TTIC 31190: Natural Language Processing

Kevin Gimpel

Spring 2018

Lecture 18: Semantics

Roadmap

- words, morphology, lexical semantics
- text classification
- simple neural methods for NLP
- language modeling and word embeddings
- recurrent/recursive/convolutional networks in NLP
- sequence labeling, HMMs, dynamic programming
- syntax and syntactic parsing
- machine translation
- semantics

Compositional Semantics

- “how should the meanings of words combine to create the meaning of something larger?”

- currently a lot of work in producing vector representations of sentences/documents

- **today**: semantic formalisms & semantic parsing

Roadmap

- semantic role labeling (SRL)
- frame-semantic parsing
- abstract meaning representation (AMR)
- combinatory categorial grammar (CCG)

Semantic Role Labeling

Who did *what* to *whom* at *where?*

$\underbrace{\text{The police officer}}_{\text{Agent}}\ \underbrace{\text{detained}}_{\text{Predicate}}\ \underbrace{\text{the suspect}}_{\text{Theme}}\ \underbrace{\text{at the scene of the crime}}_{\text{Location}}$

Can we figure out that these have the same meaning?

XYZ corporation **bought** the stock.

They **sold** the stock to XYZ corporation.

The stock was **bought** by XYZ corporation.

The **purchase** of the stock by XYZ corporation...

The stock **purchase** by XYZ corporation...

A Shallow Semantic Representation: Semantic Roles

Predicates (bought, sold, purchase) represent an **event**

semantic roles express the abstract role that arguments of a predicate can take in the event

More specific ⟷ More general

buyer **agent** **proto-agent**

Getting to semantic roles

Neo-Davidsonian event representation:

Sasha broke the window

Pat opened the door

$$\exists e,x,y\; Breaking(e) \wedge Breaker(e, Sasha) \wedge BrokenThing(e,y) \wedge Window(y)$$

$$\exists e,x,y\; Opening(e) \wedge Opener(e, Pat) \wedge OpenedThing(e,y) \wedge Door(y)$$

subjects of break and open: ***Breaker*** and ***Opener***

roles specific to each event (breaking, opening)

hard to reason about event-specific roles for downstream applications like QA

Thematic roles

- ***Breaker*** and ***Opener*** have something in common!
  - volitional actors
  - often animate
  - direct causal responsibility for their events
- thematic roles: a way to capture this semantic commonality between *Breakers* and *Openers*
  - they are both **AGENTS**
- ***BrokenThing*** and ***OpenedThing*** are **THEMES**
  - prototypically inanimate objects affected in some way by the action

A Typical Set of Thematic Roles

Thematic Role	**Definition**
AGENT	The volitional causer of an event
EXPERIENCER	The experiencer of an event
FORCE	The non-volitional causer of the event
THEME	The participant most directly affected by an event
RESULT	The end product of an event
CONTENT	The proposition or content of a propositional event
INSTRUMENT	An instrument used in an event
BENEFICIARY	The beneficiary of an event
SOURCE	The origin of the object of a transfer event
GOAL	The destination of an object of a transfer event

J&M/SLP3

A Typical Set of Thematic Roles

Thematic Role	**Example**
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The city built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	He poached catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

Problems with Thematic Roles

hard to create standard set of roles or formally define them

often roles need to be fragmented to be defined

- this quickly leads to a large number of roles!

