IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 30, 2018, accepted June 27, 2018, date of publication July 4, 2018, date of current version July 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2852004

An Ontology-Based Interpretable Fuzzy Decision Support System for Diabetes Diagnosis

SHAKER EL-SAPPAGH[1,2], JOSÉ M. ALONSO[3], FARMAN ALI[1], AMJAD ALI[1,4], JUN-HYEOG JANG[5], AND KYUNG-SUP KWAK[1], (Member, IEEE)

[1]Department of Information and Communication Engineering, Inha University, Incheon 22212, South Korea
[2]Information Systems Department, Faculty of Computer and Informatics, Benha University, Banha 13518, Egypt
[3]Centro Singular de Investigacion en Tecnoloxias da Informacion, Universidade de Santiago de Compostela, 15703 Santiago, Spain
[4]Department of Computer Science, COMSATS Institute of Information Technology, Lahore 54000, Pakistan
[5]Department of Biochemistry, School of Medicine, Inha University, Incheon 22212, South Korea

Corresponding author: Kyung-Sup Kwak (kskwak@inha.ac.kr)

This work was supported by the National Research Foundation of Korea Grant through the Korean Government (Ministry of Science, ICT and Future Planning)-undr Grant NRF-2017R1A2B2012337.

ABSTRACT Diabetes is a serious chronic disease. The importance of clinical decision support systems (CDSSs) to diagnose diabetes has led to extensive research efforts to improve the accuracy, applicability, interpretability, and interoperability of these systems. However, this problem continues to require optimization. Fuzzy rule-based systems are suitable for the medical domain, where interpretability is a main concern. The medical domain is data-intensive, and using electronic health record data to build the FRBS knowledge base and fuzzy sets is critical. Multiple variables are frequently required to determine a correct and personalized diagnosis, which usually makes it difficult to arrive at accurate and timely decisions. In this paper, we propose and implement a new semantically interpretable FRBS framework for diabetes diagnosis. The framework uses multiple aspects of knowledge-fuzzy inference, ontology reasoning, and a fuzzy analytical hierarchy process (FAHP) to provide a more intuitive and accurate design. First, we build a two-layered hierarchical and interpretable FRBS; then, we improve this by integrating an ontology reasoning process based on SNOMED CT standard ontology. We incorporate FAHP to determine the relative medical importance of each sub-FRBS. The proposed system offers numerous unique and critical improvements regarding the implementation of an accurate, dynamic, semantically intelligent, and interpretable CDSS. The designed system considers the ontology semantic similarity of diabetes complications and symptoms concepts in the fuzzy rules' evaluation process. The framework was tested using a real data set, and the results indicate how the proposed system helps physicians and patients to accurately diagnose diabetes mellitus.

INDEX TERMS Clinical decision support system, diabetes diagnosis, fuzzy inference system, ontology reasoning, fuzzy interpretability.

I. INTRODUCTION

Diabetes mellitus (DM) is a group of chronic metabolic diseases. According to the International Diabetes Federation[1] (IDF), the number of diabetics is expected to surpass 642 million by 2040. The undiagnosed cases of DM are as high as 179 million [1]. According to the IDF, an estimated US $612 billion was expended globally on diabetes care in 2014. Diabetes is associated with an increased risk of morbidity and mortality. According to the World Health Organization (WHO), 4.6 million deaths were attributed to DM in 2011, and this will be the seventh leading cause of death in 2030.[2] Unacceptable morbidity and mortality of DM continue to be recorded in all countries. DM cannot be cured, and the effectiveness of its therapy is mainly dependent on the level of accuracy and the timing of its diagnosis. Early detection of this disease allows physicians to recommend specific treatment plans. This results in reduced morbidity and mortality associated with this major health problem; however, DM is a silent, insidious, asymptomatic, ''theoryless'', and experience-based disease. It is further complicated by the imprecision, ambiguousness, and uncertainty associated with its features. A patient can suffer from diabetes

[1]*www.idf.org*

[2]*http://www.who.int/mediacentre/factsheets/fs312/en/*

for a long period without knowing [2]. Diagnosis involves a complex mental process using numerous variables, which usually makes it difficult to arrive at an accurate and timely diagnosis. In the majority of cases, patients suffer from multiple micro and macro vascular complications at the time of diagnosis; knowing this can certainly affect a physician's decision regarding the diabetes risk level. Further, the lab tests used to measure blood glucose levels such as HbA1c are affected by other medications. Moreover, patients can suffer from other symptoms at the time of diagnosis.

Consequently, DM diagnosis is a complex process and requires medical experts to make their decision based on the complete patient profile. Bickley [3] asserted that for diagnosis, physicians must first gather the signs of a disease, including a patient's past medical history, symptoms, family history, related complications, and physical examinations; finally, complementary explorations such as laboratory tests can increase or decrease the likelihood of a diagnosis. There is a lack of experts on DM, and the majority of them have insufficient time to study the patient history in detail. Further, hospitals must always optimize the diagnostic process in terms of the number and duration of patient examinations. An automated clinical decision support system (CDSS) can be used to assist non-expert physicians in the collection and analysis of the distributed patient profile from different electronic health records (EHR), which can derive from different hospitals, and require additional accurate and timely diagnostic decisions. It is sine qua non to reduce possible diabetes complications, hospitalization, morbidity, and mortality. For standalone CDSS, physicians are required to manually collect and enter these features for each patient; however, this is not acceptable because it could interrupt the clinical workflow of the physicians. To provide the correct knowledge in the proper form at the correct time, CDSS must be integrated as a plugin in the EHR ecosystem [4]. Further, patients can use the outcomes of CDSS in a mobile health environment to continuously monitor their signs.

Diabetes diagnosis is a reasoning process. The diagnosis d is a function $d = f(t_1, t_2, \ldots, t_n)$, for n features of current and historical patient conditions, where $t_1, \ldots, t_a$ are symptoms, $t_{a+1}, \ldots, t_b$ are physical examinations, $t_{b+1}, \ldots, t_c$ are lab tests, $t_{c+1}, \ldots, t_d$ are complications, $t_{d+1}, \ldots, t_e$ are drugs, and $t_{e+1}, \ldots, t_n$ are demographics, $a + b + c + d + e = n$. These features are of different types, including numerical, categorical, and semantic. The literature of diabetes diagnosis has many dimensions. The majority, if not all, of the existing data mining and machine learning (ML) studies concentrate on one type of data to diagnose DM, viz. numerical or continuous data [2], [5]; this is medically unacceptable and insufficient. Consequently, physicians do not trust the system results. Moreover, although the majority of models efficiently arrive at an accurate prediction, their structure and reasoning process are typically not transparent, complex to understand, and can only be used as "*black boxes*". The detection and inference processes cannot be explained [6]. Thus, it is essential to use models that have interpretability in knowledge formalization. Further, the diagnosis of chronic diseases is a complex problem where accurate prediction of the target from observed values (e.g., symptoms and lab tests) is impossible [7]. The complexities in medical practice make the conventional quantitative diagnosis approaches inadequate and hence call for new techniques. Fuzzy rule-based systems (FRBSs) have high interpretability and provide accuracy comparable to ML algorithms [2], [8]. They support approximate reasoning under vagueness and the computing with words paradigm [9], [10], which facilitate the explanation of the system results. However, many challenges must be considered to build an applicable FRBS. First, interpretability is not granted by the adoption of FRBS [7], [11], and a careful design methodology must be followed. Second, the system must consider all of a patient's current and historical features in his profile to provide accurate, personalized, and customized decisions. In the diabetes domain, there are two main types of data: numerical and textual (or unstructured) data. Some FRBS systems model categorical features as high-level abstract groups such as "*obesity*" and depend on the direct match between patient features and rules [12]. The resulting systems are not portable and not sufficiently intelligent because they cannot consider the semantic meaning of medical concepts. They depend mainly on their users to determine whether patients have specific characteristics. To enhance human–machine interaction, FRBS must be accompanied by ontology to represent the semantic structure of medical concepts in a hybrid intelligent system. Some studies apply this integration by designing FRBS and ontology as separate modules, which does not improve fuzzy systems [13]. Other studies have built fuzzy ontologies [14] to integrate fuzziness in ontology reasoning; however, all existing fuzzy description logics have decidability limitations as recently stated by Borgwardt *et al.* [15], and existing fuzzy ontology tools, parsers, and reasoners have not yet attained a mature state [16].

The current study integrates two mature reasoning techniques (i.e., ontology and fuzzy logic) in a novel manner. It attempts to integrate ontology reasoning in the fuzzy rule evaluation process. There is no such system in the literature. It is expected that the resulting system will demonstrate an acceptable interpretability–accuracy trade-off; it can be semantically interoperable with EHR to automatically and transparently collect patient profile information. Further, because it can dynamically collect features, it can be applied to collect patient diagnosis-related data from social media such as Facebook [17] and Twitter [18]. Consequently, for remote monitoring of a patient, he/she can be connected with a network of wireless sensors to collect their current signs; their profile can be automatically collected from EHR; and finally, the FRBS can provide them continuously with personalized and accurate decisions. These types of systems support the implementation of the chronic care model developed by Wagner *et al.* [19]. There are two main challenges that will be addressed in this paper to build a semantically interpretable FRBS—*system interpretability* and *semantic reasoning*.

A. INTERPRETABILITY CHALLENGE

The first challenge is to efficiently address numerical data. Medical data are typically imprecise and the chronic disease diagnosis problem is always complex, ill defined, and nonlinear [14]. Mansourypoor and Asadi [2] proposed an FRBS for diabetes diagnosis; however, they failed to discuss the real problem of diabetes diagnosis. They assumed that all features of a patient are numerical in nature, which is not the real environment. EHR can provide the required data to build the initial FRBS and to continuously enrich physician real-time queries. FRBSs that mimic the medical expert in both knowledge representation and reasoning process support the creation of interpretable CDSSs. A granular computing methodology such as FRBS is widely used in system modeling either for classification or regression purposes. It is a knowledge-based system, and their acknowledged originality derives from the linguistic interpretability of fuzzy rules that can attain a comparable level of accuracy to the other ML algorithms. They satisfy the universal approximation property and their inference engine implements approximate reasoning [8]. The implementation is realized in the form of a set of potentially interpretable IF-THEN rules. Interpretability means transparency and intelligibility [20]. It is essential for these systems to include high human interaction because it facilitates the understanding of the system outputs. We distinguish two trends in FRBS design. First, the expert-driven approach or linguistic fuzzy modeling, where knowledge is directly injected into the system by a human expert. It focuses on the system interpretability; the Mamdani model is commonly implemented. However, several difficulties can be encountered when the domain is complex or experts do not exist [21], [22]. Conversely, the data-driven approach or precise fuzzy modeling (PFM), where knowledge is autonomously discovered and extracted by the system from experimental data examples, focuses on system accuracy. These trends are not mutually exclusive and hybrid approaches have been explored [23]. For complex and data-intensive problems where experts cannot express the full knowledge of the domain, the PFM method is more suitable to produce accurate systems. To achieve interpretability and accuracy of data-driven FRBSs, several challenges must be addresses concerning the fuzzy partitioning and rule-based definition. According to the data used, the generated fuzzy sets and rules must be analyzed and interpreted. Consequently, rules and fuzzy sets can be modified, removed, or new rules added. Moreover, a trade-off between accuracy and interpretability is required to be measured and determined while designing an FRBS [24]. Consequently, fuzzy system design to generate accurate, fast, consistent, interpretable system is a difficult challenge [11].

B. SEMANTIC REASONING CHALLENGE

The second challenge is to address the textual features efficiently. Building an FRBS based only on the fuzzification of numerical features is not adequate, and addressing unstructured or textual features in a semantically intelligent manner is of significant importance. Anderson *et al.* [25] studied the efficiency of DM diagnosis CDSS based on the full EHR data (e.g., medications, diseases, lab tests, symptoms) and asserted the improvements of the resulting classification. Liaw *et al.* [26] stated that integration of EHR data in diabetes diagnosis CDSS improved the resulting system accuracy, and using ontology supported the development of automated systems. Addressing unstructured data in a static fashion is not efficient because the resulting system cannot be integrated into the EHR ecosystem, it cannot be used in the mobile health environment, and it provides reduced accuracy. Further, it depends on the exact matching between the concepts used in the rule base and those in the patient's profile. For example, assume a simple fuzzy rule of the form *IF <HbA1c is high and current disease is "Cardiovascular disease"> THEN <diagnosis = "diabetes mellitus">*. Now, if we have a new case with *<HbA1c = high and current disease = "Aneurysm">*, the case will not fire this rule, which is not correct.

The CDSS must provide maximal automatic information processing for physicians [27]. One possible solution to these limitations is to combine the reasoning capabilities of the stable FRBSs and crisp medical ontologies. The resulting system could exhibit highly sophisticated reasoning capabilities, and, thus, promote more efficient care practices. *As far as we know, the literature lacks a hybrid intelligent system to integrate and combine these reasoning powers, even in the diabetes domain.* Chen and Huang [28] used fuzzy and ontology-based reasoning to automatically generate the weather news using two features captured from ontology and a third statistical feature. However, the two techniques function in an isolated manner, with no cooperation. Ontology is based on the description logics formalism and reasoners capability. Patient data are frequently distributed over EHR systems and use different formats with different levels of granularities (i.e., different semantics) [4]. The automation of data collection to generate the patient's profile requires a standardization of the terminology used to solve the interoperability and portability problems [29], and enhance semantic interpretability of the resulting system. Several proposals have been published to permit the collection of CDSS patient data from EHR [30]. In the diabetes domain, the other type of data is the textual features such as a patient's current medications, diseases, and symptoms; if we consider these as regular categorical features, their semantic would be lost. Let us pose a simple example. If the patient has *hypertension*, which affects the physician's decision regarding DM diagnosis, this disease may be not stored in EHR with this name. Hypertension has many other synonyms and related diseases. For example, SNOMED CT[3] (SCT) has 178 other diseases such as "*eclampsia*" disease, which are considered subtypes or related to *hypertension*. As another example, in 2016, the Canadian Diabetes Association (CDA)

[3] *https://www.snomed.org/snomed-ct*

guidelines asserted that diabetes diagnosis rules must search for the possible existence of "*cardiovascular disease*." However, according to an SCT 2016 release, there are more than 6320 medical concepts and diseases related to "*cardiovascular disease.*" When CDSS automatically collects patient features from a distributed EHR, it must detect these semantic relationships between collected concepts (i.e., diseases, symptoms, and medications) and the concepts modeled in the CDSS knowledge base to mimic the reasoning of doctors. Chen *et al.* [31] proposed a recommendation system for anti-diabetes drug recommendation based on fuzzy logic and ontology. They tested the conditions of hypoglycemia, liver, renal, and heart to determine a condition. However, there are no semantics in their tests. For example, they tested if the heart, liver, and renal are normal or abnormal. Users must provide manually and precisely related parameters with their specific names. This design makes the resulting CDSS standalone and not interoperable with EHR systems. Tsipouras *et al.* [32] proposed a fuzzy CDSS for coronary artery disease (CAD) diagnosis; their dataset has features for the patient's history of *diabetes mellitus*, *hypertension*, and *hyperlipidemia*. However, the authors considered the data as binary (true/false), with no semantic. All these diseases have hierarchies of related diseases. For example, *hyperlipidemia* has a sub-tree of 56 sub-diseases in SCT. The resulting system would not be sufficiently smart to detect that *hyperalphalipoproteinemia* must be treated as *hyperlipidemia*. Pal *et al.* [33] proposed a fuzzy CDSS for CAD and verified if the patient had specific diseases such as *hypertension*. However, they searched for a disease by checking its medical tests, not using the disease name. To the best of our knowledge, however, the literature is insufficient concerning the integration of ontology reasoning as an embedded component in the FRBS inference process.

