国家科学技术学术著作出版基金资助

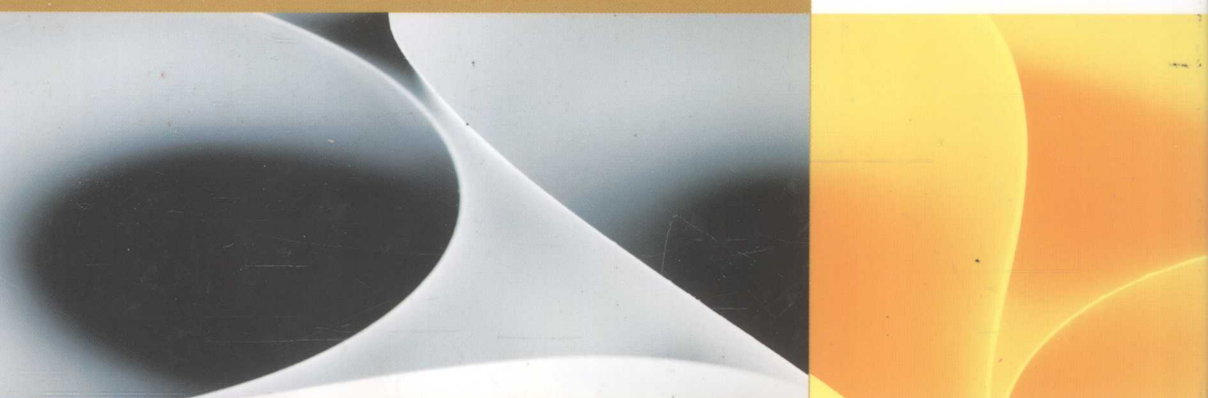
新一代信息科学与技术

Fuzzy Computational Ontologies in Contexts

Yi Cai
Ching-man Au Yeung
Ho-fung Leung

情境中的模糊计算本体

（英文版）



高等教育出版社
HIGHER EDUCATION PRESS

国家科学技术学术著作出版基金资助

新一代信息科学与技术

Yi Cai
Ching-man Au Yeung
Ho-fung Leung

Fuzzy Computational Ontologies in Contexts

情境中的模糊计算本体(英文版)

QINGJING ZHONG DE MOHU JISUAN BENTI

高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

Authors
Dr. Yi Cai
School of Software Engineering
South China University of Technology
Guangzhou, 510006, China
E-mail: ycai@scut.edu.cn

Dr. Ching-man Au Yeung
Hong Kong Applied Science and
Technology Research Institute
Hong Kong, China
E-mail: albertauyeung@gmail.com

Prof. Ho-fung Leung
Department of Computer Science and
Engineering
The Chinese University of Hong Kong
Hong Kong, China
E-mail: lhf@cuhk.edu.hk

图书在版编目(CIP)数据

情境中的模糊计算本体=Fuzzy Computational Ontologies in Contexts：英文/蔡毅，欧阳靖民，梁浩锋著. —北京：高等教育出版社，2012.1
ISBN 978-7-04-033889-8
Ⅰ.①情… Ⅱ.①蔡…②欧…③梁… Ⅲ.①人工智能-研究-英文 Ⅳ.①TP18

中国版本图书馆 CIP 数据核字(2011)第221147号

策划编辑 陈红英　　责任编辑 陈红英　　封面设计 张　楠
责任印制 朱学忠

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街4号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	涿州市星河印刷有限公司	网上订购	http://www.landraco.com
开　　本	787mm×1092mm　1/16		http://www.landraco.com.cn
印　　张	13.75	版　　次	2012年1月第1版
字　　数	350千字	印　　次	2012年1月第1次印刷
购书热线	010-58581118	定　　价	49.00元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号　33889-00

Yi Cai
Ching-man Au Yeung
Ho-fung Leung

Fuzzy Computational Ontologies in Contexts

Preface

This book originates from the research of Ching-man Au Yeung while he was a research student working with Ho-fung Leung at The Chinese University of Hong Kong. In the beginning, we were motivated to study knowledge representation mechanisms by the rapid development of the Semantic Web in recent years and the fact that ontologies had been widely used to model various concepts and objects on the Web. However, we found that existing ontology models were not fully capable of modeling concepts and objects in a way that was compatible with the human thinking process, and therefore we considered that better ontology models were needed.