Alternatives to thematic roles

1. **Fewer roles**: generalized semantic roles, defined as prototypes (Dowty 1991)

   PROTO-AGENT

   PROTO-PATIENT

   PropBank

2. **More roles**: Define roles specific to a group of predicates

   FrameNet

Semantic Role Labeling (SRL)

- The task of finding the semantic roles of each argument of each predicate in a sentence.
- FrameNet versus PropBank:

[You]	can't	[blame]	[the program]	[for being unable to identify it]
COGNIZER		TARGET	EVALUEE	REASON

[The San Francisco Examiner]	issued	[a special edition]	[yesterday]
ARG0	TARGET	ARG1	ARGM-TMP

History

- semantic roles as an intermediate semantics, used early in
  - machine translation (Wilks, 1973)
  - question-answering (Hendrix et al., 1973)
  - spoken-language understanding (Nash-Webber, 1975)
  - dialogue systems (Bobrow et al., 1977)
- early SRL systems

  Simmons 1973, Marcus 1980:

    - parser followed by hand-written rules for each verb
    - dictionaries with verb-specific case frames (Levin 1977)

Why Semantic Role Labeling?

- useful shallow semantic representation
- improves NLP tasks like:
  - question answering
    (Shen and Lapata, 2007; Surdeanu et al. 2011)
  - machine translation
    (Liu and Gildea, 2010; Lo et al. 2013)

PropBank

[The San Francisco Examiner]	issued	[a special edition]	[yesterday]
ARG0	TARGET	ARG1	ARGM-TMP

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106

PropBank Roles

Following Dowty 1991

Proto-Agent

- Volitional involvement in event or state
- Sentience (and/or perception)
- Causes an event or change of state in another participant
- Movement (relative to position of another participant)

Proto-Patient

- Undergoes change of state
- Causally affected by another participant
- Stationary relative to movement of another participant

PropBank Roles

Each verb sense has numbered arguments: Arg0, Arg1,…

- Arg0: PROTO-AGENT
- Arg1: PROTO-PATIENT
- Arg2: usually: benefactive, instrument, attribute, or end state
- Arg3: usually: start point, benefactive, instrument, or attribute
- Arg4 the end point
- *(Arg2-Arg5 are not really that consistent, causes a problem for labeling)*

PropBank Frame Files

agree.01

Arg0: Agreer
Arg1: Proposition
Arg2: Other entity agreeing

Ex1: [$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn't make an offer].
Ex2: [$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ on everything].

Advantage of a ProbBank Labeling

increase.01 "go up incrementally"

Arg0: causer of increase
Arg1: thing increasing
Arg2: amount increased by, EXT, or MNR
Arg3: start point
Arg4: end point

This would allow us to see the commonalities in these 3 sentences:

[$_{\text{Arg0}}$ Big Fruit Co.] increased [$_{\text{Arg1}}$ the price of bananas].
[$_{\text{Arg1}}$ The price of bananas] was increased again [$_{\text{Arg0}}$ by Big Fruit Co.]
[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

Modifiers or adjuncts of the predicate: Arg-M

ArgM-TMP	when?	yesterday evening, now
LOC	where?	at the museum, in San Francisco
DIR	where to/from?	down, to Bangkok
MNR	how?	clearly, with much enthusiasm
PRP/CAU	why?	because ... , in response to the ruling
REC		themselves, each other
ADV	miscellaneous	
PRD	secondary predication	...ate the meat raw

Capturing descriptions of the same event by different nouns/verbs

[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].
[$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%].
There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas].

Roadmap

- semantic role labeling (SRL)
- frame-semantic parsing
- abstract meaning representation (AMR)
- combinatory categorial grammar (CCG)

FrameNet

- Baker et al. 1998, Fillmore et al. 2003, Fillmore and Baker 2009, Ruppenhofer et al. 2006
- roles in PropBank are specific to a verb
- roles in FrameNet are specific to a **frame:** a background knowledge structure that defines a set of frame-specific semantic roles, called **frame elements**,
  - includes a set of predicates that use these roles
  - each word evokes a frame and profiles some aspect of the frame

"Change position on a scale" Frame

frame consists of words that indicate change of ITEM's position on a scale (the **ATTRIBUTE**) from starting point (**INITIAL VALUE**) to end point (**FINAL VALUE**)

[$_{\text{ITEM}}$ Oil] *rose* [$_{\text{ATTRIBUTE}}$ in price] [$_{\text{DIFFERENCE}}$ by 2%].

