Fuzzy methods on the web: A critical discussion

Steven Schockaert, Nataliya Makarytska, Martine De Cock

Abstract Gradual concepts abound in many web-related domains, ranging from the notion of relevance in information retrieval, to the strength of connection in online social networks. As a result, fuzzy set theory is often a natural choice in implementing web systems. In this chapter, we give an overview of applications of fuzzy set theory in this area, focusing in particular on information retrieval, the semantic web, and recommender systems. In each case, we contrast fuzzy methods with other approaches, analyzing why and how the ideas of fuzzy set theory may be beneficial.

1 Introduction

The world wide web has often been promoted as a key application domain for fuzzy set theory [23, 86, 113]. Indeed, it is clear that to cope with the overwhelming amount of information on the web, intelligent techniques are needed to appropriately filter and preprocess the content of web pages. In traditional search engines, users convey their information need using a textual query, which is used to rank documents according to relevance. This ranked list is then presented to the user using well-chosen snippets from each of the documents. Modern information retrieval research attempts to replace the traditional keyword-based queries by more informative information requests, such as natural language questions, and to develop more advanced ways to present search results, typically by inducing some kind of structure from the set of relevant documents using clustering techniques. It appears that

Steven Schockaert and Martine De Cock
Ghent University, Department of Applied Mathematics and Computer Science, Belgium, e-mail: {steven.schockaert,martine.decock}@ugent.be

Nataliya Makarytska
University of Washington, Institute of Technology, Tacoma, WA, USA, e-mail: natashka@u.washington.edu

fuzzy methods have a natural role to play in this process. After all, the relevance of a text document to a topic clearly is a matter of degree, as is the interest of the user in a given topic. Already in the 1980s, before the advent of the web, the importance of flexibility in querying information systems was understood, leading to a variety of information retrieval methods based on fuzzy set theory [46]. Similarly, it has long been recognized that fuzzy clustering techniques, in which the membership of an object to a cluster centre is graded, are often more appropriate than their classical counterparts [43, 84], and clustering documents is no exception to this [48, 114]. In addition to structuring document collections or lists of search results, fuzzy clustering methods have also been used to build user profiles, based on past behavior or explicit user input [3, 57]. Such user profiles are useful to help the system personalize its output according to the desires and interests of each user.

Information retrieval is not the only domain in which fuzzy methods have been proposed. The semantic web [11] — a vision of a web of interlinked machine-readable information sources — has attracted considerable attention in the last decade. At the core of the semantic web vision is the notion of ontologies, which are shared formalizations of the concepts that appear in a certain domain. Ontologies are usually encoded in a standardized language, such as OWL, which can be modeled using a particular description logic. Dedicated description logic reasoners are then used to draw conclusions. Given appropriate ontologies, the content of a web resource or a web service can be described in a machine-readable way. This makes it possible to use formal reasoning to prove that some resource is relevant to a user (semantic search), or to automatically derive what behavior results from combining certain web services. In addition to description logics, also rule-based formalisms play a central role on the semantic web. The interest in fuzzy methods for the semantic web has mainly manifested itself at the foundational level; it has led to the development of fuzzy description logics [37, 98, 100], has further stimulated the development of fuzzy logic programming [44, 104], and its integration with (extensions of) description logics [52, 55].

A third area of the web where fuzzy methods are studied are recommender systems. Recommender systems provide users with recommendations (e.g. products the user may want to buy, movies that she wants to see, reviews that she wants to read, etc.) based on information about the users' preferences and about the items (the products, the movies, the reviews, etc.). Good and accurate recommender applications that guide users through the vast amounts of online information are gaining tremendous importance, as the wealth of information makes it increasingly difficult to find exactly what you want or need; all the more because every person has her own preferences. *Content-based systems* generate recommendations based on item similarity and, as such, tend to have their recommendation scope limited to the immediate neighbourhood of a user's past purchase or rating record. The performance of these systems can be improved significantly by (additionally) using *collaborative filtering*, which typically identifies users whose tastes are similar to yours and recommends items that these so-called neighbor users have liked. A more recent addition to the family of recommendation paradigms are the *social recommender systems*; these systems make suggestions based on likes and dislikes of users in your online social network. As similarity between items and users, and strength of connections in social networks, are graded concepts, it does not come as a surprise that a variety of interesting ideas has been proposed for the use of fuzzy set theory in content-based systems [110], colloborative filtering recommender systems [69, 79], as well as social recommenders [106].

However, despite the wide enthusiasm by the fuzzy set community, it is not clear what the real impact is of fuzzy methods on the web. Often, fuzzy methods are proposed to address problems that have not previously been considered. The lack of standardized benchmark data and strong baseline techniques then makes it difficult to provide a credible evaluation, and assessing the usefulness of the proposed solutions. In more classical domains, such as document retrieval, experimental results have been mixed. Due to the lack of clear experimental evidence for the success of fuzzy methods on the web, their impact outside the fuzzy set community remains limited. On the other hand, many popular approaches that do not refer to the term "fuzzy" are strongly related to the ideas of fuzzy set theory. Tag clouds [94], for instance, are little more than a fuzzy set of keywords. Conversely, techniques such as fuzzy clustering have little in common with the core ideas of fuzzy set theory and approximate reasoning, despite their reference to the term "fuzzy". In this sense, the question of whether or not fuzzy methods currently play an important role in web research is inherently vague, and is therefore best answered in linguistic terms: to some extent.

In this chapter, we provide a personal view on the benefits of fuzzy methods in web-related applications, as well as on the challenges that arise. It is not intended as a complete survey, but rather focuses on the three key domains that were sketched above: information retrieval, semantic web, and recommendation.

The chapter is structured as follows. In the next section, we provide a general introduction to fuzzy set theory, possibility theory and multi-valued logics, focusing especially on the different intuitions underlying these frameworks. Next, we discuss the use of fuzzy methods in each of the three aforementioned application domains: information retrieval in Section 3, the semantic web in Section 4 and recommender systems in Section 5. We end the chapter with a general conclusion on the role of fuzzy methods on the web.

2 Background

This section provides a non-technical introduction to fuzzy set theory and two related frameworks: possibility theory and multi-valued (or graded) logics. With the aim of clarifying the motivation of using fuzzy methods on the web in the following sections, our focus in this section is on the different intuitions that are behind these theories.

2.1 Fuzzy sets

Fuzzy sets were introduced by Zadeh [111] with the aim of modeling human concepts. It is well-known that such concepts tend to be vague, in the sense that for some objects it is not clear whether or not they belong to the concept [62]: is architecture a science? Is food poisoning a disease? Are bookends furniture? Zadeh noted that there usually exists a continuous transition between those objects that clearly belong to some concept, and those that clearly do not. Taking this observation into account, the fuzzy set representation of a concept associates to each object a degree of membership, chosen from the unit interval [0,1]. Formally, a fuzzy set A in a universe X is any mapping from X to $[0,1]$, where $A(x) = 1$ means that x fully belongs to the concept and $A(x) = 0$ means that x does not at all belong to the concept. In addition to modeling vague concepts, fuzzy sets are also used to model the intensity to which different objects satisfy some (well-defined) property. We may define, for instance, a fuzzy set of patients with fever, such that somebody with 37.5° C receives a membership degree of, say, 0.8, even though we may consider that fever by itself is a well-defined property. The concept of fuzziness, which is related to graded membership, should thus be distinguished from vagueness, which is related to the lack of precise meaning. A fuzzy relation from a universe X to a universe Y is a fuzzy set in the cartesian product $X \times Y$. A fuzzy relation from X to X is simply called a fuzzy relation in X. Fuzzy relations are typically used to model the strength of a certain relationship between objects of X and objects of Y.

Set operations are generalized to fuzzy sets in an indirect way, by generalizing logical conjunction and disjunction to graded truth values, and relying on the intuition that e.g. an element belongs to the intersection $A \cap B$ if it belongs to A *and* it belongs to B. Conjunction is usually generalized using t-norms, which are mappings T from $[0,1]^2$ to $[0,1]$ that are symmetric, associative, increasing, and satisfy the boundary condition $T(1,a) = 1$ for all $a \in [0,1]$. Given a t-norm T, the intersection of two fuzzy sets A and B in the same universe X is defined as $(A \cap B)(x) = T(A(x), B(x))$ for all $x \in X$. Similarly as for intersection, union is defined by generalizing disjunction. Typically, disjunction is generalized using a t-conorm S, which is a symmetric, associative, increasing $[0,1]^2 - [0,1]$ mapping that satisfies the boundary condition $S(0,a) = a$ for all $a \in [0,1]$. Given a t-conorm S, it is natural to define a generalized notion of implication, by $I(a,b) = S(1-a,b)$ for all $a,b \in [0,1]^2$, thus taking advantage of the classical equivalence $p \rightarrow q$ iff $\neg p \vee q$. Such generalized implications are called S-implicators. Another way to define generalized implications starts from a left-continuous t-norm T:

$$I(a,b) = \sup\{\lambda | \lambda \in [0,1] \text{ and } T(a,\lambda) \leq b\}$$

Such operators I are called residual implicators. While their definition is less intuitive than that of S-implicators, they often turn out to be particularly useful, as they tend to preserve more properties from classical logic.

The membership degrees of a fuzzy set can essentially be interpreted in three different ways, which relate to measurement of cost, distance, and frequency [31].

