Journal of Digital Information Management

DLINE JOURNALS

ISSN: 0972-7272

JDIM 2025: 23 (3)

Fuzzy Ontology-based Identification and Interpretation of Uncertain and Imprecise Novice User Requests Approach

Omri Mohamed Nazih
MARS Research Laboratory, University of Sousse
Department of Applied Computer, National School of Engineering of Sousse, Tunisia
mohamednazih.omri@fsm.rnu.tn

ABSTRACT

Assistance with the use of technical devices is required as soon as the tasks become complex. This assistance is also needed as soon as we provide users with solutions to incidents that occur during the application of unsuitable procedures. The goal of this work is then to provide a knowledge extraction approach that can interpret and identify user requests as valid system requests, thereby responding appropriately to novice user requests. This approach is based on a fuzzy semantic network for modelling imprecise and uncertain knowledge and the automatic construction of a temporary fuzzy ontology for identifying and interpreting user requests. The proposed approach has the advantage of integrating the notion of uncertain and imprecise knowledge into the representation of system objects and procedures. The experimental results show the feasibility, efficiency and effectiveness of our approach.

Subject Categories and Descriptors: **[I.5 PATTERN RECOGNITION];** Fuzzy set [I.2.4 Knowledge Representation Formalisms and Methods]

General Terms: Fuzzy Ontology, Novice users, Semantic representation, Uncertainty

Keywords: Uncertainty, Imprecision, Knowledge Extraction, Fuzzy Semantic Network, Fuzzy Ontologies, Fuzzy Requests

Received: 19 March 2025, Revised 10 June 2025, Accepted 23 June 2025

Review Scale: 0/6, Review Score: 4.22, Inter-reviewer consistency: 72%

DOI: https://doi.org/10.6025/jdim/2025/23/3/171-189

1. Introduction

The notion of relevance is a widely held opinion, a central concept in artificial intelligence. From a cognitive perspective, relevance is also described as a key concept, and some even consider it a principle that governs our interactions with the environment or the way we communicate. The information exists, but it is necessary to be able to extract it from gigantic databases, and the recurring question is: what information is relevant?

More generally, it is understood that any problem-solving presupposes the selection of the information to be taken into consideration, even before its use in any reasoning, because taking into account all the information would make the reasoning totally ineffective.

Likewise, for systems of assistance that utilize technical devices, the concept of relevance is fundamental. Indeed, the design of these systems requires the processing of uncertain and imprecise user requests. Their performance is therefore measured by the relevance of the knowledge returned to the user.

The performance of today's conventional help systems is satisfactory in terms of both the quality and quantity of information they contain. However, their weak point lies in identifying and interpreting user requests expressed in natural language. In fact, in most cases, the user's request does not contain the terms that index the object and specify the action to be performed from the system's point of view. It is a problem of identification and correspondence between the label of the object concerned and that of the goal to be achieved contained in the request, as well as their corresponding entities in the knowledge base (KB).

The performance of today's conventional help systems is satisfactory in terms of both the quality and quantity of information they contain. However, their weak point lies in identifying and interpreting user requests expressed in natural language. In fact, in most cases, the user's request does not contain the terms that index the object and specify the action to be performed from the system's point of view. It is a problem of identification and correspondence between the label of the object concerned and that of the goal to be achieved, as well as the corresponding entities in the KB.

The goal of this work is to design a knowledge extraction model that can interpret and identify user requests as valid system requests, enabling the system to respond appropriately to novice user requests expressed in natural language. We are therefore interested in representing inaccurate and uncertain knowledge, as well as searching and extracting relevant knowledge from a semantic network.

This work is organized as follows. Section two presented the related works. In the third section, we present the Fuzzy sets, Linguistic Variables, and the Fuzzy Ontologies used in this contribution. The fourth section presents the architecture of our knowledge extraction approach, including its various components. In the fifth section, we detail the various tests and performance measures conducted, and present the results of the experiment. We will conclude and present the future perspectives for this work.

2. Background

Executing a procedure serves to reach a Goal on an Object. The underlying psychological hypothesis is that these Goals are Object properties and, as such, are generators of Object categories. Goals and procedures define the function of Objects and the way to use them. As functional properties of Objects, they enter into the construction of semantic networks in the same way as structural properties. We define a procedure as a sequence of operations whose execution serves to reach a Goal, and where the elements of the sequence are either primitive actions or subGoals which themselves call for associated procedures. Technical devices can be described through procedures in two steps. The first one is task decomposition, which is a hierarchy of Goal decomposition into subgoals, ranging from the level of the task's goal to primitive actions. The second one consists of (1) drawing up a list of possible Goals and the procedures to reach these Goals and (2) constructing the Ideal Expert Net as a classical semantic network.

Several algorithms have been proposed to address the proposed problem. [Shanshan, Meng, Yingdong, He, 2018] proposes new decision-making methods for the selection of cold chain logistics enterprises under an intuitionistic fuzzy environment based on generalized information aggregation operators. The proposed method evaluates the scaled prioritised relationships between criteria by priority labels in known and unknown situations. In [Ganeshsree Selvachandran, Madhumangal Pal, Tahani A. Abdusalam Alhawari, Abdul Razak Salleh., 2018], the authors introduce two new operations for the interval-valued complex fuzzy set model and study the fundamental algebraic properties of these new operations. The utility and applicability of the relations of this model are then demonstrated by applying it to an economics problem. [Sujit Das, Debashish Malakar, Samarjit Kar, Tandra Pal., 2018] consider that decision making using fuzzy soft-set and its extensions has become the most significant research area in the age of uncertainty. The evolution of fuzzy soft-sets during the last decade and a half (2001-2015) to analyze the impact of fuzzy soft-sets and their extension in the decision-making paradigm. Based on a selection of journals, conferences, and online databases, they categorize the decision-making process into ten distinct categories, which are based on various types of fuzzy soft-sets. They briefly explore each individual category by mentioning the theoretical/algorithmic approaches proposed by the respective authors. Furthermore, all papers are categorised with respect to publication year, published journal, application type, and decision-making criteria. This literature survey provides a platform for researchers to identify the dimensions of future research in fuzzy soft-sets by analyzing the current state and potential areas. In [T. R. Sooraj, B. K. Tripathy., 2018] authors consider that selection is a challenging task due to the presence of hundreds of varieties of seeds of each kind. Some homework is necessary for selecting suitable seeds, as new varieties and types of seeds are introduced to the market every year, each with its own strengths and weaknesses. The complexities involved in the characteristics, expressed as parameters, result in uncertainties. Consequently, some uncertainty-based models or hybrid models are required to model the scenario and make a decision. Soft sets have enough of parameterization tools to support and hence is the most suitable one for such a study. The authors proposed a model called the interval-valued fuzzy soft set and developed a decision-making algorithm for seed selection. And they concluded that the use of signed priorities and intervals for the membership of values for entities makes the study more efficient and realistic. In [W. Chebil; L. F. Soualmia; M. N. Omri., 2018] authors have proposed a new approach titled Conceptual Information Retrieval Model (CIRM). The principle of this contribution is the exploitation of possibilistic networks (PNs) and a multi-terminology approach to extract and disambiguate terms, and then retrieve documents. The two measures of possibility and necessity were used to select the relevant concept of an ambiguous term. Thus, the user request and unstructured documents are described throughout a conceptual representation. Concepts were then filtered and ranked. Finally, a possibilistic network was exploited to match documents and requests. Two biomedical terminologies were utilised, namely the MeSH thesaurus (Medical Subject Headings) and the SNOMED-CT ontology (Systematised Nomenclature of Medicine of Clinical Terms).

