Natural Language Processing and Understanding

Methods, Applications, and Frontiers

Eric Saund, Ph.D.
November, 2019

www.saund.org

$$\underbrace{\text{Natural Language}}_{a}\ \Big(\underbrace{\text{Processing}}_{x}\ \underbrace{\text{and}}_{+}\ \underbrace{\text{Understanding}}_{y}\Big)$$

$$a\,(x + y) = ax + ay$$

Distributive property of multiplication over addition.

Natural Language *Processing*

and

Natural Language *Understanding*

Natural Language *Processing*

and

Natural Language *Understanding*

Natural Language *Processing*

and

Natural Language *Understanding*

Methods, Applications, and Frontiers

NLP and NLU: Outline

- Context
  - Ingredients of Intelligence
  - Cognitive Architecture
- The Nature of Language
  - What language is for
  - Syntax and semantics
  - Speech and writing
  - Logic and statistics
- Natural Language Computing Tools
  - The NLP Pipeline
  - Part-Of-Speech, lemmas, relational frames
  - Named Entities
  - Parsers

- Applications I
  - Sentiment

- Representations
  - Bag-of-Words
- Applications II
  - Document topic models
- Knowledge Graphs
- Applications III
  - Question answering
  - Entity/Intent extraction
  - Commonsense knowledge and reasoning
- Deep Methods
  - Language Models
  - Vector embedding
  - Sequence networks
  - Transformer networks
- Applications IV
  - GLUE tasks
  - Machine Translation

- Dialogue
- Applications V
  - Google Assist
  - Call Center Assistants

Context

Context

Context: The Ingredients of Intelligence

Knowledge

- *store of information in organized form*

Pattern Matching

- *approximate fit*

- *Natural Language Processing*
- *Natural Language Understanding*

Reasoning

- *multi-step inference*

Cognitive Architecture

What is Language For?

Communication by Message Passing

Building Common Ground

Herb Clark, 1985

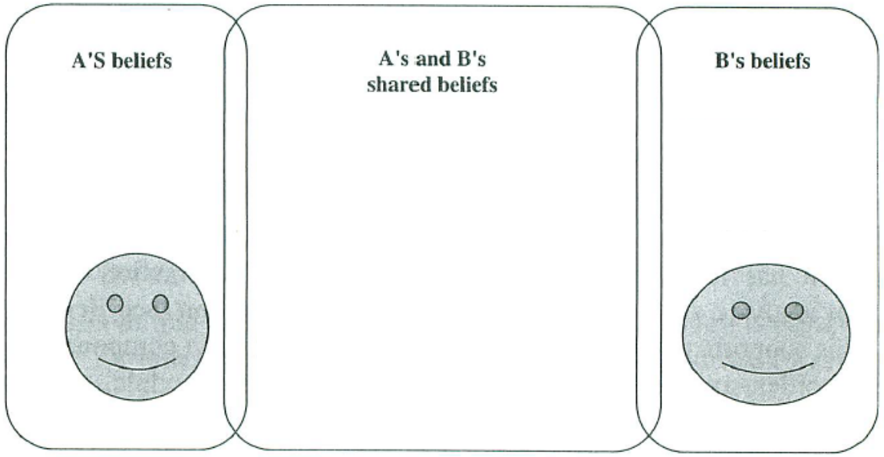
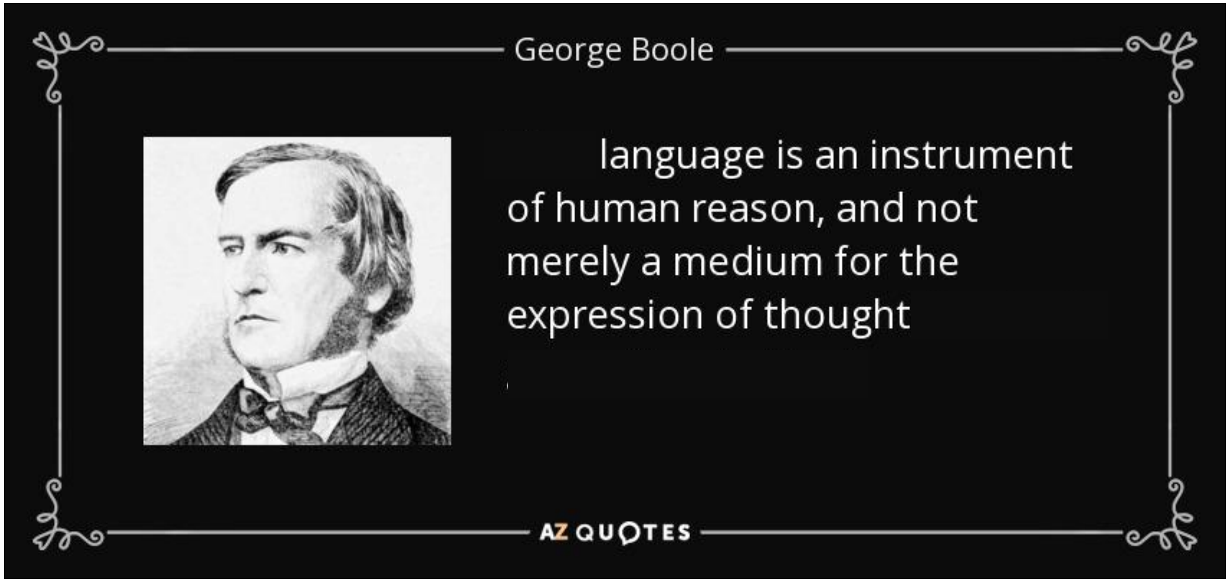


Figure 1.4 Constraint Rationality

K. Jokinen, "Constructive Dialog Modeling", 2009

What is Language For?



“mentalese”

Symbols and Meaning

"warm fire"	"ζεστή φωτιά"	"نار دافئة"	"feu chaud"
"fuego tibio"	"暖かい火"	"烈火"	"गर्म आग"

Combinations of Symbols

- Human language allows uncountably many sentences.
- Human language is *compositional*.

John saw Mary	×	start the car	×	yesterday	×	...
John told Mary to		start the fire		before dinner		
Mary expected John to		order the pizza		every Thursday		
⋮		⋮		⋮		

John saw Mary start the fire.

Mary saw John start the fire.

- Word order matters (in some languages).

- Constraints on word order and form.

POS : Part of Speech, e.g. noun, proper noun, verb, adjective, determiner, ...

* Start saw fire John the Mary.

* Mary saw John's start the fire.

inadmissibility of syntax

* John angled Mary evaporate the duration.

p. n. v. p. n. v. d. n.

inadmissibility of semantics

Speech and Writing, Formal and Informal

... then let us not forget the lessons of our forefathers...

Whoa, wipeout!

Projector

From Wikipedia, the free encyclopedia

For other uses, see Projector (disambiguation).

A **projector** or **image projector** is an optical device that projects an image (or moving images) onto a surface, commonly a projection screen. Most projectors create an image by shining a light through a small transparent lens, but some newer types of projectors can project the image directly, by using lasers. A virtual retinal display, or retinal projector, is a projector that projects an image directly on the retina instead of using an external projection screen.

Logic and Statistics

Automatic Speech Recognition

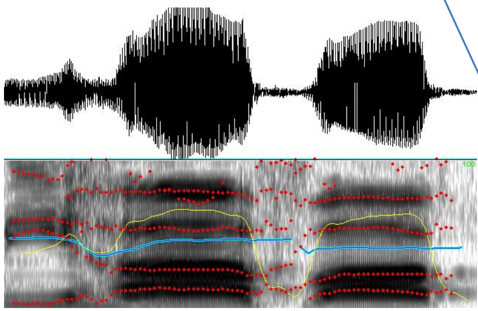
Natural Language Processing

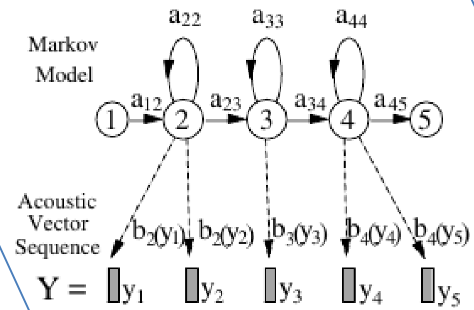
1990's

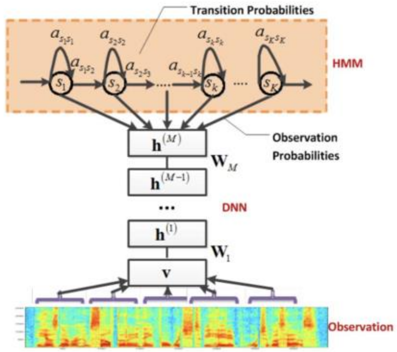
2000's

2010's

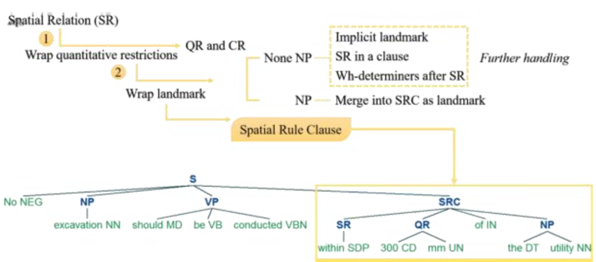
Logic and Statistics

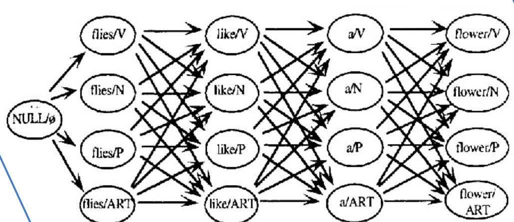


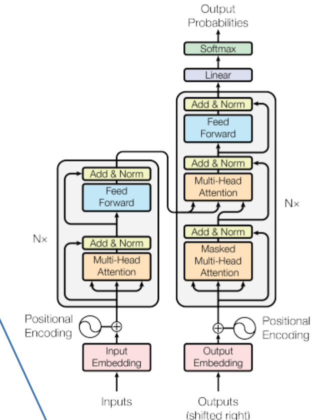




Automatic Speech Recognition







Natural Language Processing

rule-based era

statistical era

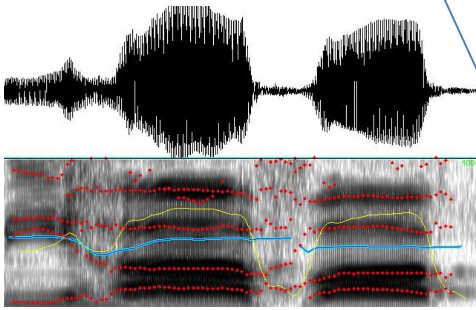
Deep NN era

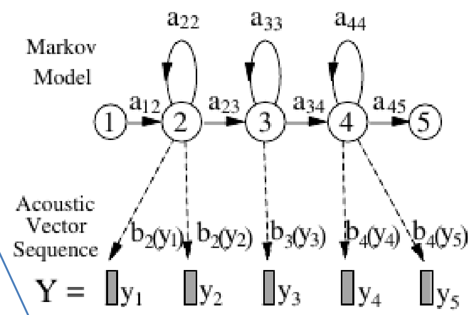
1990's

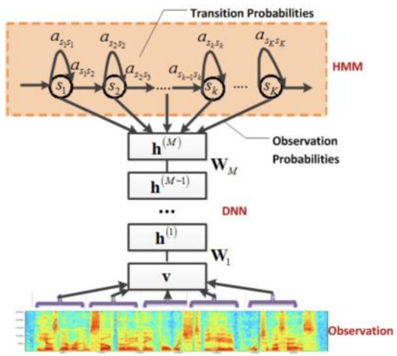
2000's

2010's

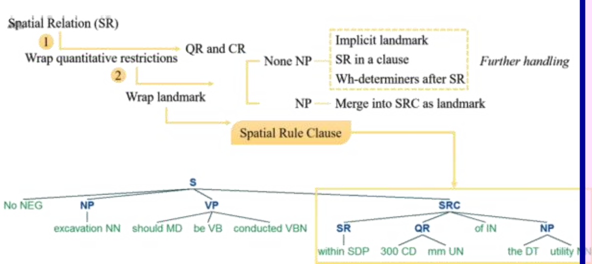
Logic and Statistics



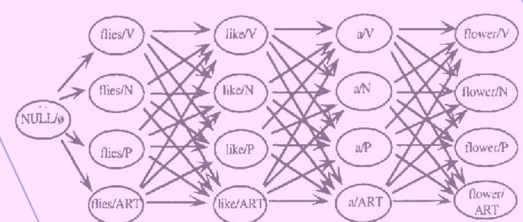


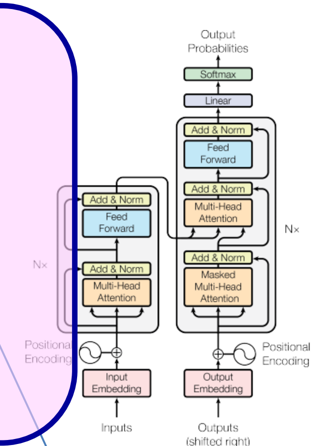


Automatic Speech Recognition



NLP Pipeline





Natural Language Processing

rule-based era | statistical era | Deep NN era

1990's | 2000's | 2010's

NLP and NLU: Outline

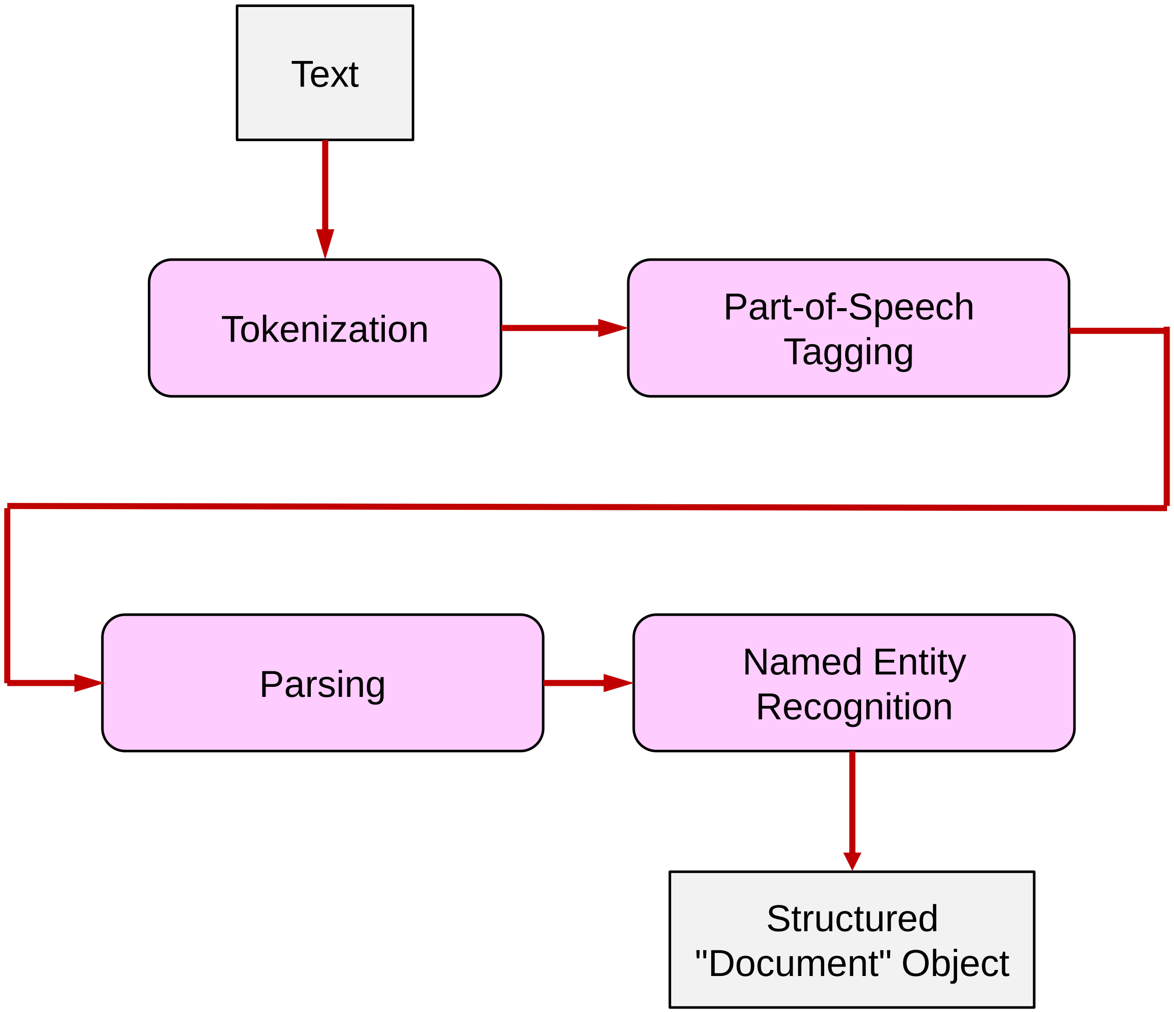
- Context
  - Ingredients of Intelligence
  - Cognitive Architecture
- The Nature of Language
  - What language is for
  - Syntax and semantics
  - Speech and writing
  - Logic and statistics
- Natural Language Computing Tools
  - The NLP Pipeline
  - Part-Of-Speech, lemmas, relational frames
  - Named Entities
  - Parsers