When fuzzy sets are used to add flexibility to a query (e.g. give me a list of cheap hotels in Ghent), membership degrees are used to encode *preference*, in the sense that objects which satisfy the query to a larger extent are more preferred. Membership degrees are then related to utility or cost. When fuzzy sets are used to provide an interface between the numerical domain and linguistic terms, the membership degrees reflect the *similarity* of an object to prototypes of the concept being modeled. The fact that architecture is a science to degree 0.3 then intuitively means that there is an object which is a clear example of the concept 'science' (e.g. physics) and which is similar to degree 0.3 with 'architecture'. In this case, membership degrees are related to distance measures. Finally, membership degrees can be used to express *uncertainty*. When being told about some user that she is young, we may consider some ages to be more plausible than others. The membership degree of a certain age in a fuzzy set modeling the concept 'young' is then interpreted as a degree of plausibility. In this case, membership degrees are related to probability theory, although different interpretations may be given to the exact relationship. Among others, the membership degree of an age λ in the fuzzy set *young* may be interpreted as the probability that somebody would assign the label young to the age λ. Fuzzy sets then correspond to likelihood functions [41]. Another way membership degrees can be related to probability is by interpreting fuzzy sets either as special cases or as approximations of random sets [30].

Regardless of the specific meaning that is given to membership degrees, it is important to note that taking the intersection of two fuzzy sets, for instance, is only meaningful if their membership degrees are *commensurable*. Let us take the example of querying a hotel reservation system. When we are interested in cheap hotels, interpreting the predicate 'cheap' is easy, as it can be done in a purely qualitative way (the cheaper the price, the more a hotel is cheap). However, when we rather ask for hotels that are at the same time 'cheap' and 'close to the city centre', the fuzzy sets modeling 'cheap' and 'close' should be such that the utility of being close to degree 0.7, for instance, is the same as the utility of being cheap to degree 0.7. Clearly, this puts strong constraints on how membership degrees should be obtained, which forms one of the most important practical difficulties in applying fuzzy set theory.

2.2 Possibility theory and approximate reasoning

Possibility theory [29, 112] is an uncertainty calculus which is tightly related to fuzzy set theory. At its basis is the notion of a possibility distribution π, which is a mapping from a universe X to the unit interval $[0,1]$, i.e. from a formal point of view, possibility distributions are fuzzy sets. A possibility distribution encodes for each $x \in X$ the degree of possibility that x is the actual value of some variable. Each possibility distribution π induces two uncertainty measures, called the possibility measure Π and the necessity measure N; for a subset $A \subseteq X$, they are defined as

$$\Pi(A) = \sup_{x \in X} \pi(x) \qquad\qquad N(A) = 1 - \Pi(coA)$$

where $coA = X \setminus A$ denotes set complement. Note that, from these definitions it follows that $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$, which is why possibility theory is called a non-additive uncertainty theory. Although possibility theory has mainly been developed as an uncertainty calculus related to fuzzy set theory, its ideas go back to the work of Shackle [93], who advocated the need for a non-additive uncertainty theory based on degrees of surprise. In this view, the possibility degree $\pi(x)$ is interpreted as the degree to which one would be surprised to learn that x is the actual value of the underlying variable. Note that possibility theory is not fully compositional. For instance, the value of $\pi(A \cap B)$ cannot be determined solely from the values of $\pi(A)$ and $\pi(B)$. This should come as no surprise as it is well-known that no uncertainty calculus can be completely compositional.

Possibility theory has played a central role in the theory of approximate reasoning. The basic idea is to interpret an assertion of the form "V is A" by the inequality $\pi_V \leq A$ (i.e. $\pi_V(x) \leq A(x)$ for all $x \in X$), where A is a fuzzy set, V a variable, and π_V a possibility distribution encoding knowledge about which values of V are possible. Such a constraint is called a flexible restriction on V. Given a number of flexible restrictions $\pi_V \leq A_1$,...,$\pi_V \leq A_n$, our state of knowledge regarding the possible values of variable V is encoded by the least specific solution to the constraints, i.e. $\pi_V(x) = \min(A_1(x), ..., A_n(x))$.

A central issue in approximate reasoning is how we can derive a flexible restriction on the value of a variable W, given a possibility distribution which encodes the possible values of variable V and an if–then rule of the form "**if** V is A **then** W is B". If–then rules are useful to encode common-sense knowledge such as "**if** the weather is nice **then** driving time to the coast will be long". Zadeh's *compositional rule of inference* suggests to derive the following possibility distribution π_W on W from the possibility distribution π_V on V and an if–then rule R:

$$\pi_W(y) = \sup_{x \in X} T(\pi_V(x), R(x,y))$$

where T is a t-norm and the if–then rule R is interpreted as a fuzzy relation. The intuition is clear: a value y for W is possible to the extent that there is a value x which is possible for V and such that the combination (x,y) does not violate the given if-then rule R. What remains to be decided is how to implement the if–then rule R itself. The most natural choice is to take $R(x,y) = I(A(x), B(y))$ for some implicator I, although sometimes a t-norm T' is used instead of an implicator. By extending this idea to multiple input variables, and multiple if–then rules, a powerful inference-scheme is obtained. Although it is computationally expensive in general, efficient methods can be obtained by assuming that the values of the input variables are precisely known and by taking fuzzy sets with membership functions that are easily processed (e.g. piecewise linear functions). Starting from the work by Mamdani [56], fuzzy reasoning with if–then rules has been widely used in industrial applications, as diverse as optimizing the program of washing machines and implementing auto-focusing techniques in digital cameras. It can be considered to be by far the most successful application of fuzzy set theory. In practice, if-then rules can be provided by an expert, or they can be derived automatically using appropriate training data. In domains where only limited training data is available, also a hybrid approach can be used: the expert provides a first version of the rules, which is subsequently refined using whatever training data that is available.

2.3 Multi-valued logic

The term fuzzy logic is used in two different senses in the literature. The first sense, often called the broad sense, mainly refers to the idea of approximate reasoning with if–then rules, as described above. The second sense, called the narrow sense, refers to formal logics in which the notion of truth is graded. This idea of graded truth is not exclusively tied to the framework of fuzzy set theory. Indeed, the notion of truth degree is already present in the three-valued logics that were developed in the first half of the 20th century, by Łukasiewicz, Gödel, Post, and Kleene, among others [28]. In the case of infinite-valued logics, truth degrees are values from [0,1] and logical connectives are interpreted as suitable $[0,1]^2 - [0,1]$ (conjunction, disjunction, implication) and $[0,1] - [0,1]$ (negation) functions. Usually, (propositional) multi-valued logics use the same syntax as classical (propositional) logic, although truth constants other than 0 are sometimes allowed in the language (e.g. in rational Pavelka logic [76]). Sound and complete proof theories for propositional fuzzy logics have been established, based on modus ponens and particular subsets of the axioms of classical logic [36]. For instance, infinite-valued Gödel logic is characterized by the axioms of intuitionistic logic together with the axiom of prelinearity: $(x \rightarrow y) \vee (y \rightarrow x)$. The semantics of Gödel logic is obtained by interpreting conjunction and disjunction by the minimum and maximum operators, and implication as the residual implicator induced by the minimum. Negation is defined as $\neg a = a \rightarrow 0$. The most popular fuzzy logics are Gödel logic, product logic and Łukasiewicz logic. In each case, conjunction and implication are interpreted in terms of some continuous t-norm and its residual implicator. For a more complete introduction to fuzzy logic, both in the narrow and the broad sense, we refer to the chapter by V. Novák and I. Perfilieva in this volume.

Note that despite the use of fuzzy logic connectives, infinite-valued logics are based on a completely different intuition than approximate reasoning. In particular, formulas from an infinite-valued logic encode a precise relationship between variables of some continuous domain. As such, propositional fuzzy logics do not deal with uncertainty or vagueness.

3 Information retrieval

3.1 Relevance models

An abundance of techniques that are based on fuzzy set theory or possibility theory have been proposed to improve the effectiveness and flexibility of search engines. Although modern web search engines are considerably more sophisticated than traditional information retrieval (IR) systems (e.g. exploiting hyperlinks to obtain more accurate relevance estimates as well as indications of authoritativeness), they are still essentially based on the same ingredients: a boolean keyword-based formulation of queries, and a bag-of-words representation of documents. In particular, documents are represented as weighted collections of terms, thus ignoring the ordering of the terms in the document, as well as any structure the document may have. Because of this simplification, a document can formally be modeled as a vector in a multidimensional space, with one dimension for each term occurring in the document collection under consideration. The weight of a component of a document vector is calculated based on the number of times the corresponding term occurs in the document (term frequency), and on the number of documents of the collection in which this term appears (inverse document frequency). The intuition is that a given term should receive a high weight when it occurs a lot in the document, while being rare in the collection as a whole. The query of the user, which is provided as a list of keywords, can also be represented as a vector, by treating it as a (short) document. A common approach to estimate the relevance of a document to a query then consists of calculating the cosine of the angle between the corresponding vectors. This *vector-space model* of information retrieval [85] has traditionally been the most popular approach to information retrieval, and is still considered state-of-the-art. However, state-of-the art performance in the vector-space model is obtained for variants of the aforementioned cosine-similarity which are difficult to interpret intuitively and rely on careful tweaking of the parameters involved [116].

More recently, *probabilistic language models*, which were first developed in the area of speech recognition, have been successfully applied to the information retrieval problem [80], combining state-of-the-art performance with intuitively appealing probabilistic models. Documents are then formally represented as probability distributions, which are used to calculate the probability that a document is relevant to the user. Retrieval models in which documents are represented as fuzzy sets have also been proposed [14, 46]. Conceptually, *fuzzy IR models* are similar in spirit to the vector-space model, using the same formulas to weigh the importance of a term in a document. The key difference is in the way queries are formulated and evaluated. Rather than representing the query as a small document, the relevance of a document is calculated using fuzzy logic connectives, measuring the degree to which a document 'implies' a query term, and subsequently combining these degrees using flexible alternatives for the operations of boolean conjunction or disjunction. Fuzzy IR models typically allow the user to specify for each keyword to what extent it is important for the query in linguistic terms (e.g. *very important*, *rather* *important*, etc.), and how the keywords should be combined using linguistic quantifiers (e.g. *most* of the keywords should be present). The main advantage of fuzzy IR models is in the flexibility they give users to specify their queries. Recently, also possibilistic approaches have been proposed [17]. Similar to probabilistic models, possibilistic models attempt to estimate the likelihood that a document is relevant to a query. In contrast to probabilistic models, however, this leads to two scores: the necessity that a document is relevant and the possibility that it is relevant. Documents are then ranked primarily based on the necessity scores, using the possibility scores to break ties, and in particular to provide meaningful results in cases where the necessity of relevance is 0 for all documents.