C. MAIN CONTRIBUTIONS

Integration of semantic ontology-based reasoning with FRBS improves interpretability and semantic interoperability of the final CDSSs. It provides suitable mechanisms of knowledge representation and a computational machinery for human-like reasoning. This provides a different viewpoint of fuzzy systems toward explainable artificial intelligence [34], [35]. This endows FRBS with the abilities to address highly complex medical problems and to share knowledge with distributed environments. The main contributions of this paper are:

1) Propose a semantically intelligent hierarchical FRBS able to provide accurate and semantically interpretable decision support by integrating fuzzy reasoning and semantic ontology reasoning in a novel manner. The proposed framework can overcome many restrictions in medical diagnosis caused by the nature of medical data in the patient profile.
2) Careful design of each single FRBS of the six subsystems by applying an accurate fuzzy modeling methodology. It combines domain experts and learns from the real medical data the set of linguistic variables, strong fuzzy partitions, and fuzzy rule bases in an accurate and consistent manner.
3) Incorporate a fuzzy analytical hierarchy process (FAHP) in the design process to determine the relative medical importance or medical weight of every sub-FRBS in the framework. These weights are integrated in the fuzzy inference process.
4) Propose an extension of the previous system by incorporating weights of FAHP for every subsystem. This weighted system supports the isolation of any number of subsystems in the decision making process.
5) Implement and evaluate the resulting system, highlight its limitations, and the reasons for integrating ontology-based semantic intelligence. The resulting system is coded based on the Java Fuzzy Markup Language (JFML),[4] a Java library for the IEEE Standard for Fuzzy Markup Language (IEEE Std 1855-2016). This is the first standard approved and supported by the IEEE Computational Intelligence Society [36], [37]. This format enables the implementation of distributed systems and enhances the interoperability between CDSS and EHR systems.
6) Propose an extension to the previously implemented weighted system by integrating the ontology reasoning capabilities in the rule evaluation process. The resulting framework improves the semantic expressivity and semantic interoperability of the interpretable FRBS. The physician or patient can select a varying number of features to describe the new case, and the system recognizes and categorizes these features according to their semantic similarities. These similarities are used in the fuzzy inference process. The rule base and real-time queries are encoded with a standard ontology based on the most globally accepted SCT medical terminology.
7) Implement and evaluate the resulting semantically intelligent FRBS and illustrate its capabilities to address real and complex problems.
8) The resulting system is accurate and interpretable from both structural and semantic viewpoints. Further, it is semantically interoperable with other EHR ecosystem components.

The remainder of the manuscript is organized as follows. Section II describes the architecture of the proposed fuzzy system. Section III describes how we designed each single FRBS and how we implemented the entire framework. Section IV discusses the experimental analysis we performed to validate our proposal. Finally, Section V provides the conclusion and future work.

II. DESIGN ARCHITECTURE OF THE PROPOSED INTERPRETABLE FRBS

The present study aims to develop a semantically intelligent hierarchical FRBS for diabetes diagnosis. The proposed system has two layers, as indicated in Fig. 1. The first

[4] *http://www.uco.es/JFML/home*

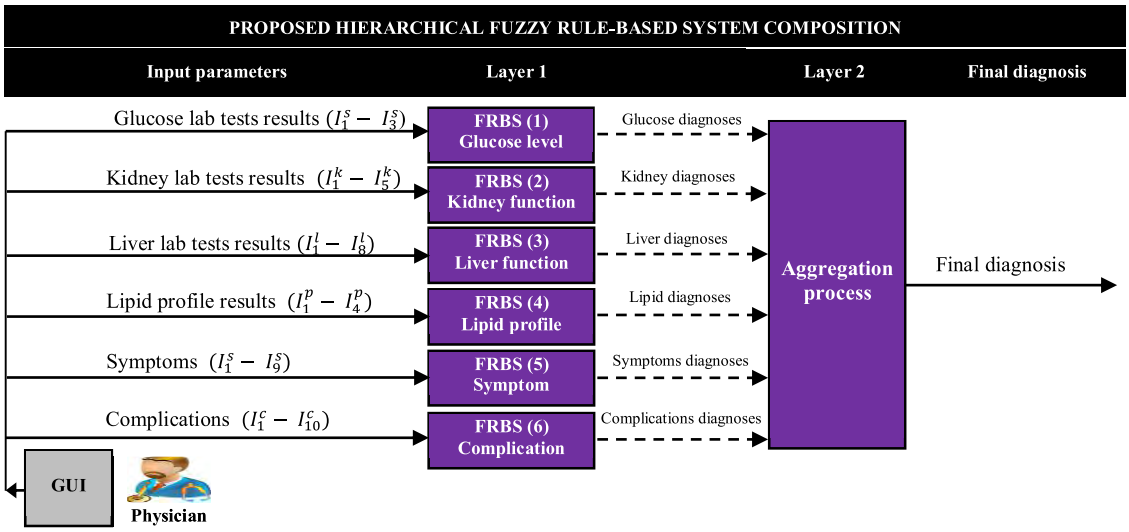

FIGURE 1. Proposed hierarchical FRBS system.

layer has six sub-FRBSs to determine the patient's risk level according to specific dimensions. According to the medical expert opinions, medical literature, and diabetes clinical practice guidelines (CPGs), the data are grouped according to their medical relationships, where dissonant data are not in the same category. FRBS 1 determines the level of risk for developing diabetes according to the patient's glucose level lab tests; it has three ($I_1^s - I_3^s$) numerical input parameters (*HbA1c*, *FPG*, and *2hPG*). FRBS 2 determines the risk level for developing diabetes based on the kidney lab tests; it has five ($I_1^k - I_5^k$) numerical input parameters (*serum potassium*, *serum sodium*, *serum creatinine*, *serum uric acid*, and *serum urea*). FRBS 3 determines the diabetes risk based on liver function tests; it has eight ($I_1^l - I_8^l$) numerical parameters (*albumin*, *total protein*, *γ GT*, *alkaline phosphatase*, *SGPT (ALT)*, *SGOT (AST)*, *direct bilirubin*, and *total bilirubin*). FRBS 4 is based on the lipid profile; it has four ($I_1^p - I_1^p$)numerical input parameters (*LDL cholesterol*, *HDL cholesterol*, *triglycerides*, and *total cholesterol*). FRBS 5 has nine input features; two are numerical (*BMI*, *age*) and seven categorical (*residence*, *gender*, *vision, fatigue, hunger, thirst, and urination frequency*). FRBS 6 has ten categorical features related to diabetes complications (e.g., *nephropathy*, *splenomegaly*, *retinopathy*, *liver cancer*, *viral hepatitis c*). We have 39 features for each patient that can influence the expert decision in the diagnosis process. However, in real situations, at times none of these data are available. Further, certain features are medically more critical than others, e.g., the level of glucose is more critical than the level of lipid, even though the two are important. Consequently, each FRBS in the first layer is assigned a weight $w_j \in (0, 1], j = 1, \ldots, 6$, $\sum_{j=1\ldots6} w_j = 1.0$. To compute the related weights, we first requested the opinion of three domain experts and then applied a well-known multi-criteria decision making (MCDM) technique called the fuzzy AHP technique [38].

The second layer is based on the Mamdani min-max inference mechanism. Each FRBS j in the first layer has a set of fuzzy rules R_i^j, where i is the number of rules and $j = 1, \ldots, 6$ is the number of FRBSs, and each rule is fired with a specific degree $D_{ij} \in [0, 1]$ for a specific input vector. D_{ij} is calculated as the minimum membership degree of all antecedents of rule R_i^j. Each FRBS j has a specific weight $w_j \in (0, 1]$. The firing degree of each rule R_i^j is multiplied by the weight of its corresponding FRBS w_j to produce the weighted degree $WD_{ij} \in [0, 1]$ of the consequent class. If we assume that the output feature has n classes, then for each FRBS j, the winner rule for each class O_k^j, $k = 1, \ldots, n$ is the one with $\max_{\forall i \in O_k^j} WD_{ij}$. For each class O_k^j, its firing degree $FD_{O_k^j}$ between FRBSs is $\max_j \max_{\forall i \in O_k^j} WD_{ij}$, where j is the number of the FRBS. Finally, the system output is $max_{k=1,\ldots,n}(FD_{O_k^j})$.

For example, assume that we have only two FRBSs. FRBS 1 has three rules R_1^1, R_2^1, and R_3^1 and FRBS 2 has three rules R_1^2, R_2^2, and R_1^3. Assume that the medical weight of FRBS 1 is $w_1 = 0.3$ and FRBS 2 is $w_2 = 0.7$. Assume the 2-class diagnosis with either "Diabetic" or "Non-Diabetic." For a specific case, if the firing degree of FRBS 1's rules were $D_{11} = 0.8$ for Diabetic by R_1^1, $D_{21} = 0.1$ for Diabetic by R_2^1, and $D_{31} = 0.1$ for Non-Diabetic by R_3^1, then the weighted firing degrees are $WD_{11} = 0.8 \times 0.3 = 0.24$ for R_1^1, $WD_{21} = 0.1 \times 0.3 = 0.03$ for R_2^1, and $WD_{31} = 0.1 \times 0.3 = 0.03$ for R_3^1. For FRBS 1, the winner rule for class Diabetic (0.24) is R_1^1 and for class Non-Diabetic (0.03) is R_3^1. The same process is performed for FRBS 2 using a different feature set and rule base. If the firing degree of FRBS 2's rules were $D_{12} = 0.3$ for Diabetic by R_1^2, $D_{22} = 0.1$ for Diabetic by R_2^2, and $D_{32} = 0.6$ for Non-Diabetic by R_3^2, then the weighted firing degrees are $WD_{12} = 0.3 \times 0.7 = 0.21$ for R_1^2, $WD_{22} = 0.1 \times 0.7 = 0.07$ for R_2^2, and $WD_{32} = 0.6 \times 0.7 = 0.42$ for R_3^2. For the FRBS 2, the winner rule

for class Diabetic (0.21) is R_1^2 and for class Non-Diabetic (0.42) is R_3^2. Now, the maximum degree for class Diabetic is the max $\{0.24, 0.21\} = 0.24$, and for class Non-Diabetic is max $\{0.03, 0.42\} = 0.42$. Consequently, the final diagnosis is Non-Diabetic with membership degree of 0.42. As each set of rules in an FRBS uses different inputs, there is no possible cooperation among the rules deriving from the different FRBSs. In this manner, if we disable one of the FRBSs, its related rules are not considered in the inference process. For example, if we disabled FRBS 2, then rules R_i^2, $i = 1, 2, 3$ are disabled. This means the winner rule would be R_1^1 (Diabetic with 0.18). The use of weights provides priority to the ''most confident'' FRBS in accordance with the experts. It could arise that a winner rule without weights is not the winner rule in the case of applying weights, as indicated in the previous example.

There are numerous reasons for building a hierarchical FRBS, including (1) the resulting reasoning is more medically intuitive because it mimics human thinking; (2) the relations between the variables can be better studied and the rule base is more interpretable; (3) the final FRBS can make decision disregarding inputs of specific FRBSs, if they are not available, even with a lower confidence; and (4) to overcome the curse of dimensionality inherent to FRBSs; we have 39 input variables and at least two fuzzy sets for each variable; hence the total number of fuzzy rules is 2^{39} for all combinations; after decomposition into six FRBSs. The total number of rules is the sum of all FRBS rules $(2^3+2^5+2^8+2^4+2^9+2^{10})$.

After implementing this hierarchical FRBS, we identified another challenge. It is intuitive that the number of symptoms and complications can be different from patient to patient. The severities of these features also differ. In Fig. 1, we test a specific set of symptoms and complications. This is extremely restricted and can cause physicians to not accept the CDSS because it cannot represent a patient's variations accurately. The CDSS decision would not be individualized or customized according to the patient's profile. More importantly, there are many features required by CDSS to make the decision. If these features can be collected transparently, automatically, yet accurately, from available resources such as distributed EHR and social media, the resulting system would be more applicable and accurate. The semantic similarity of ontology can facilitate addressing this challenge. Patient symptoms and complications are collected from different sources and, using an ontology reasoner, we can determine those related to diabetes. For these reasons, we extend the previous framework in Fig. 1 to the semantically extended FRBS as displayed in Fig. 2. The framework becomes simpler as FRBSs 5 and 6 are merged into one subsystem of FRBS56. For symptoms, the system dynamically collects them all with their severities. Then, it calculates the symptomatic level of the patient. The same procedure is performed for complications. The system users can determine any patient features manually, whereas others can be collected automatically from other sources including EHR and social media.

There are two main types of knowledge in the proposed system. *Declarative knowledge* is formulated as an ontology of diabetes-related complications and symptoms. This knowledge is implemented as an OWL 2 ontology-based SCT standard medical terminology. The resulting ontology supports the semantic interoperability between the resulting

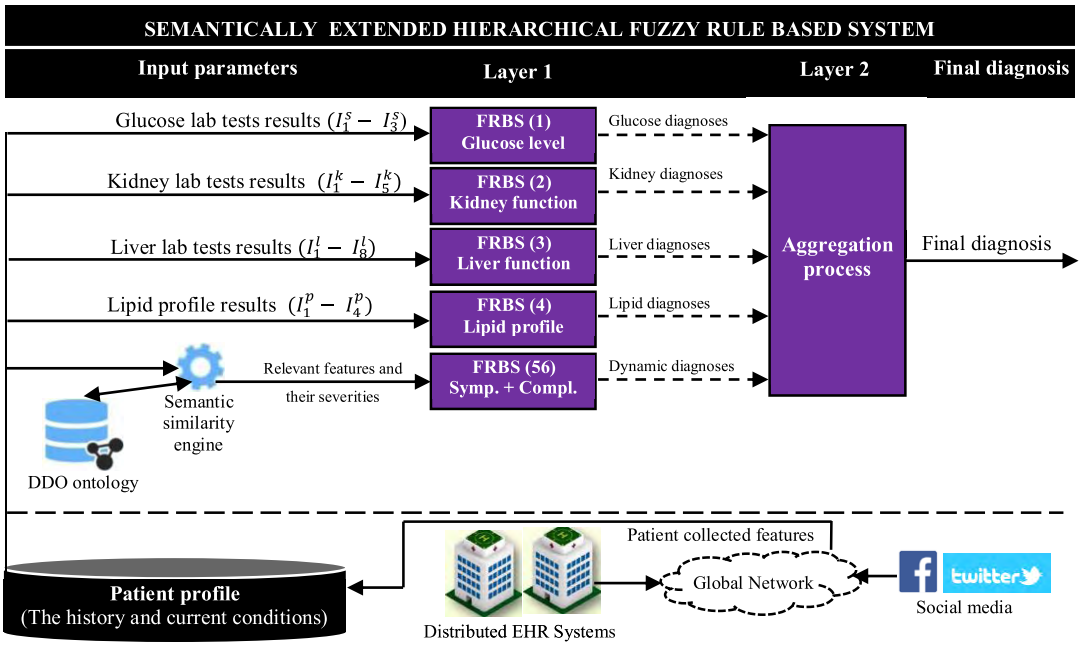

FIGURE 2. Extended semantically intelligent hierarchical FRBS system.