3. Fuzzy Sets, Linguistic Variables and Ontologies

Crusty sets are the sets that we have used for most of our lives. In a crusty set, an element is either a member of the set or not. For example, a jelly bean belongs in the class of food known as candy. Mashed potatoes do not.

Fuzzy sets, on the other hand, allow elements to be partially in a set. Each element is given a degree of membership in a set. This membership value can range from 0 (not an element of the set) to 1 (a member of the set). It is clear that if one only allowed the extreme membership values of 0 and 1, then this would actually be equivalent to crisp sets. A membership function is the relationship between the values of an element and its degree of membership in a set.

3.1 Linguistic variables: from crisp sets to fuzzy sets

Fuzzy logic is an extension of classical logic where truth propositions can take any number between 0 and 1 to model imprecise and partially true statements. Classical logic describes a binary truth proposition, where a proposition is true if an element e is a member of the set A of interest (i.e. $e \in A = 1$), and a proposition is false if an element e is not part of set A (i.e. $e \in A = 0$). These binary sets are said to be crusty, and they are represented by a step membership function μ_A, which maps the element e to the set A if A is completely representative of e. In Figure 1, the membership function assigns heights below six feet to the set 'not tall' and assigns heights above six feet to the set 'tall'. A five feet ten inches person is classified as not tall, and a six feet two inches person is classified as tall. Formally, a crusty set, or classical set A, is defined as follows.

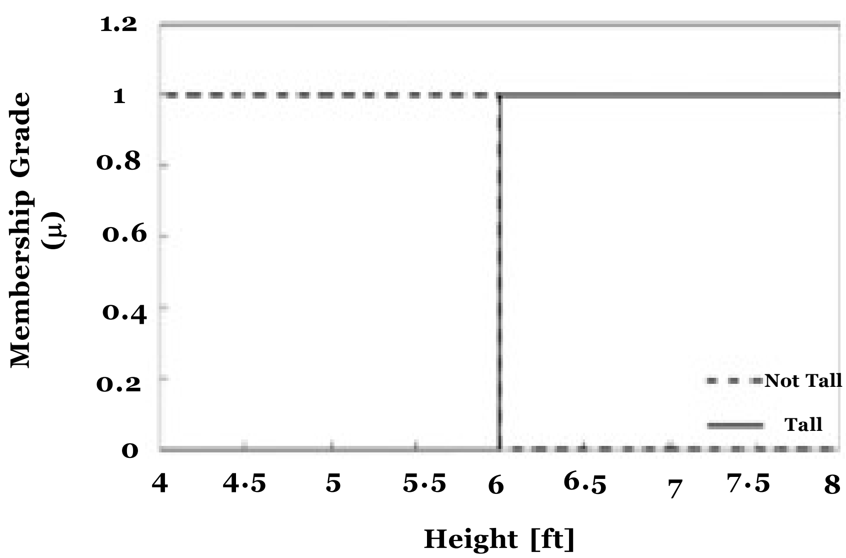

Figure 1. Crusty membership function describing human height

Definition 1.

A crisp set is comprised of a domain X of the real numbers (also called the universe of discourse of A) together with a membership function $\mu_A(x)$, where $\mu(x)$ is given by the indicator function in equation 1.

$$\mu_A(x) = \{1 \; if \; x \in A, 0 \; else \qquad (eq\ .\ 1)$$

Fuzzy set theory expands this concept to allow the degree of membership of an element to a set to lie between 0 and 1 (i.e the membership function is not restricted to an indicator function). Thus, a classical set is a special case of a fuzzy set. Two classes of fuzzy sets exist—first-type fuzzy sets and second-type fuzzy sets—which differ by their ability to handle linguistic uncertainties. A first-type fuzzy set is defined by a single deterministic, or crisp, membership function. Once it is defined, all linguistic uncertainty about the membership function disappears. In the second type, fuzzy sets membership functions are themselves uncertain, or fuzzy, and can take any value within their defined bounds.

First Type of Fuzzy Sets

The first type of fuzzy sets is the most widely used and has been successfully applied to many real-world problems. They model linguistic uncertainty using a crisp membership function, which determines the similarity between a numerical variable (i.e., element e) and a linguistic variable. Formally, a fuzzy set is defined by

Mendel et al. (2014) as follows.

Definition 2. A fuzzy set A is comprised of a domain X of the real numbers (also called the universe of discourse of A) together with a membership function μ_A: X→ [0, 1]. For each x ∈ X, the value of μ_A(x) is the degree of membership, or membership grade, of *x in A*.

If $\mu_A(x)=1$ or $\mu_A(x) = 0 \forall, x \in X$, then the fuzzy set is said to be a crisp set.

A First type of fuzzy set A is described as in Equation 2, where the integral operator denotes the collection of all values x in the universe of discourse X, with degree of membership μ_A (x). Additionally, the slash does not denote division. It associates the value x to a degree of membership in set A.

$$\mu_A(x) = \int_{x in A} \mu_A(x)/x \qquad (eq . 2)$$

The non-binary degree of membership of a numerical variable x to the set A better captures the nuances of real life and the imprecisions associated with human reasoning. The sharp boundary used by the crusty membership function to categorise human height, given in Figure 1, is unnatural, as there is no sharp boundary describing a change from tall to not tall. Instead, humans tend to think of a gradual change from the height of a tall person to the height of a person who isn't. The fuzzy membership function shown in Figure 2 captures this imprecise boundary and gradual change. The five-foot-ten-inch person and the six-foot-two-inch person are both about half tall and half not tall, with the first being slightly less tall and the second being slightly taller.

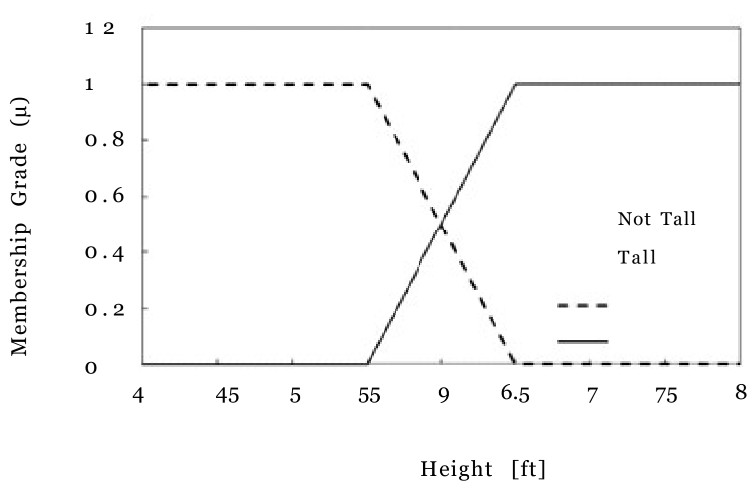

Figure 2. Example of fuzzy membership function describing human height

Second Type of Fuzzy Sets

The second type of fuzzy sets expands its first type counterpart to incorporate uncertainty in its membership function definition. A major criticism of the first sets is that they are themselves non-fuzzy, which limits their ability to handle uncertainty in the definition of their membership functions. Words have different meanings to different people and are also context-dependent. This concept cannot be adequately captured by first-type sets. Figure 2 illustrates a precise definition of the membership functions for tall (i.e., there is no uncertainty associated with its end points); however, a person's definition of the concept 'tall' tends to be vague and depends on several factors, such as nationality and age. Using a first type set with an exact membership function seems counterintuitive when describing a fuzzy quantity, which is by definition imprecise. The second type of sets, or fuzzy-fuzzy sets, provides a more accurate model of linguistic uncertainty by using a fuzzy definition of the primary membership function (Figure 3). They are defined by their primary membership μ, which is plotted in the plane of the paper, and their secondary membership u, which is plotted on the vertical axis.

