Modeling Natural Language Semantics with Learned Representations

Samuel R. Bowman

Premise: *A man speaking or singing into a microphone while playing the piano.*
Hypothesis: *A man is performing surgery on a giraffe while singing.*
Label: contradiction

Goal: Build computational models that can learn to understand and reason with human language.

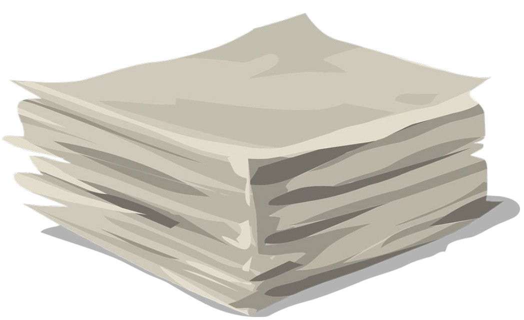
Open problems in NLP

Question answering

How old is the oldest leader of an OPEC country?

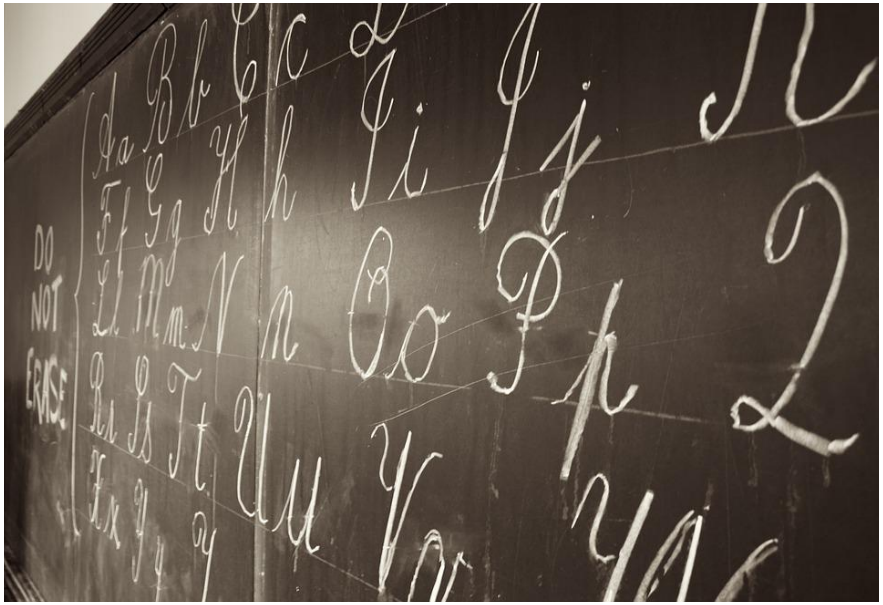
+ = ?

Summarization



= *Drug X interacts badly with drug Y.*

Open problems in linguistics



What prior knowledge must a learner have in order to fully learn language?

Open problems at the intersection

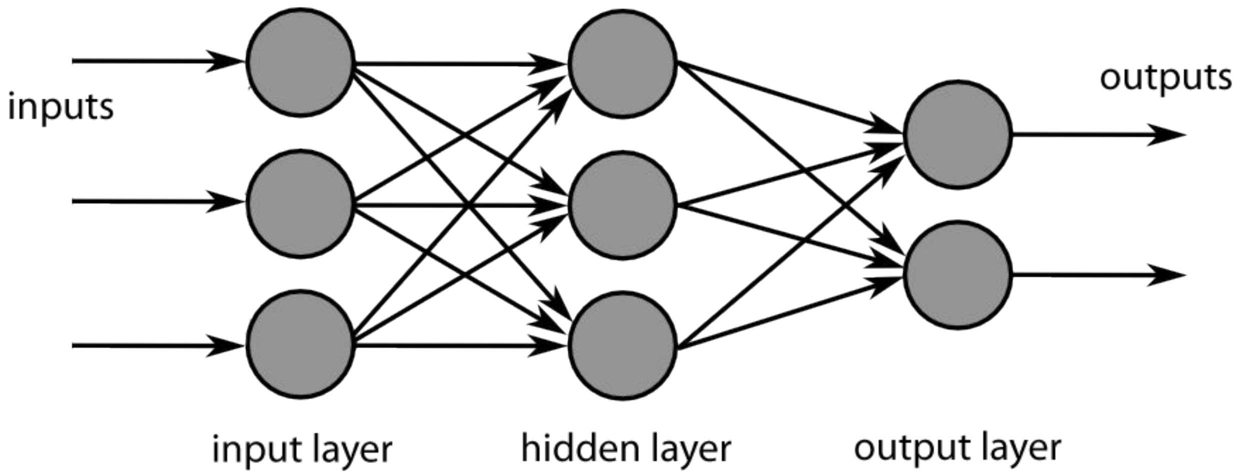
How do we combine logical approaches to meaning with a rich representations of word meaning?

$$\forall \mathbf{x}. \exists \mathbf{y} \ldots$$

If all dogs bark, do most puppies make sounds?

Is a labrador more of a dog than a chihuahua?

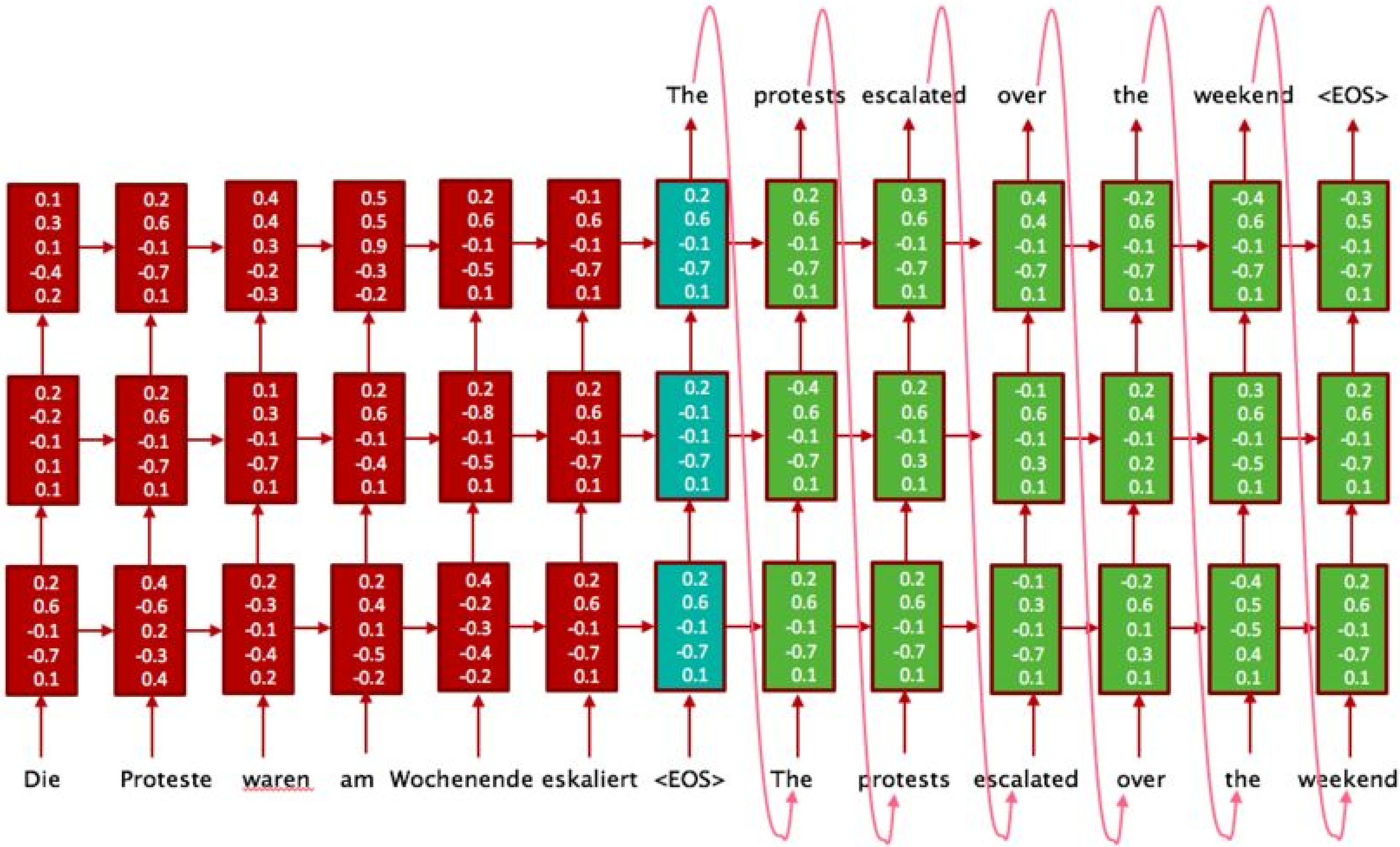
Neural networks in NLP



- **2010:** Marginal
- **2016:** Major research area
  Standard for parsing, classification, …

Figure from Christoph Burgmer.

Neural machine translation



Sutskever et al. ‘14, Bahdanau et al. ‘15, Luong et al. ‘15 (figure from Chris Manning)

Today: Some open questions

Goal: Build neural network models that can learn to understand and reason with human language.

