

# Question Answering: Research Challenges and Opportunities

Scott Wen-tau Yih  
Microsoft Research

Microsoft Research  
Faculty Summit  
**2015**



# Answer Factual Questions Using “Trusted” Sources of Documents & Databases

- Answer health and medicine related question
  - How long can a measles vaccine last?
  - What allergic reactions can be caused by the MMR vaccine?
- Answer questions using “the world’s Encyclopedia”
  - Who won the 1979 Boston Marathon?
  - Who was Katy Perry’s husband when the one that got away was released?



who was Katy Perry's husband

Web

Images

Videos

Maps

N

4,600,000 RESULTS

what's the fastest



Web

In

when

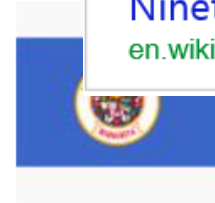
5,350,000 RESULTS

The bird  
falcon, aka

List of birds  
[en.wikipedia](http://en.wikipedia.org)

Web

4,720,



Minnesota · Founded

when was the 19th amendment ratified

Web

Images

Videos

Maps

News

Explore

2,070,000 RESULTS

Any time ▾

Susan B. Anthony and Elizabeth Cady Stanton drafted the amendment and first introduced it in 1878; it was forty-one years later, in 1919, when the Congress submitted the amendment to the states for ratification.

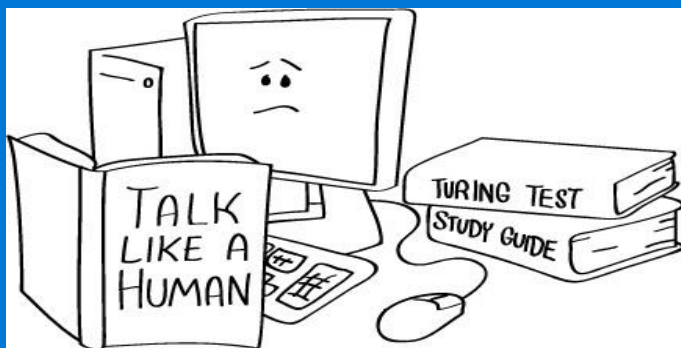
[Nineteenth Amendment to the United States ...](#)

[en.wikipedia.org/...the\\_United\\_States\\_Constitution](http://en.wikipedia.org/...the_United_States_Constitution)

Feedback

# Natural Language Understanding

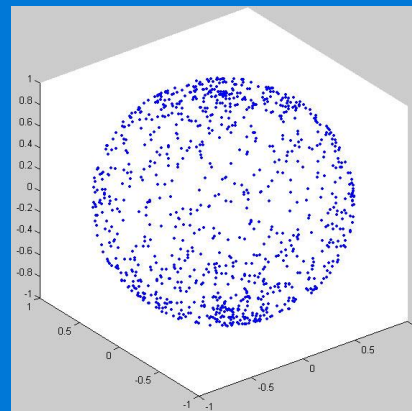
- Question answering: an intelligent system that
  - Accesses lots of structured and unstructured documents
  - Understand natural language questions from human users
  - Answer questions along with justifications



<http://csunplugged.org/turing-test>

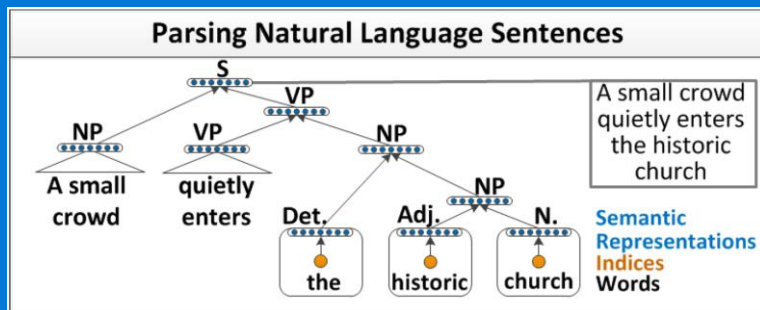
# Opportunity #1: Continuous Representations

- Word embedding models
  - Latent Semantic Analysis [Deerwester+ 90]
  - SENNA [Collobert & Weston 08]
  - Word2Vec [Mikolov+ 13]
  - GloVe [Pennington+ 14]
- Encode term co-occurrence information
- Measure semantic similarity well

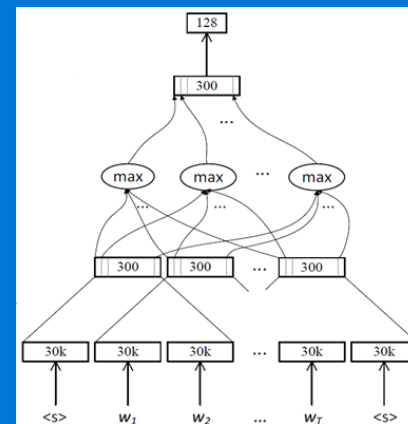


# Opportunity #1: Continuous Representations

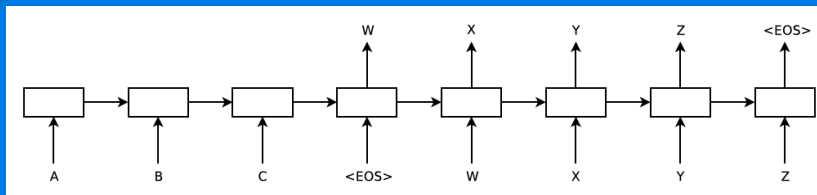
- Deep neural networks



Recursive NN [Socher+ ICML-11]