The shaded area is referred to as the footprint of uncertainty (FOU) of the primary membership function, and is bounded below by the lower membership function (LMF) and bounded above by the upper membership function (UMF). The even shading implies that all primary memberships between the LMF and UMF are equally likely and thus uniformly weighted. Such sets are denoted as interval, the second type of fuzzy sets, and make up virtually all applications of the second type of fuzzy sets. Fuzzy sets with non-uniform secondary membership functions are called general of the second type of fuzzy sets, but are still in infancy. If all uncertainty about membership function definition disappears, the secondary membership *u* reduces to μ and a second type of set reduces to a first type set, much like a random variable reduces to the mean when the variance goes to zero.

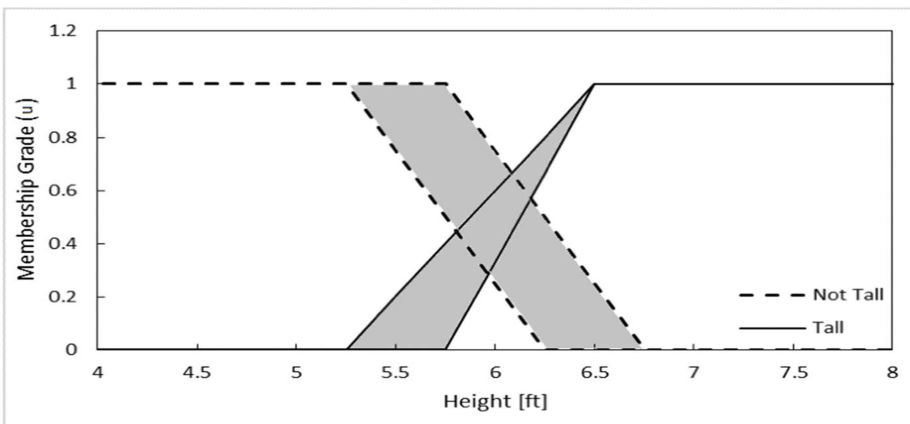

Figure 3. Example of fuzzy membership function describing human height

One can also think of a second type of fuzzy set as a collection of an infinite number of embedded first-type fuzzy sets contained within the shaded area and bounded between the LMF and UMF. For interval second-type sets, each of these embedded functions is equally weighted.

A second type of fuzzy set A can be described as in Equation 3, where the first integral operator denotes the collection of all values *x* in the universe of discourse *X*, and the second integral denotes the collection of all secondary membership $u \in U \subseteq [0, 1]$ of its embedded first-type sets.

$$\mu_A(x) = \int_{x \in X} \int_{u \in U_x} \mu_A(x, u)/(x, u) \qquad (eq\ .\ 3)$$

$$\mu_A(x) = \int_{x \in X} \int_{uin[0,1]} 1/(x, u) = 1/FOU(A) \qquad (eq\ .\ 4)$$

3.2 Fuzzy Ontologies for Extracting Relevant Information

We represent here networks used in the search for information that allow us to represent associative relations between the different terms. Thesauri allow, for example, the extension of a query to synonym terms. One can either consider binary relations between terms or consider fuzzy relationships as Radecki [Radecki, 1976] proposed. In the latter case, we speak of fuzzy ontologies. Lucarella and Morara [Lucarella & Morara, 1991] present an example of an information search system based on networks of fuzzy concepts. They consider a network of concepts in which the links are oriented and valued, and in which nodes representing the documents also appear (figure 4). Link valuation represents the degree to which the concept or document being pointed to is relevant to the node from which it is linked.

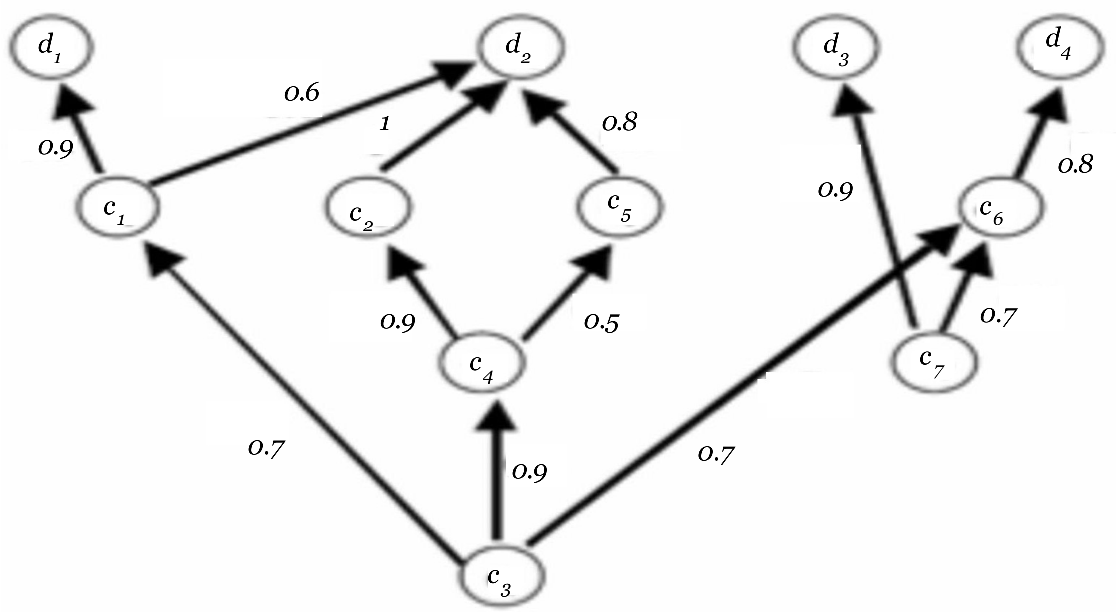

Figure 4. Example of network's fuzzy concepts

The concepts are usually words (the key words chosen for the representation of documents), and the automatic construction of ontologies most often consists of measuring the co-occurrence of words in the different documents. So when two words appear in the same documents, they are perfectly linked.

4. Proposed Approach

The example of a technical approach that we consider is an aeroplane piloting system, whose objects include the aeroplane, joystick, throttle, and procedures include takeoff, turning left, or lowering.

The model we proposed offers the user the ability to enter a user request in the following form: "How to take off the craft?" The extraction model that we propose allows building a valid system request, such as" how to take off the plane? "equivalent to the initial user request.

