Because the knowledge in the World Wide Web is continuously expanding, Web Knowledge Aggregation, Representation and Reasoning (abbreviated as KR) is becoming increasingly important. This article demonstrates how fuzzy ontologies can be used in KR to improve the interactions between humans and computers. The gap between the Social and Semantic Web can be reduced, and a Social Semantic Web may become possible. As an illustrative example, we demonstrate how fuzzy logic and KR can enhance technologies for cognitive cities. The underlying notion of these technologies is based on connectivism, which can be improved by incorporating the results of digital humanities research.

1. INTRODUCTION

Because the knowledge base of the World Wide Web (WWW) is continuously growing (i.e., big data’s volume), the velocity and variety of data (i.e., big data) flowing into companies often exceed the capacity of traditional information systems [Franks, 2012], which raises problems for knowledge management. Thus, Web Knowledge Aggregation, Representation and Reasoning (shortly KR), a form of knowledge management with the aim of presenting knowledge in an electronic format using intelligent information systems, is becoming increasingly important. Existing knowledge representation languages and reasoning methods face challenges due to the increasing volume of information [Kaltenrieder et al. 2014a; Kotoulas et al. 2012; Zeng et al. 2013]. This article presents possible ways how fuzzy logic and fuzzy ontologies can be used to enhance KR by addressing the uncertainty in data (i.e., big data’s veracity), and thus by improving the interactions between humans and computers [Portmann, 2013; Portmann & Kaltenrieder, 2015; Zadeh, 2006]. The following chapter explains basic notions, such as the Social Semantic Web, emergent semantics and fuzzy logic. In chapter 3, fuzzy logic and fuzzy ontologies are introduced into each step of KR. Cognitive cities are then presented as an illustrative example in chapter 4. Finally, we conclude this article and provide an outlook for future research in chapter 5.
2. Chapter Title

2. BACKGROUND

This chapter provides a brief overview of the basic concepts of this article. Section 2.1 elaborates on the Social Semantic Web and its issue that humans and computers do not use the same language. Section 2.2 presents the introduction of fuzzy logic and fuzzy set theory as a possible solution to this problem. The subsequent sections explain fuzzy ontologies and fuzzy grassroots ontologies (2.3), fuzzy cognitive maps and computational intelligence (2.4), and emergent semantics (2.5).

2.1 The Social Semantic Web

The Social Web (or Web 2.0) enables people to interact and create and share data over the Web [Greaves & Mika, 2008; Halpin & Tuffield, 2010], whereas the Semantic Web structures and describes the data in a manner that allows computers to process them, thus improving their cooperation with humans [Berners-Lee et al. 2001]. In the emerging Social Semantic Web (often called Web 3.0), which integrates both the Semantic Web and Social Web, social interactions create semantically rich knowledge representations that computers can understand [Blumauer & Pellegrini, 2009; Breslin et al. 2009]. The knowledge acquired by humans is often based on imprecise perceptions described by natural language, which is also imprecise. In contrast, computers require a clear and well-defined formulation to understand information, which raises difficulties for them to process natural language data [Portmann, 2013; Zadeh, 2006].

2.2 Fuzzy Logic and Fuzzy Set Theory

Fuzzy logic [Zadeh, 1988] can empower computers to understand data in natural language and can thus enable a Social Semantic Web [Portmann, 2013; Zadeh, 2006]. The term “fuzziness” refers to the ambiguity of semantics (i.e., the meaning of words) [Werro, 2008]. Fuzzy logic aims to formalize and mechanize the ability of humans to reason in an environment with imprecise concepts that are not sharply defined [Werro, 2008; Zadeh, 1988]. These concepts are captured by fuzzy set theory, which describes classes whose boundaries are not sharp [Zadeh, 1965]. In contrast, the membership function of a traditional crisp set can only take on the values 0 (i.e., does not at all belong to the set) or 1 (i.e., fully belongs to the set) [Zadeh, 1965]. A fuzzy set \( A \) in the space of objects \( X = \{x\} \) is defined by the function \( f(x) = \mu, \mu \in [0,1] \), which yields a membership grade of \( x \) in \( A \). A value of \( \mu \) that is closer to 1 has a higher membership grade.
grade [Zadeh, 1965]. The membership degree is often expressed using a linguistic variable, such as strongly, partially, or somewhat [Zadeh, 1988].

