Binding Skolem Clauses in Theorem Prover Resolution for Automated Hypothetical Question Answering

RABIAH A.K., T. M. T. SEMBOK, HALIMAH B. Z.
Department of Information Sciences
Faculty of Information Sciences & Technology
National University of Malaysia
43600 UKM Bangi, Selangor
MALAYSIA

Abstract: - This paper described a new approach of adapting an existing theorem prover to the hypothetical question in generating an automated answer to a restricted domain such as reading comprehension. The adaptation of this theorem prover involves the modification of some components from our experiment such as: Knowledge Representation and Answer Extraction Agent. Question answering systems employing Skolemized Clauses Binding as the basic reasoning technique have been used to provide hypothetical answers to questions by considering a theorem to be proven as question. Hypothetical answer is an answer which comes from the text and logical thinking that is not explicitly stated in the text. It can be shown in general form $X \Rightarrow Y$, where X cannot be proven based on the information in the knowledge base. The notion of what constitutes an answer can be expanded so that the skolemized clauses binding, added as an intermediate approach generated while finding the hypothetical answer, may be regarded as answer. When there is not enough information in a knowledge base to provide a hypothetical answer, an approach such as answer generation requires an external resource similar to world knowledge in order to obtain relevant answer for the question given.

Key-Works: - Question Answering System, Natural Language Processing, and Theorem Prover.

1. Introduction

Question answering (QA) is a type of information retrieval. Given a collection of documents (such as the World Wide Web or a local collection), the system should be able to retrieve answers to questions posed in natural language. QA is regarded as requiring more complex natural language processing (NLP) techniques than other types of information retrieval such as document retrieval, and it is sometimes regarded as the next step beyond search engines.

QA research attempts to deal with a wide range of question types including: fact, list, definition, *How*, *Why*, hypothetical, semantically-constrained and cross-lingual questions. Search collections vary from small local document collections, to internal organization documents, to compiled newswire reports, to the world wide web. There are two types of QA:

- *Closed-domain* question answering deals with questions under a specific domain (for example, medicine or automotive maintenance), and can be seen as an easier task because NLP systems can exploit domain-specific knowledge frequently formalized in ontologies.

- *Open-domain* question answering deals with questions about nearly everything, and can only rely on general ontologies and world knowledge. On the other hand, these systems usually have much more data available from which to extract the answer.

This paper focuses on open-domain QA and it considers the description of strategies involved in answering the implied or hypothetical question in reading comprehension by integrating the syntactic, semantic and pragmatic of the passage domain and world knowledge for natural linguistic in computer science. The commonsense knowledge base evolved along with the story understanding system. Whenever a piece of commonsense knowledge becomes available in the story understanding system, it is then added to the database. The database can be expanded to be utilized by the story understanding application. The story understanding and question answering through sophisticated knowledge representation, reasoning and inferential processing, require extensive prior encoding of natural language. We propose Pragmatic Skolemized Clauses Representation generated based on existing First Order Logic (FOL).

2. Background

In the early days of artificial intelligence (AI) research, "question answering system" would have fallen squarely into mainstream. With the divergence of AI into a number of subareas, different aspects of question answering migrated into various niches within AI as well as into other areas of computer science entirely. The importance of question answering to human understanding is clear from the knowledge base to structured text. The importance of large knowledge base and database has been steadily expanding with the human knowledge. Traditionally, natural language question answering (QA) systems approach the frontier of the querying methodologies for textual documents or passage and database, overtaken by keyword.

Theoretical works in question answering which can be found in AI usually refer to computational linguistic, psychology, linguistics, and philosophy. Thagard(2006), called these interdisciplinary studies as cognitive science. Cognitive science has primarily worked with the computational-representational understanding of mind: we can understand human thinking by postulating mental representations akin to computational data structures and mental procedures akin to algorithms [1]. Research in these various areas forms the basis for implementing natural language question answering systems. Natural language question answering may be considered as the most universal way to provide information access. There are several natural language question answer systems with difference purpose such as START (SynTactic Analysis using Reversible Transformations) natural language system which was developed as information retrieval system in 1993, ALICE, a chat robot, and Deep Read, a reading comprehension prototype system in 1999.