[$_{\text{ITEM}}$ It] has *increased* [$_{\text{FINAL\_STATE}}$ to having them 1 day a month].

[$_{\text{ITEM}}$ Microsoft shares] *fell* [$_{\text{FINAL\_VALUE}}$ to 7 5/8].

[$_{\text{ITEM}}$ Colon cancer incidence] *fell* [$_{\text{DIFFERENCE}}$ by 50%] [$_{\text{GROUP}}$ among men].

steady *increase* [$_{\text{INITIAL\_VALUE}}$ from 9.5] [$_{\text{FINAL\_VALUE}}$ to 14.3] [$_{\text{ITEM}}$ in dividends]

[$_{\text{DIFFERENCE}}$ 5%] [$_{\text{ITEM}}$ dividend] *increase*...

“Change position on a scale” Frame

many words can “evoke” this frame:

VERBS:	dwindle	move	soar	escalation	shift
advance	edge	mushroom	swell	explosion	tumble
climb	explode	plummet	swing	fall	
decline	fall	reach	triple	fluctuation	**ADVERBS:**
decrease	fluctuate	rise	tumble	gain	increasingly
diminish	gain	rocket		growth	
dip	grow	shift	**NOUNS:**	hike	
double	increase	skyrocket	decline	increase	
drop	jump	slide	decrease	rise	

J&M/SLP3

“Change position on a scale” Frame

Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM’s state after the change in the ATTRIBUTE’s value as an independent predication.
FINAL_VALUE	The position on the scale where the ITEM ends up.
INITIAL_STATE	A description that presents the ITEM’s state before the change in the ATTRIBUTE’s value as an independent predication.
INITIAL_VALUE	The initial position on the scale from which the ITEM moves away.
ITEM	The entity that has a position on the scale.
VALUE_RANGE	A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate.
Some Non-Core Roles	
DURATION	The length of time over which the change takes place.
SPEED	The rate of change of the VALUE.
GROUP	The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way.

J&M/SLP3

- frame-semantic parsing is generally more challenging than SRL because:
  - each frame can be evoked by many words
  - each frame has its own set of roles

Roadmap

- semantic role labeling (SRL)
- frame-semantic parsing
- abstract meaning representation (AMR)
- combinatory categorial grammar (CCG)

http://tiny.cc/amrtutorial

The Logic of AMR

Practical, Unified, Graph-Based Sentence Semantics for NLP

Nathan Schneider University of Edinburgh
Jeff Flanigan CMU
Tim O'Gorman CU-Boulder

Note: slides from this section have been removed due to large size.
Please see the original tutorial slides by Schneider/Flanigan/O'Gorman

Roadmap

- semantic role labeling (SRL)
- frame-semantic parsing
- abstract meaning representation (AMR)
- combinatory categorial grammar (CCG)

Combinatory Categorial Grammar
(Steedman, 1987)

- family of grammars that focus on **function application**
- CCGs are useful for semantic parsing and parsing to logical forms
- in one simple CCG instantiation, there are only 2 atomic types: nouns (N) and sentences (S)