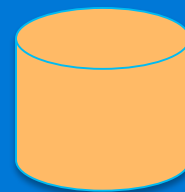
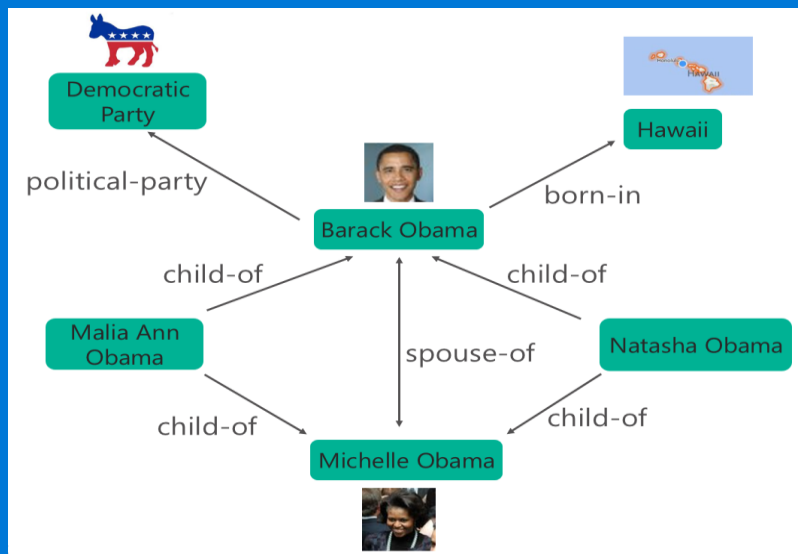
Deep Structured Semantic Model [Huang+ CIKM-13]



Sequence Learning via LSTM [Sutskever+ NIPS-14]

# Opportunity #2: Large-scale Knowledge Bases

- Encode world knowledge by storing properties of millions of entities, as well as relations among them



Satori  
Freebase  
DBpedia  
YAGO  
NELL  
OpenIE/ReVerb

# Outline

- Answer sentence selection
- Semantic parsing for QA with knowledge bases
- Other recent work from Microsoft Research
- Future research directions



# Answer Sentence Selection

- Given a factoid question, find the sentence that
  - Contains the answer
  - Can sufficiently support the answer

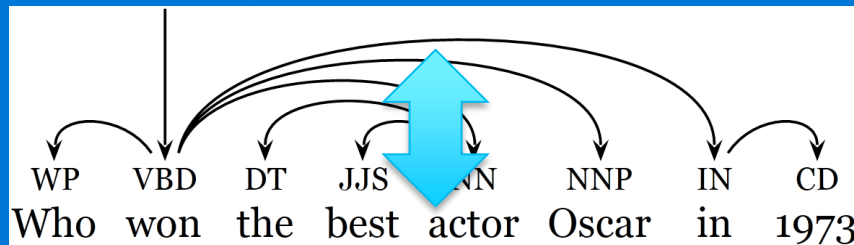
**Q:** Who won the best actor Oscar in 1973?

**S<sub>1</sub>:** Jack Lemmon was awarded the Best Actor Oscar for Save the Tiger (1973).

**S<sub>2</sub>:** Academy award winner Kevin Spacey said that Jack Lemmon is remembered as always making time for others.

# Dependency Tree Matching Approaches

Q: Who won the best actor Oscar in 1973?

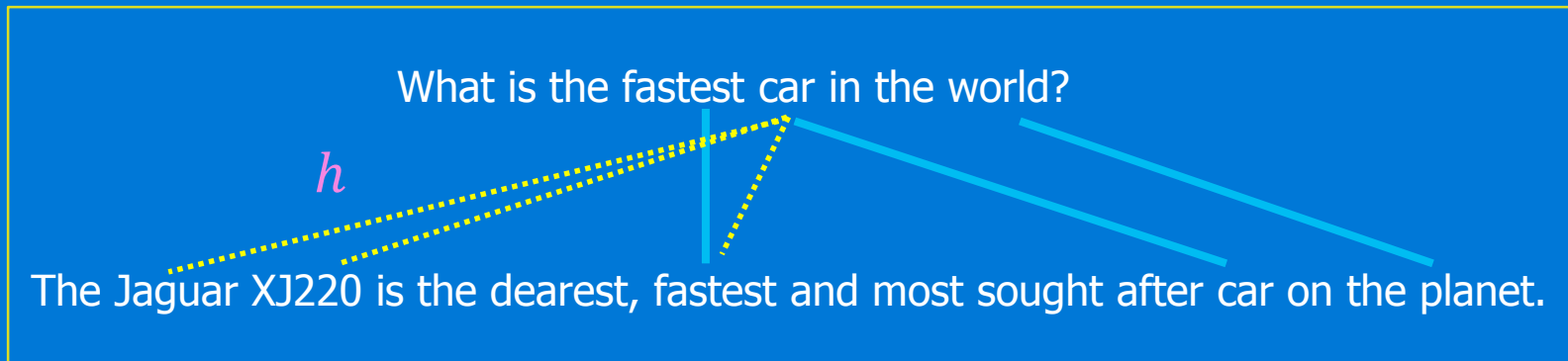


Can matching directly **Q** & **S** perform comparably?

