Question Answering

What is Question Answering?

# Question Answering

One of the oldest NLP tasks (punched card systems in 1961)
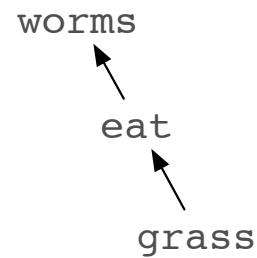
Question:
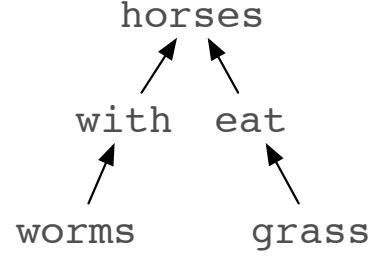
Potential Answers:
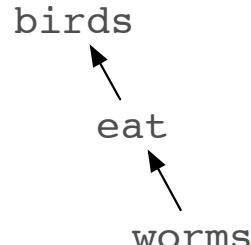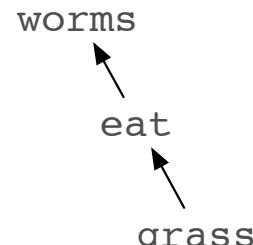
```
What do worms eat?       Worms eat grass      Horses with worms eat grass

        worms                  worms                      horses

      eat                        eat                   with   eat

    what                           grass            worms       grass


                          Birds eat worms    Grass is eaten by worms
                              birds                 worms

                            eat                    eat

                              worms                   grass
```
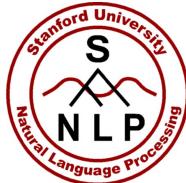
2

Dan Jurafsky

Question Answering: IBM's Watson

- Won Jeopardy on February 16, 2011!

WILLIAM WILKINSON'S
"AN ACCOUNT OF THE PRINCIPALITIES OF
WALLACHIA AND MOLDOVIA"
INSPIRED THIS AUTHOR'S
MOST FAMOUS NOVEL

→ Bram Stoker

Apple's Siri

WolframAlpha™ computational... knowledge engine

how many calories are in two slices of banana cream pie?

☰ Examples   ⤨ Random

Assuming any type of pie, banana cream | Use pie, banana cream, prepared from recipe or pie, banana cream, no-bake type, prepared from mix instead

Input interpretation:

pie	amount	2 slices	total calories
	type	banana cream	

Average result: Show details

702 Cal (dietary Calories)

Types of Questions in Modern Systems

- Factoid questions
  - *Who wrote "The Universal Declaration of Human Rights"?*
  - *How many calories are there in two slices of apple pie?*
  - *What is the average age of the onset of autism?*
  - *Where is Apple Computer based?*
- Complex (narrative) questions:
  - *In children with an acute febrile illness, what is the efficacy of acetaminophen in reducing fever?*
  - *What do scholars think about Jefferson's position on dealing with pirates?*

Dan Jurafsky

Commercial systems: mainly factoid questions

Where is the Louvre Museum located?	In Paris, France
What's the abbreviation for limited partnership?	L.P.
What are the names of Odin's ravens?	Huginn and Muninn
What currency is used in China?	The yuan
What kind of nuts are used in marzipan?	almonds
What instrument does Max Roach play?	drums
What is the telephone number for Stanford University?	650-723-2300

Paradigms for QA

- ## IR-based approaches
  - TREC; IBM Watson; Google

- ## Knowledge-based and Hybrid approaches
  - IBM Watson; Apple Siri; Wolfram Alpha; True Knowledge Evi

Many questions can already be answered by web search

Google | What are the names of Odin's ravens?

Search   About 214,000 results (0.38 seconds)

Everything

Huginn and Muninn - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Huginn_and_Muninn
The **names** of the **ravens** are sometimes modernly anglicized as Hugin and Munin. In the Poetic Edda, a disguised **Odin** expresses that he fears that they may ...
Attestations - Archaeological record - Theories - See also

Images

Maps

9

IR-based Question Answering

Google

Where is the Louvre Museum located?

Search   About 904,000 results (0.30 seconds)

Everything

Best guess for Louvre Museum Location is **Paris, France**

Images

Mentioned on at least 7 websites including wikipedia.org, answers.com and east-buc.k12.ia.us - Show sources - Feedback

Maps

Musée du Louvre - Wikipedia, the free encyclopedia
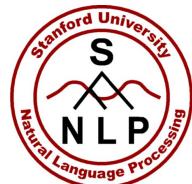
en.wikipedia.org/wiki/Musée_du_**Louvre**