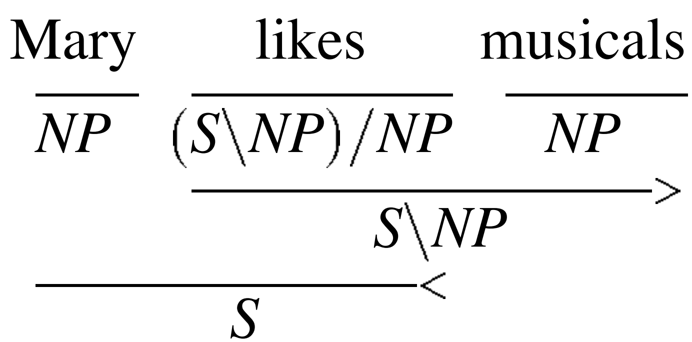
CCG

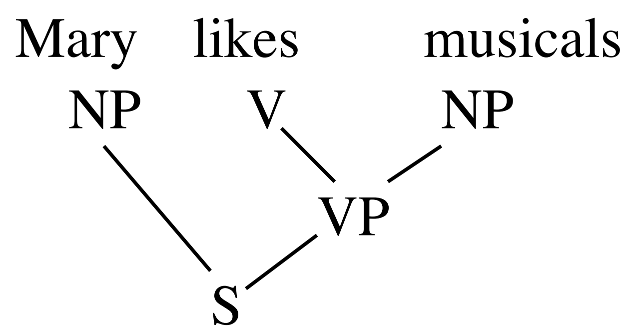
- 2 atomic types: nouns (N) and sentences (S)
- complex types created by using “slash” rules; think of these as “functions”:
  - X/Y = “something that combines with a Y **to its right** to form an X”
  - X\Y = “something that combines with a Y **to its left** to form an X”
- Consider the type S\N:
  - what are some examples of words that would have this type?
  - that is, what are some words that, when preceded by a noun, form a sentence?
  - verbs like sleeps, ate, walked

Other CCG Types

- How about (S\N)/N?
  - transitive verbs: likes, sees, ate, etc

CCG | **PCFG**

a. 

b. 

Forward Application: (>)

$$X/Y \quad Y \quad \Rightarrow \quad X$$

Backward Application: (<)

$$Y \quad X\backslash Y \quad \Rightarrow \quad X$$

Steedman (1996)

Other CCG Types

- How about N/N?
  - determiners, adjectives, nouns

Function Application as an Isomorphic Hierarchical Procedure:

likes $:= (S\backslash NP_{3s})/NP : like'$

the part after the colon (:) is the “semantic” component

Function Application as an Isomorphic Hierarchical Procedure:

We must also expand the rules of functional application in the same way:

(6) *Forward Application:* $(>)$

$$X/Y : f \quad Y : a \quad \Rightarrow \quad X : fa$$

(7) *Backward Application:* $(<)$

$$Y : a \quad X\backslash Y : f \quad \Rightarrow \quad X : fa$$

Function Application as an Isomorphic Hierarchical Procedure:

(5) likes := $(S\backslash NP_{3s})/NP : like'$

We must also expand the rules of functional application in the same way:

(6) *Forward Application:* ($>$)

$$X/Y : f \quad Y : a \quad \Rightarrow \quad X : fa$$

(7) *Backward Application:* ($<$)

$$Y : a \quad X\backslash Y : f \quad \Rightarrow \quad X : fa$$

They yield derivations like the following:

(8)

$$\begin{array}{ccc}
\text{Mary} & \text{likes} & \text{musicals} \\
\hline
NP_{3sm} : mary' & (S\backslash NP_{3s})/NP : like' & NP : musicals' \\
\cline{2-3}
 & \multicolumn{2}{c}{S\backslash NP_{3s} : like'musicals' \quad >} \\
\hline
\multicolumn{3}{c}{S : like'musicals'mary \quad <}
\end{array}$$

Other NLP Tasks and Applications

- coreference resolution
- question answering
- summarization
- dialogue systems

Other NLP Tasks and Applications

- coreference resolution
- question answering
- summarization
- dialogue systems

Coreference Resolution

- determine which pieces of text refer to the same referent:
  - President Obama selected ten delegates after receiving recommendations from his cabinet members. They spent all day Saturday working on their recommendations for him.