FRBS and EHR environments. The *procedural knowledge* is formulated as fuzzy rules regarding diabetes diagnosis. This knowledge is modeled from a prepared medical dataset extracted from distributed EHR database.

III. DESIGN AND IMPLEMENTATION DETAILS OF BASE FRBSS

First, we discuss the process of creating interpretable and accurate FRBSs from learning data for every subsystem. The resulting system has a fixed number and types of symptoms and complications. The proposed architecture is inspired by the HILK++ fuzzy modeling methodology [39] and d'Acierno et al., methodology [40]. The learning approach is decomposed into several steps: fuzzy partitioning, rule induction, simplification, and implementation. We utilize different algorithms for different tasks. Each of these subsystems is evaluated separately. Next, as indicated in Fig. 1, we build the combined hierarchical FRBS and test it. Then, we extend the resulting model by integrating ontology reasoning in the fuzzy reasoning process based on a selected semantic similarity algorithm, see Fig. 2. The resulting system is a more accurate, intuitive, interoperable, dynamic, interpretable, and semantically intelligent system. Fuzzy partitioning and rule learning are based on the automatic knowledge extracted from the data. The expert knowledge is limited to the determination of input and output variables and the number of acceptable fuzzy sets for the key variables. Fig. 3 proposes a framework for the main steps required to build a semantically interpretable FRBS.

It includes a set of sequential steps, which are discussed in detail in the next sections. In this paper, we consider Mamdani-type fuzzy rule-based classifiers generated with the GUAJE open source software [41]. We selected GUAJE because it implements HILK [39]. GUAJE assists users in the design of fuzzy systems with an acceptable interpretability–accuracy trade-off owing to the combination of expert and induced knowledge.

A. DATA PREPROCESSING

This step improves the quality of the collected medical data by performing a set of preprocessing steps including addressing missing values, identifying outliers, encoding categorical or nominal features, and semantic encoding of medical concepts according to the SCT standard terminology.

B. LINGUISTIC VARIABLES AND FUZZY PARTITIONING

This step determines the list of critical linguistic variables, their ranges, their granularities (number of fuzzy terms and membership functions (MFs)), and shapes and parameters of MFs. These numbers are of prime importance and affect the resulting systems accuracy and interpretability levels [8]. We define certain factors that, according to the physicians, can be used to evaluate the possible risk for developing diabetes. Further, the collected features are evaluated by a set of machine-learning algorithms to determine their significance according to the dataset used. An FRBS is a system that can be modeled as a function of the form displayed in Eq. (1).

$$f : X \rightarrow \Lambda \tag{1}$$

where $X = X_1 \times X_2 \times \ldots \times X_n \in R^n$ is an n-dimensional input space and $\Lambda = \{y_1, y_2, \ldots, y_c\}$ is the set of class labels. The first step in FRBS design is to define a linguistic variable for each input. For n input variables, variable j is defined as in Eq. (2).

$$V_j = (\upsilon_j, X_j, Q_j, S_j, I_j) \tag{2}$$

where υ_j is the name of the variable, X_j is the domain of the variable, $Q_j = \{q_{j1}, q_{j2}, \ldots, q_{jm}\}$ is a set of m linguistic values for the variable, $S_j = \{s_{j1}, s_{j2}, \ldots, s_{jm}\}$ is a set of fuzzy sets on X_j, $s_{jk} : X_j \rightarrow [0, 1]$, and I_j associates each linguistic value q_{jk} to a fuzzy set s_{jk}.

The first requirement in the design of interpretable FRBS is to build interpretable fuzzy partitions, which requires the definition of fuzzy sets satisfying at least the following constraints [42]:

- *Distinguishability*: Semantic integrity requires that all membership functions represent distinct linguistic concepts to be readable.
- *A justifiable number of fuzzy sets.* The number of fuzzy sets is typically maintained between two and nine owing to the limits on human capacity for processing information [43].
- *Uniformity of coverage*: Each data point x should belong significantly, $\mu(x) > \alpha$, at least to one fuzzy set, where α is called the coverage level, and at most to two fuzzy sets, where $\sum_{i=f1}^{f2} \mu_i(x) = 1$.
- *Normalization*: All fuzzy sets should be normal, $\forall i \in [1, f], \exists x \in \cup$ such that $\mu_i(x) = 1$ for f fuzzy sets.
- *Overlapping and orthogonality*: All fuzzy sets should significantly overlap such that $\forall x, \sum_{f=1}^{m} \mu_f(x) = 1$, $m = 2$, where only two adjacent fuzzy sets can overlap.

These requirements are all fulfilled by the strong fuzzy partitions (SFPs) as displayed in Eq. (3) [8]:

$$\begin{cases} \forall x & \sum_{f=1,2,3,\ldots,p} \mu_f(x) = 1 \\ \forall f & \exists x \mu_f(x) = 1 \end{cases} \tag{3}$$

where p is the number of fuzzy sets in the partition and $\mu_f(x)$ is the membership degree of x to the f^{th} fuzzy set. We select the simplest and most computationally efficient MFs; we depend on a number of semi-trapezoidal shapes at the edges and triangular MFs elsewhere to generate strong fuzzy partitions, as asserted by many authors [8], [11]. Trapezoidal and triangular MFs are displayed in Eqs. (4) and (5), respectively.

$$f(x; a, b, c, d) = \begin{cases} 0; & x \leq a \\ \dfrac{x-a}{b-a}; & a \leq x \leq b \\ 1; & b \leq x \leq c \\ \dfrac{d-x}{d-c}; & c \leq x \leq d \\ 0; & d \leq x, \end{cases} \quad \text{or}$$

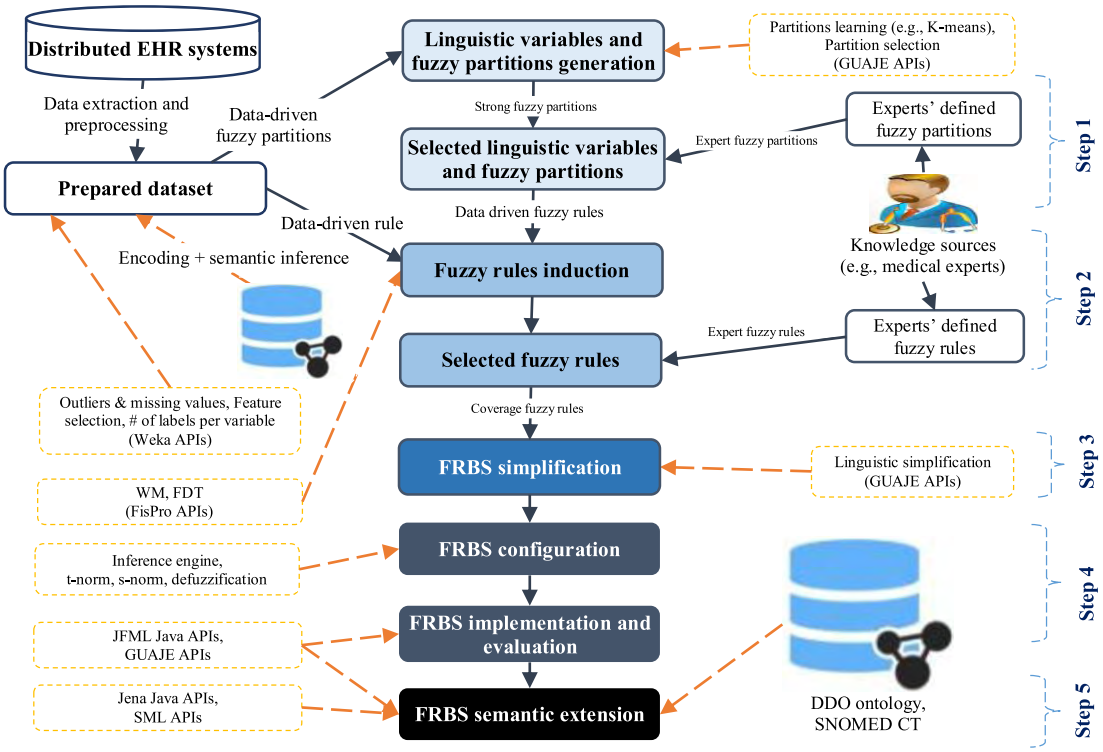

FIGURE 3. Proposed semantically interpretable FRBS.

$$f(x;a,b,c,d) = \max\left(\min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right) \quad (4)$$

$$f(x;a,b,c) = \begin{cases} 0; & x \le a \\ \dfrac{x-a}{b-a}; & a \le x \le b \\ \dfrac{c-x}{c-b}; & b \le x \le c \\ 0; & c \le x, \end{cases} \quad \text{or}$$

$$f(x;a,b,c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \quad (5)$$

The overlap between adjacent MFs is achieved as indicated in Eq. (3). This methodology ensures semantic integrity and leads to a robust system. There are many methodologies that can infer SFPs from data including *clustering* (e.g., k-means) [44] and hierarchical fuzzy partitioning (HFP) [30] methods. Clustering methods first categorize the dataset $(x_1, x_2, , , x_n)$, where each sample x_i is a d-dimensional vector, into a set of $k \le n$ clusters $(s_1, s_2, , , s_k)$, to minimize the within-cluster sum of the squares or variance. Formally, k-means determines the $\arg min_S \sum_{i=1}^{k} \sum_{x \in S_i} | |x - \varphi_i| |^2 = \arg min_S \sum_{i=1}^{k} |S_i| \, VarS_i$, where φ_i is the mean of the samples in S_i. The next step is to represent each cluster as a fuzzy set. The HFP technique is inspired from hierarchical and fuzzy clustering techniques to create univariate fuzzy partitions. The items to be clustered are fuzzy sets, not data items. HFP is performed independently for each input dimension. As a starting step, regular partitions divide the variable domain into uniform intervals according to the selected number of MFs where each fuzzy set kernel is the middle of its corresponding interval. The algorithm produces triangular and semi-trapezoidal shapes. The initial partitioning includes N fuzzy sets, where N is the number of data points (x_k, y_k), $k = 1, 2, \ldots, N$, and x_k is a p-dimensional input vector, $x_k = x_k^1, x_k^2, \ldots, x_k^p$. The k-th fuzzy set (FS) is centered at x_k^j. Each FS is assigned a weight equal to the number of points in its cluster, noted as w^f for FS f as $w^f = \sum_{x \in E} \mu_j^f(x)$. A recursive FS merging is performed according to a specific distance function, which reduces the fuzzy partition size by one at each step.

To evaluate the goodness of the generated partitions, there are several indices. This paper depends on the following indices to compare generated partitions. Let N be the data set size, $\mu_i(k)$ the membership degree of the i-th item in group i, and c the number of terms for the fuzzy partition, as in Eq. (6).

$$\begin{aligned} PC &= \frac{1}{N}\sum_{k=1}^{N}\sum_{i=1}^{c}\mu_i^2(k), \\ PE &= -\frac{1}{N}\left\{\sum_{k=1}^{N}\sum_{i=1}^{c}[\mu_i(k)log_a(\mu_i(k))]\right\}, \\ CI &= \frac{1}{N}\sum_{k=1}^{N} max_i\mu_i(k) - \frac{2}{c(c-1)} \\ &\quad \times \sum_{i=1}^{c-1}\sum_{j=i+1}^{c}\frac{1}{N}\sum_{k=1}^{N}\min\left(\mu_i(k), \mu_j(k)\right) \end{aligned} \quad (6)$$

The partition coefficient (PC) and the partition entropy (PE) were proposed by [45]; the Chen index (CI) was proposed by [46]. The best partition should minimize PE, and maximize PC and CI. The proposed framework selects the best partitioning method, the suitable number of partitions, and the partition parameters according to the results of applying these indices on the considered training datasets.

C. FUZZY RULE INDUCTION

After defining the input fuzzy partitions, we define the system's knowledge base. The naïve method for generating the complete set of rules corresponding to all fuzzy set combinations must be rejected because it would cause a fast combinatorial explosion. We must select the relevant rule premises and assign them the appropriate output. The number of rules should be small, the rule base consistent, and the rules should be general, i.e., they must not systematically include all input variables, rather, only the important ones in the rule context. These rules are frequently called incomplete, more general, rules to reduce system complexity and enhance the level of interpretability. Although the system can result in marginally reduced accuracy, this step is critical in the medical domain because medical experts must always understand the system output. The best rule base (RB) should satisfy the objectives of generalization, robustness, and accuracy.

An RB has a list of R multi-input single output rules; the general form of the r-th fuzzy rule R_r is defined in Eq. (7).

Rule R_r*:*

$$\begin{aligned} &IF \quad \underbrace{v^1 \; is \; A_r^1}_{partial\ premise\ P_1} \quad \mathrm{T} \quad \underbrace{v^2 \; is \; A_r^2}_{partial\ premise\ P_2} \quad \mathrm{T} \ldots \mathrm{T} \quad \underbrace{v^k \; is \; A_r^k}_{partial\ premise\ P_k} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{premise} \\ &THEN \quad \underbrace{y \; is \; y^r}_{conclusion} \quad with\ CF_r \end{aligned} \tag{7}$$

where P_i is a premise for $i = 1, 2, ..., k, A_r^j \in Q_j$, $CF_r \in [0, 1]$ is the certainty factor or weight of R_r, and $y^r \in \Lambda$. We concentrate on multi-input single output classification rules in this study.

For each input vector $V_i = (x_{i1}, x_{i2} \ldots x_{ik})$, the rule activation degree w for each rule R_r is obtained by the conjunction of its premise elements as: $w_r(V_i) = \mathrm{T}_{j=1}^{k} \mu_{A_r^j}(x^j) = \mu_{A_r^1}(x_{i1})\, \mathrm{T} \mu_{A_r^2}(x_{i2})\, \mathrm{T} \ldots \mathrm{T} \mu_{A_r^k}(x_{ik})$, where $\mu_{A_j^s}(x_{is})$ is the membership degree of x_s in fuzzy set A_j^s in rule j for the input sample s, $\mathrm{T} : [0, 1]^2 \rightarrow [0, 1]$ is the t-norm operator, which is selected in the configuration step (e.g., AND operator), and $i = 1, 2, \ldots, N$, and N is the number of cases. A rule r is active if $w_r > 0$. The inference function that computes the output y of the system has the form indicated in Eq. (8).

$$Fuzzyy = F(y) = \bigcup_{r=1}^{R} \left(\left(\bigcap_{i=1}^{k} \mu_r^i (x_i) \right) \bigcap \gamma_r * CF_r \right) \tag{8}$$

where $\cup_r^R$ denotes the t-conorm operator (e.g., OR), $\cap$ is the implication operator, $\cap_i^k$ is the t-norm operator (e.g., AND), $*$ is the product, and γ_r is the output activation of R_r. To obtain a numerical output, (8) is defuzzified with a function such as center of gravity as in Eq. (9).

$$y = \left(\frac{\int_Y yF(y)\, dy}{\int_Y F(y) dy} \right) \tag{9}$$

In this paper, we apply the usual min-max Mamdani inference mechanism. Thus, we consider min t-norm as the conjunction and implication operator, and max t-conorm as the aggregation operator. Moreover, output classes are represented by singletons and the winner rule procedure determines the final decision. There are numerous techniques to extract fuzzy rules from data [39]. Each method has its advantages and disadvantages. To select the most representative and accurate set of rules, we test a set of methods, including fuzzy decision tree (FDT) and the Wang and Mendel (WM) methods [39]. The resulting rule set for each applied method is validated for consistency and evaluated using the well-known classification performance evaluation measures of accuracy, precision, recall, and F-measure. To preserve the interpretability of the resulting system, the weight CF_r of rule R_r is calculated based on a fuzzy AHP.

