AnswerFinder – Question Answering by Combining Lexical, Syntactic and Semantic Information

Diego Mollá-Aliod
Mary Gardiner
8 December 2004

Outline

• Question Answering and TREC
• AnswerFinder
• Grammatical Relations
• Flat Logical Form Patterns
• Lessons Learnt and What’s Next

TREC QA 2004

• Factoid, lists, definitions
  – 231 factoid questions
  – 62 list questions
  – 66 “other”
• Questions grouped in topics
  – handling of context
• Combined evaluation
  – ½ factoid
  – ¼ list
  – ¼ “other”

Target 2: “Fred Durst”

Q 2.1 FACTOID:
What is the name of Durst’s group

Q 2.2 FACTOID:
What record company is he with?

Q 2.3 LIST:
What are titles of the group’s releases?

Q 2.4 FACTOID:
Where was Durst born?

Q 2.5 OTHER:
Other
**AnswerFinder**

**Outline**

- Question Answering and TREC
- AnswerFinder
  - Grammatical Relations
  - Flat Logical Form Patterns
  - Lessons Learnt and What’s Next

**Grammatical Relations**

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONJ(type,head+)</td>
<td>Conjunction</td>
</tr>
<tr>
<td>MOD(type,head,dependent)</td>
<td>Modifier</td>
</tr>
<tr>
<td>CMOD(type,head,dependent)</td>
<td>Clausal modifier</td>
</tr>
<tr>
<td>NCMOD(type,head,dependent)</td>
<td>Non-clausal modifier</td>
</tr>
<tr>
<td>DETMOD(type,head,dependent)</td>
<td>Determiner</td>
</tr>
<tr>
<td>SUBI(head,dependent,initial gr)</td>
<td>Subject</td>
</tr>
<tr>
<td>OBI(head,dependent,initial gr)</td>
<td>Object</td>
</tr>
<tr>
<td>DOBI(head,dependent,initial gr)</td>
<td>Direct object</td>
</tr>
<tr>
<td>XCOMP(head,dependent)</td>
<td>Clausal complement without an overt subject</td>
</tr>
</tbody>
</table>

**Grammatical Relations**

![Diagram of grammatical relations](image)
**Q:** How far is it from Mars to Earth?
(ncmod how be far)
(subj be it _)
(ncmod from be mars)
(ncmod to be earth)

**A:** It is 416 million miles from Mars to Earth.
(subj be it _)
(obj be mile)
(ncmod _ million 416)
(ncmod _ mile million)
(ncmod from be mars)
(ncmod to be earth)

---

**Q:** What is the Population of Iceland?
(subj be what _)
(obj be population _)
(ncmod of population iceland)
(detmod _ population the)

**A:** Iceland has a population of 270000.
(subj have iceland _)
(obj have population _)
(detmod _ population a)
(ncmod of population 270000)

---

**Outline**

- Question Answering and TREC
- AnswerFinder
- Grammatical Relations
  - Flat Logical Form Patterns
- Lessons Learnt and What’s Next
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

Logical Form Overlap — LFO

Q: What is the population of Iceland?

object(iceland, O6, [X6])
object(population, O4, [X1])
object(what, O1, [X1])
prop(of, P5, [X1, X6])

A: Iceland has a population of 270000

dep(270000, d6, [x6])
object(population,o4,[x4])
object(iceland,o1,[x1])
evt(have,e2,[x1,x4])
prop(of,p5,[x4,x6])

Unification:

O4 = o4
X1 = x4
P5 = p5
X6 = x6

Flat Logical Form Patterns

What is X of Y?

object(ObjY,VobjY,[VeY]),
object(what,_,[VeWHAT]),
object(ObjX,VobjX,[VeWHAT]),
prop(of,_,[VeWHAT,VeY])

Y has a X of ANSWER

dep(ANSWER,ANSW,[VeANSW]),
prop(of,_,[VeX,VeANSW]),
object(ObjY,VobjY,[VeY]),
evt(have,_,[X6,X1]),
object(ObjX,VobjX,[VeWHAT])

QA with FLF Patterns

Q: What is the population of Iceland?

dep(ANSWER,ANSW,[VeANSW])
prop(of,_,[VeY,VeANSW])
object(iceland,O6,[X6])
evt(have,_,[X6,X1])
object(population,O4,[VeY])

A: Iceland has a population of 270000

dep(270000, d6, [x6])
object(population,o4,[x4])
object(iceland,o1,[x1])
evt(have,e2,[x1,x4])
prop(of,p5,[x4,x6])
Sentence Rescoring

- 3gro+lfo
  3 times the grammatical relation overlap score added to the flat logical form pattern overlap score.
- lfo
  The flat logical form pattern overlap score.

Exact Answer Extraction

- Use information from the FLF patterns and the named entities
  - Use the score of the sentence an answer candidate is found in
  - Double the score if from an FLF pattern and NE
  - Merge (add up) scores of repeated answers

Outline

- Question Answering and TREC
- AnswerFinder
- Grammatical Relations
- Flat Logical Form Patterns
- Lessons Learnt and What’s Next

TREC Evaluation

- Our preliminary evaluation:
  - Correct top-ranking sentence: 20%
  - Correct and exact answer: 5%
- TREC evaluation:
  - Correct and exact answer: 10%
  - List F-score: 0.08
  - “Other” F-score: 0.09
Lessons Learnt

- The development of FLF patterns is time-consuming
  - Difficult for humans to understand logical forms
  - Consider methods to help development of patterns
  - Learning methods
- The similarity measures are still very crude
  - Look at the structure inside GRs and LF terms

To Do

- Refine sentence scoring
  - Assign weights to the FLF terms
  - Use graph-based comparison
- Improve exact answer scoring
  - Development of FLF patterns
  - Machine learning of FLF patterns
- Enhance readability of FLFs
  - Use of graph structures