Other NLP Tasks and Applications

- coreference resolution
- question answering
  - factoid question answering
  - machine comprehension
- summarization
- dialogue systems

IBM’s Watson


THINK
DENKE
$24,000
Who is Stoker?
(I FOR ONE WELCOME OUR NEW COMPUTER OVERLORDS)
$1,000
$77,147
Who is Bram Stoker?
$ 17,973
$21,600
WHO IS BRAM STOKER?
$5600

IBM's Watson

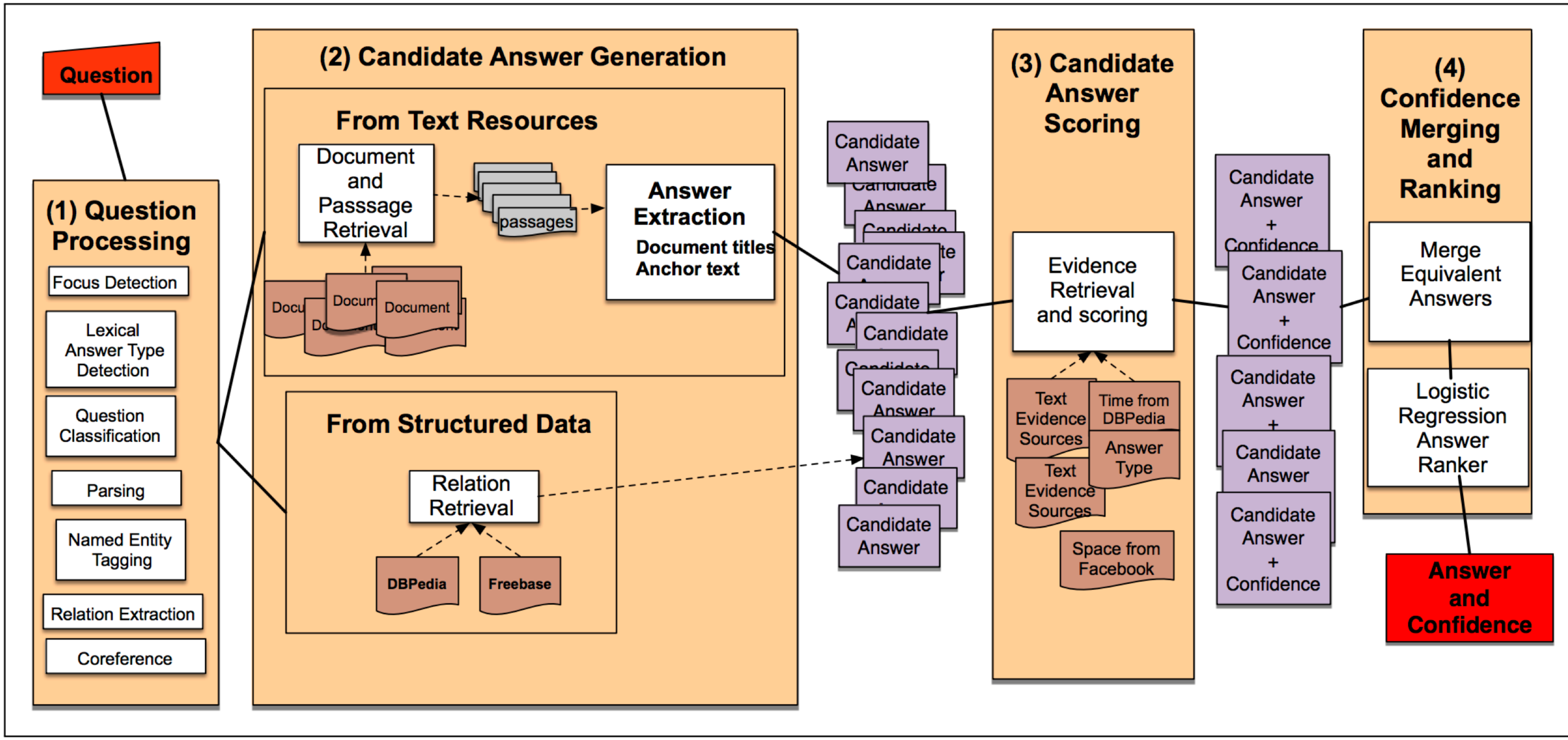


Figure 28.9 The 4 broad stages of Watson QA: (1) Question Processing, (2) Candidate Answer Generation, (3) Candidate Answer Scoring, and (4) Answer Merging and Confidence Scoring.