- Applications I
  - Sentiment

- Representations
  - Bag-of-Words
- Applications II
  - Document topic models
- Knowledge Graphs
- Applications III
  - Question answering
  - Entity/Intent extraction
  - Commonsense knowledge and reasoning
- Deep Methods
  - Language Models
  - Vector embedding
  - Sequence networks
  - Transformer networks
- Applications IV
  - GLUE tasks
  - Machine Translation

- Dialogue
- Applications V
  - Google Assist
  - Call Center Assistants

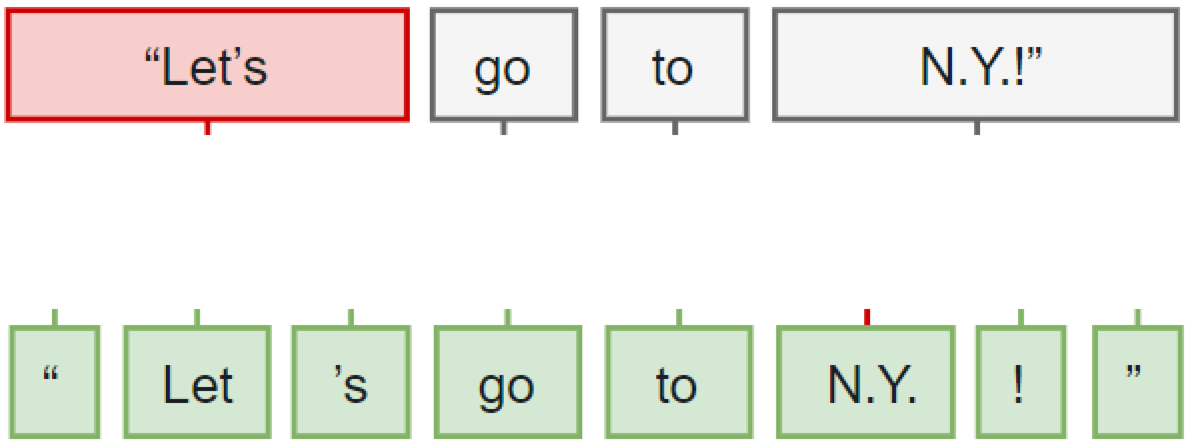
The NLP Pipeline



spaCy NLTK Natural Language Toolkit Stanford University S NLP Natural Language Processing AllenNLP

Tokenization

“Let’s go to N.Y.!”



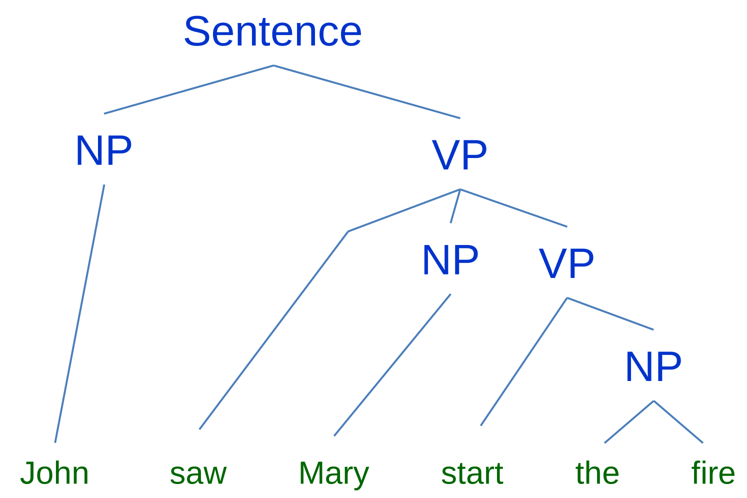
spaCy.io

Part-of-Speech Tagging

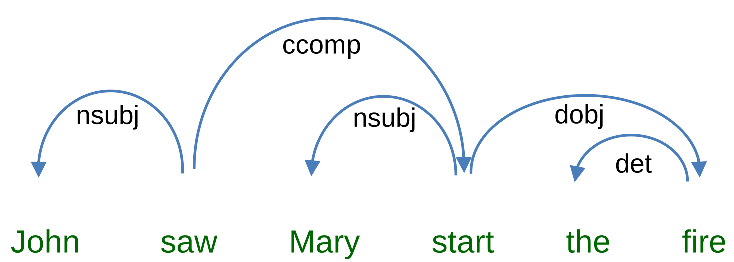
John	saw	Mary	start	the	fire
NNP Proper Noun	NN Noun	NNP Proper Noun	NN Noun	DT Determiner	NN Noun
	VBD Verb, past tense		VB Verb, base		VB Verb, base

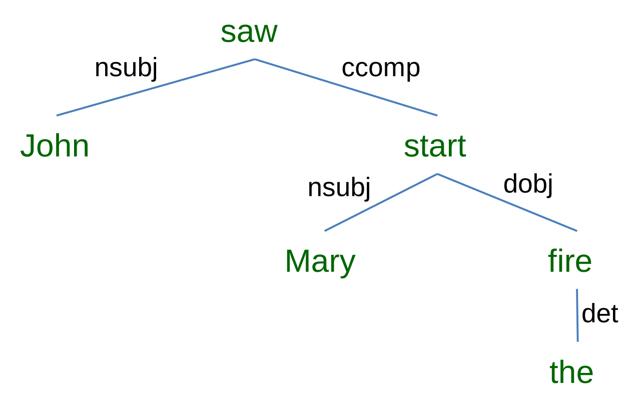
Parsing

Constituency parse
a.k.a.
Phrase Structure parse

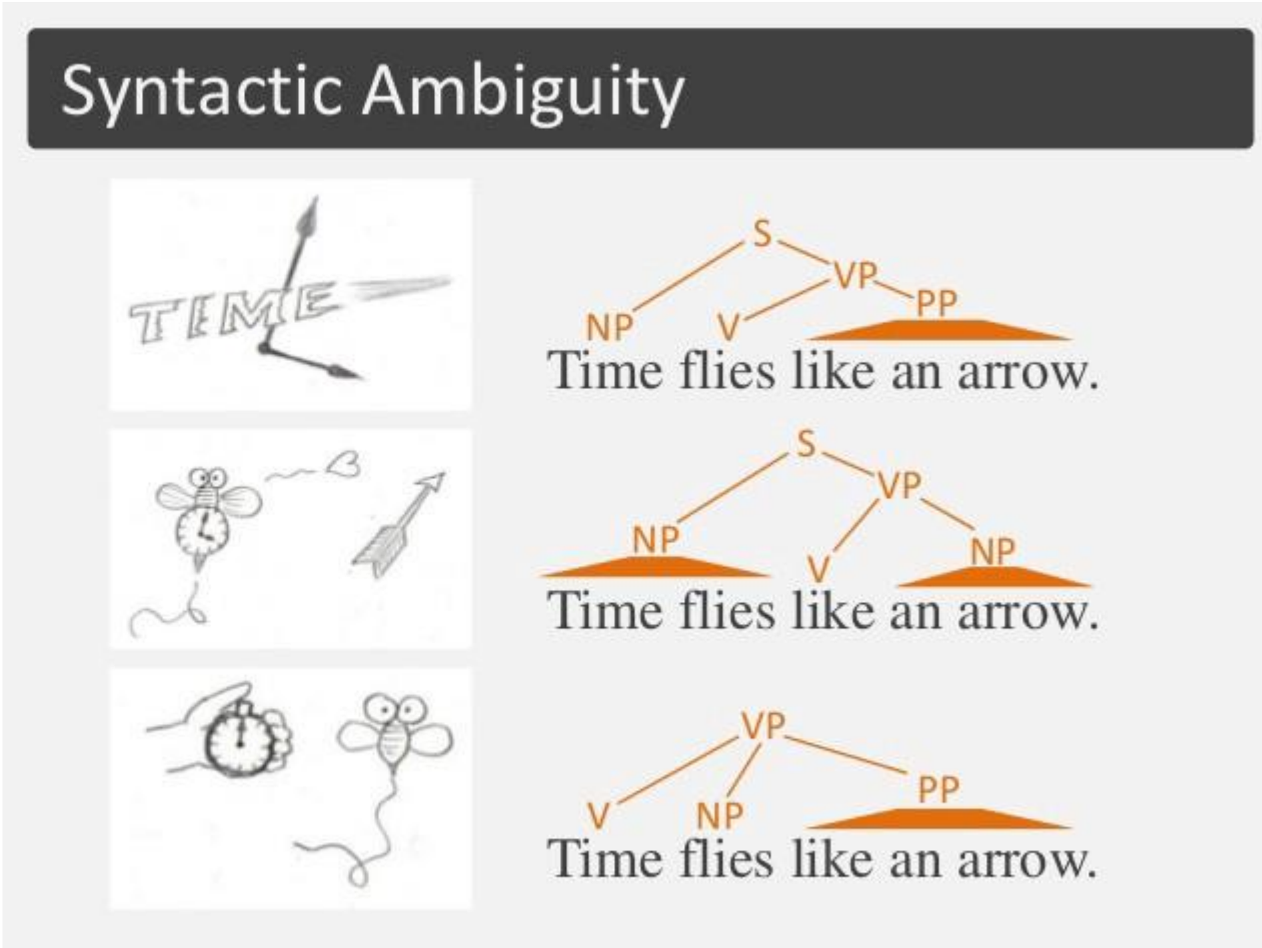


Dependency parse





Ambiguity in Parsing



Exercise: Online Parser

CMU online Constituency Parser demo:

https://www.link.cs.cmu.edu/link/submit-sentence-4.html

or search for: "cmu parse sentence"

spaCy online Dependency Parser demo:

https://explosion.ai/demos/displacy

or search for: "displacy"

Examples to try:

John saw Mary start the fire.

John angled Mary evaporate the duration.

Start saw fire John the Mary.

Time flies like an arrow.

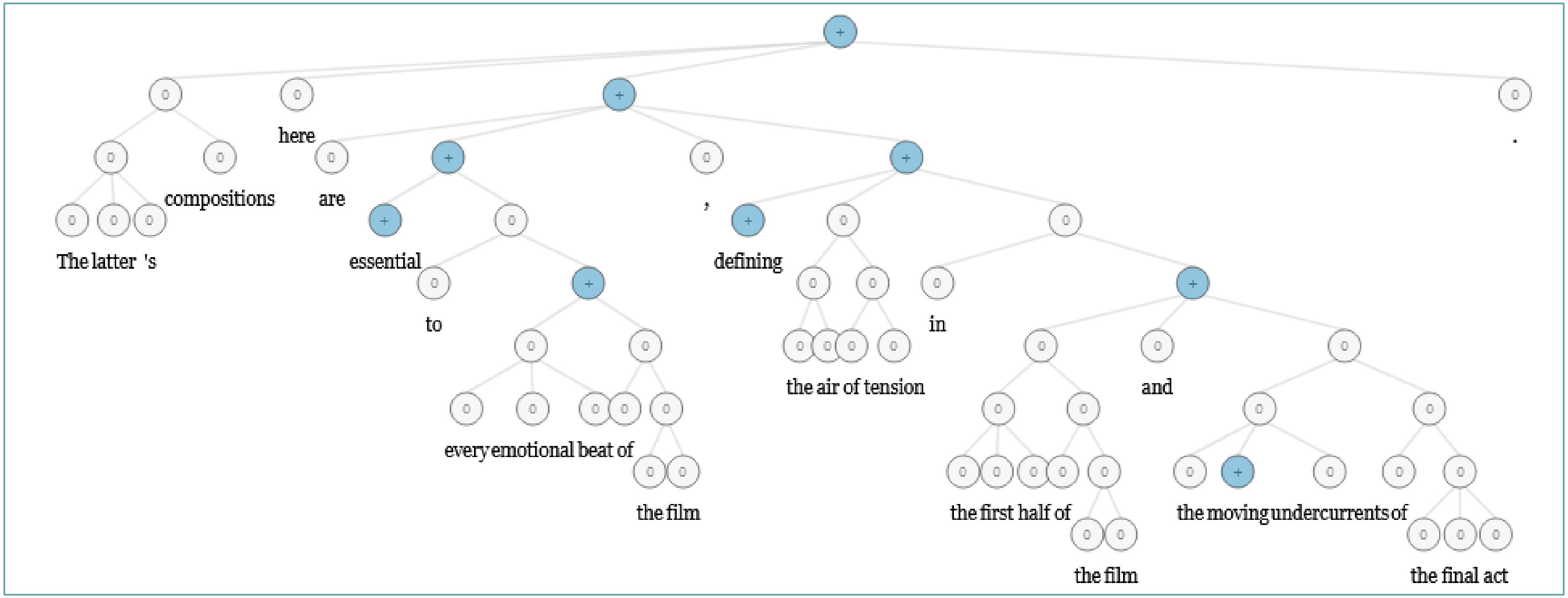
I eat spaghetti with a fork.

I eat spaghetti with meatballs.

Applications I: Sentiment Analysis

sentence from a movie review:

The latter's compositions here are essential to every emotional beat of the film, defining the air of tension in the first half of the film and the moving undercurrents of the final act.

