

Answer yes/no questions in search engines

Chunbin Lin
University of California, San Diego
chunbinlin@cs.ucsd.edu

ABSTRACT

According to the query logs in search engines, yes/no questions are issued frequently. One example yes/no question is “is Obama the 42nd president of the United States”, and the expected answer is “No”. However, the most advanced search engines only return a list of ranked documents containing the keywords within the query. Then users have to go through those returned documents in order to figure out one simple answer “No”. People may argue that this query contains semantic and it can hardly be handled by search engines. Nevertheless, current search engines have the ability to process most of the natural language questions. For instance, issuing the query “who is the 42nd president of the United States” on Google, the answer is “Bill Clinton”. Therefore, current search engines should have the ability to handle most of the yes/no queries. Based on the answer of “who is the 42nd president of the United States”, the answer of “is Obama the 42nd president of the United States” should be easily obtained. Thus, current search engines can answer most of the yes/no questions.

The ability of the answering yes/no questions will enhance the power of current search engines and will significantly improve the users’ search experiences.

Yes/no queries can be divided into two categories *objective yes/no queries (OYN questions)* and *subjective yes/no queries (SYN questions)*.

- The answers of OYN queries are simple “Yes” or “No” with evidences, which can be obtained based on existing facts. For example the answer to the OYN question “is universal studios hollywood open now?” is “Yes” if the following two facts are gained. (1) The open time of Universal Studios Hollywood is 9AM-10PM from Monday to Friday, and (2) it is 9:30AM Monday now. The answer becomes “No” if now is “11PM Monday”.
- The answers of SYN queries are (Yes, P_y) and (No, P_n) pairs with evidences, where P_y and P_n are the probability of the answers to be “yes” and “no” respectively. There are no facts for such queries, but user subjective opinions. For example the answer to the SYN question “is ios better than android” is “(Yes, 55%), (No, 45%)” based on analyzing users’ comments, tweets and blogs.

Architecture of processing yes/no queries. Figure 1 shows the architecture of processing yes/no queries. The *Query Analyzer* analyzes the query to identify the type. One possible solution is to check whether it contains subjective terms like comparison terms, e.g., “better than” and “good”. But it is not an accurate measurement, as such terms can

sometimes be a non. E.g., query “is good an English word?” is a OYN query even it contains term “good”. To improve the accuracy of query type identification, query logs should be utilized. Based on query logs, if a similar query “is bad an English word” is an OYN query, then “is good an English word?” is an OYN with high probability.

The *Context Manager* collects users’ time-spatial information and also relevant documents from search engines via keyword search. Time-spatial information is served as filter condition, while keyword search returns only relevant context of queries. For example, in order to answer the query “is it rainy today”, the following three information should be obtained: (1) today’s date, e.g., 8/1/2016, which is the time information, (2) user’s location, e.g., New York, which is the spatial information, and (3) the weather of New York on 8/1/2016.

The *OYN/SYN Processor* processes the OYN and SYN queries based on the corresponding context and query types, and returns answers in corresponding format.

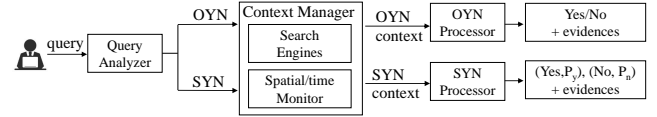


Figure 1: Architecture of processing yes/no queries.

Results exhibit Figure 2(a) shows an example output format for an OYN query “is it rainy today”, while Figure 2(b) demonstrates an example result of an SYN query “is ios better than android”. The outputs directly answer yes/no queries. Users can also see more details to see the evidences.

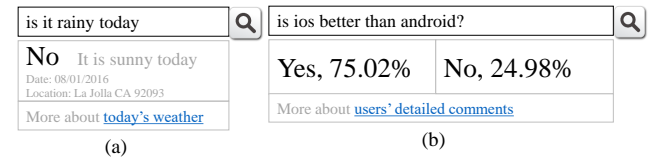


Figure 2: Example answers of a yes/no query.

Remark Though the class of yes/no queries is a member of natural language questions, it is still possible to completely handle yes/no queries. This is because (1) yes/no queries are verification questions, which are generally easier than other natural language questions, and (2) the answers of yes/no queries have fixed formats. Therefore, search engines should pay their attention to solve such queries, which are issued widely and can be completely solved.