Videos
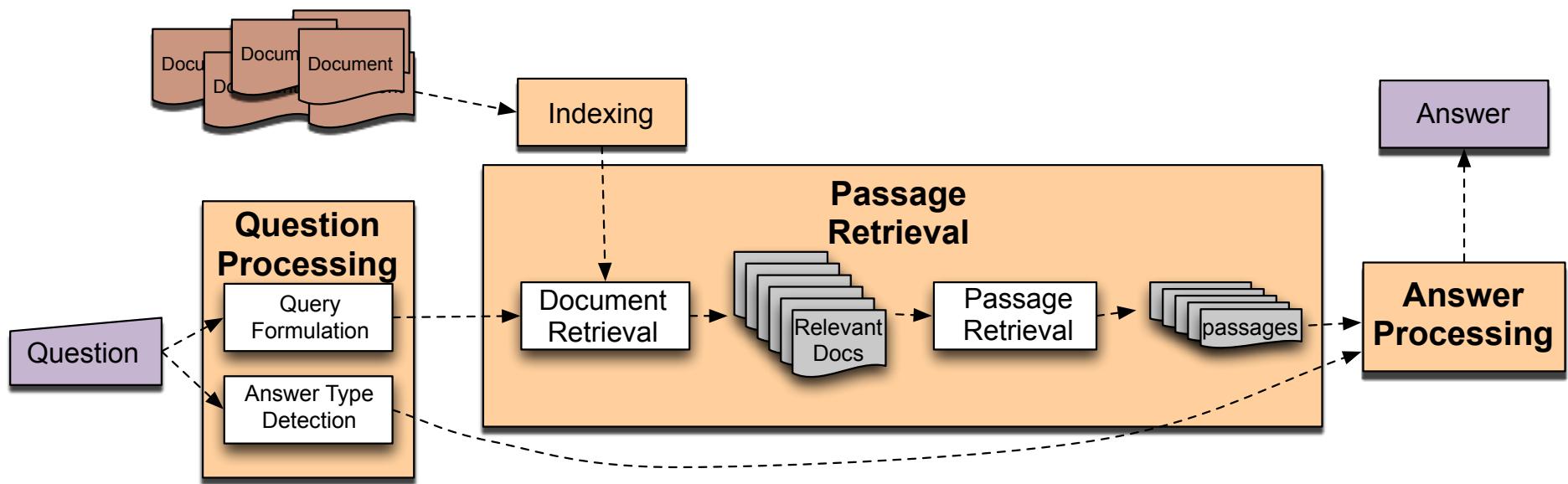
Musée du **Louvre** is **located** in Paris. **Location** within Paris. Established, 1793. **Location**,

News

Palais Royal, Musée du **Louvre**, **75001 Paris, France**. Type, Art **museum** ...

Louvre Palace - List of works in the Louvre - Category:Musée du Louvre

IR-based Factoid QA

IR-based Factoid QA

- QUESTION PROCESSING
  - Detect question type, answer type, focus, relations
  - Formulate queries to send to a search engine

- PASSAGE RETRIEVAL
  - Retrieve ranked documents
  - Break into suitable passages and rerank

- ANSWER PROCESSING
  - Extract candidate answers
  - Rank candidates
    - using evidence from the text and external sources
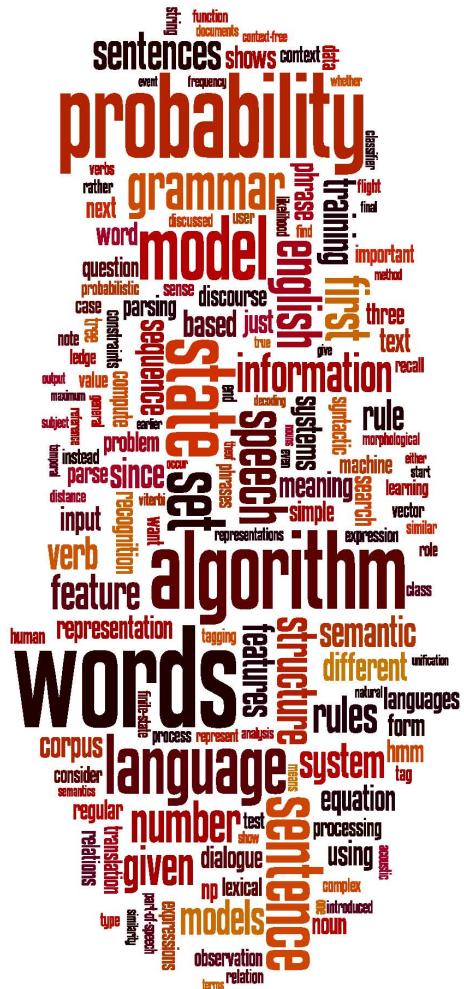
Knowledge-based approaches (Siri)

- Build a semantic representation of the query
  - Times, dates, locations, entities, numeric quantities
- Map from this semantics to query structured data or resources
  - Geospatial databases
  - Ontologies (Wikipedia infoboxes, dbPedia, WordNet, Yago)
  - Restaurant review sources and reservation services
  - Scientific databases

13
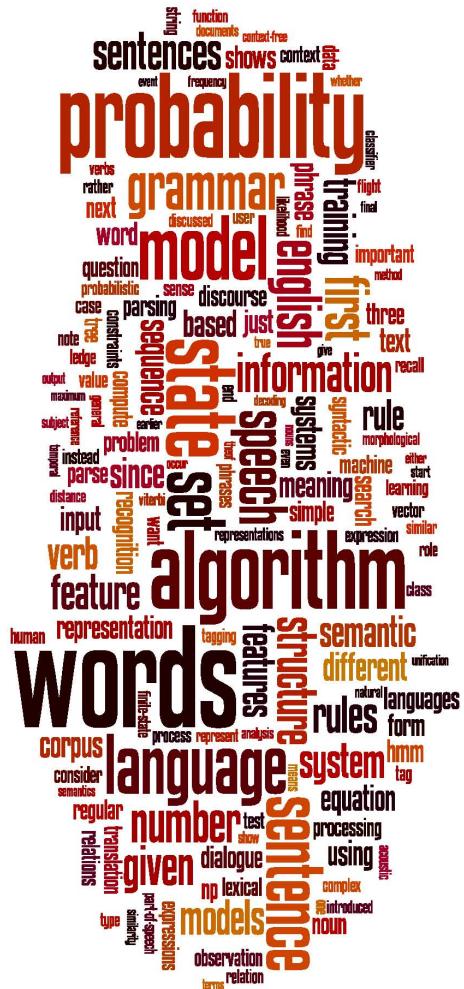
Hybrid approaches (IBM Watson)

