Towards a Hierarchical Framework for Predicting the Best Answer in a Question Answering System

Mohan John Blooma, Alton Yeow-Kuan Chua, Dion Hoe-Lian Goh,
and Zhiquan Ling

Division of Information Studies, Wee Kim Wee School of Communication & Information,
Nanyang Technological University
{bl0002hn,AltonChua,ashlgoh,ling0032}@ntu.edu.sg

Abstract. This research aims to develop a model for identifying predictive variables for the selection of the best quality answer in a question-answering (QA) system. It was found that accuracy, completeness and relevance are strong predictors of the quality of the answer.

Keywords: Question answering systems, Answer quality, Information Retrieval, Multiple Regression, Prediction model.

We developed a theoretical model from past studies on the dimensions of information quality for information resources on the web. The model developed was tested using actual data collected from a user-oriented QA system –Yahoo! Answers. The aim of the paper is to investigate if the hypothesized predictive features on the quality of the answer actually support the prediction of the best answer in the QA system.

Table 1. Predictive features of the quality of answer with their respective significance

Features			Coefficients	Sig.
Level 1	Level 2	Level 3	Beta	Std. Error
Non-textual Features	Question	Category	.028	.178
	Answerer	Reputation	.008	.701
		Authority	.012	.579
	Asker	Reputation	.010	.624
		Authority	-.066	.002
Textual Features	Answer	Accuracy	.230	**.000**
		Completeness	.375	**.000**
		Length	.027	.188
		Language	.039	.063
		Reasonableness	.370	**.000**

Features examined in this study are divided into textual and non-textual. Non-textual features cover three main factors, namely, the category of the question, the asker of the question and the answerer whose answer was selected as the best answer by the asker. Reputation of the asker, authority of the asker, reputation of the answerer, authority of the answerer, category of the question are the five dimensions

of non-textual features used as predictive variables in this study. Textual features represent different dimensions of the answer itself. Accuracy, completeness, language, reasonableness, and length of the answer are five dimensions of textual features used for evaluating the quality of answer.

Reputation and authority were derived from the user ratings while the textual features were evaluated using two human volunteers. A multiple regression analysis was used to model the dependent value quality based on its linear relationship to one or more predictors. Textual features were found to be significant predictors (Table 1.). In particular, the answer's accuracy, completeness, and reasonableness were the most significant predictors for the quality of the answer. Non-textual features were found to be least significant.

It was found that the quality of answer was most influenced by the textual features of the answer itself, rather than the non-textual features. Accuracy, completeness and reasonableness were found to be the most significant textual features highlighting the fact that the user is looking for accurate, understandable, complete and relevant information. Among these variables, completeness was found to be the strongest predictor followed by the reasonableness and accuracy of the answer. The same features were shown to be the indicators of quality of websites and information resources by various studies [1]. However, length and language were not found to be significant predictors. This study thus reveals that irrespective of the length and language of the answer, users prefer the completeness and understandability of the answer. The non-textual features like reputation of the asker, authority of the answerer, reputation of the answerer and the category were also found to be non-significant predictors. Thus these findings confirm that the user values, the understandability of the answer and the extent to which the answer satisfies his/her information needs, more than the reputation and authority of the answerer. However, recent work on evaluation of question answering system by [2] highlights the fact that the current trend points to the centrality of users' perception. This study also reveals that the satisfaction of user needs is more important than the source of the information. By promising to deliver answers, not just documents, question answering systems can more effectively fulfill users' information needs.

In conclusion, our study offers a framework for evaluating answer quality that could be used on related systems. As our study is on a community-based QA system, the results could be extended to other community-based web services and for enhancing knowledge access to digital libraries. Using this framework, researchers will be able to better understand the significant features in developing answer extraction modules for the retrieval of good quality answers in a QA system.

References

1. Fricke, M., Fallis, D.: Indicators of accuracy for answers to ready reference questions on the internet. Journal of the American Society for Information Science and Technology 55(3), 238–245 (2003)
2. Wacholder, N., Kelly, D., Kantor, P., Rittman, R., Sun, Y., Bai, B., Small, S., Yamron, B., Strzalkowski, T.A.: Model for quantitative evaluation of an end – to end question answering system. Journal of the American Society for Information Science and Technology 58(8), 1082–1099 (2007)