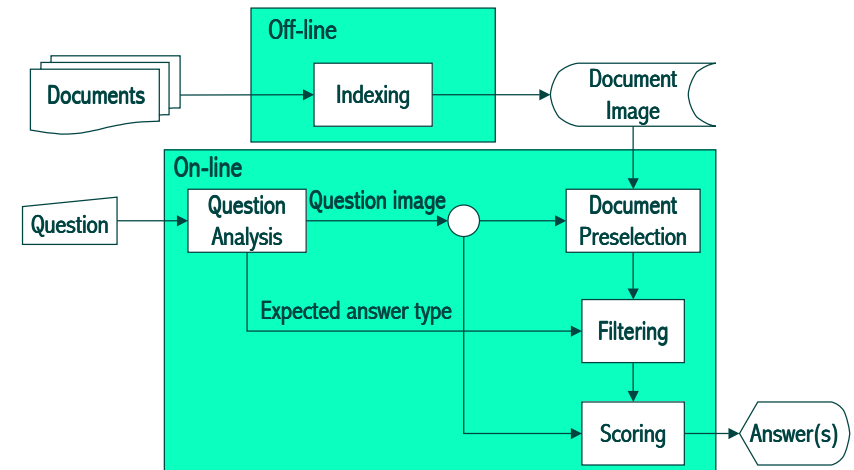


Towards Semantic-based Overlap Measures for Question Answering

Diego Mollá-Aliod
10 Dec. 2003

Architecture of a QA System

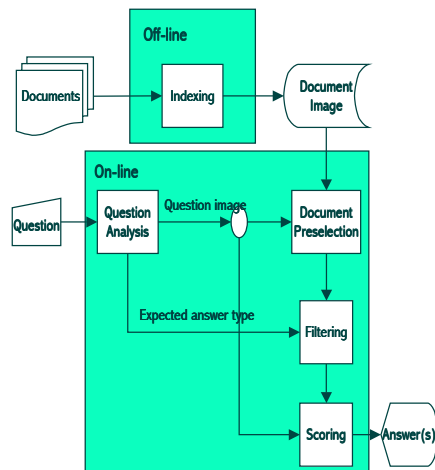


©2003 Macquarie University

Towards Semantic-based Overlap Measures for Question Answering

2/24

Architecture of a QA System



- Question Analysis
 - Question classification
 - Find answer type and focus
- Indexing, Document Preselection
 - IR techniques
 - Bag of Words
- Filtering
 - IE techniques
 - NE recognition
- Scoring
 - Pattern matching
 - Syntax, semantics



©2003 Macquarie University

Towards Semantic-based Overlap Measures for Question Answering

3/24

Grammatical Relations

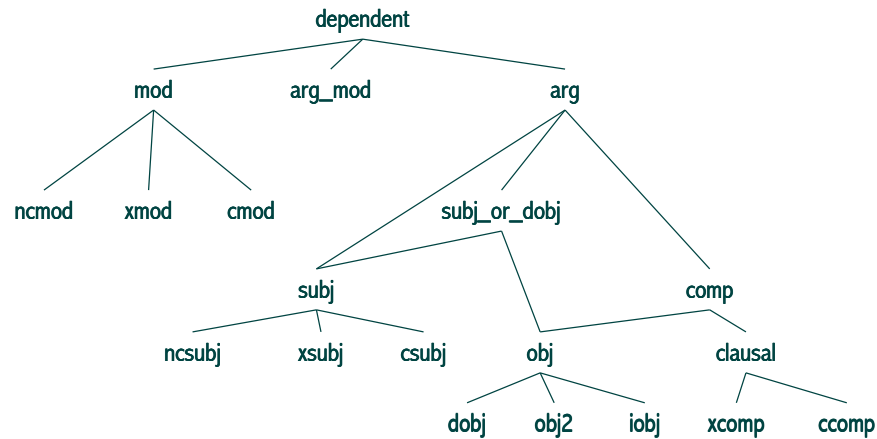
- Carroll, Briscoe, Sanfilippo “Parser evaluation: a survey and a new proposal”, Proceedings LREC98
- Purpose
 - Compare the output of different parsers
- How
 - Use relations between words (subject, object, etc)
 - Inspired in dependency theory
 - Hierarchy of grammatical relations
 - To accommodate different parser granularities

©2003 Macquarie University

Towards Semantic-based Overlap Measures for Question Answering

4/24

Grammatical Relations (Carroll et al. 1998)