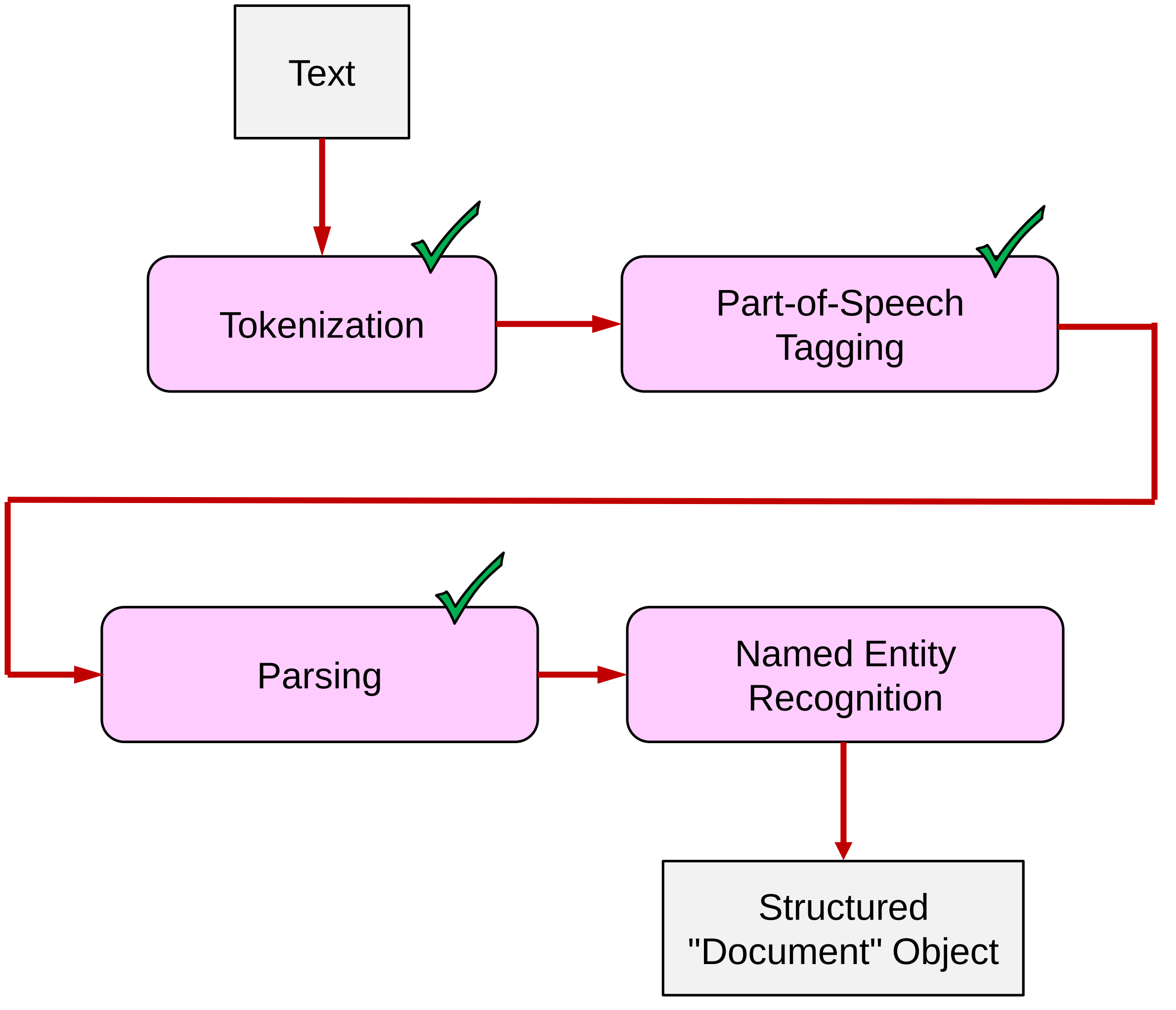


Stanford online movie sentiment analysis demo:

http://nlp.stanford.edu:8080/sentiment/rntnDemo.html

Socher, et al, "Recursive deep models for semantic compositionality over a sentiment treebank" EMNLP 2013

The NLP Pipeline



Named Entity Recognition

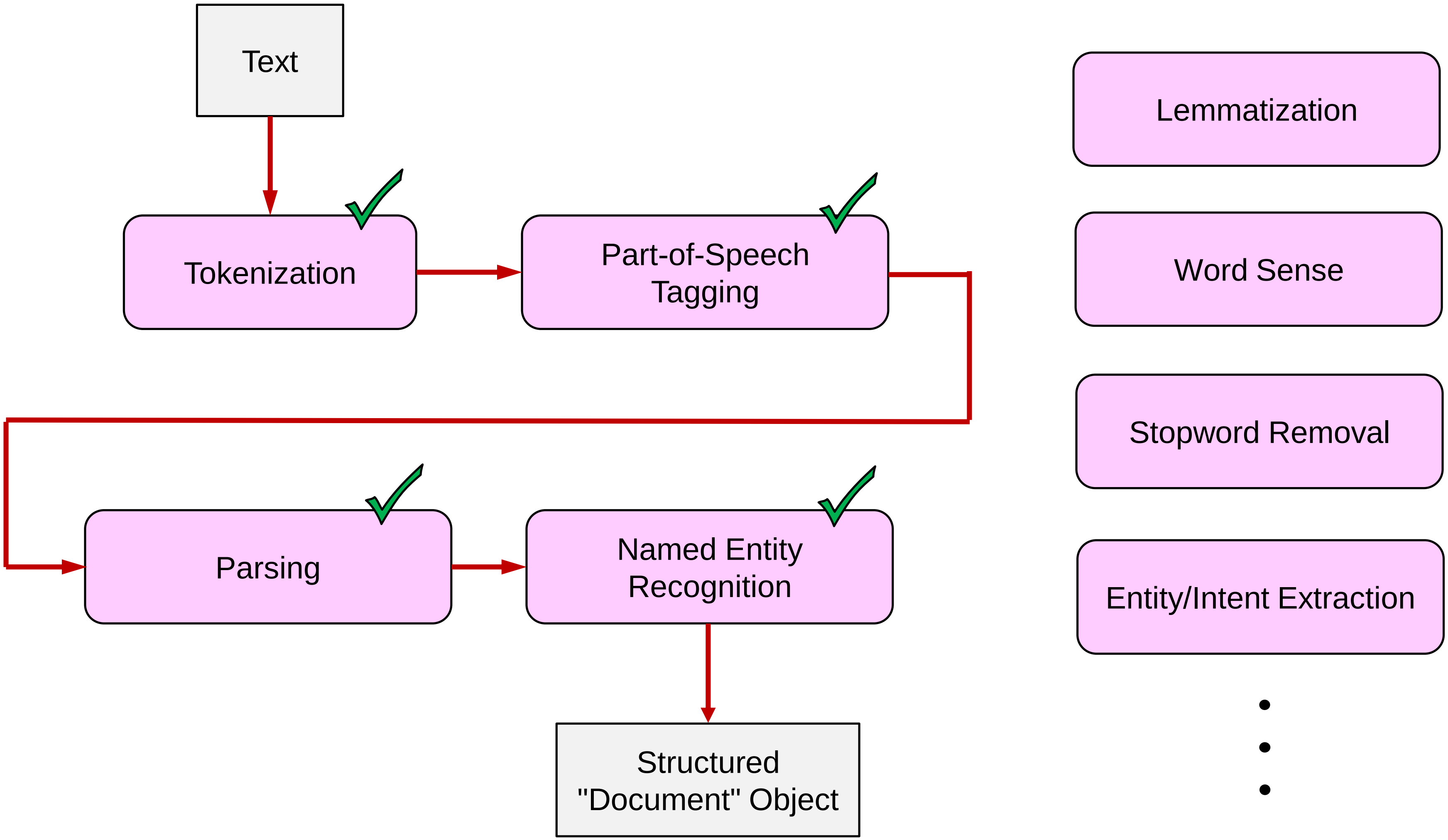
book me a ten a.m. flight from detroit to atlanta on delta next tuesday

Time-of-Day City City Airline Day-of-Week

?

i love the delta painted on that airplane's tail.

The NLP Pipeline



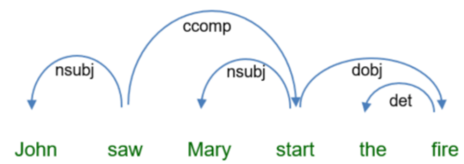
spaCy NLTK Natural Language Toolkit Stanford University NLP Natural Language Processing AllenNLP

Representations in NLP

- **Data frame** — San Francisco 's

  Part-of-Speech: pnoun
  Plurality: singular
  Gender: None
  Possessive: True
  Named Entity:
  Type: City
  DB ref: ...

- **Parse tree**



- **Logical Form** — Please start the fire.

  Request(listener,
  PerformAction(start-fire))

- **Vector embedding**

start	-0.2	0.7	0.1	-0.3	-0.3	-0.7	0.4	0.0
begin	-0.4	0.5	0.1	-0.6	0.1	-1.0	0.5	-0.3
fire	0.5	0.1	0.2	0.2	0.0	-0.5	0.4	-0.8

- **Bag-of-Words**

Topic 1: microbe 10, bacteria 6, infection 4, pathway 5, incubation 2,...
Topic 2: public 9, sanitation 4, spread 7, reports 5,...
Topic 3: meat 0, diet 2, corn 0, beef 0, obesity 1, habits 0, starches 0,...
Topic 4: securities 0, traders 0, market 0, price 0, commodities 0,...
... ...

Document Corpora and Topic Models ("corpora" is plural for "corpus")

Learn (unsupervised) groups (clusters) of words (*Topics*) that co-occur in documents from a corpus.

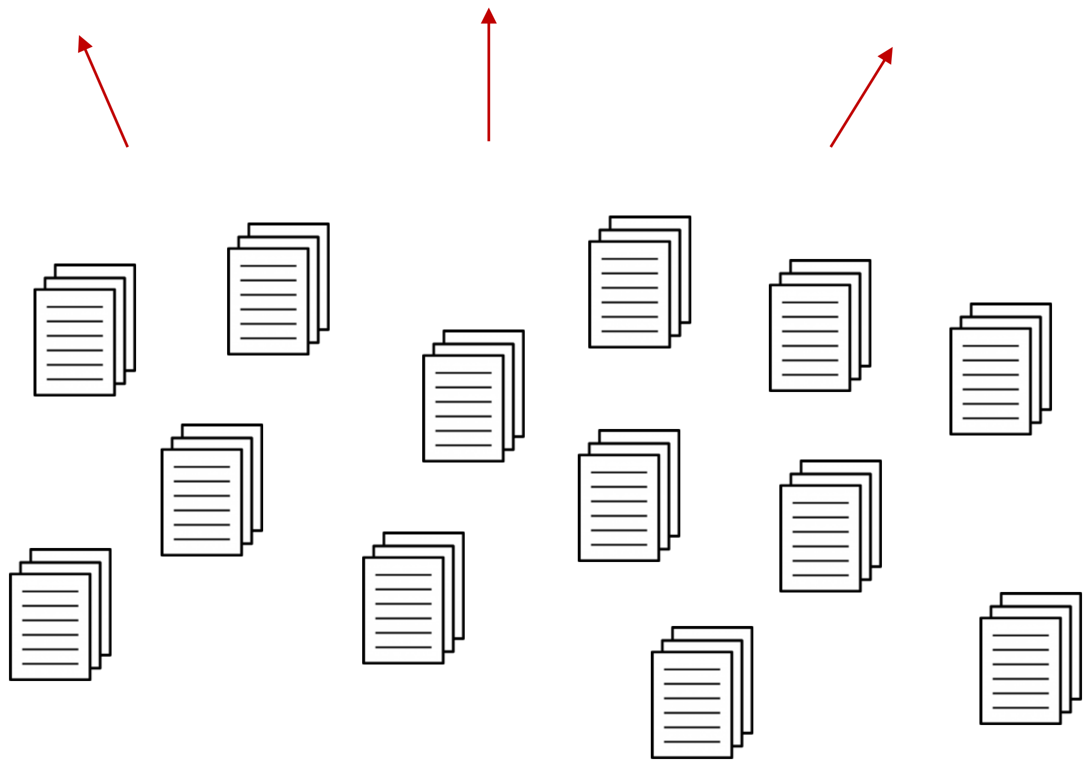
Topic 1: microbe, bacteria, infection, pathway, incubation,...

Topic 2: public, sanitation, spread, reports,...

Topic 3: meat, diet, corn, beef, obesity, habits, starches,...

Topic 4: securities, traders, market, price, commodities,...

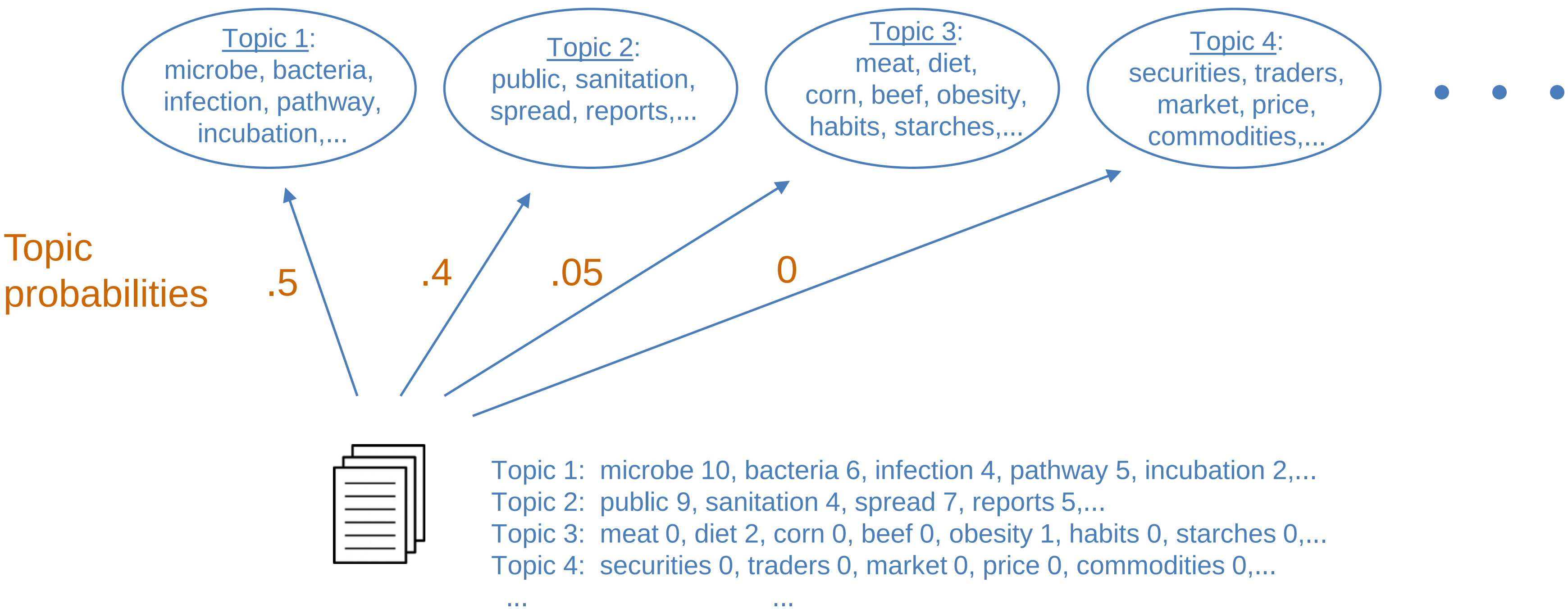
• • •



- Bag-of-Words representation
- Remove uninformative "stopwords": the, is, and, beside, ...
- Clustering Algorithm: LDA = *Latent Dirichlet Allocation*

Topics of a Single Document

For a given document, assign topic weights according to word occurrences from the topic models.



"...Infection profiles are characterized by diverse microbe types including bacteria and fungi. Reports of spreading pathways implicate public sanitation and porous surfaces..."