Recently there has been a renewed interest in question answering for reading comprehension tests within AI community, due in part to the MITRE Corporation's initiative research lead by Hirschman in 1999. We believe that reading comprehension test can be a valuable state-of-the-art tool to assist in the natural language understanding. For that particular year, MITRE Corporation defined a new research paradigm for natural language processing by developing reading comprehension system such Deep Read [2] and followed by ABCs [3]. The roots of this research are found since the past several years such as QUARC system [4] and AQUAREAS system [5]. However, there are a number of researchers interested in the same problem that was initiated by Hirshman in 1999 such as Wang(2000), and Bashir(2004) [6,7]. Reading comprehension can be defined as the level of understanding of a passage or text. It involves a process of reading a story or passage to demonstrate one's understanding of the passage by answering questions about it.

The passage and the question in reading comprehension are treated as a single package with its own difficulty levels. The computer compute score and level of difficulty adjustments in the background during a passage's reading comprehension and question package. Reading comprehension offers a new challenge and a human-centric evaluation paradigm for human language technology [8]. Reading comprehension task is a valuable state-of-the-art tool to assist in natural language understanding [4]. Question answering is a reading comprehension task used to demonstrate the understanding of the system about the passage for the purpose of showing and building up the meaning representation and to enforce syntactic and semantic agreements. The questions are based on the content of a passage. The answer in reading comprehension only comes from the short story associated with the question given. However, there are two types of question based on what is stated or implied in the passage.

- *Stated question:* (to be set forth in words)
  Questions based on specific detail included in the passage. The answers can be located in one place and are explicitly stated in the text. These are typically "who", "what", "where" and "when" questions.

- *Implied question:* (to be engaged by logical necessity)
  Questions based on the story or plot of the passage. The answers come from the text and logical thinking that is not explicitly stated in the text. The information is relevant to the passage, but does not appear in it. These are the "why", "how" and "what do you think" question types.

3. Proposed Method

For the purpose of this discussion, work on open-ended question answering requires sophisticated linguistic analysis, including discourse understanding and deals with questions about nearly everything, and can only rely on general ontologies and world knowledge. The representation of questions and answers and reasoning mechanisms for question answering are the main concerns for researchers in knowledge representation and reasoning (KR&R). Formally, mathematical approaches to question answering based on logic and theorem-proving formed a subset of KR&R approaches [9].

To achieve a question answering system that is capable of generating the automatic answers for the hypothetical question given, we proposed to implement the world knowledge and bind the skolem clauses by its argument into existing theorem prover technique. In this paper, we evaluated our approach in 5 samples of grade two children stories from Remedia Publication. The questions from each passage are chosen to measure of how well the system can understand the story. Fig.1 shows an example of a passage and its hypothetical questions of *Why…?* .

Storybook Person Found Alive!

(ENGLAND, June, 1989) - Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh.

As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read.

Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm.

Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don't know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

Question:

Why did Chris write two books of his own?

Fig.1: Sample of Remedia Passage

Why question involves a more difficult query calls, hypothetical postulations, spatially or temporally constrained questions, dialog queries and badly-worded or ambiguous questions. These type of questions need deeper understanding of the passage. Complex or ambiguous document passages likewise need more NLP techniques applied to understand the text.

Statistical QA, which introduces statistical question processing and answer extraction modules, is also growing in popularity in the research community. Many of the lower-level NLP tools used, such as part-of-speech tagging, parsing, named-entity detection, sentence boundary detection, and document retrieval, are already available as probabilistic applications.

In our QA system, we proposed to employ binding skolem clauses as an inference technique that is used to provide implicit answer to the hypothetical questions given by considering a theorem to be proven as a question. Automated theorem proving served as an early model for question answering in the field of AI [9]. And, the idea of using inference for question-answering is not new. It can be found in QA systems since the 1970s for story understanding [10].

4. Hypothetical Question Answering

Different types of questions require the use of different strategies to find the answer. A semantic model of question understanding and processing is needed, one that will recognize equivalent questions, regardless of the words, syntactic inter-relations or idiomatic forms. Hypothetical postulations such as *Why* or *How* questions need these types of deeper understanding of the question.

Before a question can be answered, knowledge sources available must be known. If the answer to a question is not present in the data sources, no matter how well we perform question processing, retrieval and extraction of the answer, a correct result cannot be obtained. Answer extraction depends on the complexity of the question, on the answer type provided by question processing, on the actual data where the answer is searched, on the search method and on the question focus and context. Since answer processing depends on such a large number of factors, research for answer processing should be tackled with a lot of care and given special importance.