3.2 Domain-specific retrieval

The traditional information retrieval models are very general. Due to the fact that they treat words as abstract entities, for instance, most models are language-independent (although effectiveness of IR models is often dependent on language-specific issues [38, 72]). By restricting attention to a narrower domain, however, additional resources may be available that can help the retrieval process. One example are thesauri, which encode semantic relationships between terms, indicating for example that two terms are related in meaning (e.g. synonyms), or that one term is a specialization of another term (e.g. 'mathematician' is a specialization of 'scientist'). It is natural to consider that relations such as 'related term' are graded, as e.g. football and FIFA are more strongly related than football and player, even though football and player are still somewhat related. Accordingly, approaches to information retrieval have been proposed which use *fuzzy thesauri* [66, 83]. While utilizing a thesaurus seems very natural, as it allows to retrieve documents that are relevant to a query without actually sharing any terms with it, experimental validations of thesaurus-based IR models have failed to show a consistent improvement over systems without thesauri [45]. Practical problems with the use of thesauri include the fact that many words have different senses, which may lead semantically unrelated documents to be considered relevant, and the difficulties and costs involved in manually building high-quality thesauri. Automatically generated thesauri, typically based on detecting co-occurrence of terms, may provide a solution to the latter problem, but such thesauri are of varying quality, and moreover, highly dependent on the collection from which they have been obtained. Going from classical thesauri to fuzzy thesauri makes some of the problems even worse: how should reliable and meaningful grades be obtained? A recent example of the use of fuzzy thesauri can be found in [96].

It is interesting to note that the idea of fuzzy thesauri is also considered in the probabilistic language modeling approach to IR, although the term 'fuzzy thesauri' is not used in this context. In particular, a document is represented as a probability distribution, which is initially obtained using maximum likelihood estimation, i.e. the probability $P(t|d)$ that a term t is generated by the language model underlying document d is estimated as $\frac{n_t}{\sum_{t'} n_{t'}}$, where n_t is the number of occurrences of term t in document d. As this leads to the undesired effect that terms which do not occur in the document receive a zero probability, different forms of *smoothing* are applied. One form of smoothing is to interpolate this initial document model with a corpus model (which models the probability that a given term appears in the collection as a whole), which has a similar effect as considering inverse document frequency in the vector-space model. Recently, however, an additional form of smoothing, called *semantic smoothing* has gained importance [115]. Essentially, semantic smoothing corresponds to using a fuzzy thesaurus to increase the probability of terms that do not occur in the document, but are related to terms that do occur.

Somewhat related to the use of thesauri is concept-based information retrieval, where documents are linked to concepts from an ontology. By abstracting away from the actual terms that appear in a document, it may be expected that documents and queries may be compared in a way which is semantically more meaningful. In [33], for instance, document are modeled as vectors of Wikipedia concepts, and experimental evidence is provided that the similarity between documents can thus be measured in a substantially more accurate way. Somewhat related, [7] proposes to represent documents and queries as subtrees of ontology concepts, and uses fuzzy logic connectives to compute relevance scores. Again, convincing experimental evidence is provided to demonstrate the usefulness of the approach. In [6] a variant based on possibilistic logic is proposed. The possibilistic view naturally allows to associate three different degrees with each pair of terms (t_1, t_2): the possibility that t_1 and t_2 refer to the same thing, the necessity that t_1 is a specialization of t_2 and the necessity that t_2 is a specialization of t_1. An important advantage of this approach is that the degrees that are involved have a clear meaning.

In addition to retrieval of text documents, there is an increasing interest in retrieving other types of objects from the web [73], such as images[1], scientific papers[2], information about people[3], events[4], products[5], etc. Due to the fact that object-based retrieval is only applied in narrow domains, focusing on one particular type of objects, semantically richer, domain-dependent techniques may be applied, which are often of a very different nature than traditional text-based retrieval. For example, image retrieval systems often use a combination of textual evidence (e.g. the text surrounding the image on a web page) and visual features, and sometimes even focus exclusively on visual features [50, 95]. Fuzzy set-based approaches have been successfully applied to measure the similarity of visual features [12, 19, 49, 68]. Due to the use of richer semantics in object-based retrieval, often new types of opportunities arise for the application of fuzzy set theory. In [92], for instance, an approach is presented for retrieving events that satisfy given temporal restrictions, using a form of fuzzy temporal reasoning [89]. Here, the use of fuzzy set theory is

---

[1] http://www.flickr.com

[2] http://scholar.google.com

[3] http://pipl.com

[4] http://upcoming.yahoo.com

[5] http://www.google.com/products

directly motivated by the fact that many real-world events are of an inherently gradual nature, lacking a precise beginning and/or ending date. Similar considerations apply in the spatial domain, where the importance of vernacular geographic regions with imprecise boundaries is widely acknowledged [2, 67, 109], making approaches based on fuzzy set theory a natural choice [88, 91]. Finally, fuzzy methods have also been advocated in the area of music retrieval [15, 16].

3.3 Manipulation of search results

Fuzzy set based methods have been proposed for a variety of problems that in one way or another manipulate the results obtained from some standard information retrieval model. In [78], for instance, a fuzzy rule based system is presented to exploit the structure of HTML documents. While several retrieval models have already been proposed that can take advantage of the fact that e.g. terms in the title of an HTML document should be considered more important than terms in the body, most existing approaches are based on an assumption of statistical independence. The approach presented in [78] does not rely on such an assumption, as the overall importance of a term for an HTML document is determined by rules of the form

if 'Title' is *High* & 'Emphasis' is *Low* & 'Others' is *Low* **then** 'Result' is *Medium*

where 'Title' is the weight of the term in the title (i.e. treating the title as a small document), 'Emphasis' is the weight of the term when considering those occurrences that are otherwise emphasized in the document, and 'Others' is the weight of the remaining occurrences; *High*, *Low* and *Medium* refer to fuzzy sets with appropriately defined membership functions. The intuition of the rule above is that words which occur in the title, but not often in the remainder of the document may very well be irrelevant, e.g. because the title is metaphorical. Clearly, this intuition is very different from other approaches to retrieval of structured documents, which would boost the importance of terms that appear in the title, regardless of whether the term also appears elsewhere in the document. A total number of 9 such rules have been manually specified, and the resulting system is experimentally shown to outperform state-of-the-art methods such as BM25 and BM25F.

The motivation for using fuzzy set theory in this way is clearly not related to uncertainty or to the modeling of vagueness. Moreover, the same intuition can be implemented using other techniques, which after careful training and tweaking, might very well outperform such fuzzy rule based approaches. What fuzzy rules offer in this context, however, is the ability to easily implement an intuitive idea, using rules that are easy to understand. If the system does not perform as expected, it is straightforward to adapt the rules until the desired behavior is obtained, while many other methods crucially depend on the availability of good training data to arrive at 'black-box' models. Moreover, if such training data is actually available, the rules that have manually been constructed can be refined in an automated way [70]. A similar use of fuzzy rules is made in [90] with the aim of clustering web search results. There, fuzzy rules are used to implement the behavior of artificial agents, called ants, that move documents in a virtual environment and put them on heaps. Again, the use of if–then rules leads to a description which is easy to understand and to adapt to different intuitions about how documents should be clustered.

Other applications where fuzzy methods have been proposed to manipulate search results are: relevance feedback [21, 59, 108], meta-search [60], and query expansion [47, 58]. The techniques that are applied to this end are as diverse as fuzzy clustering [47], fuzzy association rules [58] and again fuzzy if-then rules [108]. This further illustrates the fact that fuzzy set theory can often provide a flexible vehicle for implementing advanced systems. In many cases, however, the authors provide very little experimental evidence to demonstrate the effectiveness of the proposed techniques over sufficiently strong baseline systems.

4 Semantic web

The semantic web [11] is a vision of interlinked machine-readable resources that exist on top of the web of human-readable documents that we know today. The widespread availability of such machine-readable resources would allow for the development of a variety of intelligent systems, such as semantic search systems that can prove the relevance of an object to some query based on a semantic representation of both the resource and the query. Central in this view is the notion of an ontology, which, in this context, is essentially a formalization of a given domain, describing properties of the relevant concepts and relations. The realization of a semantic web requires that two important challenges are overcome. The first challenge is acquiring the machine-readable resources that constitute the semantic web, which could be achieved by human experts who manually build ontologies, by automated techniques based on natural language processing, or by a combination of both. The second challenge is to exploit available information in a scalable, robust and useful way. It is in addressing this second challenge that possibilistic and fuzzy methods have a key role to play.

Taken as a whole, the information that is asserted on the semantic web will inevitably be inconsistent. Uncertainty about the correctness of individual pieces of information is therefore a key issue, which could be tackled by either probabilistic or possibilistic methods. Fuzzy methods, on the other hand, serve a different, but arguably equally important purpose. In particular, when moving from classical retrieval to semantic search, we lose the idea of a ranking. Indeed, when both resources and queries are expressed using classical logic, then we cannot acquire a more refined conclusion than that a resource is relevant, or that it is not relevant. In practice, this is problematic, because it is important to discriminate between objects that best satisfy the user's information need and those that only satisfy it marginally. Moreover, when no resource completely satisfies a given query, it may still be of interest to identify resources that 'almost' satisfy it. Thus, concepts such as preference and similarity, which are at the heart of fuzzy set theory, are therefore of crucial importance.

