Question-Answering: Systems & Resources

Ling573
NLP Systems & Applications
April 8, 2010

Roadmap

- Two extremes in QA systems:
  - LCC's PowerAnswer-2
  - Insight's Patterns…
- Question classification (Li & Roth)
- Resources

PowerAnswer2

- Language Computer Corp.
  - Lots of UT Dallas affiliates
- Tasks: factoid questions
- Major novel components:
  - Web-boosting of results
  - COGEX logic prover
  - Temporal event processing
  - Extended semantic chains
- Results: “Above median”: 53.4% main

Challenges: Co-reference

- Single, basic referent:

Target 27 - *Jennifer Capriati*	
Q27.2	Who is her coach?
Q27.3	Where does she live?

Challenges: Co-reference

- Single, basic referent:

Target 27 - *Jennifer Capriati*	
Q27.2	Who is her coach?
Q27.3	Where does she live?

- Multiple possible antecedents:
  - Depends on previous correct answers

Target 136 - *Shiite*	
Q136.1	Who was the first Imam of the Shiite sect of Islam?
Q136.2	Where is his tomb?
Q136.3	What was this person's relationship to the Prophet Mohammad?
Q136.4	Who was the third Imam of Shiite Muslims?
Q136.5	When did he die?

Challenges: Events

- Event answers:
  - Not just nominal concepts

Challenges: Events

- Event answers:
  - Not just nominal concepts
  - Nominal events:
    - Preakness 1998

Challenges: Events

- Event answers:
  - Not just nominal concepts
  - Nominal events:
    - Preakness 1998
  - Complex events:
    - Plane clips cable wires in Italian resort

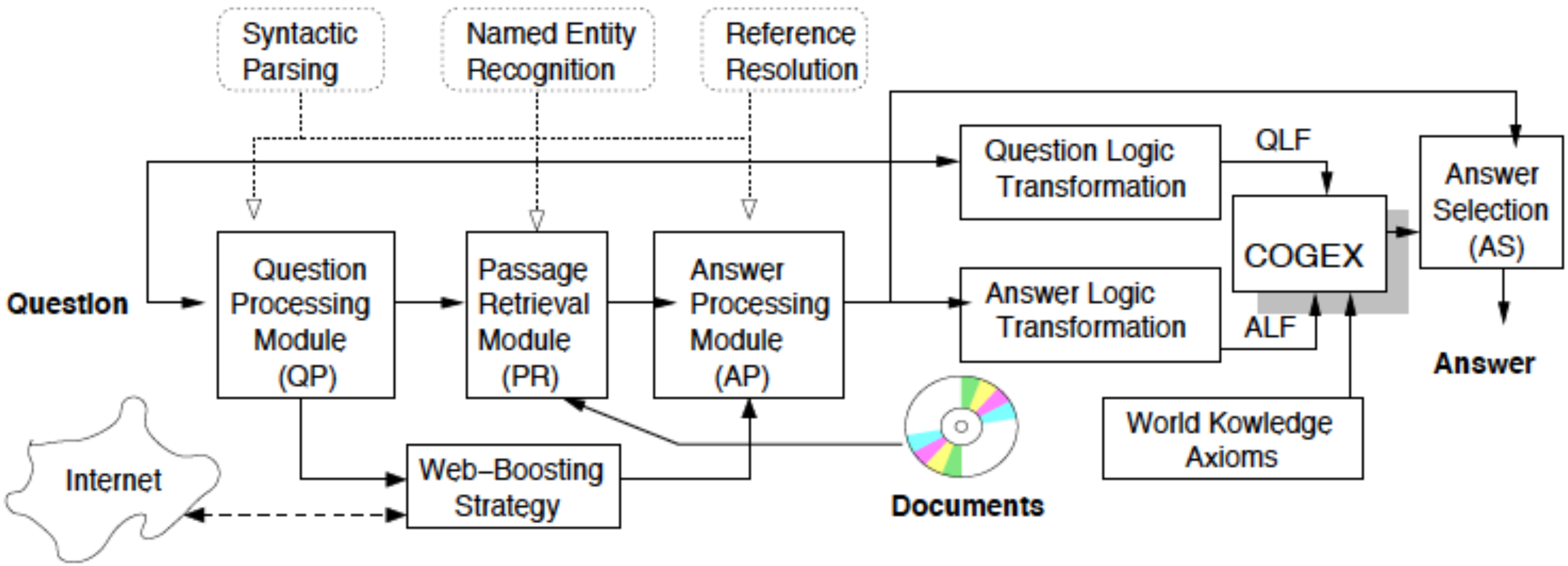
Challenges: Events

- Event answers:
  - Not just nominal concepts
  - Nominal events:
    - Preakness 1998
  - Complex events:
    - Plane clips cable wires in Italian resort

  - Establish question context, constraints

PowerAnswer-2

- Factoid QA system:



PowerAnswer-2

- Standard main components:
  - Question analysis, passage retrieval, answer processing

PowerAnswer-2

- Standard main components:
  - Question analysis, passage retrieval, answer processing
- Web-based answer boosting

PowerAnswer-2

- Standard main components:
  - Question analysis, passage retrieval, answer processing
- Web-based answer boosting
- Complex components:

PowerAnswer-2

- Standard main components:
  - Question analysis, passage retrieval, answer processing
- Web-based answer boosting
- Complex components:
  - COGEX abductive prover
  - Word knowledge, semantics:
    - Extended WordNet, etc
  - Temporal processing

