Proceedings of HLT-NAACL 2003
Demonstrations , pp. 19-20
Edmonton, May-June 2003

JAVELIN: A Flexible, Planner-Based Architecture for Question Answering

Eric Nyberg
Language Technologies Institute
Carnegie Mellon University
ehn@cs.cmu.edu

Robert Frederking
Language Technologies Institute
Carnegie Mellon University
ref@cs.cmu.edu

Abstract

The JAVELIN system integrates a flexible, planning-based architecture with a variety of language processing modules to provide an open-domain question answering capability on free text. The demonstration will focus on how JAVELIN processes questions and retrieves the most likely answer candidates from the given text corpus. The operation of the system will be explained in depth through browsing the repository of data objects created by the system during each question answering session.

1 Introduction

Simple factoid questions can now be answered reasonably well using pattern matching. Some systems (Soubbotin and Soubbotin, 2002) use surface patterns enhanced with semantic categories and question types in order to model the likelihood of answers given the question. Furthermore, Hovy et al. (Hovy et al., 2002) have obtained good results using only surface patterns pre-extracted from the web. However, pattern-based approaches don't represent the meaning of the patterns they use, and it is not clear whether they can be generalized for more difficult, non-factoid questions.

Open domain question answering is a complex, multi-faceted task, where question type, information availability, user needs, and a combination of text processing techniques (statistical, NLP, etc.) must be combined dynamically to determine the optimal answer. For more complex questions, a more flexible and powerful control mechanism is required. For example, LCC (D. Moldovan and Surdeanu, 2002) has implemented feedback loops which ensure that processing constraints are met by retrieving more documents or expanding question terms. The LCC system includes a passage retrieval loop, a lexico-semantic loop and a logic proving loop. The IBM PIQUANT system (Carroll et al., 2002) combines knowledge-based agents using predictive annotation with a statistical approach based on a maximum entropy model (Ittycheriah et al., 2001).

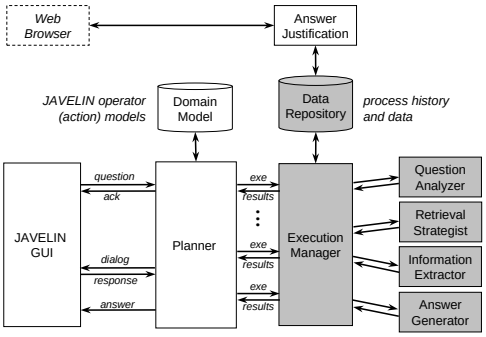

Figure 1: The JAVELIN architecture. The Planner controls execution of the individual components via the Execution Manager.

Both the LCC and IBM systems represent a departure from the standard pipelined approach to QA architecture, and both work well for straightforward factoid questions. Nevertheless, both approaches incorporate a pre-determined set of processing steps or strategies, and have limited ability to reason about new types of questions not previously encountered. Practically useful question answering in non-factoid domains (e.g., intelligence analysis) requires more sophisticated question decomposition, reasoning, and answer synthesis. For these *hard* questions, QA architectures must define relationships among entities, gather information from multiple sources, and reason over the data to produce an effective answer. As QA functionality becomes more sophisticated, the set of decisions made by a system will not be captured by pipelined architectures or multi-pass constraint relaxation, but must be modeled as a step-by-step decision flow, where the set of processing steps is determined at run time for each question.

This demonstration illustrates the JAVELIN QA architecture (Nyberg et al., 2002), which includes a general, modular infrastructure controlled by a step-by-step planning component. JAVELIN combines analysis modules, information sources, user discourse and answer synthesis as required for each question-answering interaction. JAVELIN also incorporates a global memory, or *repos-*

itory, which maintains a linked set of object dependencies for each question answering session. The repository can be used to provide a processing summary or *answer justification* for the user. The repository also provides a straightforward way to compare the results of different versions of individual processing modules running on the same question. The modularity and flexibility of the architecture provide a good platform for component-based (glass box) evaluation (Nyberg and Mitamura, 2002).

2 Demonstration Outline

The demonstration will be conducted on a laptop connected to the Internet. The demonstration will feature the JAVELIN graphical user interface (a Java application running on the laptop) and the JAVELIN Repository (the central database of JAVELIN result objects, accessed via a web browser). A variety of questions will be asked of the system, and the audience will be able to view the system's answers along with a detailed trace of the steps that were taken to retrieve the answers.