D. FRBS SIMPLIFICATION

Rules created from a domain expert are frequently simple or incomplete. They do not contain all input variables, (i.e., incomplete or general rules), where certain variables appear in some rules only. The simplification of the included rules maximizes the interpretability such that the accuracy $\geq \Delta_A$. This simplification process creates more compact, interpretable, and meaningful rules (that are closer to those typically defined by experts). Further, as a side effect, simplification can reduce the complexity of the fuzzy partitions and fuzzy variables. The two steps of simplification involve (1) rule base simplification to reduce the number of rules by deleting or merging rules, and simplifying rules premises and (2) database simplification by reducing the number of fuzzy partitions and fuzzy variables. We depend on the techniques proposed in [39] to perform the simplification process.

E. FRBS CONFIGURATION AND IMPLEMENTATION

This step determines the general configurations of the FRBSs. The disjunction operator (OR) is the maximum; the conjunction operator (AND) is the minimum; the inference mechanism is "first infer then aggregate"; the rule engine is Mamdani and the defuzzification method is "mean of max". We use a collection of tools to implement the proposed framework, including GUAJE, for building the FRBSs and Weka[5] for comparison purposes. Further, we utilize two JAVA APIs including SML[6] [40] and JFML.[7] The development environment is Eclipse IDE version Oxygen.2 Release (4.7.2) with JDK 8.

As this system is designed for usage by physicians and patients, who may not be advanced users, we aim to develop

[5] https://www.cs.waikato.ac.nz/ml/weka/

[6] *http://www.semantic-measures-library.org/sml/*

[7] *http://jfuzzylogic.sourceforge.net/html/index.html*

a user friendly and intuitive GUI interface. The system is implemented using Microsoft Access as a database engine and executes in a Microsoft Windows environment. We construct a database where the patient detailed information is stored. Patient ID number is the primary key and his/her other data including demographics, medical values, and diagnosis are stored in the database. Fig. 4 illustrates how the system and its inference are designed.

After the complete query is formulated, it is passed to the FRBS to determine the diagnosis decision. FRBS has three main modules including fuzzification module that translates classical inputs into fuzzy values through linguistic input variables, inference that applies a fuzzy reasoning mechanism to obtain fuzzy outputs based on a knowledge base, and defuzzification that translates fuzzy outputs back to classical values based on the output variables' MFs. Because we use singletons to represent the output classes, defuzzification is not required. We simply select the output with the highest activation.

F. FRBS SEMANTIC EXTENSION

In this phase, the previously designed interpretable FRBS is enhanced with additional semantic capabilities. Features with direct numerical values such as lab tests are modeled as fuzzy linguistic variables. Other features including patient symptoms, complications, and drugs are of categorical values; they cannot be modeled as a set of fixed and specific values because they can be different from one patient to another. These features are modeled using OWL 2 ontologies. There are numerous advantages of this approach. First, it enhances the semantic of the resulting system. For example, if the rule contains the condition $<...current\ complication = "Hypertension"...>$, then we must test if the patient has hypertension or any of its sub-diseases by consulting the referenced ontology. The rule would be shallow if we only had a single feature with name *Hypertension* and values *Yes*/*No* or *Normal*/*Low risk*/*Severe risk*. This design methodology was followed in [12]. The ontology semantic is used for symptoms and drugs as well because diabetes has many symptoms and drugs that can affect the glucose level. Second, the resulting system supports the integration and interoperability with the EHR system and other data sources such as social media because the system can understand not only the concepts used in FRBS but also their semantically-related concepts. Third, the number of antecedents in the resulting rules is dynamic. This is critical to build a real and medically applicable CDSS. Diabetes patients always have different numbers, types, and severities of symptoms, complications, and drugs, which can be distributed in different sources. For example, one patient can input to the system that he/she has

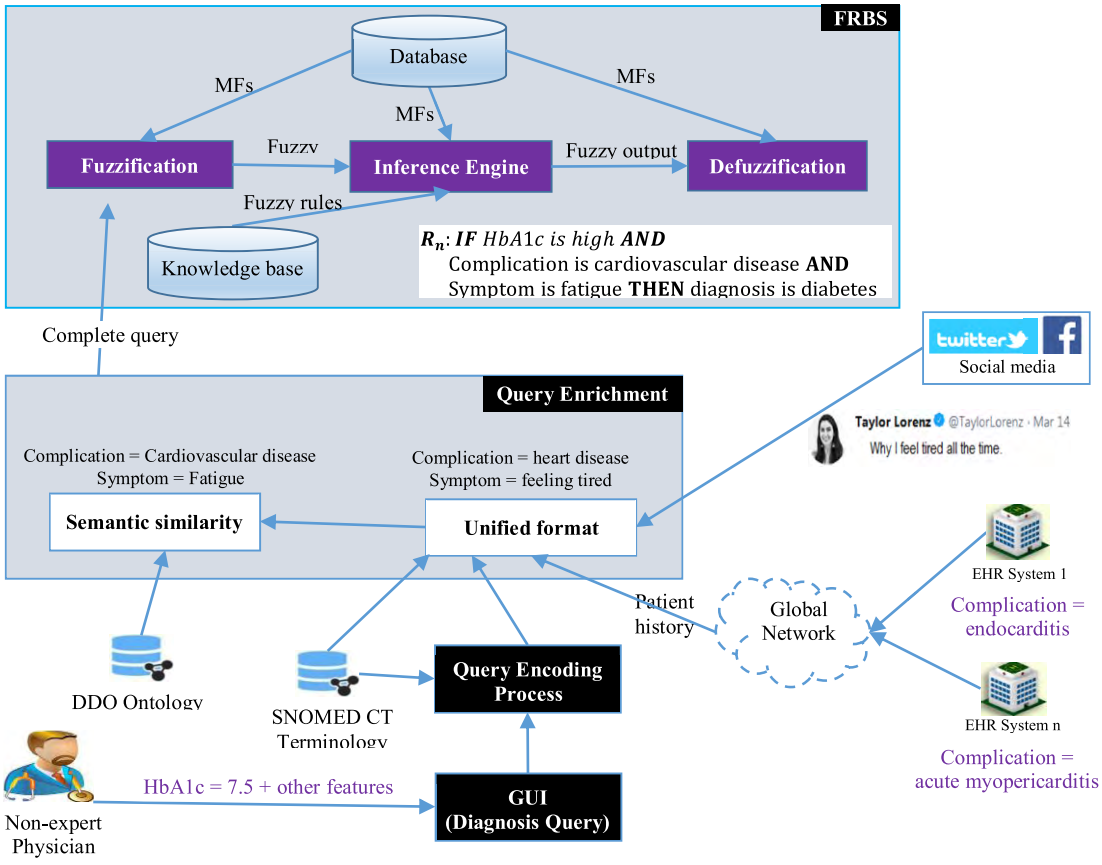

FIGURE 4. Proposed FRBS in hospital context.

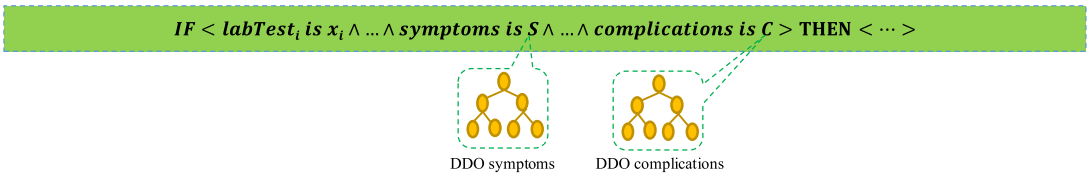

FIGURE 5. Combination of semantic and fuzzy reasoning.

fatigue and *thirst*; from his/her medical record, the system collects that he/she also had *frequent urination* and *blurry vision*; further, he/she tweeted regarding his/her *weight loss* and *numbness in the hands*. Real CDSS must consider all of these dynamic types of symptoms in its decision. Other patients could have other numbers and types of symptoms. The same thought process is used with complications and drugs. Consequently, we cannot model a fuzzy rule as *IF <Symptom 1 is "S_1" AND Symptom 2 is "S_2" AND...AND Symptom m is "S_m"> THEN <...>* because the number and types of symptoms cannot be collected in advance. This also applies to complications and drugs. A new methodology is required for designing fuzzy rules, which first collects all of the patient symptoms, complications, and drugs, and their severities from different sources. This then measures their medical similarities with those of diabetes and calculates the level of symptomatic, complication complexity, and drug complexity of the patient, and finally sends these values to the fuzzy rules to make a decision. The resulting rules are more realistic and more compatible with diabetes CPGs. One CPG model some diagnosis rules as: *IF <FPG≥7.0 mmol/L AND patient is Symptomatic> THEN <Patient is diabetic>, IF< A1C ≥ 5.7% AND patient has complications of "cardiovascular disease, hypertension, or obesity"...AND patient is Asymptomatic> THEN <patient is diabetic>*. The diagnosis rule can now be modeled as *IF <Age is "A" and BMI 2 is "B" AND ...AND Symptomatic is "S"> THEN <Diagnosis is...>*, where the *Symptomatic* feature is calculated dynamically from all the patient's symptoms. The same modeling is performed for complications and drugs. La-Ongsri and Roddick [47] attempted to implement this approach in data modeling by designing certain fields as ontology concepts, and the other fields as regular types. Torshizi *et al.* [13] attempted to combine ontology and fuzzy rule-based reasoning to diagnose and treat benign prostatic hyperplasia. However, the proposed system has two independent modules for these techniques.

Sherimon and Krishnan [48] checked for specific complications when diagnosing diabetes; however, the ontology was crisp and the checked complications had no semantics. Chen *et al.* [31] used ontology and fuzzy logic in anti-diabetic drug recommendation. However, ontology is used as a second and separate layer in the reasoning; it did not enhance the inference of the fuzzy component. This step represents the semantic features in the form of ontology and determines the measures that are used to evaluate the level of clinical similarity between compared medical concepts.

We propose a Diabetes mellitus Diagnosis OWL2 Ontology (DDO [49]) based on SCT standard medical terminology. The ontology is freely available in the BioPortal.[8] Harispe *et al.* [50] surveyed the most applicable semantic similarity measures. To select the most applicable measure, we implement and evaluate some of these measures, and select the measure with the highest accuracy. Semantic similarity is calculated based on DDO ontology using the Pellet reasoner and according to the algorithm presented in Algorithm 1. This defines how similar the meaning of concepts is according to the taxonomical evidences modeled in the ontology.

Fig. 5 illustrates an example of a fuzzy rule with the probable features including *i* lab tests, symptom complexity, and complication complexity. Regarding semantic features, these concepts have two layers of reasoning. In the first layer, they are modeled as DDO concepts. In the second layer, they are modeled as fuzzy features. It should be noted that to implement an interoperable system with the EHR environment, we must address both syntax interoperability using standard data modeling such as openEHR or HL7 RIM, standard CDSS interface such as virtual medical record (*v*MR) and clinical document architecture (CDA), and semantic interoperability using standard terminology such as SCT [51]. In this case, we address only the semantic interoperability issue; the other dimensions will be considered in a future work.

The proposed system has two layers: semantic layer and fuzzy layer. At the semantic layer, the query is coded according to SCT semantic concepts: the EHR database is searched for the additional required data regarding the patient according to the patient ID and the DDO ontology is consulted to determine the level of similarity between the patient profile and the concepts used in the target fuzzy rules. In the second layer, the patient's full profile is sent to the FRBS for evaluation according to its knowledge base.

At run time, when a physician or patient submits a query to the FRBS, the system automatically, according to the patient ID, collects the patient's other symptoms, and complications used from the distributed EHR environment and social media to enrich the query with all required information regarding the patient, see Fig. 4.

For FRBS 1–4 (see Fig. 1), the numerical values are fuzzified regularly according to their MF. For FRBS 5 and 6,

[8] *http://bioportal.bioontology.org/ontologies/DDO*

Algorithm 1 Semantic Similarity Calculation Algorithm

Input: *DDO* ontology, R ontology reasoner, Patient profile $< \ldots, S, C, \ldots >$
/$*C$ = patient's complications, S = patient's symptoms $*$/
Output: SC, CC
/$*SC$ = a number that determines the level of severity of the patient symptoms, CC = a number that determines the level of severity of the patient complications $*$/
Let $S_d = \{\}, C_d = \{\}$;
For $t = 1$ *to* u **do** /$*u$ denotes the number of symptoms in $S*$/
If $R.SIM_{Semantic}\ (DDO, S_t) > 0$ **Then** /$*$ measuring semantic similarity level between symptom S_t and DDO ontology using R and Eq. 12 $*$/
$S_d+ =\prec S_t, severity_t \succ$; //$severity_t \in (0, 1]$
End If;
End For;
/$*$ *according to the type of implementation, we may support only the number of symptoms or the number and severity of symptoms* $*$/
$SC = 0$;
If *No severities* **Then**
For $i = 1$ *to* $S_d.length$ **do**
$SC++$;
End For;
Else
For $i = 1$ *to* $S_d.length$ **do**
$SC+ = S_d\,[i]\,.severity$;
End For;
End If;
Return SC; /$*$ *The SC is the level of complexity of the patient's collection of symptoms* $*$/
For $j = 1$ *tor* **do** /$*r$ denotes the number of complications in $C*$/
If $R.SIM_{Semantic}\ (DDO, C_j) > 0$ **Then**
/$*$ measuring semantic similarity between complication C_j and DDO ontology using R and Eq. 12$*$/
$C_d+ =\prec C_j, severity_j \succ$; //$severity_j \in (0, 1]$
End If;
End For;
$CC = 0$;
If *No severities* **Then**
For $i = 1$ *to* $C_d.length$ **do**
$CC++$;
End For;
Else
For $i = 1$ *to* $C_d.length$ **do**
$CC+ = C_d\,[i]\,.severity$;
End For;
End If;
Return CC; /$*$ *The CC is the level of complexity of the patient's collection of complications* $*$/

first, the two groups of symptoms (S), and complications (C) are entered to the ontology reasoner to determine their semantic similarities with diabetes specific symptoms and complications, respectively. This step only filters the patient's diabetes-related concepts, i.e., $S_d \subset S$ and $C_d \subset C$, for all concepts where semantic similarity is greater than zero. The patient's symptom complexity is $SC = |S_d|$ and complication complexity is $CC = |C_d|$. These numbers are entered to the semantic FRBS for fuzzification.