S: Jack Lemmon was awarded the Best Actor Oscar.

# Discover Latent Word Alignment [Yih+ ACL-13]

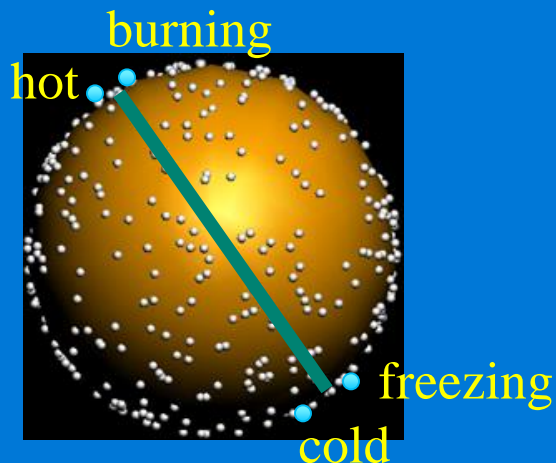
- Assume there is an underlying structure  $h$ 
  - Describes which words in *question* and *sentence* can be associated



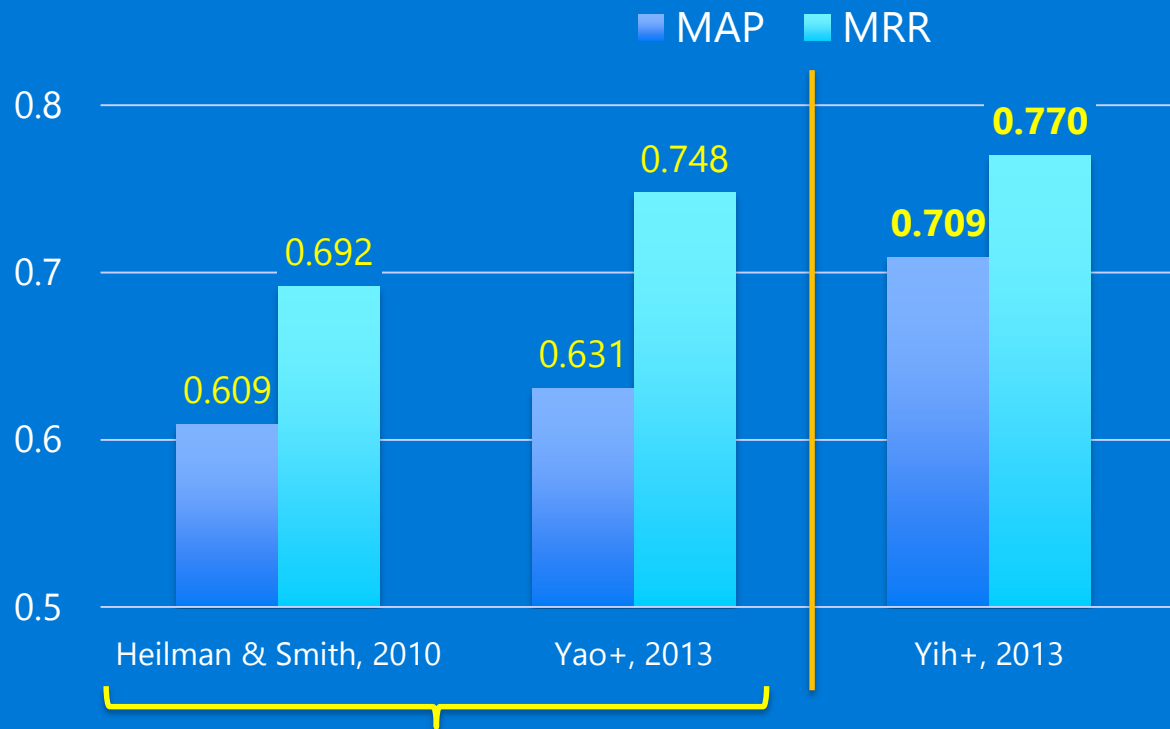
Example from [Harabagiu & Moldovan, 2001]

# Word Embedding for Lexical Semantics

- Synonyms/Antonyms
  - Polarity Inducing LSA (PILSA) [Yih+ EMNLP-CoNLL-12]
- Hypernyms (Is-A)
  - Probase [Wu+ SIGMOD-12]
- Semantic word similarity
  - Heterogeneous VSMs [Yih&Qazvinian, HLT-NAACL-12]
  - RNN language model [Mikolov+, INTERSPEECH-10]
  - Clickthrough-based latent semantic model [Gao+ SIGIR-11]

