Multistep algorithm of alternatives search in an information catalogue

B. Paliukh, I. Egereva

Tver state technical university, Tver, Russia
e-mail: boris@tstu.tver.ru, irina.egereva@gmail.com

Abstract. This article substantiates the efficient use of fuzzy sets for services search in registers and catalogues, we have set a target and provided an algorithm for multistep decision making under fuzzy conditions.

Nowadays users quite often use different functional subsystems in order to optimize the formation of task solving complexes.

There is a great number of solvers which are barely accessible due to procedural complications.

In order to allow a user to apply current technologies, familiarize themselves with the practice of managerial problem-solving, evaluate the features and their applicability to specific tasks, different catalogues are being created in subject areas, such as: Mathtree - for websites about Mathematics, developed by the Institute For Informatics n. b. A. P. Ershov of Russian Academy of Science; WolframMathWorld project, that allows not only data search but also task-solving; NigmaRF project, the catalogue of Biology services; programmableweb.com and others.

Despite the fact that data catalogues are widely used in different areas, data search that is considered as the major parts of data systematization, is being done with strictly formulated search requests that result in a search results list.

However, quite often it is impossible to state a specific search request, keywords do not match service description tags. It is also impossible to state search criteria specifically. Users have to choose a search area to find a result that matches their task, which they quite often cannot do because they don't know, for example, which subject areas of Mathematics a service developer has used. To avoid restricting the access to fill information available in given catalogues, it is recommended to provide quality system access, also this kind of data is better stated through fuzzy sets.

Our targets are as follows: there is a system that has a catalogue of services for a manager.

A manager, the user of an information system, states a request with a problem that is subject to solve. The system addresses a catalogue set of services in order to make a ranked list of alternatives (figure 1). Also, it is assumed that number of steps is fixed.
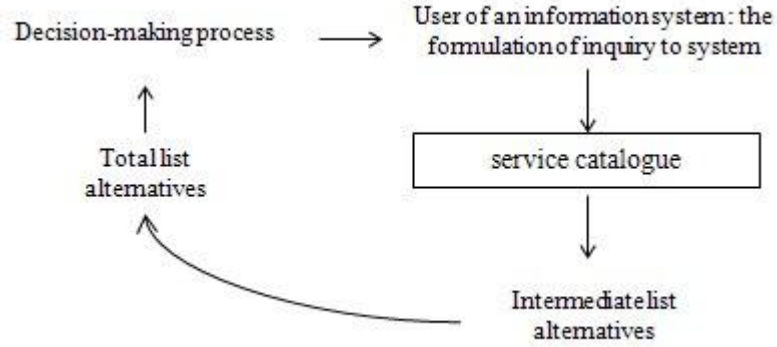
Figure 1. Multistep decision making, general view.

The task is to develop a data management system allowing the formation of a search results list that suits a search request as much as possible providing that number of steps is fixed.

We will address the algorithm of problem solving. If $X=\{x\}$ is a service catalogue that incorporates services $\{x_1, x_2, ..., x_n\}$. A fuzzy set $Y_x=\{y_x\}$ has semantic description of every service in the catalogue. Fuzzy set $H_x=\{h_x\}$ has description of services parameters (figure 2).
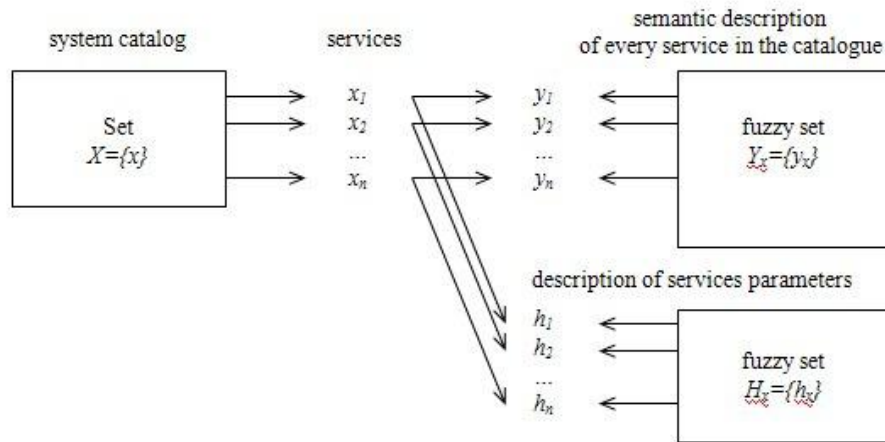

Figure 2. Brief scheme of catalogue structure.

