating Weights for Collapsed Vertex

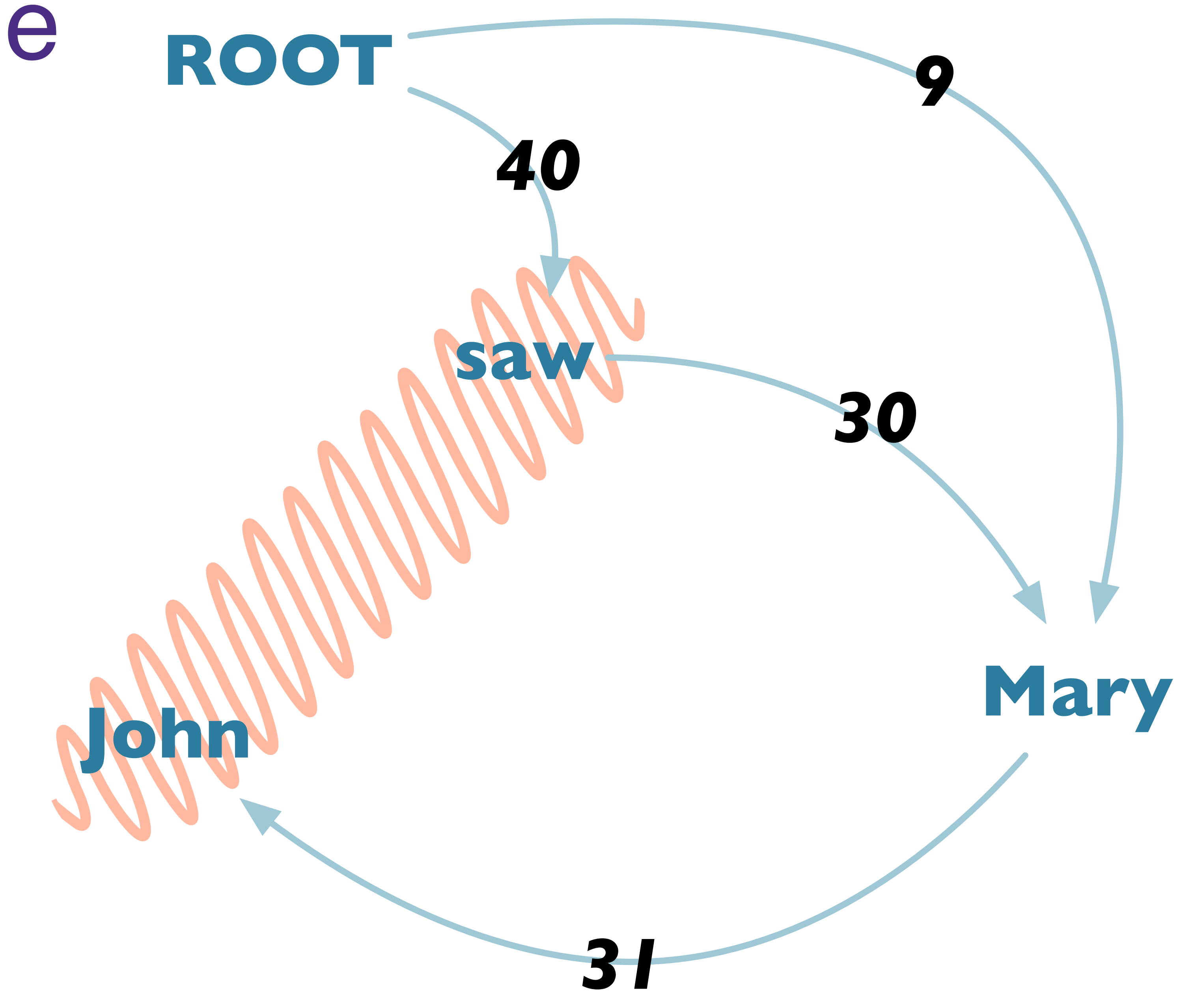
- Since there's a cycle:
  - Contract cycle & reweight
  - John+saw as single vertex
  - Calculate weights in & out as:
- Recurse

s(ROOT, C) 10 + 30 = 40



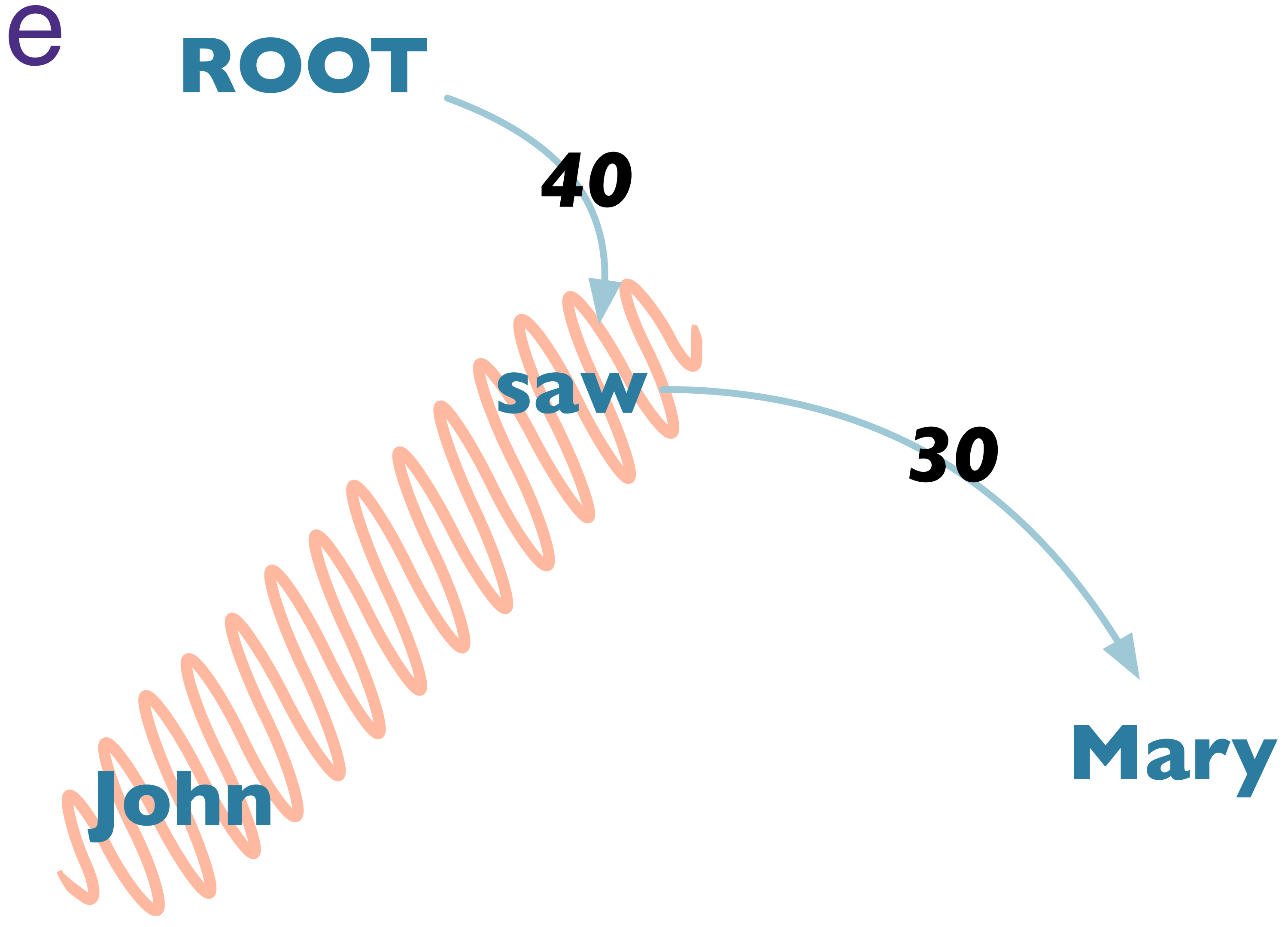
Step 3

- With cycle collapsed, recurse on step 1:
- Keep highest weighted incoming edge for each edge