Web-Based Boosting

- Create search engine queries from question

Web-Based Boosting

- Create search engine queries from question
- Extract most redundant answers from search
  - Cf. Dumais et al - AskMSR

Web-Based Boosting

- Create search engine queries from question
- Extract most redundant answers from search
  - Cf. Dumais et al - AskMSR
- Increase weight on TREC candidates that match
  - Higher weight if higher frequency

Web-Based Boosting

- Create search engine queries from question
- Extract most redundant answers from search
  - Cf. Dumais et al - AskMSR
- Increase weight on TREC candidates that match
  - Higher weight if higher frequency
- Intuition:
  - Common terms in search likely to be answer
  - QA answer search too focused on query terms

Web-Based Boosting

- Create search engine queries from question
- Extract most redundant answers from search
  - Cf. Dumais et al - AskMSR
- Increase weight on TREC candidates that match
  - Higher weight if higher frequency
- Intuition:
  - Common terms in search likely to be answer
  - QA answer search too focused on query terms
  - Reweighting improves
- Web-boosting improves significantly: 20%

Deep Processing: Query/Answer Formulation

- Preliminary shallow processing:
  - Tokenization, POS tagging, NE recognition, Preprocess

Deep Processing: Query/Answer Formulation

- Preliminary shallow processing:
  - Tokenization, POS tagging, NE recognition, Preprocess
- Parsing creates syntactic representation:
  - Focused on nouns, verbs, and particles
    - Attachment

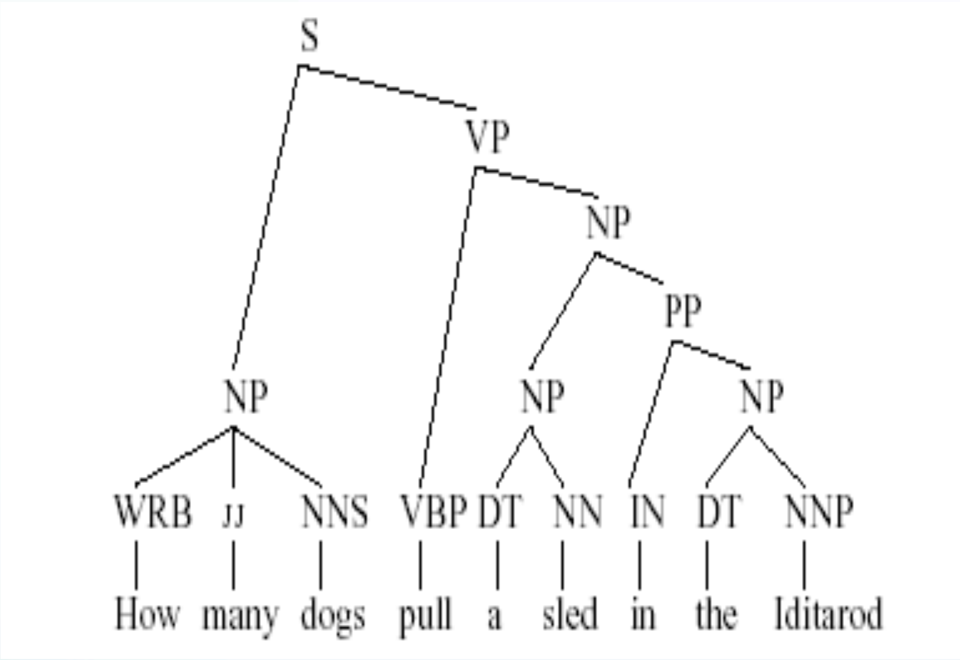
Deep Processing: Query/Answer Formulation

- Preliminary shallow processing:
  - Tokenization, POS tagging, NE recognition, Preprocess
- Parsing creates syntactic representation:
  - Focused on nouns, verbs, and particles
    - Attachment
- Coreference resolution links entity references

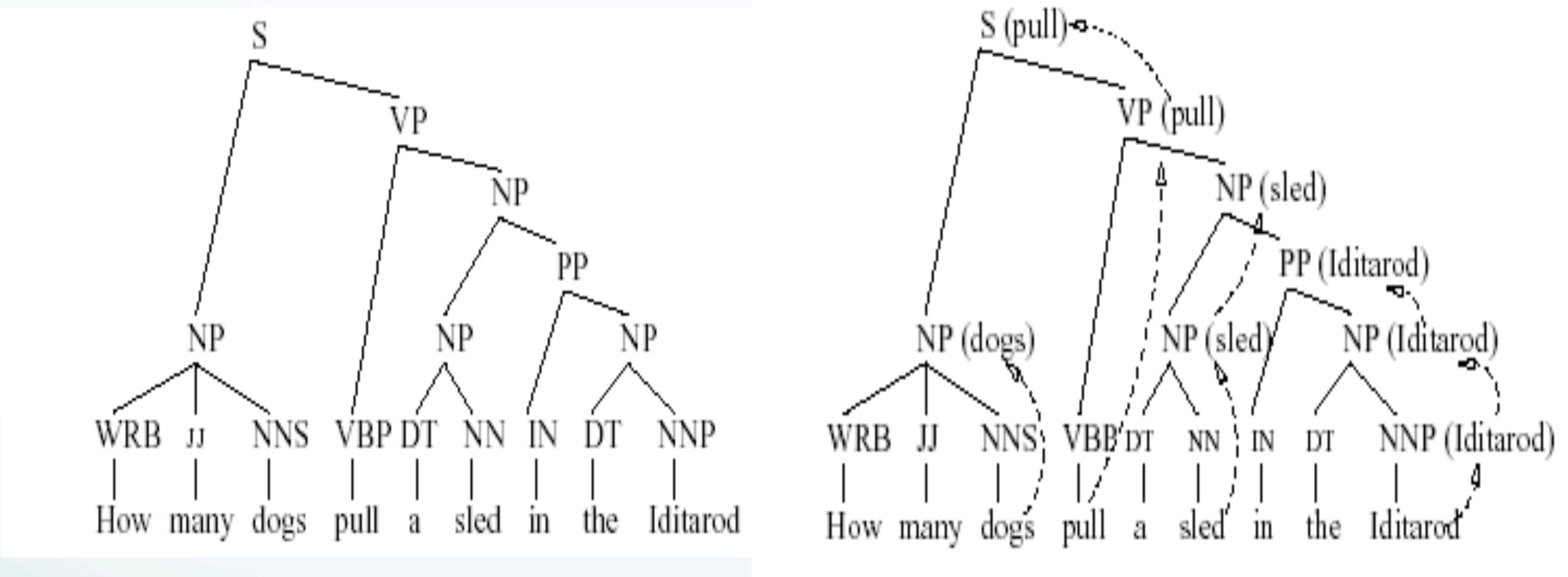
Deep Processing: Query/Answer Formulation