This analysis of the user request is based on the search and retrieval of relevant knowledge from the fuzzy semantic network that forms the core of the KB. It also uses fuzzy ontology construction to estimate the relevance of user requests against equivalent system requests (Figure 5).

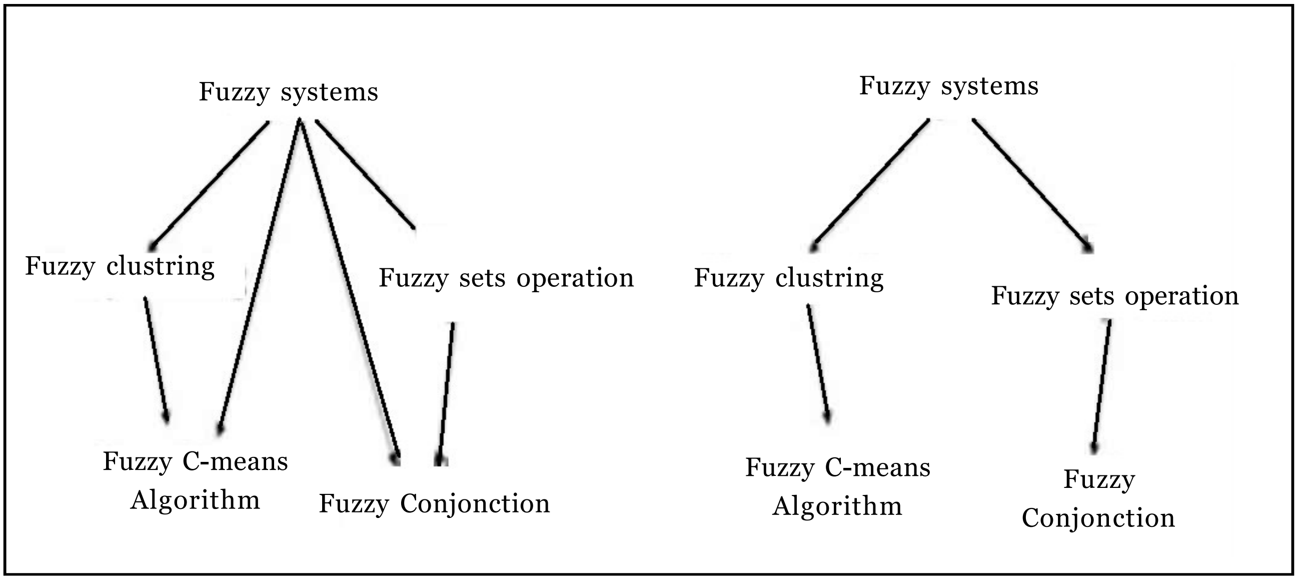

Figure 5. Example showing the construction of ontologies

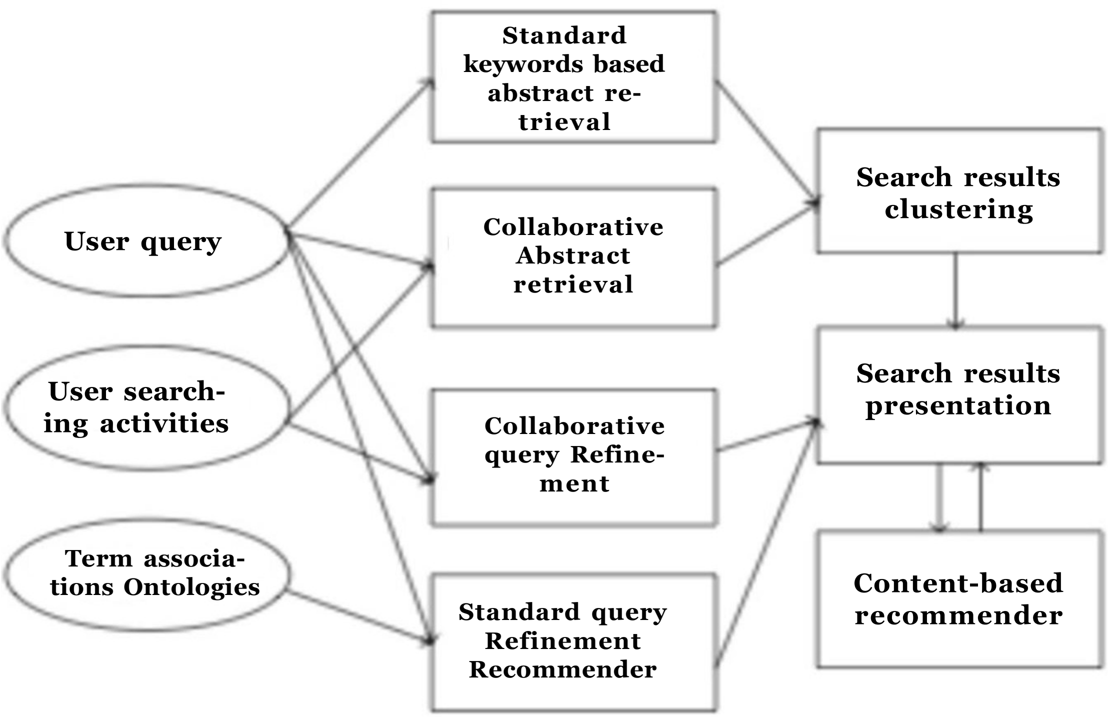

Figure 6. General architecture of the proposed approach

4.1 Structures of System Objects and Procedures

Structures of the System Objects

Let O_S be a system object and O_U be the user object. We associate with each object of the system a list of users Z_{Os} extracted from novice user requests expressed in natural language. For each object in the list, we associated a degree of belonging $\mu_{Oui/OS}$, *which reflects the degree to which we believe that the semantic meaning of* O_{ui} is near to the semantic meaning of the targeted system object O_S.

$$Z_{Os} = \{(O_{u_1}, d_{11}), (O_{u_2}, d_{12}), \ldots, (O_{u_n}, d_{1n})\}$$

EXAMPLE:

$$Z_{air\text{-}plain} = \{(machine, 0.5), (device, 0.6), \ldots, (engin, 0.2)\}$$

Structure of the System Procedures

As with system objects, we have associated with each system procedure belonging to a system class a "user zone" consisting of a list of procedures used by users to designate the system procedure and the degrees with which we believe these procedures have been performed. A semantic meaning close to the meaning of the intended system procedure.

EXAMPLE:

$$Z_{remove} = \{(start, 0.3), (move, 0.6), (leave\text{-}the\text{-}ground, 0.8)\}$$

To enhance the performance of search engines on the Internet, new methods are implemented that not only look for terms present in the request, but also allow for finding other terms that are related to those being sought. These methods use what is called "fuzzy ontologies"

Ontology Construction Process

The principle of our approach algorithm for the automatic construction of fuzzy ontologies is based on two principal steps. Let us give some definitions that will allow us to construct the ontology.

