Chapter 11

Numeric Query Answering on the Web

Steven O'Hara
University of Texas at San Antonio, USA

Tom Bylander
University of Texas at San Antonio, USA

ABSTRACT

Query answering usually assumes that the asker is looking for a single correct answer to the question. When retrieving a textual answer this is often the case, but when searching for numeric answers, there are additional considerations. In particular, numbers often have units associated with them, and the asker may not care whether the raw answer is in feet or meters. Also, numbers usually denote a precision. In a few cases, the precision may be explicit, but normally, there is an implied precision associated with every number. Finally, an association between different reliability levels to different sources can be made. In this paper, the authors experimentally show that, in the context of conflicting answers from multiple sources, numeric query accuracy can be improved by taking advantage of units, precision, and the reliability of sources.

MOTIVATION

Suppose you wanted to know the answer to a specific question, such as “What is the atomic mass of Copper?” or “What is the diameter of Pluto?” Certainly, you could Google (www.google.com) it and get many different answers. You could also ask experts, or go to specific websites that you expect to know an answer, or run software to get an answer. Now, suppose you were able to ask all of these sources at the same time, and wanted to consolidate those answers into a single answer. You’d probably find some answers that are incorrect; and some answers that are more precise than others; and there may even be multiple “correct” answers.

DOI: 10.4018/978-1-4666-3610-1.ch011
Expecting any single source to have the best answers for a wide range of questions is quite unreasonable. Sources like Wikipedia (www.wikipedia.org), Google Calculator, Yahoo! Answers (answers.yahoo.com), Wolfram Alpha (www.wolframalpha.com) and others are viable sources, but not to the exclusion of the other possible sources. One of the major tenets of this research is that having a range of sources for answers is likely to increase overall accuracy.

Many questions follow this same form: given an object, what is the value of a specified attribute? What is the radius of Mars? What is the population of Rome? What is the length of the Nile River? How tall is Mt. Everest? And so forth. The literature uses the term “factoid” for these types of questions (Lin & Katz, 2003).

With the immense amount of information available on the internet, and the advent of systems like GROK (described herein), it is possible to send a single question and receive many different answers quickly. The GROK architecture provides the mechanism for broadcasting questions to multiple agents (Singh & Huhns, 2005), which in turn search websites for answers. Next, those answers are consolidated into a “best” answer and returned back to the user. Natural Language processing (such as understanding, paraphrasing or generation) is beyond the scope of this project. See the Text Retrieval Conference website (trec.nist.gov) for some recent research.

HYPOTHESIS

Numeric query answering is different than ordinary query answering (Katz et al., 2002; Kwok, Etzioni, & Weld, 2001; Lin & Katz, 2003; Roussinov, Fan, & Robles-Flores, 2008) in two fundamental ways: unit conversions and implied precision. Query accuracy can be improved by taking advantage of these distinct differences. It can also take advantage of source reliability, which is considered in this paper.

Numeric Values Often Have Units

When searching for numeric values, many of them will be expressed in units, such as feet or meters. When a candidate value is found, the units can be converted automatically. So, if you are searching for the radius of Mars, some answers may come back in meters, some in miles, and some in kilometers. If the search is expecting only miles, it may miss some good answers that were expressed in kilometers, for example.

Numbers Have Implicit Precision

In a few cases, numbers have explicit precision. Wikipedia, for example, often shows values in the form “2100 ± 40 grams” which means 2060 to 2140 grams. However, most resources will report values like “2100 grams” or “2110 grams”, etc. Based on common interpretation of numbers ending in zeros, we assume that the 2100 value has an implied precision of 2050 to 2150, and that the 2110 value means 2105 to 2115. These interval values can be compared against each other to see if they reinforce each other.

Naturally, this assumption can fail in both directions. One might find a value such as “1000 meters” that really means exactly 1000.000 meters. And one might find a value such as “123.45 miles” that is not really accurate to one hundredth of a mile. Part of this research is testing to see if overall accuracy improves with this assumption, even knowing that it will occasionally fail.

Sources Have Different Reliability

Sources like NASA are tightly controlled and very likely to have accurate answers to most astronomy-related questions. Wikipedia is loosely controlled (for a discussion on reliability, see en.wikipedia.org/wiki/Reliability_of_Wikipedia); information is often entered by non-experts and is often invalid. Google, in general, is completely uncontrolled. When comparing answers, it seems reasonable to consider the reliability of the source.