4.1 Description logics

Ontologies for the semantic web are usually modeled in description logics [4]. In such logics, knowledge is encoded in two separate knowledge bases, called the T-box and the A-box. The core idea is to describe properties of concepts and relations in the T-box and to describe in the A-box which objects are instances of which concepts, and which pairs of objects belong to which relations; usually relations are called roles in this context. Typically, atomic concepts are denoted by upper case letters A, B, etc. From such atomic concepts, complex concepts can be formed such as $A \sqcap B$, $A \sqcup B$, and $\neg A$, where e.g. $A \sqcap B$ is the concept whose instances are those objects that both belong to A and to B. The formal semantics is defined in terms of interpretations $\mathscr{I}$ that map concepts to sets of objects from a given domain $\Delta^{\mathscr{I}}$, e.g. $(A \sqcap B)^{\mathscr{I}} = A^{\mathscr{I}} \cap B^{\mathscr{I}}$. Similarly, the interpretation of roles is as relations in $\Delta^{\mathscr{I}} \times \Delta^{\mathscr{I}}$. In addition to the basic constructs, $\sqcap$, $\sqcup$, and $\neg$, also the concepts $\exists R\,.A$ and $\forall R\,.A$ are commonly used, where A is a concept and R a role; their semantics is as follows:

$$(\exists R\,.A)^{\mathscr{I}} = \{x \in \Delta^{\mathscr{I}} \mid \exists y \in \Delta^{\mathscr{I}}\,.\,(x,y) \in R^{\mathscr{I}} \wedge y \in A^{\mathscr{I}}\}$$
$$(\forall R\,.A)^{\mathscr{I}} = \{x \in \Delta^{\mathscr{I}} \mid \forall y \in \Delta^{\mathscr{I}}\,.\,(x,y) \in R^{\mathscr{I}} \rightarrow y \in A^{\mathscr{I}}\}$$

Thus, intuitively, $\exists R\,.A$ is the concept which contains all objects that are related (w.r.t. R) with some object in A, whereas $\forall R\,.A$ contains the objects that are only related to objects in A. Various variants of description logics can be defined, based on which type of constructs are allowed.

The T-box of a description logic theory encodes how different concepts relate to each other, using assertions of the form $A \sqsubseteq B$, where A and B are (not necessarily atomic) concepts, e.g.

$$\mathit{Professor} \sqsubseteq \mathit{FacultyMember} \qquad \exists \mathit{authorOf}\,.\mathit{ResearchPaper} \sqsubseteq \mathit{Researcher} \qquad (1)$$

encoding that professors are faculty members, and that all individuals who have authored at least one research paper are researchers. The A-box of a description logic theory contains assertions about individual objects of the form $x : A$, where x is an object and A is a concept, as well as assertions of the form $(x_1, x_2) : r$, where x_1 and x_2 are objects and r is a role; e.g.

$$\mathit{etienne} : \mathit{Professor} \qquad p_1 : \mathit{ResearchPaper} \qquad (\mathit{etienne}, p_1) : \mathit{authorOf}$$

Together (1) and (2) entail e.g. that $\mathit{etienne} : \mathit{FacultyMember} \sqcap \mathit{Researcher}$.

4.2 Fuzzy description logics

The main idea of fuzzy description logics [100] is to interpret concepts as fuzzy sets, acknowledging that many of the concepts that appear in real-world ontologies are vague. The most popular approach, initially proposed by Straccia [100], is based on a direct generalization of the semantics of classical description logics, e.g. the semantics of the concept $\exists R.A$ becomes for $x \in \Delta^{\mathscr{I}}$

$$(\exists R.A)(x) = \sup_{y \in \Delta^{\mathscr{I}}} T(R^{\mathscr{I}}(x,y), A^{\mathscr{I}}(y))$$

where T is a t-norm, $A^{\mathscr{I}}$ is a fuzzy set in $\Delta^{\mathscr{I}}$ and R is a fuzzy relation in $\Delta^{\mathscr{I}}$. Thus, each object x belongs to a concept such as $\exists R.A$ to some degree in $[0,1]$. The T-box now contains assertions of the form $\langle A \sqsubseteq B \geq \lambda \rangle$, for A and B concepts or roles, and $\lambda \in [0,1]$. In the case where A and B are concepts, for instance, the semantics of this assertion is as follows:

$$\mathscr{I} \models \langle A \sqsubseteq B \geq \lambda \rangle \quad \text{iff} \quad \inf_{x \in \Delta^{\mathscr{I}}} I(A(x), B(y)) \geq \lambda \tag{2}$$

where I is an implicator. Similarly, the A-box contains assertions of the form $\langle x : A \geq \lambda \rangle$, which, semantically, correspond to the condition that $A^{\mathscr{I}}(x) \geq \lambda$. Sound and complete reasoning procedures were introduced in [100] for a basic fuzzy description logic and a particular choice for the fuzzy logic connectives. More recently, among others, more expressive description logics have been considered [97, 99], larger classes of fuzzy logic connectives [37], and more complex reasoning tasks [54]. Of particular interest are fuzzy description logics with concrete domains [101], which allow to explicitly define fuzzy predicates which can then be used in the definition of concepts. For instance, in such logics, we could define a prolific researcher as a researcher who has published many papers as follows:

$$\mathit{Researcher} \sqcap \exists \mathit{numberOfPapers}.\mathit{Many} \sqsubseteq \mathit{ProlificResearcher}$$

together with an appropriate fuzzy set in $\mathbb{N}$ that encodes the predicate 'many', e.g.

$$\mathit{Many}(n) = \begin{cases} \frac{n-1}{n} & \text{if } n > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

As we have already indicated, adding fuzziness to description logics serves two rather distinct purposes. First, the fact that concepts are fuzzy sets naturally leads to flexibility in the querying process. For instance, when a user indicates that he is interested in a list of prolific researchers, it suffices to rank all instances of the concept *Researcher* according to the degree to which they belong to $\exists \mathit{numberOfPapers}.\mathit{Many}$. The definition of *Many* which was chosen in (3) then essentially means that the ordering of researchers according to their membership degree in *ProlificResearcher* is identical to the ordering based on their number of publications. Thus, the use of fuzzy sets allows for flexibility, as it eliminates the need for a crisp threshold on the required number of publications, and it naturally allows the system to rank the objects that (partially) satisfy the query. Second, when specifying a knowledge base, the fuzziness naturally allows to encode the intensity by which certain properties are satisfied. For instance, rather than specifying that Etienne is a prolific researcher, we can specify to what extent he is a prolific researcher, e.g. by asserting

$$\langle \mathit{etienne} : \mathit{ProlificResearcher} \geq \frac{389}{390} \rangle \tag{4}$$

It is important to note, however, that fuzzy description logics are *not* suitable for modeling vague knowledge, despite a wide number of claims to the contrary in the literature. Indeed, asserting (4) is exactly the same as asserting that Etienne has published at least 390 papers, which is clearly not vague at all. Modeling vague knowledge, such as "Etienne has published many papers" requires a mechanism for dealing with uncertainty, which is not present in standard fuzzy description logics. In other words, fuzzy description logics are suitable to deal with information which is naturally graded, but which is precisely known.

Most work on fuzzy description logics has been theoretical, developing more expressive formalisms, or more scalable reasoning mechanisms [13, 75]. One notable application of fuzzy description logics is in multimedia retrieval [65], where fuzzy description logic are used to encode both semantic annotations of multimedia documents and the result of e.g. image processing analyses. When it comes to the semantic web, it is not clear which is the role to be played by fuzzy description logics. A core requirement on the semantic web is the ability to link two ontologies that have been developed independently from each other. This, however, introduces a problem of commensurability. How should we compare what is called a prolific researcher to degree 0.4 in one fuzzy description logic base to what is called a young researcher to degree 0.7 in another fuzzy description logic base? Which guarantees do we have that it actually makes sense to combine these two degrees, to answer the query $\mathit{YoungResearcher} \sqcap \mathit{ProlificResearcher}$? In such a case, it seems more reasonable to explicitly encode the number of publications and the age of the researcher (or the constraints on these values that are known), rather than to encode degrees of membership. Thus, in contexts where interoperability plays a role, it seems that the use of fuzzy description logics may be problematic. Along the same lines, what may be problematic for certain types of applications is that different users may have a different view on concepts such as 'young' or 'prolific'. In traditional approaches to flexible querying, it is indeed the user who (implicitly or explicitly) determines how such concepts should be understood and how the degrees to which somebody is 'young' or 'prolific' should influence the ranking of the results. In fuzzy description logics, such concepts have a fixed meaning, which is independent of the preferences of an individual user. To some extent, it seems that the need for flexible approaches to querying ontologies and web information systems has been confused with a need for fuzziness at the knowledge representation level.

4.3 Possibilistic description logics

Possibilistic description logics [42, 27, 82] keep the crisp representation of concepts from classical description logics, adding a mechanism for handling uncertainty instead. Essentially, a T-box axiom then takes the form $(A \sqsubseteq B, \lambda)$, meaning that it is certain to degree λ that the concept inclusion $A \sqsubseteq B$ holds. Similarly, an A-box axiom takes the form $(x : A, \lambda)$, meaning that it is certain to degree λ that x is an instance of A. Semantically, possibilistic description logics generalize classical description logics, in a similar way as possibilistic logic generalizes propositional logic. Specifically, let $\mathscr{W}$ be the set of all description logic interpretations $\mathscr{I}$. An interpretation of a possibilistic description logic theory then is a possibility distribution π on $\mathscr{W}$. For each classical interpretation $\mathscr{I}$, $\pi(\mathscr{I})$ expresses how plausible it is that $\mathscr{I}$ corresponds to the real world. Note that the notion of plausibility that is considered here is purely qualitative. Writing $[\![\alpha]\!] \subseteq \mathscr{W}$ for the set of models of a description logic formula α (i.e. either a concept inclusion axiom in the T-box or an A-box assertion), the possibilistic description logic formula (α, λ) is interpreted as $N(\alpha) \geq \lambda$, i.e. the possibilistic description logic models of (α, λ) are those possibility distributions π whose induced necessity measure N is such that $N(\alpha) \geq \lambda$.

