ARTICLE IN PRESS

Engineering Science and Technology, an International Journal xxx (2017) xxx–xxx

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

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

Full Length Article

# A fuzzy ontology modeling for case base knowledge in diabetes mellitus domain

Shaker El-Sappagh [a], Mohammed Elmogy [b,*]

[a] Faculty of Computers & Information, Minia University, Egypt
[b] Faculty of Computers & Information, Mansoura University, Egypt

ARTICLE INFO

*Article history:*
Received 29 December 2016
Revised 22 March 2017
Accepted 27 March 2017
Available online xxxx

*Keywords:*
Case-based reasoning
Semantic retrieval
Case base representation
Fuzzy ontology
Diabetes diagnosis

ABSTRACT

Knowledge-Intensive Case-Based Reasoning Systems (KI-CBR) mainly depend on ontologies. Ontology can play the role of case-base knowledge. The combination of ontology and fuzzy logic reasoning is critical in the medical domain. Case-base representation based on fuzzy ontology is expected to enhance the semantic and storage of CBR knowledge-base. This paper provides an advancement to the research of diabetes diagnosis CBR by proposing a novel case-base fuzzy OWL2 ontology (CBRDiabOnto). This ontology can be considered as the first fuzzy case-base ontology in the medical domain. It is based on a case-base fuzzy Extended Entity Relation (EER) data model. It contains 63 (fuzzy) classes, 54 (fuzzy) object properties, 138 (fuzzy) datatype properties, and 105 fuzzy datatypes. We populated the ontology with 60 cases and used SPARQL-DL for its query. The evaluation of CBRDiabOnto shows that it is accurate, consistent, and cover terminologies and logic of diabetes mellitus diagnosis.



## 1. Introduction

Knowledge Intensive Case-Based Reasoning (KI-CBR) success depends on its Case Base (CB) structure and content. Ontology constructs domain knowledge in a machine-readable format that humans are capable of understanding. It can be used in CBR systems as a knowledge representation formalism for specifying domain background knowledge and CB knowledge [6]. KI-CBR enables automatic reasoning with semantic knowledge in addition to the syntactic properties of cases. Ontology supports the creation of semantic retrieval algorithms to enhance the intelligence of CBR systems [14]. It provides a conceptualization of the domain, which consists of concepts, properties, and axioms. It has been utilized in many medical [15] and non-medical [6] CBR systems.

For diabetes mellitus (DM) diagnosis, ontology has not been utilized in CB, background knowledge, and case retrieval. A crisp CB OWL 2 ontology for DM diagnosis is proposed by El-Sappagh et al. [22], a background domain ontology based on SNOMED CT[1] (SCT) [23], a CB preparation model from Electronic Health Record (HER) data [24], and encoded the CB based on a proposed encoding methodology [27]. Nevertheless, the knowledge imprecision issue has not been handled in KI-CBR paradigm yet [3]. Medical data, such as DM diagnosis, are mostly imprecise and experience-based. The success of CBR in this domain depends on how this issue is handled.

According to Zadeh [63], much of the human experience knowledge is imprecise and inaccurate. This knowledge has to be manipulated in a suitable way to prevent the poor effectiveness of knowledge-based systems. Fuzzy sets have been integrated with CBR to generate Fuzzy-CBR in many studies as [31]. However, they did not utilize Fuzzy Ontology (FO) for case representation and case retrieval causes CBR systems, which can lose many semantic reasoning capabilities [3]. A mechanism is required to utilize ontology as the "vehicle" for the introduction of fuzzy semantics to KI-CBR. FO integrates fuzzy set theory into crisp ontology logic, tools, and languages [18]. Bobillo and Straccia [9] proposed an extension for crisp OWL 2 to generate fuzzy OWL 2 ontologies. The fuzzification of a crisp ontology needs translation into a supported language of an FO reasoner such as *fuzzyDL* [9]. We assert that FO supports crisp aspects as well as fuzzy aspects. For example, in DM diagnosis domain, there are *crisp components* (such as sex and residence), *fuzzy components* (such as age and lab tests), and *semantic components* (such as diseases), which are related to other ontology, such as SCT.

