Solving Logic Puzzles: From Robust Processing to Precise Semantics

Iddo Lev, Bill MacCartney,
Christopher D. Manning,
and Roger Levy

Purpose

- Build a system that brings together Natural Language Processing and Automated Reasoning.
- Accomplish this by building a system that solves logic puzzles from the LSAT and GRE

Why Logic Puzzles?

- For humans, understanding the language of a logic puzzle is simple, the puzzle hard, for machines, the reverse. Research can focus on the NLP problem, rather than puzzle solving, because of this.
- Texts employ everyday language.
- Answers to puzzle questions never appear explicitly in the text. This means that they must be logically inferred. Superficial analysis methods cannot replace a deep understanding of the question.
- The task has a clear evaluation metric. Was the question correctly answered?
- Despite the fact that this is not a real world application, it is felt that methods developed here could be used in real world applications, for example, an office assistant application.

Example of a Logic Puzzle

Preamble: Six sculptures – C, D, E, F, G, and H – are to be exhibited in rooms 1, 2, and 3 of an art gallery. The exhibition conforms to the following conditions:

(1) Sculptures C and E may not be exhibited in the same room.
(2) Sculptures D and G must be exhibited in the same room.
(3) If sculptures E and F are exhibited in the same room, no other sculpture may be exhibited in that room.
(4) At least one sculpture must be exhibited in each room, and no more than three sculptures may be exhibited in any room.

Question 1: If sculpture D is exhibited in room 3 and sculptures E and F are exhibited in room 1, which of the following may be true?

(A) Sculpture C is exhibited in room 1.
(B) No more than 2 sculptures are exhibited in room 3.
(C) Sculptures F and H are exhibited in the same room.
(D) Three sculptures are exhibited in room 2.
(E) Sculpture G is exhibited in room 2.

Question 2: If sculptures C and G are exhibited in room 1, which of the following may NOT be a complete list of the sculpture(s) exhibited in room 2?

(A) Sculpture D (B) Sculptures E and H (C)...

Adapted from (Weber, 1999).

Figure 1: Example of a Puzzle Text

System Overview

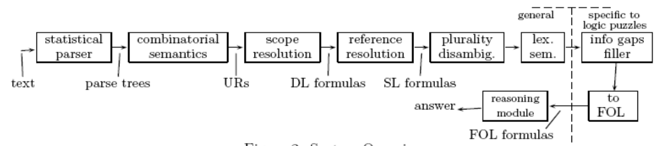

Figure 2: System Overview

Reasoning

- First-Order Logic
  - First-Order Logic offers an obvious way of solving logic puzzles using off-the-shelf FOL reasoners such as model builders and theorem provers.
- Semantic Logic
  - Semantic Logic is intended to be a general purpose semantic representation language.
  - Extends FOL
    - Includes event and group variables
    - Includes operators for "necessarily" and "possibly"
    - Includes Generalized Quantifiers
- Non-Determinism
  - Logic puzzles are designed to reduce ambiguity, such that there will only be one answer per question. Despite this, there will be many ambiguities in analysis. To accommodate this, each module ranks possible output representations, allowing for backtracking.

Logical Representations

First Order Logic

constraint (4) in Figure 1 is: $\forall x.room(x) \rightarrow \exists y.sculpture(y) \wedge exhibit(y,x)$. (The treatment of the modal 'must' is explained in §9.2).

Semantic Logic

$$\Box Q(\forall, x_1, room(x_1), Q(\geq_1, x_2, sculpture(x_2), \exists e.exhibit(e) \wedge subj(e, x_2) \wedge in(e, x_1)))$$

Challenges

- Combinatorial Semantics
- Scope Ambiguities
- Reference Resolution
- Plurality Disambiguation
- Lexical Semantics
- Information Gaps
- Presuppositions and Implicatures

Combinatorial Semantics

- Work in NLP has shifted from hand-built grammars to, more robust, statistical parsing.
- Grammars that involve precise semantics are still mostly hand-built.
- One goal is to make semantics more robust.
- Work starts with the compositional semantics framework (Blackburn and Bos, 2000; Bos, 2001)
- In rule-to-rule systems, syntactic rules have separate associated semantic rules.
- To deal with verbs and VPs in a more robust way, a new system is developed that only creates an event variable.
- NP modifiers are combined to the event using only generic roles.
- For the domain of puzzles, this treatment is appropriate because only the relationships between objects are important. Their semantic roles are not.

Information Gaps

- Natural language texts assume some knowledge implicitly.
  - Example, Figure 1. does not explicitly specify that a sculpture may be in only one room at a time.
- Filling in these gaps is difficult.
- Presuppositions
  - Pieces of information that are assumed in a sentence.
- Implicatures
  - Pieces of information that are suggested by saying or not saying something.
  - Two maxims from (Grice, 1989) dictate that a sentence should be consistent and informative.
  - Another says that the sentence should say only as much as is required. Therefore, saying “No more than three sculptures may be exhibited in any room” implies that there is a solution in which three sculptures are exhibited in one room.

Paul H. Grice. 1989. *Studies in the way of Words*. Harvard University Press.

Progress

- An end-to-end system has been built, using sculptures as a prototype.
- The system correctly solves all questions in this puzzle.
  - There is still no understanding of “complete list”, from question 2.
- The backend reasoning module is finished and works for any puzzle formalized in FOL + modals.
- The scope resolution module achieves accuracy of 94%, when trained on 259 two-quantifier sentences and tested over 46 unseen instances, extracted from puzzles.
- In 60% of unseen puzzles, the parser’s output is accurate enough to be used for computation of semantics. It has not yet been trained on puzzle text.
- To-date, the system has only worked end-to-end on one unseen puzzle.

Conclusions & Further Work

- Sufficiently robust and general methods for building semantic representations must be found in order to avoid hand-built translation rules for special phenomena.
- Work in the immediate future will focus on:
  - Coverage of syntax-to-semantics mappings
  - Incorporating classifiers for suggesting likely coreference resolutions and operator scopings
  - Developing methods for calculating presuppositions and inferences
- It is possible that it may be necessary to develop more robust models of syntactic to semantic transductions.