Classifying Questions into “Lexical Answer Types”

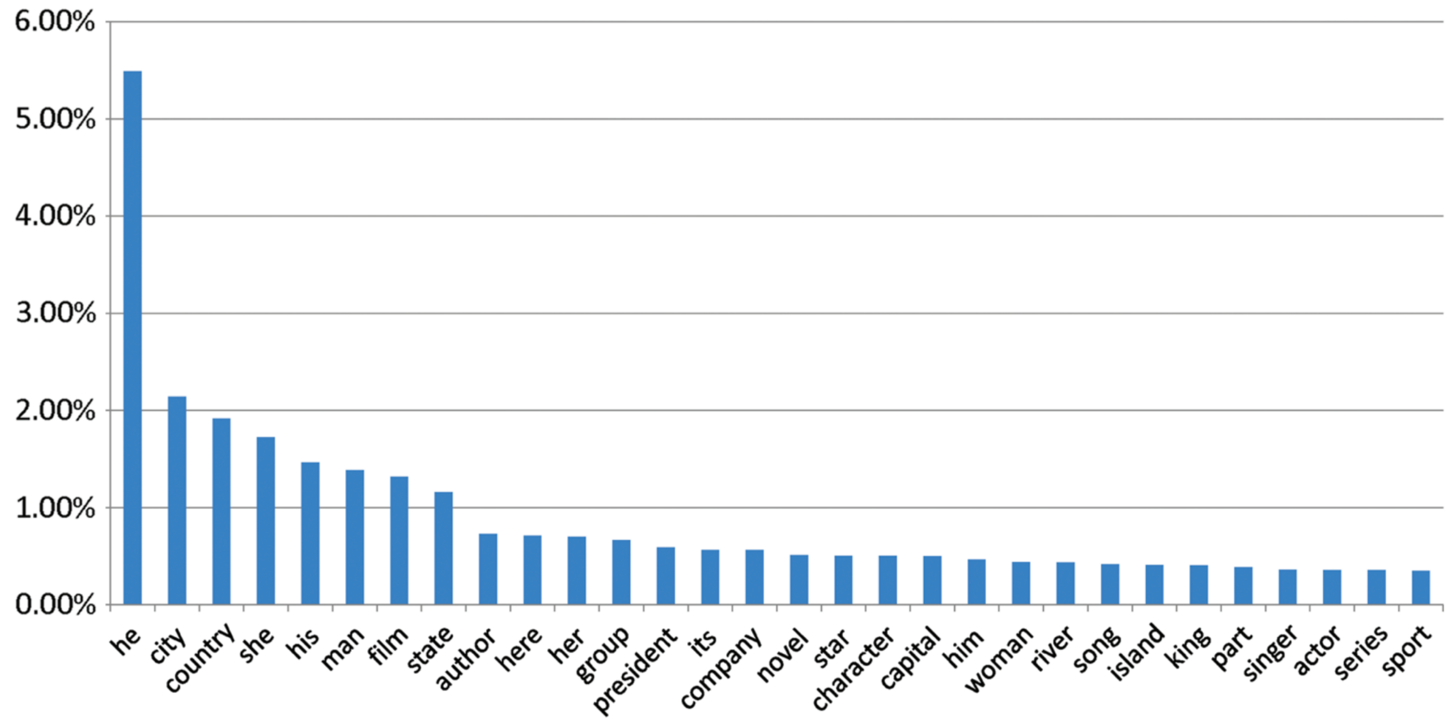


Figure 1

Distribution of the 30 most frequent lexical answer types in 20,000 Jeopardy! questions.

Machine Comprehension

Can a machine read a document and answer questions about it?