- Preliminary shallow processing:
  - Tokenization, POS tagging, NE recognition, Preprocess
- Parsing creates syntactic representation:
  - Focused on nouns, verbs, and particles
    - Attachment
- Coreference resolution links entity references
- Translate to full logical form
  - As close as possible to syntax

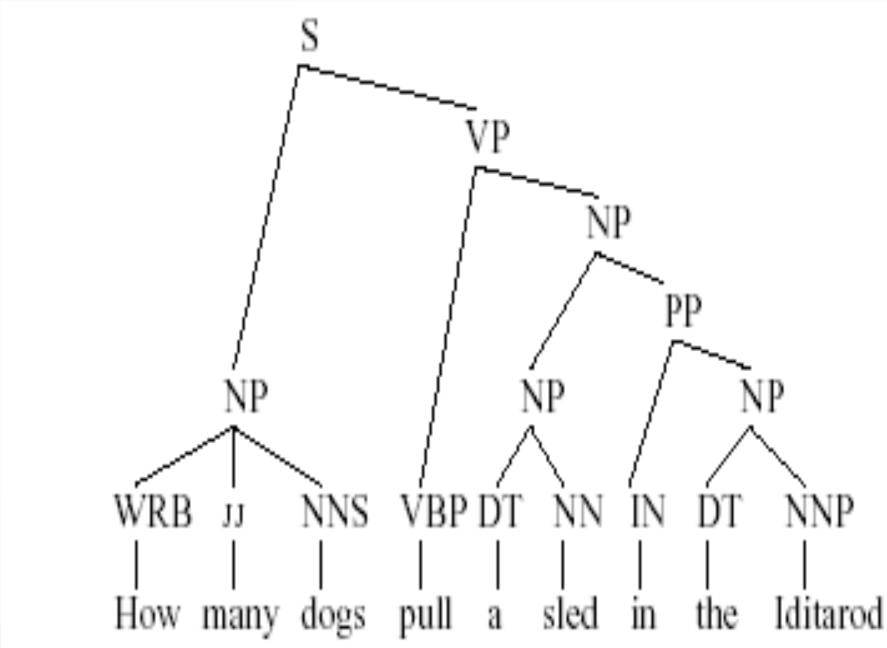
Syntax to Logical Form

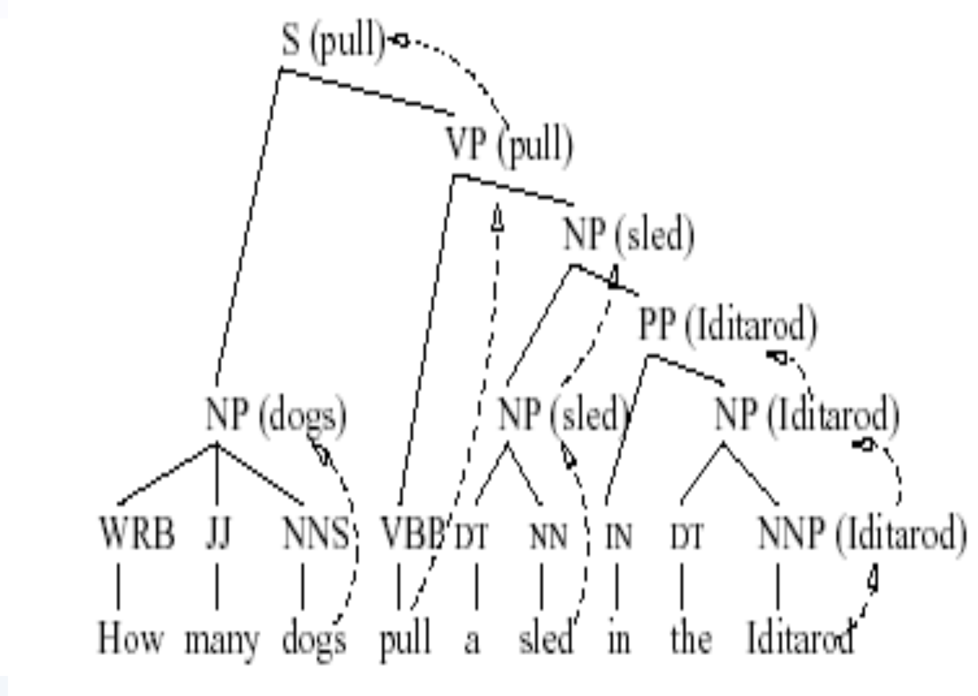


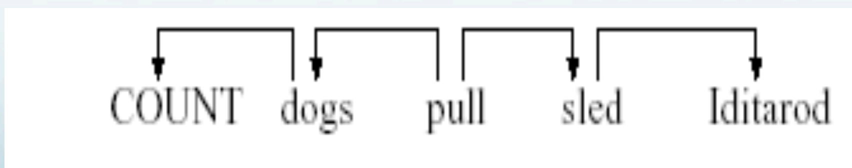
Syntax to Logical Form



Syntax to Logical Form







Deep Processing: Answer Selection

- Lexical chains:
  - Bridge gap in lexical choice b/t Q and A
    - Improve retrieval and answer selection