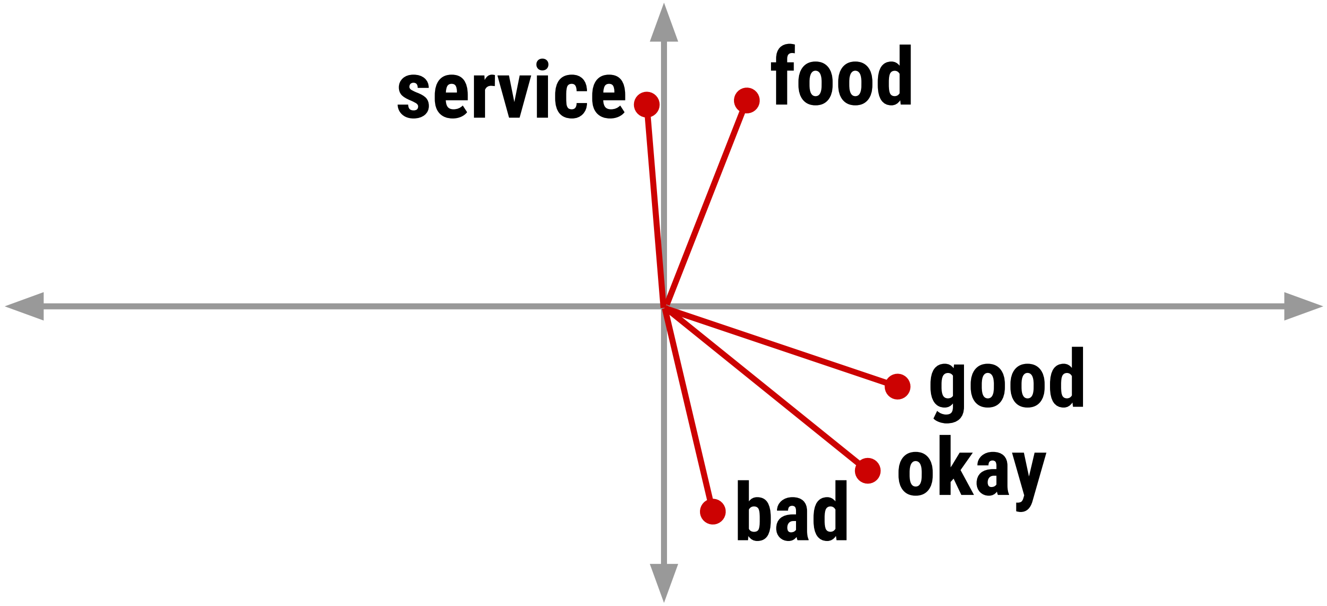
- Can continuous models do symbolic reasoning?
- Can they learn to understand real language?
- What can formal semantics teach them?

Background:
Neural networks and natural language

Distributed feature vectors for words

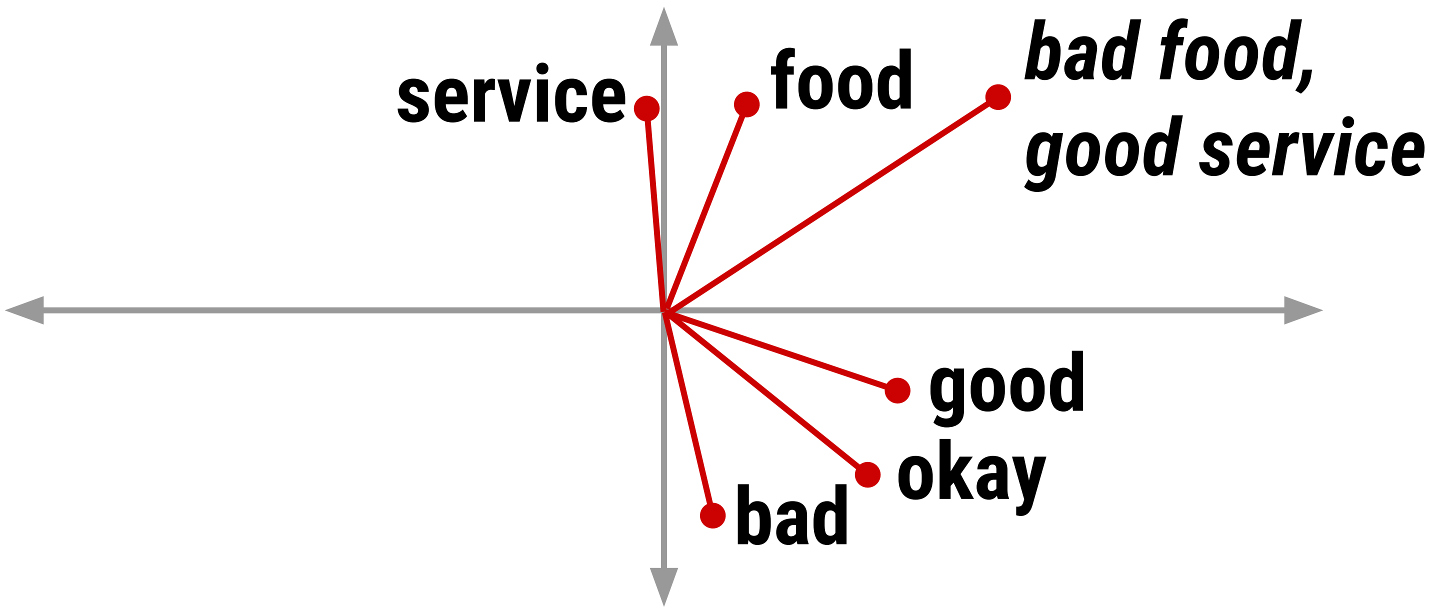
good	⇒	< 0.9, -0.2 >
okay	⇒	< 0.8, -0.5 >
bad	⇒	< 0.2, -0.7 >

...



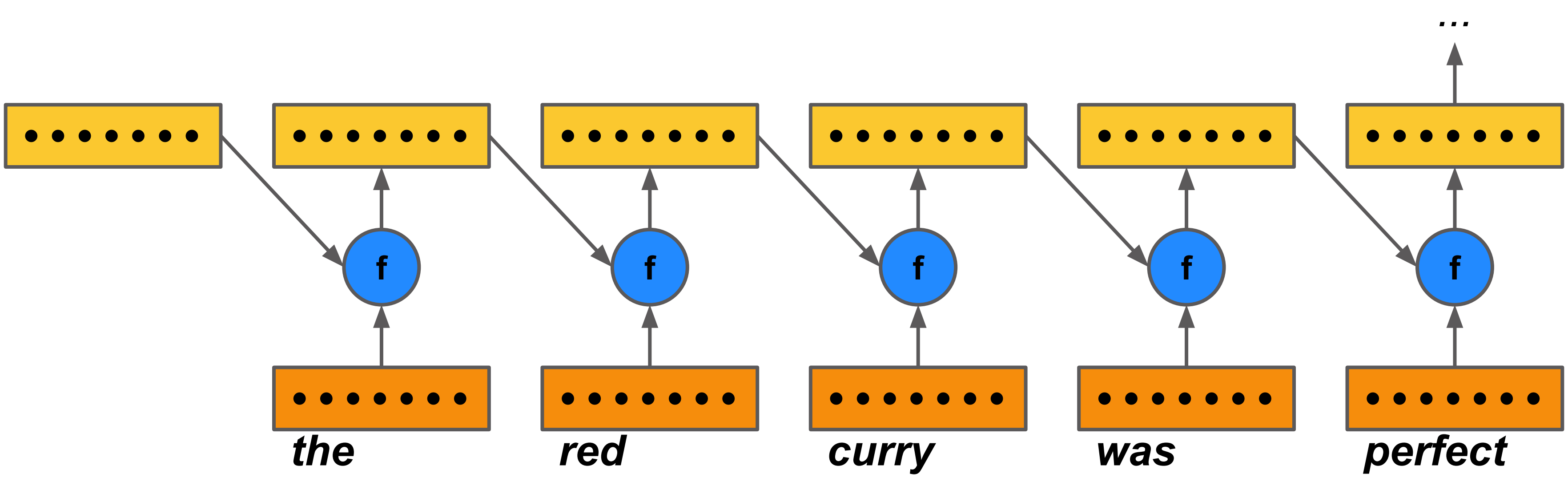
Composition: From words to sentences

How do we construct sentence representations from word representations?