Parser Accuracy

Corpus and methodology: Carroll and Briscoe

• 500 sentences, 10,000 words

		<i>With Link Grammar</i>	<i>With Conexor FDG</i>
<i>Precision</i>	SUBJ	50.3%	73.6%
	OBJ	48.5%	84.8%
	XCOMP	62.2%	76.2%
	MOD	57.2%	63.7%
	Average	54.6%	74.6%
<i>Recall</i>	SUBJ	39.1%	64.5%
	OBJ	50%	53.4%
	XCOMP	32.1%	64.7%
	MOD	53.7%	56.2%
	Average	43.7%	59.7%

QA with Grammatical Relations

– *A man named Richard Sears has been playing a joke on shoppers.*

(detmod man a),

(subj name man), (dobj name richard_sears),

(detmod joke a), (subj play man), (aux play have),

(aux play be),

(ncmod shopper play on), (dobj play joke)

– *Who played a joke on shoppers?*

(subj play who), (dobj play joke),

(ncmod shopper play on), (detmod joke a)

Flat Logical Forms

- Avoid the use of nested expressions
 - Conjunction of predicates with all variables existentially bounded
 - Reify objects, events, and properties
- Also called minimal logical forms because they encode minimal information for the task of question answering
 - Incremental extensibility

QA with Flat Logical Forms

- *A man named Richard Sears has been playing a joke on shoppers.*

```
holds(v_o10), object('man',v_o2,[v_x2]),
evt('name',v_e3,[v_X3,v_x2,v_x4]),
object('play_on',v_e5,[v_x2,v_x10,v_x12]),
object('richard_sears',v_o4,[v_x4]),
object('joke',v_o10,[v_x10]),
object('shopper',v_o12,[v_x12])
```

- *Who played a joke on shoppers?*

```
holds(v_e2), object('who',v_o1,[v_x1]),
evt('play_on',v_e2,[v_x1,v_x4,v_x6]),
object('joke',v_o4,[v_x4]), object('shopper',v_o6,[v_x6])
```

The Evaluation – Document Set

- Remedia Publications' Reading Comprehension Tests
- Levels 2, 3, 4, and 5
- Every document contains 5 questions (*who, what, when, where, why*)
- The answer is always in the text
- Files:
 - Original text (*.txt)
 - Coreferences (*.txt.coref)
 - Named Entities (*.txt.ne)
 - Answers marked-up (*.txt.snra)
 - Answers extracted (*.txt.wdra.xml)

Corpus Used – Sample of file rm4-16.txt

1989 Remedia Publications, Comprehension/5Ws-4
Watch Out for Sears!

(North Redwood, Minn., September, 1889) A man named Richard Sears has been playing a joke on shoppers.

Sears likes to sell items by mail. Not long ago, he ran an ad in some newspapers in small towns. The ad showed a drawing of lovely furniture. There was a sofa and two chairs.

The ad said the furniture was for sale. It said the pieces were made of fine metal frames and were beautiful to see. The ad said that for a short time only, these chairs would be shipped to all who paid 95 cents.

This message sounded too good to be true. Still, a lot of people sent in their money. Imagine their surprise when they received the furniture. The furniture was made for a doll house! They were tiny pieces.

Some people complained. That's when Sears showed them the tiny print in his ad.

In very small letters, he had included the word "miniature." That means the furniture was not full size. Sears says he did this to get attention.

1. Who played a joke on shoppers?
2. What does the Sears ad offer?
3. When did Sears play this joke?
4. Where is the word "miniature"?
5. Why did Sears play this joke?

Corpus Used – Sample of file rm4-16.txt.ne

1989 Remedia Publications, Comprehension/5Ws-4

Watch Out for <ENAMEX TYPE="PERSON">Sears</ENAMEX>!

(<ENAMEX TYPE="LOCATION">North Redwood</ENAMEX>, <ENAMEX TYPE="LOCATION">Minn.</ENAMEX>, <TIMEX TYPE="DATE">September, 1889</TIMEX>) A man named <ENAMEX TYPE="PERSON">Richard Sears</ENAMEX> has been playing a joke on shoppers.

<ENAMEX TYPE="PERSON">Sears</ENAMEX> likes to sell items by mail. Not long ago, he ran an ad in some newspapers in small towns. The ad showed a drawing of lovely furniture. There was a sofa and two chairs.

The ad said the furniture was for sale. It said the pieces were made of fine metal frames and were beautiful to see. The ad said that for a short time only, these chairs would be shipped to all who paid <NUMEX TYPE="MONEY">95 cents</NUMEX>.

This message sounded too good to be true. Still, a lot of people sent in their money. Imagine their surprise when they received the furniture. The furniture was made for a doll house! They were tiny pieces.

