Information Sciences 180 (2010) 1390–1406

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A framework for reasoning with soft information

Ronald R. Yager

Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, United States

ARTICLE INFO

Article history:
Received 31 August 2009
Received in revised form 25 November 2009
Accepted 4 December 2009

Keywords:
Soft information
Fuzzy sets
Deduction
Fusion

ABSTRACT

In order to provide for the representation and manipulation of human sourced soft information we turn to the fuzzy set based theory of approximate reasoning. We describe how approximate reasoning provides a framework for representing and manipulating a wide body linguistically expressed information. We then suggest a number of extensions of the theory to enhance its representational capacity. One such extension focuses on the ability to model imprecise variables as well as imprecise values for the variable. We consider the representation of possible qualified propositions. We look at the issue of deduction in the face of conflict in our knowledge base and suggest an approach compatible with human behavior.

1. Introduction

The ability to represent and manipulate human sourced soft information is crucial to the modern technological agenda. It is important to the construction of information fusion systems as well to the vast variety of intelligent Web based systems. As much of human sourced knowledge is expressed using words this requires the use of granular technologies to model the imprecision inherent in human language. Our objective here is to look at one such technology, the fuzzy set based theory of approximate reasoning (AR). We first introduce this theory and describe how it provides a framework for representing and manipulating a wide body linguistically expressed information. We then suggest a number of extensions of the theory to enhance its representational capacity.

In typical applications we express knowledge with statements such as Age of Mary is young. Here we assign some imprecise value young to the variable Mary's Age and we use a fuzzy set to define young. In this work we extend the representational capability of this format by considering expressions in which the object about whose age we a commenting is imprecisely described. An example of this is the statement the age of the tall woman near the bridge is young. Thus here rather then specifying a specific object Mary we have indicated am imprecise linguistically described object, the tall woman near the bridge. We suggest a formulation for representing this type of knowledge that allows it to be used within the AR deduction process. We next consider the representation of possibility-qualified statements. Here we make use of interval valued and intuitionistic fuzzy sets to supply this representation. One issue we investigate involves deduction in the face conflicting information. Most reasoning systems, including the standard binary logic deductive system, when faced with a knowledge base containing conflicting knowledge, allow the inference of everything. We suggest a deduction mechanism that provides a much more circumspect behavior in this situation, as the degree of conflict in the knowledge base increases this system reduces the facts that can be deduced. Finally we suggest a methodology for making inferences in the face of a similarity-based context.

E-mail address: yager@panix.com

0020-0255/$ - see front matter © 2009 Published by Elsevier Inc.
doi:10.1016/j.ins.2009.12.004

2. The theory of approximate reasoning

To provide a technology for manipulating human sourced linguistically expressed soft knowledge Zadeh introduced the fuzzy set based theory of Approximate Reasoning (AR) and the generalized constraint language [16–18]. Fundamental to this approach is the realization that much of our knowledge can be viewed as a constraint on some explicit or implicit variable. The process introduced by Zadeh for manipulating this type of knowledge can be seen to involve four basic steps (see Fig. 1). The first step involves the representation of the available knowledge in terms of the formal language of the theory of approximate reasoning. This can be seen as a step of translation. In this step each of the pieces of knowledge are translated into constraints. The second step involves the combining or fusing of the individual pieces of knowledge to form a global constraint that constitutes our knowledge base. The third step involves the making of inferences in response to some objective. The final step is one of retranslation. This step involves taking the inferred knowledge, which is a constraint expressed in the formal language of approximate reasoning, and turning it into some natural language statement that is easier for a human to understand.

Central to this approach is the task of translation. Here we take knowledge and using the formal language of AR, convert it into constraints on some implicit or explicit variable. Considerable effort has been devoted to the issue of translation. Numerous rules have been suggested for accomplishing this translation [6,13].

An important class of constraints are the possibility constraints. These constraints arise from propositions or statements such as ***John is tall***. Here we are imposing some constraints of the possible ages of John. At a formal level if V is a variable with a domain of discourse X then a formal representation of a possibilistic constraint is

$$V \text{ is } A.$$

Here A is a fuzzy subset of the domain X. Often A is a representation of some linguistic value associated with V.

The effect of this constraint is to induce a possibility distribution $\prod$ over X so that $\prod(x) = A(x)$, is the possibility that x is the value of V. Here $A(x)$ is the membership grade of x in A.

Two special cases of V *is* A are worth noting. The first is the case where $A(x) = 1$ for all $x \in X$, here $A = X$. In this case we see we know nothing about V. We can see this corresponds the to a linguistic value of **unknown**. Another special case is where $A(x^*) = 1$ and $A(x) = 0$ for $x \neq x^*$. In this case we have that $V = x^*$, it equals the value x^*. There is no uncertainty.

We shall say a fuzzy set, possibility distribution, is normal if there exists at least one element with membership one. If this is not the case we say it is subnormal. In the framework of Zadeh's theory of approximate reasoning subnormality is some indication of conflict or inconsistency. An extreme case of this occurs when $A = \varnothing$, the null set. Generally constraints involving subnormal fuzzy subsets do not appear in primary or source knowledge but only appear as a result of the combination or fusion of multiple pieces of knowledge that may be conflicting. For example the combination of knowledge that **John is old** and **John is young** will induce some subnormality.

Here we describe some operations for managing information of this type.

$$V \text{ is } A \textbf{ and } V \text{ is } B \Rightarrow V \text{ is } D,$$

where $D = A \cap B$. Here we have $D(x) = \text{Min}[A(x), B(x)]$.

$$V \text{ is } A \textbf{ or } V \text{ is } B \Rightarrow V \text{ is } E,$$

where $E = A \cup B$. Here we have $E(x) = \text{Max}[A(x), B(x)]$

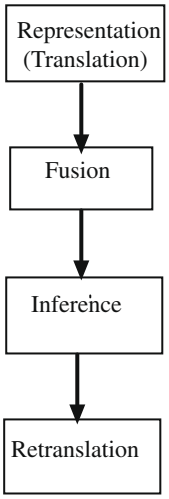

Fig. 1. Framework of approximate reasoning.