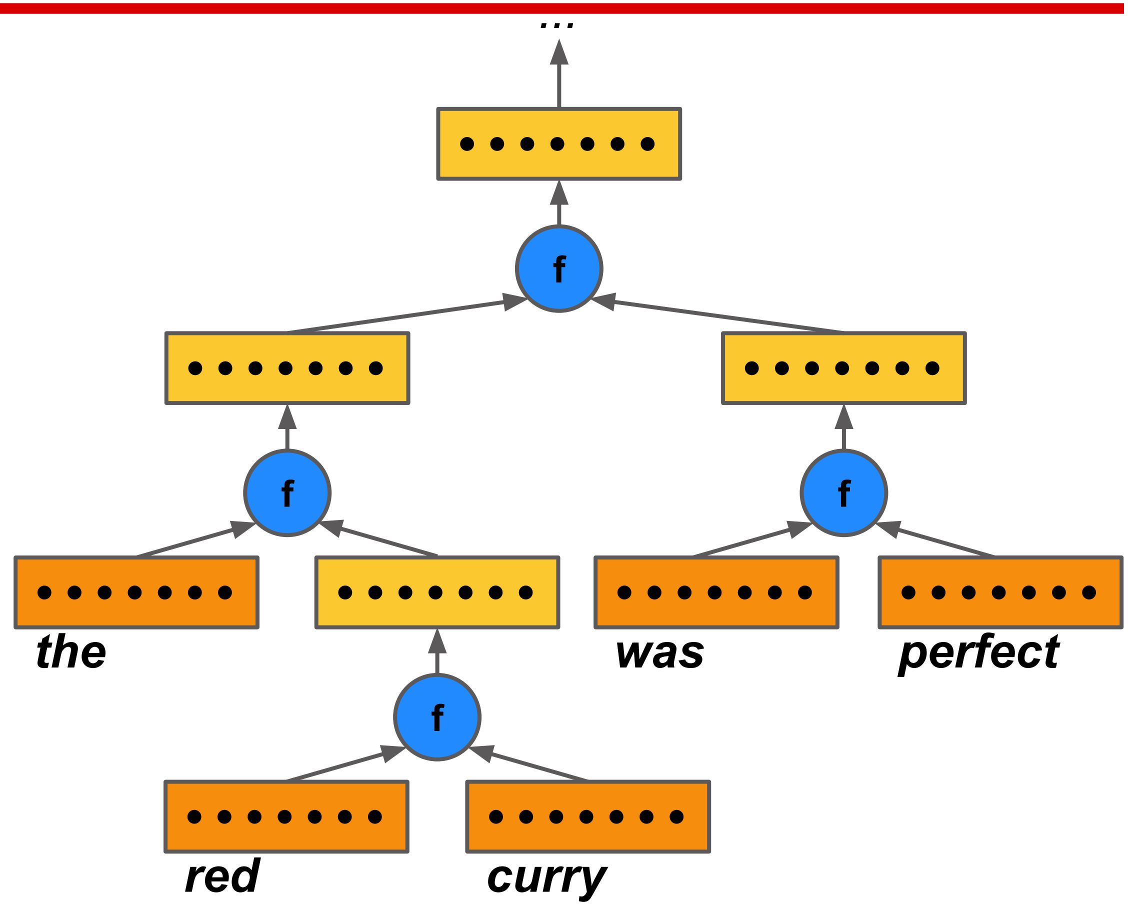
Composition: From words to sentences

Sequence-based *(recurrent) neural network encoder*



Rumelhart et al., ‘86; Werbos, ‘90; Mikolov, ‘10

Composition: From words to sentences



Alternative model:

Tree-structured *(recursive) neural network encoder*

Goller & Küchler ’96; Socher et al. ‘10

Some open questions

Goal: Build neural network models that can learn to understand and reason with human language.

- **Can continuous models do symbolic reasoning?**
- Can they learn to understand real language?
- What can formal semantics teach them?

Measuring success

Goal: Build neural network models that can learn to understand and reason with human language.

What does success look like?

Where does supervision come from?

Natural language inference (NLI)

or recognizing textual entailment *(RTE)*

James Byron Dean refused to move without blue jeans

{entails, contradicts, neither}

James Dean didn't dance without pants

MacCartney, '09

Natural language inference (NLI)

or recognizing textual entailment *(RTE)*

James Byron Dean refused to move without blue jeans

{**entails**, contradicts, neither}

James Dean didn't dance without pants

MacCartney, '09

Why natural language inference?

James Byron Dean refused to move without blue jeans

{**entails**, contradicts, neither}

James Dean didn't dance without pants

- *move vs. dance* (hypernymy and hyponymy)
- *refused to vs. didn't* (factives and implicatives)
- *James B. Dean vs. James Dean* (coreference)

...

Why natural language inference?

Natural language inference
is a major sub-problem of:

- Question answering
- Semantic web search
- Summarization
- Machine translation

and more!

NLI and Natural Logic

Research in **Natural Logic** formally characterizes sound inference patterns over natural language.

dance $\sqsubset$ *move*

so...

James Dean danced $\sqsubset$ *James Dean moved*

but...

James Dean didn't dance $\sqsupset$ *James Dean didn't move*

Sánchez-Valencia, ‘91; MacCartney, ‘09; Icard & Moss ‘13

Reasoning with words

Building a learning problem

Training data

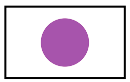
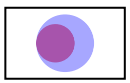
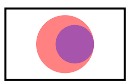
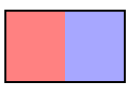
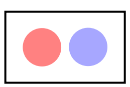
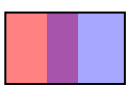
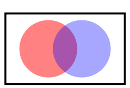
dance	***entails***	*move*
tango	***entails***	*dance*
sleep	***contradicts***	*dance*
waltz	***entails***	*dance*

Test data

sleep	***?***	*waltz*

Natural logic: The seven relations

Seven possible relations between phrases/sentences:

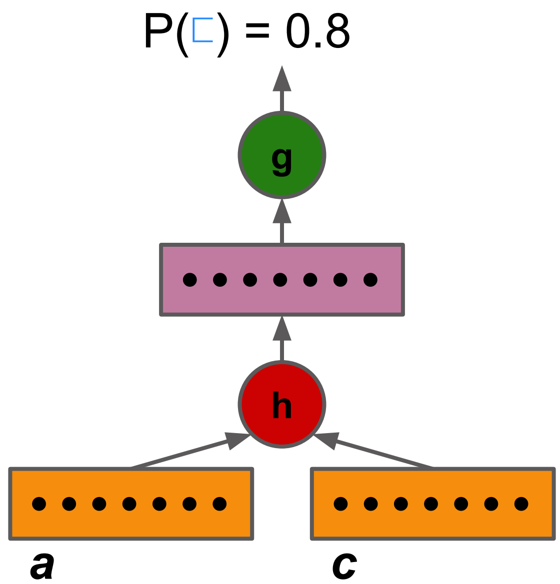
	equivalence	*couch*	$\equiv$	*sofa*
	forward entailment (strict)	*crow*	$\sqsubset$	*bird*
	reverse entailment (strict)	*European*	$\sqsupset$	*French*
	negation (exhaustive exclusion)	*human*	$\wedge$	*nonhuman*
	alternation (non-exhaustive exclusion)	*cat*	$\mid$	*dog*
	cover (exhaustive non-exclusion)	*animal*	$\smile$	*nonhuman*
	independence	*hungry*	$\#$	*hippo*

Slide from Bill MacCartney

Lexical relation data

TRAIN	TEST
a $\equiv$ a	a $\equiv$ b
a $\wedge$ f	a $\smile$ d
b $\smile$ c	a $\sqsupset$ e
b $\smile$ d	b $\sqsupset$ e