NLP and NLU: Outline

- Context
  - Ingredients of Intelligence
  - Cognitive Architecture
- The Nature of Language
  - What language is for
  - Syntax and semantics
  - Speech and writing
  - Logic and statistics
- Natural Language Computing Tools
  - The NLP Pipeline
  - Part-Of-Speech, lemmas, relational frames
  - Named Entities
  - Parsers

- Applications I
  - Sentiment

- Representations
  - Bag-of-Words
- Applications II
  - Document topic models

- Knowledge Graphs
- Applications III
  - Question answering
  - Entity/Intent extraction
  - Commonsense knowledge and reasoning
- Deep Methods
  - Language Models
  - Vector embedding
  - Sequence networks
  - Transformer networks
- Applications IV
  - GLUE tasks
  - Machine Translation

- Dialogue
- Applications V
  - Google Assist
  - Call Center Assistants

The Ingredients of Intelligence

Knowledge

- Knowledge Graph

Application:
Q&A Conversational Agent

Pattern Matching

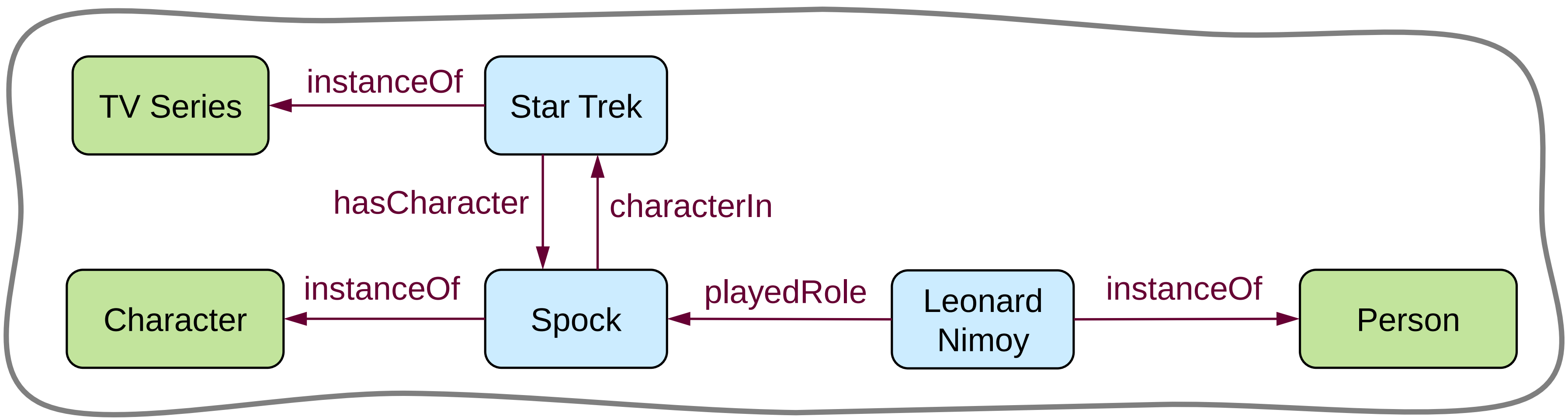
- Entity/Intent Recognition

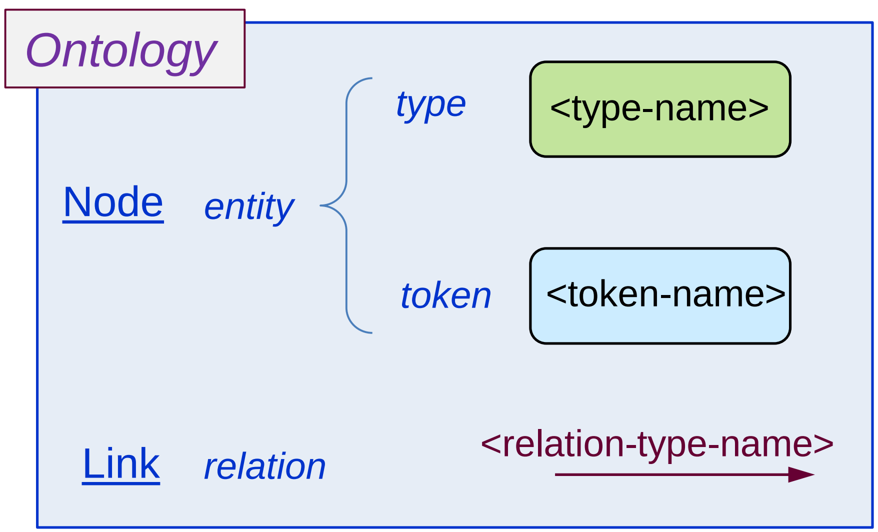
- *Natural Language Understanding*

Reasoning

Knowledge Graphs

- *Knowledge: store of information in organized form*





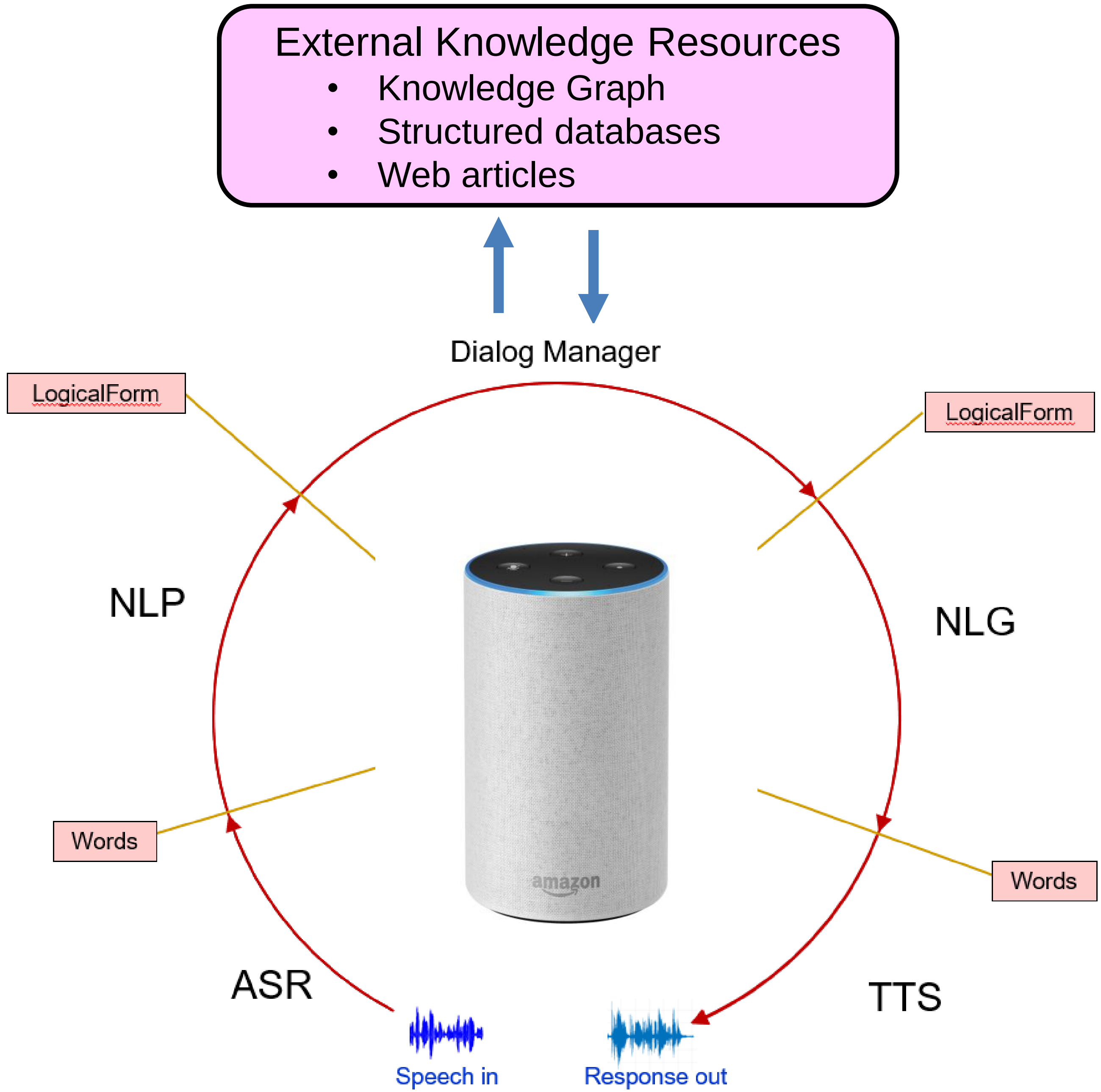
SIZE OF SOME SCHEMA-BASED KNOWLEDGE BASES

Knowledge Graph	Number of Entities	Number of Relation Types	Number of Facts
Freebase[3]	40 M	35,000	637 M
Wikidata[4]	18 M	1,632	66 M
DBpedia (en)[5]	4.6 M	1,367	538 M
YAGO2 [6]	9.8 M	114	447 M
Google Knowledge Graph[7]	570 M	35,000	18,000 M

Nickel etal, 2015

Who played Spock in Star Trek?

Applications III: Question Answering by a Conversational Agent



NLP: From Words to LogicalForm (Meaning)

Who played Spock in Star Trek?

POS,
Named Entity
Recognition:

?		Spock		Star Trek
Person	v.	Character	p.	TV Series

Entity/Intent Extraction

- Question-word
- Command-word

Intent terms

- Entity-Name-word
- Relation Word

Entity/Relation terms

Syntactic Pattern:

Q-Entity1 relation1 Entity2 relation2 Entity3

Logical Form Pattern:

```
{Query(relation1,
       Entity1Type(entity1),
       Entity2Type(entity2,
                   relation2(entity2, entity3))}
```

LogicalForm:

```
Query( playedRole,
       Person ( ? ),
       Character ( Spock,
                   characterIn ( Spock,
                                 TVSeries ( Star Trek ))))
```

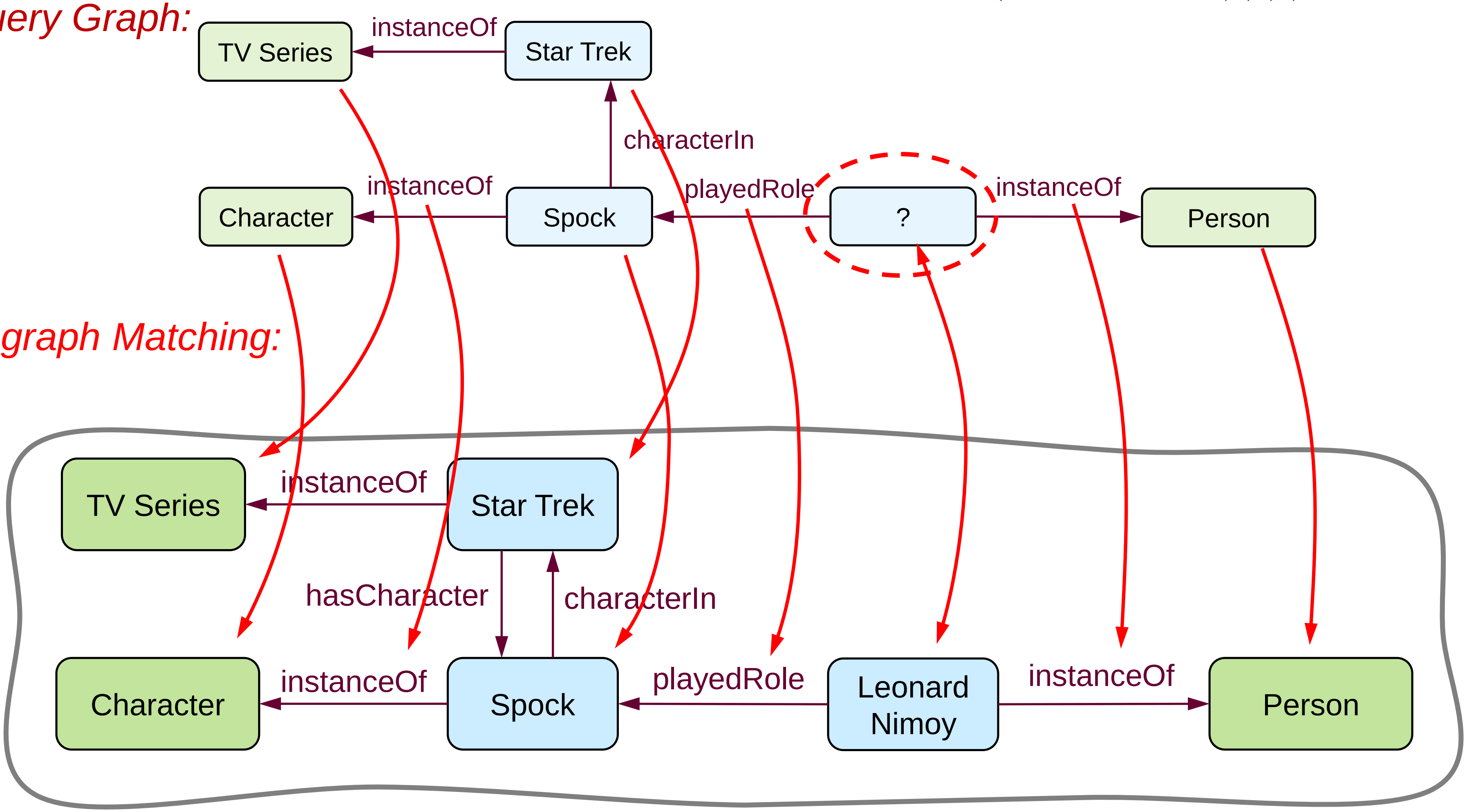
Querying a Knowledge Graph

LogicalForm:

```
Query( playedRole,
       Person ( ? ),
       Character ( Spock,
                   characterIn ( Spock,
                                 TVSeries ( Star Trek ))))
```

Query Graph:

Subgraph Matching:



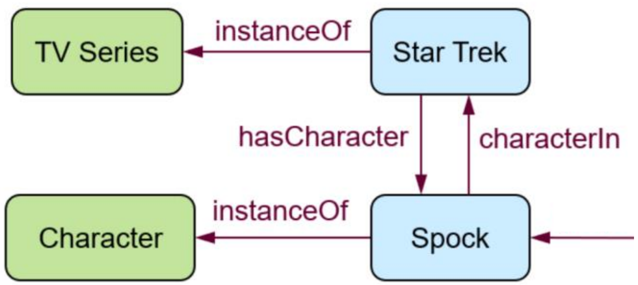
Knowledge Queries in a Conversational Agent (1)

Who played Spock in Star Trek? ✓ Spock was played by Leonard Nimoy in Star Trek, the original series.

Who did Leonard Nimoy play in Star Trek? ✓ ...

Is Leonard Nimoy a Person? ✓ Yes. Leonard Nimoy was an American actor...

Knowledge Queries in a Conversational Agent (2)



Is Star Trek a television series? — Yes. Star Trek, ,the original series...

Is Spock a character in Star Trek? — Yes. Spock is a fictional character...

Is Spock a character in a television series? — Here's something I found on Wikipedia...

Is Star Trek a pineapple? — No. Star Trek is not a pineapple.

Is Star Trek a pineapple in Mexico? — Hmm. I'm not sure

Is Los Angeles a City? — ...

Is Los Angeles a Person? — The Rams are third in the NFC West...

Knowledge Queries and Reasoning

Is a refrigerator heavier than a peanut?

Yes. A refrigerator is heavier than a peanut.

1M Entities × 10 Relations × 1M Entities = 10^{13}

Is Barack Obama taller than William Shatner?

Barack Obama is 6'1" tall, which is 3.48 inches taller than William Shatner. William Shatner is 5'9" tall.

Is a Barack Obama taller than a toaster oven?

Sorry, I don't know that one.

How tall is a toaster oven?

A Black & Decker toaster oven is 9.4 inches tall.

Will a baseball float in beer?

Can you tune a violin with a piano?

- Commonsense knowledge
- Commonsense reasoning

Can you tune a violin with a refrigerator?

The Ingredients of Intelligence

Knowledge

- Knowledge Graph

Pattern Matching

- Entity/Intent Recognition

- *Commonsense knowledge*
- *Commonsense reasoning*

- *Natural Language*
- *Understanding*

Reasoning

NLP and NLU: Outline

- Context
  - Ingredients of Intelligence
  - Cognitive Architecture
- The Nature of Language
  - What language is for
  - Syntax and semantics
  - Speech and writing
  - Logic and statistics
- Natural Language Computing Tools
  - The NLP Pipeline
  - Part-Of-Speech, lemmas, relational frames
  - Named Entities
  - Parsers

- Applications I
  - Sentiment

- Representations
  - Bag-of-Words
- Applications II
  - Document topic models
- Knowledge Graphs
- Applications III
  - Question answering
  - Entity/Intent extraction
  - Commonsense knowledge and reasoning
- Deep Methods
  - Language Models
  - Vector embedding
  - Sequence networks
  - Transformer networks
- Applications IV
  - GLUE tasks
  - Machine Translation

- Dialogue
- Applications V
  - Google Assist
  - Call Center Assistants

Language Models

hand ___?___ predict the next word

Language Models

hand ________

the hand ___?___

Language Models

hand ________

the hand ________

in the hand ___?___

Language Models

hand ________

the hand ________

in the hand ________

bird in the hand ___?___

Language Models

hand ________

the hand ________

in the hand ________

bird in the hand ________

A bird in the hand ? ? ? ? ? ?