MCTest: A Challenge Dataset for the Open-Domain Machine Comprehension of Text

Matthew Richardson
Microsoft Research
One Microsoft Way
Redmond, WA 98052
mattri@microsoft.com

Christopher J.C. Burges
Microsoft Research
One Microsoft Way
Redmond, WA 98052
cburges@microsoft.com

Erin Renshaw
Microsoft Research
One Microsoft Way
Redmond, WA 98052
erinren@microsoft.com

- 660 fictional stories, written at a 4th grade reading level
- 4 multiple choice questions per story

research on the machine comprehension of text. Previous work on machine comprehension (e.g., semantic modeling) has made great strides, but primarily focuses either on limited-domain datasets, or on solving a more re-

evaluated individually, rather than by how much they advance us towards the end goal. On the other hand, the goal of semantic parsing is the machine comprehension of text (MCT), yet its evaluation requires adherence to a specific knowledge repre-

Once there was a boy named Fritz who loved to draw. He drew everything. In the morning, he drew a picture of his cereal with milk. His papa said, “Don’t draw your cereal. Eat it!”

After school, Fritz drew a picture of his bicycle. His uncle said, “Don't draw your bicycle. Ride it!”

…

What did Fritz draw first?

A) the toothpaste

B) his mama

C) cereal and milk

D) his bicycle

Once there was a boy named Fritz who loved to draw. He drew everything. In the morning, he drew a picture of his cereal with milk. His papa said, “Don’t draw your cereal. Eat it!”

After school, Fritz drew a picture of his bicycle. His uncle said, “Don't draw your bicycle. Ride it!”

…

What did Fritz draw first?

A) the toothpaste

B) his mama

C) cereal and milk

D) his bicycle

SQuAD

The Stanford Question Answering Dataset

Super Bowl 50 was an American football game to determine the champion of the National Football League (NFL) for the 2015 season. The American Football Conference (AFC) champion Denver Broncos defeated the National Football Conference (NFC) champion Carolina Panthers 24–10 to earn their third Super Bowl title. The game was played on February 7, 2016, at Levi's Stadium in the San Francisco Bay Area at Santa Clara, California. As this was the 50th Super Bowl, the league emphasized the "golden anniversary" with various gold-themed initiatives, as well as temporarily suspending the tradition of naming each Super Bowl game with Roman numerals (under which the game would have been known as "Super Bowl L"), so that the logo could prominently feature the Arabic numerals 50.

Which NFL team represented the AFC at Super Bowl 50?

Ground Truth Answers: Denver Broncos | Denver Broncos | Denver Broncos

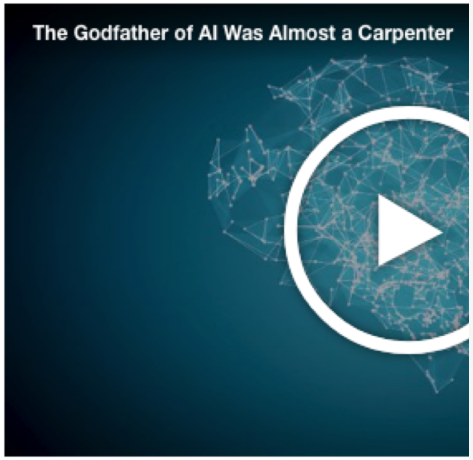
Bloomberg

WATCH LIVE
8:40 AM
Boeing Leads the Declines in the D

Alibaba's AI Outgun Reading Test

By Robert Fenner
January 14, 2018, 11:16 PM CST

- Its natural-language processing A humans
- Alibaba says it's the first time a m people



The Godfather of AI Was Almost a Carpenter

Alibaba has developed an artificial int scored better than humans in a Stanfo comprehension test.