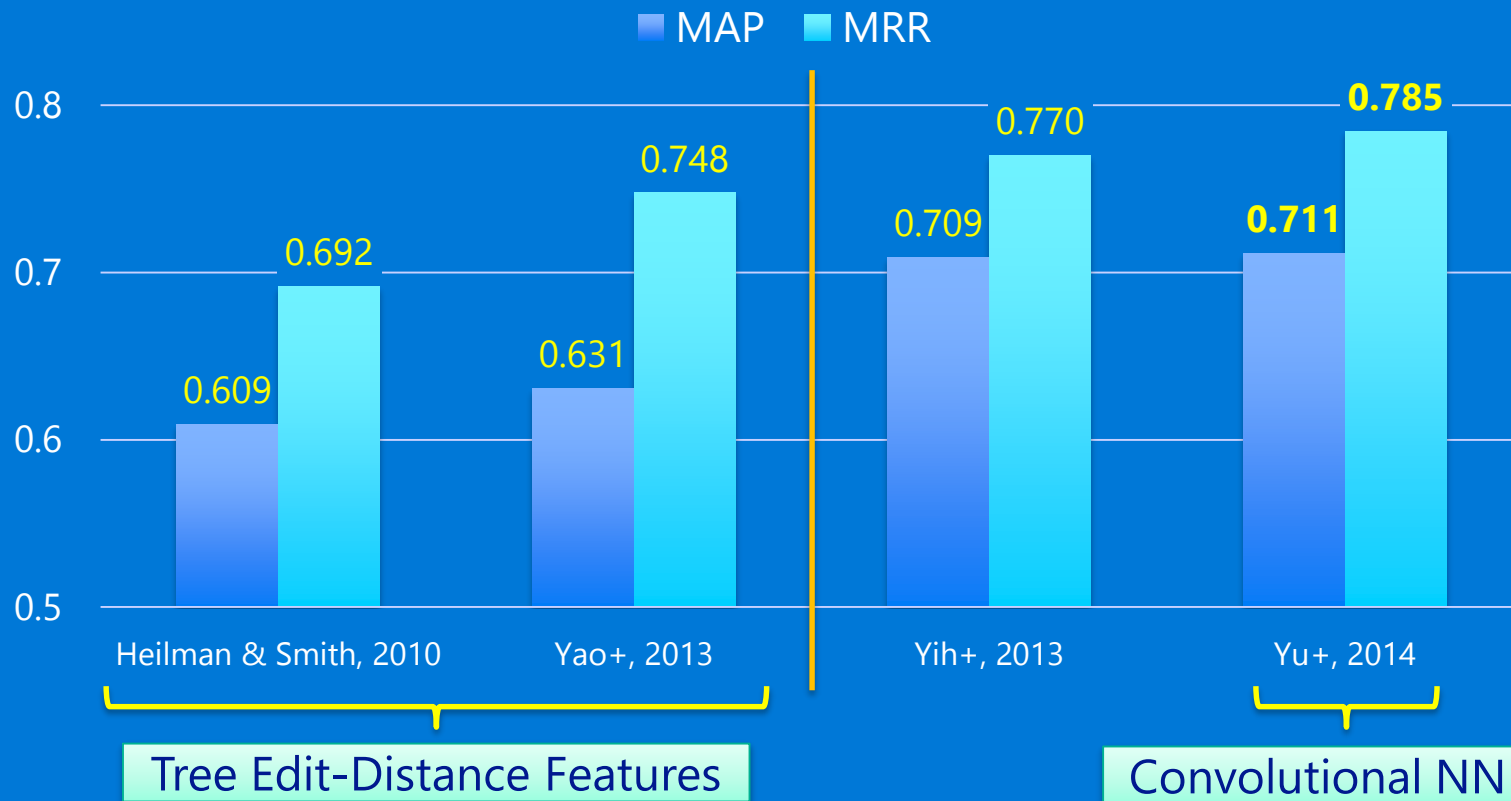


# Results on TREC-QA Dataset



Tree Edit-Distance Features

# Results on TREC-QA Dataset



# Limitation of Word Matching Models

- Sources of errors
  - Uncovered or inaccurate entity relations
  - Lack of robust question analysis
  - Need of high-level semantic representation and inference

**Q:** In what film is Gordon Gekko the main character?

**S:** He received a best actor Oscar in 1987 for this role as Gordon Gekko in “Wall Street”.

# Limitation of Word Matching Models

- Sources of errors
  - Uncovered or inaccurate entity relations
  - Lack of robust question analysis
  - Need of high-level semantic representation and inference

Q: In what film is Gordon Gekko the main character?