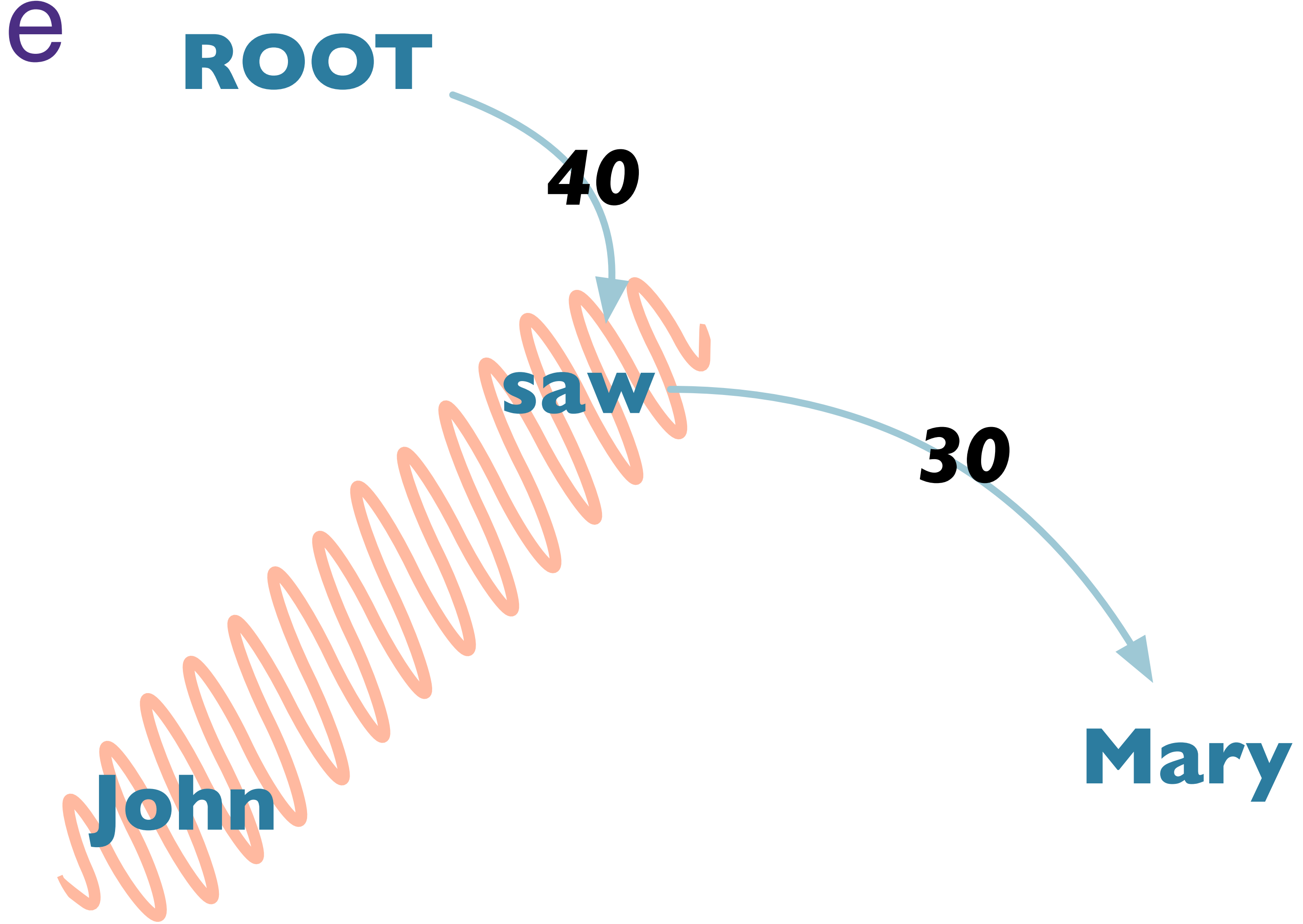
Step 3

- With cycle collapsed, recurse on step 1:
- Keep highest weighted incoming edge for each edge



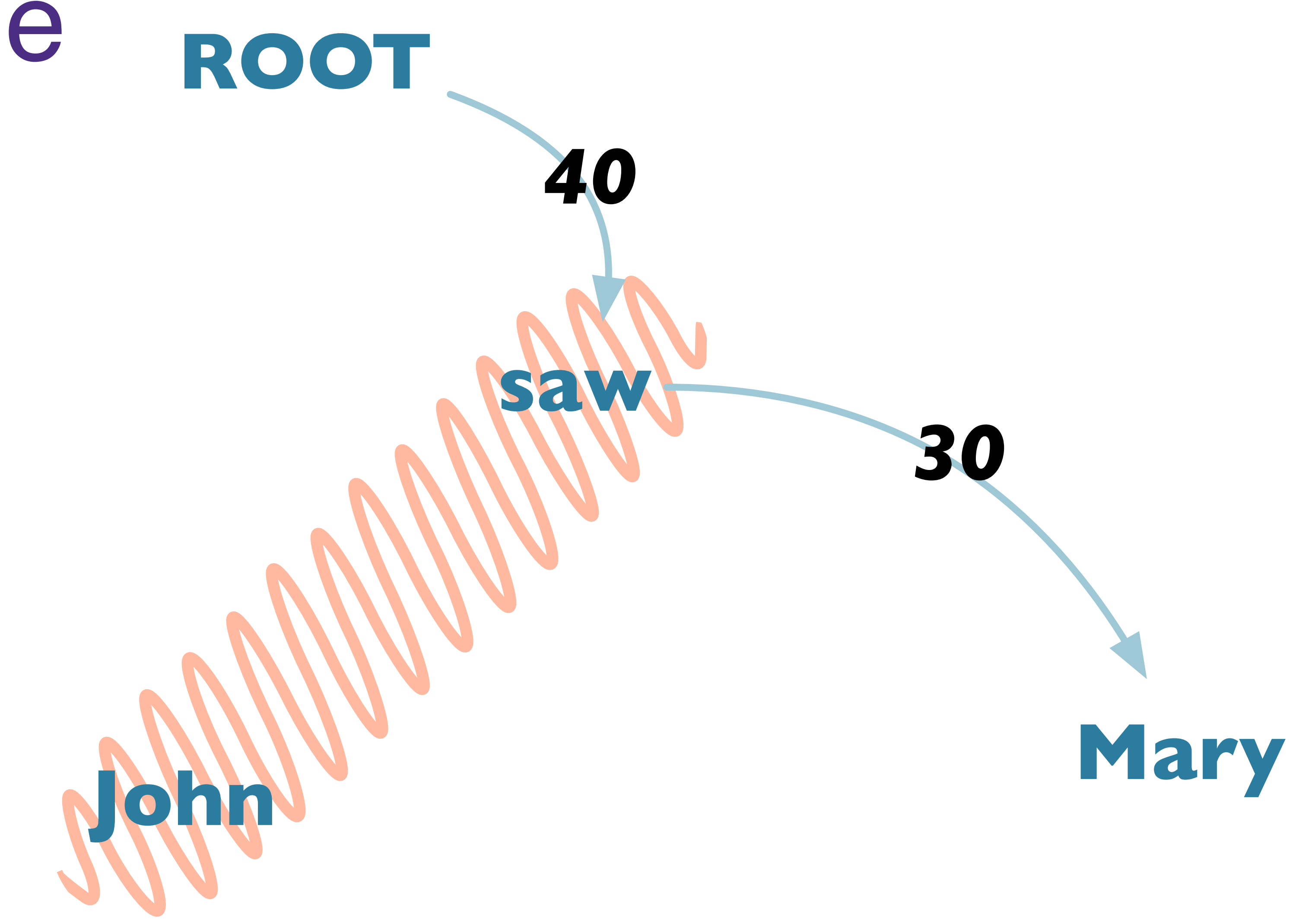
Step 3

- With cycle collapsed, recurse on step 1:
- Keep highest weighted incoming edge for each edge
- Is it a tree?