More sophisticated questioners expect answers which are outside the scope of written texts or structured databases. To upgrade a Q&A system with such capabilities, we need to integrate reasoning components operating on a variety of knowledge bases, encoding world knowledge and common-sense reasoning mechanisms as well as knowledge specific to a variety of domains. Hypothetical answers have received some attention since 1980s [11], in part due to their utilisation when reasoning took place in the absence of a Closed World Assumption, as in abductive reasoning.

We identified hypothetical answers with a set of skolemized clauses containing at least one skolem constant that represent an entity that is bound to other skolemized clauses. Written in semantic relation that ass represented in Pragmatic Skolemized Clauses, a hypothetical answer consisting of an answer literals must bind to each other in order for the accompanying answer to be considered as an answer.

5. Generating Answer with Skolemized Clauses Binding

This exploration of answer generation has been done using a skolemized clauses binding and theorem prover as a reasoning mechanism. Meanwhile, a semantic relation rule was being specified in pragmatic skolemized clauses as a knowledge representation. In the example provided, this can be seen as binding process proceeds. If the semantic relation rule being searched contains rules that are unified to a question through its skolem constant, hypothetical answers will be produced. Consider the following sample as a semantic relation rule used as an illustration that was originally based on children passage entitled “Storybook Person Found Alive!” from Remedia Publications.

Semantic relation rules:

- cl([two(g46)],[])
- cl([book(g46)],[])
- cl([own(his)],[])
- cl([writes(chris,g46)],[])
- cl([famous(g52)],[])
- cl([be(like(_37214 ^ isa(tells(g46,it),_37214)),g52)],[])

Given above simple semantic relation rules, and the question *Why did Chris write two books?*, the following answer are produced (each rule is binded by a skolem constant, *g46*).

~ cl([two(g46)],[]) # ~ cl([book(g46)],[]) # ~ writes(chris,g46))

g46 is unified with the semantic relation rules in knowledge based

~ two(g46) :- two(g46)

~ book(g46) :- book(g46)

~ writes(chris,g46) :- writes(chris,g46)

then, bind *g46* and *g52* to any relevant semantic relation rule to find the answer.

1. writes(chris,g46)
2. two(g46)
3. book(g46)
4. famous(g52)
5. be(like(_37214 ^ isa(tells(g46,it),_37214)),g52)

The skolemized clauses 1 to 5 are a collection of hypothetical answer set that are unified to the question given because each clause is bound with at least one skolem constant. The semantic relation rule base indicates that *g46 (two book)* is bound to clause 5, which contain skolem a constant *g52* that is bound to clause 4. Skolem constant g52 stands for *famous* predicate.

The example is perhaps better motivated by showing what happened when the fact *two book* is bound to other clauses or semantic relation rules. The resulting answer is:

1. famous(g52)

2. be(like(_37214 ^ isa(tells(g46,it),_37214)),g52)

Both the skolemized clauses are considered as a set of hypothetical answer that is relevant to the question, and they may be the best information available. Another examples are shown in Table 1. Each example begins with part of a collection of semantic rules in knowledge base, represented in skolemized clauses. In this research, a question *Q* is represented as a proposition, and a traditional proof initiated by adding the negation of the clause form of *Q* to a consistent knowledge base *K*. If an inconsistency is unified, then skolemized clauses binding process proceed to find the relevant answer.

Table 1: An Examples of Question Answering Process

1st Example	2nd Example
Semantic relation rules (*K*)	
cl([now(g1)],[]) cl([new(f16)],[]) cl([faster(f16)],[]) cl([way(f16)],[]), cl([sents(g1,f16)],[]) cl([now(g1)],[]) cl([end(r(pony & express),g1)],[])	cl([pledge(f25)],[]), cl([young(g37)],[]), cl([people(g37)],[]), cl([proud(g38)],[]), cl([feels(g37,g38)],[]), cl([makes(f25,g37)],[]), cl([writes(r(frances & bellamy),f25)],[])
Proposition (*Q*)	
~ end(r(pony & express),g1) # answer(g1)	~ pledge(f25) # ~ writes(r(frances & bellamy),f25) # answer(f25)
Unifying process	
~ end(r(pony & express),g1) :- end(r(pony & express),g1)	~ pledge(f25) :- pledge(f25) ~ writes(r(frances & bellamy),f25) :- writes(r(frances & bellamy),f25)
Relevant answer (after skolemized clauses binding process)	
now(g1) mail(g1) new(f1) faster(f1) way(f1)] sents(g1,f1)	makes(f25,g37) young(g37) people(g37) feels(g37,g38) proud(g38)

6. Relevant Answer

A relevant answer to a particular question can be generally defined as an answer that implies all clauses to that question. Relevance for answers has been defined as unifying the skolem constant by the question.