# Semantic Parsing for Question Answering

*In what film is Gordon Gekko the main character?*



*Wall Street  
(1987 Film)*



semantic parsing

$\lambda x. \text{character\_in}(\text{gordon\_gekko}, x)$

matching

$\lambda x. \text{character\_portrayed\_in\_films}(\text{gordon\_gekko}, x)$

query

# WebQuestions Dataset [Berant+ EMNLP-2013]

- What character did Natalie Portman play in Star Wars? ⇒ Padme Amidala
- What currency do you use in Costa Rica? ⇒ Costa Rican colon
- What did Obama study in school? ⇒ political science
- What do Michelle Obama do for a living? ⇒ writer, lawyer
- What killed Sammy Davis Jr? ⇒ throat cancer

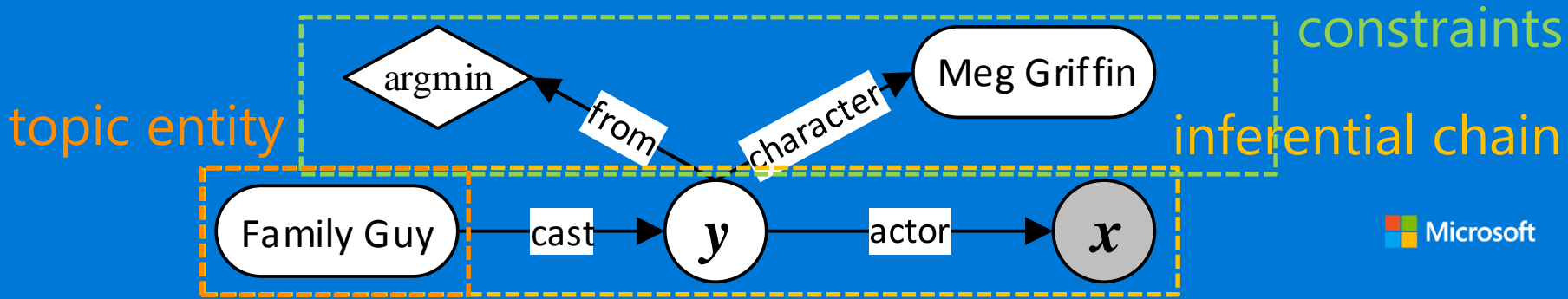
- 5,810 questions crawled from Google Suggest API
- Answers found in Freebase by Amazon MTurk
- 3,778 training, 2,032 testing
- A question may have multiple answers → using Avg. F1 (~accuracy)

# Key Challenge – Language Mismatch

- Lots of ways to ask the same question
  - *“What was the date that Minnesota became a state?”*
  - *“Minnesota became a state on?”*
  - *“When was the state Minnesota created?”*
  - *“Minnesota's date it entered the union?”*
  - *“When was Minnesota established as a state?”*
  - *“What day did Minnesota officially become a state?”*
- Need to map them to the predicate defined in KB
  - `location.dated_location.date_founded`

# Staged Query Graph Generation [Yih+ ACL-15]

- Query graph
  - Resembles subgraphs of the knowledge base
  - Can be directly mapped to a logical form in  $\lambda$ -calculus
  - Semantic parsing: a search problem that *grows* the graph through actions
- Who first voiced Meg on Family Guy?
- $\lambda x. \exists y. \text{cast}(\text{FamilyGuy}, y) \wedge \text{actor}(y, x) \wedge \text{character}(y, \text{MegGriffin})$

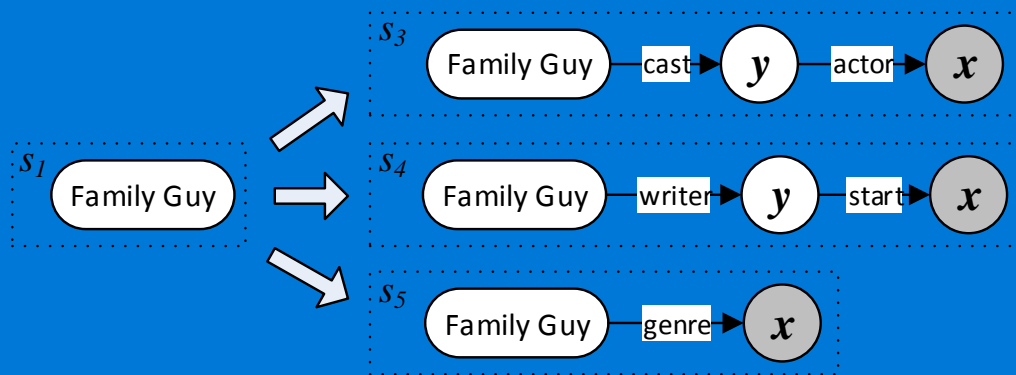
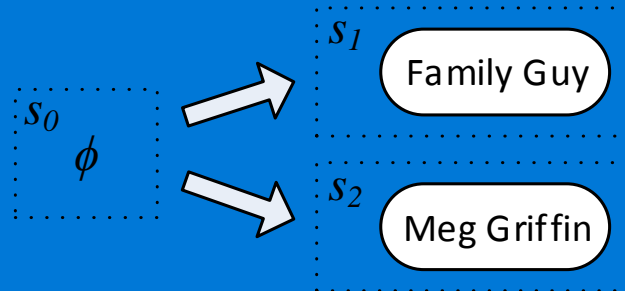


# Graph Generation Stages

- Who first voiced Meg on Family Guy?