- Build a shallow semantic representation of the query

- Generate answer candidates using IR methods

  - Augmented with ontologies and semi-structured data

- Score each candidate using richer knowledge sources

  - Geospatial databases

  - Temporal reasoning

  - Taxonomical classification

Question Answering

What is Question Answering?
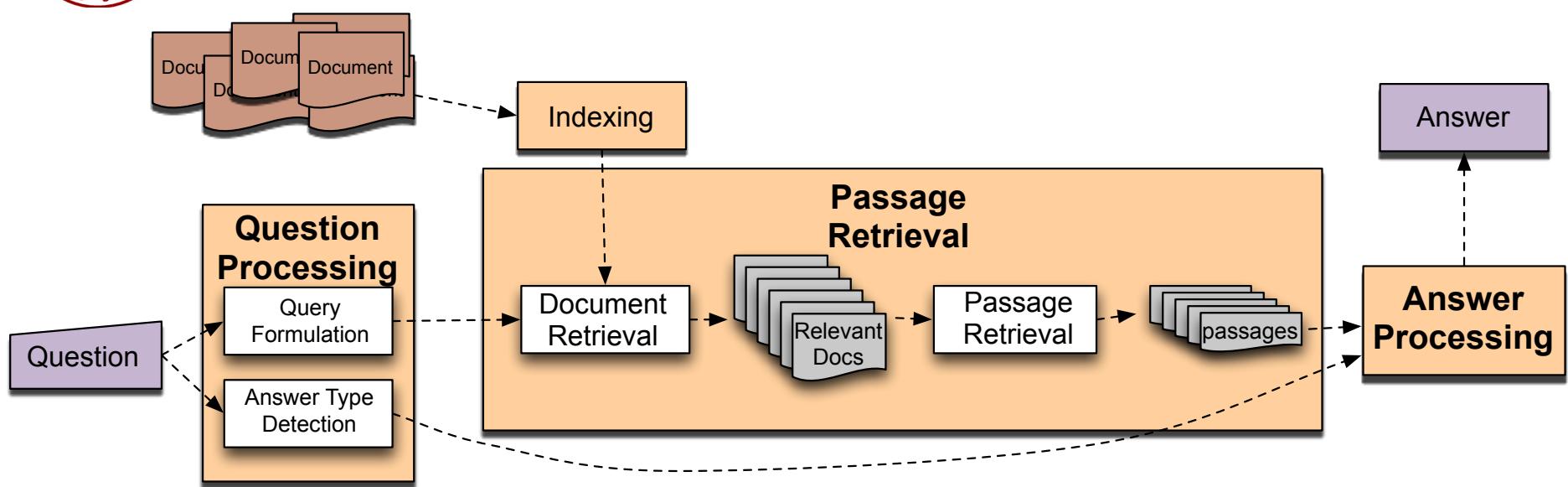
Question Answering

Answer Types and Query Formulation

Factoid Q/A

Document Document Document

Document Document

Indexing

Question Processing

Query Formulation

Answer Type Detection

Question

Passage Retrieval

Document Retrieval

Relevant Docs

Passage Retrieval

passages

Answer Processing

Answer

Question Processing
Things to extract from the question

- Answer Type Detection
  - Decide the **named entity type** (person, place) of the answer
- Query Formulation
  - Choose **query keywords** for the IR system
- Question Type classification
  - Is this a definition question, a math question, a list question?
- Focus Detection
  - Find the question words that are replaced by the answer
- Relation Extraction
  - Find relations between entities in the question

18

Question Processing

They're the two states you could be reentering if you're crossing Florida's northern border

- Answer Type: US state

- Query: two states, border, Florida, north

- Focus: the two states

- Relations: borders(Florida, ?x, north)