In a rule base consisting solely of skolem constants, the unifying of a single skolem constant to a question would be considered a relevant answer. When rules are added, the experiment becomes more complicated. When taxonomic relationship is represented in a rule base, a relevant answer can be defined as an interconnection of all clauses that unify and bind the same skolem constants. Thus, in first example in Table 1, *g1* is considered as a skolem constant to be unified to a skolemized clause in knowledge base, ~ *end(r(pony & express),g1) :- end(r(pony & express),g1)*. Then *g1* binds to any skolemized clauses consisting the same skolem constant, and tracks all possible skolemized clauses in knowledge base by binding skolem constant exists, *f1*, until all skolem constants binding are complete, *sents(g1,f1); now(g1); mail(g1); new(f1); faster(f1); and way(f1)]*.

Providing information in a form of pragmatic skolemized clauses is just a method to collect the keywords of relevant answers. The issues relates to the problem of providing an answer in correct English phrases can be considered another important area of research in question answering. In this research this problem has been considered, but thus far it has taken the form of observations rather than formal theories. This represents an area of further research interest.

7. Conclusion

As a conclusion, we insist on the interest of our approach in open-domain question answering addressed to the semantic relations rule base. Such approach is interesting when we attempted to provide answers which do not depend on a particular state of fact base.

The algorithm presented is based on a theoretical study, and provides all the answers, under the assumptions that were made. As already mentioned, this work also involved in making an additional algorithm to theorem prover so that the hypothetical answers in the absence of constraining information are considered. As the dominant paradigm in theorem proving remains that of providing only extensional answers, it is easy to enhance a theorem prover by applying additional method to produce valuable information in the form of hypothetical answer.

The limitation of the current implementation is the limited scope of the syntactic category, which causes answers to be produced in incomplete English phrases. We planned to extend our syntactic category by taking into consideration the adverb and preposition categories. Reading comprehension passages heavily use various form of co-reference, such as anaphora, definite description, etc. In this paper, we intend to adapt the world knowledge for this purpose, but a general solution to resolving the issues is considered a topic for ongoing research.

References:

[1] Thagard, P. Computation and the Philosophy of Science. *http://cogsci.uwaterloo.ca/Articles/Pages/comp.phil.sci.html* (June, 30, 2006), 2006.

[2] Hirschman, L., Light, M., Breck, E., & Burger, J.D. Deep Read: A Reading Comprehension System. *Proc. of the 37th Annual Meeting of the Association for Computational Linguistics*, College Park, Maryland, 1999.

[3] Ferro, L., Greiff, W., Hirschman, L., & Wellner, B. Reading, Learning, Teaching. *http://www.mitre.org/news/events* (April 02, 2004).

[4] Riloff, E., and Thelen, M. A Rule-based Question Answering System for Reading Comprehension Tests, *Proceeding of ANLP/NAACL 2000 Workshop on Reading Comprehension Tests as Evaluation for Computer-Based Language Understanding Systems*, 2000, pp. 13 - 19.

[5] Ng, H. T., Teo, L.H., & Kwan, J.L.P. A Machine Learning Approach to Answering Questions for Reading Comprehension Tests. *Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora* (EMNLP/VLC-2000), 2000.

[6] Wang, J., Auer, J., Parasumaran, R., Zubarev, I., Brandyberry, D., & Harper, M.P. A Question Answering System Development, *Project in a Natural Language Processing Course*, 2000.

[7] Bashir, A., Kantor, A., Ovesdotter C. A., Ripoche, G., Le, Q., & Atwell, S. Story Comprehension. *http://12r.cs.uiuc.edu/~danr/Teaching/CS598-04/Projects/First/Term2.pdf*. (April 11, 2005), 2004

[8] Hirschman, L. Reading Comprehension: Reading, Learning, *http://www.mitre.org/news/events/tech03/briefings/human_language/hirschman.pdf*, (April 2004), 2003.

[9] Burhans, D. T. A Question Answering Interpretation of Resolution Refutation, *PhD. Dissertation*, Faculty of the Graduate School. State University of New York, Buffalo, 2002.

[10] Lehnert, W.G. *The Process of Question Answering*. Lawrence Erlbaum. 1978.

[11] Burhans, D. T., & Shapiro S.C. Abduction and Question Answering, *Proceeding of the IJCAI Workshop on Abduction Reasoning, UJCAI-01*, 2001.