Definition 3. Let C = $\{a_1, a_2, \ldots, a_n\}$ be a collection of articles a_i and each article $a_i = (t_1, t_2, \ldots, t_m)$ is represented by a set of terms *tj*. Let $\mu_{Occ}(t_{ij}, a_i)$ be the function which characterizes the occurrences of term t_{ij} in article a_i and NT(t_i, t_j) be the function which expresses *Narrow-than* t_j. We defined the belonging degree $\mu_{NT}(t_i, t_j)$ by the next equation (eq. 5).

$$\mu_{RT}(t_i, t_j) = \frac{\sum_{a \in C} min}{\sum_{a \in C} max} \qquad \text{(eq. 5).}$$

$\mu_{NT}(t_i, t_j) \approx 1$ when t_i and t_j are present together in the articles or when the occurrences of t_i are much smaller than those of t_j.

Definition 4. Let $\mu_{BT}(t_i, t_j)$ be the *broader-than* function that denotes t_i is wider than t_j. Because *"broader" is the opposite of "narrower," this measure is given by the following* equation (eq. 6).

$$\mu_{BT}(t_i, t_j) = \mu_{NT}(t_j, t_i) \qquad (eq.\ 6)$$

Definition 5. The third relationship, called *Related-Term* (RT) is defined by the next equation (eq. 7).

$$\mu_{RT}(t_i, t_j) = \frac{\sum_{a \in C} min(}{\sum_{a \in C} max(} \qquad (eq.\ 7)$$

First Step of Ontology Construction Process

The first step is to develop a comprehensive ontology based on NT relationships. For each pair of terms *(t_i, t_j)*, we compute the membership degrees of two NT relations, $\forall t_i, t_j \in a_i \,/\, t_i \neq t_j$ calculate $\mu_{NT}(t_i, t_j)$ and $\mu_{NT}(t_i, t_j)$.

We keep the NT relation which has the highest degree, *max* ($\mu_{NT}(t_i, t_j)$, $\mu_{NT}(t_i, t_j)$) and we eliminate the other NT relationship. In this way, the NT redundant relationships are eliminated and the stored information is halved.

When the relationships between the terms are not strong, the values of their NT relations have low degrees and, therefore, cannot be incorporated into the ontology. One fixes a certain limit a, then applies the (a-cut), which consists of removing all the NT relations that are below this limit. If an NT relation has a degree equal to zero $\mu_{NT}(t_i, t_j) = 0$, then the two terms are unrelated.

The membership degrees assigned to user terms in the "zones" are based on corpus analysis.

Second Step of Ontology Construction Process

This step involves reducing the ontology by removing all non-essential NT relations. Figure 3 illustrates two ontologies: one with excessive relationships and the other with normal relationships. From this, it can be observed that two terms can have a relation either directly or indirectly via other terms, referred to as "indirect paths". We begin by finding all the indirect paths between the two terms t_i *and* t_j.

Definition 6: Let P be one of these paths. $P = \{NT(t_i, t_{m1}), NT(t_{m1}, t_{m2}), \ldots, NT(t_{mn}, t_j)\}$. NT(P). *NT(P)* is assumed to be an *NT* alternative relation for (t_i, t_j). The degree of adhesion of *NT(P)*, which is defined by the next equation (eq.8) as follows:

$$\mu_{NT}(P) = \min\{NT(t_i, t_{m1}), NT(t_{m1}, t_{m2}), \ldots, NT(t_{mn}, t_j)\} \qquad \text{(eq. 8):}$$

4.2 Identification Algorithm of User Objects and Procedures

The identification process involves returning the most appropriate system object to the user object, expressed in natural language. The idea we propose is not only to exploit the degree of belonging of the user object of the request to the different system objects, but to build a S consisting of objects fuzzy ontology to connect this object with other domain objects. The goal is to find the most relevant path for linking the user object to the target system object.

Different Steps of the Algorithm

Let o_{Ux} be the user object contained in the request, o_{Si} be the system object, o_{Ui} be the user object, and Z_{OSi} be the user area of the system object o_S.

The first step of the algorithm is to extract the user objects that will be added to the ontology. These objects must check the condition $\mu_{o_{Ui}/o_{Sk}} > \mu_{o_{Ux}/o_{Sk}}$ and o_{Ux} et $o_{Ui} \in Z_{o_{S_k}}$. This condition is necessary to select only the user objects that have a narrower meaning than the initial object of the user. The extracted objects are added to the $L = \{o_{Ux}, o_{U_1}, o_{U_2}, \ldots, o_{U_m}\}$.

The second step of the algorithm consists of two different phases. In the first phase, we will build the ontology from the list L of user objects using the NT relationship (narrower than) between the two objects o_{Ui} *and* o_{Uj}. The narrower than degree between two terms t_i *and* t_j noted $\mu_{NT}(o_{Ui}, o_{Ui})$ is defined by the following equation (eq. 9):

$$\mu_{NT}(t_i, t_j) = \frac{\sum_k min(}{\sum_k \mu OUi, OSk} \qquad \text{(eq. 9):}$$

With $(\mu_{NT}(O_{Ui}, O_{Uj})$ is the relevance degree, which means that the object o_{Ui} is more relevant than o_{Uj}. In this step, the relevant system objects are added to the ontology to get the final fuzzy ontology.

In the third step, we proceed to measure the relevance degree of the different system objects with respect to the user object to build the list of equivalent system objects sorted in order of decreasing relevance. The input parameter of our algorithm is therefore a user object (or a procedure), and the output is a set m of objects (or procedures) systems sorted by degree of relevance in decreasing order.

Building of the equivalent system requests

The construction of the equivalent system requests proposed to the user is based on the identification algorithm (see Algorithm 1). After identifying the object and the user procedure, the algorithm for constructing equivalent system requests proposes to the user, based on the set of identified objects and procedures, a set of possible requests associated with degrees of relevance.

Algorithm. Construction of the ontology

Begin
Step 1: /*extraction of user objects*/
$L=\{O_{Ui}\}$
for each $Z_{O_{Si}}$, with $O_{Ui} \in Z_{O_{Si}}$ **do**
if $(O_{Uj} \in Z_{O_{Si}}$ and $\mu_{NT}(O_{uj}, O_{Si}) \geq \mu_{NT}(O_{Ux}, O_{Si})$ **then** $L = L \cup \{O_{uj}\}$
Step 2 : /*Construction of the ontology*/
Phase 1: /*adding user objects to the ontology*/
ontology= {}
for each O_{ui}, $O_{Uj} \in L$, and $O_{Ui} \neq O_{uj}$ **do**
Calculate $NT(O_{ui}, O_{uj})$ and $NT((O_{uj}, O_{ui})$
Select $(O_{Ui},O_{Uj})=\max\{\mu_{NT}(O_{Ui}, O_{Uj}), \mu_{NT}(O_{Uj}, O_{Ui})\}$ and $\mu(O_{Up}, O_{Uq}) \geq \alpha$
Add $\{NT(O_{Up},O_{Uq}), \mu_{NT}(O_{Up},O_{Uq})\}$ to Ontology
Phase 2 : /*adding system objects to the ontology*/
for $\forall\, O_{Ui} \in L$ **do**
if $\nexists\, O_{Ui} \in L$ such that $\{NT(O_{Ui},O_{Uk}), \mu_{NT}(O_{Ui},O_{Uk})\} \in$ Ontology **then**
for each O_{Sj} such that $O_{Ui}, \in Z_{sj}$ **do**
Add $\{NT(O_{Ui}, O_{Uk})), \mu_{NT}(O_{Ui}, O_{Uk})\}$ to Ontology
if $\exists\, O_{Sj}$ such that $O_{Ui} \in Z_{O_{Sj}}$ **and** $\nexists\, S_{O_{U_i}} \longrightarrow_{O_{Sj}}$ **then**
Add $\{NT(O_{U_i}, O_{S_j}), \mu_{NT}(O_{Ui}, O_{Sj})\}$ to Ontology
Step 3: /*Computation of the relevance degrees*/
Let $S_{O_{U_x}} \longrightarrow_{O_{S_i}}$ be the maximum path that binds O_{Ux} to O_{Si}
$Pert(O_{Ux}, O_{Si}) = \frac{\sum_{i=1,n}(}{n}$
end.