Deep Processing: Answer Selection

- Lexical chains:
  - Bridge gap in lexical choice b/t Q and A
    - Improve retrieval and answer selection
  - Create connections between synsets through topicality
    - *Q: When was the internal combustion engine invented?*
    - *A: The first internal-combustion engine was built in 1867.*
    - invent → create_mentally → create → build

Deep Processing: Answer Selection

- Lexical chains:
  - Bridge gap in lexical choice b/t Q and A
    - Improve retrieval and answer selection
  - Create connections between synsets through topicality
    - *Q: When was the internal combustion engine invented?*
    - *A: The first internal-combustion engine was built in 1867.*
    - invent → create_mentally → create → build
- Perform abductive reasoning b/t QLF & ALF
  - Tries to justify answer given question

Deep Processing: Answer Selection

- Lexical chains:
  - Bridge gap in lexical choice b/t Q and A
    - Improve retrieval and answer selection
  - Create connections between synsets through topicality
    - *Q: When was the internal combustion engine invented?*
    - *A: The first internal-combustion engine was built in 1867.*
    - invent → create_mentally → create → build

- Perform abductive reasoning b/t QLF & ALF
  - Tries to justify answer given question
  - Yields 10% improvement in accuracy!

Temporal Processing

- 16% of factoid questions include time reference

Temporal Processing

- 16% of factoid questions include time reference
- Index documents by date: absolute, relative

Temporal Processing

- 16% of factoid questions include time reference
- Index documents by date: absolute, relative
- Identify temporal relations b/t events
  - Store as triples of (S, E1, E2)
    - S is temporal relation signal – e.g. during, after

Temporal Processing

- 16% of factoid questions include time reference
- Index documents by date: absolute, relative
- Identify temporal relations b/t events
  - Store as triples of (S, E1, E2)
    - S is temporal relation signal – e.g. during, after
- Answer selection:
  - Prefer passages matching Question temporal constraint
  - Discover events related by temporal signals in Q & As
  - Perform temporal unification; boost good As

Temporal Processing

- 16% of factoid questions include time reference
- Index documents by date: absolute, relative
- Identify temporal relations b/t events
  - Store as triples of (S, E1, E2)
    - S is temporal relation signal – e.g. during, after
- Answer selection:
  - Prefer passages matching Question temporal constraint
  - Discover events related by temporal signals in Q & As
  - Perform temporal unification; boost good As
- Improves only by 2%
  - Mostly captured by surface forms

Results

	PowerAnswer-2
Factoid	0.713
List	0.468
Other	0.228
Overall	0.534

Table 2: Results in the main task.

Overview

- Key sources of improvement:
  - Shallow processing:
    - Web-boosting: +20%

Overview

- Key sources of improvement:
  - Shallow processing:
    - Web-boosting: +20%

  - Deep processing:
    - COGEX logic prover + semantics: 10%
    - Temporal processing: 2%

Overview

- Key sources of improvement:
  - Shallow processing:
    - Web-boosting: +20%
  - Deep processing:
    - COGEX logic prover + semantics: 10%
    - Temporal processing: 2%
  - Relation queries:
    - All relatively shallow:
      - Biggest contributors: Keyword extraction, Topic signatures

Patterns of Potential Answer Expressions…

- “Insight”
- Shallow-pattern-based approach
  - Contrasts with deep processing techniques

Patterns of Potential Answer Expressions…

- "Insight"
- Shallow-pattern-based approach
  - Contrasts with deep processing techniques

- Intuition:
  - Some surface patterns highly correlated to information

Patterns of Potential Answer Expressions…

- “Insight”
- Shallow-pattern-based approach
  - Contrasts with deep processing techniques
- Intuition:
  - Some surface patterns highly correlated to information
    - E.g. Mozart (1756-1791)

Patterns of Potential Answer Expressions…

- “Insight”
- Shallow-pattern-based approach
  - Contrasts with deep processing techniques
- Intuition:
  - Some surface patterns highly correlated to information
    - E.g. Mozart (1756-1791)
    - Person – birthdate, death date
      - Pattern: Capitalized word; paren, 4 digits; dash; 4 digits; paren
      - Attested 850 times in a corpus

Pattern Library

- Potentially infinite patterns

Pattern Library

- Potentially infinite patterns
- Pattern structure:
  - Fixed components:
    - Words, characters, symbols

Pattern Library

- Potentially infinite patterns
- Pattern structure:
  - Fixed components:
    - Words, characters, symbols
  - Variable components:
    - Usually query terms and answer terms

Pattern Library

- Potentially infinite patterns
- Pattern structure:
  - Fixed components:
    - Words, characters, symbols
  - Variable components:
    - Usually query terms and answer terms
  - List of 51 pattern elements – combined for patterns
    - Ordered or unordered

Pattern Library

- Potentially infinite patterns
- Pattern structure:
  - Fixed components:
    - Words, characters, symbols
  - Variable components:
    - Usually query terms and answer terms
  - List of 51 pattern elements – combined for patterns
    - Ordered or unordered
  - More complex patterns are typically more indicative

Other Examples

- Post questions: Who is the Queen of the Netherlands?