In a possibilistic description logic, we may write, for instance, that

$$(etienne : Professor, 0.9) \qquad\qquad (etienne : AboveFourty, 0.6)$$

which means that we are quite certain that Etienne is a professor, and rather certain that his age is above 40. Note that by combining the ideas of fuzzy and possibilistic description logics, vague knowledge may be encoded, e.g. writing assertions as

$$(\langle etienne : ProlificResearcher \geq 0.7\rangle, 0.9)$$
$$(\langle etienne : ProlificResearcher \geq 0.8\rangle, 0.6)$$
$$(\langle etienne : ProlificResearcher \geq 0.9\rangle, 0.3)$$

When it comes to semantic web applications, possibilistic description logics share with their fuzzy counterparts the problem of commensurability. The certainty weights that appear in possibilistic description logic bases that have been developed independently cannot be compared. One solution would be to resort to possibilistic logic variants that can deal with partially ordered certainty weights [9]. Another avenue for applications is the combination of several classical description logic bases. Possibilistic certainty weights could then be added to the assertions that are made by each source, based on their reliability. In this way, when different sources are combined, the ones that are considered least reliable are discarded. Although this idea has not yet been considered for possibilistic description logics, similar ideas have been extensively studied for merging conflicting propositional knowledge bases [8, 10]. Especially when utilizing variants which do not suffer from the so-called ‘drowning effect’, such as the approach presented in [82], this seems to be a promising direction.

4.4 Logic programming

Logic programming deals with inferring knowledge from rules of the form

$$c \leftarrow a_1, ..., a_n, not\, b_1, ..., not\, b_m \tag{5}$$

which encode the intuition that unless one of the terms $b_1, ..., b_m$ can be derived, it holds that $a_1 \wedge ... \wedge a_n$ implies c. In its simplest form, terms are restricted to atomic propositions and the semantics of logic programs may be given in a purely declarative way using the notion of stable models [34]; this approach is often referred to as *answer set programming*. The idea of logic programming in general, and answer set programming in particular, has been generalized to deal with graded properties. The intuition of (5) is then that the truth degree of c is at least as high as the truth degree of $a_1 \wedge .... \wedge a_n$, unless one of the atoms $b_1, ..., b_m$ can be derived to a high degree. Note however, that there exist several ways to implement this intuition, leading to different semantics of fuzzy logic programming [44, 64, 102, 107]. In parallel, some possibilistic extensions to logic programming have been considered [5, 20, 26, 71, 74], in which it is possible to encode that a given rule or fact is more or less certain (or important, preferred, etc.). Although the idea of uncertainty or preference is clearly different from the idea of graded truth, at the formal level, extensions of answer set programming that deal with uncertainty are closely related to extensions dealing with graded truth [5, 24].

An interesting application of fuzzy logic programming for the semantic web is discussed in [63], where grades are used to encode similarity between terms. Among others, this is useful to deal with the fact that different resources may use a different terminology to refer to the same or similar concepts, for instance to tackle problems related to interoperability on the semantic web. In the proposed approach, a given set of logic programming rules is augmented with additional rules that encode which terms can be considered similar and to what degree. What is not entirely clear, from an application point of view, however, is whether these degrees relate to the certainty that two terms describe the same property/concept/object, or to the strength of the similarity between the two terms. Although the formal treatment may be analogous in both cases, in order to obtain meaningful results, a clear operational semantics of the grades is needed in applications, which may put constraints, for instance, on which fuzzy logic connectives can be used to combine the grades.

Motivated by the needs of semantic web applications, logic programming rules have also been combined with description logics, leading to *description logic programs* [32]. Accordingly, fuzzy description logic programs have been developed, which combine fuzzy description logics with fuzzy answer set programming [53, 103]. In [40], as an application of such fuzzy description logic programs, the problem of ranking web services according to the preferences of a given user is considered. Again, there is no clear distinction between uncertainty and graded truth, in the sense that the motivation of the paper is given in terms of graded truth, while the application example that is presented essentially deals with uncertainty.

5 Recommendation and personalization

The wealth of information available on the web has made it increasingly difficult to find what one is really looking for. This is particularly true for exploratory queries where one is searching for opinions and views, not because it is difficult to look up this kind of information, but because there is simply so much of it that one does not know where to start consuming it. Hence, it comes at no surprise that personalization systems that guide the search process are gaining importance. On the popular consumer review site Epinions[6] for instance, the order in which reviews are presented to the user is personalized and depends on the user's previous ratings of other reviews (in terms of helpfulness) and the user's social network information. Another example is Google News[7], a computer-generated news site that aggregates headlines from news sources worldwide, groups similar stories together and displays them according to each reader's personalized interests. From an e-commerce perspective too, the value of a good recommender system cannot be underestimated: Cinematch, the recommender of the American online movie rental system Netflix[8], delivers two thirds of Netflix's rented movies, and Amazon.com claims that 35% of their sales result from recommendations [51]. In essence, the recommendation problem consists of predicting the extent to which a particular user, the so-called *target user*, will like a particular item, called the *target item*, which can be a review, a news article, a movie, a book, a song, a research paper, etc. The predicted degree is usually taken from a linear scale (for instance from 1 to 5 stars) which can, without loss of generality, be mapped to $[0,1]$. Hence the predicted degree is a fuzzy membership degree that encodes preference. This preference degree is however an outcome of (and not an input to) the recommendation process, and can be arrived at through various methods discussed below. Before we go on, note that a solution to the canonical recommendation problem also implies a solution to the problem of presenting the target user with a personalized list of items, as these can be ranked in order of preference degree.

5.1 Content-based recommendations

The content-based approach to recommendation has its roots in information retrieval and employs many of the same techniques [77]. All content-based recommender systems take the *content* of items into account, which are usually described by vectors of attributes. In a movie recommender system, for example, a movie is typically represented by a vector that contains the title, the genre, the director, the lead actors, etc., while a personalized news website can use a term frequency–inverse document frequency (TF-IDF) representation of every news article. Furthermore,

[6] www.epinions.com
[7] news.google.com
[8] www.netflix.com

content-based recommenders rely on a *profile* of the target user, which can be either manually entered by the user or derived from past behavior, such as previous ratings or purchases. Some content-based recommender systems provide an interface that allows users to construct a representation of their own interests. In this case, the recommendation process, which compares the available items with the user profile, very much resembles information retrieval as discussed in Section 3, with the user profile playing the role of a query. This is especially so when the items contain textual information, such as news articles or research papers, and the user profile consists of keywords or topics that the target user is interested in. In addition, in the fuzzy research community, systems have been proposed in which users state their information need with linguistic labels, asserting for instance that *weight* is a *very important* consideration in a new laptop they want to buy, or which *research topics* are *more or less compatible* with their interests [18, 81, 110]. These linguistic labels are then mapped to fuzzy sets which are compared to a similar fuzzy set representation of the available items (consumer products, research funding opportunities,...). This approach's achilles heel for large scale deployment seems to be the need for domain experts to evaluate the features of every item and to establish item descriptions as vectors of linguistic labels (fuzzy sets).

Other content-based recommender systems learn the user profile automatically from past behavior and recommend items that are similar to items purchased or rated highly by the target user in the past. Implementing this requires a technique to compute the similarity between items, which varies with the domain. A common approach is to use the cosine similarity between the attribute vectors that describe the items, especially for textual items represented by TF-IDF vectors in term space. To this end, items are sometimes also represented as vectors in user space, with the p-th entry of the vector for an item containing the rating that the p-th user gave to the item, or, alternatively, a 1 if the p-th user purchased the item and a 0 otherwise. In this case, items are considered similar to the extent to which they have a common customer or fan base. Regardless of which of these techniques is used to compute it, let us denote the similarity of items i and j as $Sim(i,j)$, which, without loss of generality, can be thought of as a fuzzy relation in the set of items $\mathscr{I}$. The predicted rating $\widehat{P}_{\mathrm{CB}}(u,i)$ for target user u and target item i can then be computed as the weighted mean [87]

$$\widehat{P}_{\mathrm{CB}}(u,i) = \frac{1}{\sum\limits_{j \in \mathscr{I}_u} Sim(i,j)} \sum_{j \in \mathscr{I}_u} (Sim(i,j) \cdot P(u,j)) \tag{6}$$

in which $\mathscr{I}_u$ is the set of items previously rated by u, and $P(u,j)$ denotes the rating that user u previously gave to item j. In the fuzzy set community, proposals have been made to generalize the product in (6) to an arbitrary t-norm, and to replace the weighted mean by the supremum as the aggregation operation, resulting in prediction formulas such as [79, 110]

$$\widehat{P}_{\mathrm{CB}}(u,i) = \sup_{j \in \mathscr{I}_u} T(Sim(i,j), P(u,j)) \tag{7}$$

To the best of our knowledge, no experimental studies exist on which t-norm/aggregation combination performs best on benchmark datasets.

In content-based recommender systems, items for which no description is available can not be recommended, and the accuracy of the recommendations heavily relies on the quality of the representations. Furthermore, the technique to compute the similarities is domain dependent. For instance, a content-based system developed for recommendation of reviews or news articles in English requires adaption before it can be used for other languages as well. Another drawback of content-based systems is that they tend not to explore interests of the user besides those expressed in his rating record. In this sense, they can be improved significantly by (additionally) using collaborative methods, which do not require item descriptions.