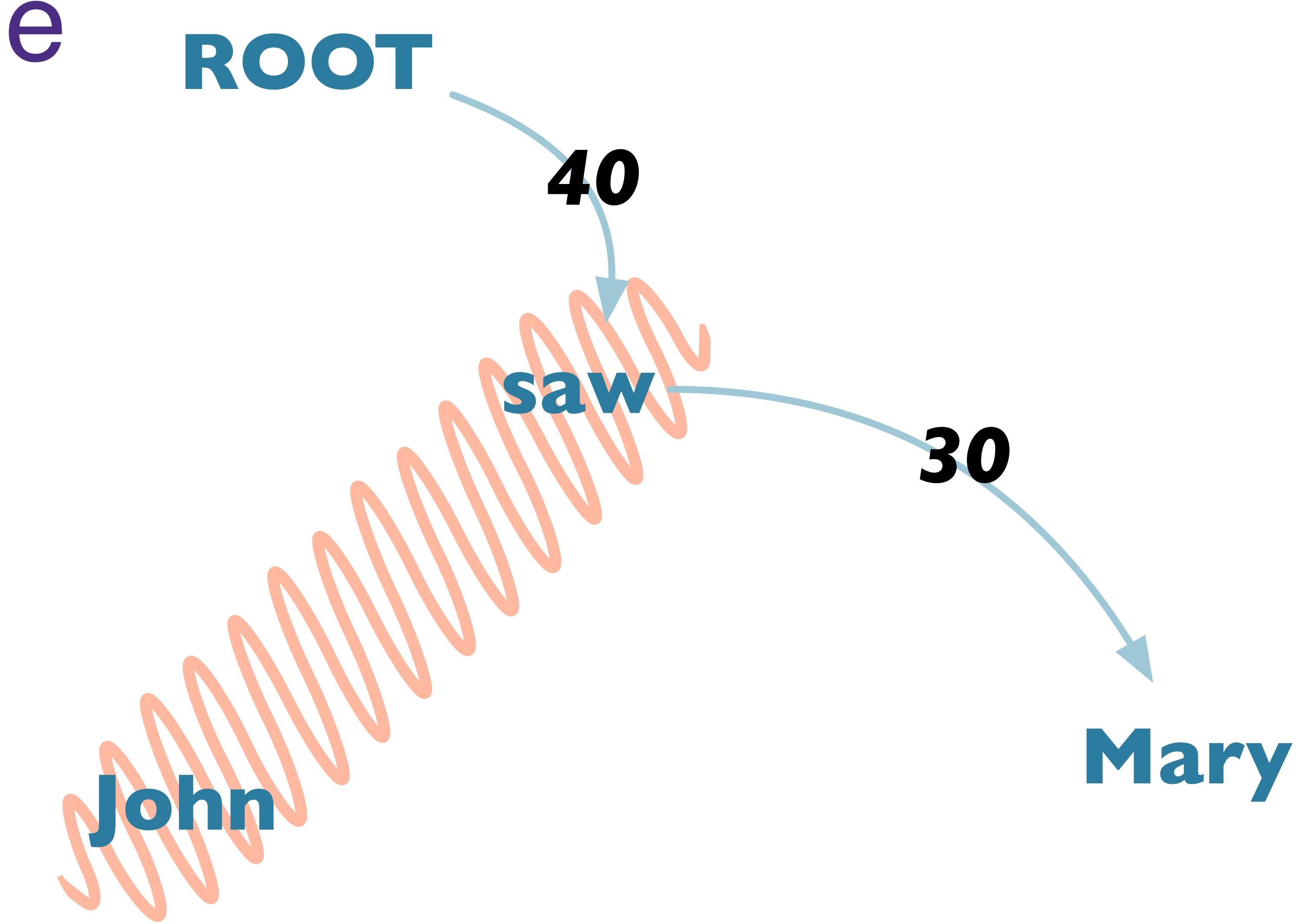
Step 3

- With cycle collapsed, recurse on step 1:
- Keep highest weighted incoming edge for each edge
- Is it a tree?
  - Yes!



