A New Approach to Query Suggestion for Bipolar Queries via Linguistic Data Summaries

Janusz Kacprzyk[0000-0003-4187-5877] and Sławomir Zadrożny[0000-0002-6642-0927]

Systems Research Institute, Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
kacprzyk@ibspan.waw.pl, zadrozny@ibspan.waw,pl

Abstract. We are concerned with an extension of our novel approach to the bipolar relational database queries represented by the mandatory and optional conditions like: „find a house which is inexpensive (mandatory) and, if possible, is close to city transportation (optional)"- cf. Zadrożny and Kacprzyk (2009 – 2023). These conditions are then aggregated by some specific operator, „… and, if possible,…" beyond the traditional „and", or similarly by another bipolar type aggregation „…or, if impossible…".
The bipolar database queries have a much higher expressive power by being able to more adequately express information needs, intentions and preferences of the human user. They have found many applications showni in the authors' works, and have many extensions, notably by the authors' addition of context.
However, in real life querying, also in the bipolar type, for an average user the formulation of a proper query to be able to retrieve what is really needed can be too difficult conceptually and technically. In such a case, a solution can be to use query suggestion which has enjoyed much popularity both in more theoretical and application oriented works on database querying or Web search.
We propose a new approach to query suggestion which boils down to an automatic generation of some proper additional queries. We use two patterns. We start with an initial query formulated by the user which is maybe not ideal for retrieving information sought, and should be augmented or modified. We use here the general approach used in virtually all traditional approaches to query suggestion, i.e. to, analyze the history of queries posed by the particular user, or a class of users or concerning some aspect, and then use machine learning tools and technique, to subsequently formulate new augmented or modified queries.
The novelty of our approach is that to determine the above modification/change of the initial query, we use an intuitively appealing linguistic summary, exemplified by „in most of recent queries posed by the user, the mandatory condition on price is milder and optional condition on city transportation is stricter".
We either generate such suggested modification one by one, or a set of them in one shot, as a list of suggestions, all by using linguistic data summaries, too, augmented by measures of similarity between the linguistic summaries.
For both these generation options we propose a metric for the ranking of query suggestions via our concept of similarity between the linguistic summaries.

Keywords: Fuzzy query, Bipolar query, Query Suggestion, Linguistic summary