5.2 Collaborative filtering

While content-based methods depend on the computation of similarity between items, collaborative filtering relies on similarity between users. The main idea is to recommend items that have been rated highly by users similar to the target user. Similarity between users is typically assessed based on rating behavior, i.e. users are considered similar if they (dis)like the same items, and can be computed in the same way across different domains. The similarity between users u and v is commonly measured with Pearson's correlation coefficient (PCC) [39]

$$Sim(u,v) = \frac{\sum\limits_{j \in \mathscr{I}_u \cap \mathscr{I}_v} (P(u,j) - \overline{P}(u)) \cdot (P(v,j) - \overline{P}(v))}{\sqrt{\left(\sum\limits_{j \in \mathscr{I}_u \cap \mathscr{I}_v} (P(u,j) - \overline{P}(u))^2\right) \cdot \left(\sum\limits_{j \in \mathscr{I}_u \cap \mathscr{I}_v} (P(v,j) - \overline{P}(v))^2\right)}} \tag{8}$$

where the summations range over all items j previously rated by both u and v, and $\overline{P}(u)$ and $\overline{P}(v)$ are the average ratings given by u and v so far. The PCC ranges between -1 and 1. A positive PCC means that both users have similar taste in the sense that, when one of them rates an item above/below average, the other one does so too. The more negative the coefficient, the more the rating behaviors are opposites, and a correlation coefficient of 0 means that there is no relationship between the two sets of ratings. In practice, most often only users with a positive correlation with target user u and who have rated target item i are considered in the recommendation process. We denote this set by $\mathscr{U}_i$. The predicted rating $\widehat{P}_{\mathrm{CF}}(u,i)$ for target user u and target item i can then be computed as the weighted mean [1]

$$\widehat{P}_{\mathrm{CF}}(u,i) = \frac{1}{\sum\limits_{v \in \mathscr{U}_i} Sim(u,v)} \sum_{v \in \mathscr{U}_i} (Sim(u,v) \cdot P(v,i)) \tag{9}$$

Formula (9) does not take into account the fact that not every user exhibits the same rating behavior, in the sense that some users might be easy to please and regularly give high ratings, while others have a more pronounced taste and tend to give lower ratings more often. The classic collaborative filtering formula accounts for this [1]

$$\widehat{P}_{\mathrm{CF}}(u,i) = \overline{P}(u) + \frac{1}{\sum\limits_{v\in\mathscr{U}_i} Sim(u,v)} \sum_{v\in\mathscr{U}_i} \left(Sim(u,v)\cdot(P(v,i)-\overline{P}(v))\right) \tag{10}$$

Note, however, that such an adaption still has its limitations. For instance, as there is no correction based on the standard deviation of the scores, users whose scores are almost always around 3 will influence recommendations to a lesser extent than users who make use of the entire range from 1 to 5 on a regular basis. More fundamentally, the theoretical justification of formulas such as (10) is very loose. In principle, user ratings can only be interpreted in an ordinal way, and imposing any kind of metric on these scores is always to some extent arbitrary. Approaches which are based on difference in ratings, average ratings, etc., should therefore be seen as heuristics rather than well-founded methods. One might imagine alternative, more principled techniques which are more in the spirit of qualitative decision making [25], e.g. ranking an item i_1 higher than an item i_2 iff the set of users who have rated i_1 higher than i_2 is more similar to the target user than the set of users who have rated i_2 higher than i_1. Such methods, however, would probably suffer from other issues, such as scalability. As in the domain of information retrieval, it thus seems that there is a trade-off between looking for techniques that make sense from a theoretical point of view, and exclusively relying on experimental studies to arrive at techniques that are efficient and effective in practice.

Similarly as with content-based recommendation, in the fuzzy set community proposals have been made to replace the product in (9) by an arbitrary t-norm and to use the supremum to aggregate over the users in the neighborhood of u, resulting in formulas such as [22, 69, 79]

$$\widehat{P}_{\mathrm{CF}}(u,i) = \sup_{v\in\mathscr{U}_i} T(Sim(u,v), P(v,i)) \tag{11}$$

Formulas (7) and (11) are very similar in structure. In (7) the supremum ranges over all items j previously rated by target user u, and their similarity with target item i is taken into account. In (11) on the other hand, the supremum ranges over all users v who are already familiar with the target item i; in this case the similarity between users u and v is an important factor.

In [69], Formula (11) is used for a web page recommender system that dynamically appends a set of links to the contents of a web document returned in response to the most recent query of an ongoing user session. Recommendations are made based on access data instead of rating behavior. User sessions are represented as attribute vectors with the p-th attribute equal to 1 if the p-th url was accessed during the session, and 0 otherwise. To limit the number of user sessions over which (11) ranges, the set of user sessions from the access log files is replaced by a smaller set of prototypical user sessions, which represent clusters found in the original set. The attribute vector describing such a prototypical user session or cluster has values between 0 and 1; the p-th attribute indicates the relative frequency with which the p-th url was visited in all user sessions belonging to the clusters. $Sim(u,v)$ is computed as the cosine similarity of the vectors for u and v instead of the PCC, and min is used as the t-norm in (11). The authors compare their approach with the results of a nearest profile based recommendation approach (recommend the urls visited in the prototypical user session that is most similar to the ongoing user session) and with a k-nearest neighbor approach followed by top-n recommendations (recommend the n most frequently visited urls from the k most similar prototypical user sessions). They report a small drop in precision which is more than compensated for by an increase in recall. The question whether perhaps even better results could be obtained with Formula (9) remains open.

5.3 Social recommenders

When a web application with a built-in recommender offers a social networking component which enables its users to form a trust network, it can generate more personalized recommendations by combining data from the user profiles (ratings) with information from the social network. These are the so-called trust-enhanced or social recommendation systems. Ratings are predicted in a style similar to collaborative filtering, with the similarity score $Sim(u,v)$ replaced by a trust score $Trust(u,v)$ corresponding to the degree to which user u trusts user v. The trust-based versions of (9) and (10) are at the heart of the trust-enhanced recommendation algorithms of Golbeck et al. [35] and Massa et al. [61] respectively. If no direct trust score is available (because u does not know v), then it can often still be derived through *trust propagation* and *aggregation* in the online network, inspired by the way in which humans often seek recommendations in real life. For instance, the trust score of u in v can be estimated as a weighted mean of the trust scores of other users in v, weighted by the trust of v in those other users [35, 61]

$$\widehat{Trust}(u,v) = \frac{1}{\sum\limits_{w\in\mathscr{U}} Trust(u,w)} \sum_{w\in\mathscr{U}} \left(Trust(u,w)\cdot Trust(w,v)\right) \tag{12}$$

Formula (12) only considers one step propagation, i.e., where u and v are directly connected through a third party w; extensions that take into account longer propagation paths are possible as well. Propagation is modeled in (12) by the product. A proposal has been made to generalize this to an arbitrary t-norm and to use ordered weighting averaging operators that can deal with gradual trust as well as distrust [106, 105]. Even though some of the initial experimental results are promising, a proper evaluation of the effect of different t-norms on the performance of a trust-enhanced recommender system is currently hampered by the lack of a publicly available benchmark dataset that contains both item ratings as well as a social network with gradual trust relations.

Moreover, one may wonder what the precise meaning of a trust degree is, why a formula such as (12) is compatible with this meaning, and how such degrees can be acquired in practice. The basic intuition seems to be that friends are more likely to have similar interests than random users, which would suggest to use trust mainly to adapt the Pearson correlation in the collaborative filtering model, such that the degree of similarity between friends is boosted. The notion of trust then takes a role which is similar in spirit to that of a prior probability in Bayesian decision theory.

6 Conclusions

In this chapter, we have looked at the use of fuzzy set theory in three research areas that are related to the world wide web: information retrieval, the semantic web, and recommender systems. While the motivation for using fuzzy techniques is very natural in each of these domains, the most commonly used techniques are nonetheless still based on other approaches. This can partly be explained by the fact that more experimental evidence is needed to demonstrate whether fuzzy methods are really able to outperform state-of-the-art approaches. In addition, the assessment of the impact of fuzzy methods on the web is obscured by the fact that sometimes methods are used which are based on its ideas, without making use of its vocabulary. To stimulate the future impact of fuzzy approaches to web intelligence, we believe that more efforts are needed to lay bare what fuzzy set theory really has to offer in this domain, beyond the (important) fact that it allows to develop elegant and intuitively appealing methods.

Information retrieval research is dominated by algebraic (vector space model) and probabilistic (language models) approaches. Fuzzy set theory has mainly been applied to implement more flexible ways of formulating queries, and to develop semantically informed retrieval models for particular narrow domains. In addition, fuzzy rule based methods have sometimes proven useful for translating human intuitions on how search results should be manipulated, in domains where sufficient training data is missing.

In the last decade, Tim Berners-Lee's vision of a semantic web has drawn many researchers to work on fuzzy versions of its main components. In particular, research on fuzzy description logics has substantially progressed, both at the theoretical (more expressive formalisms) and at the practical level (more efficient reasoners). More recently, there has also been a renewed interest in fuzzy logic programming, in relation to the semantic web. There exists some confusion, however, between the need for flexible querying, the presence of vague concepts, the presence of uncertainty, and the need for fuzziness at the representation level. While convincing applications of fuzzy description logics have already been developed, we are not aware of any applications that are in the spirit of the semantic web, e.g. dealing with problems that result from linking different fuzzy description logics that have been developed independently. More work is needed to clarify the advantages of fuzzy description logics over extensions of traditional web information systems that are endowed with flexible querying capabilities.

Recommender systems aim to solve a problem that is familiar to the fuzzy set community, namely predicting the degree to which a target user might like a target item. Most solutions proposed in the fuzzy set community are very similar in structure to those proposed outside. One potential advantage that fuzzy set theory has to offer is its wider variety of operators, compared with traditional approaches that tend to limit themselves to the use of the product for conjunction and the use of the mean for aggregation. The proof of the pudding is in the eating though, in this case, whether some of these other operators can lead to more and better recommendations in practice. Since the first proposals for fuzzy logic recommendation techniques were made, a variety of benchmark datasets have become available. Even though the nature of these datasets does not allow yet to empirically evaluate the use of fuzzy methods for trust-enhanced recommender systems, an evaluation of fuzzy methods for content-based and collaborative filtering seems a feasible and logical next step.

Acknowledgements Steven Schockaert was funded as a postdoctoral fellow of the Research Foundation – Flanders.