A request (figure 3) is stated semantically and resembles sets: $Z_{x1}=\{z_{x1}\}$ - set of search tags, $Z_{x2}=\{z_{x2}\}$ - extra set of synonyms, $Z_{x3}=\{z_{x3}\}$ - additional set of words similar or close in meaning and concludes the set :

$$Z_x=Z_{x1}\cup Z_{x2}\cup Z_{x3} \tag{1}$$

set of search tags

extra set of synonyms

additional set of words
similar or close in meaning

$Z_{x1} = \{z_{x1}\}$
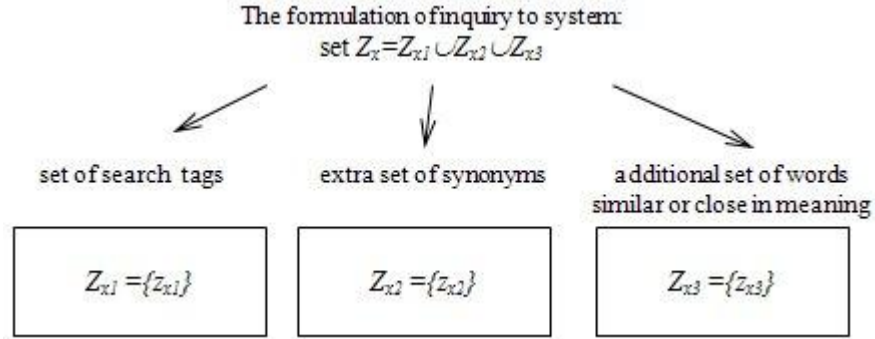
$Z_{x2} = \{z_{x2}\}$

$Z_{x3} = \{z_{x3}\}$

Figure 3. Formation of a semantic set of alternatives

Solving is a set of services D_x formed by a multistep alternatives selection system that best matches the request. Following the Bellman - Zade formula we will represent solving as integration of tasks and limitations.

If we put a set match $I_x = \{i_x\}, \ 0 \leq i_x \leq 1$ including indexed matches of a objects descriptions $Y_x = \{y_x\}$ to requests $Z_x = \{z_x\}$ formed by calculating the percentage of match.

Services with the lowest index of match are not considered, so it is reasonable to limit the set I_x with the desired percentage of semantic match to a description a_x. Using this service identifier in the set $I_x|a_x$ of the set I_x, by limiting the original set $X = \{x\}$ we get the set $X_{I_x|a_x} \subset X$.

Next step is required if the limitation $I_x|a_x$ is not sufficient. Besides standard repeating of the I_x set and the $X_{I_x|a_x}$ set it is also possible to a fine semantic set B_x, on the basis of the semantic set of alternatives Z_x and limitations a_x, which alows not to change the whole request. You only need to input some corrections for the search in a previously formed set. This stage will result in a new set $X_{I_x|a_x|b_x} = X_{I_x|a_x}$.

When selecting the most appropriate service for a given task we will need to use C_x, the system of parameters that assess content of services and their features. Users set primary requirements to services.

After comparing data in H_x, the set of parameters description, with a user's request, the $C_x \cap H_x$ set is formed.

It's also possible to form a ranked list of services from the set $X_{I_x|a_x|b_x} = X_{I_x|a_x}$ in compliance with given parameters. Further $X_{I_x|a_x|b_x}$ is limited by the system of parameters $C_x \cap H_x$.

The result of our search is as follows

$$D_x = X_{I_x|a_x|b_x} \cap C_x \cap H_x, \tag{2}$$

representing the crossing of our targets and limitations.

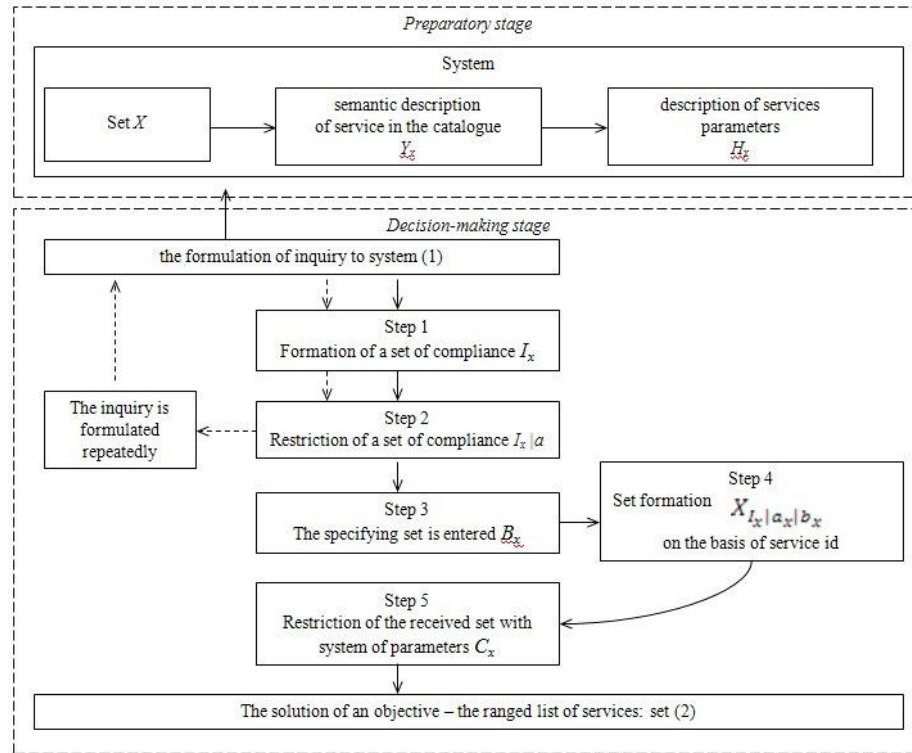Figure 4 shows a multistep decision making process chart.

Figure 4. Flowchart of functioning of system of multistep process decision-making

At the moment Tver State Technical University is developing a data management system capable of assessing innovative projects that will allow analyzing options of resolving a user's task and conclude about applicability of a given service.