Further, each feature severity level can be considered. In such a case, the severities of each group are aggregated and entered to the fuzzy rule for fuzzification. For example, the urination frequency is 0.8. For all symptom $S_d^i, i \in |S_d|$ and the $SC = \sum_{i=1}^{|S_d|} severity(S_d^i)$, where $severity(S_d^i) \in (0, 1]$; for all complications C_d^i, $i \in |C_d|$ and the $CC = \sum_{i=1}^{|C_d|} severity(C_d^i)$, where $severity(C_d^i) \in (0, 1]$. The calculated values of SC and CC are sent to FRBS for fuzzification. Other parameters can be used to calculate the values of SC and CC including the level of semantic similarity between the patient concept and DDO concept, and the medical weight (i.e., level of importance) assigned to a feature. For example, a complication such as heart attack must have a higher weight than flu. We consider these enhancements as future works. Calculating semantic similarity is based on our proposed measure [14]. It estimates the clinical similarity between two concepts using Eq. (10).

$$SIM_{Semantic}\ (DDO(u), v) = W_1 SIM_{path}\ (DDO(u), v) + W_2 SIM_{feature}\ (DDO(u), v) \tag{10}$$

where

$$SIM_{path}\ (DDO\ (u)\ , v) = \frac{2 \times depth\ (lca\ (DDO\ (u)\ , v))}{depth\ (DDO\ (u)) + depth\ (v)},$$

and, $SIM_{feature}\ (DDO\ (u)\ , v)$, as shown at the top of the next page, and $W_1, W_2 \in (0, 1]$, $W_1 + W_2 = 1$, $lca(DDO(u), v)$ is the least common ancestor, and $A\ (x) = \{y | x \sqsubseteq y\}$.

The Query Enrichment process collects the entire patient's missing data, which can help making accurate decisions, including lab tests, complications, and symptoms according to the patient medical ID.

These data can be collected from several different resources, including distributed EHRs and social media. The semantic features are encoded according to SCT; then, their semantic similarity to diabetes complications and symptoms are measured. In this paper, we only enrich the query from our local encoded database according to SCT. A fully integrated CDSS requires the addressing of semantic interoperability with distributed EHRs and sentiment analysis to extract features from social media. However, this paper concentrates only on building an accurate and semantically interpretable FRBS. The connection to social networks remains as a future work. After preparing all the single FRBSs for the hierarchical frameworks, they are combined as discussed in Section II to build the full system. In the next section,

$$SIM_{feature}\left(DDO\left(u\right),v\right)=1-log_2\left(1+\frac{\left|A\left(DDO\left(u\right)\right)\backslash A\left(v\right)\right|+\left|A\left(v\right)\backslash A\left(DDO\left(u\right)\right)\right|}{\left|A\left(DDO\left(u\right)\right)\backslash A\left(v\right)\right|+\left|A\left(v\right)\backslash A\left(DDO\left(u\right)\right)\right|+\left|A\left(DDO\left(u\right)\right)\cap A\left(v\right)\right|}\right)$$

we discuss the evaluation process of all subsystems and the generated hierarchical FRBSs.

IV. RESULTS AND DISCUSSION

This section discusses the experimental analysis of the proposed model. We implemented the proposed framework based mainly on the JFML and Jena JAVA APIs. JFML supports the generation of fuzzy systems in XML format based on the IEEE Standard for Fuzzy Markup Language (IEEE Std 1855-2016). This unified format is critical to build distributed CDSSs and we expect it will add great value to the next generation of CDSSs. Jena APIs support the implementation of ontology semantic similarity measure.

A. EXPERIMENTAL SETUP

We concentrate in this study on the diabetes diagnosis process. Our dataset was obtained from the hospitals of Mansoura University, Mansoura, Egypt from January 2010 through August 2013. It addresses the diabetes mellitus diagnosis problem. It included 39 features that can add values in diabetes diagnosis, see Table 1. There are two main types of features, i.e., numerical and categorical. The output variable is a binary categorical variable with two values: Diabetic and Non-Diabetic. The dataset included 60 patients distributed as 53% Diabetic and 47% Non-Diabetic.

B. DESIGNING OF SINGLE FISS

As explained previously, for several reasons, it is not possible to build a single FRBS with 39 input features. These reasons include: rules would be overly long; there would be a large number of rules; and finally, this is not medically intuitive. To overcome this problem, we build a hierarchical system, following the methodology described in Section III. The dataset is divided into medically related feature groups. According to this distribution, we carefully build the proposed two-layer hierarchical FRBS (see Fig. 1). Each of the six FRBSs in the first layer is carefully designed with the aim of achieving an acceptable interpretability–accuracy trade-off. The resulting system is both highly interpretable and acceptably accurate.

The first step in the design process is to determine the strong fuzzy partitions related to all the linguistic features. We evaluated two partitioning methods (k-means and HFP), which were thoroughly compared with regular partitions uniformly distributed in the given input range. Following the suggestion of psychologists, we considered a small odd number of fuzzy sets for each fuzzy partition. Thus, *PRn* represents a partition that is composed of *n* fuzzy sets (with *n* in {3, 5, 7}). Table 2 summarizes the reported results (which are averaged over 10-fold cross-validation) for the first five FRBSs in the proposed hierarchical system. Note that we only learn fuzzy partitions for those numeric variables included in Table 1. In the case of categorical variables, we set up as many singleton membership functions as the number of categories that where identified by the physicians. FRBS6 only contains categorical variables. Accordingly, it is not included in Table 2. The quality of each fuzzy partition was measured in terms of three quality indexes (PC, PE, and CI), which were introduced in Section III.B, see Eq. (6). Given a number of fuzzy sets, the best partition should maximize PC and CI while minimizing PE. Accordingly, we report in the table the

TABLE 1. Data Set Description.

Feature groups for FRBSs	Feature name	Data type	Unit of measurement	Min-mean-max
Glucose lab tests	HbA1C	N	%	5-6.373-7.4
	2h PG	N	mg/dl	165-202.733-235
	FPG	N	mg/dl	96-129.633-156
Kidney Function Lab tests	Serum potassium	N	mEq/L	2.4-3.767-4.3
	Serum urea	N	mg/dL	17-31.56-67
	Serum Uric acid	N	mg/dL	3-4.237-7.9
	Serum creatinine	N	mg/dL	0.9-1.35-3.6
	Serum sodium	N	mEq/L	134-137.833-158
Liver function tests	S. albumin	N	g/dL	1.9-4.082-5.4
	Total bilirubin	N	mg/dL	0.8-1.317-3
	Direct bilirubin	N	mg/dL	0.3-0.533-1.6
	SGOT (AST)	N	U/L	35-54.567-165
	SGPT (ALT)	N	U/L	35-57.317-183
	Alk. phosphatase	N	U/L	170-214.2-360
	γ GT	N	U/L	18-35.833-98
	Total protein	N	g/dL	3.1-4.858-8.7
Lipid profile	LDL cholesterol	N	mg/dL	50-94.917-170
	Total cholesterol	N	mg/dL	158-209.367-275
	Triglycerides	N	mg/dL	78-144.767-189
	HDL cholesterol	N	mg/dL	30-55.533-65
Symptoms	Urination frequency	C	-	{normal, +, ++}
	Vision	C	-	{normal, +, ++}
	Thirst	C	-	{normal, +, ++}
	Hunger	C	-	{normal, +, ++}
	Fatigue	C	-	{normal, +, ++}
	Residence	C	-	{ Urban, Rural}
	Gender	C	-	{Male, Female}
	Age	N	year	29-48-74
	BMI	N	kg/m^2	20-33.117-45
Complications	Ten features for patient's current and historical complications	C	-	Collection of diseases
Diagnosis	Diabetes diagnosis	C	-	{Diabetic, Non-Diabetic}

TABLE 2. Fuzzy partitions design for FRBSS 1–5.

			Uniform				Kmeans				HFP			
			PC	PE	CI	#	PC	PE	CI	#	PC	PE	CI	#
FRBS 1	2hPG	PR3	0.633686	0.538616	0.633649	0	0.767037	0.352409	0.771967	10	0.693592	0.456702	0.692	0
		PR5	0.633797	0.537964	0.689382	0	0.708744	0.442139	0.767144	2	0.727905	0.408885	0.772118	8
		PR7	0.663184	0.505053	0.74508	0	0.752155	0.38764	0.830116	9	0.730964	0.403521	0.788361	1
	FPG	RP3	0.671853	0.495892	0.68	0	0.734507	0.399832	0.747804	10	0.69984	0.449506	0.702413	0
		RP5	0.647408	0.519012	0.701777	0	0.779451	0.3546	0.84071	10	0.708065	0.439327	0.758288	0
		RP7	0.633334	0.540658	0.699678	0	0.763187	0.380646	0.840352	10	0.718779	0.419128	0.770962	0
	HbA1c	RP3	0.698611	0.45254	0.707407	0	0.777194	0.346197	0.795532	9	0.766193	0.359819	0.777483	1
		RP5	0.672222	0.479404	0.712776	0	0.700454	0.450806	0.751773	2	0.726354	0.401891	0.763068	8
		RP7	0.704166	0.431348	0.75556	0	0.694519	0.464291	0.767927	2	0.723803	0.40821	0.774252	8
FRBS 2	Serum Potassium	RP3	0.780056	0.330958	0.789474	0	0.825492	0.268911	0.833142	10	0.802447	0.293278	0.796146	0
		RP5	0.751802	0.360055	0.780001	0	0.843131	0.250508	0.881846	10	0.769123	0.334037	0.796261	0
		RP7	0.76085	0.350062	0.802423	0	0.856036	0.227601	0.899853	10	0.783339	0.318918	0.829253	0
	Serum Sodium	RP3	0.838887	0.263912	0.862964	0	0.907421	0.150972	0.917878	10	0.741507	0.377124	0.733325	0
		RP5	0.766667	0.349227	0.816667	0	0.896418	0.168599	0.926809	10	0.732784	0.386893	0.782014	0
		RP7	0.749999	0.360732	0.78175	0	0.856048	0.220437	0.883271	10	0.737123	0.378196	0.79054	0
	Serum Creatinine	RP3	0.812986	0.303621	0.838683	0	0.918175	0.136577	0.929506	10	0.767965	0.341425	0.728079	0
		RP5	7298.863	0.397019	0.775925	0	0.86552	0.203617	0.896923	10	0.768282	0.352268	0.823673	0
		RP7	0.716873	0.422549	0.792417	0	0.817596	0.293002	0.877878	6	0.802808	0.294562	0.850754	4
	Serum Uric Acid	RP3	0.739081	0.414214	0.779594	0	0.850311	0.247497	0.873878	10	0.680414	0.468662	0.6544	0
		RP5	0.617548	0.552813	0.664388	0	0.83439	0.26308	0.878839	10	0.717129	0.415001	0.75932	0
		RP7	0.635403	0.532815	0.696046	0	0.873758	0.188489	0.907786	10	0.743563	0.378723	0.781575	0
	Serum Urea	RP3	0.72992	0.402688	0.720889	0	0.779653	0.332725	0.795843	6	0.773803	0.34239	0.774235	4
		RP5	0.757013	0.376078	0.815934	0	0.882085	0.204145	0.917959	10	0.722731	0.410349	0.761184	0
		RP7	0.70528	0.435357	0.76254	0	0.860426	0.240438	0.911177	10	7633.369	0.377836	0.797943	0
FRBS 3	Serum Albumin	RP 3	0.568273	0.610258	0.497777	0	0.865504	0.234997	0.892787	10	0.811774	0.289312	0.827026	0
		RP 5	0.779754	0.34824	0.819283	0	0.815457	0.318682	0.878536	10	0.731275	0.379969	0.738182	0
		RP 7	0.605877	0.568514	0.66975	0	0.838611	0.282346	0.900317	10	0.747044	0.362739	0.770963	0
	Total Protein	RP 3	0.676489	0.465127	0.659524	0	0.769364	0.341476	0.771815	4	0.772197	0.352089	0.788615	6
		RP 5	0.727381	0.409674	0.783931	0	0.768424	0.362704	0.83225	10	0.715939	0.416179	0.759817	0
		RP 7	0.786012	0.335336	0.844728	0	0.883029	0.191512	0.91679	10	0.70982	0.422072	0.761133	0
	Gamma GT	RP 3	0.691041	0.475169	0.723333	0	0.921017	0.144061	0.937797	10	0.716542	0.416759	0.709165	0
		RP 5	0.594166	0.584176	0.625998	0	0.7942	0.300099	0.817228	10	0.750523	0.373495	0.786109	0
		RP 7	0.642707	0.52801	0.718014	0	0.772746	0.3569	0.832756	6	0.773711	0.327489	0.810795	4
	Alkaline Phosphatase	RP 3	0.765946	0.365846	0.786666	0	0.848952	0.232952	0.859809	10	0.711348	0.422542	0.695467	0
		RP 5	0.703787	0.437417	0.766879	0	0.859307	0.226797	0.902883	10	0.72017	0.408424	0.746503	0
		RP 7	0.665448	0.477628	0.701152	0	0.846793	0.242943	0.897746	10	0.714087	0.413843	0.755776	0
	SGPT (ALT)	RP 3	0.837089	0.274151	0.865164	0	0.904993	0.16387	0.919235	10	0.80216	0.298255	0.797323	0
		RP 5	0.752861	0.378783	0.813694	0	0.908556	0.162387	0.937219	10	0.734609	0.397083	0.778659	0
		RP 7	0.715785	0.420851	0.780801	0	0.862108	0.222475	0.907411	10	0.789055	0.307744	0.822629	0
	SGOT (AST)	RP 3	0.853507	0.25919	0.883076	0	0.926735	0.131051	0.937497	10	0.751439	0.359665	0.740662	0
		RP 5	0.764797	0.376665	0.835001	0	0.937025	0.118235	0.955349	10	0.757881	0.352349	0.795599	0
		RP 7	0.730793	0.413156	0.814652	0	0.914688	0.15877	0.949532	10	0.769655	0.336092	0.81546	0
	Direct Bilirubin	RP 3	0.669334	0.470529	0.644445	0	0.843651	0.274776	0.88045	0	1	0	1	10
		RP 5	0.744	0.389977	0.809332	0	0.843651	0.274776	0.88045	0	1	0	1	10
		RP 7	0.744	0.389977	0.818416	0	0.843651	0.274776	0.88045	0	1	0	1	10
	Total Bilirubin	RP 3	0.693112	0.472574	0.727273	0	0.892206	0.180797	0.908275	10	0.680729	0.471532	0.674028	0
		RP 5	0.590634	0.587091	0.648332	0	0.797371	0.334129	0.859959	10	0.751079	0.378826	0.793951	0
		RP 7	0.631956	0.536477	0.680953	0	0.700267	0.447479	0.779506	0	0.754973	0.367536	0.81669	10
FRBS 4	LDL Cholesterol	RP 3	0.725417	0.414044	0.717407	0	0.825123	0.263544	0.834321	10	0.791065	0.329436	0.793212	0
		RP 5	0.749443	0.385138	0.801389	0	0.775945	0.344685	0.819018	9	0.732451	0.404009	0.781238	1
		RP 7	0.728751	0.409483	0.789607	0	0.806264	0.310481	0.866128	10	0.731022	0.402924	0.787027	0
	HDL Cholesterol	RP 3	0.812243	0.275144	0.812698	0	0.89052	0.172744	0.903037	10	0.862115	0.214702	0.8737	0
		RP 5	0.810883	0.274122	0.829763	0	0.846529	0.229248	0.868098	9	0.826325	0.256302	0.854242	1
		RP 7	0.830203	0.252701	0.871291	0	0.853496	0.22291	0.891093	10	0.831042	0.247113	0.86139	0
	Triglycerides	RP 3	0.670146	0.493425	0.673072	0	0.779369	0.341836	0.798923	10	0.728325	0.421682	0.74267	0
		RP 5	0.661364	0.506723	0.730287	0	1.679641	0.649316	1.765015	10	1.63268	0.701924	1.715452	0
		RP 7	0.670953	0.485201	0.734792	0	0.725188	0.426748	0.808188	6	0.72239	0.420606	0.783476	4
	Total Cholesterol	RP 3	0.607316	0.574791	0.617475	0	0.868839	0.223415	0.878964	10	0.670307	0.476428	0.664276	0
		RP 5	0.576838	0.602734	0.609046	0	0.854961	0.245931	0.892654	9	0.653011	0.492021	0.659862	1
		RP 7	0.713697	0.449246	0.797192	0	0.869964	0.221991	0.908998	10	0.643171	0.507908	0.671406	0
FRBS 5	BMI	RP 3	0.64224	0.525253	0.628443	0	0.767719	0.350524	0.772049	10	0.686703	0.464674	0.680639	0
		RP 5	0.683627	0.476526	0.738934	0	0.693001	0.463045	0.750156	4	0.698313	0.446866	0.743643	6
		RP 7	0.686827	0.473669	0.760445	0	0.683913	0.476414	0.758019	0	0.71306	0.425644	0.771034	10
	Age	RP 3	0.664923	0.496342	0.652838	0	0.780496	0.334645	0.791305	10	0.6895	0.461027	0.684378	0
		RP 5	0.70117	0.455151	0.761259	0	0.708545	0.440326	0.771763	5	0.708567	0.435458	0.758076	5
		RP 7	0.664296	0.498186	0.740424	0	0.702682	0.450812	0.779689	3	0.7049	0.437735	0.758158	7

number of folds (#) for which the related partition (Uniform, Kmeans, and HFP) becomes the best in the light of the values of the quality indexes.

