Why Should You Trust Answers from Web?

Deborah L. McGuinness

Knowledge Systems, Artificial Intelligence Laboratory, Stanford University

Abstract

In order to trust answers obtained from arbitrary applications, users will need to understand how answers were obtained and what information they depended upon. Particularly in web applications that may use question answering systems that may be heuristic or incomplete or data that is either of unknown origin or may be out of date, it becomes more important to have information about how answers were obtained. We propose that future web systems will return answers augmented with Meta information about how answers were obtained. In this position paper, we explore an approach that can improve trust in answers generated from web applications by making the answer process more transparent. The added information is aimed to provide users (humans or agents) with answers to questions of trust, reliability, recency, and applicability.

Keywords: Explanation, Trust, Semantic Web.

1. Trustable Answers

As the web has grown in terms of its distributed nature, scope, and diversity, answer quality has become more difficult to judge. Today's recipients of answers may be obtaining results from multiple question answering systems that may obtain raw input from many sources, some of which may have origins that are either unknown or questionable. Additionally, raw data may vary in recency and may have unknown or unreliable origins. It has become difficult in many settings to judge how reliable or trustworthy an answer is and thus it may be difficult to determine how one should use answers, and when one should act on them. Our vision of the future is one where answers are annotated to include information about:

- the sources used to obtain the answer (along with provenance information about the sources),
- the question answering systems used to generate the answers and any information manipulation steps they performed,
- any assumptions relied on (including logic-related notions such as closed world reasoning, negation as failure as well as domain-oriented assumptions such as "birds fly").

This annotation would also include trust values for information that is used to compute the answer. Thus, if some of the information was obtained from the New York Times, trust values for either the user's or the community's trust rating of the NYT would be included. Additionally, if text extraction routines were run over the NYT to obtain "facts" used to populate a knowledgebase, then trust values for the extraction routines (as well as other meta information about the extraction routines such as author, release status, etc.) would be potentially included.

These optional annotations may contain quite a bit of information that may be presented through interfaces that filter the information at a level appropriate to the user. Some users may want to see a lot of detail but other users may want to see high level abstractions or summaries. Some users may need interactive graphical displays while others may need text summaries. Sometimes the same user will need different amounts of information for the same question depending upon the context. One example summarization strategy may include a description of the sources relied on to return the answer (e.g., all sources were considered authoritative and were updated in the last six months).

We believe that much of the underlying infrastructure to support this future answer annotation is provided by the InferenceWeb [1]. Inference Web is an explanation toolkit for the web. It includes a proof markup language (PML) [2] that may be used to encode justifications for any answer. PML may be used to encode source information as well as information manipulation information. PML is encoded in W3C's recommended web ontology language, OWL [3], and interoperates with distributed applications in W3C recommended standards such as XML and RDF. Sources may be registered in the Inference Web registry so they may be accessed and presented in detailed views of the justifications. Inference Web includes a browser that may be used to display justifications in proof markup language in multiple formats including natural language, graphical views, and logical views. The toolkit also includes multiple strategies for presenting information so that end users may have summary views or detailed views and views may be customized according to context. IW includes services to help applications generate and check PML. It also has database registry services. Inference Web has also been expanded with a trust component (IWTrust [4]) so that it may access networks of trust values and may present those trust values, combine them, propagate them, and use them for filtering answers.

2. Towards Trustable Answers

One can see pieces of our vision of answers augmented with explanation support by either looking at prototype implementations or initial program deployments. We began this theme with research on explaining description logics [5, 6], its implementation in the CLASSIC description logic system [7] and its use to explain results in applications such as the PROSE/QUESTAR family of configurators [8]. Since then, we began work evolving that approach to be better suited to distributed web applications (with Pinheiro da Silva) and generated today's Inference web infrastructure. We are using Inference Web to explain recommendations to intelligence analysts [9] in ARDA's Novel Intelligence for Massive Data program. One aspect of this work focuses on explaining how text extraction techniques were used to generate facts in a knowledge base. The explanations can expose the raw sources used and the text analytic techniques. Another aspect of the work focuses on explaining the theorem-prover style reasoning that is used to generate conclusions of interest from the knowledge bases and presents them to the user. The explanations can expose assumptions relied on as well as reasoning techniques used and it can display multiple styles of presentations (from summaries to graphs to natural language). In both focus areas, the internal application generates answers with optional proof markup language that contains all of the information required to generate explanations for the answer, how it was obtained, what it depended on, and anything known about the raw source used. The Inference Web toolkit is used to compose, display, integrate, and manipulate the explanations.

One other prototype application, the KSL Wine Agent [10, 11], shows how wine and food pairing recommendations can be explained. The interface prompts users to give some description of the meal they are planning to eat and then the application determines either a description of the recommended wine or a listing of specific wine recommendations (from local web sites). The internal application uses a theorem prover run over a knowledge base of wine and food descriptions along with suggested pairings descriptions. It uses the OWL query language (OWL-QL [12]) to ask questions of the reasoner for recommendations. All recommendations come back with a proof markup language justification for how the answer was obtained including information about what sources it relied on (e.g., if it obtained some information about wines from one of the web sites it knows about or if it only used its internal wine knowledge base). The wine agent uses Inference Web to generate, display, and provide follow-up explanations.

We have also explored the vision of augmenting answers to semantic matching applications [13, 14]. Here satisfiability algorithms are used to determine when two terms may mean the same thing. Inference Web and JSAT have been integrated so that the application can generate PML for its matching recommendations and Inference Web may be used to display justifications for its recommendations.

The Inference Web-based explanation approach has also been tested in designs for explaining task-oriented processing such as that done in SRI's SPARK system that is used in DARPA's Personalized Assistant that Learns program [15]. Inference Web is being integrated with SPARK so that answers to questions such as "what are you doing?" and "why are you still doing that?" may be answered with information such as what a goal is in service of, who told the agent to do it, what it is waiting for, etc. This is being demonstrated in the Year 2 demonstration system from SRI.

3. Conclusion

In this paper, we have discussed the topic of users' (humans and agents) need for information to support their decisions concerning trusting and acting on answers from web applications. We view the future of applications as one where all answers are optionally annotated with justification information. We provided some examples of this vision using the Inference Web explanation toolkit and the Proof Markup Language supporting interoperable web justifications.

4. References

[1] Deborah L. McGuinness and Paulo Pinheiro da Silva. Explaining Answers from the Semantic Web: The Inference Web Approach. Journal of Web Semantics. Vol.1 No.4, pages 397-413, October 2004.

[2] Paulo Pinheiro da Silva, Deborah L. McGuinness and Richard Fikes. A Proof Markup Language

for Semantic Web Services. Information Systems. Accepted for publication.

[3] Deborah L. McGuinness and Frank van Harmelen. OWL Web Ontology Language Overview. World Wide Web Consortium (W3C) Recommendation. February 10, 2004. Available from http://www.w3.org/TR/owl-features/ .

[4] Ilya Zaihrayeu, Paulo Pinheiro da Silva and Deborah L. McGuinness. IWTrust: Improving User Trust in Answers from the Web. Proceedings of 3rd International Conference on Trust Management (iTrust2005), Springer, Rocquencourt, France, 2005.

[5] Deborah L. McGuinness. ``Explaining Reasoning in Description Logics''. Ph.D. Thesis, Rutgers University, 1996. Technical Report LCSR-TR-277.

[6] Deborah L. McGuinness and Alex Borgida. ``Explaining Subsumption in Description Logics.'' In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Canada, August, 1995.

[7] Ronald J. Brachman, Alex Borgida, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lori Alperin Resnick. ``Living with CLASSIC: When and How to Use a KL-ONE-Like Language.'' In *Principles of Semantic Networks: Explorations in the representation of knowledge*, ed. John Sowa. San Mateo, California: Morgan Kaufmann, 1991, pages 401—456.

[8] Deborah L. McGuinness and Jon Wright. ``An Industrial Strength Description Logic-based Configurator Platform''. *IEEE Intelligent Systems*, Vol. 13, No. 4, July/August 1998, pp. 69-77.

[9] Christopher Welty, J. William Murdock, Paulo Pinheiro da Silva, Deborah L. McGuinness, David Ferrucci, Richard Fikes. Tracking Information Extraction from Intelligence Documents. In Proceedings of the 2005 International Conference on Intelligence Analysis (IA 2005), McLean, VA, USA, 2-6 May, 2005.

[10] http://www.ksl.stanford.edu/people/dlm/webont/wineAgent/

[11] Eric Hsu and Deborah L. McGuinness. KSL Wine Agent: Semantic Web Testbed Application, Proceedings of the 2003 International Workshop on Description Logics (DL2003). Rome, Italy, September 5-7, 2003.

[12] http://ksl.stanford.edu/projects/owl-ql/

[13] Deborah L. McGuinness, Pavel Shvaiko, Fausto Giunchiglia and Paulo Pinheiro da Silva. Towards Explaining Semantic Matching. In 2004 International Workshop on Description Logics (DL2004), Whistler, British Columbia, Canada, CEUR Workshop Proceedings Vol. 104, 2004.

[14] Pavel Shvaiko, Fausto Giunchiglia, Paulo Pinheiro da Silva and Deborah L. McGuinness. Web Explanations for Semantic Heterogeneity Discovery. In Proceedings of the 2nd European Semantic Web Conference (ESWC 2005), Heraklion, Greece, May 29 to June 1, 2005 Springer, 2005.

[15] http://www.darpa.mil/ipto/Programs/pal/index.htm