Fuzzy ontologies have been used in many domains. For example, Rodríguez et al. [57] proposed an FO to model human behavior. Ali et al. [5] used it in opinion mining. Lee et al. [43,44] used FO and fuzzy set for modeling diabetes application in diet and diagnosis.

* Corresponding author.
*E-mail address:* melmogy@mans.edu.eg (M. Elmogy).
Peer review under responsibility of Karabuk University.
[1] *http://www.ihtsdo.org/snomed-ct/*.

http://dx.doi.org/10.1016/j.jestch.2017.03.009
2215-0986/© 2017 Karabuk University. Publishing services by Elsevier B.V.

Moreover, it has been managed in some AI systems as rule-based systems [45]. The usage of FO with rule-based systems is less applicable than other systems because these systems require collecting explicit models of domains. For DM, it is a challenge to collect the set of rules that model it. Therefore, it is possible to develop CBs that avoid the knowledge-acquisition bottleneck. FO extends the capabilities of the crisp one and improves the accuracy and applicability of CBR in medical domain [3].

Park et al. [52] utilized crisp ontology in a fuzzy CBR system for the prevention ships collision. Alexopoulos et al. [3] tried to build an FO for CBR using fuzzy algebra. Their proposed ontology has been designed for electronic libraries. However, utilizing relational databases and conceptual data models for building fuzzy ontologies is the most suitable form to build CB fuzzy ontology. Consequently, the patient CB is often gotten from EHR database of patient histories [10]. There is little research in FO in any medical CBR systems. As far as we know, there are no studies in the literature that proposed a DM diagnosis fuzzy CB ontology.

Fuzzy CB ontology can be created from many sources [64]. The translation of fuzzy EER model other than its relational schema is better because EER is richer in semantic, and EER model is closely related to ontology structure [64]. Moreover, many studies proposed mappings from fuzzy EER model to FO [46]. Besides, chronic diseases management requires collecting cases from the patient's (distributed) EHR. El-Sappagh et al. [25] proposed a case base data model using HL7 RIM.[2] A method is required to fuzzify this model and map it to a fuzzy CB ontology. Because ontology provides a formal representation of a shared conceptualization, the CB FO can be combined with DM standard domain ontology to support the sharing and interoperability with EHR environment [23]. Although there are many studies in fuzzy EER models creation and translation into fuzzy ontologies and fuzzy DLs [46,47], few studies utilized their results in CBR especially in medical domains as DM diagnosis. Most fuzzification strategies of EER models depend on one linguistic value to represent a feature [12]. These approaches limit the flexibility of query problem description. They limit the retrieval algorithms efficiency. The paper proposes a strategy to solve this issue.

In this paper, we designed a new ontology called CBRDiabOnto. It is an OWL 2 fuzzy CB ontology to represent DM diagnosis knowledge for any CBR system. The paper concentrates on the knowledge representation formalism of CB fuzzy ontology. This ontology has been used to build a whole CBR system [26]. Different types of features including text, ordinal, semantic, numerical, and fuzzy types can be modeled by the proposed ontology. As a result, the resulting knowledge base supports many types of reasoning, such as semantic, crisp, and fuzzy reasoning, in fuzzy CBR system. The rest of this paper is organized as follows. Section 2 provides the process of constructing CBRDiabOnto fuzzy CB ontology. Section 3 is the querying method for the CB ontology. Section 4 evaluates the proposed fuzzy ontology. Finally, Section 5 presents the conclusion and future work.