To understand how human perceives concepts and objects, we sought inspirations from studies in cognitive psychology. Cognitive psychology is a subject in psychology that explores human mental processes, and has a close relationship with artificial intelligence. Much research has been carried out in cognitive psychology to study how human perceives concepts and performs categorization. We were in particular attracted to the concept of typicality.

According to studies in cognitive psychology, human usually considers some objects to be better examples of a given concept than the others. Even though objects in a particular concept all possess the necessary properties that define the concept, some objects may still be perceived as more typical examples than the others with respect to the concept in question. This is known as the 'typicality effect.' This effect is particularly interesting because its nature is mostly psychological, and in many cases it has almost nothing to do with the definition of a concept. Hence, typicality effect is different from concept fuzziness, which concerns with concepts that have no clear boundaries, such as 'tall man,' 'high building' and 'high temperature.'

After reviewing the cognitive psychology literature, we believed that it would be beneficial to incorporate typicality into existing ontology models, as a complementary measure to existing fuzzy membership degrees. The result of this research is the first fuzzy ontology model we described in this book. This model provides a number of features. Firstly, it allows concepts and objects to be defined by properties with different weights. It also provides a mechanism to measure object membership when the concept is fuzzy. Moreover, it provides a mechanism to construct a prototype for a concept, which can be used to measure typicality of different objects with respect to the given concept. The model is the first to allow modeling of both concept fuzziness and typicality effect at the same time.

This first model was further developed by Yi Cai while he was a PhD student of Leung at The Chinese University of Hong Kong. After some investigations, we found that there was a limitation in the first model. More specifically, we found that a single-prototype model of concepts was not always enough to represent a concept and a single-characteristic vector was not always enough to define a concept. For example, for the concept of 'vehicle,' there are at least three different prototypes: land vehicle, watercraft, and aircraft. It would be odd to artificially create a single prototype for 'vehicle' that combines possesses all the salient features of land vehicles, watercrafts, and aircrafts, which would only result in something very odd. Cai and Leung found in the literature of cognitive psychology the idea of multiple-prototype concepts, which was used to extend the first model to a fuzzy ontology model that could overcame the limitation of the first model. Other issues, such as property hierarchy, context effects, property importance and priority, were investigated and solutions were formally incorporated into the models.

To investigate the usefulness of object typicality in real life applications, we looked into recommender systems from new perspectives that take into account object typicality. We noticed that existing collaborative filter approaches recommend items to users based on either user similarity or item similarity. We explored new mechanisms that recommend typical items in a category to typical users that like items in that category, and proposed a typicality-based recommendation system named ROT and a typicality-based collaborative filtering approach named TyCo. Experiments showed that these approaches have superior performance. They can improve the recommendation quality, i.e., obtaining more accurate predictions with less big-error predictions while comparing previous recommendation methods, especially for sparse training data sets.

To the best of our knowledge, this book is the first that introduces object membership and typicality in fuzzy ontologies. It is also the first to apply object typicality to extending recommendation algorithms. This book contains a lot of examples to assist readers to understand the proposed models. We also include the necessary background knowledge from cognitive psychology in this book.

The main target readers of this book are graduate students and researchers who are interested into fuzzy ontology modeling, and engineers who are working on knowledge representation, the Semantic Web and recommender systems. We hope that this book will be of considerable use to the Semantic Web community by providing original and distinct views on this important interdisciplinary subject, and by contributing to a better understanding among individuals in this research area.

Finally, we are grateful to the Higher Education Press and Springer for encouraging us to put together our research outputs with some extensions as a monograph.