Language Models

hand ________

the hand ________

in the hand ________

bird in the hand ________

A bird in the hand ________

A bird in the tropics ___?___

is
can
under
will
sings
flies
.
.
.

Language Models

hand ________

the hand ________

in the hand ________

bird in the hand ________

A bird in the hand ________

A bird in the tropics ________

birds in the tropics ___?___

are	~~is~~
can	can
under	under
will	will
sing	~~sings~~
fly	~~flies~~
.	.
.	.
.	.

Language Models

hand ________

the hand ________

in the hand ________

bird in the hand ________

A bird in the hand ________

A bird in the tropics ________

birds in the tropics ________

in the tropics ___?___

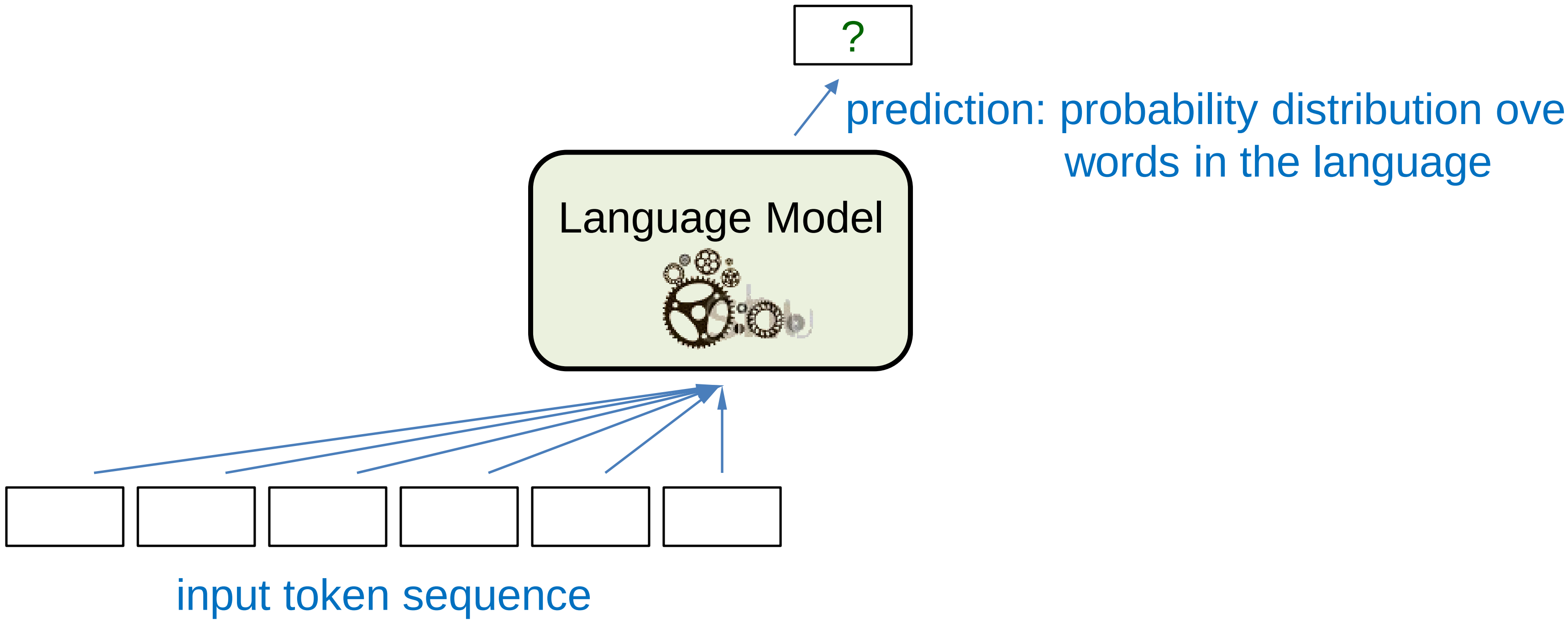
are
can
under
will
sing
fly
.
.
.
there
people
most
of
.
.
.

Language Models

hand ________
the hand ________
in the hand ________
bird in the hand ________
A bird in the hand ________
A bird in the tropics ________
birds in the tropics ________
in the tropics ________
the tropics ___?___

are
can
under
will
sing
fly
.
.
.
there
people
most
of
.
.
.
support
provide
don't
frequently
.
.

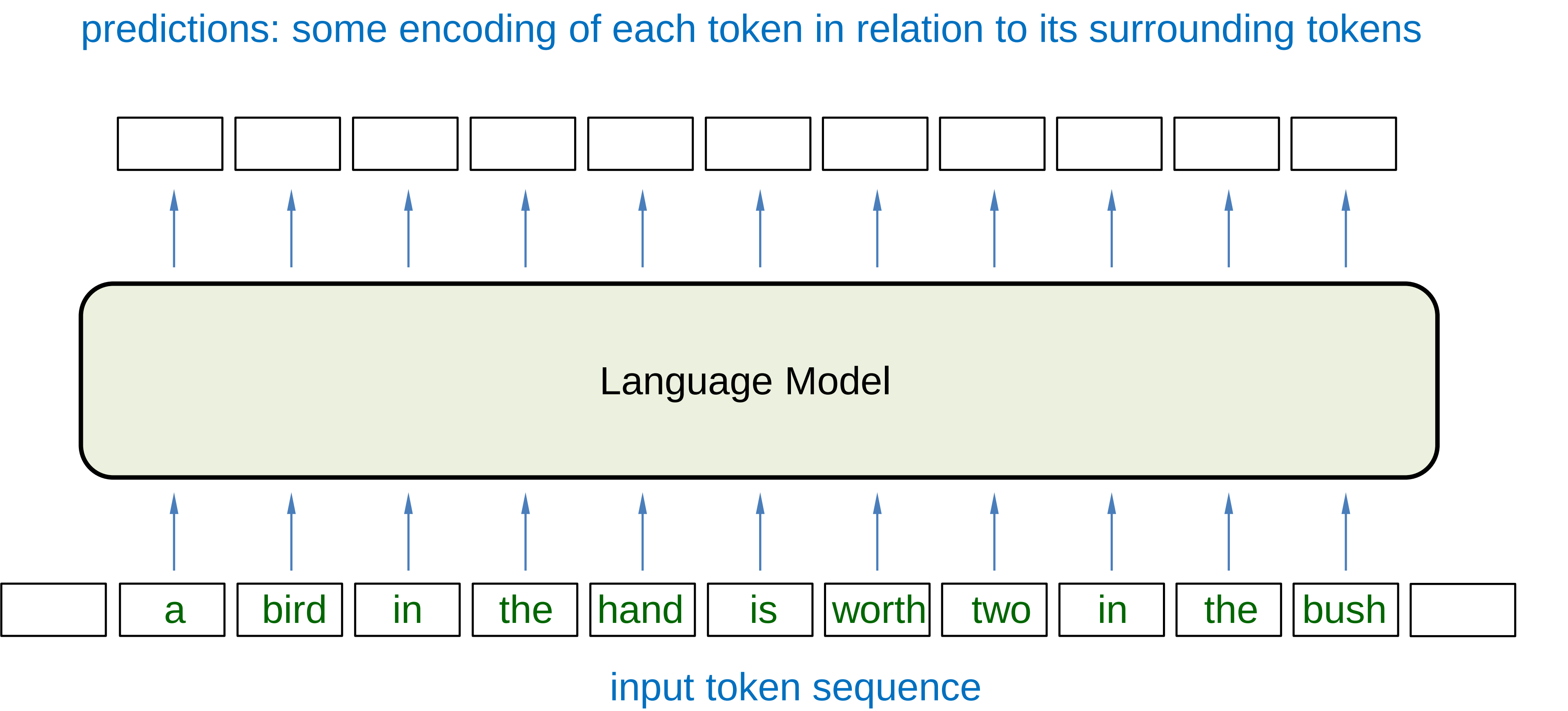
Language Models



A language model captures structure and pattern in token sequences.

- syntax -- rules of grammar
- semantics -- topically-related words
- common phrases
- what else...?

Language Models

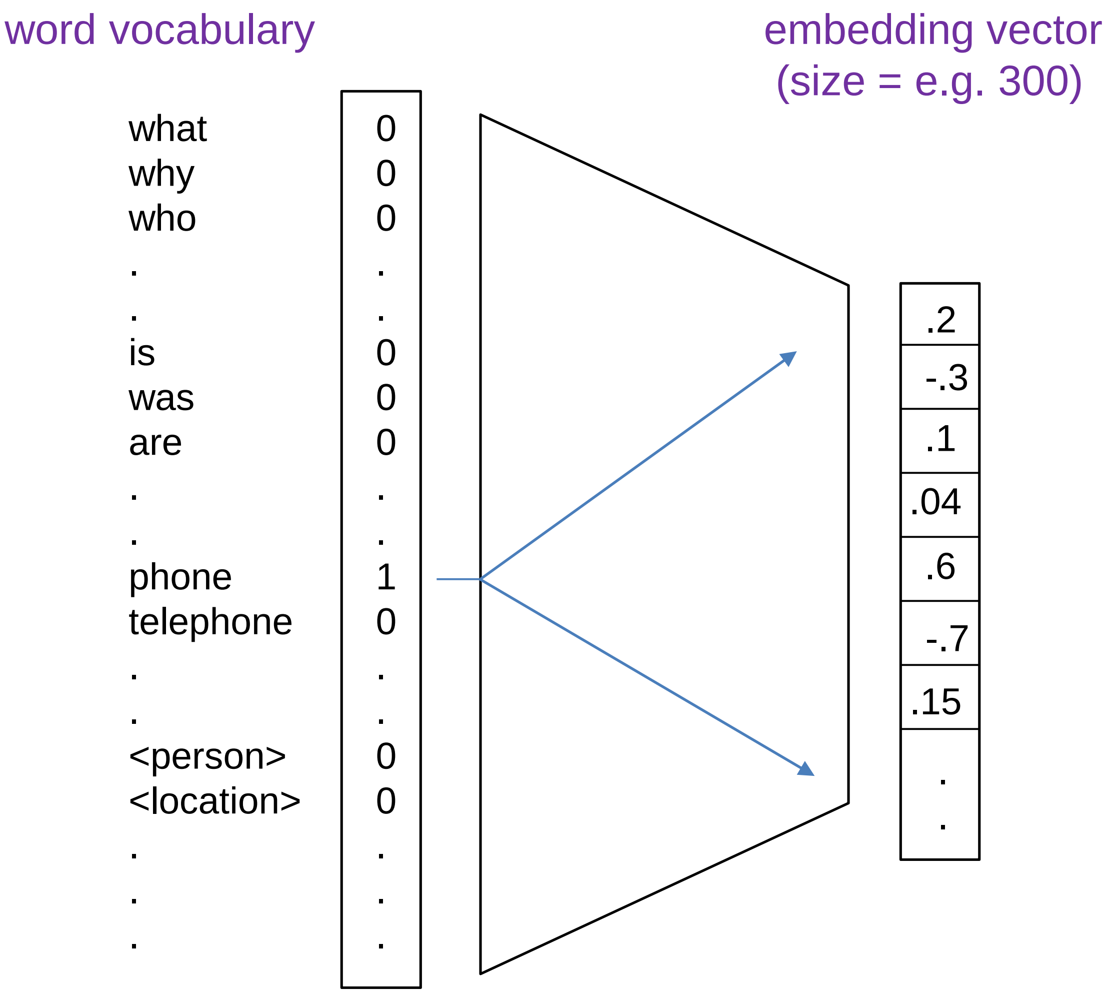


Media for Building World Models

Few model parameters
"Parametric" methods

Many model parameters
Deep Networks

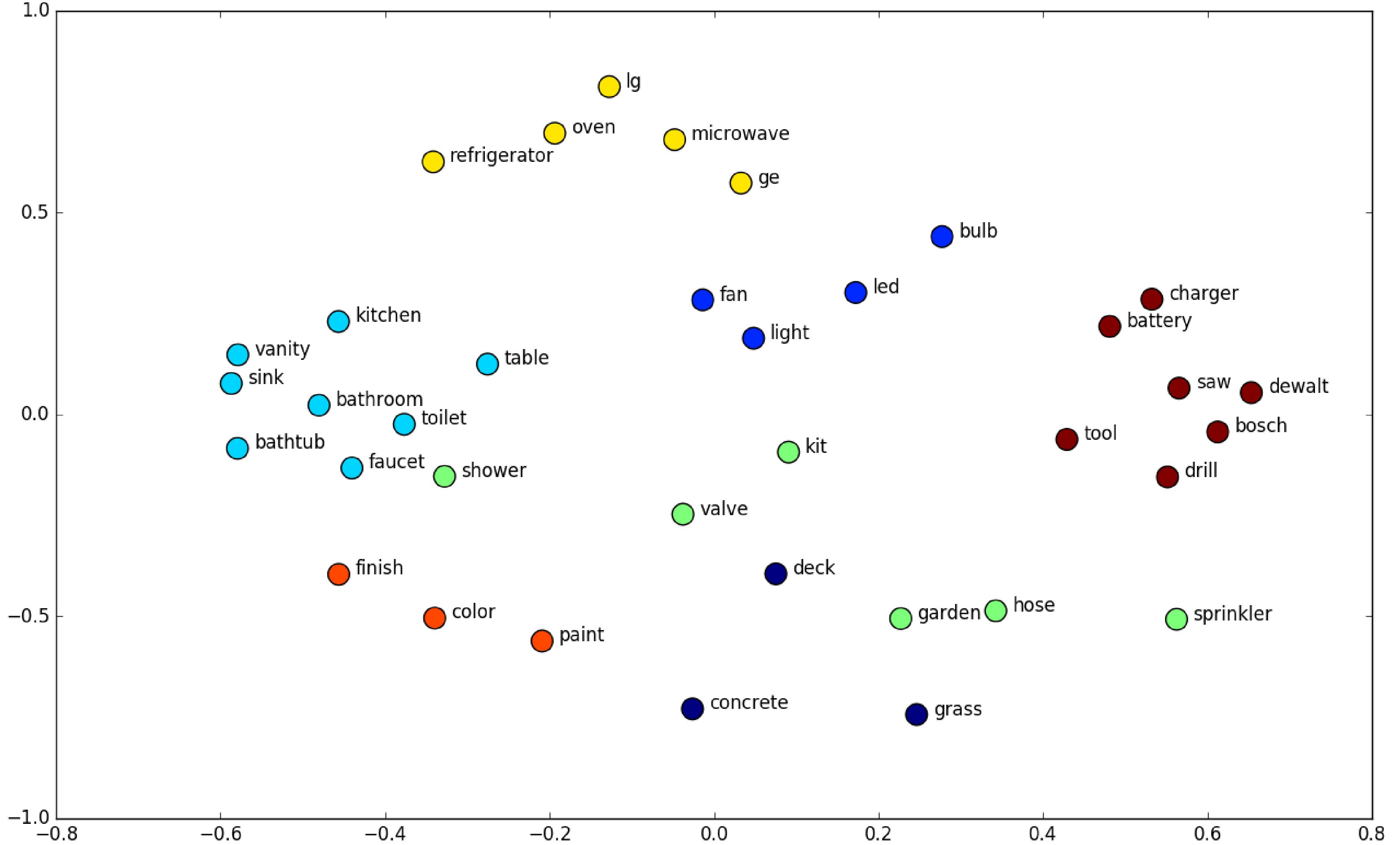
Word Vector Embedding



Purpose:

- lower dimensionality reduces data sparsity for learning
- similar words :: similar vectors   leads to improved generalization

Semantic Similarity in Word Vector Embedding

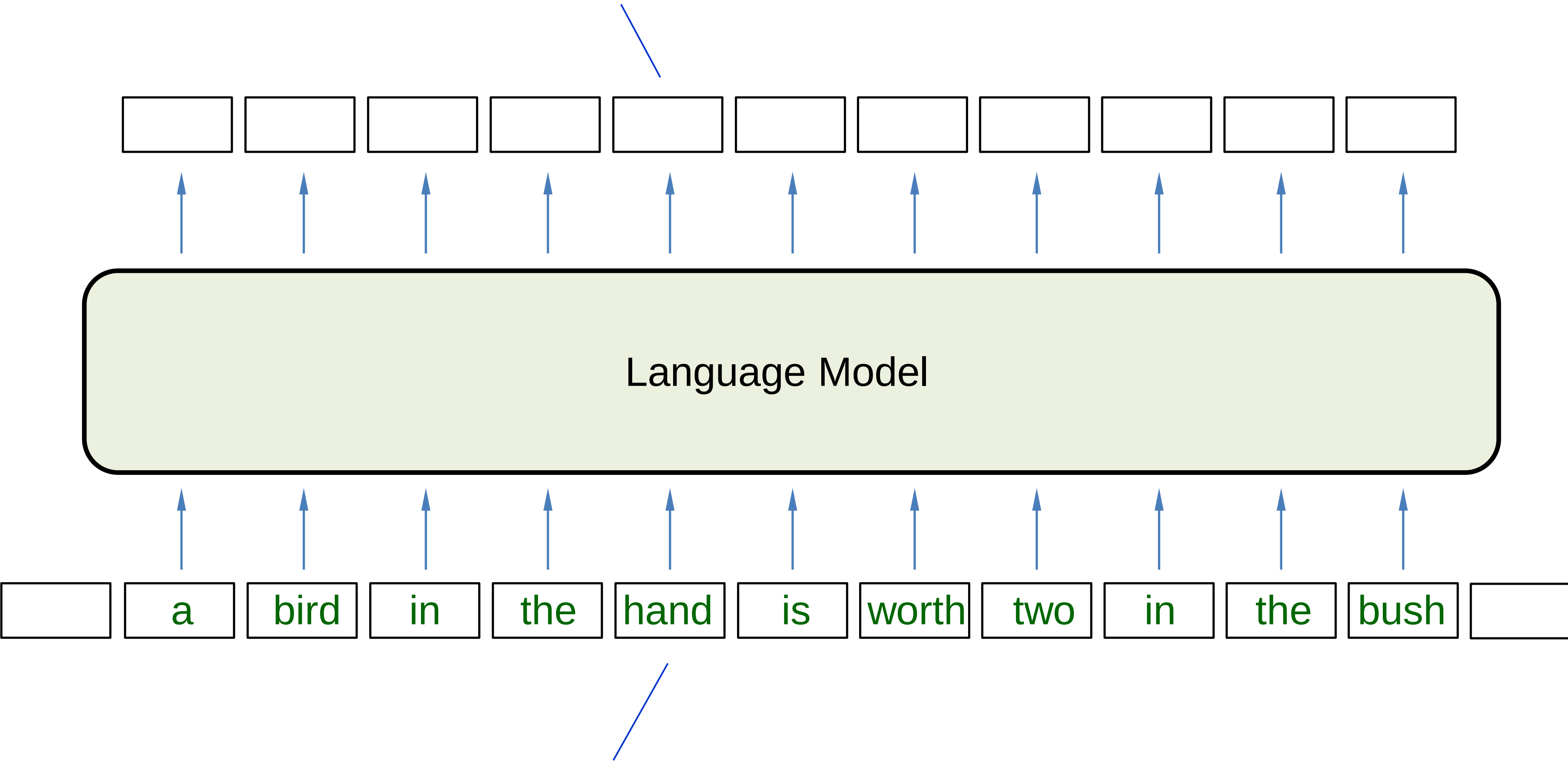


.

.

Language Models

Contextual embedding vector:
"hand", noun form, singular, body part, container of bird, related to worth



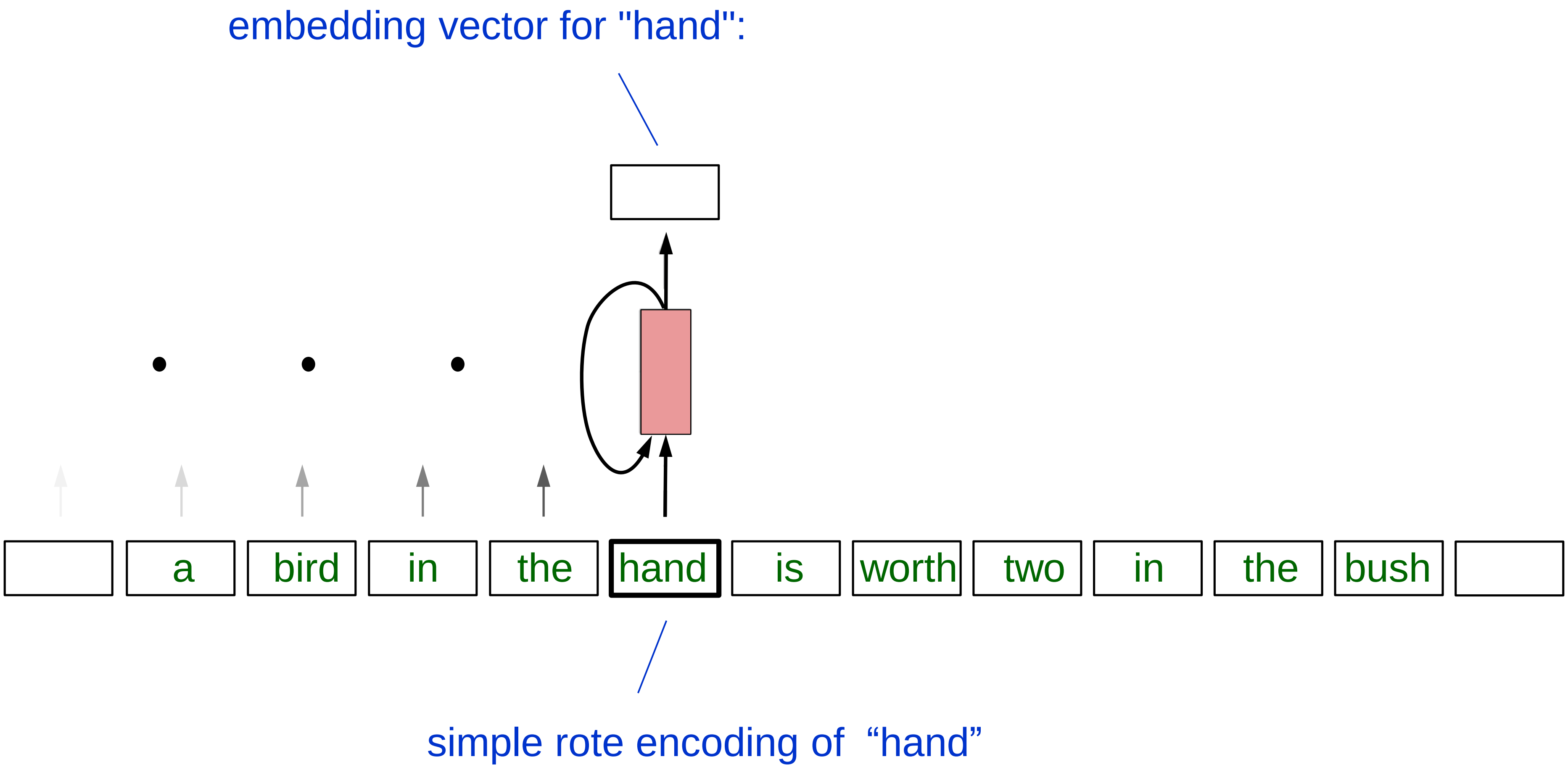
simple rote encoding of “hand”

Language Models using Deep Networks

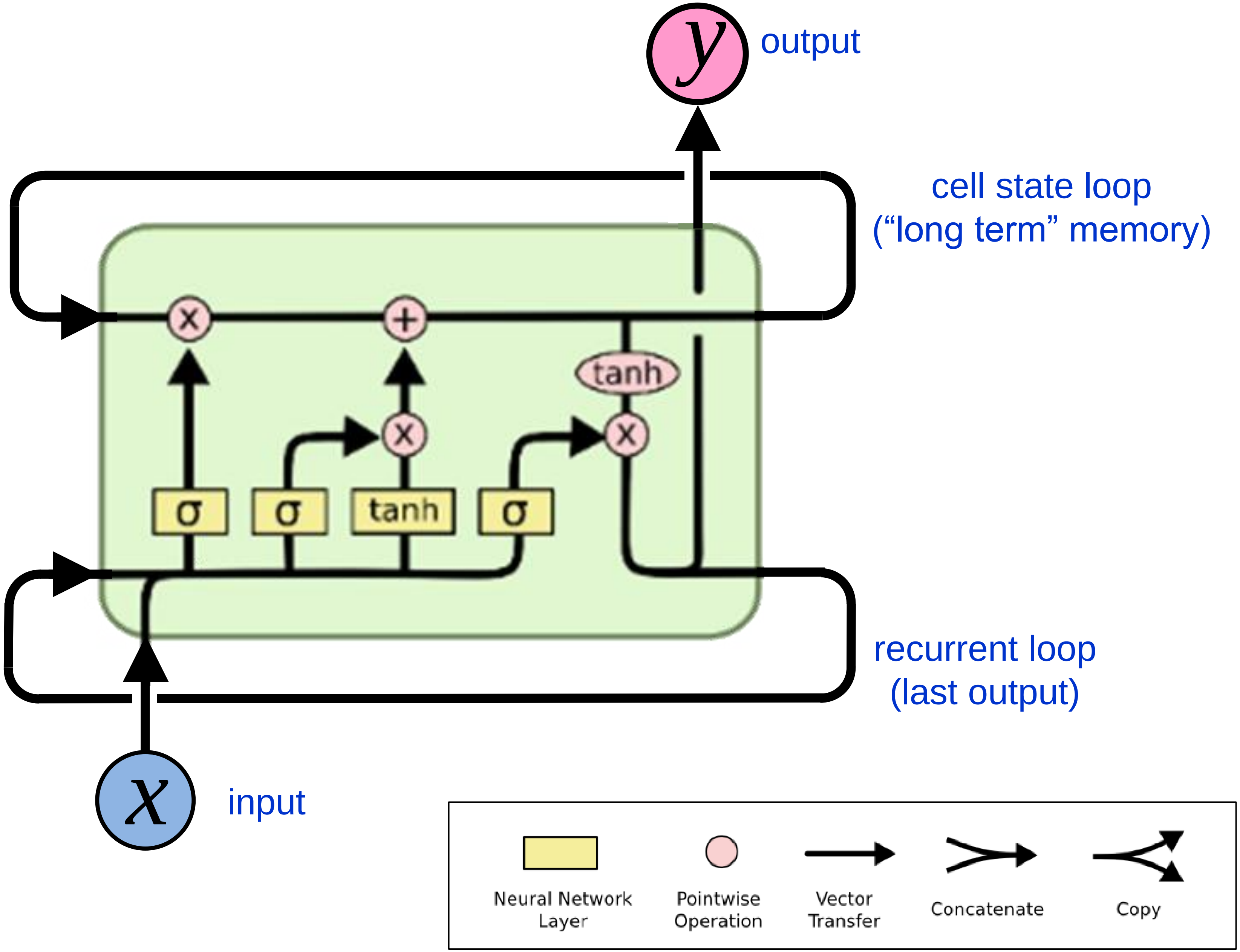
- Sequence Networks
  - ELMo:
    Embeddings from Language Models

- Transformer Networks
  - BERT:
    Bidirectional Encoder Representations from Transformers

Deep Methods: Sequence Networks



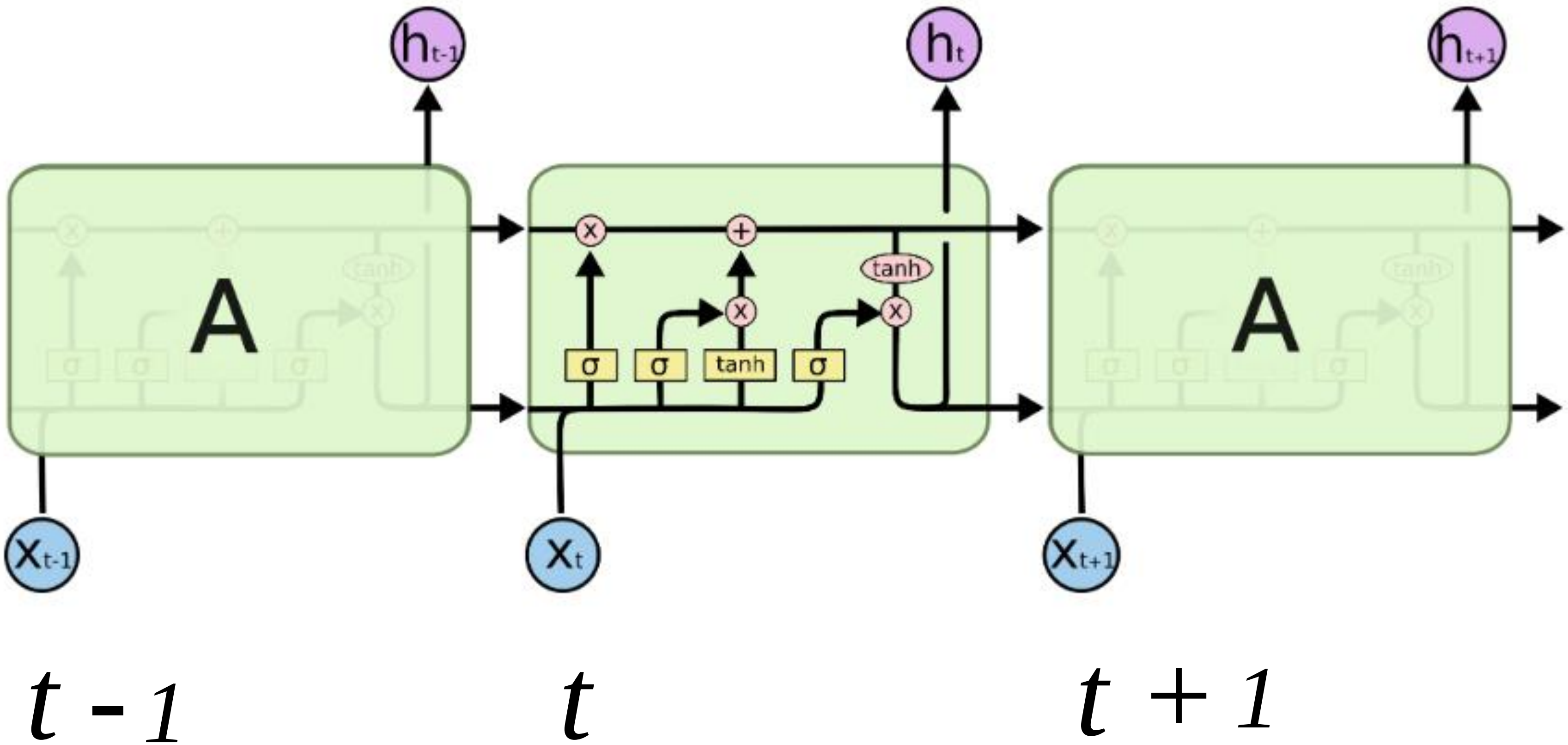
LSTM Module (Long Short Term Memory)