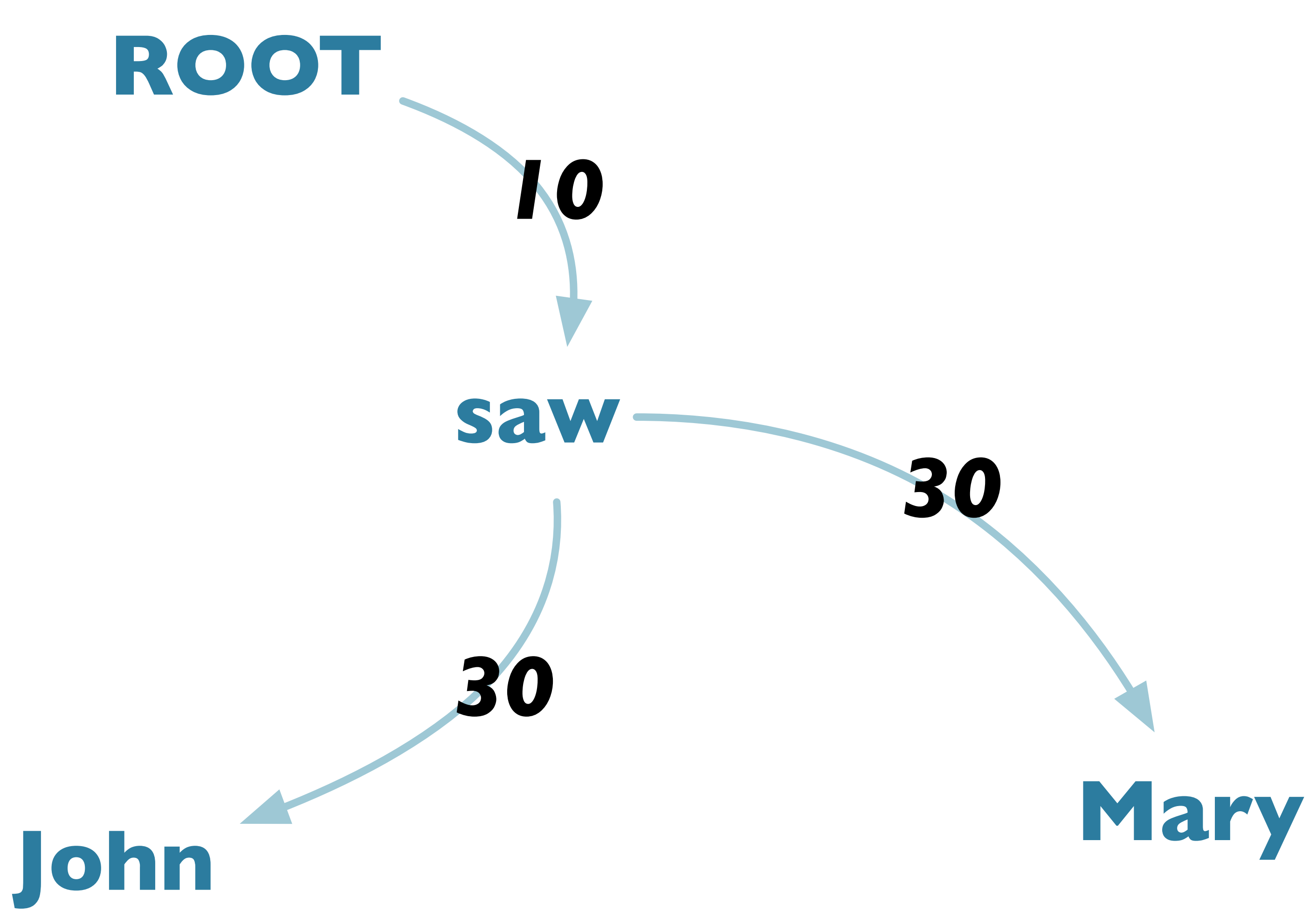
Step 3

- With cycle collapsed, recurse on step 1:
- Keep highest weighted incoming edge for each edge
- Is it a tree?
    - Yes!
    - …but must recover collapsed portions.



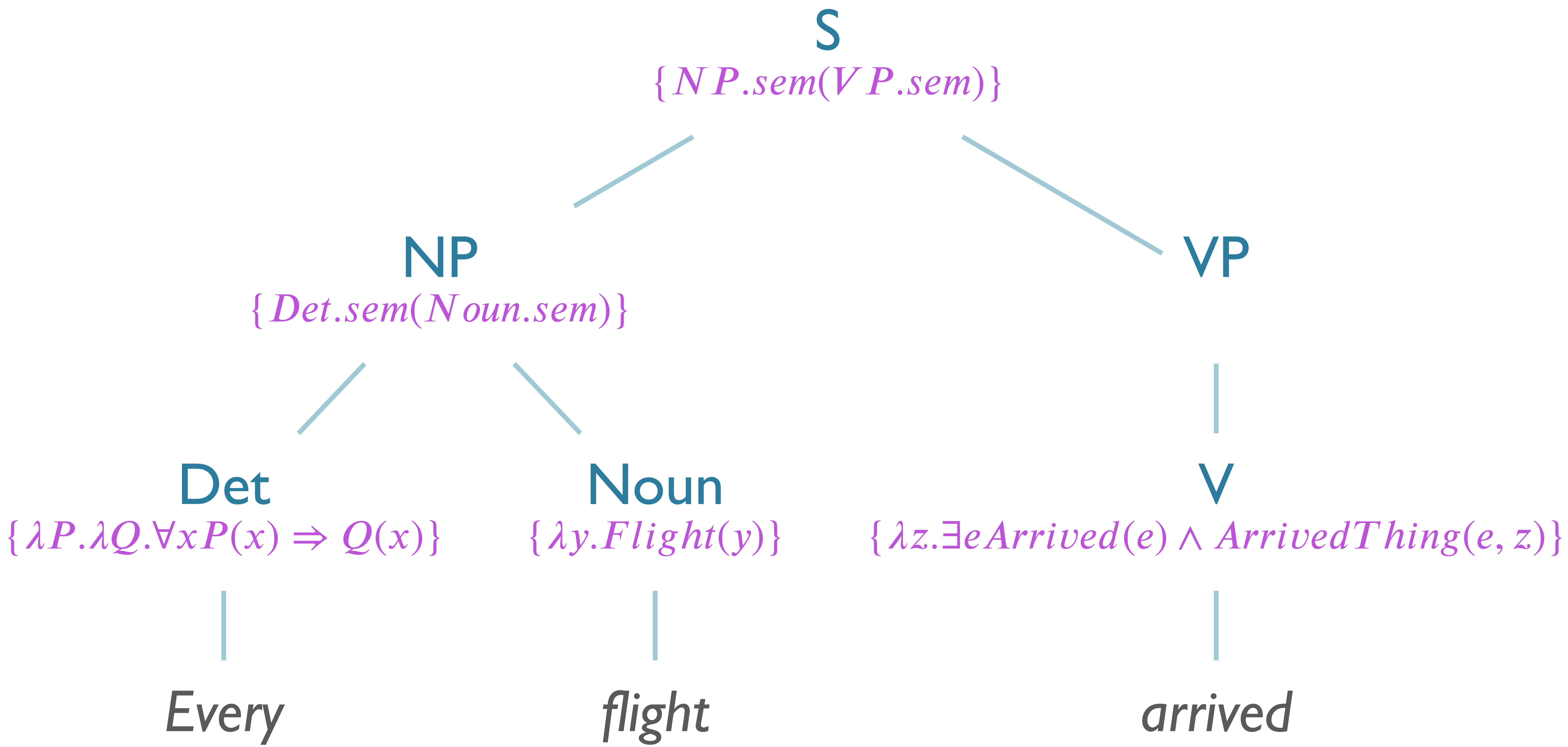
Step 3

- With cycle collapsed, recurse on step 1:
- Keep highest weighted incoming edge for each edge
- Is it a tree?
  - Yes!
  - …but must recover collapsed portions.

