

FERRET: Interactive Question-Answering for Real-World Environments

Andrew Hickl, Patrick Wang, John Lehmann, and Sanda Harabagiu

Language Computer Corporation
1701 North Collins Boulevard
Richardson, Texas 75080 USA
ferret@languagecomputer.com

Abstract

This paper describes FERRET, an interactive question-answering (Q/A) system designed to address the challenges of integrating automatic Q/A applications into real-world environments. FERRET utilizes a novel approach to Q/A – known as *predictive questioning* – which attempts to identify the questions (and answers) that users need by analyzing how a user interacts with a system while gathering information related to a particular scenario.

1 Introduction

As the accuracy of today’s best factoid question-answering (Q/A) systems (Harabagiu et al., 2005; Sun et al., 2005) approaches 70%, research has begun to address the challenges of integrating automatic Q/A systems into real-world environments. A new class of applications – known as interactive Q/A systems – are now being developed which allow users to ask questions in the context of extended dialogues in order to gather information related to any number of complex scenarios. In this paper, we describe our interactive Q/A system – known as FERRET – which uses an approach based on *predictive questioning* in order to meet the changing information needs of users over the course of a Q/A dialogue.

Answering questions in an interactive setting poses three new types of challenges for traditional Q/A systems. First, since current Q/A systems are designed to answer single questions in isolation, interactive Q/A systems must look for ways to foster interaction with a user throughout all phases of the research process. Unlike traditional Q/A applications, interactive Q/A systems must do more

than cooperatively answer a user’s single question. Instead, in order to keep a user collaborating with the system, interactive Q/A systems need to provide access to new types of information that are somehow relevant to the user’s stated – and unstated – information needs.

Second, we have found that users of Q/A systems in real-world settings often ask questions that are much more complex than the types of factoid questions that have been evaluated in the annual Text Retrieval Conference (TREC) evaluations. When faced with a limited period of time to gather information, even experienced users of Q/A may find it difficult to translate their information needs into the simpler types of questions that Q/A systems can answer. In order to provide effective answers to these questions, interactive question-answering systems need to include *question decomposition* techniques that can break down complex questions into the types of simpler factoid-like questions that traditional Q/A systems were designed to answer.

Finally, interactive Q/A systems must be sensitive not only to the content of a user’s question – but also to the context that it is asked in. Like other types of task-oriented dialogue systems, interactive Q/A systems need to model both what a user knows – and what a user wants to know – over the course of a Q/A dialogue: systems that fail to represent a user’s knowledge base run the risk of returning redundant information, while systems that do not model a user’s intentions can end up returning irrelevant information.

In the rest of this paper, we discuss how the FERRET interactive Q/A system currently addresses the first two of these three challenges.