Other Examples

- Post questions: Who is the Queen of the Netherlands?
- Beatrix, Queen of the Netherlands

Other Examples

- Post questions: Who is the Queen of the Netherlands?
- Beatrix, Queen of the Netherlands
- Pattern elements:
  - Country name
  - Post name
  - Person name
  - Title (optional)
    - In some order

Basic Approach

- Question analysis:
  - Identify detailed question type

Basic Approach

- Question analysis:
  - Identify detailed question type
- Passage retrieval
  - Collect large number of retrieval snippets
    - Possibly with query expansion

Basic Approach

- Question analysis:
  - Identify detailed question type
- Passage retrieval
  - Collect large number of retrieval snippets
    - Possibly with query expansion
- Answer processing:
  - Find matching patterns in candidates
    - 10s of patterns/answer type

Results

- Best result in TREC-10
- MRR (strict) 0.676:
  - Correct: 289; 120 unanswered
- Retrieval based on shallow patterns
  - Bag of patterns, and sequences
  - Still highly effective

Question Classification: Li&Roth

Roadmap

- Motivation:

Why Question Classification?

Why Question Classification?

- Question classification categorizes possible answers

Why Question Classification?

- Question classification categorizes possible answers
  - Constrains answers types to help find, verify answer

    Q: What Canadian city has the largest population?
    - Type?

Why Question Classification?

- Question classification categorizes possible answers
  - Constrains answers types to help find, verify answer

    Q: What Canadian city has the largest population?
    - Type? -> City
    - Can ignore all non-city NPs

Why Question Classification?

- Question classification categorizes possible answers
  - Constrains answers types to help find, verify answer

    Q: What Canadian city has the largest population?

    - Type? -> City
    - Can ignore all non-city NPs

  - Provides information for type-specific answer selection
    - *Q: What is a prism?*
    - Type? ->

Why Question Classification?

- Question classification categorizes possible answers
  - Constrains answers types to help find, verify answer

    Q: What Canadian city has the largest population?
    - Type? -> City
    - Can ignore all non-city NPs

  - Provides information for type-specific answer selection
    - *Q: What is a prism?*
    - Type? -> Definition
      - Answer patterns include: 'A prism is…'

Challenges

Challenges

- Variability:
  - What tourist attractions are there in Reims?
  - What are the names of the tourist attractions in Reims?
  - What is worth seeing in Reims?
    - Type?

Challenges

- Variability:
  - What tourist attractions are there in Reims?
  - What are the names of the tourist attractions in Reims?
  - What is worth seeing in Reims?
    - Type? -> Location

Challenges

- Variability:
  - What tourist attractions are there in Reims?
  - What are the names of the tourist attractions in Reims?
  - What is worth seeing in Reims?
    - Type? -> Location
- Manual rules?

Challenges

- Variability:
  - What tourist attractions are there in Reims?
  - What are the names of the tourist attractions in Reims?
  - What is worth seeing in Reims?
    - Type? -> Location
- Manual rules?
  - Nearly impossible to create sufficient patterns
- Solution?

Challenges

- Variability:
  - What tourist attractions are there in Reims?
  - What are the names of the tourist attractions in Reims?
  - What is worth seeing in Reims?
    - Type? -> Location
- Manual rules?
  - Nearly impossible to create sufficient patterns
- Solution?
  - Machine learning – rich feature set

Approach

- Employ machine learning to categorize by answer type
  - Hierarchical classifier on semantic hierarchy of types
    - Coarse vs fine-grained
      - Up to 50 classes

  - Differs from text categorization?

Approach

- Employ machine learning to categorize by answer type
  - Hierarchical classifier on semantic hierarchy of types
    - Coarse vs fine-grained
      - Up to 50 classes

  - Differs from text categorization?
    - Shorter (much!)
    - Less information, but
    - Deep analysis more tractable

Approach

- Exploit syntactic and semantic information
  - Diverse semantic resources

Approach

- Exploit syntactic and semantic information
  - Diverse semantic resources
    - Named Entity categories
    - WordNet sense
    - Manually constructed word lists
    - Automatically extracted semantically similar word lists

Approach

- Exploit syntactic and semantic information
  - Diverse semantic resources
    - Named Entity categories
    - WordNet sense
    - Manually constructed word lists
    - Automatically extracted semantically similar word lists
- Results:
  - Coarse: 92.5%; Fine: 89.3%
  - Semantic features reduce error by 28%

Question Hierarchy

Class	#	Class	#
ABBREVIATION	18	term	19
abbreviation	2	vehicle	7
expression	16	word	0
DESCRIPTION	153	**HUMAN**	171
definition	126	group	24
description	13	individual	140
manner	7	title	4
reason	7	description	3
ENTITY	174	**LOCATION**	195
animal	27	city	44
body	5	country	21
color	12	mountain	5
creative	14	other	114
currency	8	state	11
disease/medicine	3	**NUMERIC**	289
event	6	code	1
food	7	count	22
instrument	1	date	146
lang	3	distance	38
letter	0	money	9
other	19	order	0
plant	7	other	24
product	9	period	18
religion	1	percent	7
sport	3	speed	9
substance	20	temp	7
symbol	2	vol.size	4
technique	1	weight	4