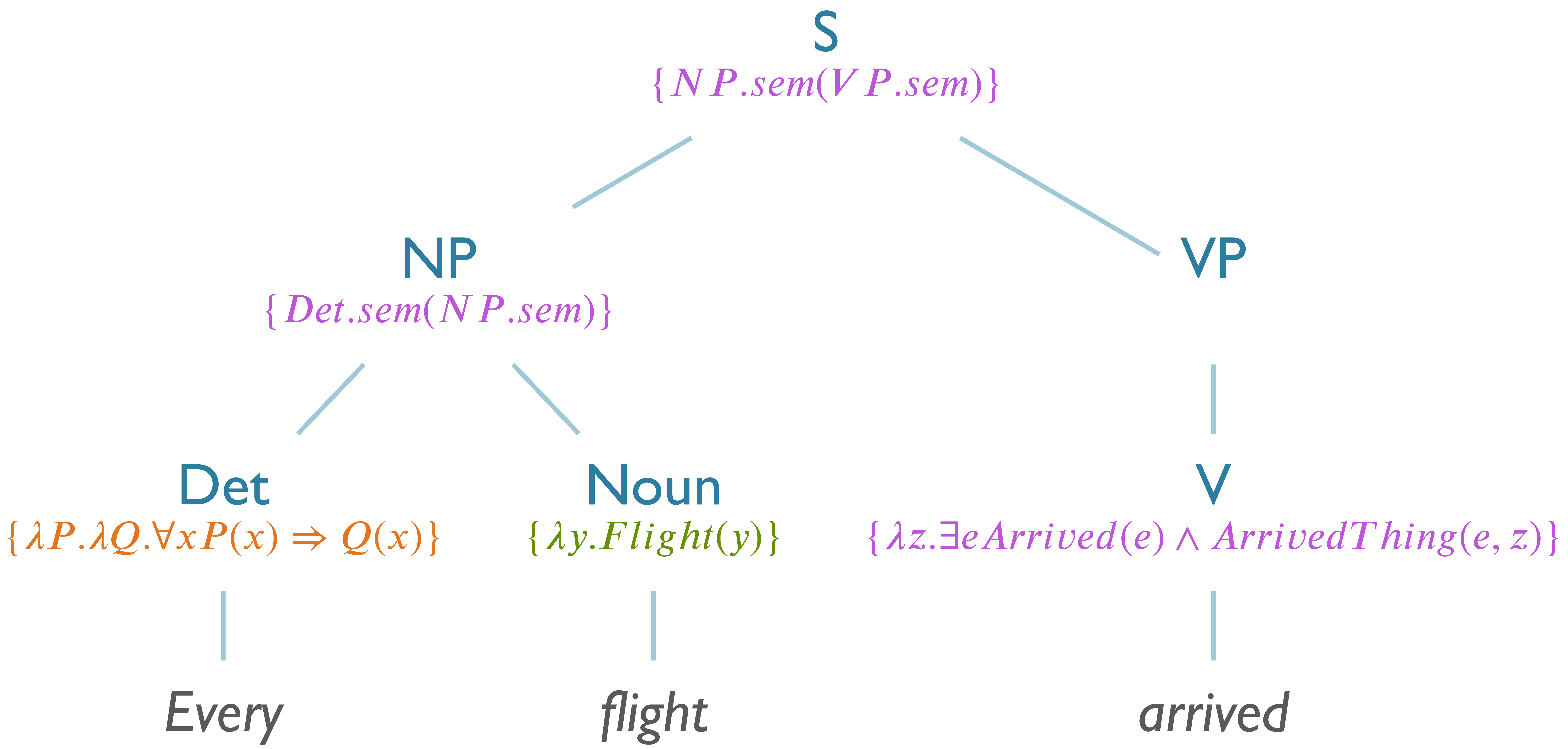


Semantics

- First order logic + lambda calculus
- Neo-Davidsonian event semantics
- Parsing via features
- Distributional Semantics + word embeddings
- Word Sense Disambiguation
- Semantic Role Labeling

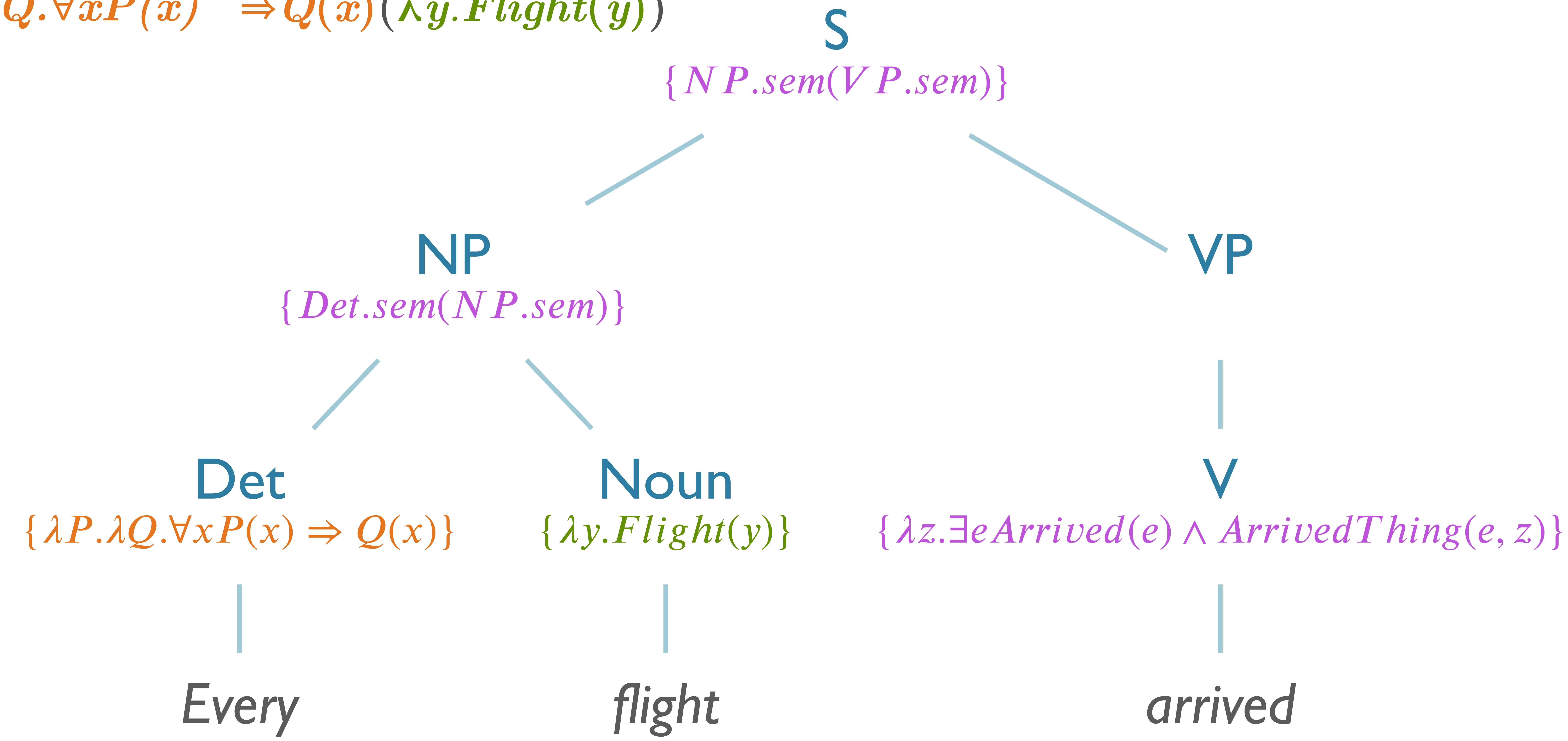

S
$\{NP.sem(VP.sem)\}$
NP
$\{Det.sem(Noun.sem)\}$
VP
Det
$\{\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)\}$
Noun
$\{\lambda y.Flight(y)\}$
V
$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$
Every
flight
arrived

$$NP \rightarrow Det.sem(NP.sem)$$



$$NP \rightarrow Det.sem(NP.sem)$$

$$\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)(\lambda y.Flight(y))$$



$$NP \rightarrow Det.sem(NP.sem)$$
$$\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)(\lambda y.Flight(y))$$
$$\lambda Q.\forall x \lambda y.Flight(y)(x) \Rightarrow Q(x)$$