2.3 Fuzzy Ontologies and Fuzzy Grassroots Ontologies

An ontology specifies a common vocabulary to formally represent knowledge [Gruber, 1993; Gruber, 2008] and to classify it hierarchically, which supports the process of deciding [Parry, 2005]. Parry introduced the notion of fuzzy ontologies, which is based on fuzzy set theory [Zadeh, 1965] and assumes that each concept is related to every other concept with a certain degree of membership \( \mu \) [Parry, 2005]. Because fuzzy ontologies allow for the transformation of human perceptions into computer artifacts, computers should be able to perform tasks in the same manner as humans; therefore, a Web 3.0 could be realized [Portmann & Kaltenrieder, 2015].

Fuzzy Grassroots Ontologies (FGOs) are special cases of fuzzy ontologies that emerge bottom-up from Social Web data [Portmann, 2013]. A fuzzy clustering algorithm can be implemented to discover relationships between different data sources and transform data into computer-understandable ontologies [Gruber, 1993]. To accomplish this goal, Web agents constantly search the Web for terms and their connections. A metric is assigned to the terms based on the similarity of their semantic content (i.e., the distance \( d(x,y) \) between two terms \( x \) and \( y \)), which allows for a gradation of the relationships [Benedetto, 2013]. Then, the information is aggregated iteratively by granulating the terms to clusters using a fuzzy clustering algorithm (e.g., Fuzzy C-Means, Gustafson-Kessel, or FLAME [Portmann, 2013]). As the distance \( d \) to the cluster center increases, the term’s membership degree to the cluster decreases [Kaufmann et al. 2012; Portmann, 2013].

2.4 Fuzzy Cognitive Maps and Computational Intelligence

Fuzzy Cognitive Maps (FCMs) [Kosko, 1986] are highly suitable tools for managing fuzzy ontologies. An FCM is a signed weighted digraph, consisting of concepts (i.e., nodes) \( N_1, N_2, ..., N_n \) and connections between them (i.e., edges) represented by the adjacency matrix \( A \). In its simplest form, the concepts can take on values within \([0,1] \). The element \( a_{kl} \) of matrix \( A \) denotes the degree of the causal relationship (i.e., the weight of the edge) between \( N_k \) and \( N_l \) and lies between -1 and 1. If the causality between \( N_k \) and \( N_l \) is positive, \( a_{kl} > 0 \) (i.e., \( N_k \) increases \( N_l \)). If it is negative, \( a_{kl} < 0 \) (i.e., \( N_k \) decreases \( N_l \)). If there is no relationship, \( a_{kl} = 0 \). Multiple FCMs can be aggregated into...
a new FCM by combining their matrices if they have identical dimensions and related entries (i.e., if the FCMs have the same concepts). If there are \( z \) FCMs, each having a matrix \( A \), the aggregated matrix is \( A^{agg} = f(\sum_{i=1}^{z} A_i) \), where \( f \) is a sigmoid function that assures that the matrix entries stay between -1 and 1 [Kaltenrieder et al. 2014a; Kosko, 1997]. By adding a credibility factor (or weight) \( w_i \) to each FCM, the new aggregated FCM has the matrix \( A^{agg} = f(\sum_{i=1}^{z} A_i w_i) \) [Groumpos, 2010; Kaltenrieder et al. 2014a]. Because FCMs are dynamic, fuzzy-neural systems [Kosko, 1986; Salmeron, 2010; Zadeh, 1965], they can, for instance, be utilized as the learning part of Computational Intelligence (CI) methods [Portmann & Kaltenrieder, 2015]. CI combines biomimetic methods to process information with the goal of building intelligent systems. It combines neural networks, which are able to develop adaptive systems, evolutionary computation, which can solve structural optimization problems, and fuzzy logic [Engelbrecht, 2007; Pedrycz, 2002]. Granular Computing (GrC) is the underlying part of CI that allows computers in an information system to obtain different levels of knowledge and thus a better understanding of the knowledge structure (cf. emergent semantics in section 2.5) [Lin, 1997; Zadeh, 1998]. The GrC paradigm can be implemented in a straightforward manner using fuzzy ontologies [Pedrycz, 2002]. While searching the ontology, weights between 0 and 1 are assigned, which leads to different knowledge granules [Portmann & Kaltenrieder, 2015]. Granules are collections of objects that are grouped together due to their similarity [Pedrycz, 2013; Zadeh, 1998].