Learning a Hierarchical Question Classifier

- Many manual approaches use only :

Learning a Hierarchical Question Classifier

- Many manual approaches use only :
  - Small set of entity types, set of handcrafted rules

Learning a Hierarchical Question Classifier

- Many manual approaches use only :
  - Small set of entity types, set of handcrafted rules
    - Note: Webclopedia’s 96 node taxo w/276 manual rules

Learning a Hierarchical Question Classifier

- Many manual approaches use only :
  - Small set of entity types, set of handcrafted rules
    - Note: Webclopedia's 96 node taxo w/276 manual rules
- Learning approaches can learn to generalize
  - Train on new taxonomy, but

Learning a Hierarchical Question Classifier

- Many manual approaches use only :
  - Small set of entity types, set of handcrafted rules
    - Note: Webclopedia's 96 node taxo w/276 manual rules
- Learning approaches can learn to generalize
  - Train on new taxonomy, but
    - Someone still has to label the data…
- Two step learning: (Winnow)
  - Same features in both cases

Learning a Hierarchical Question Classifier

- Many manual approaches use only :
  - Small set of entity types, set of handcrafted rules
    - Note: Webclopedia's 96 node taxo w/276 manual rules
- Learning approaches can learn to generalize
  - Train on new taxonomy, but
    - Someone still has to label the data…
- Two step learning: (Winnow)
  - Same features in both cases
    - First classifier produces (a set of) coarse labels
    - Second classifier selects from fine-grained children of coarse tags generated by the previous stage
    - Select highest density classes above threshold

Features for Question Classification

- Primitive lexical, syntactic, lexical-semantic features
  - Automatically derived
  - Combined into conjunctive, relational features
  - Sparse, binary representation

Features for Question Classification

- Primitive lexical, syntactic, lexical-semantic features
  - Automatically derived
  - Combined into conjunctive, relational features
  - Sparse, binary representation
- Words
  - Combined into ngrams

Features for Question Classification

- Primitive lexical, syntactic, lexical-semantic features
  - Automatically derived
  - Combined into conjunctive, relational features
  - Sparse, binary representation
- Words
  - Combined into ngrams
- Syntactic features:
  - Part-of-speech tags
  - Chunks
  - Head chunks : 1^{st} N, V chunks after Q-word

Syntactic Feature Example

- Q: Who was the first woman killed in the Vietnam War?

Syntactic Feature Example

- Q: Who was the first woman killed in the Vietnam War?
- POS: [Who WP] [was VBD] [the DT] [first JJ] [woman NN] [killed VBN] {in IN] [the DT] [Vietnam NNP] [War NNP] [? .]

Syntactic Feature Example

- Q: Who was the first woman killed in the Vietnam War?
- POS: [Who WP] [was VBD] [the DT] [first JJ] [woman NN] [killed VBN] {in IN] [the DT] [Vietnam NNP] [War NNP] [? .]
- Chunking: [NP Who] [VP was] [NP the first woman] [VP killed] [PP in] [NP the Vietnam War] ?

Syntactic Feature Example

- Q: Who was the first woman killed in the Vietnam War?
- POS: [Who WP] [was VBD] [the DT] [first JJ] [woman NN] [killed VBN] {in IN] [the DT] [Vietnam NNP] [War NNP] [? .]
- Chunking: [NP Who] [VP was] [NP the first woman] [VP killed] [PP in] [NP the Vietnam War] ?
- Head noun chunk: ‘the first woman’

Semantic Features

- Treat analogously to syntax?

Semantic Features

- Treat analogously to syntax?
  - Q1:What's the semantic equivalent of POS tagging?

Semantic Features

- Treat analogously to syntax?
  - Q1:What's the semantic equivalent of POS tagging?
  - Q2: POS tagging > 97% accurate;
    - Semantics? Semantic ambiguity?

Semantic Features

- Treat analogously to syntax?
  - Q1:What's the semantic equivalent of POS tagging?
  - Q2: POS tagging > 97% accurate;
    - Semantics? Semantic ambiguity?

- A1: Explore different lexical semantic info sources
    - Differ in granularity, difficulty, and accuracy

Semantic Features

- Treat analogously to syntax?
  - Q1:What's the semantic equivalent of POS tagging?
  - Q2: POS tagging > 97% accurate;
    - Semantics? Semantic ambiguity?

- A1: Explore different lexical semantic info sources
    - Differ in granularity, difficulty, and accuracy
  - Named Entities
  - WordNet Senses
  - Manual word lists
  - Distributional sense clusters

Tagging & Ambiguity

- Augment each word with semantic category

- What about ambiguity?
  - E.g. ‘water’ as ‘liquid’ or ‘body of water’

Tagging & Ambiguity

- Augment each word with semantic category

- What about ambiguity?
  - E.g. ‘water’ as ‘liquid’ or ‘body of water’
  - Don’t disambiguate
    - Keep all alternatives
    - Let the learning algorithm sort it out
    - Why?

Semantic Categories

- Named Entities
  - Expanded class set: 34 categories
    - E.g. Profession, event, holiday, plant,...