1. Topic Entity Linking [Yang&Chang ACL-15]

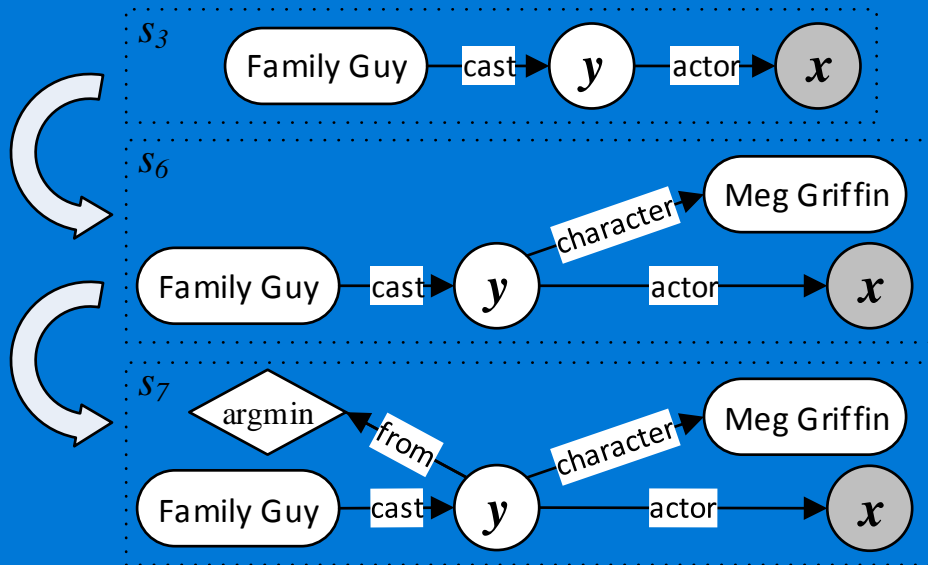
2. Identify the core inferential chain



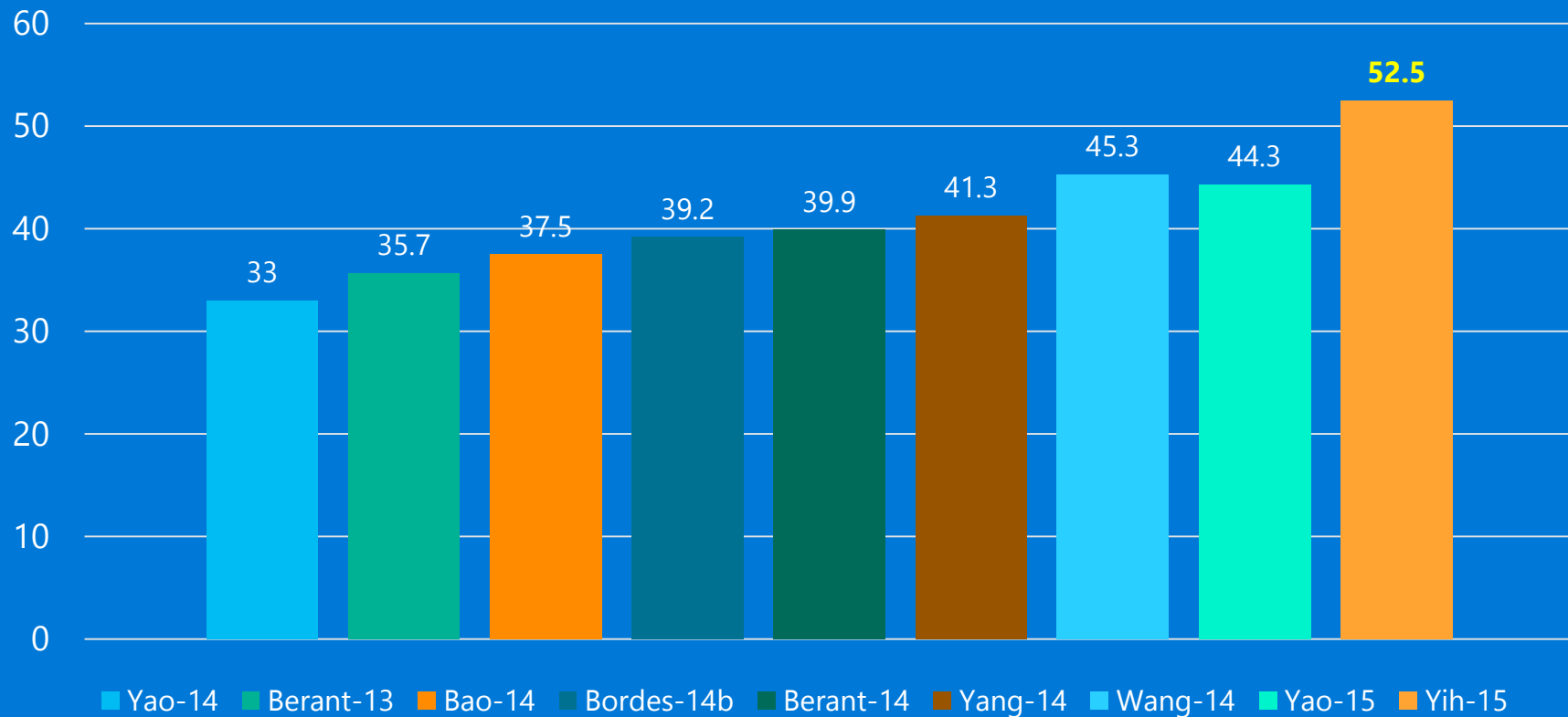
# Graph Generation Stages (cont'd)