As can be observed in the table, typically Kmeans was the best partitioning method. Further, in some cases, HFP was selected. Uniform partitions was never selected. Notice that there are numerous papers where uniform partitions are used when no expert partitions are available. Our experimental results demonstrate how learning fuzzy partitions from data with Kmeans and HFP can yield improved results. Hence, we design fuzzy partitions that fit the data distribution to guarantee an acceptable accuracy. Further, the partitioning approach respects all interpretability constraints that are typically demanded to build interpretable fuzzy partitions. It for this reason that all the generated partitions are strong fuzzy partitions (see Eq. 3). Table 3 lists the final strong MFs that were designed for the involved features. We utilized trapezoidal functions at the two ends and a triangular function in the middle. Once all the partitions were designed, it was time to generate the related fuzzy rule bases.

For each FRBS, and for the same folds considered in the partitioning stage, we applied two rule learning methods (WM and FDT) with the best partitions ($RP3, RP5, RP7$) previously selected. Then, we applied a linguistic simplification procedure aimed at producing a rule base as compact as possible while preserving the initial accuracy. This simplification stage requires merging and/or removing rules. Further, it can yield to remove those linguistic terms that were initially defined but not used in any rule. In Table 4, we report four quality indexes (Accuracy, Precision, Recall, and F-measure) for training and testing (averaged over 10-fold cross-validation), and two basic interpretability indexes (Rules and Premises). The first column in the table identifies the FRBS; the second column is related to the rule base design procedure (it includes rule learning and simplification).

For each of the first five systems, we generated 12 rule bases. Their names follow the pattern $RPn - SP - RLM(-S)$

TABLE 3. Sample of linguistic variables and fuzzy sets used for diabetes diagnosis.

Category	Linguistic variable	Linguistic fuzzy set	Shape	Parameters
FRBS 1 (Glucose level lab test)	2hPG	Low	Trapezoidal	[165, 165, 185, 195.195]
		Average	Triangular	[185, 195.195, 219.509]
		High	Trapezoidal	[195.195, 219.509, 235, 235]
	FPG	Low	Trapezoidal	[96, 96,129.75, 138.312]
		High	Trapezoidal	[129.75, 138.312, 156, 156]
	HbA1C	Low	Trapezoidal	[5.5, 5.991, 6.4]
		Average	Triangular	[5.991, 6.4, 6.609]
		High	Trapezoidal	[6.4, 6.609, 7.4,7.4]
FRBS 2 (Kidney lab tests)	Serum Potassium	Low	Trapezoidal	[2.4, 2.4, 2.629, 3.455]
		Average	Triangular	[2.629, 3.455, 4.197]
		High	Trapezoidal	[3.455, 4.197, 4.3, 4.3]
	Serum Sodium	Low	Trapezoidal	[134, 134, 135.02, 146]
		Average	Triangular	[135.02, 146, 158]
		High	Trapezoidal	[146, 158, 158, 158]
	Serum Creatinine	Low	Trapezoidal	[0.9, 0.9, 1.015, 2.56]
		Average	Triangular	[1.015, 2.56, 3.35]
		High	Trapezoidal	[2.56, 3.35, 3.6, 3.6]
	Serum Uric Acid	Low	Trapezoidal	[3, 3, 3.347, 5.196]
		Average	Triangular	[3.347, 5.196, 7.85]
		High	Trapezoidal	[5.196, 7.85, 7.9, 7.9]
	Serum Urea	Low	Trapezoidal	[17, 17, 22.243, 40.533]
		Average	Triangular	[22.243, 40.533, 58.5]
		High	Trapezoidal	[40.533, 58.5, 67, 67]
FRBS 3 (Liver lab tests)	S. Albumin	Low	Trapezoidal	[1.9, 1.9, 2.358, 4.424]
		Average	Triangular	[2.358, 4.424, 5.133]
		High	Trapezoidal	[4.424, 5.133, 5.4, 5.4]
	Total Protein	Low	Trapezoidal	[3.1, 3.1, 3.118, 4.433]
		Average	Triangular	[3.118, 4.433, 8.7]
		High	Trapezoidal	[4.433, 8.7, 8.7, 8.7]
	Gamma GT	Low	Trapezoidal	[18, 18, 25.604, 70.375]
		Average	Triangular	[25.604, 70.375, 89.5]
		High	Trapezoidal	[70.375, 89.5, 98, 98]
	Alkaline Phosphatase	Low	Trapezoidal	[170, 170, 183.683, 265.688]
		Average	Triangular	[183.683, 265.688, 356.667]
		High	Trapezoidal	[265.688, 356.667, 360, 360]
	SGPT(ALT)	Low	Trapezoidal	[35, 35, 40.521, 102.75]
		Average	Triangular	[40.521, 102.75, 168]
		High	Trapezoidal	[102.75, 168, 183, 183]
	SGOT(AST)	Low	Trapezoidal	[35, 35, 39.146, 96.375]
		Average	Triangular	[39.146, 96.375, 156]
		High	Trapezoidal	[96.375, 156, 165, 165]
	Direct Bilirubin	Low	Trapezoidal	[0.0, 0.0, 0.3, 0.4]
		Average	Triangular	[0.3, 0.4, 0.5]
		High	Trapezoidal	[0.4, 0.5, 1, 1]
	Total Bilirubin	Low	Trapezoidal	[0.8, 0.8, 1.044, 2.071]
		Average	Triangular	[1.044, 2.071, 2.88]
		High	Trapezoidal	[2.071, 2.88, 3, 3]
FRBS 4 (Lipid profile)	LDL Cholesterol	Low	Trapezoidal	[50, 50, 58.333, 86.9]
		Average	Triangular	[58.333, 86.9, 159.812]
		High	Trapezoidal	[86.9, 159.812, 170, 170]
	HDL Cholesterol	Low	Trapezoidal	[30, 30, 34.571, 53]
		Average	Triangular	[34.571, 53, 64.081]
		High	Trapezoidal	[53, 64.081, 65, 65]
	Triglycerides	Low	Trapezoidal	[78, 78, 103.857, 147.219]
		Average	Triangular	[103.857, 147.219, 180.071]
		High	Trapezoidal	[147.219, 180.071, 189, 189]
	Total Cholesterol	Low	Trapezoidal	[158, 158, 173, 200.711]
		Average	Triangular	[173, 200.711, 250.917]
		High	Trapezoidal	[200.711, 250.917, 275, 275]
FRBS 5 (Symptoms)	Vision	{Non, Blurred-vision, Allergy-redness}	Singleton	{1, 2, 3}
	Fatigue, Hunger, Thirst, Urination Frequency	{Normal, High, Very High}	Singleton	{1, 2, 3}
	Residence	{Urban, Rural}	Singleton	{1, 2}
	Gender	{Female, Male}	Singleton	{1, 2}
	BMI	Very Low	Trapezoidal	[20, 20, 24.585, 28.941]
		Low	Triangular	[24.585, 28.941, 32.275]
		Average	Triangular	[28.941, 32.275, 37.813]
		High	Triangular	[32.275, 37.813, 42.071]
		Very High	Trapezoidal	[37.813, 42.071, 45, 45]
	Age	Very Low	Trapezoidal	[29, 29, 29, 35.266]
		Low	Triangular	[29, 35.266, 40.862]
		Average	Triangular	[35.266, 40.862, 51.738]
		High	Triangular	[40.862, 51.738, 74]
		Very High	Trapezoidal	[51.738, 74, 74, 74]
FRBS 6 (Complications)	Nephropathy, Shrunken Kidney, Splenomegaly, Retinopathy, Hypercholesterolemia, Ovarian Cancer, Viral Hepatitis C, Liver Cancer, Bleeding Gum, Fatty Liver	{False, True}	Singleton	{0, 1}
Diagnosis	Diabetes Diagnosis	{Normal, Diabetic}	Singleton	{1, 2}

where $RPn - SP$ (with n in $\{3, 5, 7\}$) identifies the fuzzy partitions previously selected, RLM is WM or FDT; we compare results with $(-S)$ without running the simplification procedure. Notice that in the case of FRBS 6, because there was no partition learning, only four rule bases are generated in accordance with the given options (two different rule learning methods with and without simplification). The best rule base for each FRBS, i.e., the one with the highest Accuracy on testing, is highlighted in bold font and shadowed background. For FRBS 1, the best results were provided by $RP5 - SP - FDT - S$ (98.33% of classification rate regarding testing and 96.83% with respect to training data; in addition to a small number of rules, 3.2, and premises, 4.4). Accordingly, we first selected the best fuzzy partitions for each input variable, we generated rules with FDT, and finally we performed the simplification procedure. In this case, the accuracy increased after simplification regarding both training and test, and the interpretability clearly improved because the number of rules and premises was dramatically reduced. For FRBS 2, we selected$RP3 - SP - WM$, which produced higher testing accuracy (48.35%). In this case, the remainder of the generated rule bases were affected by an overfitting effect. For FRBS 3, $RP3 - SP - WM - S$ was the best rule base (56.66% of testing accuracy). For FRBS 4, $RP3 - SP - FDT - S$ was the selected rule base. For FRBS 5, $RP5 - SP - FDT$ reported the highest training accuracy (65.01%). Finally, $FDT - S$ produced the best interpretability–accuracy trade-off (46.67% on testing, 68.15% on training, 11.2 rules)

TABLE 4. Fuzzy rule base design.