Semantic Categories

- Named Entities
  - Expanded class set: 34 categories
    - E.g. Profession, event, holiday, plant,...
- WordNet: IS-A hierarchy of senses
  - All senses of word + direct hyper/hyponyms

Semantic Categories

- Named Entities
  - Expanded class set: 34 categories
    - E.g. Profession, event, holiday, plant,…
- WordNet: IS-A hierarchy of senses
  - All senses of word + direct hyper/hyponyms
- Class-specific words
  - Manually derived from 5500 questions
    - E.g. Class: Food
      - {alcoholic, apple, beer, berry, breakfast brew butter candy cereal champagne cook delicious eat fat ..}
      - Class is semantic tag for word in the list

Semantic Types

- Distributional clusters:
  - Based on Pantel and Lin
  - Cluster based on similarity in dependency relations
  - Word lists for 20K English words

Semantic Types

- Distributional clusters:
  - Based on Pantel and Lin
  - Cluster based on similarity in dependency relations
  - Word lists for 20K English words
    - Lists correspond to word senses
    - Water:
      - Sense 1: { oil gas fuel food milk liquid}
      - Sense 2: {air moisture soil heat area rain}
      - Sense 3: {waste sewage pollution runoff}

Semantic Types

- Distributional clusters:
  - Based on Pantel and Lin
  - Cluster based on similarity in dependency relations
  - Word lists for 20K English words
    - Lists correspond to word senses
    - Water:
      - Sense 1: { oil gas fuel food milk liquid}
      - Sense 2: {air moisture soil heat area rain}
      - Sense 3: {waste sewage pollution runoff}
  - Treat head word as semantic category of words on list

Evaluation

- Assess hierarchical coarse->fine classification
- Assess impact of different semantic features
- Assess training requirements for diff’t feature set

Evaluation

- Assess hierarchical coarse->fine classification
- Assess impact of different semantic features
- Assess training requirements for diff't feature set
- Training:
  - 21.5K questions from TREC 8,9; manual; USC data
- Test:
  - 1K questions from TREC 10,11

Evaluation

- Assess hierarchical coarse->fine classification
- Assess impact of different semantic features
- Assess training requirements for diff’t feature set
- Training:
  - 21.5K questions from TREC 8,9; manual; USC data
- Test:
  - 1K questions from TREC 10,11
- Measures: Accuracy and class-specific precision

Results

- Syntactic features only:

Classifier	Word	POS	Chunk	Head(SYN)
Coarse	85.10	91.80	91.80	**92.50**
Fine	82.60	84.90	84.00	**85.00**

  - POS useful; chunks useful to contribute head chunks
  - Fine categories more ambiguous

Results

- Syntactic features only:

Classifier	Word	POS	Chunk	Head(SYN)
Coarse	85.10	91.80	91.80	**92.50**
Fine	82.60	84.90	84.00	**85.00**

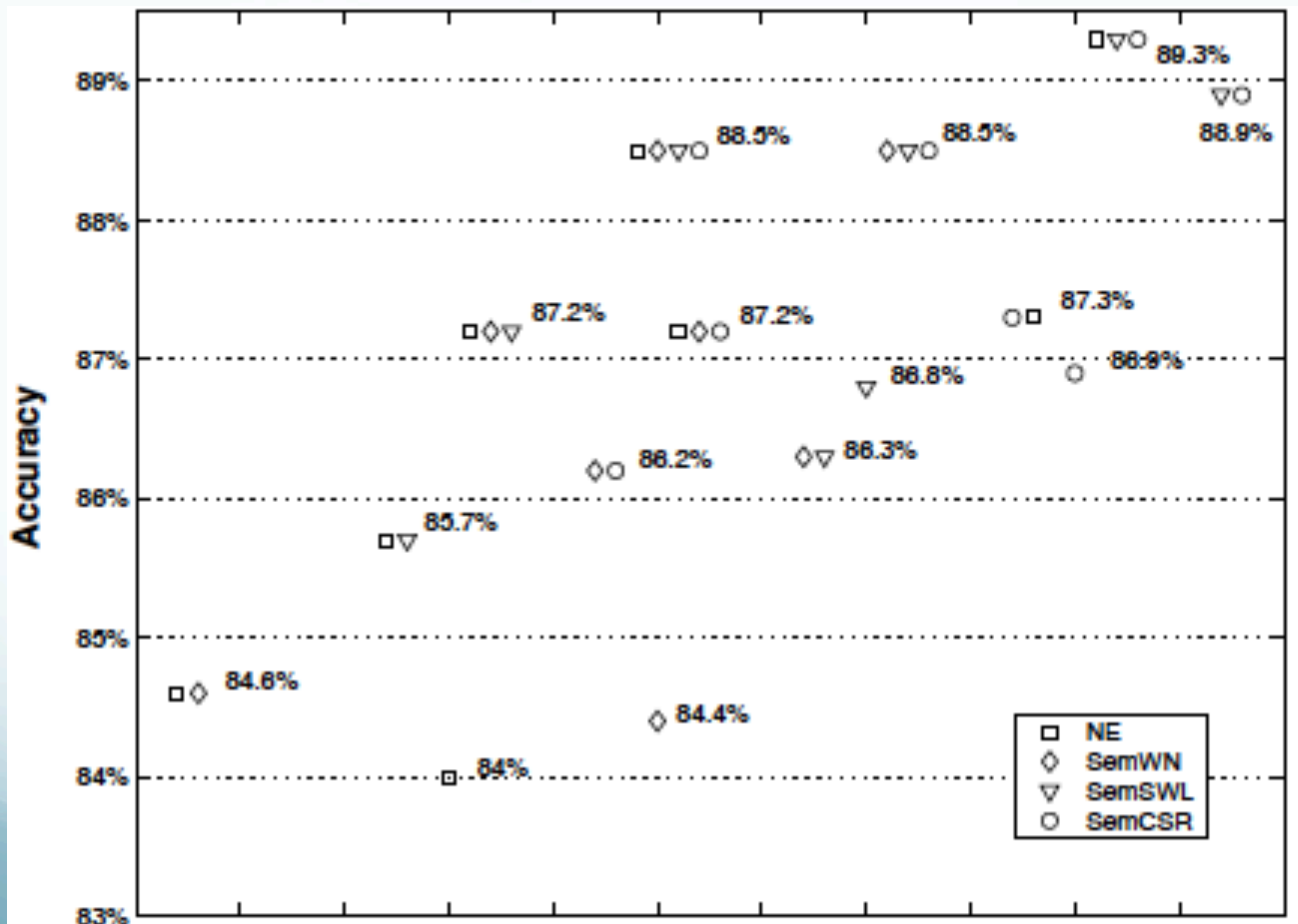
  - POS useful; chunks useful to contribute head chunks
  - Fine categories more ambiguous
- Semantic features:
  - Best combination: SYN, NE, Manual & Auto word lists
    - Coarse: same; Fine: 89.3% (28.7% error reduction)