- Who first voiced Meg on Family Guy?

## 3. Augment constraints

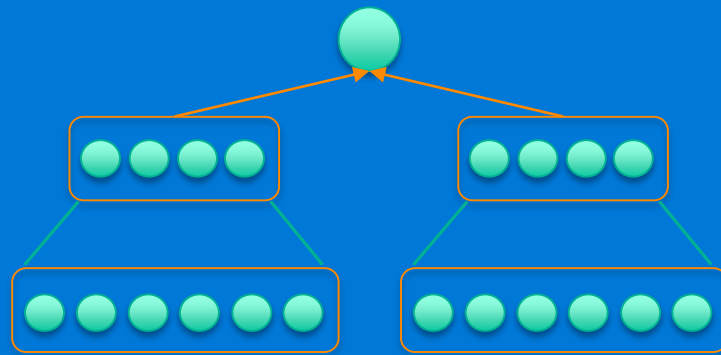
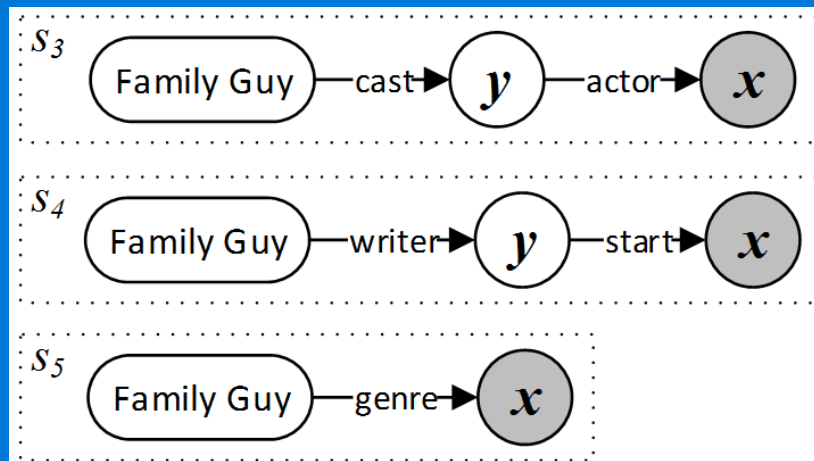


# Avg. F1 (Accuracy) on WebQuestions Test Set



# Identify Inferential Chain using Deep NN

- Who first voiced Meg on **Family Guy**?