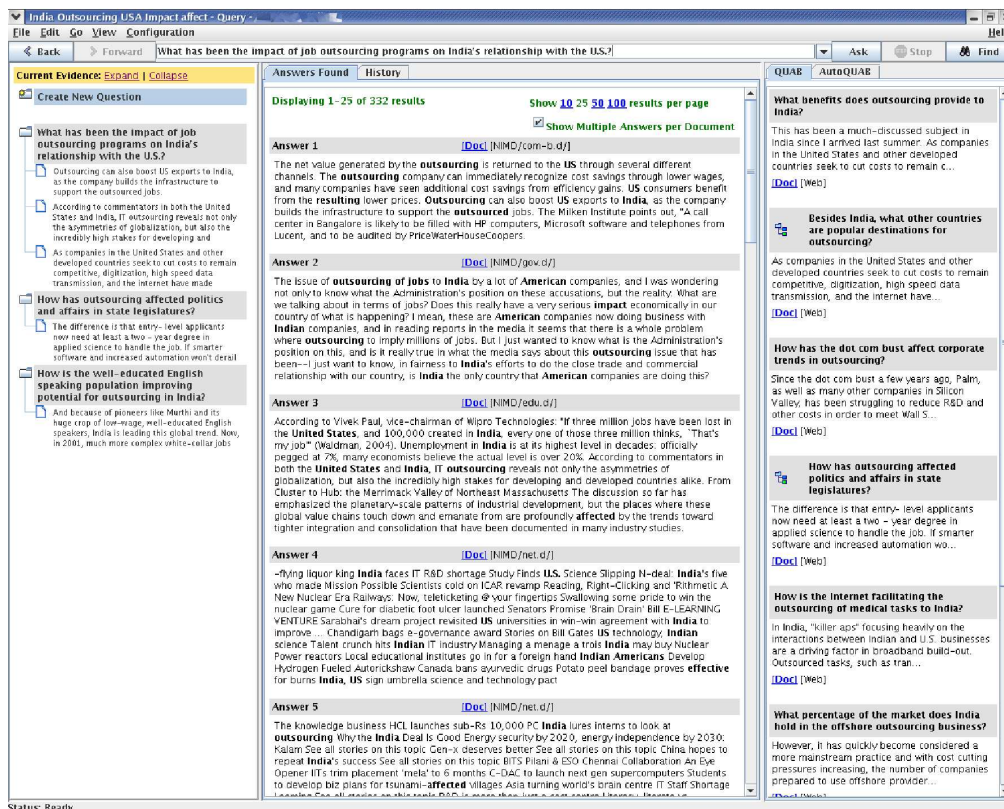


Figure 1: The FERRET Interactive Q/A System

2 The FERRET Interactive Question-Answering System

This section provides a basic overview of the functionality provided by the FERRET interactive Q/A system.¹

FERRET returns three types of information in response to a user's query. First, FERRET utilizes an automatic Q/A system to find answers to users' questions in a document collection. In order to provide users with the timely results that they expect from information gathering applications (such as Internet search engines), every effort was made to reduce the time FERRET takes to extract answers from text. (In the current version of the system, answers are returned on average in 12.78 seconds.²)

In addition to answers, FERRET also provides information in the form of two different types of *predictive question-answer pairs* (or QUABs). With FERRET, users can select from QUABs that

¹For more details on FERRET's question-answering capabilities, the reader is invited to consult (Harabagiu et al., 2005a); for more information on FERRET's predictive question generation component, please see (Harabagiu et al., 2005b).

²This test was run on a machine with a Pentium 4 3.0 GHz processor with 2 GB of RAM.

were either generated automatically from the set of documents returned by the Q/A system or that were selected from a large database of more than 10,000 question-answer pairs created offline by human annotators. In the current version of FERRET, the top 10 automatically-generated and hand-crafted QUABs that are most judged relevant to the user's original question are returned to the user as potential continuations of the dialogue. Each set of QUABs is presented in a separate pane found to the right of the answers returned by the Q/A system; QUABs are ranked in order of relevance to the user's original query.

Figure 1 provides a screen shot of FERRET's interface. Q/A answers are presented in the center pane of the FERRET browser, while QUAB question-answer pairs are presented in two separate tabs found in the rightmost pane of the browser. FERRET's leftmost pane includes a "drag-and-drop" clipboard which facilitates note-taking and annotation over the course of an interactive Q/A dialogue.

3 Predictive Question-Answering

First introduced in (Harabagiu et al., 2005b), a *predictive questioning* approach to automatic

question-answering assumes that Q/A systems can use the set of documents relevant to a user’s query in order to generate sets of questions – known as *predictive questions* – that anticipate a user’s information needs. Under this approach, topic representations like those introduced in (Lin and Hovy, 2000) and (Harabagiu, 2004) are used to identify a set of text passages that are relevant to a user’s domain of interest. Topic-relevant passages are then semantically parsed (using a PropBank-style semantic parser) and submitted to a question generation module, which uses a set of syntactic rewrite rules in order to create natural language questions from the original passage.