2.5 Emergent Semantics

Semantics (i.e., the study of meaning) is necessary to enable computers to generate, process and understand natural language [Portmann & Kaltenrieder, 2015]. Rapaport’s concept of syntactic semantics demonstrates that syntax can be sufficient to achieve this goal because it is able to adequately capture the meaning of signs [Rapaport, 2000; Rapaport, 2003]. Syntax studies the rules of the formation and manipulation of symbols [Posner, 1992] within a single domain (i.e., the syntactic domain), whereas semantics examines the relationships between symbols and their interpretations (i.e., the semantic domain). According to Rapaport [Rapaport, 2000; Rapaport, 2003], semantics is turned into syntax if the two domains are unified into one domain. Emergent semantics builds on this concept and exposes meaning from the way in which a sign is used [Cudré-Mauroux, 2009]. Emergence can occur if new structures arise spontaneously as a consequence of
the interaction between entities (e.g., if semantic knowledge bases arise automatically from Web 2.0 interactions via bottom-up ontologies [Kaufmann et al. 2012; Portmann et al. 2012]). Emergent semantics can be created using semiotic, inductive and approximate reasoning combined. A semiotic approach assumes that the representation of the meaning of the signs in the Web imitates the way in which the signs are processed and interpreted in the brain [Kaufmann et al. 2012]. An inductive approach expands the data by applying inferences that are not certain but only probable [Hawthorne, 2008]. The combination of these two approaches with approximate reasoning, which uses fuzzy variables, can generate diagrams of likely relationships between signs from natural language Web data and enables emergent semantics [Kaufmann et al. 2012].

3 WEB KNOWLEDGE AGGREGATION, REPRESENTATION AND REASONING

The processes of KR are shown in Figure 1. Knowledge from different sources is aggregated and represented in a computer-understandable format in a knowledge base, allowing for reasoning by humans as well as for automatic reasoning by computers [Kaltenrieder et al. 2014a]. Because only humans are capable of assessing their interests appropriately, a human-centered design should be pursued in every one of these steps [Portmann & Kaltenrieder, 2015]. Sections 3.1 and 3.2 explain different approaches to enhance knowledge aggregation and representation, respectively. Finally, section 3.3 elaborates on the concept of reasoning with knowledge by both humans and computers.

Fig. 1. The Web Knowledge Aggregation, Representation and Reasoning process.
6. Portmann, Kaltenrieder and Zurlinden

3.1 Knowledge Aggregation

One possible method of aggregating information is to use an FGO (see 2.3) managed by a graph database management system, which has been implemented into a search engine in online reputation analysis [Portmann, 2013].

Another possible approach is to combine top-down and bottom-up knowledge engineering, which has already been used to successfully create an enterprise search platform [Kaufmann et al. 2014]. Top-down methods are able to project existing ontologies from the conceptual level to the document level to structure the data and explicitly represent the corporate context. However, conventional software tools have difficulty automatically assigning resources to fuzzy concepts, and they require extensive manual maintenance. Bottom-up knowledge approaches are able to generate ontologies from data by extracting knowledge from different resources and clustering concepts in a concept graph. However, they cannot fully describe the semantics of concepts and their relationships, and they do not account for the organizational context. To overcome their shortcomings and to automatically generate a knowledge map, the two approaches can be combined by, for example, domain and enterprise ontologies [Kaufmann et al. 2014]. Domain ontologies capture the content of the information in an organization, and enterprise ontologies capture the organizational context of this information [Abecker et al. 1998]. The enterprise ontology can be used as a basic structure graph that is gradually extended with information extracted bottom-up by assigning a degree of membership [Zadeh, 1965] to each concept predefined in this ontology. This provides the flexibility in the concept mapping that is needed for the integration of the two ontologies and allows for a simple ranking of search results [Kaufmann et al. 2014].