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

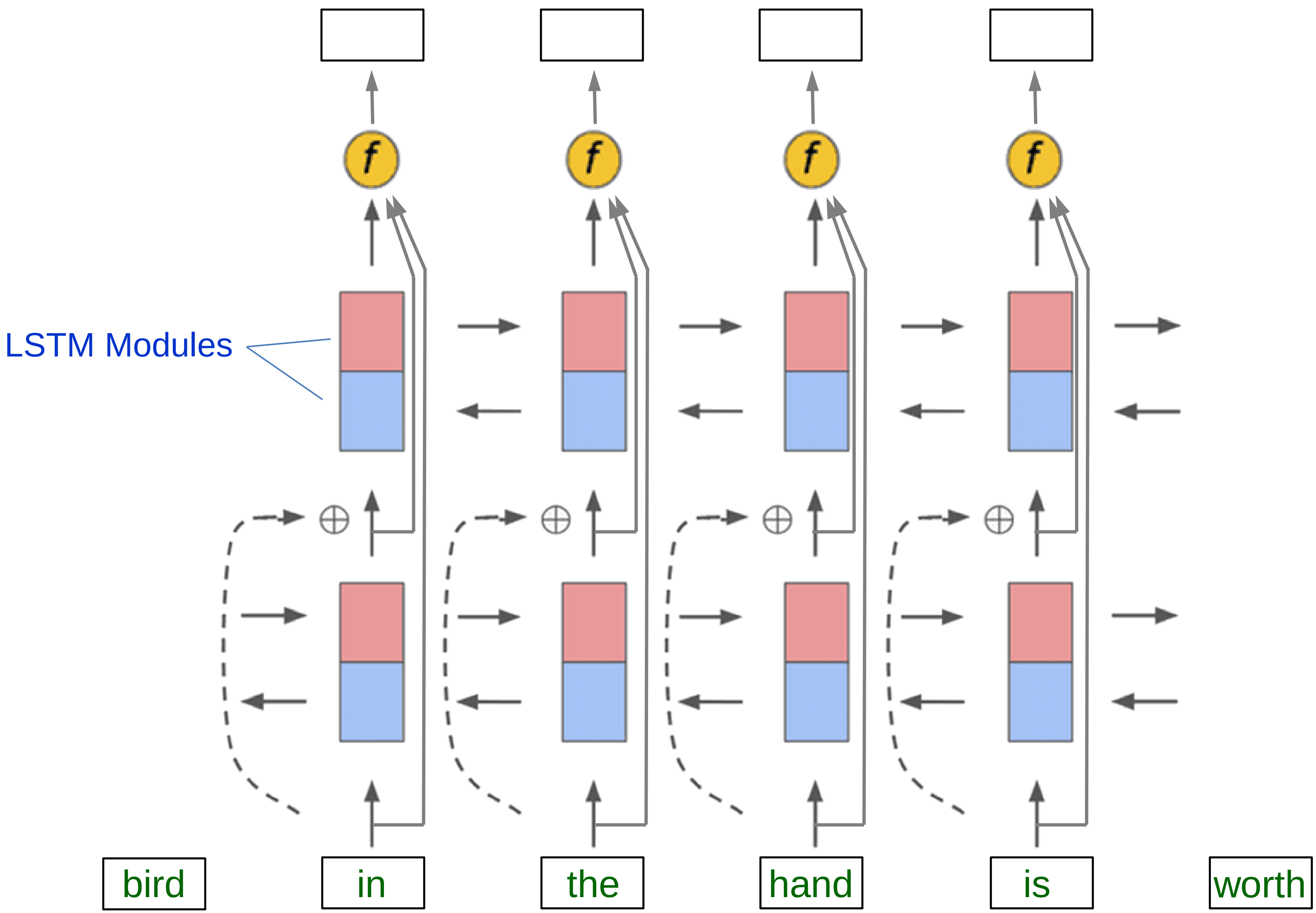
LSTM Module

"Unrolled" depiction



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Deep Methods: ELMo Unrolled



Deep Methods: BERT Transformer Network

10^8 parameters

N = 12 layers

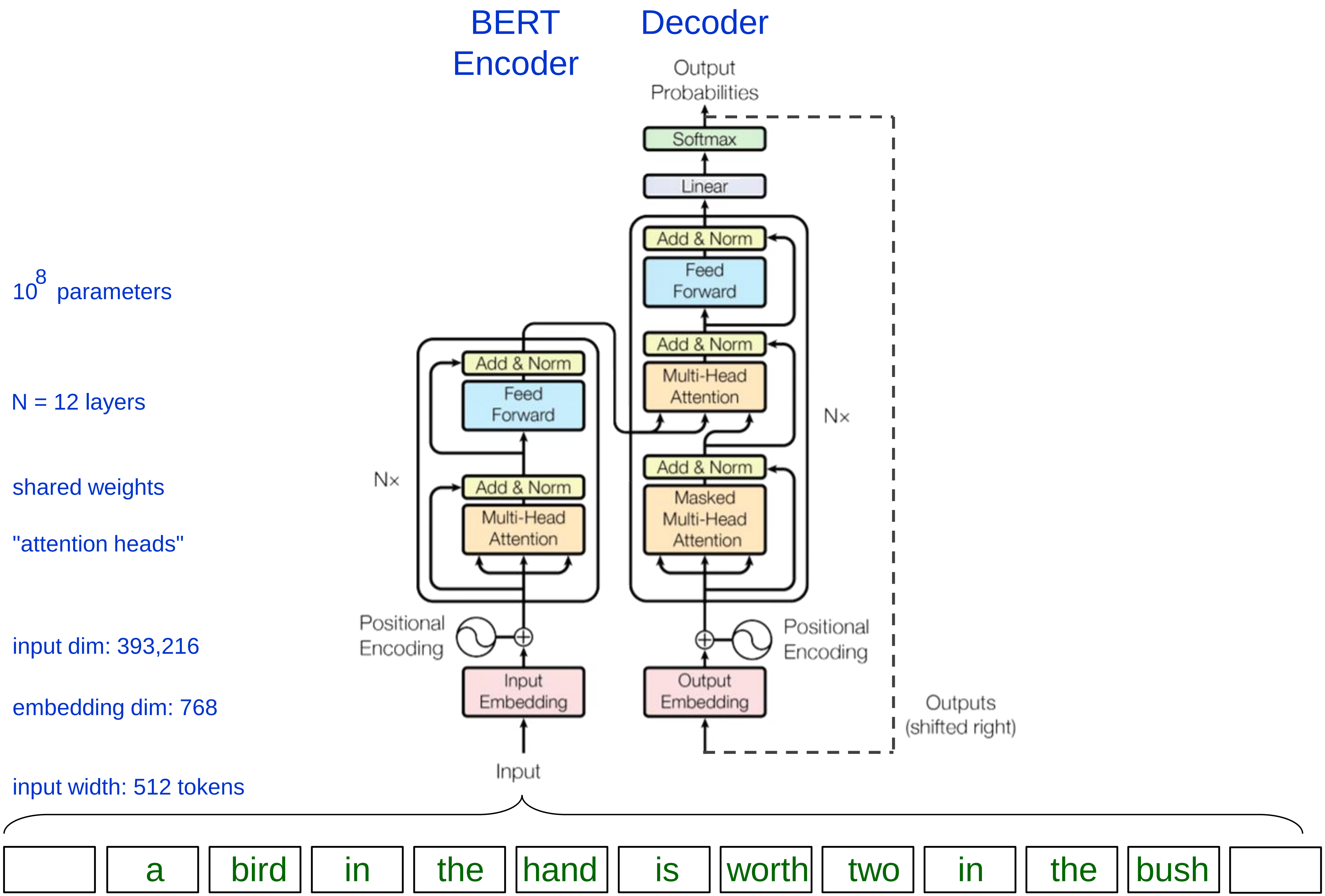
shared weights

"attention heads"

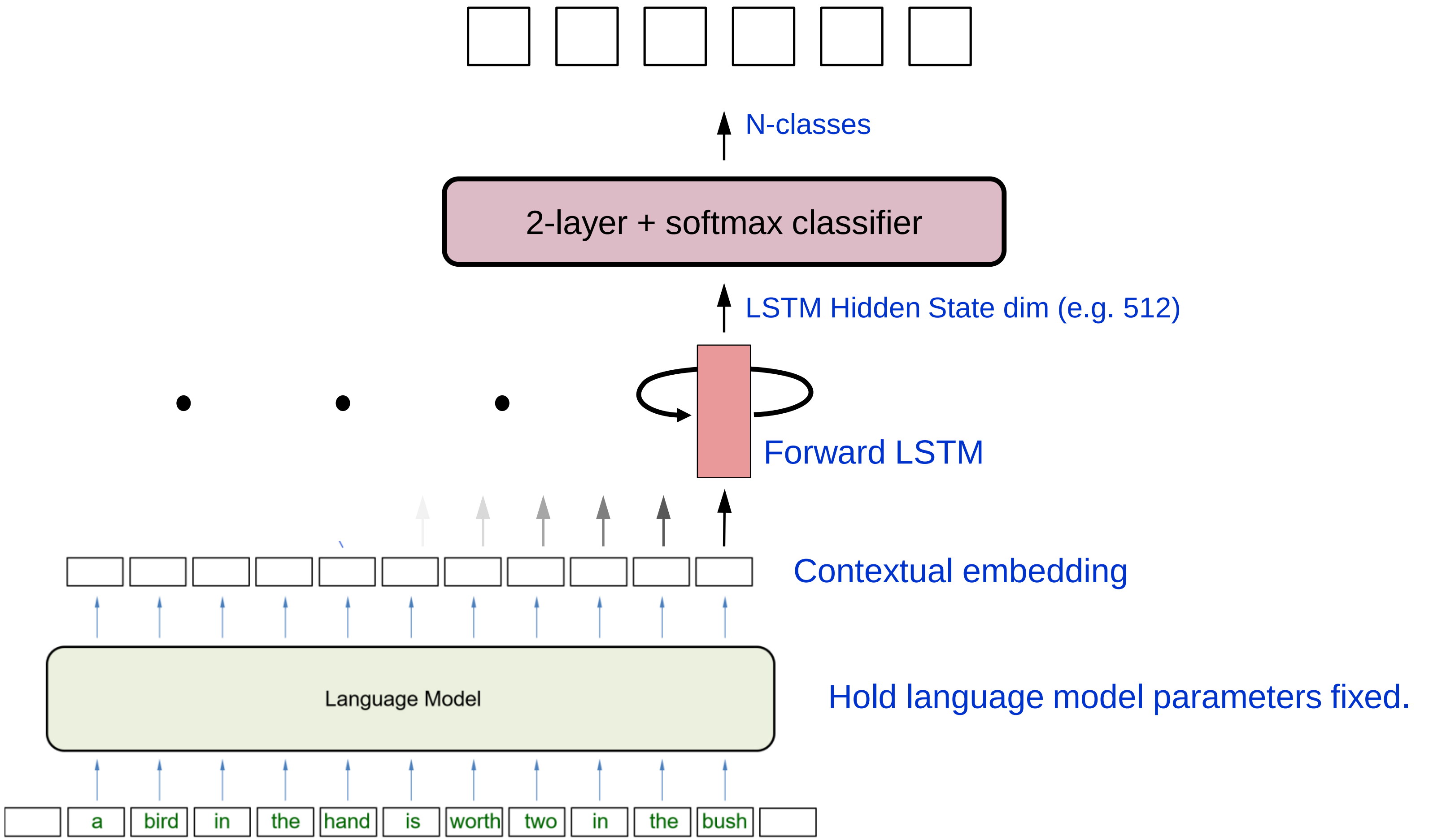
input dim: 393,216

embedding dim: 768

input width: 512 tokens



Deep Methods: Using a Language Model



Evaluation Tasks

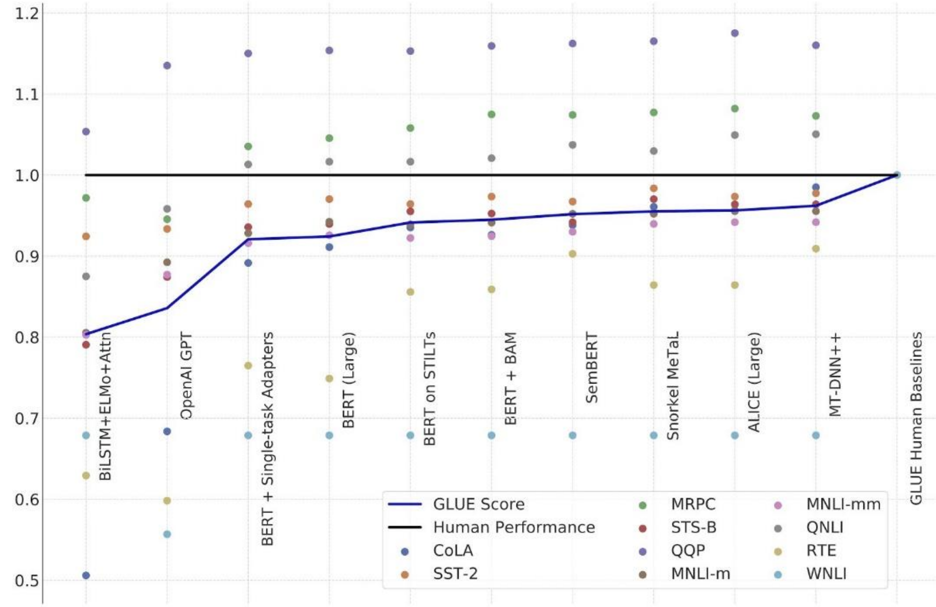
GLUE: General Language Understanding Evaluation

- Part-of-Speech tagging
- word sense
- sentiment
- question answering
- semantic entailment

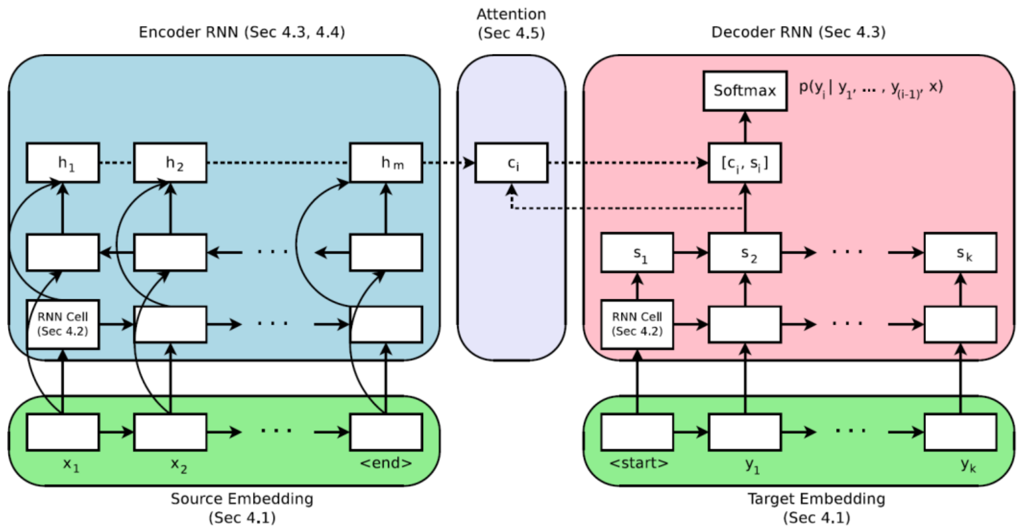
Entailment question examples: Output range: [entailment / neutral / contradicted]

Premise: Your gift is appreciated by each and every student who will benefit from your generosity.
Hypothesis: Hundreds of students will benefit from your generosity.
Label: neutral

Premise: At the other end of Pennsylvania Avenue, people began to line up for a White House tour.
Hypothesis: People formed a line at the end of Pennsylvania Avenue.
Label: entailment