$\nexists S_{O_{U_i} \longrightarrow O_{Sj}}$: means that there is not a path that binds O_{ui} *to* O_{Sj} and *n*: is the number of lies in attributes to the link.

Example

Let's take the example where the user request is: How to start the craft? We seek to identify the system object that corresponds to the user object "engin", for this reason we will use the proposed identification algorithm.

The different user areas of the objects *airplane*, *motor* and *steering wheel* are defined as follows.

$Z_{air\text{-}plain}$ = {(engin, 0.5), (apparatus, 0.6), (machine, 0.2)}

Z_{motor} = {(turbine, 0.4), (engin, 0.5), (machine, 0.6)}

$Z_{steering\ wheel}$ = {(device, 0.6), (joystick, 0.4), (machine, 0.2)

• Step 1: Extract the user objects

Initially we have $L = \{o_{ui}\} = \{machine\}$. Since *engine* belongs to the user zones Zaircraft and Zmotor, we will add the user objects, that have a membership degree *greater than* or equal to *engin,* to *L*.

$L = \{engin, device, machine\}$

• Step 2: Build the fuzzy ontology

Phase 1: Adding User Objects to the Ontology

In this step, we will build the fuzzy ontology from the constructed list *L*, and we will use equation (eq. 2).

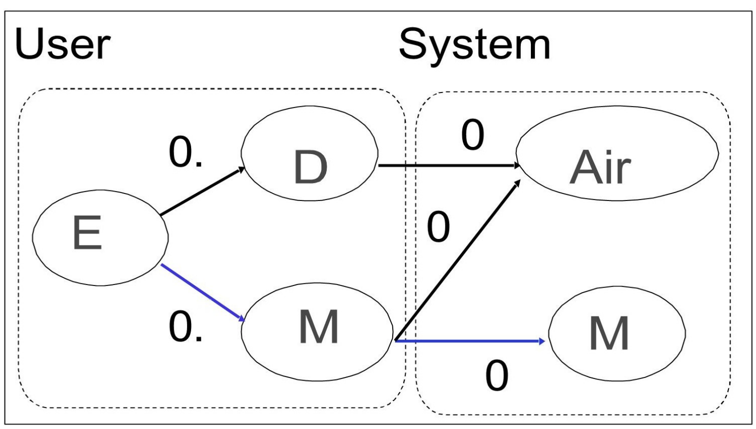

Figure 7. final fuzzy ontology obtained from the engin user object

$$\mu_{NT}(engin, device) = \frac{min(0.5,0.6) + min(0.5,0)}{0.5 + 0.5} = 0.5$$

$$\mu_{NT}(device, engin) = \frac{min(0.6,0.5) + min(0,0.5)}{0.6} = 0.83$$

$$\mu_{NT}(engin, machine) = \frac{min(0.5{,}0.2) + min(0.5{,}0.2)}{0.5 + 0.5} = 0.8$$

$$\mu_{NT}(machine, engin) = \frac{min(0.2{,}0.5) + min(0.6{,}0.5)}{0.2 + 0.6} = 0.87$$

$$\mu_{NT}(device, machine) = \frac{min(0.6{,}0.2) + min(0{,}0.6)}{0.6} = 0.33$$

$$\mu_{NT}(machine, device) = \frac{min(0.2, 0.6) + min(0.6, 0)}{0.2 + 0.6} = 0.25$$

Phase 2: Adding system objects to the ontology

We notice that the user object and *machine* do not have NT relations to other user objects of *L*. They can therefore be linked to the *aeroplane* and *engine* system objects since they appear in their zones. We obtain the final ontology (see Figure 7).

• Step 3: Counting of the Relevance Degrees

The relevance of a system object o_{Si} with regard to a user object o_{Ux} is calculated as follows:

$$Pert(S_{engin,airplaine}) = \frac{0.83 + 0.6}{2} = 0.715 \text{ and } Pert(S_{engin,motor}) = \frac{0.8\bar{7} + 0.6}{2} = 0.73$$

5. Experimental Study and Results Analysis

To evaluate the performance qualities of our approach, we considered a number of relevant and irrelevant requests that were rendered during the analysis of a user's requests. We used three standard relevance coefficients: recall, precision, and F-measure. Suppose a system has returned ten (10) objects (or procedures) in response to a request. Suppose also that a judge has been shown and that he considers that only five (5) objects (or procedures) are relevant. We will say that the system has an accuracy of 50%. This would work perfectly, except for a minor issue. This is due to the fact that only five (5) objects (or procedures) are relevant throughout the collection. Finding the five objects (or procedures) would be a perfect result, and should not be seen as just 50% success. Therefore, one needs to define another measure as the ratio of relevant objects (or procedures) found in the collection to the total number of relevant objects (or procedures) throughout the collection.

For this measure, our example will be 5/5, so 100%. This measure is commonly referred to as a reminder, whereas the first one measures the ratio of relevant objects (or procedures) found to the total number of objects (or procedures) found, and is called precision.

Precision and recall measurements are obtained by partitioning all the objects (or procedures), restored by the system into two categories: the relevant objects (or procedures) and the irrelevant objects (or proce dures).

The recall coefficient measures the system's ability to retrieve all relevant system requests in response to the initial request of the novice user. It is given by the ratio of the relevant found system requests to the set of system requests in the KB.

Definition 7. Let R be the number of system requests in the KB, and R+ be the number of relevant system requests returned from the KB during the analysis of a user request (q_i). We defined the recall coefficient by the following equation (eq. 10).

$$Recall(q_i) = \frac{R_+}{R} \quad (eq.\ 10)$$

Precision measures the system's ability to reject all irrelevant system requests in response to a user's request expressed in natural language. It is given by the ratio between the set of relevant selected requests and the set of selected system requests.

Definition 8. Let R+ be the number of relevant system requests returned from the KB during the analysis of a user request q_i, and M be the number of system requests returned from the KB that are relevant to the user request q_i. We define the precision degree of the request q_i by the following equation (Eq. 11).

$$Precision(q_i) = \frac{R_+}{M} \quad (eq.\ 11)$$

Definition 9. Let Precision and Recall be respectively the precision degree and the recall degree of the request q_i. We define the F-measure coefficient of the request qi as a measure that combines precision and recall, which is their harmonic mean, by the following equation (Eq. 12).

$$F-measure(q_i) = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (eq.\ 12)$$

To evaluate the performance of the model, we have developed our own KB for the flight control system from the semantic network structure inspired by flight simulation software (flight simulator, 2004) and validated by an experienced pilot of the air force (Table 1).