		TRAINING (%)				TESTING (%)				INTERPRETABILITY	
		Accuracy	Precision	Recall	F-measure	Accuracy	Precision	Recall	F-measure	Rules	Premises
FRBS 1	**RP3-SP-WM**	86.46	88.56	86.46	86.39	84.98	87.78	84.98	84.78	8	24
	RP3-SP-WM-S	86.46	88.56	86.46	86.39	84.98	87.78	84.98	84.78	2	2.1
	RP3-SP-FDT	87.93	89.47	87.93	87.9	83.31	86.31	83.31	82.84	9.5	22.9
	RP3-SP-FDT-S	88.3	89.74	88.3	88.27	81.64	85.06	81.64	81.13	3.7	5.4
	RP5-SP-WM	94.61	95.49	94.61	94.58	93.32	96.25	93.32	93.87	14.5	43.5
	RP5-SP-WM-S	94.98	95.81	94.98	94.95	94.98	95.81	94.98	94.95	2.5	4.9
	RP5-SP-FDT	94.24	94.69	94.24	94.22	91.66	92.5	91.66	91.44	10.3	19.8
	RP5-SP-FDT-S	**96.83**	**97.17**	**96.83**	**96.85**	**98.33**	**98.75**	**98.33**	**98.29**	**3.2**	**4.4**
	RP7-SP-WM	99.26	99.35	99.26	99.25	91.67	97.64	91.67	93.4	22.2	66.6
	RP7-SP-WM-S	99.26	99.35	99.26	99.25	96.66	98.75	96.66	97.29	2.7	4.2
	RP7-SP-FDT	99.81	99.82	99.81	99.82	96.66	97.5	96.66	96.58	5	10
	RP7-SP-FDT-S	99.81	99.82	99.81	99.82	96.66	97.5	96.66	96.58	2.9	3.7
FRBS 2	**RP3-SP-WM**	**60.57**	**62.29**	**60.57**	**58.53**	**48.35**	**47.01**	**48.35**	**46.64**	**13.8**	**69**
	RP3-SP-WM-S	60.75	64.64	60.75	57.53	45.01	44.93	45.01	43.68	6.6	27.4
	RP3-SP-FDT	66.31	66.84	66.31	65.82	38.34	38.67	38.34	39.5	23.7	86.4
	RP3-SP-FDT-S	66.5	67.23	66.5	65.84	38.34	38.39	38.34	39.78	8.2	25.6
	RP5-SP-WM	66.86	67.63	66.86	66.15	43.34	49.08	43.34	43.65	21.7	108.5
	RP5-SP-WM-S	66.86	67.56	66.86	65.79	38.33	40.29	38.33	36.64	11.7	45.4
	RP5-SP-FDT	73.52	74.02	73.52	73.48	35.01	35.75	35.01	38.98	31.8	99.8
	RP5-SP-FDT-S	73.52	73.9	73.52	73.47	35.01	34.42	35.01	39.4	13.6	36.9
	RP7-SP-WM	66.85	67.5	66.85	66.41	38.34	48.53	38.34	40.18	22.1	110.5
	RP7-SP-WM-S	66.85	67.75	66.85	65.91	36.67	37.42	36.67	35.49	12.9	42
	RP7-SP-FDT	74.08	75.26	74.08	73.98	33.35	35.76	33.35	40.27	32.2	91.1
	RP7-SP-FDT-S	74.08	74.54	74.08	74	35.02	34.5	35.02	40.18	13.1	29.2
FRBS 3	**RP3-SP-WM**	69.08	72.57	69.08	67.49	45	53.6	45	51.22	20.9	167.2
	RP3-SP-WM-S	**69.08**	**72.8**	**69.08**	**67.43**	**56.66**	**60.49**	**56.66**	**59.08**	**6.1**	**20.1**
	RP3-SP-FDT	72.04	75.26	72.04	71.54	55	57.08	55	58.02	26.6	97.5
	RP3-SP-FDT-S	72.22	76.38	72.22	71.55	53.33	56.24	53.33	57.93	6.4	16
	RP5-SP-WM	70.55	72.34	70.55	69.51	36.68	40.31	36.68	41.18	26.8	214.4
	RP5-SP-WM-S	70.55	73.78	70.55	69.02	53.33	44.5	53.33	48.38	9.5	37.4
	RP5-SP-FDT	76.12	76.86	76.12	76.11	40	40.46	40	42.91	32	117.5
	RP5-SP-FDT-S	76.12	76.62	76.12	76.09	41.67	40.16	41.67	41.74	13.3	42.4
	RP7-SP-WM	69.63	72.36	69.63	68.14	23.32	27.16	23.32	31.64	26.9	215.2
	RP7-SP-WM-S	69.82	74.45	69.82	67.58	38.34	42.49	38.34	42.86	13	51.1
	RP7-SP-FDT	74.44	76.18	74.44	74.3	43.33	46.63	43.33	45.38	41.6	136.9
	RP7-SP-FDT-S	74.63	77.03	74.63	74.27	50	53.06	50	51.36	11.9	29
FRBS 4	**RP3-SP-WM**	66.12	75.09	66.12	61.6	53.34	39.94	53.34	43.09	12.9	51.6
	RP3-SP-WM-S	66.12	75.09	66.12	61.6	58.34	51.69	58.34	50.97	5	13.9
	RP3-SP-FDT	71.47	73.59	71.47	70.24	65	58.94	65	59.26	18.6	65.2
	RP3-SP-FDT-S	**71.47**	**73.59**	**71.47**	**70.24**	**65**	**58.94**	**65**	**59.26**	**5.9**	**16.2**
	RP5-SP-WM	68.72	71.38	68.72	68.23	46.66	48.33	46.66	51.11	27.7	110.8
	RP5-SP-WM-S	68.72	71.59	68.72	68.18	43.33	46.42	43.33	50.27	13.3	52.4
	RP5-SP-FDT	74.08	77.18	74.08	72.78	50	45.5	50	45.26	42.9	139.1
	RP5-SP-FDT-S	74.45	77.5	74.45	73.22	48.33	45.58	48.33	47.45	11.2	31.4
	RP7-SP-WM	77.22	77.95	77.22	76.99	31.67	38.8	31.67	37.6	35.2	140.8
	RP7-SP-WM-S	77.22	78.21	77.22	76.9	35.01	41.05	35.01	40.82	17.1	67.4
	RP7-SP-FDT	80	81.94	80	79.48	50	55.91	50	48.36	50	150.3
	RP7-SP-FDT-S	80	81.92	80	79.49	48.32	54.58	48.32	47.92	16.1	45.1
FRBS 5	**RP3-SP-WM**	99.81	99.82	99.81	99.82	0	0	0	0	53.9	485.1
	RP3-SP-WM-S	99.81	99.82	99.81	99.82	30	63.34	30	44.11	20	155.2
	RP3-SP-FDT	99.81	99.82	99.81	99.82	56.67	58.21	56.67	56.26	44.2	197.1
	RP3-SP-FDT-S	99.81	99.82	99.81	99.82	41.67	52.08	41.67	47.84	22.8	96.6
	RP5-SP-WM	99.05	99.1	99.05	99.07	0	0	0	0	53.5	481.5
	RP5-SP-WM-S	99.05	99.1	99.05	99.07	38.33	59.83	38.33	47.94	20.8	166
	RP5-SP-FDT	**99.05**	**99.1**	**99.05**	**99.07**	**65.01**	**72.32**	**65.01**	**65.14**	**52.4**	**206.8**
	RP5-SP-FDT-S	99.24	99.28	99.24	99.27	56.67	72.02	56.67	60.22	21.2	77.2
	RP7-SP-WM	100	100	100	100	0	0	0	0	54	486
	RP7-SP-WM-S	100	100	100	100	20.01	33.05	20.01	24.23	24.3	196.9
	RP7-SP-FDT	100	100	100	100	63.33	73.93	63.33	64.15	56.9	188.9
	RP7-SP-FDT-S	99.81	100	99.81	99.91	58.34	67.59	58.34	60.56	21.9	66
FRBS 6	**WM**	65.95	69.89	65.95	64.25	39.99	49.44	39.99	49.58	15.1	151
	WM-S	65.95	69.89	65.95	64.25	46.67	48.89	46.67	49.94	7.9	47.7
	FDT	68.15	73.62	68.15	67.08	46.67	45.67	46.67	52.03	12.8	71.3
	FDT-S	**68.15**	**73.62**	**68.15**	**67.08**	**46.67**	**45.67**	**46.67**	**52.03**	**11.2**	**59.4**

for FRBS 6. Considering these reported results, it can be appreciated how simplification always preserves accuracy on training, and in some cases, improves accuracy in testing. However, the generalization effect of the simplification procedure depends on each rule base and the complexity of the data input space. WM typically produces a large number of rather specific rules, whereas FDT produces more compact and general rules. The initial number of generated rules is greater for partitions with a larger number of fuzzy sets. Moreover, in some cases, we observed how increasing the number of fuzzy sets and, thus the number of rules, yielded extremely high accuracy on training at the cost of overfitting on testing. Table 5 provides samples of generated fuzzy rules for each FRBS. These rules have been simplified to preserve the system interpretability.

The results are calculated based on four well-known performance evaluation metrics: accuracy, precision, recall, and F-measure. The testing results of the six FRBSs are displayed in Table 6. The evaluation is based on a 10-fold cross-validation technique. For comparison with other approaches, we applied the same dataset to a set of seven well-known machine-learning techniques including naïve Bayes, support vector machine (SVM), logistic regression, k-nearest neighbor (KNN) with k = 3, decision tree based on C4.5, artificial neural network (ANN), and random forest. From these results, it is clear that in the majority of cases, the proposed FRBS subsystems produced superior performance compared to the other ML techniques. Further, the proposed systems are highly preferable for medical experts because they provide interpretable systems with results that can be easily explained. These base systems are combined to build an entire system, which can increase the confidence of the resulting decisions, and hence improve the level of expert acceptance.

C. DESIGN OF THE ENTIRE SYSTEM

Making a decision based on only one of the previous FRBSs is not sufficient and, even though it could provide high performance from a machine-learning point of view, it would not be medically acceptable. This is because diabetes is a

TABLE 5. Example of fuzzy rules generated for every FRBS.

Glucose level FRBS fuzzy rules (Number of rules is 4 rules)
IF HbA1c is Low **THEN** Diagnosis is Normal [*Weight=0.2253*]
IF 2hPG is Low **AND** HbA1C is Average **THEN** Diagnosis is Normal [*Weight=0.2253*]
IF 2hPG is Average **AND** FPG is Low **AND** HbA1C is Average **THEN** Diagnosis is Normal [*Weight=0.2253*]
IF HbA1c is High **THEN** Diagnosis is Diabetic [*Weight=0.2253*]
Kidney lab tests FRBS fuzzy rules (Number of rules is 14 rules)
IF Serum Sodium is Low **AND** Serum Creatinine is Average **AND** Serum Uric Acid is High **AND** Serum Urea is High **THEN** Diagnosis is Normal [*Weight= 0.2119*]
IF Serum Sodium is Average **AND** Serum Creatinine is Average **AND** Serum Uric Acid is Average **AND** Serum Urea is Average **THEN** Diagnosis is Diabetic [*Weight= 0.2119*]
IF Serum Sodium is Average **AND** Serum Creatinine is Average **AND** Serum Uric Acid is High **AND** Serum Urea is High **THEN** Diagnosis is Diabetic [*Weight= 0.2119*]
IF Serum Potassium is Average **AND** Serum Sodium is High **AND** Serum Creatinine is High **AND** Serum Uric Acid is Average **AND** Serum Urea is High **THEN** Diagnosis is Diabetic [*Weight= 0.2119*]
...
Liver lab tests FRBS fuzzy rules (Number of rules is 7 rules)
IF Total Protein is Low **AND** Gamma GT is Low **AND** Alkaline Phosphatase is Low **AND** SGPT(ALT) is Low **AND** SGOT(AST) is Low **AND** Direct Bilirubin is NOT(Low) **AND** Total Bilirubin is Low **THEN** Diagnosis is Diabetic [*Weight= 0.1472*]
IF Alkaline Phosphatase is NOT(High) **AND** Direct Bilirubin is High **THEN** Diagnosis is Diabetic [*Weight= 0.1472*]
IF Serum Albumin is Low **AND** Total Protein is High **AND** Gamma GT is Average **AND** Alkaline Phosphatase is High **AND** SGPT(ALT) is Average **AND** SGOT(AST) is Average **AND** Direct Bilirubin is High **AND** Total Bilirubin is NOT(Low) **THEN** Diagnosis is Normal [*Weight= 0.1472*]
...
Lipid profile FRBS fuzzy rules (Number of rules is 6 rules)
IF Total Cholesterol is Low **THEN** Diagnosis is Normal [*Weight=0.0972*]
IF LDL Cholesterol is Low **AND** HDL Cholesterol is High **AND** Triglycerides is NOT(Low) **AND** Total Cholesterol is Average **THEN** Diagnosis is Diabetic [*Weight=0.0972*]
...
Symptoms FRBS fuzzy rules (Number of rules is 53 rules)
IF Vision is Allergy-redness **AND** BMI is Very High **THEN** Diagnosis is Diabetic [Weight= 0.1703]
IF Urination Frequency is Normal **AND** Residence is Urban **AND** BMI is Low **AND** Gender is Female **AND** Age is very High **THEN** Diagnosis is Normal [*Weight= 0.1703*]
...
Complications FRBS fuzzy rules (Number of rules is 12 rules)
IF Ovarian Cancer is False **AND** Liver Cancer is False **AND** Bleeding Gum is True **AND** Fatty Liver is False **THEN** Diagnosis is Diabetic [*Weight= 0.1481*]
IF Nephropathy is False **AND** Retinopathy is False **AND** Ovarian Cancer is False **AND** Liver Cancer is False **AND** Bleeding Gum is True **AND** Fatty Liver is False **THEN** Diagnosis is Normal [*Weight= 0.1481*]
...

complicated disease and its diagnosis process must consider many parameters to be applicable. We combined the proposed FRBS using JAVA APIs and produced the hierarchical FRBS (H-FRBS). The system has only two layers as discussed previously (see Fig. 1). In an actual hospital environment, although all features are required, certain features are more critical than others. For example, diabetes symptoms are more important than lipid profile. Moreover, in some cases, not all features are known for a given patient, and physicians must make a decision with the available information. To consider these requirements, we extended the previous H-FRBS to weighted H-FRBS (WH-FRBS). We utilized the MCDM technique of fuzzy AHP to determine the medical weights of each subsystem.

Table 7 presents the pairwise comparison of the six subsystems after conversion of the given evaluations to fuzzy triangular numbers for the three domain experts. We calculated the weight of each subsystem according to each domain expert. To be more accurate, we considered the average of the three weights. The resulting weights are glucose level = 0.2253, kidney function = 0.2119, liver function = 0.1472, lipid profile = 0.0972, symptoms = 0.1703, and complications = 0.1481. Blood glucose level has the highest impact in diabetes diagnosis; lipid profile has the least impact. Fig. 6 displays the Java interface that medical experts use to provide the values representing symptoms, complications, and laboratory investigations of a particular patient during the diagnosis process. This interface is connected to the EHR database and WH-FRBS. The database stores all patient medical data and medical diagnosis outcomes.

We implemented the entire system including access control and authentication; adding, updating, and deleting patients; searching for a specific patient; and searching for a specific diagnosis. The system was based on a Microsoft Access backend database. After ''clicking'' the diagnose button (see Fig. 6), the patient diagnosis was calculated. If the physician desired to ignore a specific subsystem, he/she could select the corresponding ''Not Available'' check box and the final decision would be affected by this absent information. The confidence level is calculated based on the medical weight of the absent subsystem. The confidence level is $1 - \sum_{i \in Absent\ FRBS} w_i$. For example, in Fig. 6, the system determined that the patient was Diabetic with a confidence level of 80%. Using weights assigned to each FRBS, we can isolate any number of subsystems, and the system can continue to a make decision (see Table 8). The resulting decisions depend on the weights of the absent FRBSs. These results could provide higher accuracies; however, it would be with a reduced confidence level because of the missing some data. In Table 8, we display a list of 25 combinations concentrating on the effect of the absence of glucose tests, which have the highest weight. The system indicated an accuracy of 90% and confidence of 100% using all subsystems (C1). Because glucose level tests are critical in the diagnosis process, the system demonstrated an accuracy of 71.6% (C2) and confidence of 77.47%. The absence of patient symptoms has the second negative effect after glucose, even though it has less weight than kidney (C5). This is medically intuitive because diabetes symptoms are frequently considered with glucose tests. Consequently, the absence of both glucose and symptoms has a further reduced accuracy of 65% (C11). The absence of glucose, symptoms, and complications demonstrated the poorest results (C25); this also is medically intuitive. For comparison purposes, we also measured the performance of random forest and ANN using the entire dataset and after isolating glucose features (please see bottom of Table 8). We must mention that the system exhibited acceptable accuracy compared to other ML techniques while preserving the interpretability level by the simplification process achieved in the base subsystems' design.

The resulting WH-FRBS is medically more intuitive. It can make decisions with different combinations of data sources and with different confidence levels. However, the resulting system only functions well with numerical features. Categorical features with medical semantics such as complications and symptoms are dynamic in nature. Different patients can have different types and numbers of these features. These features have been implemented in a static form in the WH-FRBS, i.e., five categorical symptoms (vision, fatigue, hunger, thirst, and urination frequency) and ten complications (nephropathy, shrunken kidney, splenomegaly, retinopathy, hypercholesterolemia, ovarian cancer, viral hepatitis C, liver cancer, bleeding gum, and fatty liver). WH-FRBS has no solution if a patient has features other than these values that are also related to diabetes. Further, it cannot understand the semantic relations between these concepts and

TABLE 6. Results of FRBS subsystems compared with machine-learning techniques.