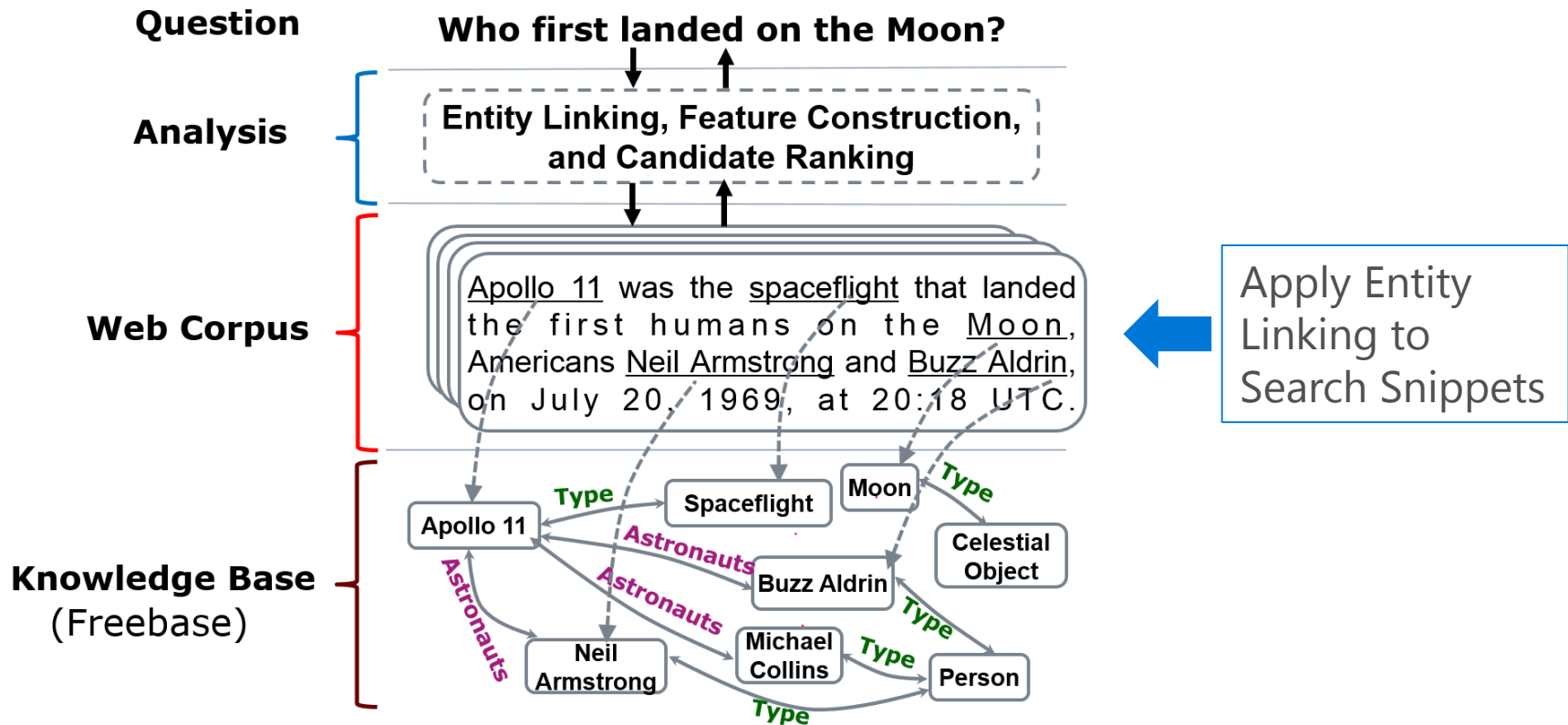
C-DSSM [Shen+14; Gao+14]

- Semantic match (“Who first voiced Meg on  $\langle e \rangle$ ”, “cast-actor”)
- Single pattern/relation matching model: 49.6% F<sub>1</sub> (vs. 52.5%)



# Question Answering via Semantic Enrichment

[Sun+ WWW-15]



# Machine Comprehension Test (MCTest)

[Richardson+ EMNLP-13]

- A reading comprehension task
- Fiction: Answers are found only in the story
- Multiple-choice: Objective/offline evaluation
- Grade-school level
  - Limited Vocabulary (8000 words)
  - Limits scope of necessary world knowledge

# MCTest – Sample Story

James the Turtle was always getting in trouble. Sometimes he'd reach into the freezer and empty out all the food. Other times he'd sled on the deck and get a splinter. His aunt Jane tried as hard as she could to keep him out of trouble, but he was sneaky and got into lots of trouble behind her back.

One day, James thought he would go into town and see what kind of trouble he could get into. He went to the grocery store and pulled all the pudding off the shelves and ate two jars. Then he walked to the fast food restaurant and ordered 15 bags of fries. He didn't pay, and instead headed home.

His aunt was waiting for him in his room. She told James that she loved him, but he would have to start acting like a well-behaved turtle.

After about a month, and after getting into lots of trouble, James finally made up his mind to be a better turtle.

- 1) What is the name of the trouble making turtle?
  - A) Fries
  - B) Pudding
  - C) James
  - D) Jane
- 2) What did James pull off of the shelves in the grocery store?
  - A) pudding
  - B) fries
  - C) food
  - D) splinters
- 3) Where did James go after he went to the grocery store?
  - A) his deck
  - B) his freezer
  - C) a fast food restaurant
  - D) his room
- 4) What did James do after he ordered the fries?
  - A) went to the grocery store
  - B) went home without paying
  - C) ate them
  - D) made up his mind to be a better turtle

# Moving Forward – Conversational QA

- Who founded AFL?
- Samuel Gompers, if by AFL you mean American Federation of Labor.
- No, I meant football.
- Lamar Hunt and Bud Adams are the co-founders of the American Football League. AFL could also mean the Arena Football League.

Know an entity mention is ambiguous

# Moving Forward – Conversational QA

- Arena football? That's interesting. How is it different from regular American football?
- It is played indoors on a smaller field than American football, resulting in a faster-paced and higher-scoring game. [Wikipedia]
- How many teams are there in this Arena Football League?
- 12. [Database]

Integrate information from different sources;  
Provide interpretable explanations to users

# Moving Forward – Conversational QA

- Any of them in the Pacific Northwest?
- Please define "Pacific Northwest".
- Washington and Oregon.
- Yes, Portland Thunder and Spokane Shock.

Know when it  
does not know

Learn through  
interactions

# Conclusions

- Question Answering – an exciting & active field
  - Research stimulated by the real user need
  - Application of natural language understanding
- Significant progress
  - MSR: Deep QA, MSRA, AskMSR [Brill+ 02; Agichtein&Cucerzan 05]
  - Rapid development at Bing: entity/paragraph/table answers
- Research challenges
  - Knowledge representation
  - Composition in languages