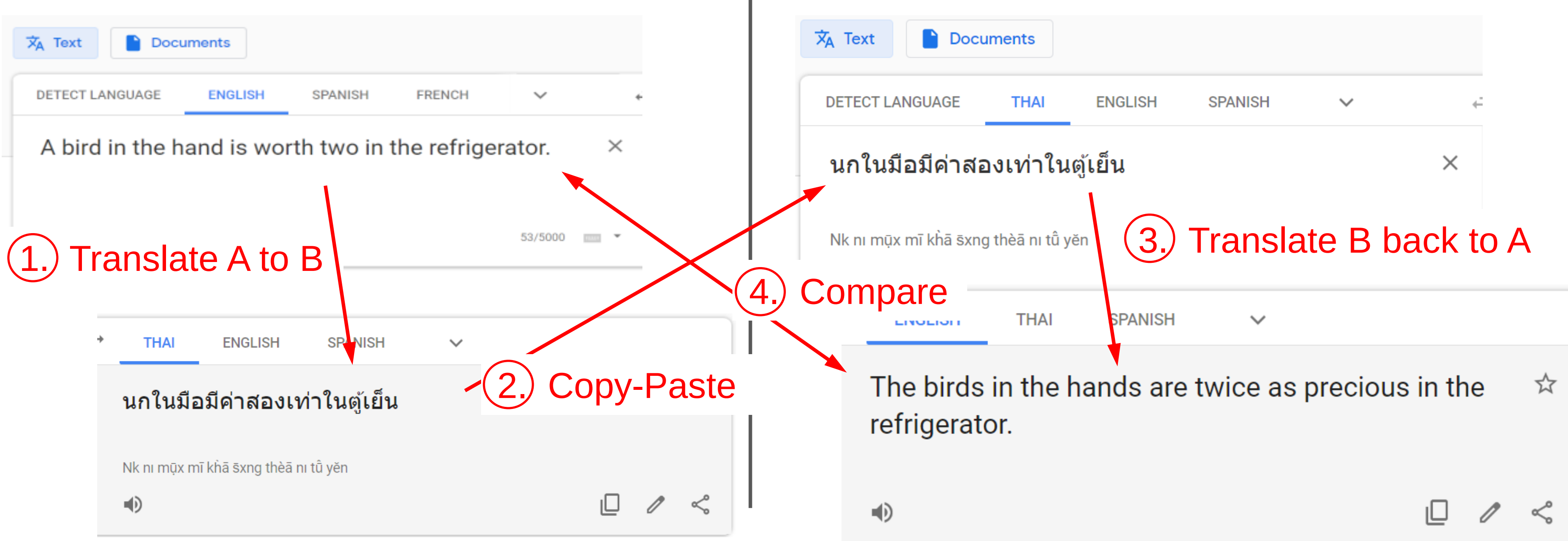
Applications IV: Neural Machine Translation



Britz etal, 2017

https://translate.google.com/

https://translate.google.com/



NLP and NLU: Outline

- ✓ Context
  - Ingredients of Intelligence
  - Cognitive Architecture
- ✓ The Nature of Language
  - What language is for
  - Syntax and semantics
  - Speech and writing
  - Logic and statistics
- ✓ Natural Language Computing Tools
  - The NLP Pipeline
  - Part-Of-Speech, lemmas, relational frames
  - Named Entities
  - Parsers

- ✓ Applications I
  - Sentiment

- ✓ Representations
  - Bag-of-Words
- ✓ Applications II
  - Document topic models
- ✓ Knowledge Graphs
- ✓ Applications III
  - Question answering
  - Entity/Intent extraction
  - Commonsense knowledge and reasoning
- ✓ Deep Methods
  - Language Models
  - Vector embedding
  - Sequence networks
  - Transformer networks
- ✓ Applications IV
  - GLUE tasks
  - Machine Translation

- Dialogue
- Applications V
  - Google Assist
  - Call Center Assistants

Dialogue Understanding

Alex: Well, good morning. I get a...a vanilla latte?

Jill: Uh huh. Sure can. Do you want that for here...or do you want it in a mug or...?

Alex: For here, for here.

Jill: In a mug or in a...

Alex: No no just in a paper cup.

Jill: Okay. All right, and what else?

Alex: And uh...a blueberry bagel.

Jill: Sliced and...

Alex: *Sliced and toasted.*

Jill: ...toasted.

Alex: And with a ... butter and uh jam please.

Jill: Ok. And we have grape jelly. Is that okay?

Alex: Hmm?

Jill: Grape jelly. Is that okay?

Alex: Ya whatever, whatever.

Jill: Alright.

(28 seconds)

Conversation Analysis

- Structured
  - opening, closing
  - Q & A
  - turn-taking
  - confirmations
  - request repeat
  - echoing
  - repair
- Disfluencies, repetition
- Overtalk
- Adaptive
- Reactive
- Contextual

Dialogue Management in Conversational Agents and Chatbots

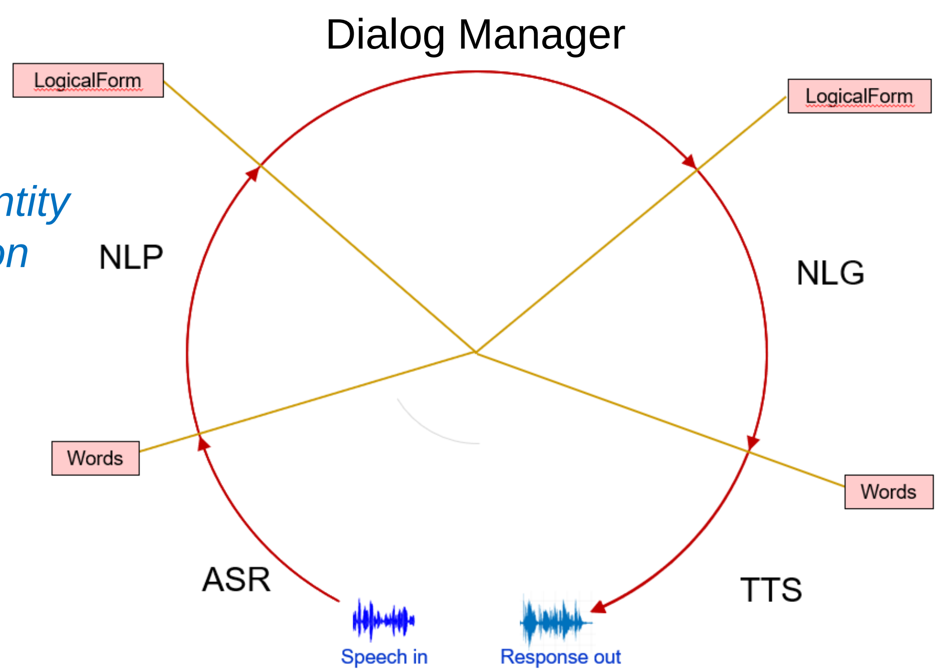
Pipeline Tools

- Amazon Lex
- Google Dialogflow
- IBM Watson Assistant
- Microsoft Azure Bot Service
- Rasa.ai
- …

Slot	Value
Service:	woman's haircut
Name:	Lisa
Date:	May 3
Time:	12:00pm, morning

- *Slot filling*
- *Custom rule logic*
- *Learned policy*

- *Intent/Entity extraction*



Applications V: Google Assistant (Duplex)

Hair Salon: Hello how can I help you?

Google Assistant: Hi, I'm calling to book a women's haircut for a client. Um, I'm looking for something on May 3rd.

Hair Salon: Sure, give me one second.

GA: Mm-hmm.

Hair Salon: Sure what time are you looking for around?

GA: At 12 pm.

Hair Salon: We do not have a 12 pm available. The closest we have to that is a 1:15.

GA: Do you have anything between 10 am and uh12 pm?

Hair Salon: Depending on what service she would like. What service is she looking for?

GA: Just a woman's haircut, for now.

Hair Salon: Okay we have a 10 o'clock.

GA: 10 a.m. is fine.

Hair Salon: Okay, what's her first name?

GA: The first name is Lisa.

Hair Salon: Okay perfect. So I will see Lisa at 10 o'clock on May 3rd.

GA: Okay great, thanks.

Hair Salon: Great. Have a great day. Bye.

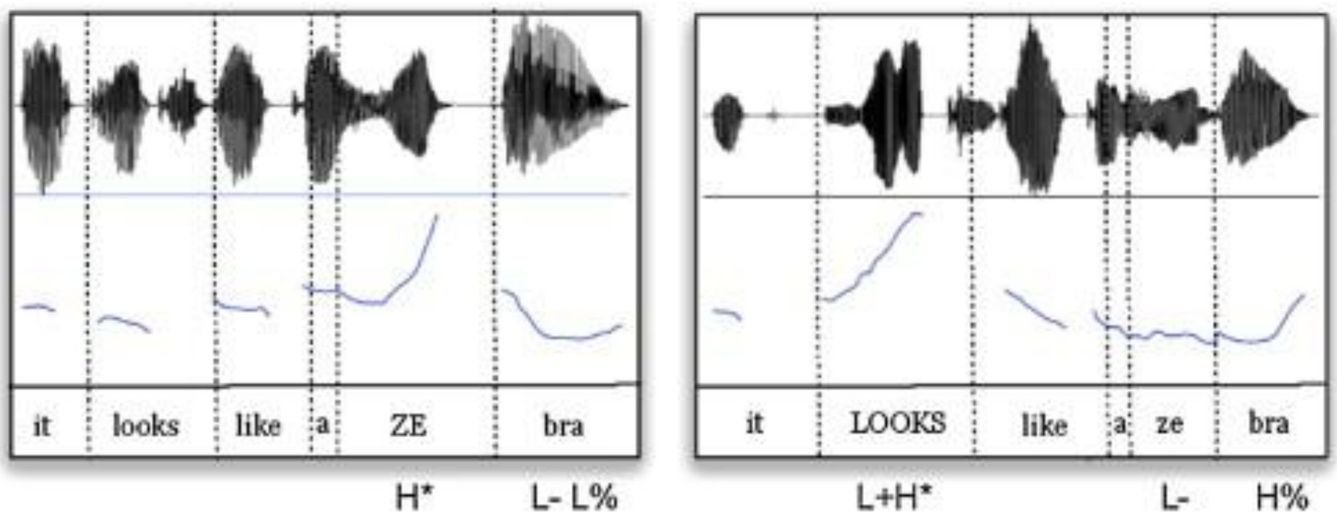
(56 seconds)

What is Google Duplex?
https://www.cnet.com/how-to/what-is-google-duplex/

- Human intervention:
  2019: 60% of reservations handled end-to-end by GA

Prosody, Intonation, Emotion

It looks like a zebra.



(a) Noun-focus prosody

(b) Verb-focus prosody

Kurumada etal, 2014

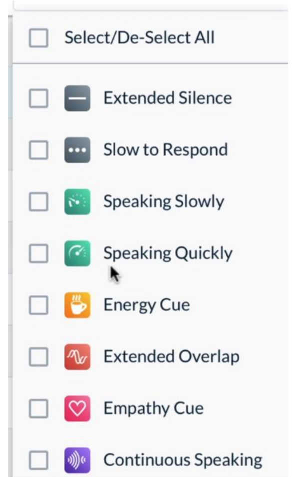
Applications VI: Call Center Assist

- Audio processing only, no NLP (yet)

cogito

Real-time conversational guidance

Cogito detects human signals and provides live behavioral guidance to improve the quality of every interaction.



Call Evaluation Features

Recommendation to Agent

- ASR, NLP, audio

Gridspace

Take command of your contact center

Gridspace Sift gives you the power to capture, understand, and handle conversations in real-time. We process tens of millions of interactions for top businesses every year.

NLP and NLU: Outline

- Context
    - Ingredients of Intelligence
    - Cognitive Architecture
- The Nature of Language
    - What language is for
    - Syntax and semantics
    - Speech and writing
    - Logic and statistics
- Natural Language Computing Tools
    - The NLP Pipeline
    - Part-Of-Speech, lemmas, relational frames
    - Named Entities
    - Parsers

- Applications I
    - Sentiment

- Representations
    - Bag-of-Words
- Applications II
    - Document topic models
- Knowledge Graphs
- Applications III
    - Question answering
    - Entity/Intent extraction
    - Commonsense knowledge and reasoning
- Deep Methods
    - Language Models
    - Vector embedding
    - Sequence networks
    - Transformer networks
- Applications IV
    - GLUE tasks
    - Machine Translation

- Dialogue
- Applications V
    - Google Assist
    - Call Center Assistants

Summary

Natural Language…

Processing

Understanding

<u>Methods & Technologies</u>

NLP Pipeline

Statistical/ML Tools

Knowledge Graphs

Language Models

Deep Networks

<u>Applications</u>

Document Processing

Question Answering

Machine Translation

Conversational Assistants

<u>Frontier</u>

<u>Knowledge</u>

<u>Pattern Matching</u>

<u>Reasoning</u>

Cognitive Architecture

Eric Saund

- *Research scientist in Cognitive Science and AI.*
- *Conversational Agents, Visual Perception, Cognitive Architectures.*
- *I build stuff.*

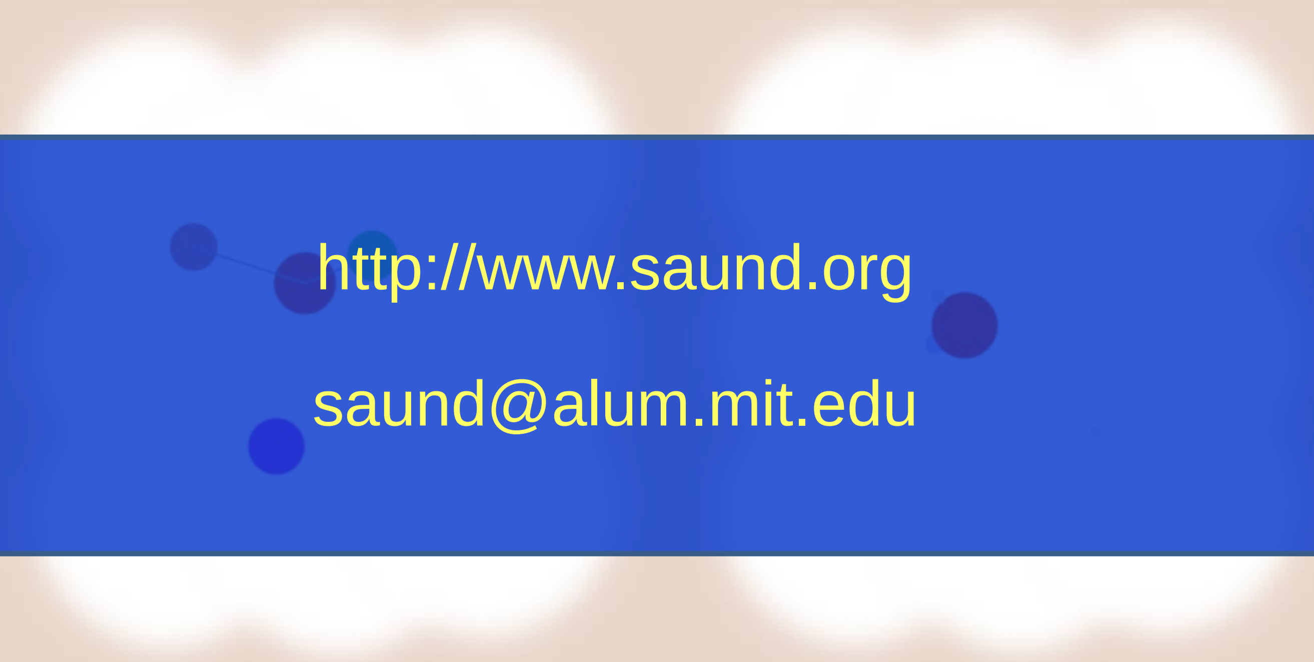
Projects

Papers

Curiosities

Links

Contact



http://www.saund.org

saund@alum.mit.edu

Conversation