A third possible approach is to wrap the underlying ontologies into FCMs, which has so far been used for stakeholder management [Portmann & Kaltenrieder, 2015]. The FCMs are aggregated to a new FCM model (see 2.4) that contains the knowledge of the initial, unaltered ontologies [Portmann & Pedrycz, 2014].

3.2 Knowledge Representation

Figure 2 shows the Semantic Web Stack, which presents the languages used to model knowledge in the Semantic Web [Hitzler et al. 2010]. The technologies in the lowest layer build the basis for the Semantic Web, whereas the middle layers contain standardized technologies to build Semantic Web applications. The Resource Description Framework (RDF) creates statements in the form of triples, consisting of two entities and their relationship.
their binary relationship, and presents information as graphs. RDF Schema (RDFS) is the basic vocabulary for RDF and is extended by the Web Ontology Language (OWL), which is able to describe the semantics of RDF statements. The Rule Interchange Format (RIF) provides rules to define relationships that cannot be described directly using OWL. RDF triples can be stored in triplestores, which are queried by the SPARQL Protocol and RDF Query Language (SPARQL) to retrieve information. The top layer contains concepts that have yet to be implemented to realize the Semantic Web [Portmann, 2013].

Because fuzzy logic [Zadeh, 1988] accounts for the imprecision in human reasoning, it can address this concept in a more humanlike fashion and thus enable an improved Semantic Web or even a Social Semantic Web. Fuzzy logic consists of fuzzy sets (see OWL/RDF(S)/SPARQL in Fig. 2) and fuzzy rules (see RIF in Fig. 2), whereby the information obtained from fuzzy sets is combined using fuzzy rules. Fuzzy sets and thus FCMs [Kosko, 1986] can enhance RDF(S)/OWL/SPARQL, whereas fuzzy rules can improve RIF.

3.3 Reasoning with Knowledge

Intelligence amplification via human-computer interactions (HCI) could enable a (Social) Semantic Web. Humans can learn from computers with the help of Web semantics maps, which are based on emergent ontologies and visualized as graphs, whereas computers can learn from humans by processing data in OWL/RDF(S) formats [Kaufmann et al. 2012; Portmann et al. 2012].

The Semantic Web enables both knowledge-based reasoning by humans and automatic reasoning by computers. RDF links are suitable for knowledge-based reasoning because
they allow for the browsing of knowledge structures, instead of only hyperlinks, and the navigation between data sources [Bizer et al. 2009]. FCMs [Kosko, 1986] can be useful for knowledge-based reasoning because conclusions can be drawn from the knowledge stored in the ontologies [Portmann & Kaltenrieder, 2015]. A Graphical User Interface (GUI) enables interactions between the user and the FCM, so that the user can browse it, search it by keyword, change it, and view the resources related to the concepts [Kaufmann et al. 2014; Portmann, 2013]. Functions inspired by GrC [Lin, 1997; Zadeh, 1998] (e.g., zoom or drag-and-drop) support the efficient investigation of information and the automatic addition of new analytical content [Portmann, 2013].

Based on fuzzy ontologies stored in a graph database as FCMs, reasoning can enable automatic logical consequences and thus automatic reasoning [Portmann & Pedrycz, 2014]. Graph databases are powerful tools for graph-like queries, such as finding (the shortest) paths between two nodes, finding clusters, or computing diameters. Therefore, a simple implementation of an automatic reasoner (e.g., interactive and iterative approaches to decision making) could be built on a query language. Clicks made by the user on the dynamic user interface would be a learning input for the reasoning system. Operations including fuzzy logic [Zadeh, 1988] imitate human reasoning and could thus improve the underlying reasoning interfaces [Portmann & Kaltenrieder, 2015].

4 COGNITIVE CITIES – A POSSIBLE APPLICATION

This chapter first presents connectivism, a cognition theory, as the underlying theory for cognitive cities that, for example, use KR technologies to improve the interactions with their citizens. Thereby, digital humanities is emphasized as important supplier of information for the implementation of cognition theories and thus for the development of cognitive cities technologies.

