

New Directions in Question Answering

Mark T. Maybury (editor)
(MITRE Corporation)

Menlo Park, CA: AAAI Press and
Cambridge, MA: The MIT Press, 2004,
xi+336 pp; paperbound, ISBN
0-262-63304-3, \$40.00, £25.95

Reviewed by
Marius Paşca
Google Inc.

With goals as intuitive and desirable as they are challenging, the field of automated question answering has generated growing interest in the past few years. The increased momentum is apparent in the spread of research groups working on the topic, the number of relevant Ph.D. theses, papers that are now a common occurrence in the proceedings of top-rated conferences related to the topic, the scale of commercial endeavors, and the steady financial commitments of various agencies to government-sponsored evaluations and research programs. The last of these constitutes the main justification of *New Directions in Question Answering*, a relatively recent contribution to the field. As shown in Table 1, the book suggests new milestones and directions of research which concurrently increase:

- the requirements imposed on the system (e.g., time-sensitive search and detection of obscure relations) (Chapter 2);
- scope (e.g., open or restricted application domains) (Chapter 6);
- complexity (e.g., fact-seeking vs. exploratory questions with opinion answers) (Chapter 7);
- the granularity of information sources (e.g., databases) (Chapter 9) versus unstructured text (Chapter 17) versus full-fledged knowledge bases (Chapter 19);
- generally, the expectations of the users of the system (e.g., regular users versus experts), including intelligence analysts.

Even if considered in isolation rather than combined together, the majority of these goals are already very far reaching. Most readers will have an immediate intuition as to how difficult it would be in practice to answer, with reliable consistency, questions of seemingly unbounded complexity such as *Has there been any change in the official opinion from China toward the 2001 annual U.S. report on human rights since its release?* (page 83), *How has pollution in the Black Sea affected the fishing industry, and what are the sources of this pollution?* (page 134), or *What part did ITT (International Telephone and Telegraph) and Anaconda Copper play in the Chilean 1970 election?* (page 210). Even if the question complexity is limited to the factual type, current fact-seeking question-answering technology has only a moderate impact on global-scale information-seeking environments such as Web search. The fact that quite a few chapters of the book are motivated as departures from

Table 1

List of chapters in *New Directions in Question Answering*.

- 1 Question Answering: An Introduction (*Mark Maybury*)
- 2 Software Architectures for Advanced QA (*Eric Nyberg, John Burger, Scott Mardis, and David Ferrucci*)
- 3 Bringing Commercial Question Answering to the Web (*Brian Ulicny*)
- 4 Answering Definitional Questions: A Hybrid Approach (*Sasha Blair-Goldensohn, Kathleen R. McKeown, and Andrew Hazen Schlaikjer*)
- 5 A Hybrid Approach to Answering Biographical Questions (*Ralph Weischedel, Jinxi Xu, and Ana Licuanan*)
- 6 Question Answering in Terminology-Rich Technical Domains (*Fabio Rinaldi, Michael Hess, James Dowdall, Diego Mollá, and Rolf Schwitter*)
- 7 Low-Level Annotations and Summary Representations of Opinions for Multiperspective QA (*Claire Cardie, Janyce M. Wiebe, Theresa Wilson, and Diane J. Litman*)
- 8 Representing Temporal and Event Knowledge for QA Systems (*James Pustejovsky, Roser Saurí, José Castaño, Dragomir R. Radev, Robert Gaizauskas, Andrea Setzer, Beth Sundheim, and Graham Katz*)
- 9 Answering Questions about Moving Objects in Videos (*Boris Katz, Jimmy Lin, Chris Stauffer, and Eric Grimson*)
- 10 A Data Driven Approach to Interactive QA (*Sharon Small, Tomek Strzalkowski, Ting Liu, Nobuyuki Shimizu, and Boris Yamrom*)
- 11 Finding Answers to Complex Questions (*Anne R. Diekema, Ozgur Yilmazel, Jiangping Chen, Sarah Harwell, Lan He, and Elizabeth D. Liddy*)
- 12 Experiences from Combining Dialogue System Development with Information Extraction Techniques (*Arne Jönsson, Frida Andén, Lars Degerstedt, Annika Flycht-Eriksson, Magnus Merkel, and Sara Norberg*)
- 13 Reuse in Question Answering: A Preliminary Study (*Marc Light, Abraham Ittycheriah, Andrew Laffo, and Nancy McCracken*)
- 14 Retrieval Models and Q and A Learning with FAQ Files (*Noriko Tomuro and Steven L. Lytinen*)
- 15 Holistic Query Expansion Using Graphical Models (*Daniel Mahler*)
- 16 Viewing the Web as a Virtual Database for Question Answering (*Boris Katz, Sue Felshin, Jimmy Lin, and Gregory Marton*)
- 17 Toward Answer-Focused Summarization Using Search Engines (*Harris Wu, Dragomir R. Radev, and Weiguo Fan*)
- 18 A Multi-agent Approach to Using Redundancy and Reinforcement in Question Answering (*John Prager, Jennifer Chu-Carroll, and Krzysztof Czuba*)
- 19 Deductive Question Answering from Multiple Resources (*Richard Waldinger, Douglas Appelt, Jennifer L. Dungan, John Fry, Jerry Hobbs, David J. Israel, Peter Jarvis, David Martin, Susanne Riehemann, Mark E. Stickel, and Mabry Tyson*)
- 20 Advanced Relaxation for Cooperative Question Answering (*Farah Benamara and Patrick Saint-Dizier*)
- 21 Trusting Answers from Web applications (*Deborah L. McGuinness and Paulo Pinheiro da Silva*)