Generated questions are then assembled into question-answer pairs – known as QUABs – with the original passage serving as the question’s “answer”, and are then returned to the user. For example, two of the predictive question-answer pairs generated from the documents returned for question Q_0 , “*What has been the impact of job outsourcing programs on India’s relationship with the U.S.?*”, are presented in Table 1.

Q_0	What has been the impact of job outsourcing programs on India’s relationship with the U.S.?
PQ_1	How could India respond to U.S. efforts to limit job outsourcing?
A_1	U.S. officials have countered that the best way for India to counter U.S. efforts to limit job outsourcing is to further liberalize its markets.
PQ_2	What benefits does outsourcing provide to India?
A_2	India’s prowess in outsourcing is no longer the only reason why outsourcing to India is an attractive option. The difference lies in the scalability of major Indian vendors, their strong focus on quality and their experience delivering a wide range of services”, says John Blanco, senior vice president at Cablevision Systems Corp. in Bethpage, N.Y.
PQ_2	Besides India, what other countries are popular destinations for outsourcing?
A_2	A number of countries are now beginning to position themselves as outsourcing centers including China, Russia, Malaysia, the Philippines, South Africa and several countries in Eastern Europe.

Table 1: Predictive Question-Answer Pairs

While neither PQ_1 nor PQ_2 provide users with an exact answer to the original question Q_0 , both QUABs can be seen as providing users information which is complementary to acquiring information on the topic of *job outsourcing*: PQ_1 provides details on how India could respond to anti-outsourcing legislation, while PQ_2 talks about other countries that are likely targets for outsourcing.

We believe that QUABs can play an important role in fostering extended dialogue-like interactions with users. We have observed that the incorporation of predictive-question answer pairs into an interactive question-answering system like FERRET can promote dialogue-like interactions

between users and the system. When presented with a set of QUAB questions, users typically selected a coherent set of follow-on questions which served to elaborate or clarify their initial question. The dialogue fragment in Table 2 provides an example of the kinds of dialogues that users can generate by interacting with the predictive questions that FERRET generates.

User Q_1 :	What has been the impact of job outsourcing programs on India’s relationship with the U.S.?
QUAB $_1$:	How could India respond to U.S. efforts to limit job outsourcing?
QUAB $_2$:	Besides India, what other countries are popular destinations for outsourcing?
User Q_2 :	What industries are outsourcing jobs to India?
QUAB $_3$:	Which U.S. technology companies have opened customer service departments in India?
QUAB $_4$:	Will Dell follow through on outsourcing technical support jobs to India?
QUAB $_5$:	Why do U.S. companies find India an attractive destination for outsourcing?
User Q_3 :	What anti-outsourcing legislation has been considered in the U.S.?
QUAB $_6$:	Which Indiana legislator introduced a bill that would make it illegal to outsource Indiana jobs?
QUAB $_7$:	What U.S. Senators have come out against anti-outsourcing legislation?

Table 2: Dialogue Fragment

In experiments with human users of FERRET, we have found that QUAB pairs enhanced the quality of information retrieved that users were able to retrieve during a dialogue with the system.³ In 100 user dialogues with FERRET, users clicked hyperlinks associated with QUAB pairs 56.7% of the time, despite the fact the system returned (on average) approximately 20 times more answers than QUAB pairs. Users also derived value from information contained in QUAB pairs: reports written by users who had access to QUABs while gathering information were judged to be significantly ($p < 0.05$) better than those reports written by users who only had access to FERRET’s Q/A system alone.

4 Answering Complex Questions

As was mentioned in Section 2, FERRET uses a special dialogue-optimized version of an automatic question-answering system in order to find high-precision answers to users’ questions in a document collection.

During a Q/A dialogue, users of interactive Q/A systems frequently ask complex questions that must be decomposed syntactically and semantically before they can be answered using traditional Q/A techniques. Complex questions submitted to

³For details of user experiments with FERRET, please see (Harabagiu et al., 2005b).

