

IT Systems Engineering | Universität Potsdam

Seminar Question Answering

Question Analysis

Oliver Richter

Agenda



_

- Surface Text Patterns
 - Pattern Learning
- Named Entity Annotation
- Semantic Role Labeling
 - FrameNet
 - Dependency Parsing

Surface Text Patterns



- Characteristic answer phrases
- "When was X born?"
 - "Mozart was born in 1756."
 - □ "Ghandi (1869 1948) ..."
- Regular expressions
 - "<NAME> was born in <BIRTHDATE>"
 - "<NAME> (<BIRTHDATE> "

Surface Text Patterns – Learning of Patterns



- Pattern-learning algorithm by example
- 1.) Example query: BIRTHYEAR
 - "Mozart 1756"
- 2.) Retrieve top 1000 documents
- 3.) Create suffix tree from normalized sentences
 - □ "The great composer Mozart (1756–1791) achieved fame at a young age"
 - □ "Mozart (1756–1791) was a genius"
 - □ "The whole world would always be indebted to the great music of Mozart (1756–1791)"
 - \square "Mozart (1756 1791)", 3 times



Surface Text Patterns – Learning of Patterns (2)

- 4.) Filter resulting substrings
- 5.) Replace ...
 - ... question term by "<NAME>"
 - ... answer term by "<ANSWER>"
- 6.) Repetition for different examples
 - □ born in <ANSWER> , <NAME>
 - <NAME> was born on <ANSWER> ,
 - □ <NAME> (<ANSWER> -
 - □ <*NAME*> (<*ANSWER*)

HPI Hasso Plattne Institu

Surface Text Patterns – Learning of Patterns (3)

- 7.) Calculate Precision of patterns
 - Query with question term
 - Compare results that fit pattern with answer
- 8.) Remove bad performing patterns

7

Surface Text Patterns - Shortcomings



/

- Best performance, when question type is known
- "Where are the Rocky Mountains located?"
 - "Denver's new airport, topped with white fiberglass cones in imitation of the Rocky Mountains in the background, continues to lie empty."
- "the <NAME> in <ANSWER>"
 - No knowledge that "background" is no location
- Regular expressions → problems with long-distance dependencies

Named Entity Annotation



- Gain knowledge on type of answer
- "When" implies date
 - \Box \rightarrow Filter answer candidates

example:

Steven Paul Jobs, co-founder of Apple, was born in 1955.

person organization year

Semantic Role Labeling





Operate_vehicle

Alice destroys the item with a plane. Destroyer Undergoer Instrument

Destroying

10

Semantic Role Labeling - FrameNet



- Lexical database of English
- More than 170,000 manually annotated sentences
- **Semantic frame**: "Description of type of event, relation, or entity and the participants in it."

 $_{\text{Cook}}$ [The boys] GRILL $_{\text{Food}}$ [their catches] $_{\text{Heating instrument}}$ [on an open fire].

Semantic Role Labeling – Example (1)



- __
 - Find the questions head verb
 - □ (Who **purchased** YouTube?)
 - FrameNet: *purchase*
 - Commerce_buy Frame
 - Buyer [Subj,NP] VERB Goods [Obj,NP]
 - Buyer [Subj,NP] VERB Goods [Obj,NP] Seller [Dep,PP-from]
 - □ ...

- PP Prepositional Phrases
- NP Noun Phrases

Semantic Role Labeling – Example (2)



- Who purchased YouTube?
 - Buyer [Subj,NP] VERB Goods [Obj,NP]
- "Who" is subject → Buyer
- "YouTube" is object → Goods
- Question asks for Buyer role

Semantic Role Labeling – Answer Templates



- Creation of pattern
 - ANSWER [NP] purchased YouTube
 - YouTube was purchased by ANSWER [NP]
- Makes use of synonyms & tenses

Semantic Role Labeling – Example (3)



14

- Query:
 - "YouTube was purchased by"
- Result:
 - "On October 9, 2006, YouTube was purchased by Google for an incredible US\$1.65 billion"
- "Google" fits *Buyer* role

Semantic Role Labeling - Advantages

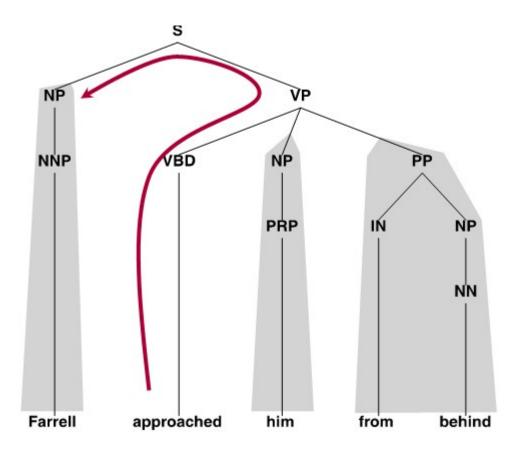


- "When was YouTube purchased?"
 - Answer is no NP but adjunct
 - \square \rightarrow can have different positions
 - Abstract frame structure for query
 - Buyer [Subj, NP, unknown] VERB Goods [Obj, NP, "YouTube"]
 - □ → "has purchased YouTube"
 - All PPs with correct type and position are candidates
 - Rules to derive roles from question words

Dependency Parsing



Identify grammatical structure and syntactical relations



Dependency Parsing - Scoring



- 1a) Same head verb?
- 1b) Same path to head?
- 2) Example's path to answer role?
- 3) Paths to other roles shared?
- 4) Surface strings match?
- If 1a) and 2) holds, answer can be extracted
- Higher score the more tests are passed

Dependency Parsing – Example (1)



- "Who purchased YouTube?"
 - □ role = ARG0, paths = $\{ \downarrow s, \downarrow subj \}$ ("Who")
 - □ role = ARG1, paths = $\{ \downarrow obj \}$ ("YouTube")
 - □ role = TMP, paths = { \downarrow mod}
- "Their aim is to compete with YouTube, which Google recently purchased for more than \$1 billion."
 - □ phrase = "Google", paths = $\{ \downarrow s, \downarrow subj \}$
 - □ phrase = "which", paths = { \ obj}
 - □ phrase = "YouTube", paths = $\{ \uparrow i \uparrow rel \}$
 - □ phrase = "for more than \$1 billion", paths = { \ \ mod}





- 1a) Same head verb? ✓
- 1b) Same path to head? X
- 2) Example's path to answer role? ✓
- 3) Paths to other roles shared?
- 4) Surface strings match?

Sources



- Deepak Ravichandran, Eduard Hovy, Learning Surface Text
 Patterns for a Question Answering System, ACL Conference, 2002
- Michael Kaisser, Bonnie Webber, Question Answering based on Semantic Roles, ACL workshop on Workshop on Important Unresolved Matters, 2005
- Preceding Presentations
 - Named Entity Annotation Stefan Klauck
 - Dependency Parsing Cindy Fähnrich
 - Semantic Role Labeling Sebastian Oergel
- FrameNet
 - https://framenet.icsi.berkeley.edu/fndrupal/about