19

Answer Type Detection: Named Entities

- *Who founded Virgin Airlines?*
  - PERSON
- *What Canadian city has the largest population?*
 - *CITY.*
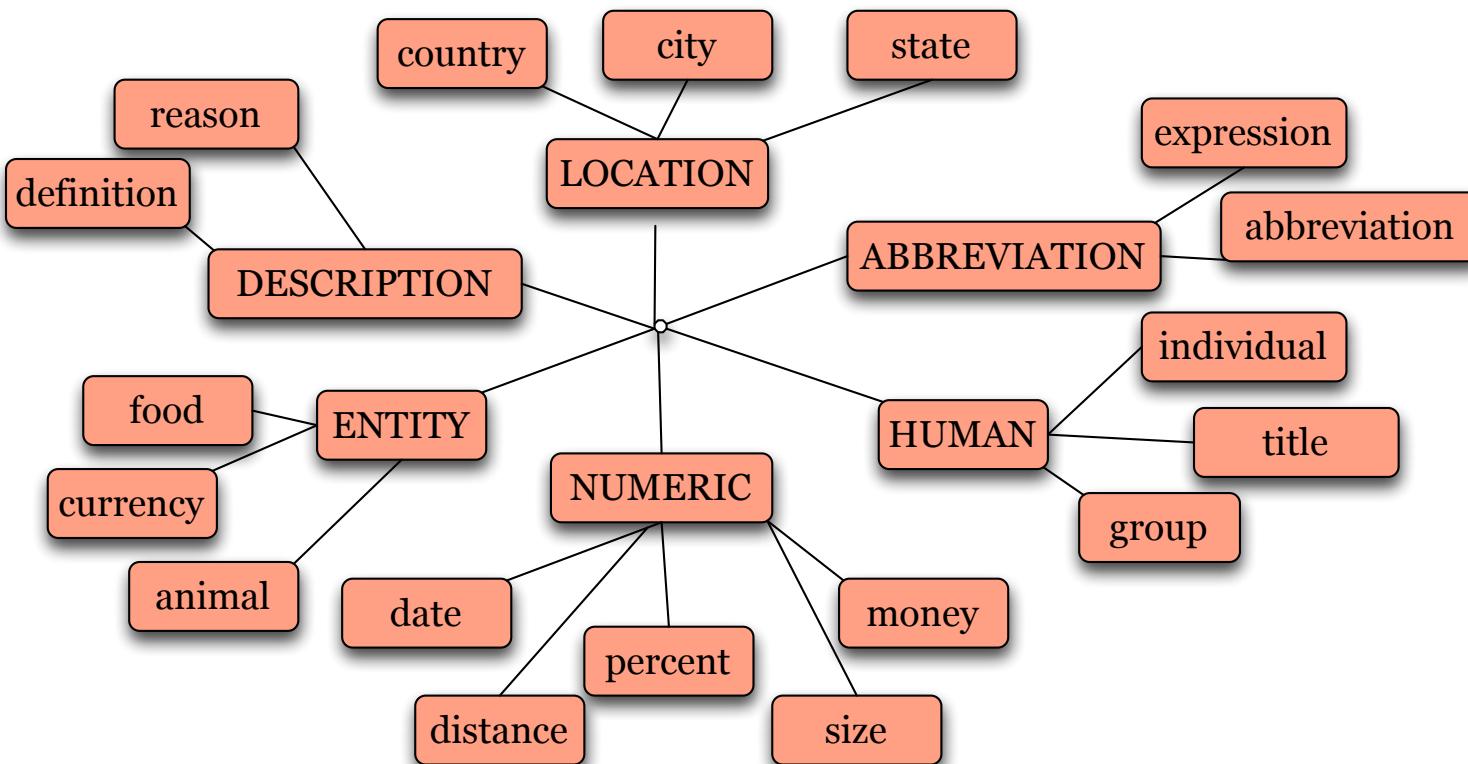
Answer Type Taxonomy

Xin Li, Dan Roth. 2002. Learning Question Classifiers. COLING'02

- 6 coarse classes
  - ABBEVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION, NUMERIC
- 50 finer classes
  - LOCATION: city, country, mountain…
  - HUMAN: group, individual, title, description
  - ENTITY: animal, body, color, currency…

21

Part of Li & Roth's Answer Type Taxonomy
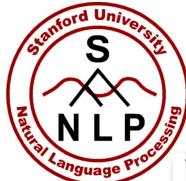
Answer Types

ENTITY	
animal	What are the names of Odin's ravens?
body	What part of your body contains the corpus callosum?
color	What colors make up a rainbow ?
creative	In what book can I find the story of Aladdin?
currency	What currency is used in China?
disease/medicine	What does Salk vaccine prevent?
event	What war involved the battle of Chapultepec?
food	What kind of nuts are used in marzipan?
instrument	What instrument does Max Roach play?
lang	What's the official language of Algeria?
letter	What letter appears on the cold-water tap in Spain?
other	What is the name of King Arthur's sword?
plant	What are some fragrant white climbing roses?
product	What is the fastest computer?
religion	What religion has the most members?
sport	What was the name of the ball game played by the Mayans?
substance	What fuel do airplanes use?
symbol	What is the chemical symbol for nitrogen?
technique	What is the best way to remove wallpaper?
term	How do you say " Grandma " in Irish?
vehicle	What was the name of Captain Bligh's ship?
word	What's the singular of dice?

23

More Answer Types

HUMAN	
description	Who was Confucius?
group	What are the major companies that are part of Dow Jones?
ind	Who was the first Russian astronaut to do a spacewalk?
title	What was Queen Victoria's title regarding India?
LOCATION	
city	What's the oldest capital city in the Americas?
country	What country borders the most others?
mountain	What is the highest peak in Africa?
other	What river runs through Liverpool?
state	What states do not have state income tax?
NUMERIC	
code	What is the telephone number for the University of Colorado?
count	About how many soldiers died in World War II?
date	What is the date of Boxing Day?
distance	How long was Mao's 1930s Long March?
money	How much did a McDonald's hamburger cost in 1963?
order	Where does Shanghai rank among world cities in population?
other	What is the population of Mexico?
period	What was the average life expectancy during the Stone Age?
percent	What fraction of a beaver's life is spent swimming?
speed	What is the speed of the Mississippi River?
temp	How fast must a spacecraft travel to escape Earth's gravity?
size	What is the size of Argentina?
weight	How many pounds are there in a stone?

Dan Jurafsky

Answer types in Jeopardy

- 2500 answer types in 20,000 Jeopardy question sample
- The most frequent 200 answer types cover < 50% of data
- The 40 most frequent Jeopardy answer types

he, country, city, man, film, state, she, author, group, here, company,
president, capital, star, novel, character, woman, river, island, king,
song, part, series, sport, singer, actor, play, team, show,
actress, animal, presidential, composer, musical, nation,
book, title, leader, game

Dan Jurafsky

Answer Type Detection

- Hand-written rules

- Machine Learning

- Hybrids

Answer Type Detection

- Regular expression-based rules can get some cases:
  - Who {is|was|are|were} PERSON
  - PERSON (YEAR – YEAR)

- Other rules use the **question headword**:

  (the headword of the first noun phrase after the wh-word)

  - Which **city** in China has the largest number of foreign financial companies?
  - What is the state **flower** of California?
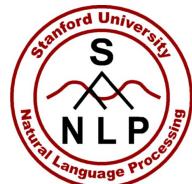
Answer Type Detection

- Most often, we treat the problem as machine learning classification

  - **Define** a taxonomy of question types

  - **Annotate** training data for each question type

  - **Train** classifiers for each question class
    using a rich set of features.