S
$\{NP.sem(VP.sem)\}$

NP
$\{Det.sem(NP.sem)\}$

VP

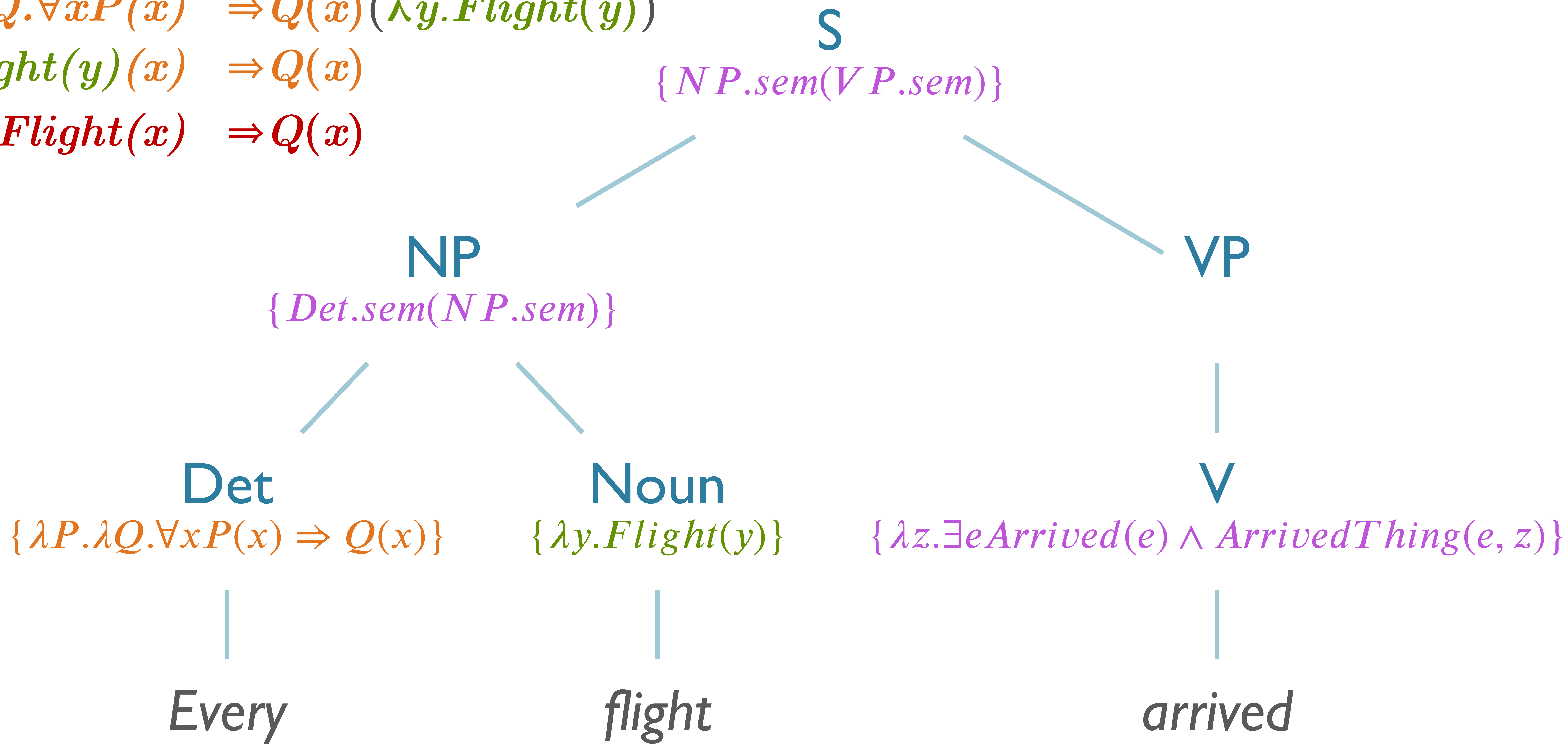
Det
$\{\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)\}$

Noun
$\{\lambda y.Flight(y)\}$

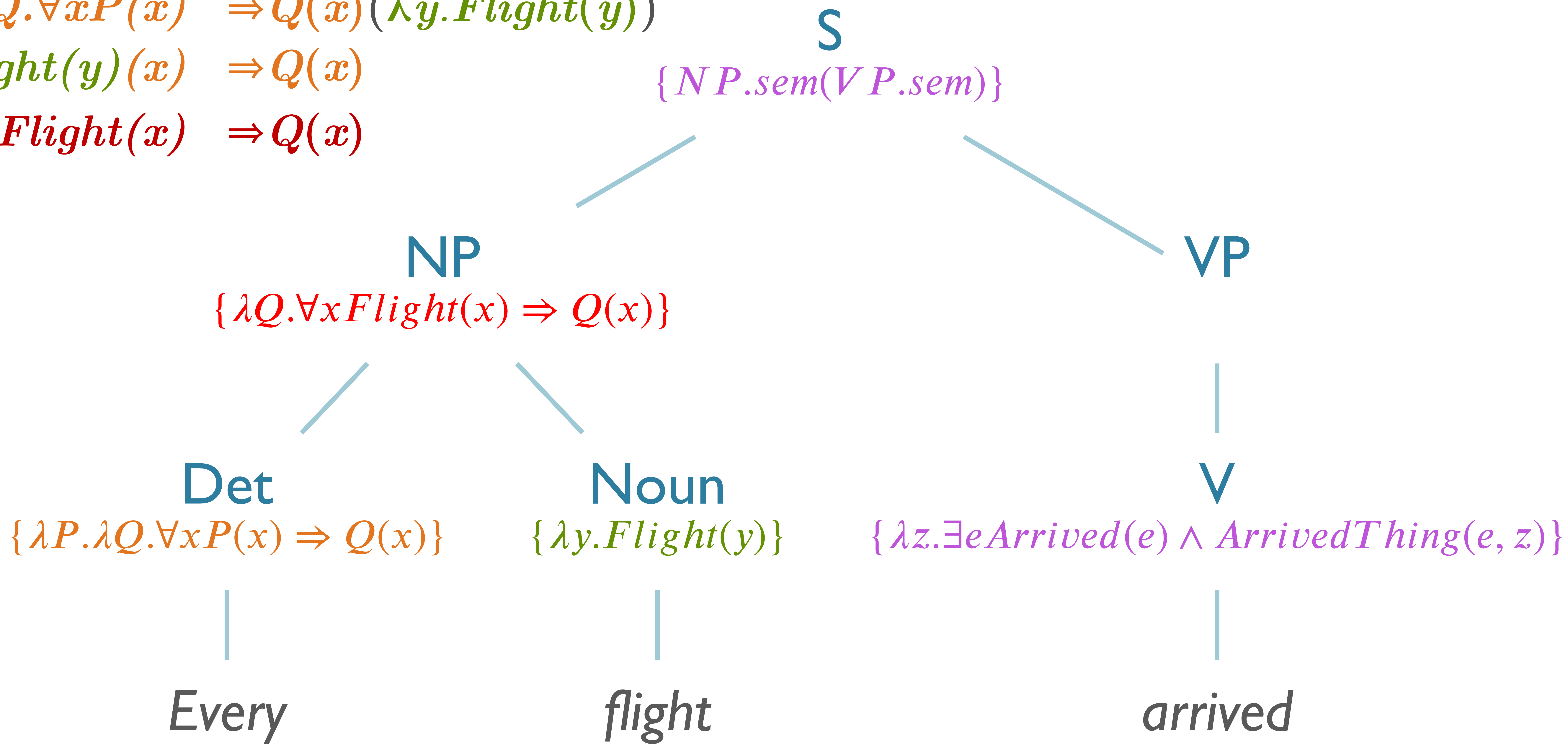
V
$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$

Every *flight* *arrived*

$$NP \rightarrow Det.sem(NP.sem)$$
$$\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)(\lambda y.Flight(y))$$
$$\lambda Q.\forall x \lambda y.Flight(y)(x) \Rightarrow Q(x)$$
$$\lambda Q.\forall x Flight(x) \Rightarrow Q(x)$$



$$
\begin{aligned}
NP &\rightarrow Det.sem(NP.sem) \\
\lambda P.\lambda Q.\forall x P(x) &\Rightarrow Q(x)(\lambda y.Flight(y)) \\
\lambda Q.\forall x \lambda y.Flight(y)(x) &\Rightarrow Q(x) \\
\lambda Q.\forall x Flight(x) &\Rightarrow Q(x)
\end{aligned}
$$