References

1. G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.
2. A. Arampatzis, M. van Kreveld, I. Reinbacher, C. Jones, S. Vaid, P. Clough, H. Joho, and M. Sanderson. Web-based delineation of imprecise regions. *Computers, Environment and Urban Systems*, 30(4):436–459, 2006.
3. D. Arotaritei and S. Mitra. Web mining: a survey in the fuzzy framework. *Fuzzy Sets and Systems*, 148(1):5–19, 2004.
4. F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider. *The Description Logic Handbook*. Cambridge University Press, 2007.
5. K. Bauters, S. Schockaert, M. De Cock, and D. Vermeir. Possibilistic answer set programming revisited. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence*, to appear.
6. M. Baziz, M. Boughanem, Y. Loiseau, and H. Prade. Fuzzy logic and ontology-based information retrieval. In P. Wang, D. Ruan, and E. Kerre, editors, *Fuzzy Logic: A Spectrum of Theoretical & Practical Issues*, pages 193–218. Springer-Verlag, 2007.
7. M. Baziz, M. Boughanem, H. Prade, and G. Pasi. A fuzzy logic approach to information retrieval using a ontology-based representation of documents. In *Fuzzy logic and the Semantic Web*, pages 363–377. Elsevier, 2006.
8. S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Possibilistic merging and distance-based fusion of propositional information. *Annals of Mathematics and Artificial Intelligence*, 34(1-3):217–252, 2002.
9. S. Benferhat, S. Lagrue, and O. Papini. Reasoning with partially ordered information in a possibilistic logic framework. *Fuzzy Sets and Systems*, 144(1):25–41, 2004.

10. S. Benferhat and C. Sossai. Reasoning with multiple-source information in a possibilistic logic framework. *Information Fusion*, 7(1):80–96, mar 2006.
11. T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, 284(5):28–37, 2001.
12. E. Binaghi, I. Gagliardi, and R. Schettini. Image retrieval using fuzzy evaluation of color similarity. *International Journal of Pattern Recognition and Artificial Intelligence*, 8:945–945, 1994.
13. F. Bobillo, M. Delgado, and J. Gómez-Romero. Optimizing the crisp representation of the fuzzy description logic SROIQ. In *Proceedings of the 3rd ISWC Workshop on Uncertainty Reasoning for the Semantic Web*, 2007.
14. G. Bordogna and G. Pasi. A fuzzy linguistic approach generalizing boolean information retrieval: A model and its evaluation. *Journal of the American Society for Information Science*, 44(2):70–82, 1993.
15. K. Bosteels and E. E. Kerre. Fuzzy audio similarity measures based on spectrum histograms and fluctuation patterns. In *Proceedings of the 2007 International Conference on Multimedia and Ubiquitous Engineering*, pages 361–365, 2007.
16. K. Bosteels, E. Pampalk, and E. E. Kerre. On the benefits of representing music objects as fuzzy sets. In *Proceedings of the IFSA/EUSFLAT Conference*, pages 25–29, 2009.
17. M. Boughanem, A. Brini, and D. Dubois. Possibilistic networks for information retrieval. *International Journal of Approximate Reasoning*, 50(7):957–968, 2009.
18. Y. Cao and Y. Li. An intelligent fuzzy-based recommendation system for consumer electronic products. *Expert Systems with Applications*, 33:230–240, 2007.
19. Y. Chen and J. Wang. A region-based fuzzy feature matching approach to content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1252–1267, 2002.
20. R. Confalonieri, J. C. Nieves, and J. Vázquez-Salceda. Pstable semantics for logic programs with possibilistic ordered disjunction. In *Proceedings of the Conference of the Italian Association for Artificial Intelligence*, pages 52–61, 2009.
21. O. Cordón, F. Moya, and C. Zarco. A new evolutionary algorithm combining simulated annealing and genetic programming for relevance feedback in fuzzy information retrieval systems. *Soft Computing*, 6(5):308–319, 2002.
22. C. Cornelis, J. Lu, G. X., and Z. G. One-and-only item recommendations with fuzzy logic techniques. *Information Sciences*, 177:4906–4921, 2007.
23. F. Crestani and G. Pasi. Handling vagueness, subjectivity, and imprecision in information access: an introduction to the special issue. *Information Processing & Management*, 39(2):161–165, 2003.
24. C. V. Damásio and L. M. Pereira. Hybrid probabilistic logic programs as residuated logic programs. In *Proceedings of the 7th European Workshop on Logics in Artificial Intelligence*, pages 57–72, London, UK, 2000. Springer-Verlag.
25. D. Dubois, H. Fargier, and H. Prade. Decision-making under ordinal preferences and comparative uncertainty. In *Proceedings of the 13h Conference on Uncertainty in Artificial Intelligence*, pages 157–164, 1997.
26. D. Dubois, J. Lang, and H. Prade. Towards possibilistic logic programming. In *Proceedings of the International Conference on Logic Programming*, pages 581–595, 1991.
27. D. Dubois, J. Mengin, and H. Prade. Possibilistic uncertainty and fuzzy features in description logic: A preliminary discussion. *Fuzzy logic and the semantic web*, pages 101–113, 2006.
28. D. Dubois, W. Ostasiewicz, and H. Prade. Fuzzy sets: History and basic notions . In D. Dubois and H. Prade, editors, *Fundamentals of Fuzzy Sets* , The Handbooks of Fuzzy Sets Series, pages 21–124. Kluwer Academic Publishers, 2000.
29. D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, 1988.
30. D. Dubois and H. Prade. Fuzzy sets, probability and measurement. *European Journal of Operational Research*, 40(2):135 – 154, 1989.

31. D. Dubois and H. Prade. The three semantics of fuzzy sets. *Fuzzy Sets and Systems*, 90(2):141–150, 1997.
32. T. Eiter, G. Ianni, T. Lukasiewicz, R. Schindlauer, and H. Tompits. Combining answer set programming with description logics for the semantic web. *Artificial Intelligence*, 172(12-13):1495–1539, 2008.
33. E. Gabrilovich and S. Markovitch. Computing semantic relatedness using wikipedia-based explicit semantic analysis. In *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, pages 1606–1611, 2007.
34. M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In *Proceedings of the Fifth International Conference and Symposium on Logic Programming*, pages 1081–1086, 1988.
35. J. Golbeck and J. Hendler. Filmtrust: movie recommendations using trust in web-based social networks. In *Proceedings of the 3rd IEEE Consumer Communications and Networking Conference*, pages 282–286, 2006.
36. P. Hájek. *On the Metamathematics of Fuzzy Logic*. Kluwer Academic Press, 1998.
37. P. Hájek. Making fuzzy description logic more general. *Fuzzy Sets and Systems*, 154(1):1 – 15, 2005.
38. T. Hedlund, A. Pirkola, and K. Järvelin. Aspects of swedish morphology and semantics from the perspective of mono- and cross-language information retrieval. *Information Processing & Management*, 37(1):147 – 161, 2001.
39. J. Herlocker, J. Konstan, L. Terveen, and J. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22:5–53, 2004.
40. S. Heymans and I. Toma. Ranking services using fuzzy hex programs. In *Proceedings of the 2nd International Conference on Web Reasoning and Rule Systems*, pages 181–196, 2008.
41. E. Hisdal. Are grades of membership probabilities? *Fuzzy Sets and Systems*, 25(3):325 – 348, 1988.
42. B. Hollunder. An alternative proof method for possibilistic logic and its application to terminological logics. *International Journal of Approximate Reasoning*, 12(2):85 – 109, 1995.
43. F. Höppner, F. Klawonn, R. Kruse, and T. Runkler. *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*. Wiley, 1999.
44. J. Janssen, S. Schockaert, D. Vermeir, and M. De Cock. General fuzzy answer set programs. In *Proceedings of the 8th International Workshop on Fuzzy Logic and Applications (WILF)*, pages 352–359, 2009.
45. Y. Jing and W. Croft. An association thesaurus for information retrieval. In *Proceedings of RIAO*, number 1994, pages 146–160, 1994.
46. E. Kerre, R. Zenner, and R. De Caluwe. The use of fuzzy set theory in information retrieval and databases: A survey. *Journal of the American Society for Information Science*, 37(5):341–341, 1986.
47. D. Kraft, J. Chen, M. Martin-Bautista, and M. Vila. Textual information retrieval with user profiles using fuzzy clustering and inferencing. *Studies In Fuzziness And Soft Computing*, pages 152–165, 2003.
48. R. Krishnapuram, A. Joshi, and L. Yi. A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering. In *IEEE International Fuzzy Systems Conference Proceedings*, volume 3, pages 1281 –1286, 1999.
49. R. Krishnapuram, S. Medasani, S. Jung, Y. Choi, and R. Balasubramaniam. Content-based image retrieval based on a fuzzy approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(10):1185–1199, 2004.
50. M. La Cascia, S. Sethi, and S. Sclaroff. Combining textual and visual cues for content-based image retrieval on the world wide web. In *IEEE Workshop on Content-Based Access of Image and Video Libraries*, pages 24–28. Citeseer, 1998.
51. P. Lamere and O. Celma. Music recommendation tutorial, 2007.
52. T. Lukasiewicz. Fuzzy description logic programs under the answer set semantics for the semantic web. *Fundamenta Informaticae*, 82(3):289–310, 2008.