    - features include those hand-written rules!

28
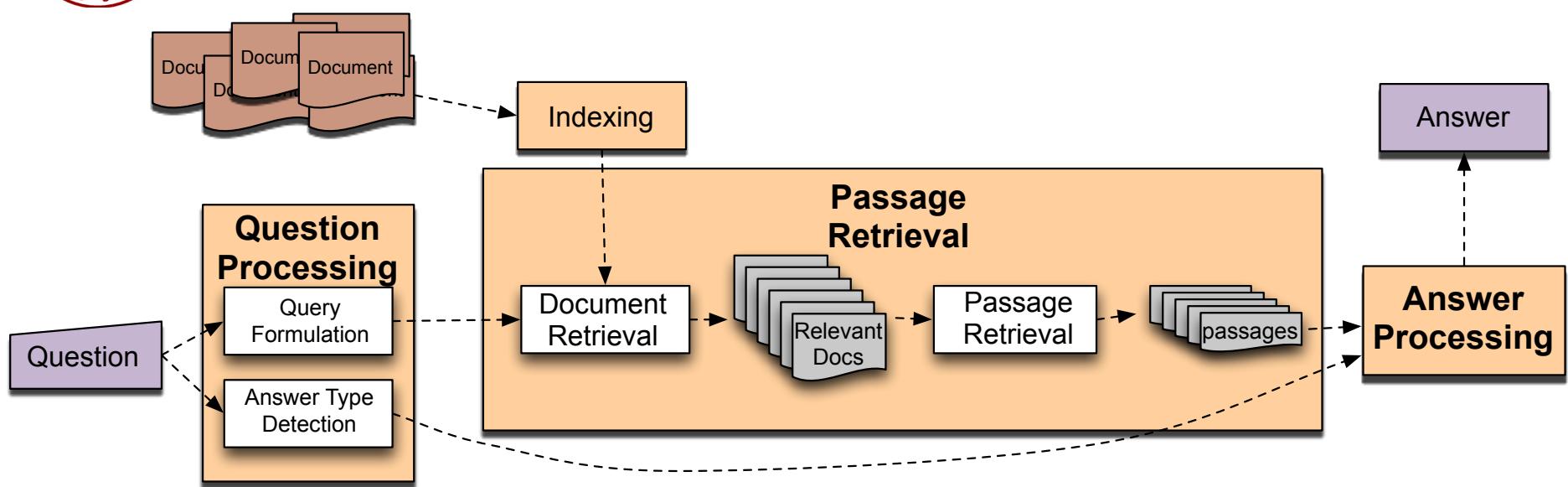
Features for Answer Type Detection

- Question words and phrases

- Part-of-speech tags

- Parse features (headwords)

- Named Entities

- Semantically related words

Factoid Q/A

Keyword Selection Algorithm

Dan Moldovan, Sanda Harabagiu, Marius Paca, Rada Mihalcea, Richard Goodrum, Roxana Girju and Vasile Rus. 1999. Proceedings of TREC-8.
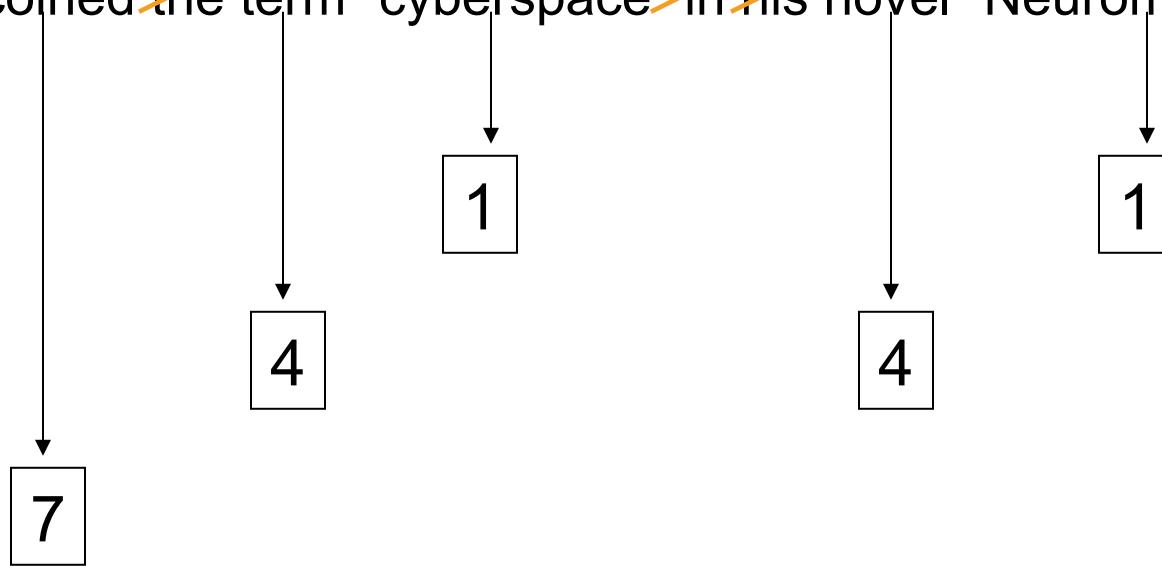
1. Select all non-stop words in quotations
2. Select all NNP words in recognized named entities
3. Select all complex nominals with their adjectival modifiers
4. Select all other complex nominals
5. Select all nouns with their adjectival modifiers
6. Select all other nouns
7. Select all verbs
8. Select all adverbs
9. Select the QFW word (skipped in all previous steps)
10. Select all other words
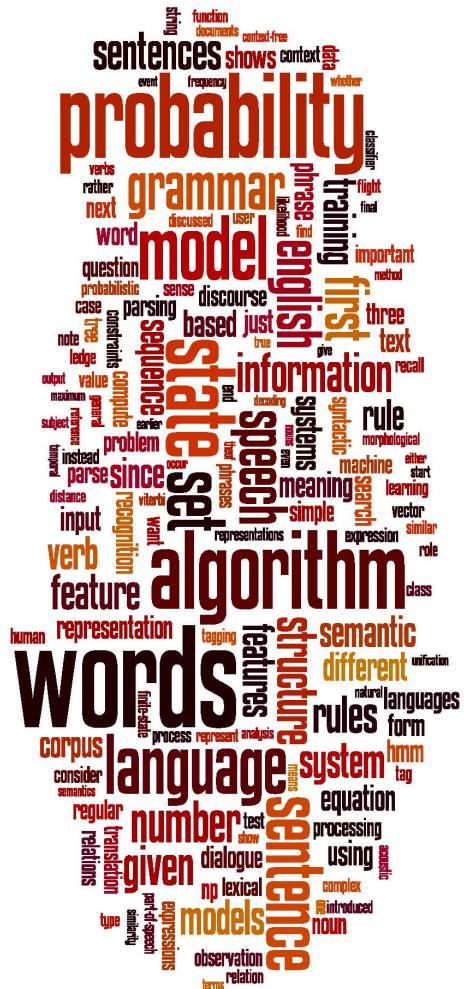
Choosing keywords from the query