FERRET are first subject to a set of syntactic decomposition heuristics which seek to extract each overtly-mentioned subquestion from the original question. Under this approach, questions featuring coordinated question stems, entities, verb phrases, or clauses are split into their separate conjuncts; answers to each syntactically decomposed question are presented separately to the user. Table 3 provides an example of syntactic decomposition performed in FERRET.

CQ ₁	What industries have been outsourcing or offshoring jobs to India or Malaysia?
QD ₁	What industries have been outsourcing jobs to India?
QD ₂	What industries have been offshoring jobs to India?
QD ₃	What industries have been outsourcing jobs to Malaysia?
QD ₄	What industries have been offshoring jobs to Malaysia?

Table 3: Syntactic Decomposition

FERRET also performs semantic decomposition of complex questions using techniques first outlined in (Harabagiu et al., 2006). Under this approach, three types of semantic and pragmatic information are identified in complex questions: (1) information associated with a complex question’s expected answer type, (2) semantic dependencies derived from predicate-argument structures discovered in the question, and (3) and topic information derived from documents retrieved using the keywords contained the question. Examples of the types of automatic semantic decomposition that is performed in FERRET is presented in Table 4.

CQ ₂	What has been the impact of job outsourcing programs on India’s relationship with the U.S.?
QD ₅	What is meant by India’s relationship with the U.S.?
QD ₆	What outsourcing programs involve India and the U.S.?
QD ₇	Who has started outsourcing programs for India and the U.S.?
QD ₈	What statements were made regarding outsourcing on India’s relationship with the U.S.?

Table 4: Semantic Question Decomposition

Complex questions are decomposed by a procedure that operates on a Markov chain, by following a random walk on a bipartite graph of question decompositions and relations relevant to the topic of the question. Unlike with syntactic decomposition, FERRET combines answers from semantically decomposed question automatically and presents users with a single set of answers that represents the contributions of each question. Users are notified that semantic decomposition has occurred, however; decomposed questions are displayed to the user upon request.

In addition to techniques for answering complex questions, FERRET’s Q/A system improves performance for a variety of question types by employing separate question processing strategies in

order to provide answers to four different types of questions, including factoid questions, list questions, relationship questions, and definition questions.

5 Conclusions

We created FERRET as part of a larger effort designed to address the challenges of integrating automatic question-answering systems into real-world research environments. We have focused on two components that have been implemented into the latest version of FERRET: (1) predictive questioning, which enables systems to provide users with question-answer pairs that may anticipate their information needs, and (2) question decomposition, which serves to break down complex questions into sets of conceptually-simpler questions that Q/A systems can answer successfully.

6 Acknowledgments

This material is based upon work funded in whole or in part by the U.S. Government and any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the U.S. Government.

References

- S. Harabagiu, D. Moldovan, C. Clark, M. Bowden, A. Hickl, and P. Wang. 2005a. Employing Two Question Answering Systems in TREC 2005. In *Proceedings of the Fourteenth Text REtrieval Conference*.
- Sanda Harabagiu, Andrew Hickl, John Lehmann, and Dan Moldovan. 2005b. Experiments with Interactive Question-Answering. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL’05)*.
- Sanda Harabagiu, Finley Lacatusu, and Andrew Hickl. 2006. Answering complex questions with random walk models. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Seattle, WA.
- Sanda Harabagiu. 2004. Incremental Topic Representations. In *Proceedings of the 20th COLING Conference*, Geneva, Switzerland.
- Chin-Yew Lin and Eduard Hovy. 2000. The automated acquisition of topic signatures for text summarization. In *Proceedings of the 18th COLING Conference*, Saarbrücken, Germany.
- R. Sun, J. Jiang, Y. F. Tan, H. Cui, T.-S. Chua, and M.-Y. Kan. 2005. Using Syntactic and Semantic Relation Analysis in Question Answering. In *Proceedings of The Fourteenth Text REtrieval Conference (TREC 2005)*.