The simplest viable neural inference model



Learning lexical relations

Generalization (test) accuracy 99.6%

Training

dance ***entails*** *move*

Test

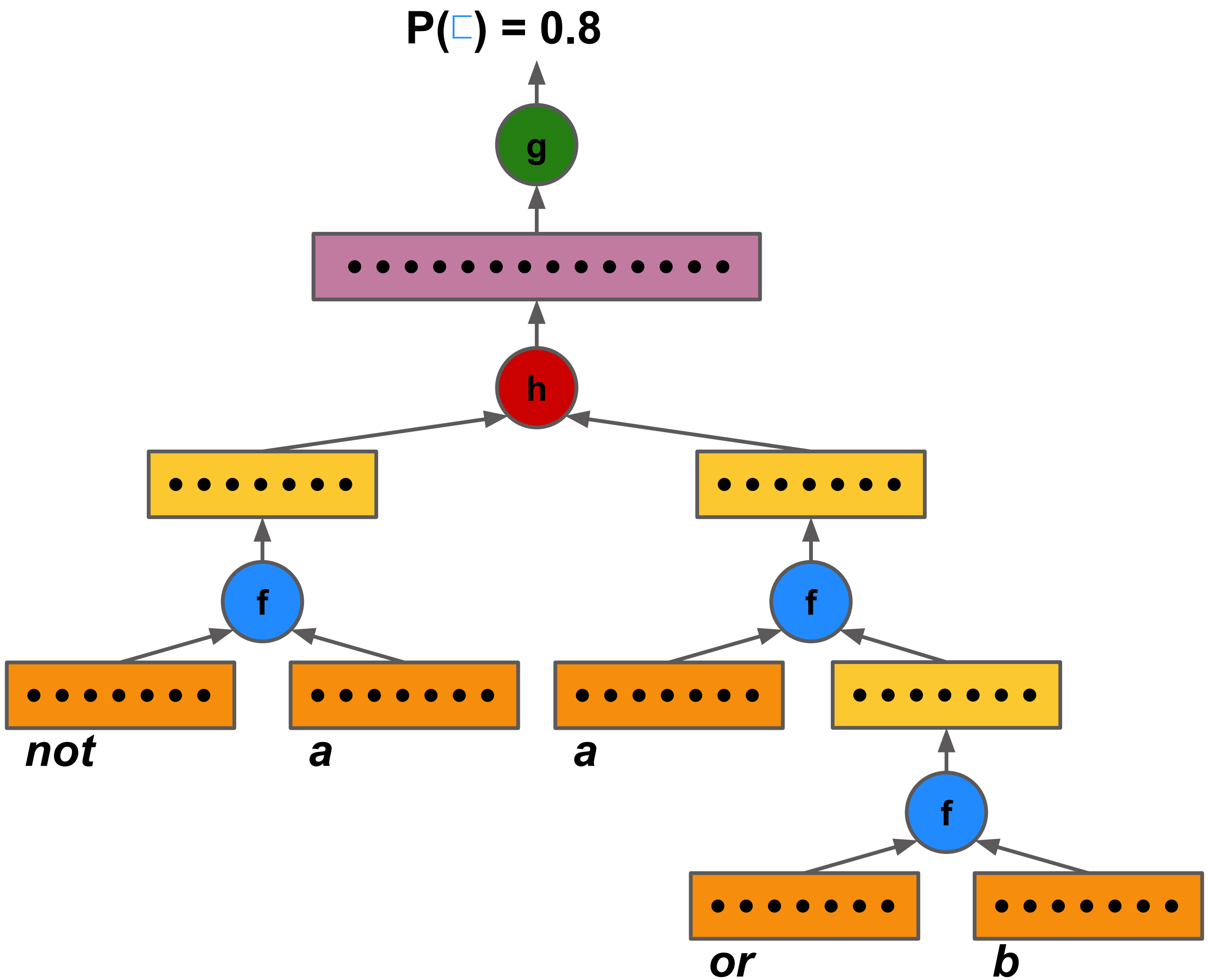
sleep ***?*** *waltz*

Reasoning with novel sentences

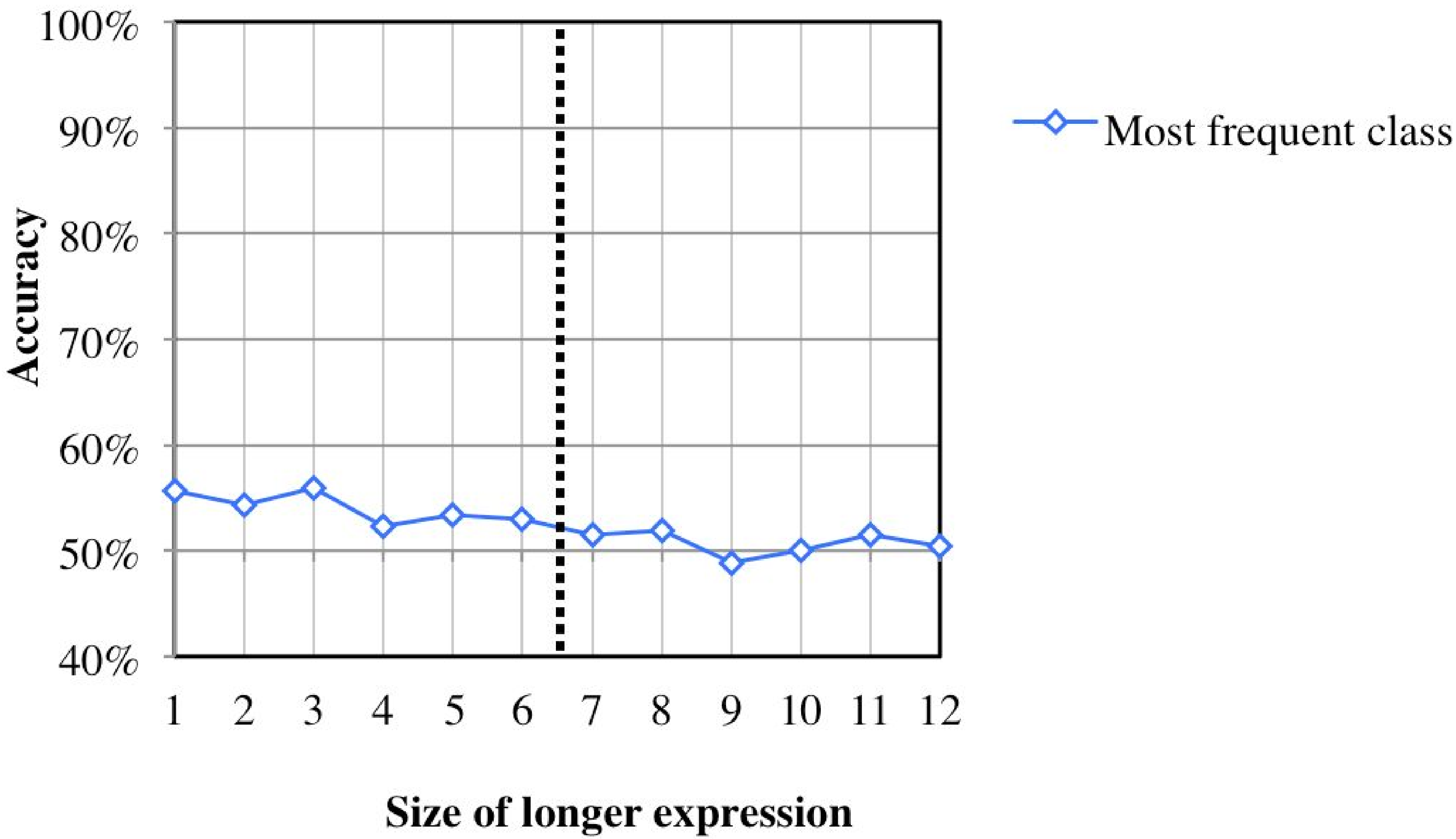
Function words and infinite languages

TRAIN			TEST		
b	≡	b	not a	^	a
not (not a)	≡	a	c or d	⊐	d
c	⊐	b and c	not not b	≡	b
–			not (not a and not d)	≡	a or d

