Fuzzy Logic as a Model of User Preference

Veronika Vaneková\textsuperscript{1} and Peter Vojtáš\textsuperscript{2}

\textsuperscript{1} Institute of Computer Science
Pavol Jozef Šafárik University
Košice, Slovakia
veronika.vanekova@upjs.sk
\textsuperscript{2} Faculty of Mathematics and Physics
Charles University
Prague, Czech Republic
peter.vojtas@mff.cuni.cz

1 Introduction

From its birth in 1965, fuzzy logic has spread to many areas. It has influenced logic, databases, neural networks, expert systems, natural language processing, artificial intelligence and other fields of theoretical research. It has also industrial application in fuzzy controllers.

Fuzzy logic allows us to deal with imprecise, vague or uncertain information. In this paper we point out a different application of fuzzy logic: we use fuzzy sets to represent a notion of user preference in the process of web search. The main source of fuzziness is not in the data itself, but in the imprecise way of human thinking and specifying requirements. We present a system for preferential search and describe our model of user preferences.

2 Towards a System Connecting Web and User

Users search the web nowadays in order to solve problems, not only to find documents. A typical problem is to buy some item online, make a flight reservation, find a transportation line, apply for a job. Most of these problems actually require searching for objects from some domain (e.g. flights, books, notebooks, job offers). The search depends on various attributes of objects (e.g. price, destination, date of departure, airlines and class in the domain of flights).

We describe a system providing access to objects from one arbitrary domain (see figure 1). Objects can be acquired from heterogeneous sources on the web. This phase is called \textit{WIE – Web Information Extraction} (see [3, 4]). The extracted information about objects is stored in a standard database or ontology. The system allows users to find such objects that meet their specified requirements. It is possible to acquire preferences directly from users by means of a special graphical interface. Other option is to let a user rate a small sample of objects and learn preferences from these ratings.
3 User Preference Representation

Let us consider a user looking for a notebook. The information can be extracted from big online stores as well as local advertisements. User can be interested in attributes like memory size, disk capacity, price, screen size, resolution, manufacturer, etc. Different users may prefer different values of these attributes. One user can look for at most 12.1” screen, other can require at least 14.5”. Preference to price is usually “the less the better”, with possible exceptions (e.g. an employee is allowed to spend at most 800$ for a new notebook and he wants to spend it all). Attributes like disk capacity and memory belong to the type “the more the better”, again with possible exceptions.

Preference over one individual attribute is called local preference in our model. We represent local preferences as fuzzy subsets of all possible attribute values. Every attribute value is given a membership degree from $[0, 1]$. If attribute $A_i$ has a domain $D_i$, then $f_i : D_i \rightarrow [0, 1]$ is a fuzzy set membership function for attribute $A_i$. We use mostly monotonic and trapezoidal membership functions. Figure 2 shows an example of non-decreasing fuzzy set membership function $\text{good disk}$. Values from 160 GB higher are fully preferred by the user, while values between 80 GB and 160 GB are not so good and values below 80 GB are unacceptable. Note that every local preference defines a linear ordering of all objects based on their membership values. The ordering generated by the fuzzy set $\text{good disk}$ would preserve weak value ordering of attribute values. Some values are considered equally good, e.g. 160 GB and 161 GB.

User requirements are often conflicting (like looking for both cheap and high performance notebook). Local preferences cheap and high performance based on price and processor speed will create different object orderings. Therefore we global preference (use fuzzy aggregation operators) to acquire overall ordering. Global preference is n-ary monotone aggregation function $u : [0, 1]^n \rightarrow [0, 1]$. If an object $x$ has attribute values $(x_1, x_2, \ldots, x_n)$, and local preferences $f_1, f_2, \ldots, f_n$, then the global preference degree will be $u(f_1(x_1), f_2(x_2), \ldots, f_n(x_n))$. Typical aggregation function is weighted average or classification rules [5]. Weighted average reflects the idea that some attributes can be more important than the others. The following function shows that price is more important than speed:

$$\text{good_notebook}(x) = \frac{2 \cdot \text{cheap}(x) + \text{high_performance}(x)}{3}$$

Classification rules [5] reflect global preferences in more natural way. The following example shows a set of classification rules:
Every rule consists of a head (e.g., $\text{good_notebook}(x) \geq 0.8$) and a body, which is a conjunction of condition clauses (like $\text{good_price}(x) \geq 0.8$). If some object fulfills these conditions, the global preference value will be greater or equal to the value of head. One object may fulfill more rules. In such case the greatest value is taken as a result and all inequalities still hold.

This model is designed to support user-dependent search with top-k algorithm. It is suitable for More details can be found in [6] and [7]. A testing implementation of this system can be found on [8].

4 Sources of fuzziness in web search

As we already stated, the main source of fuzziness is vague statement of user requirements. This process is illustrated on Figure 1 with the arrow “user input”. However, other parts of the searching process may also extend the model with some uncertainty.

Issues connected with WIE process and uncertainty are thoroughly analysed in [1], [2]. Every extraction method can generate a level of uncertainty about its results. Other important process is preference learning – it means learning local and global preferences from a sample of objects rated by user. We can determine the accuracy of learned fuzzy preferences and determine whether we use them in the search process.

5 Conclusion

We present a system for user-dependent web search. User preferences are represented as fuzzy sets and fuzzy aggregation functions. They generate an ordering of objects based on their degrees of preference.
Uncertainty in our model stems in the representation of human thinking and in the process of automatic learning. This model has been implemented and tested [6]. It shows both good performance and very close similarity to real user preferences.

References

8. http://x64.ics.upjs.sk:8080/nazou/