Slide from Mihai Surdeanu

Who coined the term "cyberspace" in his novel "Neuromancer"?

1

1

4

4

7
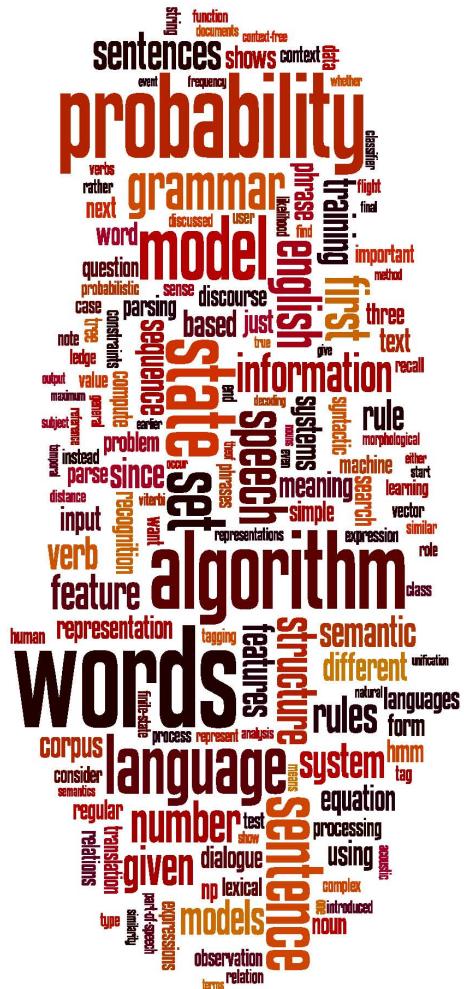
cyberspace/1 Neuromancer/1 term/4 novel/4 coined/7

32
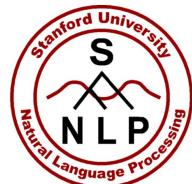
Question Answering

Answer Types and Query Formulation
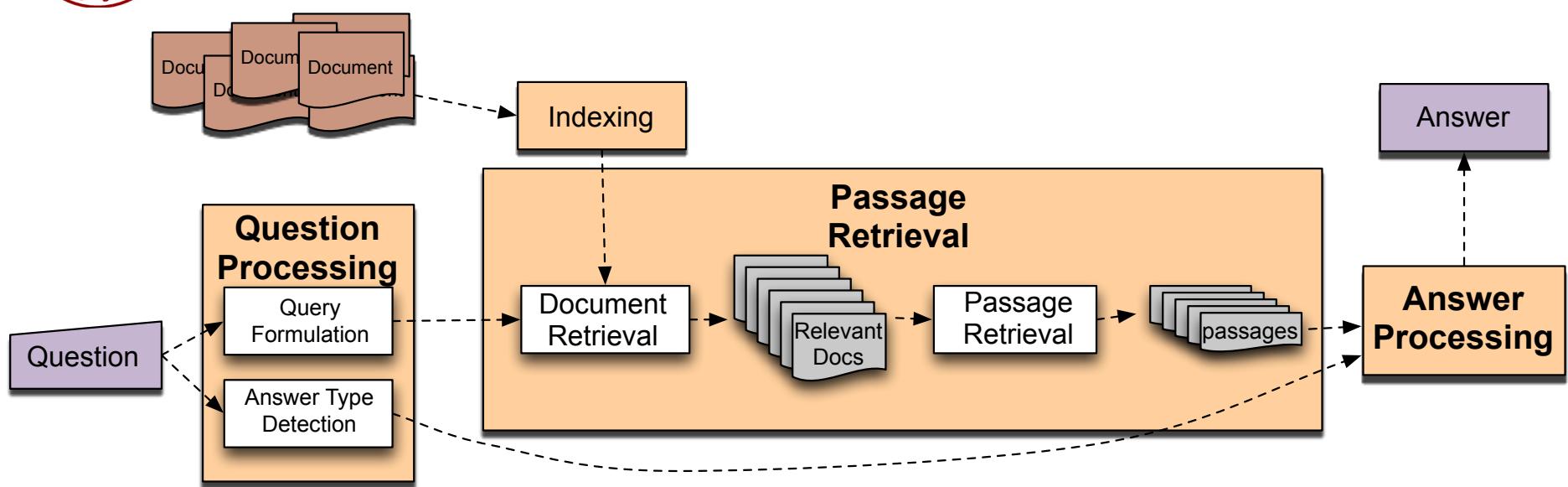
Question Answering