According to connectivism, the increasing knowledge base makes learning building solely on one’s own experiences impossible. Instead, people must rely on experiences of others, making technology as well as connections with others essential for learning [Siemens, 2005]. Connectivism is closely linked to the theory of complex systems that investigates how collective behaviors of a system arise from the relationships between its parts and how the system interacts with the environment [Zadeh, 1973; Mostashari et al. 2011]. The idea of cognitive cities is based on these theories [Mostashari et al. 2011; Kaltenrieder et al. 2014b]. Fuzzy-logic-based KR technologies [Zadeh, 1988] can be used
to improve the interactions between a city and its citizens. The urban ontology could be the basic structure of the semantic net, from which the knowledge structure would emerge bottom-up from user inputs (e.g., from social media data about citizens activities) and machine learning. The result would be a Web-based, interactive knowledge map of the city. The user can browse all concepts relevant to his urban life in his city and expand the knowledge graphs directly [Kaufmann et al. 2014]. Another useful technology for cognitive cities could be a mobile application (app) based on intelligence amplification through HCI and GrC [Kaltenrieder et al. 2014b]. This app would process large amounts of data in a short time, aggregate the data, manage the data’s fuzziness and visualize the data using FCMs, which would allow for human-like reasoning and improved citizen management [Kaltenrieder et al. 2014b].

Within digital humanities research, it is possible to develop cognitive structures that include knowledge from the past, which may be important for cognition theories and thus for the development of cognitive cities technologies. The digital humanities cover all research at the intersection of computing and the humanities [Burdick et al. 2012] based on the notion that digital technologies provide new ways to represent and process data. Because data in the humanities is heterogeneous and ambiguous, whereas computational methods must be homogeneous and contain clearly defined data, some information must be neglected, which introduces the possible danger of oversimplification [Berry, 2011]. Because fuzzy logic [Zadeh, 1988] processes information described in natural language, it can account for the ambiguity of knowledge in the humanities. Thus, it is able to reduce this danger and combine humanistic research and computation.

5 CONCLUSION AND OUTLOOK

In this article, several approaches to enhance KR using fuzzy logic (more specifically, fuzzy ontologies and FCMs) are considered. Fuzzy logic can facilitate computers to process natural language data, which may enable a Social Semantic Web. Successful implementations thus far have been in online reputation analysis, intranet search platforms or stakeholder management. Cognitive cities are presented as an illustrative example of a complex system for applications of KR technologies, which are based on connectivism and can be enhanced via digital humanities research.

The accelerating creation of data results in both opportunities and difficulties for businesses. New technologies and new paradigms are necessary to handle and utilize
large amounts of data. The approaches presented in this article constitute first steps in this direction, but they must be pursued further, and solutions that lead to automatic reasoning must be developed.

We would be happy to answer any questions and provide more information about the presented research areas and projects (portmann.iwi.unibe.ch).

ACKNOWLEDGMENTS

We want to thank the Fuzzy Marketing Methods Research Center (FM², fmsquare.org) at the University of Fribourg; the Berkeley Initiative in Soft Computing (BISC, www-bisc.cs.berkeley.edu/) at the University of California at Berkeley; the Swiss National Science Foundation (SNF, www.snf.ch); the Commission for Technology and Innovation (CTI, www.cti.admin.ch); and the Institute of Information Systems (IWI, www.iwi.unibe.ch) at the University of Bern, Switzerland, for their valuable help.

REFERENCES


Fuzzy Ontologies and the Social Semantic Web


12. Portmann, Kaltenrieder and Zurlinden


Prof. Dr. Edy Portmann is assistant professor of information science at the University of Bern in Switzerland. His research deals with questions relating to information systems, processing and procurement.

Patrick Kaltenrieder, a PhD student, and Noémie Zurlinden, a Master student, work at the Institute of Information Systems at the University of Bern. Patrick Kaltenrieder’s research deals with knowledge management, especially KR in the fields of stakeholder management and cognitive cities.

SIGWEB Newsletter Autumn 2007