Microso

Microsoft creates AI tha document and answer q it as well as a person

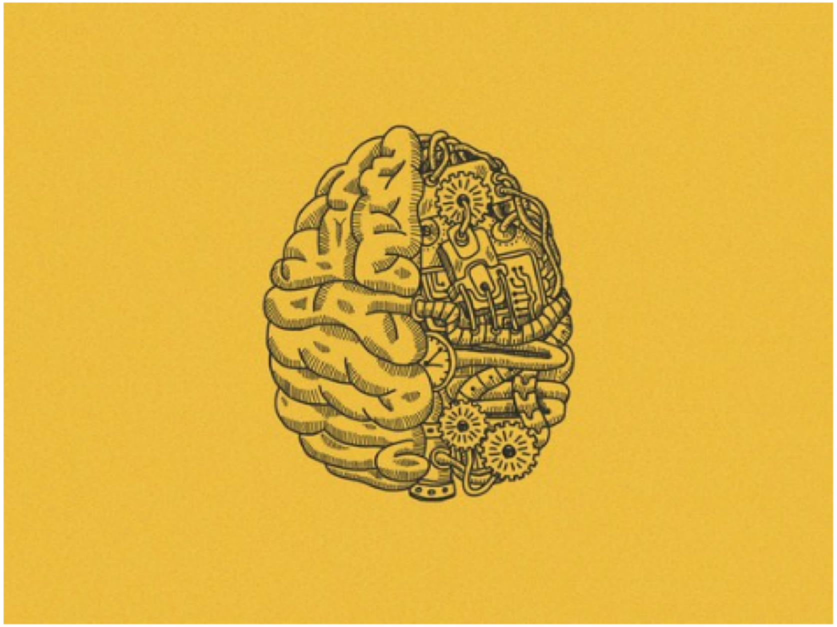
January 15, 2018 | Allison Linn

Microsoft researchers have created tec intelligence to read a document and ar about as well as a human.

WIRED SUBSCRIBE

TOM SIMONITE BUSINESS 01.18.18 03:35 PM

AI BEAT HUMANS AT READING! MAYBE NOT



GETTY IMAGES

NEWS SPREAD MONDAY of a remarkable breakthrough in artificial intelligence. Microsoft and Chinese retailer Alibaba independently announced that they had made software that matched or outperformed

Article: Super Bowl 50

Paragraph: "*Peyton Manning became the first quarterback ever to lead two different teams to multiple Super Bowls. He is also the oldest quarterback ever to play in a Super Bowl at age 39. The past record was held by John Elway, who led the Broncos to victory in Super Bowl XXXIII at age 38 and is currently Denver's Executive Vice President of Football Operations and General Manager. Quarterback Jeff Dean had jersey number 37 in Champ Bowl XXXIV.*"

Question: "*What is the name of the quarterback who was 38 in Super Bowl XXXIII?*"

Original Prediction: John Elway

Prediction under adversary: Jeff Dean

Jia and Liang (2017)

Other NLP Tasks and Applications

- coreference resolution
- question answering
- summarization
- dialogue systems

Automatic Summarization

- given a document, produce a summary of a provided length
- most systems are **extractive**: they extract content from the document
  - this is safer, since the document is presumably grammatical
  - but this limits applicability
- recent work tries to do **abstractive** summarization
  - typically based encoder-decoder models but also some based on intermediate semantic representations

Automatic Text Summarization of Newswire: Lessons Learned from the Document Understanding Conference

Ani Nenkova
Columbia University
1214 Amsterdam Ave
New York, NY 10027
ani@cs.columbia.edu

AAAI 2005

baseline = take first 100 words of document

regarding the first two years of DUC:

Both years, none of the systems outperforms the baseline (and the systems as a group do not outperform the baseline) and in fact the baseline has better coverage than most of the automatic systems (see the first row in table 1). It has often been noted that this baseline is indeed quite strong, due to journalistic convention for putting the most important part of an article in the initial paragraphs. But the fact that human summarizers (with the exception of F and J) significantly outperform the baseline shows that the task is meaningful and that better-than-baseline performance is possible. The