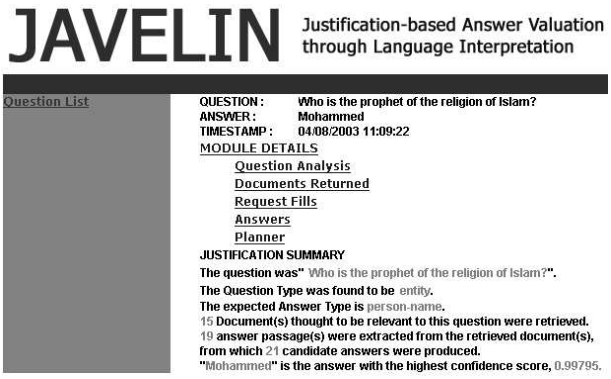

Figure 2: An Answer Justification.

Figure 2 shows the top-level result returned by JAVELIN. The preliminary answer justification includes the selected answer along with a variety of hyperlinks that can be clicked to provide additional detail regarding the system's analysis of the question, the documents retrieved, the passages extracted, and the full set of answer candidates. The justification also provides drill-down access to the steps taken by the Planner module in reasoning about how to best answer the given question. Figure 3 shows additional detail that is exposed when the "Documents Returned" and "Request Fills" links are activated.

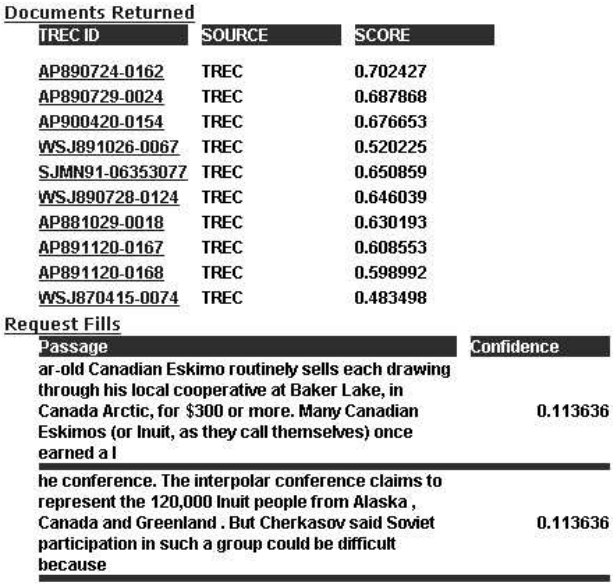

Figure 3: Partial Answer Detail.

Acknowledgements

The research described in this paper was supported in part by a grant from ARDA under the AQUAINT Program Phase I. The current version of the JAVELIN system was conceived, designed and constructed with past and current members of the JAVELIN team at CMU, including: Jamie Callan, Jaime Carbonell, Teruko Mitamura, Kevyn Collins-Thompson, Krzysztof Czuba, Michael Duggan, Laurie Hiyakumoto, Ning Hu, Yifen Huang, Curtis Huttenhower, Scott Judy, Jeongwoo Ko, Anna Kupść, Lucian Lita, Stephen Murtagh, Vasco Pedro, David Svoboda, and Benjamin Van Durme.

References

J. Carroll, J. Prager, C. Welty, K. Czuba, and D. Ferrucci. 2002. A multi-strategy and multi-source approach to question answering.

S. Harabagiu D. Moldovan, M. Pasca and M. Surdeanu. 2002.

E. Hovy, U. Hermjakob, and D. Ravichandran. 2002. A question/answer typology with surface text patterns.

A. Ittycheriah, M. Franz, W. Zhu, and A. Ratnaparkhi. 2001. Question answering using maximum-entropy components.

E. Nyberg and T. Mitamura. 2002. Evaluating qa systems on multiple dimensions.

E. Nyberg, T. Mitamura, J. Carbonell, J. Callan, K. Collins-Thompson, K. Czuba, M. Duggan, L. Hiyakumoto, N. Hu, Y. Huang, J. Ko, L. Lita, S. Murtagh, V. Pedro, and D. Svoboda. 2002. The javelin question-answering system at trec 2002.

M. Soubbotin and S. Soubbotin. 2002. Use of patterns for detection of likely answer strings: A systematic approach.