S

$\{NP.sem(VP.sem)\}$

NP

$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$

VP

$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$

Det

$\{\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)\}$

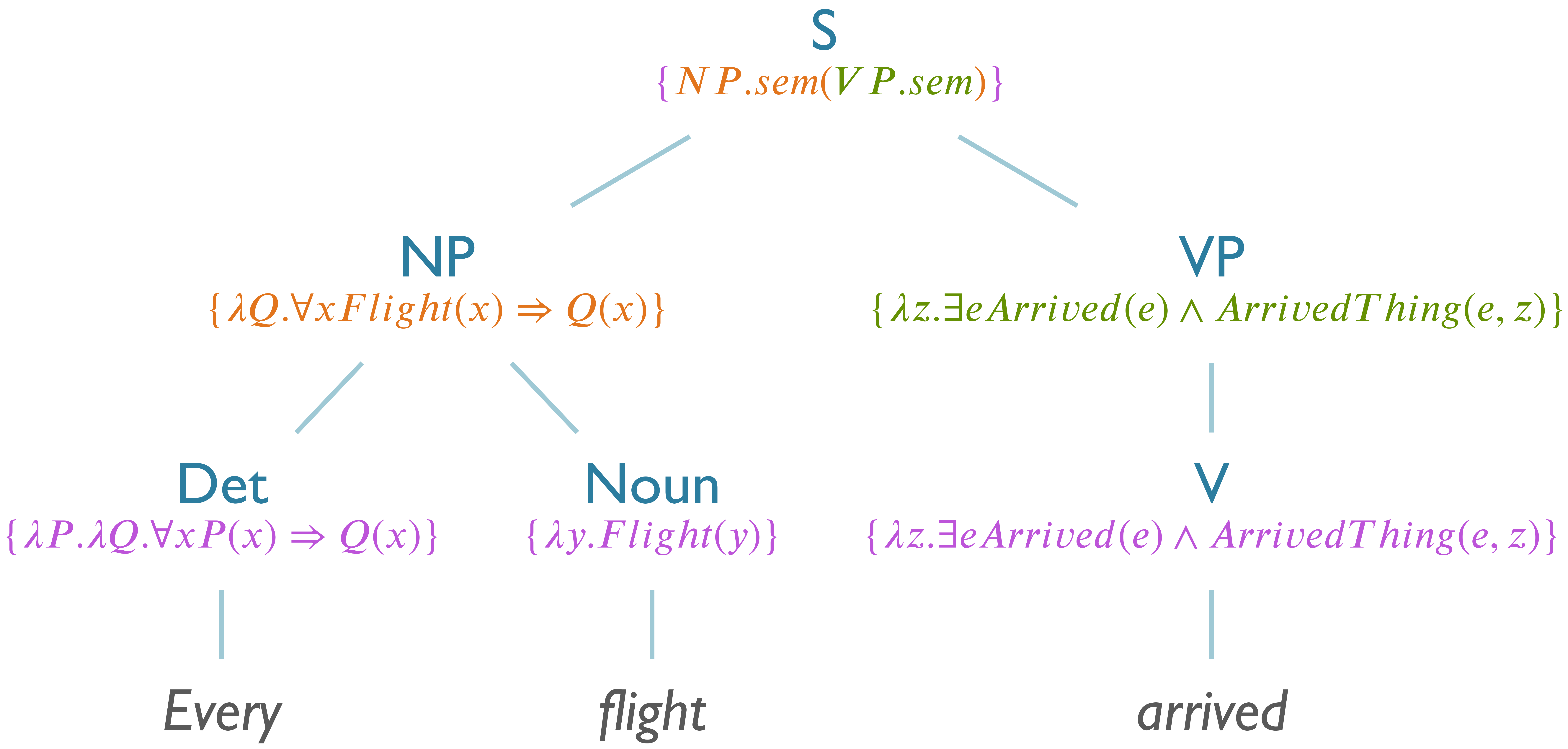
Noun

$\{\lambda y.Flight(y)\}$

V

$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$

Every *flight* *arrived*


S
$\{NP.sem(VP.sem)\}$
NP
$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$
VP
$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$
Det
$\{\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)\}$
Noun
$\{\lambda y.Flight(y)\}$
V
$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$
Every
flight
arrived

S

$\{NP.sem(VP.sem)\}$

NP

$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$

VP

$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z)\}$

S

$$\{\forall x Flight(x) \Rightarrow \exists e Arrived(e) \wedge ArrivedThing(e, x)\}$$

NP

$$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$$

VP

$$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z)\}$$

S

$\{\forall x Flight(x) \Rightarrow \exists e Arrived(e) \wedge ArrivedThing(e, x)\}$

NP

$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$

VP

$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z)\}$

$$\lambda Q.\forall x Flight(x) \Rightarrow Q(x)(\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z))$$

S

$\{\forall x Flight(x) \Rightarrow \exists e Arrived(e) \wedge ArrivedThing(e, x)\}$

NP

$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$

VP

$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z)\}$

$$\lambda Q.\forall x Flight(x) \Rightarrow Q(x)(\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z))$$

$$\forall x Flight(x) \Rightarrow \lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z)(x)$$

S

$\{\forall x Flight(x) \Rightarrow \exists e Arrived(e) \wedge ArrivedThing(e,x)\}$