Data	Number
System objects	26
System procedures	38
System attributes	10
Possibilistic system requests	38

Table 1. Characteristics of the test KB

We have prepared a set of user requests that we will apply to our approach (Table 2).

Data	Number
User requests	30
Requests with known objects and unknown procedure	8
Requests with unknown objects and known procedures	8
Requests with unknown objects and procedures	14

Table 2. Characteristics of user requests

To study the behavior of the proposed approach, following is the insertion of new objects by learning, we tested user requests on three sets (see Table 3).

Data Base	Number of user's objects	Number of user's procedures
First data set: Set 1	26	36
Second data set: Set 2	52	72
Third data set: Set 3	78	108

Table 3. Characteristics of Data Bases used sets of user requests

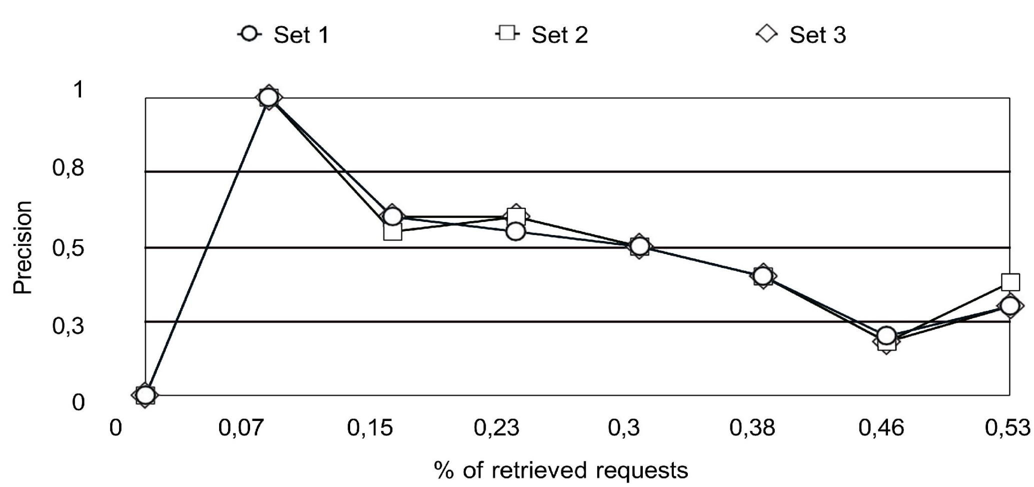

Figure 8. Measurement of the precision coefficient

We have measured the average accuracy of our extraction model for the different sets defined in Table 3. We have observed that our model achieves good precision (> 0.4) when the number of proposed requests is low, specifically when the percentage of retrieved requests falls within the range [0.07, 0.3]. We have also observed that when the number of extracted requests is high, the accuracy of the proposed approach decreases, yet it remains above 0.1, ensuring that relevant requests are always among those proposed to the user (see Figure 8).

We also observed an improvement in the accuracy of the proposed approach for the curves of data set 1 and data set 2, especially when the number of requests selected by our extraction model is high. We explained this phenomenon by the fact that the enrichment of the KB allowed our model to perfect the extraction process.

We measured the recall rate of our proposed model for the different sets of the KB. Figure 9 illustrates that the recall rate varies within the range of [0.45, 1]. This indicates that the majority of system requests equivalent to the initial request of the novice user were selected and extracted by our extraction model. Notably, for the curves of sets 2 and 3, we observe that the recall rate exceeds 0.6.

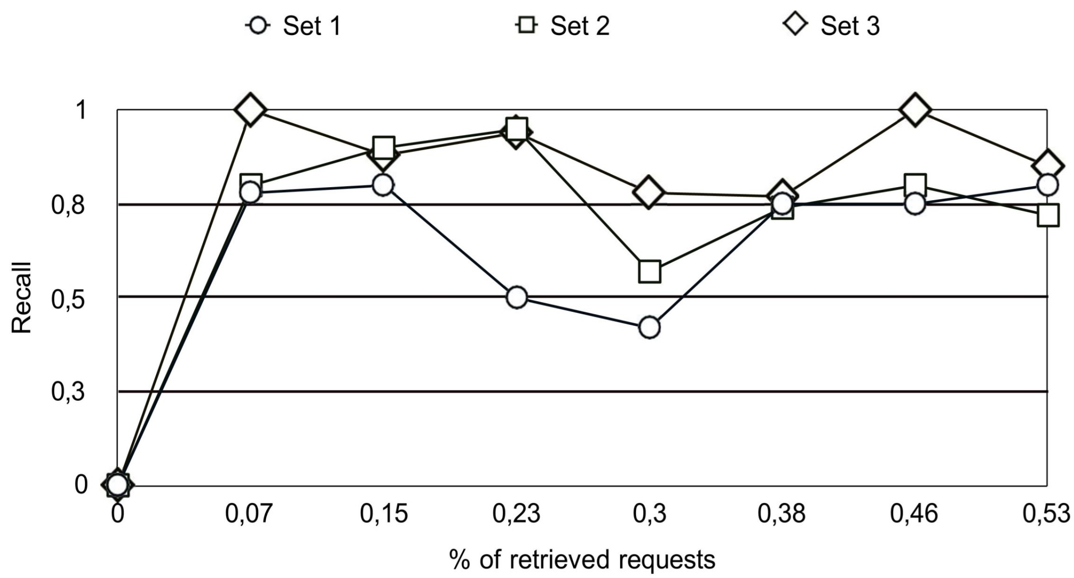

Figure 9. Measurement of the recall coefficient

We also measured the F-measure rate to prove the efficiency of our proposed model for the different sets of the KB. Figure 10 shows that the F-measure rate varies in the range [0.32, 1]. These values indicate that the majority of system requests, equivalent to the initial request of the novice user, were selected and extracted by our extraction model. Notably, for the three curves, we observe that the recall rate exceeds 0.6.

6. Conclusion

In this paper, we present a relevant knowledge extraction approach based on the automatic construction of a temporary fuzzy ontology for identifying and interpreting user requests. The fuzzy ontology enabled us to automatically construct equivalent system requests and estimate their relevance. Additionally, we can demonstrate that it also enables our model to refine the initial user request by presenting the user with a list of