Passage Retrieval and Answer Extraction

Factoid Q/A

Passage Retrieval
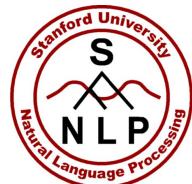
- Step 1: IR engine retrieves documents using query terms

- Step 2: Segment the documents into shorter units

  - something like paragraphs

- Step 3: Passage ranking

 - Use answer type to help rerank passages
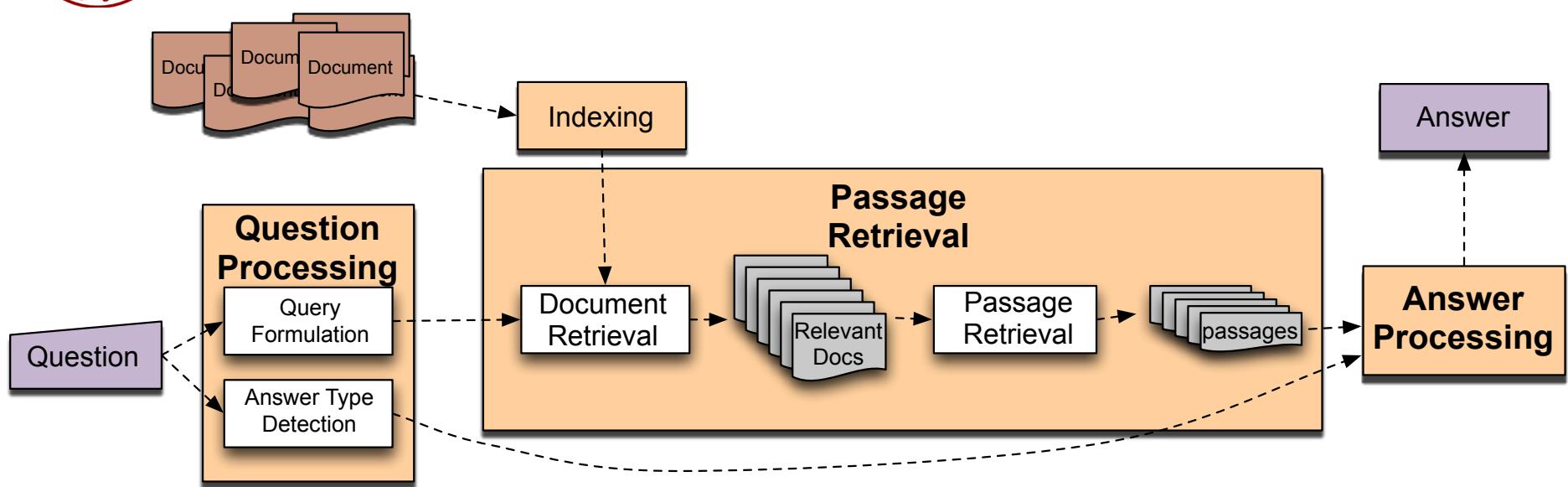
Features for Passage Ranking

Either in rule-based classifiers or with supervised machine learning

- Number of Named Entities of the right type in passage
- Number of query words in passage
- Number of question N-grams also in passage
- Proximity of query keywords to each other in passage
- Longest sequence of question words
- Rank of the document containing passage

Factoid Q/A

Question Processing
- Query Formulation
- Answer Type Detection

Question

Indexing

Passage Retrieval

Document Retrieval → Relevant Docs → Passage Retrieval → passages

Answer Processing

Answer

38

Dan Jurafsky

Answer Extraction

- Run an answer-type named-entity tagger on the passages
  - Each answer type requires a named-entity tagger that detects it
  - If answer type is CITY, tagger has to tag CITY
    - Can be full NER, simple regular expressions, or hybrid
- Return the string with the right type:
  - `Who is the prime minister of India` (PERSON)
    `Manmohan Singh``, Prime Minister of India, had told`
    `left leaders that the deal would not be renegotiated.`
  - `How tall is Mt. Everest?` (LENGTH)
    `The official height of Mount Everest is` **`29035 feet`**

Ranking Candidate Answers

- But what if there are multiple candidate answers!

Q: Who was Queen Victoria's second son?

- Answer Type: **Person**

- Passage:

The Marie biscuit is named after Marie Alexandrovna, the daughter of Czar Alexander II of Russia and wife of Alfred, the second son of Queen Victoria and Prince Albert

Ranking Candidate Answers

- But what if there are multiple candidate answers!

Q: Who was Queen Victoria's second son?

- Answer Type: **Person**

- Passage:

The Marie biscuit is named after **Marie Alexandrovna**, the daughter of **Czar Alexander II of Russia** and wife of **Alfred**, the second son of **Queen Victoria** and **Prince Albert**

Dan Jurafsky

Use machine learning:
Features for ranking candidate answers

Answer type match: Candidate contains a phrase with the correct answer type.

Pattern match: Regular expression pattern matches the candidate.

Question keywords: # of question keywords in the candidate.

Keyword distance: Distance in words between the candidate and query keywords

Novelty factor: A word in the candidate is not in the query.

