ACL-2002 Demonstration, Philadelphia, PA, U.S.A., July 7-12, 2002.

Automated Question Answering in Webclopedia – A Demonstration

Ulf Hermjakob, Eduard Hovy and Chin-Yew Lin

USC Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292-6695
{ulf,hovy,cyl}@isi.edu
Tel: 310-448-8476, 310-448-8731 and 310-448-8711

In this demonstration we present Webclopedia, a semantics-based question answering system accessible via the web (Hovy et al. 2002, 2001, 2000). Through a live interface (Figure 1), users can type in their questions or select a predefined question. The system returns its top 5 candidate answers, drawn from NIST's TREC corpus, a collection of 1 million newspaper texts.

Some key points:

Webclopedia integrates **IR and NLP components**. Both symbolic and statistical techniques are employed. For example, the CONTEX parser's grammar, the intermediate result ranking rules, and answer matching patterns are created by machine learning; the answer pinpointer uses hand-crafted matching rules.

Like almost all modern QA systems, Webclopedia uses a **taxonomy of question/answer types**. The QA Typology (Hovy et al 2002), one of the most extensive used in the literature, contains over 180 types, and is based on an analysis of 17,384 questions, plus subsequent extensions.
The typology is at http://www.isi.edu/natural-language/projects/Webclopedia/Taxonomy/taxonomy_toplevel.html.

Webclopedia took part in NIST's TREC **QA evaluations**, achieving MRR (mean reciprocal rank) scores of 31% in TREC9 (tried second place) and 45% in TREC10.

Recent work at ISI has focused on developing **Korean and Mandarin Chinese** versions of Webclopedia, allowing the user to ask English questions and receive English answers from foreign-language text sources.

Instead of using the TREC corpus as source, Webclopedia is being extended to also **query the web**, using commercial web search engines to provide documents with likely answer candidates.

The system works as follows:

- **Question parsing:** Using BBN's IdentiFinder (Bikel et al., 1999), the CONTEX parser (Hermjakob 1997, 2001) produces a syntactic-semantic analysis of the question and determines the QA type(s) sought.
- **Query formation**: Single- and multi-word units (content words) are extracted from the analysis, and WordNet synsets are used for query expansion. A series of Boolean queries is formed.
- **IR**: The IR engine MG (Witten et al., 1994) returns the top-ranked *N* documents.
- **Selecting and ranking sentences**: For each document, the most promising $K<<N$ sentences are located and scored using a formula that rewards word and phrase overlap with the question and its expanded query words. Results are ranked.
- **Parsing sentences**: CONTEX parses the top-ranked 300 sentences.
- **Pinpointing**: Each candidate answer sentence parse tree is matched against the parse of the question, with particular attention to the QA type(s) sought. The matching patterns were built by hand; additional patterns are learned off the web (Ravichandran and Hovy, 2002). As a fallback the window method is used.

- **Ranking of answers**: The candidate answers' scores are computed and the topmost 5 are output as final answers.

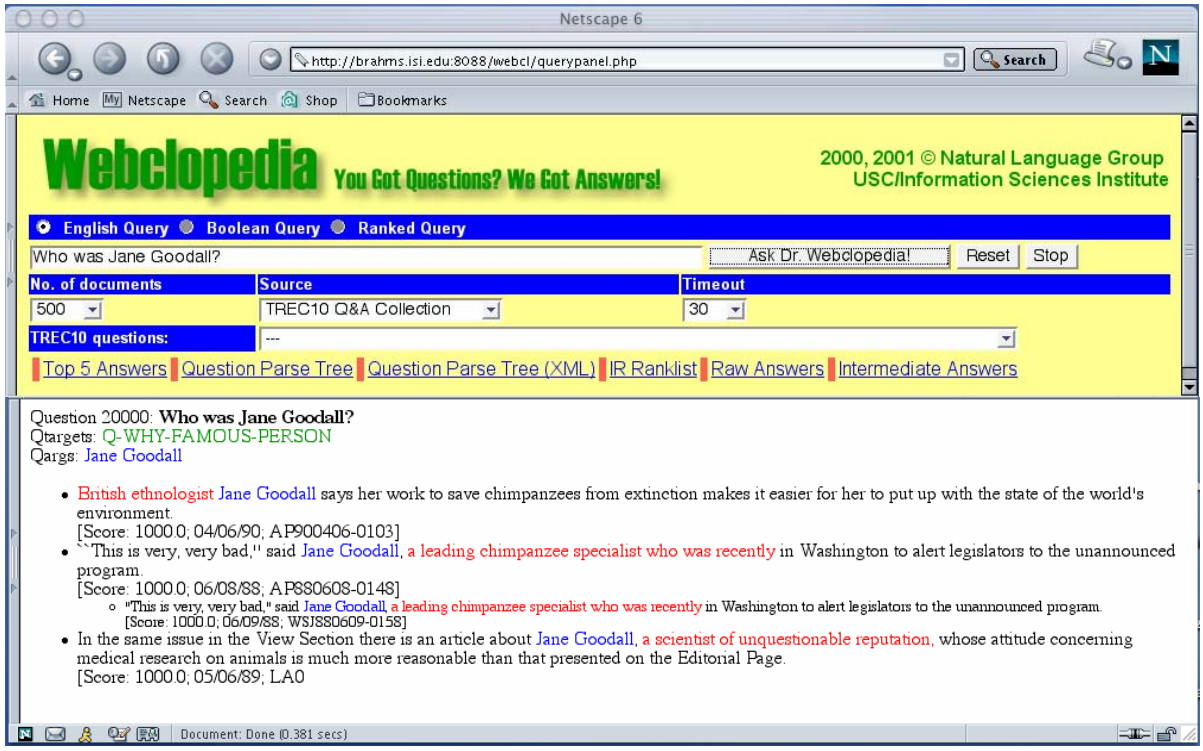

Figure 1. Webclopedia web interface (answers in red, matched portions in blue).

References

Hermjakob, U. 1997. *Learning Parse and Translation Decisions from Examples with Rich Context*. Ph.D. dissertation, University of Texas, Austin. file://ftp.cs.utexas.edu/pub/~mooney/papers/hermjakob-dissertation- 97.ps.gz.

Hermjakob, U. 2001. Parsing and Question Classification for Question Answering. *Proceedings of the ACL Workshop on Question Answering*. Toulouse, France.

Hovy, E.H., L. Gerber, U. Hermjakob, M. Junk, and C.-Y. Lin. 2000. Question Answering in Webclopedia. *Proceedings of the TREC-9 Conference*. NIST, Gaithersburg, MD.

Hovy, E.H., L. Gerber, U. Hermjakob, C.-Y. Lin, and D. Ravichandran. 2001. Toward Semantics-Based Answer Pinpointing. *Proceedings of the DARPA Human Language Technology Conference (HLT)*. San Diego, CA.

Hovy, E.H., U. Hermjakob, and D. Ravichandran. 2002. A Question/Answer Typology with Surface Text Patterns. Poster in *Proceedings of the Human Language Technology Conference*. San Diego, CA.

Ravichandran, D. and E.H. Hovy. 2002. Learning Surface Text Patterns for a Question Answering System. *Proceedings of the ACL Conference*. Philadelphia, PA.

Witten, I. H. and A. Moffat and T. Bell 1994. *Managing Gigabytes*. New York: Van Nostrand Reinhold.