Hong Kong
August 2011

Yi Cai
Ching-man Au Yeung
Ho-fung Leung

Acknowledgements

We thank Juanzi Li and Jie Tang of Tsinghua University for their enthusiastic help in this research work. The work presented in this book is partially supported by CUHK Research Committee Funding (Direct Grants) (Project Codes: EE07672, EE08562 and EE10513), and by Guangdong Natural Science Foundation (No.S2011040002222).

Contents

Chapter 1 Introduction

Since the seminal Dartmouth Conference in 1956 [1], artificial intelligence (AI) has grown into an independent field of research, drawing ideas as well as techniques from various fields, including philosophy, mathematics, computer science and engineering, economics, neuroscience, psychology and linguistics [2]. Within this large field of research, there are many areas, such as problem solving, searching, knowledge representation and reasoning, planning and decision making, statistical learning and neural networks, and robotics. There is no doubt that each of these areas has contributed to the advancement of artificial intelligence and has constituted a lot of useful applications, and each area has its own importance and significance. Here, we single out the area of knowledge representation.

Knowledge representation and reasoning is an area in artificial intelligence that concerns with how human knowledge, including abstract concepts, categories, method of classifications, procedural knowledge and relations between different entities, can be represented symbolically and in a structured way, so that a computer is able to manipulate the knowledge, and other relevance information in an automated and efficient way, to perform reasoning tasks and to draw conclusions from known facts and knowledge [3]. We consider knowledge representation as one of the most important areas in the field of artificial intelligence. The ultimate aim of artificial intelligence is to realize intelligence in artificial entities such as computers. It has been a general view that human beings behave intelligently because of what they know and understand, and because of their ability to apply their knowledge to solve problems they encountered, to adapt to their continuously changing environment and to achieve their goals [3]. Therefore, to allow artificial software entities to behave intelligently or appear to have intelligence, it becomes inevitable that there must be effective and efficient ways for the representation of knowledge, which can be used as the basis for further intelligent tasks such as reasoning and decision making.

Research in the area of knowledge representation has generated quite a number of research topics, such as formal logics and logical reasoning, categorization and classification, analogical reasoning, and expert systems. Different methods and formalisms for representing human knowledge in computers in a structured and organized way have been developed, including first order logic, semantic networks, object-oriented models, description logics and ontologies, each with its own characteristics, advantages and limitations [3]. Among these formalisms, ontologies have attracted more and more attention in the last decade. Ontologies are now widely used as a means of conceptual modeling or domain modeling in various areas of application including knowledge management, natural language processing, e-commerce, information retrieval, bio-informatics, and the new emerging Semantic Web [4]. In particular, the Semantic Web [5] and the development of multi-agent systems [6] have accelerated research on ontologies and ontological engineering.

In this book, we focus on the issue of knowledge representation with the use of ontologies in the context of the Semantic Web. We discuss the challenges facing knowledge representation in ontologies, identify problems as well as other desirable features of ontologies in the Semantic Web, and propose possible solutions to the problems and challenges. In the following sections, we give an overview of the Semantic Web and the use of ontologies as a knowledge representation formalism, and discuss the motivations as well as our objectives of our research work.

1.1 Semantic Web and Ontologies

Ontology is originally a philosophical discipline [7]. It is a major and fundamental branch of metaphysics that tries to give a systematic explanation of being. It studies the problems of being, existence and their basic categorizations and relationships [4]. The word *ontology* has been adopted into the field of computer science, especially by researchers in artificial intelligence, to refer to the specification of the objects, properties and relations that one would encounter in a particular domain of discourse. One of the first definitions noted in Ref. [4] was given by Neches et al. [8]:

> An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary.

Another mostly quoted definition of ontology was given by Gruber [9]:

> An ontology is an explicit specification of a conceptualization.