## 2. Related work

Traditional CBR has been used for diabetes diagnosis in many studies [48,40]. An evolution of this paradigm is the ontology-based or KI-CBR [6]. Ontology plays many roles to enhance CBR semantics [20]. This approach is generally more effective in representing, indexing, maintaining, adapting, and retrieving similar cases than traditional ones [30,37]. Case semantic retrieval can be enhanced by using case-base and domain background knowledge in the form of ontologies [30]. Amailef and Lu [6] proposed a domain ontology for m-government emergency response services based on CBR. El-Sappagh et al. [23] proposed a diabetes domain ontology for CBR system based on SNOMED CT. However, the CBR system fundamentally depends on the case-base knowledge representation [30]. As a result, Heras et al. [38] proposed a crisp case-base ontology for case-based argumentation system entitled ArgCBROnto. Zhukova et al. [65] proposed a crisp case-base ontology for a human resource management system. In the diabetes domain, ontologies have been used in many CDSS [61,54]. However, regarding diabetes diagnosis, none of the existing ontologies is designed for CBR, and few studies have used ontology in CBR [61,40]; these studies are very abstract and use ontology only for domain knowledge. Jaya and Uma [39] listed the roles of ontology in a diabetes diagnosis CBR. El-Sappagh et al. [22] proposed a crisp OWL 2 ontology for diabetes diagnosis case-base. This ontology can be used to store and retrieve cases semantically. On the other hand, diabetes diagnosis depends on the physician's experience and the patient's description of his case. Vagueness in medical domains can be handled using fuzzy logic [63], which has been used in diabetes diagnosis rule-based systems [45]. Moreover, fuzzy logic has been integrated with CBR in hybrid systems [1] and used for calculating the fuzzy similarity between cases [41]. Recently, Sohn et al. [59] integrated fuzzy CBR reasoning with crisp ontology reasoning for personalized service in a smart home environment. However, this hybrid system did not benefit from fuzzy ontology reasoning capabilities.

After the success of crisp ontologies in CBR environment, fuzzy ontologies can extend the benefits of crisp ontologies [4]. Building a case-base fuzzy ontology is a challenge [52]. Alexopoulos et al. [3] proposed a fuzzy case-base ontology by utilizing fuzzy algebra. With respect to diabetes, it has utilized fuzzy ontologies in many domains [45,44]. Lee and Wang [45] proposed a five-layer fuzzy ontology and utilized it in a fuzzy expert system for diabetes management. Lee et al. [44] proposed a type-2 fuzzy ontology for diet management in diabetics. Although CBR is more suitable for managing ill-formed problems as diabetes diagnosis, there is no fuzzy ontology-based CBR for diabetes, and there are no studies for case-base fuzzy ontology construction. Crisp ontologies are not suitable to address imprecise and vague knowledge, which is inherent in real-world domains [57].

There are many techniques for developing crisp ontology [22]. However, they are not sufficient for fuzzy ontology construction [64]. Many techniques are available to generate fuzzy ontology including: extend an existing crisp ontology [4], map relational database schema, map fuzzy ER model [64], map fuzzy EER model [64], fuzzy formal concept analysis [19], and other data sources. The difference between these approaches is mainly in what aspects of the crisp ontology are being fuzzified, and these aspects depend on the domain needs. Representation languages such as OWL 2 need to be able to represent fuzzy ontologies. There are two ways either extend the language itself [60], which changes the structure of the language or using the current language to represent fuzzy aspects. Bobillo and Straccia [9] have used annotations to represent fuzzy in regular OWL 2.[3] We will depend on the second choice by proposing a fuzzy case-base ontology structure (TBOX) based on a case base EER model. This EER model is extended to fuzzy EER model based on fuzzy set theory and then converted to fuzzy OWL 2 ontology. Moreover, we populate the resulting ontology with the case from EHR database (ABOX), and we provide a way for semantically querying the resulting case base ontology. The paper has the following contributions:

1. We proposed an EER data model for the DM diagnosis CB according to our CB schema [26].

[2] http://www.hl7.org/implement/standards/rim.cfm.

[3] http://www.straccia.info/software/FuzzyOWL/.