NP

$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$

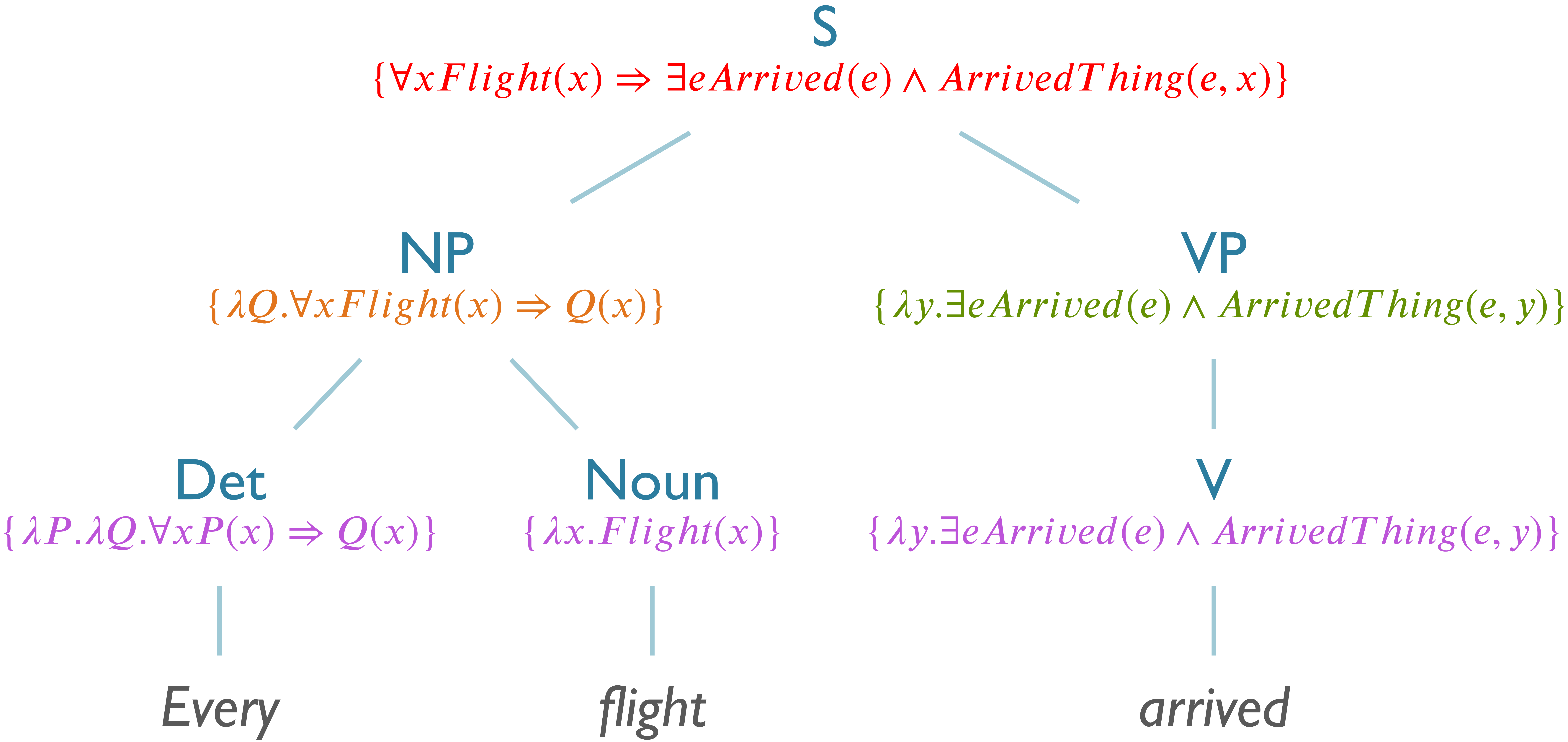
VP

$\{\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e,z)\}$

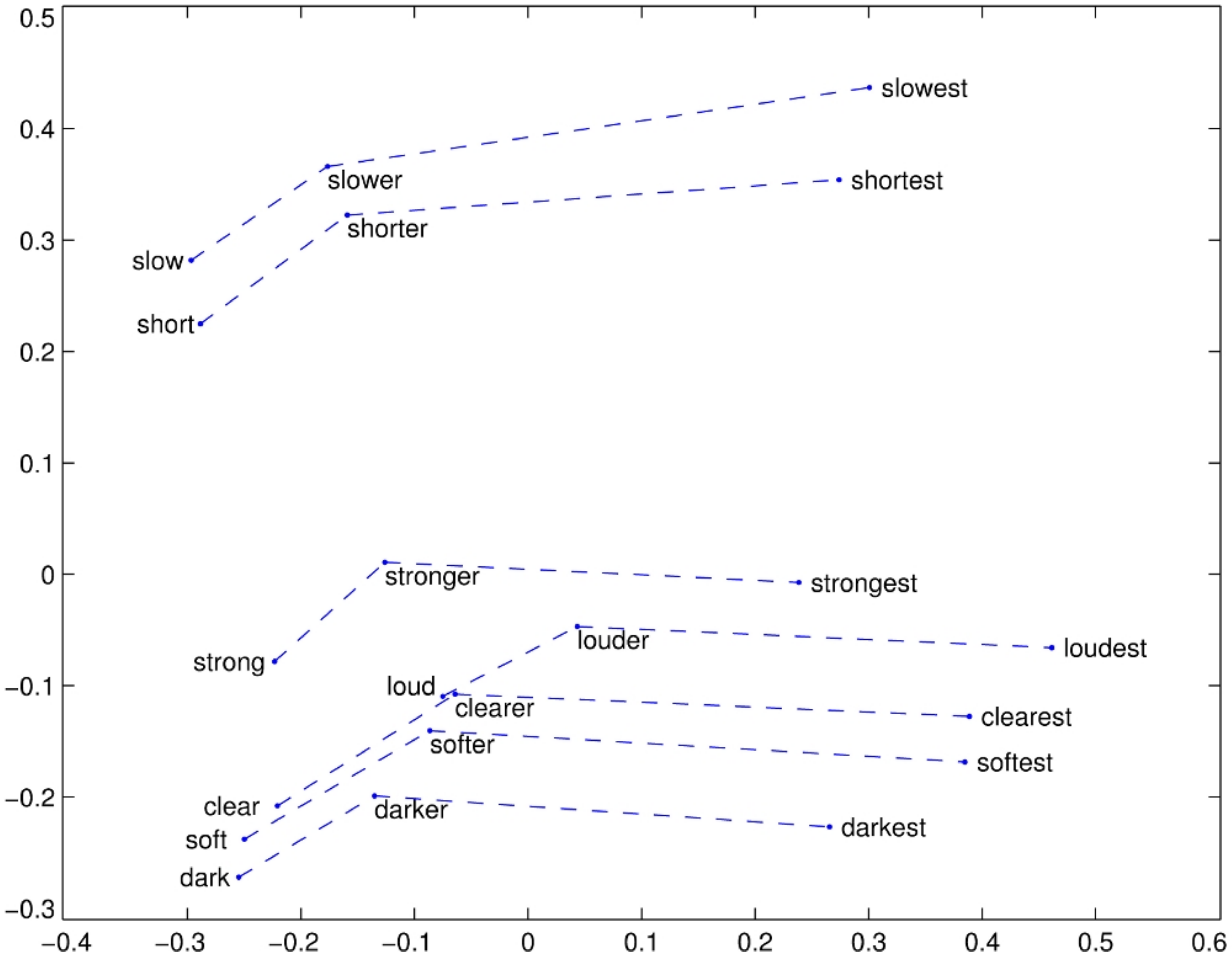
$$\lambda Q.\forall x Flight(x) \Rightarrow Q(x)(\lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z))$$

$$\forall x Flight(x) \Rightarrow \lambda z.\exists e Arrived(e) \wedge ArrivedThing(e, z)(x)$$

$$\forall x Flight(x) \Rightarrow \exists e Arrived(e) \wedge ArrivedThing(e, x)$$


S
$\{\forall x Flight(x) \Rightarrow \exists e Arrived(e) \wedge ArrivedThing(e, x)\}$
NP
$\{\lambda Q.\forall x Flight(x) \Rightarrow Q(x)\}$
VP
$\{\lambda y.\exists e Arrived(e) \wedge ArrivedThing(e, y)\}$
Det
$\{\lambda P.\lambda Q.\forall x P(x) \Rightarrow Q(x)\}$
Noun
$\{\lambda x.Flight(x)\}$
V
$\{\lambda y.\exists e Arrived(e) \wedge ArrivedThing(e, y)\}$
Every
flight
arrived

Word Vectors



Pragmatics

- Discourse phenomena
- Coreference resolution [esp. pronominal]
  - Hobbs' Algorithm

- Segmentation / Cohesion
- Discourse parsing: hierarchical structure of coherence relations
  - PDTB discourse parsing

Summary

- Deep Processing techniques for NLP
  - Parsing, semantic analysis, logical forms, reference, etc
  - Create richer computational models of natural language
    - Closer to language understanding
- Shallow processing techniques have dominated many areas
  - IR, QA, MT, WSD, etc
    - More computationally tractable, fewer required resources
- Deep processing techniques experience resurgence
  - Some big wins – e.g. QA
  - Improved resources: treebanks (syntactic/discourse, FrameNet, Propbank)
  - Improved learning algorithms: structured learners, neural nets
  - Increased computation: cloud resources, Grid, etc
    - Current goal: leveraging these resources to do deep processing [e.g. semi-supervised learning]

Open Floor for Discussion

Thank you!

Course evaluations:

https://uw.iasystem.org/survey/279980