Some people complained. That's when <ENAMEX TYPE="PERSON">Sears</ENAMEX> showed them the tiny print in his ad.

In very small letters, he had included the word "miniature." That means the furniture was not full size. <ENAMEX TYPE="PERSON">Sears</ENAMEX> says he did this to get attention.

1. Who played a joke on shoppers?
2. What does the <ENAMEX TYPE="PERSON">Sears</ENAMEX> ad offer?
3. When did <ENAMEX TYPE="PERSON">Sears</ENAMEX> play this joke?
4. Where is the word "miniature"?
5. Why did <ENAMEX TYPE="PERSON">Sears</ENAMEX> play this joke?

Corpus Used – Sample of file rm4-16.txt.snra

1989 Remedia Publications, Comprehension/5Ws-4

Watch Out for Sears!

<ANSQ3>(North Redwood, Minn., September, 1889)</ANSQ3> <ANSQ1>A man named Richard Sears has been playing a joke on shoppers.</ANSQ1>

Sears likes to sell items by mail. <ANSQ3>Not long ago, he ran an ad in some newspapers in small towns.</ANSQ3> The ad showed a drawing of lovely furniture. There was a sofa and two chairs.

<ANSQ2>The ad said the furniture was for sale.</ANSQ2> It said the pieces were made of fine metal frames and were beautiful to see. <ANSQ2>The ad said that for a short time only, these chairs would be shipped to all who paid 95 cents.</ANSQ2>

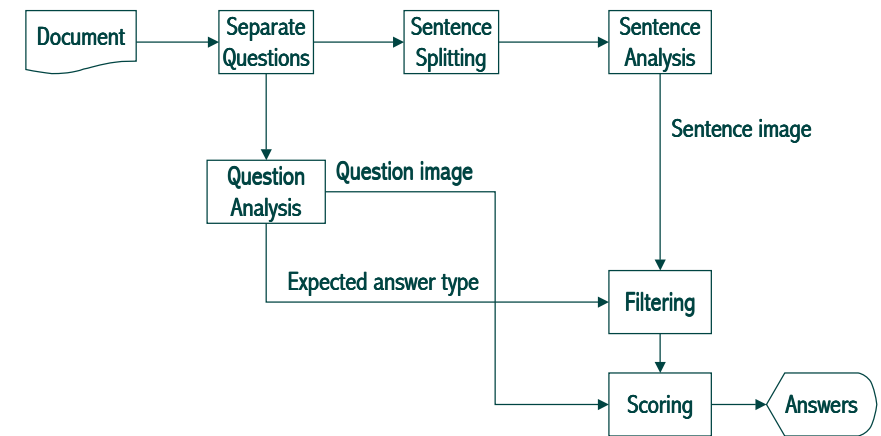
This message sounded too good to be true. Still, a lot of people sent in their money. Imagine their surprise when they received the furniture. The furniture was made for a doll house! They were tiny pieces.

Some people complained. That's when Sears showed them the tiny print in his ad.

In very small letters, he had included the word "miniature." That means the furniture was not full size. <ANSQ5>Sears says he did this to get attention.</ANSQ5>

1. Who played a joke on shoppers?
2. What does the Sears ad offer?
3. When did Sears play this joke?
4. Where is the word "miniature"?
5. Why did Sears play this joke?

Architecture of Our QA System



©2003 Macquarie University

Towards Semantic-based Overlap Measures for Question Answering

14/24

Our QA Prototype

• Question Analysis

Regex	Expected Answer Type
^Who	person, organization
^What	any
^When	date, time
^Where	location
^Why	any

• Filtering

- No expected answer: penalization of -100

• Scoring

- Overlap of GramRels or FLFs

©2003 Macquarie University

Towards Semantic-based Overlap Measures for Question Answering

15/24

The Evaluation

• Mean Reciprocal Rank (MRR)

- Used in past QA tracks of TREC
- Examine the 5 highest-ranking answer candidates of each question

$$RR = \frac{1}{\text{rank of first correct answer}}, \text{ or } 0$$

- MRR = mean of all RR

©2003 Macquarie University

Towards Semantic-based Overlap Measures for Question Answering

16/24

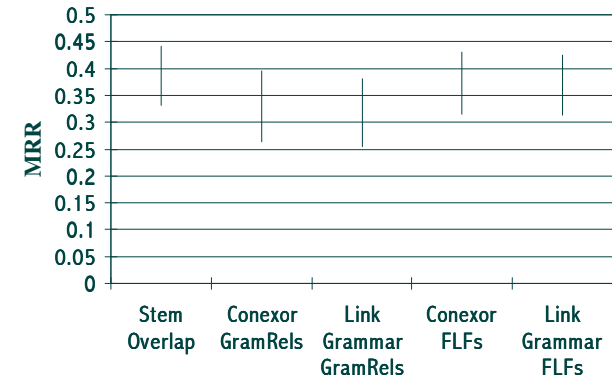
Best and Worst Cases

- The overlap scores in the GramRels and MLFs are very low
- The right answer is competing with other sentences

What does the Sears ad offer?