fact-seeking question answering should ensure that the book will attract the attention of intrigued and avid readers.

Most of the 21 chapters of the book are extended versions of papers from a 2003 AAAI symposium on question answering (Maybury 2003). As a strength, the chapters provide a multitude of refreshing, often orthogonal perspectives on question answering. The chapters are mostly independent units, with few cross-references. It is therefore possible to focus first on the chapters that seem most relevant and read other chapters later. In fact, out-of-order perusal may reduce the occasional overlap in the sections that motivate various chapters individually, as they often share, as a common theme, the desire to move beyond fact-seeking question answering.

It is impractical to describe all chapters individually. For an audience interested in more-practical methods that are accompanied by thorough evaluations, I recommend chapters 5, 17, and 18. In particular, Chapter 5 (Ralph Weischedel, Jinxi Xu, and Ana Licuanan) takes a look at answering biographical questions. Since such questions often take very simple forms (e.g., *Who is Edgar Froese?*), the challenge relative to other classes of questions does not lie in question analysis. Instead, the inherent difficulty of biographical questions, and of the related class of definitional questions (e.g., *What is a metronome?*), is in selecting and assembling an answer from text fragments that are inherently scattered across multiple documents. The chapter combines clear explanations of the method and system architecture with a nice walk-through example that shows the information flow among different stages of processing of a sample question. Most notably, the evaluation section more than makes up for the relatively small size of the test set (less than 30 questions) by analyzing the system output from different angles and pursuing both a subjective and an automated evaluation of the results.

Chapter 17 (Harris Wu, Dragomir Radev, and Weiguo Fan) is among the few that consider fact-seeking questions. Following a pragmatic approach, the chapter introduces a series of experiments and comparative evaluations of a task that is more relaxed than those that motivate other chapters and yet has high potential impact in practice. More precisely, the authors efficiently address the problem of extracting more-accurate document snippets or summaries in response to users' questions, in addition to output returned by Web search engines. Chapter 18 (John Prager, Jennifer Chu-Carroll, and Krzysztof Czuba) is one of my favorites. It is a very effective and enlightening tour through a collection of different answer extraction techniques embedded in a single system, which could also be seen as a meta-question-answering system. The various extraction techniques (statistical, pattern-based, definitional, dossier-based) are often radically different, yet complement one another well, resulting in one of the most modular architectures for question answering described to date.

If the main reason to approach the book is an interest in exploratory work and attempts to formalize various problems related to question answering, without a focus on complete evaluation, Chapters 2, 7, 8, and 13 offer an enjoyable lecture. Chapter 2 (Eric Nyberg et al.) aims at capturing the requirements of advanced question answering and their impact on system design. The challenges that face the push toward question-answering systems of increased complexity become apparent to the attentive reader after reading the chapter, especially challenges of practicability and scalability. Indeed, such issues become important for any system that would actually attempt to perform AI-style planning in a broad domain (Section 2.3.2). Similarly, it may be very challenging to find common linguistic representations (Section 2.3.1) to use across highly modular systems for encoding internal information, as the information sources themselves can vary widely from unstructured text on one end of the spectrum to full-blown knowledge bases on the other end. Chapter 7 (Claire Cardie