Apposition features: The candidate is an appositive to question terms

Punctuation location: The candidate is immediately followed by a comma, period, quotation marks, semicolon, or exclamation mark.

Sequences of question terms: The length of the longest sequence of question terms that occurs in the candidate answer.

Candidate Answer scoring in IBM Watson

- Each candidate answer gets scores from >50 components
  - (from unstructured text, semi-structured text, triple stores)

  - logical form (parse) match between question and candidate
  - passage source reliability
  - geospatial location
    - California is "southwest of Montana"
  - temporal relationships
  - taxonomic classification

43
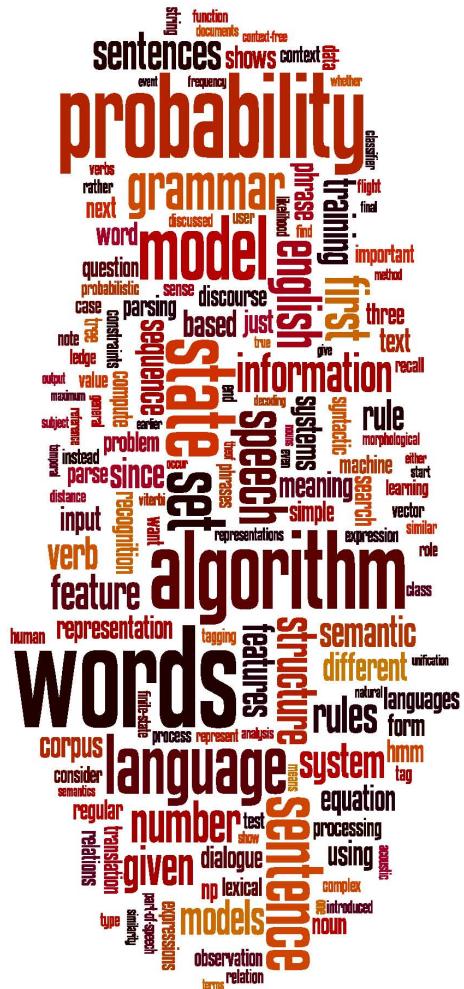
Common Evaluation Metrics

1. Accuracy (does answer match gold-labeled answer?)

2. Mean Reciprocal Rank
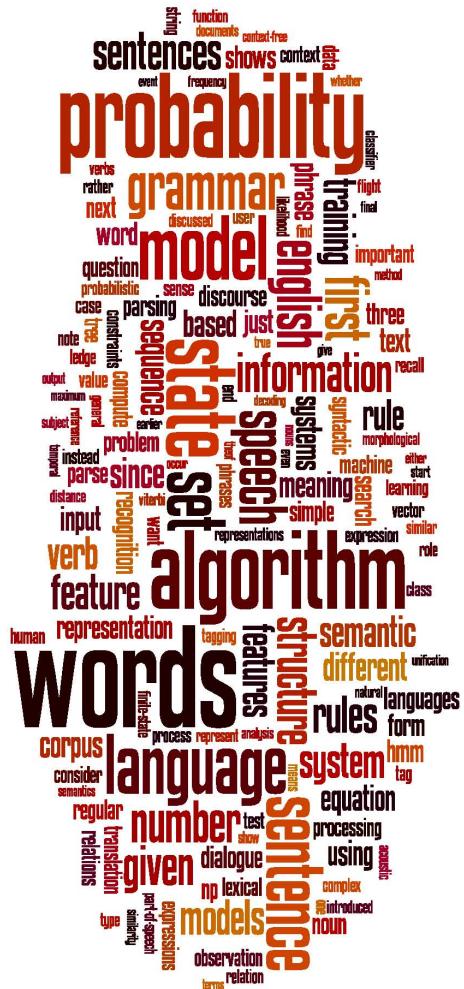
- For each query return a ranked list of M candidate answers.
- Query score is 1/Rank of the first correct answer
  - *If first answer is correct: 1*
  - *else if second answer is correct: ½*
  - *else if third answer is correct: ⅓, etc.*
  - *Score is 0 if none of the M answers are correct*
- Take the mean over all N queries

$$MRR = \frac{\sum_{i=1}^{N} \frac{1}{rank_i}}{N}$$

44

Question Answering

Passage Retrieval and Answer Extraction

Question Answering

Using Knowledge in QA

Relation Extraction

- Answers: Databases of Relations
  - born-in("Emma Goldman", "June 27 1869")
  - author-of("Cao Xue Qin", "Dream of the Red Chamber")
  - Draw from Wikipedia infoboxes, DBpedia, FreeBase, etc.

- Questions: Extracting Relations in Questions

  Whose granddaughter starred in E.T.?

  ```
  (acted-in ?x "E.T.")
     (granddaughter-of ?x ?y)
  ```

47

Temporal Reasoning

- Relation databases
  - (and obituaries, biographical dictionaries, etc.)

- IBM Watson

  "In 1594 he took a job as a tax collector in Andalusia"

  Candidates:
  - Thoreau is a bad answer (born in 1817)
  - Cervantes is possible (was alive in 1594)

48

Dan Jurafsky

Geospatial knowledge (containment, directionality, borders)

- **Beijing** is a good answer for **"Asian city"**
- **California** is **"southwest of Montana"**
- geonames.org:

49

Context and Conversation in Virtual Assistants like Siri

- Coreference helps resolve ambiguities

  U: "Book a table at Il Fornaio at 7:00 with **my mom**"

  U: "Also send **her** an email reminder"

- Clarification questions:

  U: "Chicago pizza"

  S: "Did you mean pizza restaurants in Chicago
  or Chicago-style pizza?"

Question Answering

Using Knowledge in QA