In summary, an ontology can be considered as a formal specification of basic concepts (terms), properties, relations between different entities, as well as rules governing the relations and interdependencies between the entities in a particular domain of discourse. Ontologies can be modeled with different knowledge representation formalisms and can be implemented in different formal languages. For example, at the beginning of the 1990s, ontologies were modeled mainly by techniques based on frames and first-order logic [4]. In recent years, description logics have been used to model ontologies [10, 11]. It has also been suggested that other techniques that are widely used in software engineering and databases for conceptual modeling are also appropriate for building lightweight ontologies [4].

In recent years, the development of ontological engineering has been propelled and accelerated by the advancement of the World Wide Web and the emergence of the Semantic Web. As Berners-Lee et al. pointed out [5], ontology is an indispensable component of the Semantic Web. The Semantic Web enables more efficient information processing by describing resources on the World Wide Web with meta-data, so that the semantics of the resources as well as the relations between different resources can be understood by autonomous software agents which carry out information processing tasks on behalf of their human users. Ontologies play an important role in this technology, because they provide structures or models of known knowledge [12]. They specify the standard vocabularies for describing the available resources, and define the concepts and properties involved. With a suitable reasoning engine, software agents will be able to process information, discover implicit knowledge, or draw conclusions with the help of the definitions of concepts and relations in ontologies [12].

Since ontologies are so important in enabling the Semantic Web, the ability of ontologies to represent human knowledge of a particular domain in a precise and flexible way becomes a crucial aspect. In fact, there are quite a number of ontology models or ontology languages available when one wants to build an ontology [12]. In particular, it has been reported [13] that the DARPA Agent Markup Language and the Ontology Inference Layer (DAML+OIL) [14], the Resource Description Framework and Schema (RDF(S)) [14] and the Web Ontology Language (OWL) [16] are the three major ontology languages that are currently commonly used in the World Wide Web. These different ontology languages are characterized with different expressiveness and inference mechanisms. In general, a more expressive language or ontology model allows the ontology to model concepts and relations of higher complexity in a more efficient and flexible way. However, there is also tradeoff between expressiveness and tractability (computational complexity) in these models [3].

While these ontology models or languages provide standard methods for modeling knowledge of a particular domain, it is not difficult to note that these models suffer from certain limitations which avoid systems from providing better services on the Semantic Web. In this book, we investigate the limitations in conceptual modeling in existing ontology models, and propose possible extensions and solutions to these problems.

1.2 Motivations

There is no doubt that by using the ontology languages and models mentioned above we are able to model the known knowledge of a particular domain and are able to describe concepts and individual objects so that the underlying semantics become more explicit. For example, by using OWL, we can model the domain of publications, specify the common properties of the concept of 'publication', define 'magazine' and 'book' as subclasses of 'publication', so that they inherit all the properties of the concept [17]. Figure 1.1 shows a part of an OWL file describing the relationships among 'book', 'magazine' and 'publication', and Figure 1.2 represents the relationships among 'book', 'magazine' and 'publication' by a graph. Such an ontology will facilitate the task of processing information about publications with the help of autonomous software agents. Nevertheless, we notice that these ontology models are not without disadvantages or limitations.

```
-<owl:Class rdf:about="#Book">
 -<rdfs:sub Class Of>
   <owl:Class rdf:ID="Publication"/>
  </rdfs:sub Class Of>
  <rdfs:commentrdf:datatype="http://www.w3.org/2001/XMLSchema#string">This
  Class represents the concept of books.</rdfs:comment>
 -<owl:disjoint With>
  <owl:Class rdf:about="Magazine"/>
  </owl:disjoint With>
 </owl:Class>

 -<owl:Class rdf:about="#Magazine">
  -<rdfs:sub Class Of>
    <owl:Class rdf:ID="Publication"/>
   </rdfs:sub Class Of>
   <rdfs:commentrdf:datatype="http://www.w3.org/2001/XMLSchema#string">This
   Class represents the concept of magazines.</rdfs:comment>
 -<owl:disjoint With>
   <owl:Class rdf:about="Book"/>
   </owl:disjoint With>
 </owl:Class>
```

Fig. 1.1 Representing the relationships among 'book', 'magazine' and 'publication' in an OWL file.