Rank	Sentence	Score	Overlap	Correct
1	1989 Remedia Publications, Comprehension/5W<92>s-4	1	compound_noun(v_x2,v_x3)	no
2	(North Redwood, Minn	1	compound_noun(v_x2,v_x3)	no
3	Not long ago, he ran an ad in some newspapers in small towns	1	object(ad,v_o7,[v_x7])	no
4	The ad showed a drawing of lovely furniture	1	object(ad,v_o2,[v_x2])	no
5	The ad said the furniture was for sale	1	object(ad,v_o2,[v_x2])	yes

Results (average of all levels)



Combining Scores

- Weight on Grammatical Relations
Score = 27 x GRO + 9 x FO + 3 x DO + WO
- Weight on Flat Logical Forms
Score = 27 x FO + 9 x GRO + 3 x DO + WO

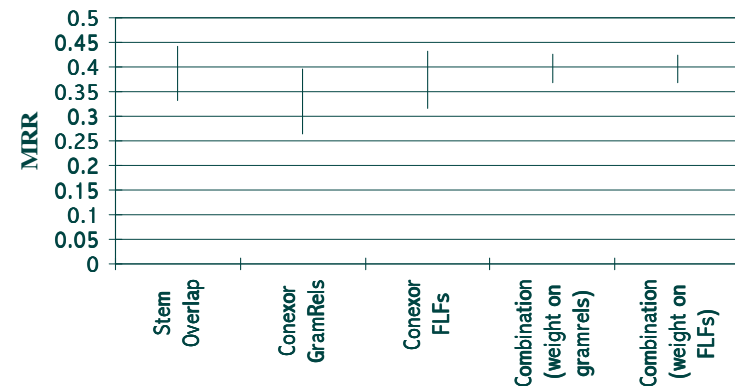
WO = word form overlap

DO = dependency overlap (using the output of Conexor FDG)

FO = overlap of FLFs

GRO = overlap of GramRels

New Results



Conclusions

- Overlap of FLFs is better than overlap of GramRels
- The impact of the parser for QA is small, but...
- Overlap of word forms is better!
- Questions
 - The Reading Comprehension corpus is too simple?
 - The parser and semantic interpreter introduces too many errors?
 - More complex overlap measures are perhaps better?
 - Logical forms are not required for simple (fact-based) questions?

Epilogue – Using Another Corpus

- TREC 2003 Passage Task
 - Document set: 3Gb of newswire text
 - Factoid questions (used in TREC 2002)
 - Retrieve a 250-word passage containing the answer

- Results:

<i>Run</i>	<i>Formula</i>	<i>Without</i>	<i>With NE</i>
answfind1	3wo+gro	16.8%	19.1%
answfind2	9wo+3gro+fo	16.8%	19.3%
answfind3	9wo+3fo+gro	15.6%	18.4%
	fo	10.8%	13.0%
	wo	14.8%	16.4%
	gro	9.0%	12.4%

QA with Grammatical Relations – In Reality

- *A man named Richard Sears has been playing a joke on shoppers.*
(xmod _ man name), (detmod _ man a),
(subj name man), (dobj name richard _),
(detmod _ joke a), (subj sear man _), (subj play sear _),
(aux _ play have), (aux _ play be), (ncmod _ play on),
(xcomp _ play joke)
- *Who played a joke on shoppers?*
(subj play who _), (dobj play joke _), (ncmod _ play on),
(detmod _ joke a)

QA with Flat Logical Forms – In Reality

- *A man named Richard Sears has been playing a joke on shoppers.*
holds(v_o10), object('man',v_o2,[v_x2]),
evt('name',v_e3,[v_x3,v_x4,v_x2]),
object('joke_on',v_o10,[v_e5,v_x12]),
object('richard',v_o4,[v_x4]), evt('sear',v_e5,[v_x2]),
object('shopper',v_o12,[v_x12])
- *Who played a joke on shoppers?*
holds(v_e2), object('who',v_o1,[v_x1]),
evt('play_on',v_e2,[v_x1,v_x4,v_x6]),
object('joke',v_o4,[v_x4]), object('shopper',v_o6,[v_x6])