53. T. Lukasiewicz and U. Straccia. Tightly integrated fuzzy description logic programs under the answer set semantics for the semantic web. In *Proceedings of the 1st international conference on Web reasoning and rule systems*, pages 289–298, 2007.
54. T. Lukasiewicz and U. Straccia. Top-k retrieval in description logic programs under vagueness for the semantic web. In *Proceedings of the First International Conference on Scalable Uncertainty Management*, pages 16–30, 2007.
55. T. Lukasiewicz and U. Straccia. Description logic programs under probabilistic uncertainty and fuzzy vagueness. *International Journal of Approximate Reasoning*, 50(6):837–853, 2009.
56. E. Mamdani and S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1):1–13, 1975.
57. M. Martin-Bautista, D. Kraft, M. Vila, J. Chen, and J. Cruz. User profiles and fuzzy logic for web retrieval issues. *Soft Computing*, 6(5):365–372, 2002.
58. M. Martin-Bautista, D. Sanchez, J. Chamorro-Martı?nez, J. Serrano, and M. Vila. Mining web documents to find additional query terms using fuzzy association rules. *Fuzzy Sets and Systems*, 148(1):85–104, 2004.
59. M. Martin-Bautista, M. Vila, and H. Larsen. A fuzzy genetic algorithm approach to an adaptive information retrieval agent. *Journal of the American Society for Information Science*, 50(9):760–771, 1999.
60. M. J. Martin-Bautista, D. Sanchez, M. A. Vila, and M. Delgado. A new fusion strategy for web metasearch. *Soft Computing*, 14(8):847–855, 2009.
61. P. Massa and P. Avesani. Trust metrics on controversial users: balancing between tyranny of the majority and echo chambers. *International Journal on Semantic Web and Information Systems*, 3:39–64, 2007.
62. M. McCloskey and S. Glucksberg. Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, 11(1):1 – 37, 1979.
63. J. Medina, M. Ojeda-Aciego, and P. Vojtá. Similarity-based unification: a multi-adjoint approach. *Fuzzy Sets and Systems*, 146(1):43–62, 2004.
64. J. Medina, M. Ojeda-Aciego, and P. Vojtáš. Multi-adjoint logic programming with continuous semantics. *Lecture Notes in Computer Science*, 2173:351–364, 2001.
65. C. Meghini, F. Sebastiani, and U. Straccia. A model of multimedia information retrieval. *Journal of the ACM*, 48(5):909–970, 2001.
66. S. Miyamoto. Information retrieval based on fuzzy associations. *Fuzzy Sets and Systems*, 38(2):191–205, 1990.
67. D. Montello, M. Goodchild, J. Gottsegen, and P. Fohl. Where's downtown?: behavioral methods for determining referents of vague spatial queries. *Spatial Cognition and Computation*, 3(2-3):185–204, 2003.
68. M. Nachtegael, D. van der Weken, V. de Witte, S. Schulte, T. Melange, and E. Kerre. Color image retrieval using fuzzy similarity measures and fuzzy partitions. In *IEEE International Conference on Image Processing*, pages 13–16, 2007.
69. P. C. Nasraoui, O. Combining web usage mining and fuzzy inference for website personalization. In *Proceedings of WebKDD*, pages 37–46, 2003.
70. D. Nauck and R. Kruse. A neuro-fuzzy method to learn fuzzy classification rules from data. *Fuzzy Sets and Systems*, 89(3):277–288, 1997.
71. P. Nicolas, L. Garcia, I. Stéphan, and C. Lefèvre. Possibilistic uncertainty handling for answer set programming. *Annals of Mathematics and Artificial Intelligence*, 47(1–2):139–181, 2006.
72. J.-Y. Nie and F. Ren. Chinese information retrieval: using characters or words? *Information Processing & Management*, 35(4):443 – 462, 1999.
73. Z. Nie, Y. Ma, S. Shi, J.-R. Wen, and W.-Y. Ma. Web object retrieval. In *Proceedings of the 16th International Conference on World Wide Web*, pages 81–90, 2007.
74. J. C. Nieves, M. Osorio, and U. Cortés. Semantics for possibilistic disjunctive programs. In *Proceedings of the International Conference on Logic Programming and Nonmonotonic Reasoning*, pages 315–320, 2007.

75. J. Z. Pan, G. Stamou, G. Stoilos, S. Taylor, and E. Thomas. Scalable querying services over fuzzy ontologies. In *Proceedings of the 17th International Conference on World Wide Web*, pages 575–584, 2008.
76. J. Pavelka. On fuzzy logic. *Zeitschr. f. Math. logik und Grundl. der Math.*, 25:45–52, 119–134, 447–464, 1979.
77. M. Pazzani and D. Billsus. Content-based recommendation systems. *Lecture Notes in Computer Science*, 4321:325–341, 2007.
78. J. Pérez-Iglesias, V. Fresno, and J. Pérez-Agüera. FuzzyFresh: A Fuzzy Logic Approach to the Ranking of Structured Documents. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 755–758, 2008.
79. Z. J. Perny, P. Collaborative filtering methods based on fuzzy preference relations. In *Proceedings of EUROFUSE-SIC*, pages 279–285, 1999.
80. J. M. Ponte and W. B. Croft. A language modeling approach to information retrieval. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 275–281, 1998.
81. C. Porcel, A. López-Herrera, and E. Herrera-Viedma. A recommender system for research resources based on fuzzy linguistic modeling. *Expert Systems with Applications*, 36:5173–5183, 2009.
82. G. Qi, J. Z. Pan, and Q. Ji. Extending description logics with uncertainty reasoning in possibilistic logic. In *Proceedings of the 9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 828–839, 2007.
83. T. Radecki. Mathematical model of information retrieval system based on the concept of fuzzy thesaurus. *Information Processing & Management*, 12(5):313–318, 1976.
84. E. Ruspini. Numerical methods for fuzzy clustering. *Information Sciences*, 2(3):319–350, 1970.
85. G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.
86. E. Sanchez, editor. *Fuzzy Logic and the Semantic Web*. Capturing Intelligence. Elsevier, 2006.
87. B. Sarwar, G. . Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web*, pages 285–295, 2001.
88. S. Schockaert and M. De Cock. Neighborhood restrictions in geographic IR. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 167–174, 2007.
89. S. Schockaert and M. De Cock. Temporal reasoning about fuzzy intervals. *Artificial Intelligence*, 172:1158–1193, 2008.
90. S. Schockaert, M. De Cock, C. Cornelis, and E. Kerre. Clustering web search results using fuzzy ants. *International Journal of Intelligent Systems*, 22(5):455–474, 2007.
91. S. Schockaert, M. De Cock, and E. Kerre. Automatic acquisition of fuzzy footprints. In *Proceedings of the International Workshop on Semantic Based Geographic Information Systems, LNCS 3762*, pages 1077–1086, 2005.
92. S. Schockaert, M. De Cock, and E. E. Kerre. Reasoning about fuzzy temporal information from the web: towards retrieval of historical events. *Soft Computing*, 14(8):869–886, 2010.
93. G. Shackle. *Decision, Order and Time in Human Affairs*. Cambridge University Press, 1961.
94. J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34(1):15–29, 2008.
95. A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1349–1380, 2000.
96. A. Soto, J. Olivas, and M. Prieto. Fuzzy Approach of Synonymy and Polysemy for Information Retrieval. *Granular Computing: At the Junction of Rough Sets and Fuzzy Sets*, pages 179–198, 2008.

97. G. Stoilos, G. Stamou, J. Z. Pan, V. Tzouvaras, and I. Horrocks. Reasoning with very expressive fuzzy description logics. *Journal of Artificial Intelligence Research*, 30(1):273–320, 2007.
98. G. Stoilos, G. Stamou, V. Tzouvaras, J. Pan, and I. Horrocks. Reasoning with the fuzzy description logic f-SHIN: theory, practice and applications. In *Uncertainty Reasoning for the Semantic Web, LNCS 5327*, pages 262–281, 2008.
99. G. Stoilos, U. Straccia, G. Stamou, and J. Z. Pan. General concept inclusions in fuzzy description logics. In *Proceeding of the European Conference on Artificial Intelligence*, pages 457–461, 2006.
100. U. Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 4:137–166, 2001.
101. U. Straccia. Description logics with fuzzy concrete domains. In *21st Conference on Uncertainty in Artificial Intelligence*, pages 559–567, 2005.
102. U. Straccia. Annotated answer set programming. In *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, (IPMU'06)*, 2006.
103. U. Straccia. Fuzzy description logic programs. In *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 1818–1825, 2006.
104. U. Straccia, M. Ojeda-Aciego, and C. V. Damásio. On fixed-points of multivalued functions on complete lattices and their application to generalized logic programs. *SIAM Journal on Computing*, 38(5):1881–1911, 2009.
105. P. Victor. *Trust Networks for Recommender Systems*. PhD thesis, Ghent University, 2010.
106. P. Victor, C. Cornelis, M. De Cock, and P. Pinheiro da Silva. Gradual trust and distrust in recommender systems. *Fuzzy Sets and Systems*, 160(10):1367–1382, 2009.
107. P. Vojtás. Fuzzy logic programming. *Fuzzy Sets and Systems*, 124(3):361–370, 2001.
108. S. Vrettos and A. Stafylopatis. A Fuzzy Rule-Based Agent for Web Retrieval-Filtering. In *Proceedings of the 1st Asia-Pacific Conference on Web Intelligence*, volume 448, page 453. Springer, 2001.
109. T. Waters and A. Evans. Tools for web–based GIS mapping of a "fuzzy" vernacular geography. In *Proceedings of the 7th International Conference on GeoComputation*, 2003.
110. R. R. Yager. Fuzzy logic methods in recommender systems. *Fuzzy Sets and Systems*, 136(2):133–149, 2003.
111. L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
112. L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3–28, 1978.
113. L. A. Zadeh. A note on web intelligence, world knowledge and fuzzy logic. *Data & Knowledge Engineering*, 50(3):291–304, 2004.
114. Y. Zhao and G. Karypis. Soft clustering criterion functions for partitional document clustering: a summary of results. In *Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management*, pages 246–247, 2004.
115. X. Zhou, X. Hu, X. Zhang, X. Lin, and I.-Y. Song. Context-sensitive semantic smoothing for the language modeling approach to genomic ir. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 170–177, 2006.
116. J. Zobel and A. Moffat. Exploring the similarity space. *SIGIR Forum*, 32(1):18–34, 1998.