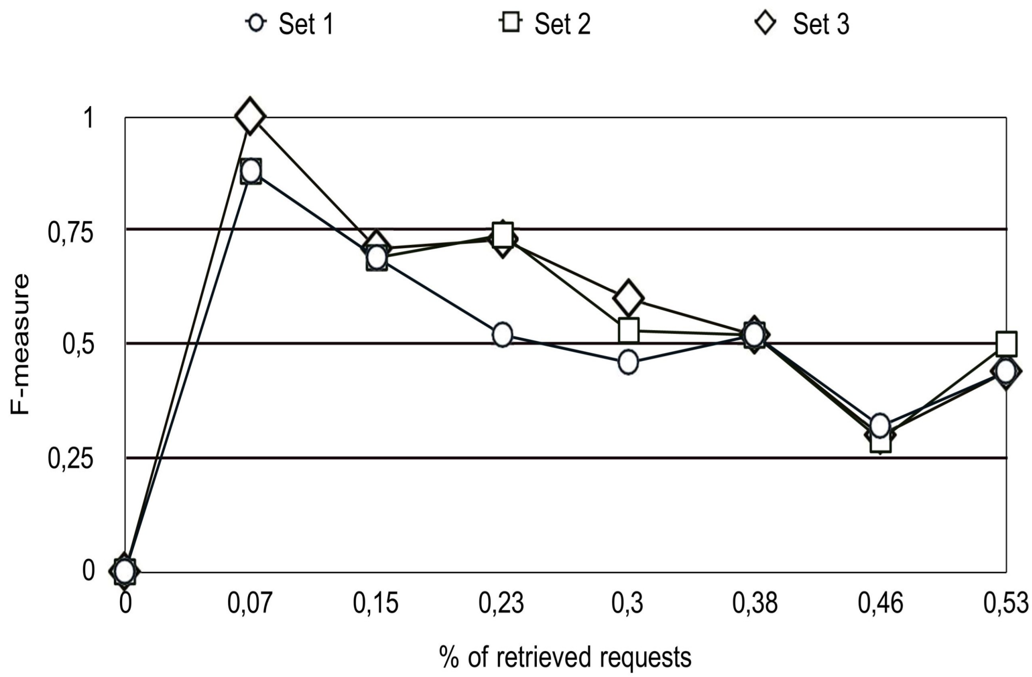

Figure 10. Measurement of the F-measure coefficient

equivalent system requests, leveraging the semantic relations between different objects and user procedures within the ontology.

The implementation of the different algorithms enabled us to test the model's performance, measure its accuracy, and evaluate its recall coefficient. We have observed that our approach achieves significant precision when the number of proposed requests is low and maintains acceptable precision when the number of retrieved requests is quite high. We have also noticed that the recall rate of our model is quite high, especially after enrichment of the KB.

Regarding our work perspectives, we first aim to conduct different tests using a standard KB to verify and confirm the previous results. The model we proposed, in its current version, can solve user requests with a specific format. A first improvement of the model would be to handle more complex requests by exploiting the user areas and the fuzzy values of the system attributes. A second improvement would be to incorporate a learning process that enriches the user areas of objects and procedures while preserving the consistency and stability of the KB.

References

[1] Bouchon, B., Omri, M.N., Tijus. C.A., (1994). A Fuzzy Sets Based System for User's Assistance: How SIFADE Diagnoses User's Goals. *The Second Wold Congress on Expert Systems: Moving towards Expert Systems Globally in the 21st Century. Macmillian ed. CD Rom. Estoril, Lisbonne*. (vol. 21, pp. 493-498).

[2] Chebil, W. Soualmia, L. F., Omri, M.N. (2018). Possibilistic Information Retrieval Model Based on a Multi-

Terminology. *The 14th International Conference on Advanced Data Mining and Applications (ADMA)*. (vol. 14. pp 66-79). Nanjing, China.

[3] Chebil, W., Soualmia, L., Omri, M N., Darmoni, S. J. (2017). Biomedical Documents Indexing with Bayesian Networks and Terminologies. *In*: 12th International Conference on Intelligent Systems and Knowledge Engineering (vol. 12. pp. 1-6). *IEEE,* Nanjing, China.

[4] Chebil, W., Soualmia, L. F., Omri, M. N., Darmoni, S. J. (2016). Indexing biomedical documents with a possibilistic network. *Journal of the Association for Information Science and Technology*. (vol. 67 (4), pp. 928–941).

[5] Chebil, W., Soualmia, L. F., Omri, M N., Darmoni, S. J. (2014). Extraction possibiliste de concepts MeSH a partir de documents biomedicaux. *Revue d Intelligence Artificielle*. 28(6), 729–752.

[6] Fkih, F., Omri, M.N. (2016). IRAFCA: An O(n) Information Retrieval Algorithm based on Formal Concept Analysis. *Knowledge and Information Systems*. 48(2), 465-491.

[7] Omri, M.N. (2012). Effects of Terms Recognition Mistakes on Requests Processing for Interactive Information Retrieval. *In: International Journal of Information Retrieval Research,* 2(3), 19-35).

[8] Omri, M.N. (2004). Possibilistic Pertinence Feedback and Semantic Networks for Goal's Extraction. *Asian Journal of Information Technology (AJIT)*. 3(4), 258-265.

[9] Omri, M.N. (2004). Pertinent Knowledge Extraction from a Semantic Network: Application of Fuzzy Sets Theory. *In International Journal on Artificial Intelligence Tools (IJAIT)*. 13(3), 705-719.

[10] Omri, M.N. (2004). Relevance Feedback for Goal's Extraction from Fuzzy Semantic Networks. *Asian Journal of Information Technology (AJIT)*. 3(6), 434-440.

[11] Omri, M.N. (2004). Possibilistic pertinence feedback and semantic networks for goal extraction, *Asian journal of information Technology*. 3(4), 258-265.

[12] Omri, M.N., Chouigui, N. (2001). Measure of similarity between fuzzy concepts for identification of fuzzy user requests in fuzzy semantic networks. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. 9(6), 743-748.

[13] Mendel, J. M., Hagras, H., Tan, W. W., Melek, W. W., Ying, H. (2014). Introduction to Type-2 Fuzzy Logic Control: Theory and Applications. *IEEE Press Series on Computational Intelligence*. John Wiley & Sons, Hoboken, NJ.

[14] Naouar, F., Hlaoua, L., Omri, M.N. (2013). Relevance Feedback for Collaborative Retrieval Based on Semantic Annotations. In the International Conference on Information and Knowledge Engineering . (vol. 13, pp.54-60), Las Vegas, Nevada, USA.

[15] Naouar, F., Hlaoua, L., Omri, M.N. (2013). Relevance Feedback in Collaborative Information Retrieval based on Validated Annotation. *In International Conference on Reasoning and Optimization in Information Systems*. (vol. 1). Sousse, Tunisia.

[16] Parry, D. (2002). A fuzzy ontology for medical retrieval. Auckland university of technology. ACSW Frontiers '04 Proceedings of the second workshop on Australasian information security, *Data Mining and Web Intelligence, and Software Internationalisation*. New Zealand. (vol. 32, pp. 121-126).

[17] Selvachandran, G., Pal,, M., Tahani A. Alhawari, A., Salleh. A. (2018). Interval-Valued Complex Fuzzy Sets and its Application to the Malaysian Economy. *International Journal of Fuzzy System Applications*. 7 (1), 22-31.

[18] Shanshan, M., Yingdong, H. (2018). Generalized Scaled Prioritized Intuitionistic Fuzzy Geometric Interaction Aggregation Operators and Their Applications to the Selection of Cold Chain Logistics Enterprises. *International Journal of Fuzzy System Applications*. 7(2), 1-21.

[19] Sooraj, T. R. Tripathy. B. K. (2018). An Interval Valued Fuzzy Soft Set Based Optimization Algorithm for High Yielding Seed Selection. *International Journal of Fuzzy System Applications*. 7(2), 44-61.

[20] Sujit, D, Debashish M., Samarjit K., Tandra P. (2018). A Brief Review and Future Outline on Decision Making Using Fuzzy Soft Set. *International Journal of Fuzzy System Applications*. 7(2), 1-43.

[21] Widyantoro D., Yen J. (2001). A fuzzy Ontology-based Abstract Search Engine and its User Studies, Texas University, Pennsylvania State University.