Results

- Syntactic features only:

Classifier	Word	POS	Chunk	Head(SYN)
Coarse	85.10	91.80	91.80	**92.50**
Fine	82.60	84.90	84.00	**85.00**

  - POS useful; chunks useful to contribute head chunks
  - Fine categories more ambiguous
- Semantic features:
  - Best combination: SYN, NE, Manual & Auto word lists
    - Coarse: same; Fine: 89.3% (28.7% error reduction)
- Wh-word most common class: 41%



Accuracy
89%
88%
87%
86%
85%
84%
83%
89.3%
88.9%
88.5%
88.5%
87.2%
87.2%
87.3%
86.8%
86.9%
86.2%
86.3%
85.7%
84.6%
84.4%
84%
NE
SemWN
SemSWL
SemCSR

Class	#	Precision[c]	Class	#	Precision[c]
abb	2	100%	desc	25	36%
exp	17	94.11%	manner	8	87.5%
animal	27	85.18%	reason	7	85.71%
body	4	100%	gr	19	89.47%
color	12	100%	ind	154	90.25%
cremat	13	76.92%	title	4	100%
currency	6	100%	desc	3	100%
dismed	4	50%	city	41	97.56%
event	4	75%	country	21	95.23%
food	6	100%	mount	2	100%
instru	1	100%	LOC:other	116	89.65%
lang	3	100%	state	14	78.57%
ENTY:other	24	37.5%	count	24	91.66%
plant	3	100%	date	145	100%
product	6	66.66%	dist	37	97.29%
religion	1	100%	money	6	100%
sport	4	75%	NUM:other	15	93.33%
substance	21	80.95%	period	20	85%
symbol	2	100%	perc	9	77.77%
termeq	22	63.63%	speed	8	100%
veh	7	71.42%	temp	4	100%
def	125	97.6%	weight	4	100%
TOTAL	**1000**	**89.3%**			

Observations

- Effective coarse and fine-grained categorization
  - Mix of information sources and learning
  - Shallow syntactic features effective for coarse
  - Semantic features improve fine-grained
    - Most feature types help
      - WordNet features appear noisy
      - Use of distributional sense clusters dramatically increases feature dimensionality

NE	0.23
SemWN	16
SemCSR	23
SemSWL	557

Software Resources

- Build on existing tools
  - Focus on QA specific tasks
- General: Machine learning tools

Software Resources

- General: Machine learning tools
  - Mallet: http://mallet.cs.umass.edu
  - Weka toolkit: *www.cs.waikato.ac.nz/ml/weka/*

Software Resources

- General: Machine learning tools
  - Mallet: http://mallet.cs.umass.edu
  - Weka toolkit: *www.cs.waikato.ac.nz/ml/weka/*

- NLP toolkits, collections:
  - GATE: http://gate.ac.uk
  - NLTK: http://www.nltk.org
  - LingPipe: *alias-i.com/lingpipe/*
  - Stanford NLP tools: http://nlp.stanford.edu/software/

Software Resources: Specific

- Information retrieval:
  - Lucene: http://lucene.apache.org (on patas)
    - Standard system, tutorials
  - Indri/Lemur: http://www.lemurproject.org/indri/
    - High quality research system
  - Managing Gigabytes: http://ww2.cs.mu.oz.au/mg//
    - Linked to textbook on IR

Software Resources: Cont'd

- POS taggers:
  - Stanford POS tagger
  - Treetagger
  - Maxent POS tagger
  - Brill tagger
- Stemmers: http://snowball.tartarus.org
  - Implementations of Porter stemmer in many langs
- Sentence splitters
  - NIST

Software Resources

- Parsers:
  - Constituency parser
    - Stanford parser
    - Collins/Bikel parser
    - Charniak parser
  - Dependency parsers
    - Minipar
- WSD packages:
  - WordNet::Similarity

Software Resources

- Semantic analyzer:
  - Shalmaneser
- Databases, ontologies:
  - WordNet
  - FrameNet
  - PropBank

Information Resources

- Proxies for world knowledge:
  - WordNet: Synonymy; IS-A hierarchy

Information Resources

- Proxies for world knowledge:
  - WordNet: Synonymy; IS-A hierarchy
  - Wikipedia

Information Resources

- Proxies for world knowledge:
  - WordNet: Synonymy; IS-A hierarchy
  - Wikipedia
  - Web itself
  - ....
- Training resources:
  - Question classification sets (UIUC)
  - Other TREC QA data (Questions, Answers)