et al.) describes a representation scheme for annotating opinions, as opposed to facts, as they occur in textual documents. The intuition is that identifying and extracting opinions automatically from text would be essential in any attempts to answer questions with multiple possible perspectives, such as *What is the general opinion from the press on the recent U2 world tour?* The proposed annotation scheme and its motivations are described in detail, exemplified in a few examples, and shown to provide promising interannotator agreement scores despite the difficulty of the task. Chapter 8 (James Pustejovsky et al.) provides an overview of an annotation standard for representing events and temporal expressions as they occur in text. The representation language is the product of significant research efforts and numerous iterative improvements. The current version of the language was fueled by a series of government-sponsored workshops dedicated to the topic of temporal awareness in the context of question answering (Pustejovsky 2002). In Chapter 13 (Marc Light et al.), an empirical analysis of a corpus of questions enables the authors to identify examples of reuse scenarios, in which future questions could be answered better by using information previously available to the system (e.g., in the form of previously submitted questions or answers already returned to users). The authors acknowledge that some of the proposed categories of reuse are very difficult to implement in working system modules.

Chapters 19 (Richard Waldinger et al.), 20 (Farah Benamara and Patrick Saint-Dizier), and 21 (Deborah McGuinness and Paulo Pinheiro da Silva) delve into knowledge-based question answering and supporting inferential processes for verifying candidate answers and providing justifications. For the audience looking for an insider's view into existing question-answering systems, Chapters 14 (Noriko Tomuro and Steven Lytinen) and 16 (Boris Katz et al.) will be particularly interesting. Rather than proposing any new directions, they share some of the practical lessons learned while providing users with answers from FAQ files (Chapter 14) and from the Web seen as a database of facts (Chapter 16).

Some of the advanced methods proposed in the book are likely to become relatively more language-dependent, as they require larger and more-complex resources of various kinds. Given the variety of topics already covered by its chapters, the book limits the scope of the discussion to questions and answer sources in English. Many lessons learned through experiments on question answering in different languages can be found in Chen and Lin (2002), Kando and Ishikawa (2004), and Peters and Borri (2004).

Question answering is situated at the confluence of a large number of related areas: information retrieval (Gaizauskas, Hepple, and Greenwood 2004), natural language processing (Ravin, Prager, and Harabagiu 2001; de Rijke and Webber 2003), information extraction, and knowledge representation and reasoning (Harabagiu and Chaudhri 2002), to name only a few. Overall, the book represents a useful reference for readers interested in possible future developments in the field. There are a few scattered typos and very rare inconsistencies, including a paragraph on page 173 that somehow slipped into the final manuscript while still marked as *To review*. It will be interesting to compare the present and near-future progress in question answering against the aggressive milestones from the proposal in the first chapter of the book.

References

Chen, Hsin-Hsi and Chin-Yew Lin, editors.
2002. *Proceedings of the Workshop on
Multilingual Summarization and Question*

Answering at COLING-02, Taipei,
Taiwan, August.
de Rijke, Maarten and Bonnie Webber,
editors. 2003. *Proceedings of the Workshop on*

- Natural Language Processing for Question Answering at EACL-03*, Budapest, Hungary, April.
- Gaizauskas, Robert, Mark Hepple, and Mark A. Greenwood, editors. 2004. *Proceedings of the Workshop on Information Retrieval for Question Answering at SIGIR-04*, Sheffield, UK, July.
- Harabagiu, Sanda and Vinay Chaudhri, editors. 2002. *Proceedings of the AAAI Spring Symposium on Mining Answers from Texts and Knowledge Bases*, Stanford, CA, March.
- Kando, Noriki and Haruko Ishikawa, editors. 2004. *Working Notes of the Fourth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Summarization (NTCIR-04)*, Tokyo, Japan, June. National Institute of Informatics.
- Maybury, Mark, editor. 2003. *Proceedings of the AAAI Spring Symposium on New Directions in Question Answering*, Stanford, CA, March.
- Peters, Carol and Francesca Borri, editors. 2004. *Working Notes of the Fifth Cross-Language Evaluation Forum (CLEF-04)*, Bath, UK, September.
- Pustejovsky, James, editor. 2002. *Final Report of the Workshop on TERQAS: Time and Event Recognition in Question Answering Systems*, Bedford, MA, January–July. ARDA.
- Ravin, Yael, John Prager, and Sanda Harabagiu, editors. 2001. *Proceedings of the Workshop on Open-Domain Question Answering at ACL-01*, Toulouse, France, July.

Marius Paşca is a senior research scientist at Google Inc. He earned a Ph.D. in computer science from Southern Methodist University in 2001, with a thesis on open-domain question answering from large text collections. His current research interests include natural language processing and text mining for information retrieval. His address is Google Inc, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA.