

An Analysis of the AskMSR Question-Answering System

Eric Brill, Susan Dumais and Michele Banko July 2002

Aditya Vadrevu

I. DESCRIPTION OF THE STUDY

The main purpose of this research is to provide a system in which a question is answered correctly. The AskMSR system predicts an answer using the internet as a giant corpus to find data. This problem is significant because it can help us formulate answers to fact based questions and provides a way to properly retrieve data from a large corpus. The main objectives of this research is to find ways to increase accuracy of the question answering system by finding a proper answer to a question as well as predicting when an answer is incorrect and not answering.

II. METHODS AND DESIGN

The architecture that the authors use to model the system is as follows:

- Query Reformulation
 - Query Reformulation is where the question string is rewritten as an answer by taking all the non stop words and AND-ing them to produce an answer-like statement. For example, "Where is the Eiffel Tower located?" into "The Eiffel Tower is located..." The authors currently do not use a part of speech tagger or a parser to do this but have created rules and weights to do this task.
- N-gram Mining
 - Using the rewritten questions the system then sends the question to a search engine and gets the results. Using the results that the search engine gives, not the full page responses, the system generates uni-gram, bi-gram, and tri-gram models. The models are then weighted by the question that retrieved that answer as well as the n-gram it uses.
- N-gram Filtering
 - Using one of seven question types the n-gram is reweighted. The n-grams are then analyzed and reweighted again based on how many features are relevant to the question being asked. 15 different filters were added to help reweight the n-grams properly.
- N-gram Tiling
 - The final step in this system is to merge similar answers together to form a longer answer. Answers

with similar parts will be joined together and n-grams with different answers will be removed from the list of possible answers.

This seems like a very efficient process not only because the system reads only search engine results but because the small answers merge to form a longer answer. This makes it more likely to produce a correct answer.

III. ANALYSIS AND RESULTS

There are 2 different data corpus that the system could use, the web and TREC databases. The main way that the system is used is by using Google as the back end search engine because it provided summaries of a website so that the system could access that instead of reading through the whole page. The accuracy of the default system was 61% and they were able to increase this number by knowing when not to answer the question. Removing or modifying any of the 4 parts of the system caused a significant drop in accuracy with no filtering causing the largest drop of 33%. All of the parts of the system are necessary in order to achieve a high accuracy to provide a correct answer. The most important feature seems to be the tiling of n-grams. This makes sense because it is taking small and concise answers and adding on similar answers to form a more in-depth and detailed answer. This will most likely lead to an answer that is correct.

IV. LIMITATIONS

The biggest limitation that the authors have faced is number retrieval. The input for a numerical retrieval question itself is the issue, the authors report. They state that in-order to get numerical answers we must have numerical queries which they don't do in the system. The authors state that only 12% of failures of the system are actually incorrect and the others can be fixed with minor enhancements to the system.

V. SIGNIFICANCE

This paper is significant to the Information Retrieval and Information Extraction communities because it helps explore the ways in which data can be manipulated in-order to form answers to queries that are received. This can also be significant for NLP communities because of the way that the query is rewritten. The rewriting of the question into a simpler answer query is arguably one of the most important parts of the whole system.

VI. CONCLUSION

Overall this paper gives an in-depth insight into the process that the AskMSR system goes through when it needs to find an answer to a query. The authors go into great detail for each step which is very important and outline where their shortcomings were. They also provide next steps in this project expanding from short answers to full fledged information retrieval. The main point of focus for the paper is that they were able to find a way to predict if the question would produce an answer that was accurate enough, and if it did not then the system would not answer the question. They were able to explain this system well enough to say that this paper is worth reading.

REFERENCES

- [1] Brill, Eric, et al. An Analysis of the AskMSR Question-Answering System. Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - EMNLP 02, July 2002, doi:10.3115/1118693.1118726.