The model: A TreeRNN for NLI



Function words and infinite languages

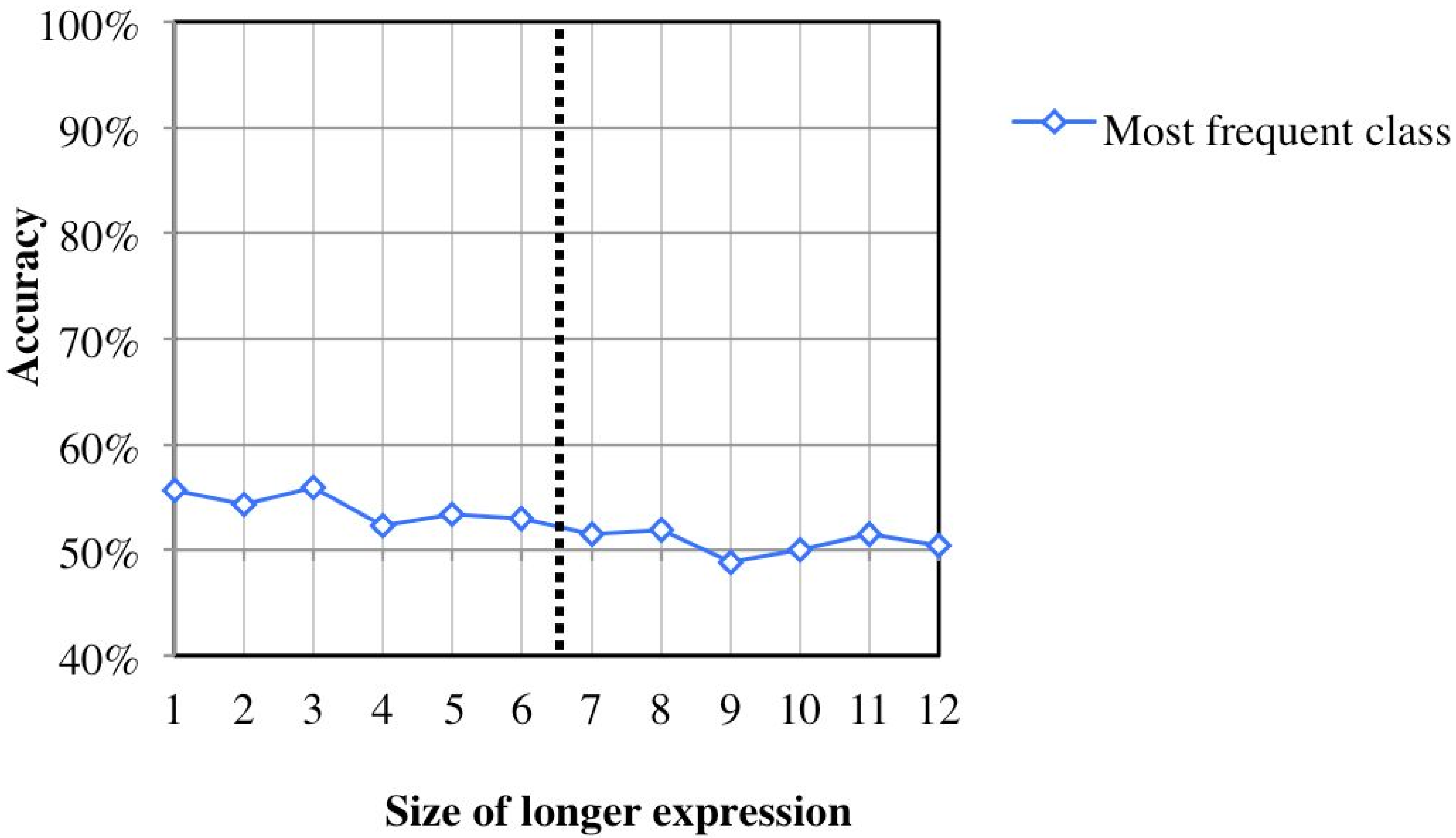


An example with twelve connectives

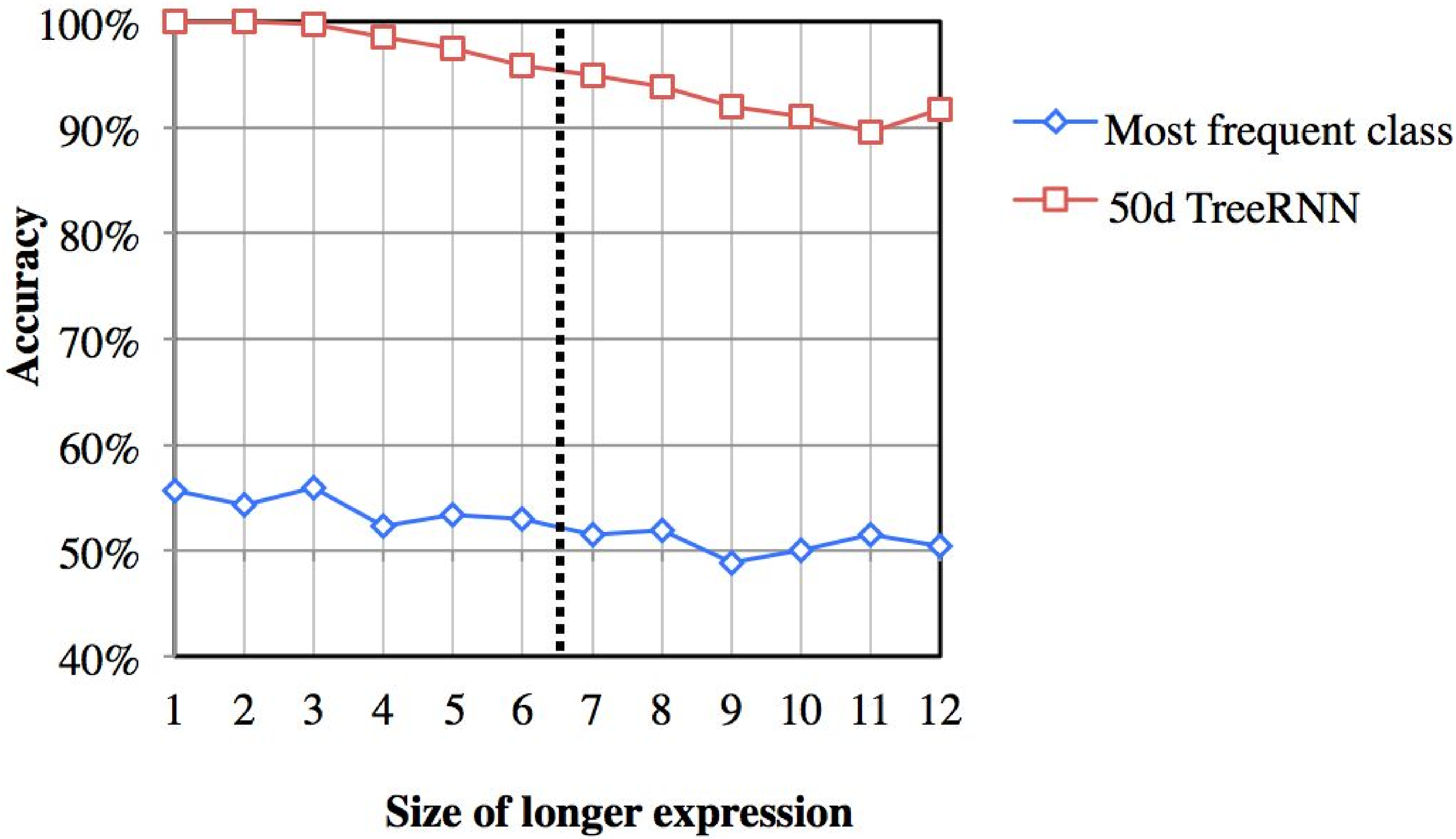
((not d) or (not ((not (b or e)) and (b or (not b)))))

(not ((not ((b and (not b)) or (not (d and b))))
or (not (((not e) or d) and (d or c)))))

Function words and infinite languages



Function words and infinite languages



Some open questions

Goal: Build neural network models that can learn to understand and reason with human language.

- Can continuous models do symbolic reasoning?
- **Can they learn to understand real language?**
- What can formal semantics teach them?

What data can we learn from?

Corpus	Complete Sentences	Human Labeled	Size (num. pairs)
FraCaS	✓	✓	**.3k**
RTE 1-5	✓	✓	**7k**
SICK	✓	✓	**10k**
DenotationGraph	✗	✗	**728k**
Levy Graphs	✗	✗	**1,500k**
PPDB 2.0	✗	✗	**100,000k**

What data can we learn from?

Corpus	Complete Sentences	Human Labeled	Size (num. pairs)
FraCaS	✓	✓	**.3k**
RTE 1-5	✓	✓	**7k**
SICK	✓	✓	**10k**
DenotationGraph	✗	✗	**728k**
Levy Graphs	✗	✗	**1,500k**
PPDB 2.0	✗	✗	**100,000k**

Training neural networks on existing data

A little girl is looking at a woman in costume

{entailment, contradiction, **neutral**}

The little girl is looking at a man in costume

Approach	**SICK test acc.**
Just guessing ‘neutral’	56.7%
Best NN model	76.9%
Best prior non-NN model	**84.5%**

What data can we learn from?

Corpus	Complete Sentences	Human Labeled	Size (num. pairs)
FraCaS	✓	✓	**.3k**
RTE 1-5	✓	✓	**7k**
SICK	✓	✓	**10k**
DenotationGraph	✗	✗	**728k**
Levy Graphs	✗	✗	**1,500k**
PPDB 2.0	✗	✗	**100,000k**

Our large, human-labeled NLI corpus

Corpus	Complete Sentences	Human Labeled	Size (num. pairs)
FraCaS	✓	✓	**.3k**
RTE 1-5	✓	✓	**7k**
SICK	✓	✓	**10k**
SNLI	✓	✓	**570k**
DenotationGraph	✗	✗	**728k**
Levy Graphs	✗	✗	**1,500k**
PPDB 2.0	✗	✗	**100,000k**

Bowman et al. '15: "A large annotated corpus for learning natural language inference"

The Stanford NLI Corpus

EMNLP '15
Best New Data
Set Award

Girl in a red coat, blue head wrap and jeans is making a snow angel.

{**entailment**, contradiction, neutral}

A girl outside plays in the snow.

- Typical examples require:
  - Full sentence understanding.
  - Common sense world knowledge.
- Outside the scope of pure natural logic.

How do we collect this data?

Prompt for Mechanical Turk annotators:

> We will show you the caption for a photo. We will not show you the photo. Using just the caption and what you know about the world, write a new caption for the same photo that is {definitely accurate, definitely inaccurate, possibly accurate}.

Initial machine learning results

Model	**Test acc.**
Just guessing ‘entailment’	33.7%
Big simple classifier	**78.2%**
Recurrent (sequence) NN model	**77.6%**

Extramural results

- **Sep. 2015:** Corpus release
- **Sep. 2015:** Google DeepMind/UCL/Oxford
- **Nov. 2015:** U. of Toronto
- **Dec. 2015:** Peking U./Baidu
- **Dec. 2015:** Singapore Management U.
- **Jan. 2016:** U. of Edinburgh
- **Feb. 2016:** Unbabel Lda./IT/INESC-ID (Pt.)

Some open questions

Goal: Build neural network models that can learn to understand and reason with human language.

- Can continuous models do symbolic reasoning?
- Can they learn to understand real language?
- **What can formal semantics teach them?**

Working assumptions in formal semantics

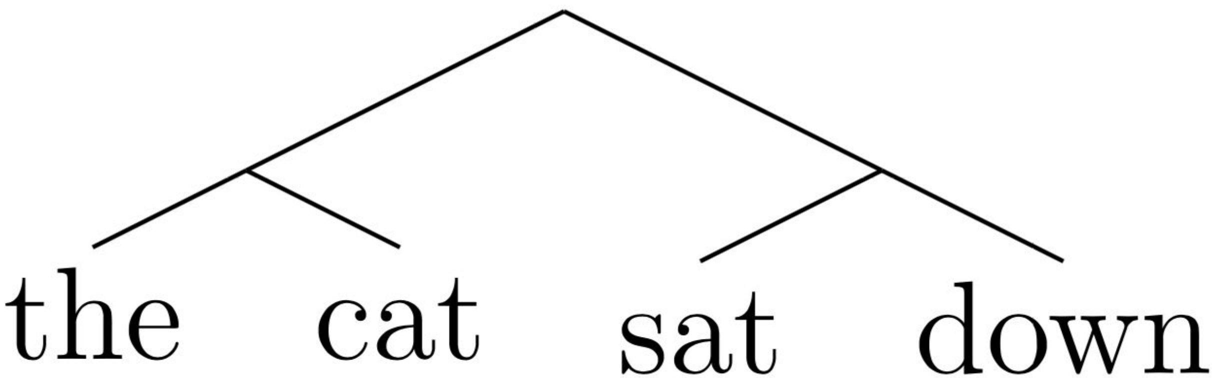
Loosely, *the principle of compositionality*:

the cat sat down

Sentence meanings are constructed incrementally by composing together word meanings.

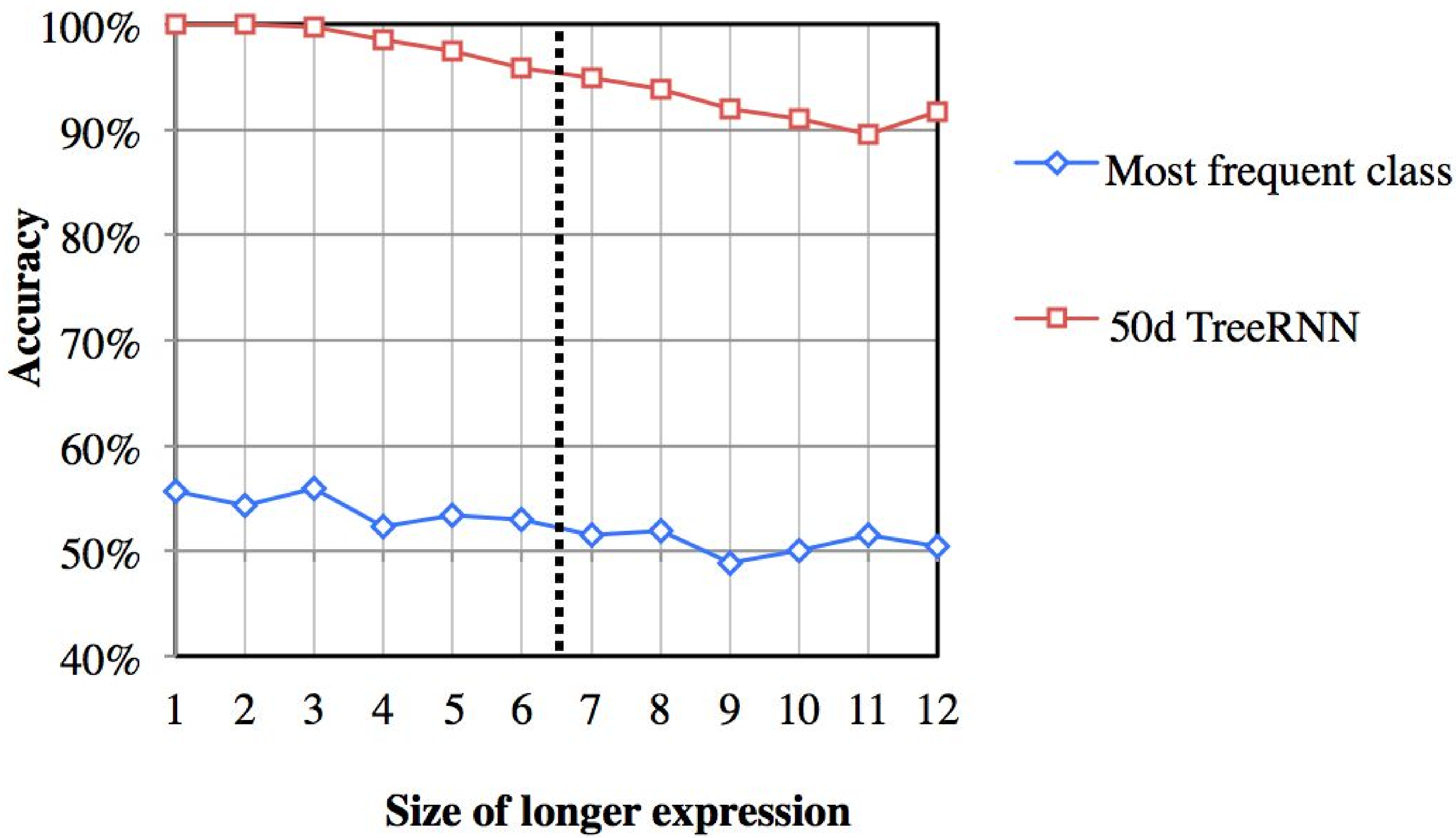
Working assumptions in formal semantics

Loosely, *the principle of compositionality*:

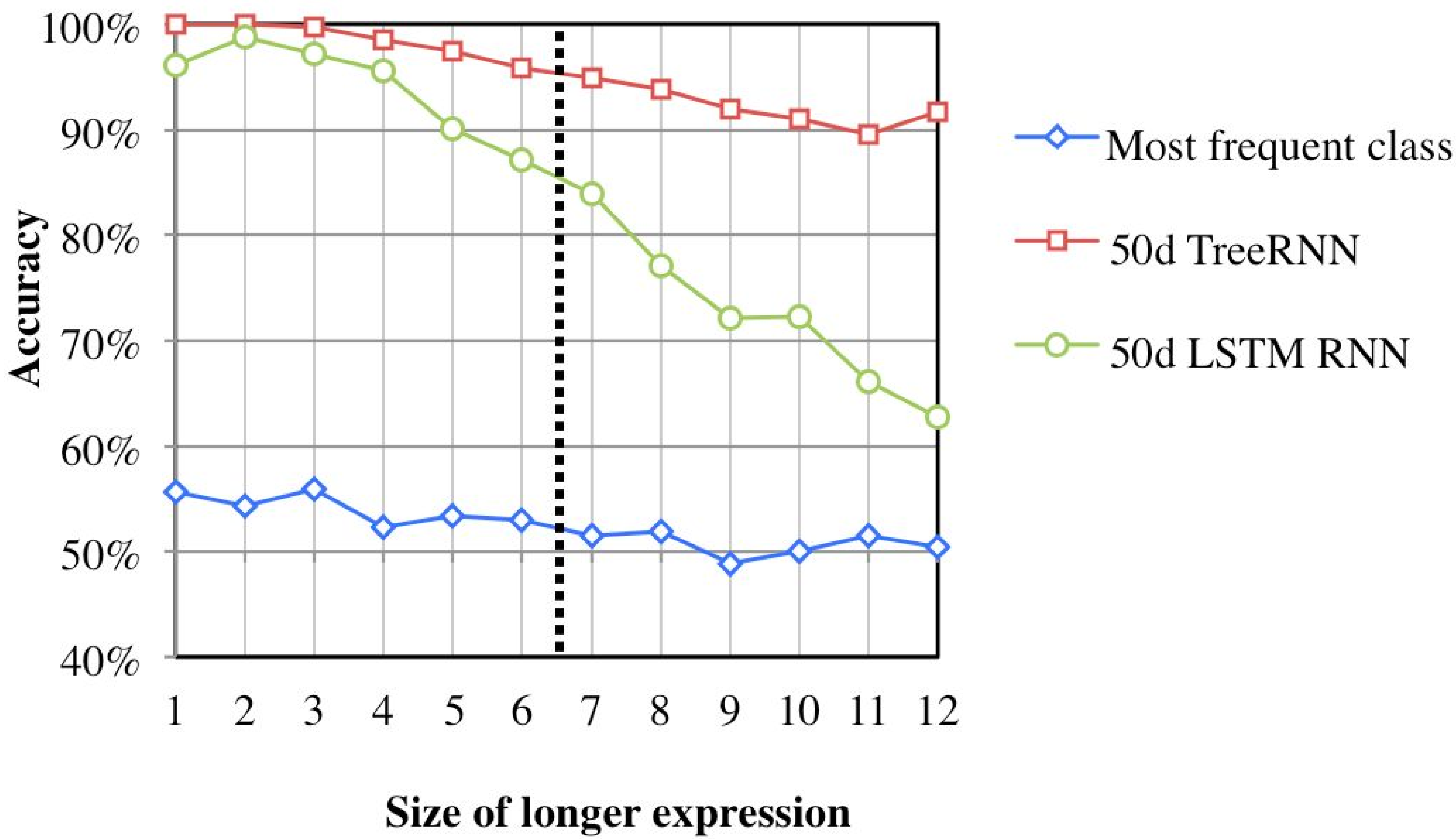


This composition process can be most concisely described using a phrase structure that *roughly* follows the phrase structure used in syntax.

Recursion with propositional logic



Recursion with propositional logic

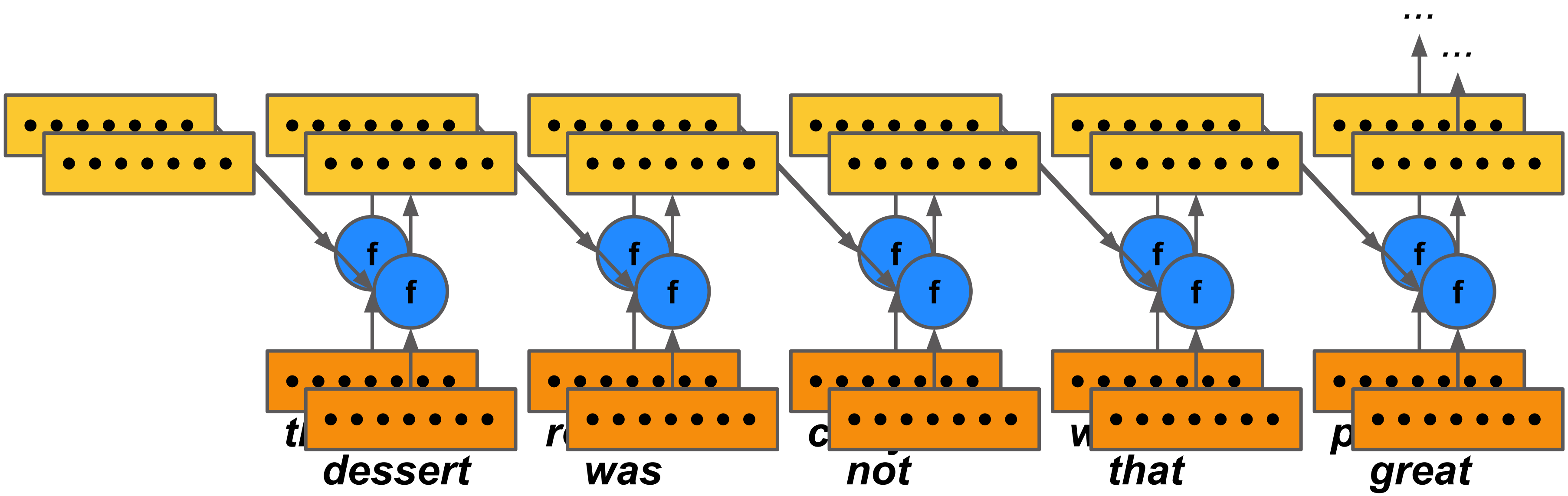


Tree structured models in practice

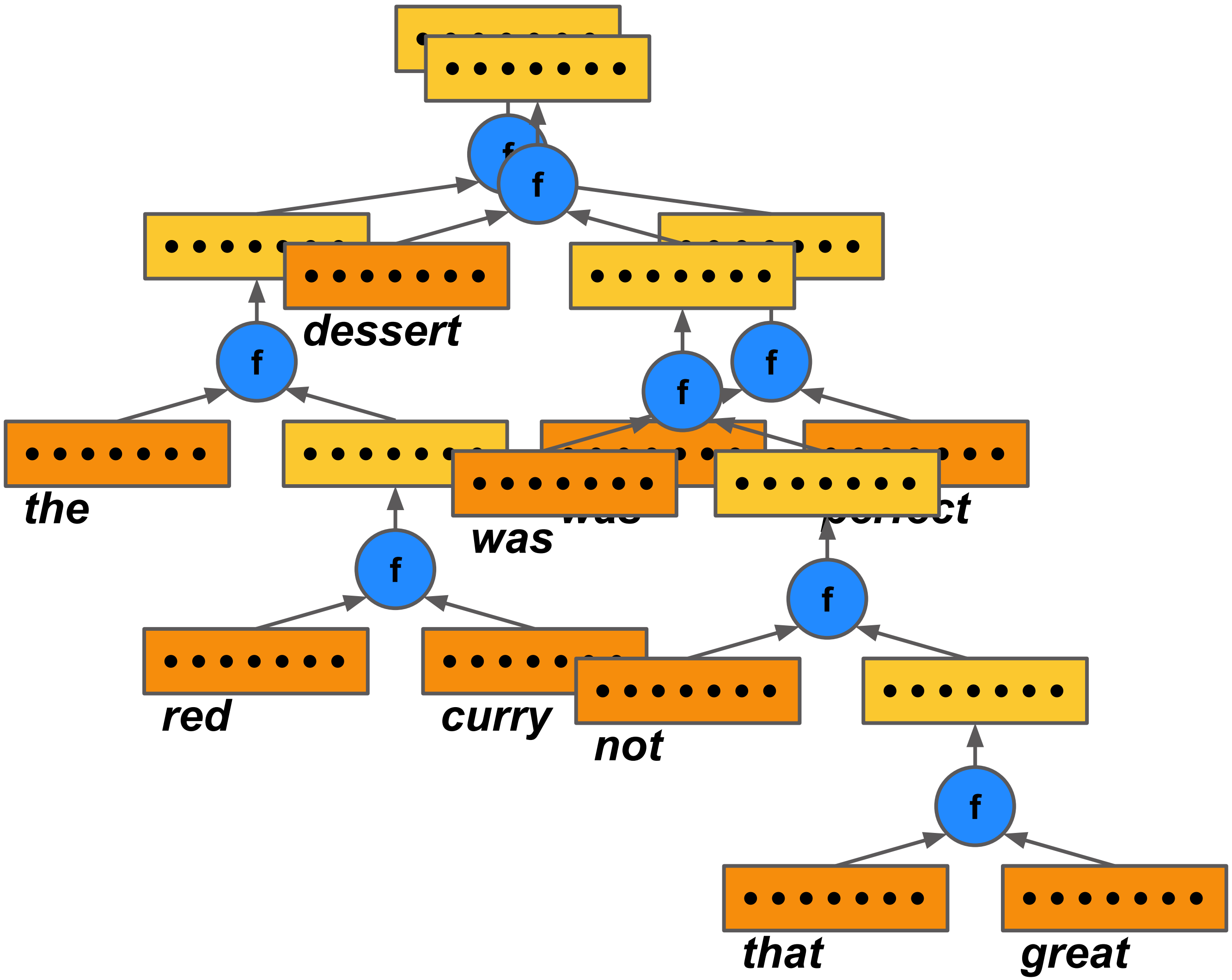
Robust successes on NLP for tasks with smaller datasets: sentiment analysis, paraphrase detection...

Larger datasets? Too slow.

Batched computation



Batched computation



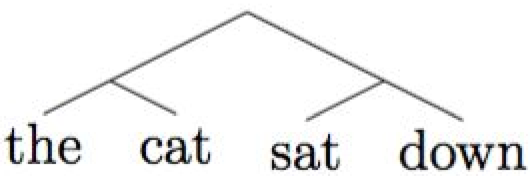
Transition-based parsing

Is it possible to do tree-structured compositionality in an efficient model?

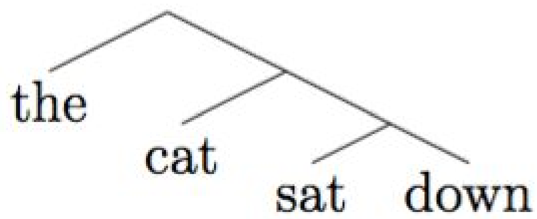
Transition-based parsing offers a clue.

Bowman et al. '16: "A Fast Unified Model for Parsing and Sentence Understanding"

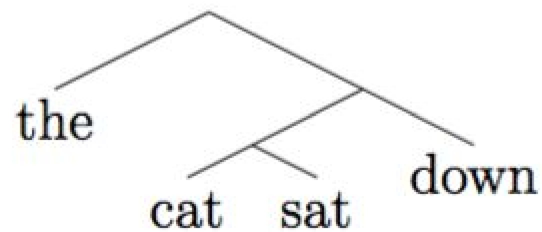
Transition-based parsing



SHIFT SHIFT
REDUCE SHIFT
SHIFT REDUCE
REDUCE

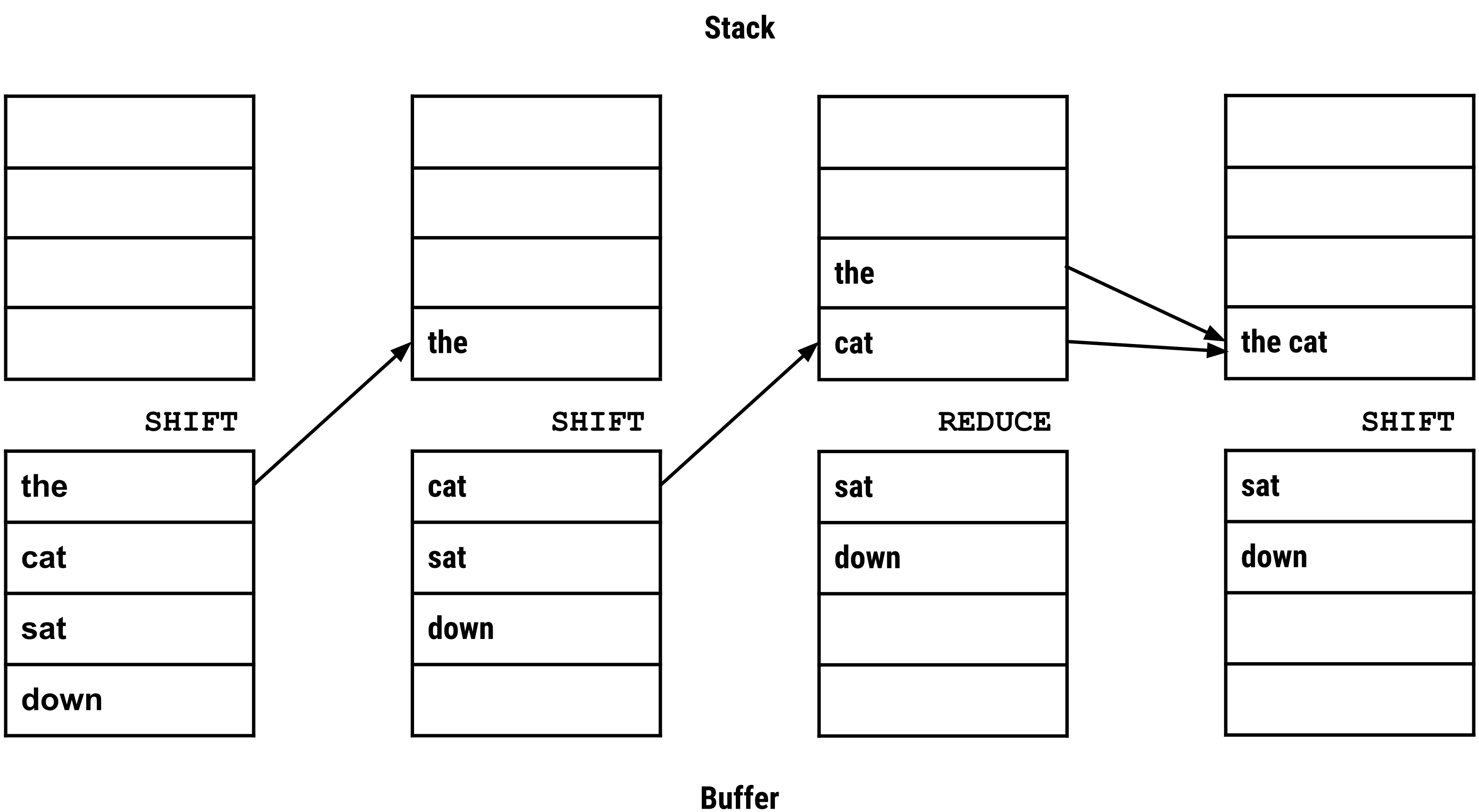


SHIFT SHIFT
SHIFT SHIFT
REDUCE REDUCE
REDUCE

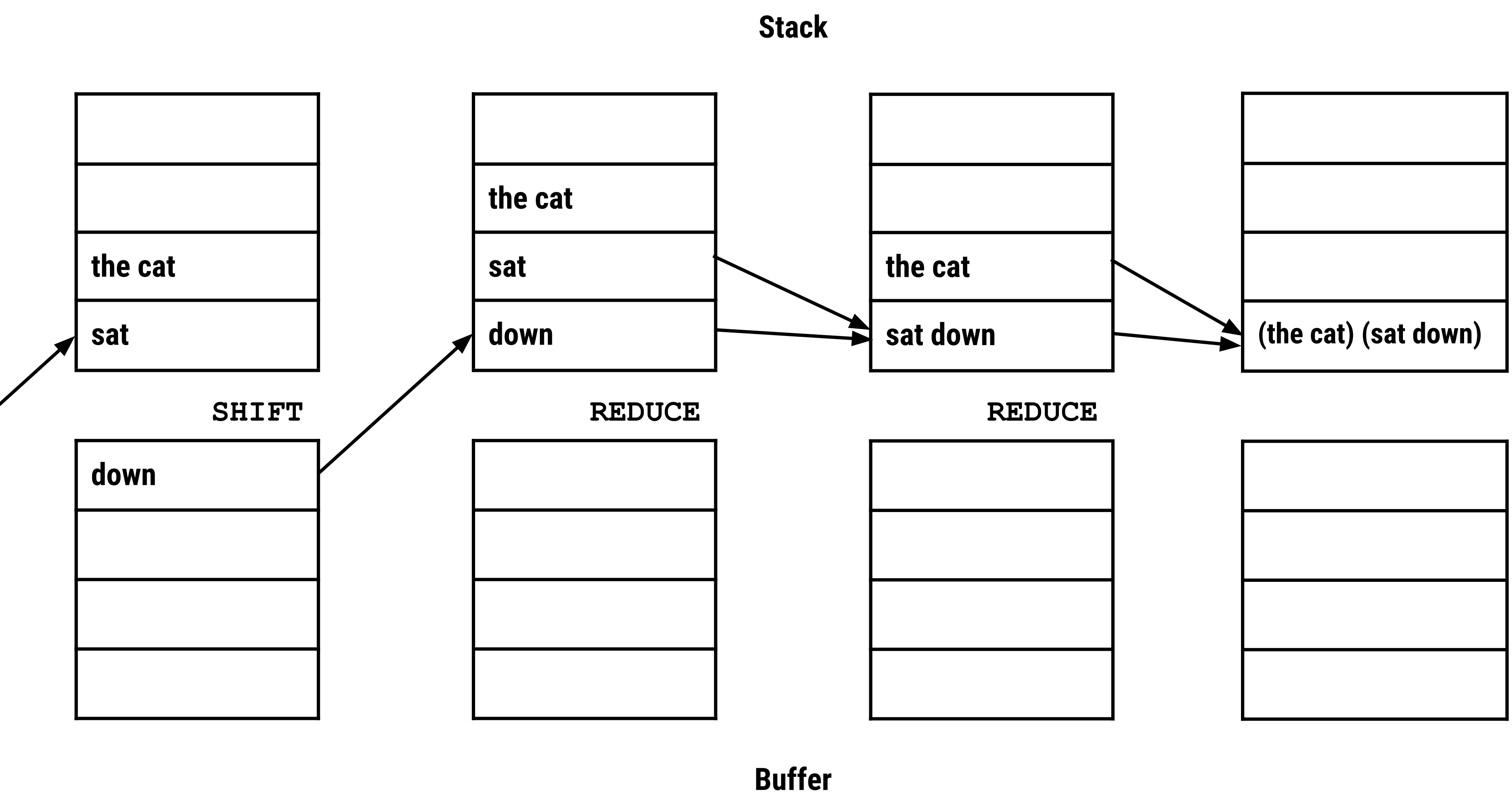


SHIFT SHIFT
SHIFT REDUCE
SHIFT REDUCE
REDUCE

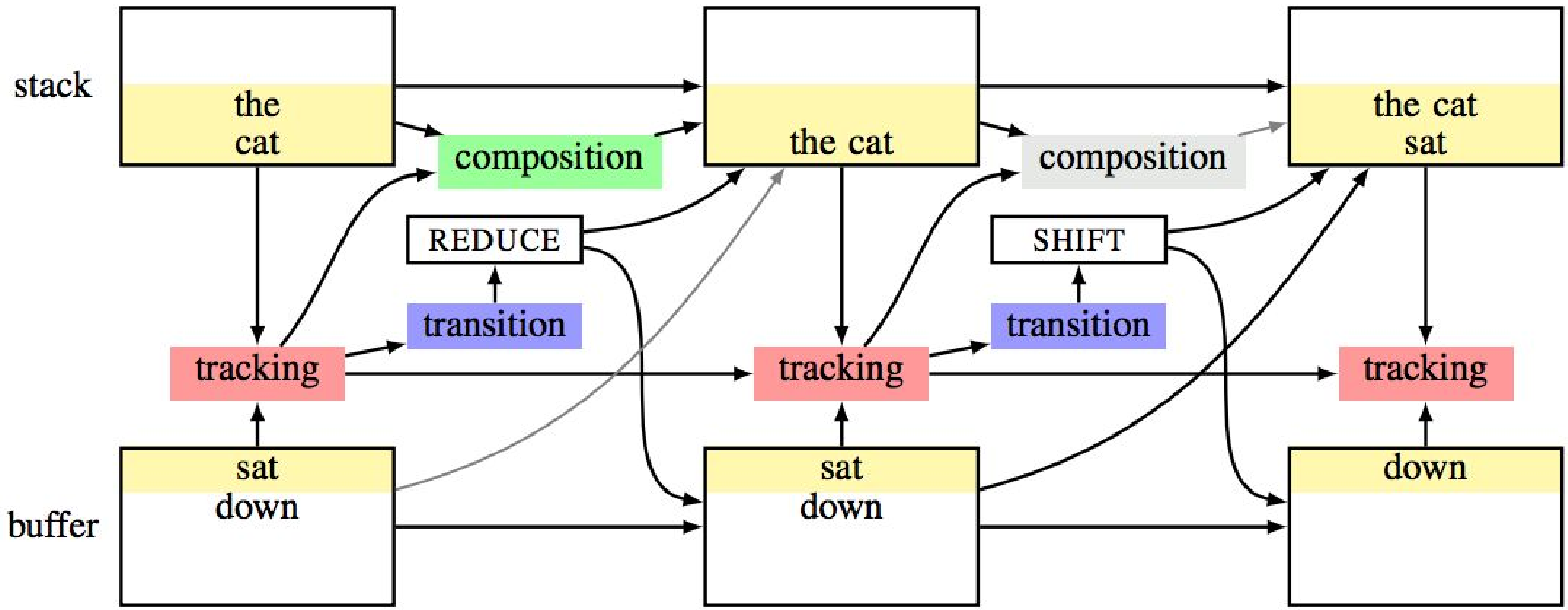
Transition-based parsing



Transition-based parsing



Stack-augmented Parser-Interpreter NN



The shift-reduce model on SNLI

Model	**Test acc.**
Sequence model (our prev. impl.):	77.6%
Best comparable model:	82.1%
Sequence model (our new impl.):	80.6%
SPINN (purely tree-structured):	80.9%
SPINN (hybrid):	**83.2%**

Ongoing work: Future directions

Neural attention

State-of-the-art *attention-based* model: **89.0%**

Attention-based SPINN: ?

Learning syntax from semantics

Build models that can learn to use whatever parse structure best supports the task at hand

Some open questions, and some answers

Goal: Build neural network models that can learn to understand and reason with human language.

Can continuous models do symbolic reasoning?

- Yes, e.g., lexical relations, recursive functions...

((not d) or (not ((not (b or e)) and (b or (not b)))))

(not ((not ((b and (not b)) or (not (d and b))))
or (not (((not e) or d) and (d or c)))))

Some open questions, and some answers

Goal: Build neural network models that can learn to understand and reason with human language.

Can they learn to understand real language?

- Not perfectly yet, but at the state of the art and making rapid progress.

Girl in a red coat, blue head wrap and jeans is making a snow angel.

{**entailment**, contradiction, neutral}

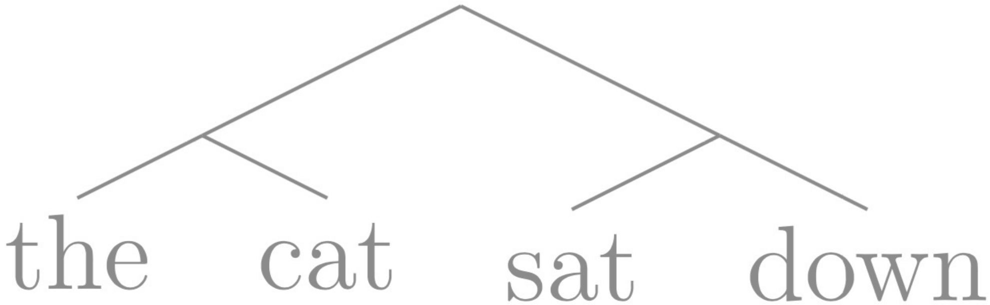
A girl outside plays in the snow.

Some open questions, and some answers

Goal: Build neural network models that can learn to understand and reason with human language.

What can formal semantics teach them?

- Compositionality, at least: yields huge gains on artificial data, and significant gains on English.



Where we are now

Neural networks are the most effective tool we have for learning to understand natural language, but our models are still far from human-level understanding.

Future work

To fill the gap, more work is needed into:

- Discovering what aspects of meaning these models learn to use in practice.
- Applying our theoretical understanding of language to build helpful learning biases.
- Building models that can learn to refine their representations of meaning using raw text or other kinds of labeled data.

Modeling Natural Language Semantics with Learned Representations

Samuel R. Bowman

Premise: *A man speaking or singing into a microphone while playing the piano.*
Hypothesis: *A man is performing surgery on a giraffe while singing.*
Label: contradiction