Dataset	Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
Glucose level dataset	Naïve Bayes	96.67	96.7	96.7	96.7
	SVM	96.67	96.7	96.7	96.7
	Logistic regression	95	95.5	95	95
	KNN (k=3)	95	95.1	95	95
	Decision tree (C4.5)	100	100	100	100
	ANN	91.67	92.1	91.7	91.7
	Random forest	98.33	98.4	98.3	98.3
	FRBS 1: RP5-SP-FDT-S (3.2 rules)	**98.33**	**98.75**	**98.33**	**98.29**
Kidney function dataset	Naïve Bayes	38.3	31	38.3	32.9
	SVM	45	26.2	45	33.1
	Logistic regression	41.67	38.5	41.7	38.8
	KNN (k=3)	43.3	43.1	43.3	39.1
	Decision tree (C4.5)	51.67	28	51.7	36.3
	ANN	43.33	41.7	43.3	41.8
	Random forest	35	35.2	35	34.6
	FRBS 2: RP3-SP-WM (13.8 rules)	**48.35**	**47.01**	**48.35**	**46.64**
Liver function dataset	Naïve Bayes	53.3	58.9	53.3	48.9
	SVM	50	46.4	50	43.6
	Logistic regression	46.67	46.7	46.7	46.7
	KNN (k=3)	48.3	49.1	48.3	48
	Decision tree (C4.5)	48.3	47	48.3	46.6
	ANN	51.6	52.5	51.7	51.4
	Random forest	46.67	471	46.7	46.7
	FRBS 3: RP3-SP-WM-S (6.1 rules)	**56.66**	**60.49**	**56.66**	**59.08**
Lipid profile	Naïve Bayes	51.67	53.6	51.7	50.1
	SVM	51.67	48.2	51.7	43.1
	Logistic regression	53.3	53	53.3	52.9
	KNN (k=3)	46.67	46.1	46.7	46.2
	Decision tree (C4.5)	58.3	61.4	58.3	52.4
	ANN	50	48.7	50	47.9
	Random forest	55	54.5	55	54
	FRBS 4: RP3-SP-FDT-S (5.9 rules)	**65**	**58.94**	**65**	**59.26**
Symptoms	Naïve Bayes	56.67	56.7	56.7	56.7
	SVM	51.67	51.4	51.7	51.4
	Logistic regression	61.67	61.5	61.7	61.4
	KNN (k=3)	45	44.9	45	44.9
	Decision tree (C4.5)	60	60	60	60
	ANN	60	60.5	60	60
	Random forest	60	59.9	60	59.9
	FRBS 5: RP5-SP-FDT (52.4)	**65.01**	**72.32**	**65.01**	**65.14**
Complications	Naïve Bayes	46.67	47.3	46.7	46.5
	SVM	41.67	38.5	41.7	38.8
	Logistic regression	45	45.6	45	44.7
	KNN (k=3)	41.67	39.4	41.7	35.4
	Decision tree (C4.5)	38.3	35.7	38.3	36.2
	ANN	48.3	491	48.3	48
	Random forest	48.3	49.3	48.3	47.7
	FRBS 6: FDT-S (11.2 rules)	**46.67**	**45.67**	**46.67**	**52.03**

TABLE 7. Pairwise comparisons of elements for criteria (FRBSS) based on three domain experts.

		Glucose lab test	Kidney lab test	Liver lab test	Lipid profile	Symptoms	Complications
Glucose lab test	Domain expert 1	(1,1,1)	(2, 5/2,3)	(1,3/2,2)	(3/2,2,5/2)	(3/2,2,5/2)	(3/2,2,5/2)
	Domain expert 2	(1,1,1)	(1/2,1,3/2)	(1,3/2,2)	(1/2,1,3/2)	(2,5/2,3)	(3/2,2,5/2)
	Domain expert 3	(1,1,1)	(5/2,3,7/2)	(5/2,3,7/2)	(5/2,3,7/2)	(3/2,2,5/2)	(3/2,2,5/2)
Kidney lab test	Domain expert 1	(1/3,2/5,1/2)	(1,1,1)	(2,5/2,3)	(3/2,2,5/2)	(3/2,2,5/2)	(1,3/2,2)
	Domain expert 2	(2/3,1,2)	(1,1,1)	(1,3/2,2)	(1,3/2,2)	(3/2,2,5/2)	(1/2,1,3/2)
	Domain expert 3	(2/7,1/3,2/5)	(1,1,1)	(3/2,2,5/2)	(1/2,1,3/2)	(3/2,2,5/2)	(1,3/2,2)
Liver lab test	Domain expert 1	(1/2,2/3,1)	(1/3,2/5,1/2)	(1,1,1)	(1,3/2,2)	(1,1,1)	(1/2,1,3/2)
	Domain expert 2	(1/2,2/3,1)	(1/2,2/3,1)	(1,1,1)	(3/2,2,5/2)	(1,3/2,2)	(1,3/2,2)
	Domain expert 3	(2/7,1/3,2/5)	(2/5,1/2,2/3)	(1,1,1)	(3/2,2,5/2)	(2,5/2,3)	(1,3/2,2)
Lipid profile	Domain expert 1	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1/2,2/3,1)	(1,1,1)	(1/2,1,3/2)	(1/2,1,3/2)
	Domain expert 2	(2/3,1,2)	(1/2,2/3,1)	(2/5,1/2,2/3)	(1,1,1)	(1,3/2,2)	(1,3/2,2)
	Domain expert 3	(2/7,1/3,2/5)	(2/3,1,2)	(2/5,1/2,2/3)	(1,1,1)	(3/2,2,5/2)	(1/2,1,3/2)
Symptoms	Domain expert 1	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1,1,1)	(2/3,1,2)	(1,1,1)	(1/2,1,3/2)
	Domain expert 2	(1/3,2/5,1/2)	(2/5,1/2,2/3)	(1/2,2/3,1)	(1/2,2/3,1)	(1,1,1)	(3/2,2,5/2)
	Domain expert 3	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1/3,2/5,1/2)	(2/5,1/2,2/3)	(1,1,1)	(1/2,1,3/2)
Complications	Domain expert 1	(2/5,1/2,2/3)	(1/2,2/3,1)	(2/3,1,2)	(2/3,1,2)	(2/3,1,2)	(1,1,1)
	Domain expert 2	(2/5,1/2,2/3)	(2/3,1,2)	(1/2,2/3,1)	(1/2,2/3,1)	(2/5,1/2,2/3)	(1,1,1)
	Domain expert 3	(2/5,1/2,2/3)	(1/2,2/3,1)	(1/2,2/3,1)	(2/3,1,2)	(2/3,1,2)	(1,1,1)

others. For example, WH-FRBS has no ability to discover that *Steatohepatitis* is a *Fatty Lever* disease, or *Diabetic Glomerulonephritis* is a *Nephropathy* disease. Fortunately, this type of knowledge can be represented in the form of ontologies, and the inference of these relations is based on the description logic formalism. Ontology reasoners such as Pellet or Hermit can be used to discover these types of relationships.

In this project, we utilized a light version of our OWL 2 ontology called DDO; we also used the Pellet reasoner in the Jena API to implement our semantic similarity measure. We extended the proposed WH-FRBS with the capabilities of ontology reasoning and produced the semantically intelligent WH-FRBS (SWH-FRBS).

D. EVALUATION OF THE SEMANTIC EXTENSION

SWH-FRBS is based on only five subsystems in the first layer and the same technique of aggregation in the second layer. As indicated in Fig. 2, FRBS56 is the combination of FRBS 5 and FRBS 6 in Fig. 1. This subsystem has weight = 0.3184, and FRBSs 1–4 have their previous weights. The main difference of SWH-FRBS is in FRBS56, where the patient can suffer from any number and type of symptoms and complications. FRBS56 is based on six features (i.e., residence, BMI, gender, number of complications (NoCs), number of symptoms (NoSs), and age). We generated fuzzy partitions for the new FRBS following the same procedure previously explained for the other of FRBSs. Table 9 extends the results that were reported in Table 2. Kmeans and HFP

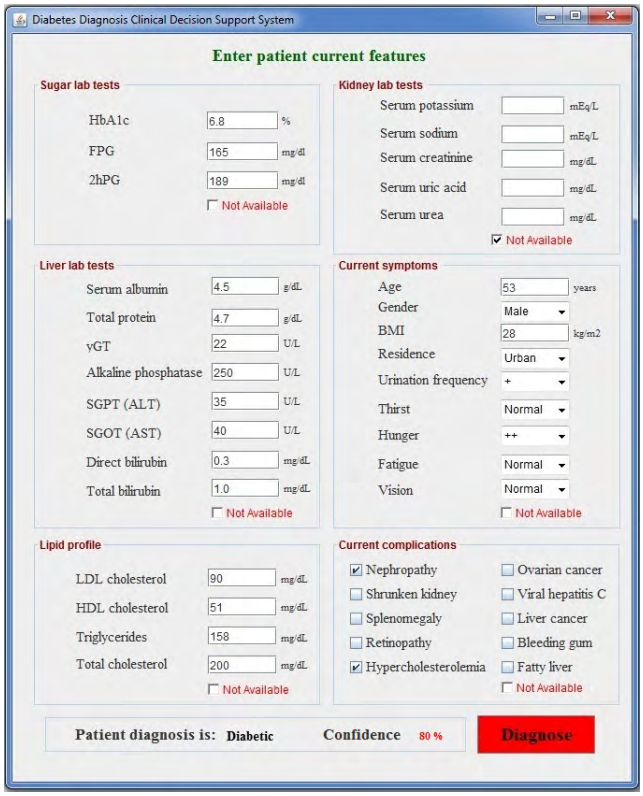

FIGURE 6. Query interface form for combined FRBS.

produce superior strong fuzzy partitions (with respect to the quality indexes PC, PE, and CI) compared to only considering uniform partitions. In this case, HFP outperformed Kmeans for the majority of the generated partitions.

The process for fuzzy rule base induction and simplification was performed for FRBS56, as displayed in Table 10. The system's best rule base (RP3-SP-FDT-S) achieved testing accuracy of 61.67% with approximately 16 rules. The FRBS56 learned with RP3-SP-FDT-S over the entire dataset as follows. The NoCs were modeled as three fuzzy sets: *Low* was trapezoid (0, 0, 0, 2), *Average* was triangular (0, 2, 4), and *High* was trapezoid (2, 4, 4, 4). The NoSs were also modeled as three fuzzy sets: *Low* was trapezoid (0, 0, 0.625, 2.5), *Average* was triangular (0.625, 2.5, 4.222), and *High* was trapezoid (2.5, 4.222, 5, 5). This FRBS56 had a rule base of 17 fuzzy rules. The following are examples of these rules:

Rule 1: *IF* BMI is High *AND* Number of Complications is High *THEN* Diagnosis is Diabetic *[Weight = 0.3184]*

Rule 2: *IF* Residence is Urban *AND* BMI is High *AND* NoCs is Average *AND* NoSs is High *AND* Age is Low *THEN* Diagnosis is Diabetic *[Weight = 0. 3184]*

Rule 3: *IF* Residence is Rural *AND* BMI is NOT (High) *AND* Gender is Male *AND* NoCs is Low *THEN* Diagnosis is Normal *[Weight = 0. 3184]*

The performance of FRBS56 is presented in Table 11 compared with other ML techniques. These are testing results based on 10-fold cross-validation. From the obtained results, it is clear that this FRBS achieved a comparable performance with other ML techniques. Further, it can provide interpretation for its decisions because it is designed considering the interpretability–accuracy trade-off. The most critical contribution of this subsystem is the ability to understand the semantic relationship between patient characteristics and diabetes medical concepts. These features constitute a major advantage of the proposed CDSS for DM diagnosis.

In this version of the project, we considered the number of symptoms and complications. This is because our dataset only supports these types of data, and this approach has been used to diagnose diabetes in other studies [52]. Fig. 7 displays the physician interface to enter the patient characteristics. Regarding symptoms, physicians can enter, in the text box, any types of symptoms for each patient. Further, he/she can enter the severity of each symptom. Clicking the Add Symptom button adds this symptom to the list. In addition to entering the symptom manually, physicians can click the

TABLE 8. WH-FRBS performance after isolating specific subsystems.

# of utilized subsystems or ML	Combinations	Type of subsystems	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)
6	C1	With All subsystems	90	88.24	93.75	90.91
5	C2	Without glucose	71.6	67.44	90.63	77.33
	C3	Without kidney	90	96.43	84.38	89.99
	C4	Without liver	90	88.24	93.75	90.9
	C5	Without symptoms	86.67	85.29	90.63	87.88
	C6	Without lipid	90	88.24	93.75	90.9
	C7	Without complications	91.67	86.49	100	92.75
4	C8	Without glucose and kidney	78.33	88	68.75	77.19
	C9	Without glucose and liver	71.67	67.44	90.63	77.33
	C10	Without glucose and lipid	71.67	67.44	90.63	77.33
	C11	Without glucose and symptoms	65	63.41	81.25	71.23
	C12	Without glucose and complications	65	60.78	96.87	74.69
	C13	Without symptoms and lipid	86.66	85.29	90.62	87.87
	C14	Without symptoms and liver	86.66	85.29	90.62	87.87
	C15	Without symptoms and kidney	93.33	100	87.5	93.33
3	C16	Without glucose and kidney and liver	78.33	88	68.75	77.19
	C17	Without glucose and kidney and lipid	78.33	88	68.75	77.19
	C18	Without glucose and kidney and symptoms	71.67	80	62.5	70.17
	C19	Without glucose and kidney and complications	88.33	100	78.12	87.72
	C20	Without glucose and liver and lipid	71.67	67.44	90.63	77.33
	C21	Without glucose and liver and symptoms	65	63.41	81.25	71.23
	C22	Without glucose and liver and complication	66.67	61.54	100	76.19
	C23	Without glucose and lipid and symptoms	65	63.41	81.25	71.23
	C24	Without glucose and lipid and complication	65	60.78	96.87	74.69
	C25	Without glucose and symptoms and complication	55	54.54	93.75	68.96
Random Forest with all data			96.66	96.7	96.7	96.7
Random forest without glucose data			61.67	61.6	61.7	61.6
ANN with all data			85	85	85	85
ANN without glucose data			65	65.1	65	65

TABLE 9. Fuzzy partitioning of FRBS56.

			Uniform				Kmeans				HFP			
			PC	PE	CI	#	PC	PE	CI	#	PC	PE	CI	#
FRBS56	BMI	RP 3	0.64224	0.525253	0.628443	0	0.767719	0.350524	0.772049	10	0.686703	0.464674	0.680639	0
		RP 5	0.683627	0.476526	0.738934	0	0.693001	0.463045	0.750156	2	0.698313	0.446866	0.743643	8
		RP 7	0.686827	0.473669	0.760445	0	0.683913	0.476414	0.758019	0	0.71306	0.425644	0.771034	10
	Sum Complications	RP 3	0.816665	0.254155	0.755556	0	0.789067	0.315779	0.798015	0	0.7707	0.335222	0.777949	10
		RP 5	1	0	1	0	1	0	1	0	1	0	1	10
		RP 7	0.816665	0.254155	0.80794	0	1	0	1	0	1	0	1	10
	Sum Symptoms	RP 3	0.666667	0.496684	0.679999	0	0.694042	0.465319	0.727534	10	0.669114	0.490512	0.665227	0
		RP 5	0.626666	0.539834	0.681001	0	0.748589	0.360222	0.792705	0	0.771994	0.32713	0.812331	10
		RP 7	0.626666	0.539834	0.696186	0	1	0	1	0	1	0	1	10
	Age	RP 3	0.664923	0.496342	0.652838	0	0.780496	0.334645	0.791305	10	0.6895	0.461027	0.684378	0
		RP 5	0.70117	0.455151	0.761259	0	0.708545	0.440326	0.771763	5	0.708567	0.435458	0.758076	5
		RP 7	0.664296	0.498186	0.740424	0	0.702682	0.450812	0.779689	2	0.7049	0.437735	0.758158	8

TABLE 10. Fuzzy rule induction and simplification process (10-fold cross-validation).

	TRAINING				TESTING				INTERPRETABILITY	
	Accuracy	Precision	Recall	F-measure	Accuracy	Precision	Recall	F-measure	Rules	Premises
RP3-SP-WM	87.39	89.45	89.06	89.1	45.01	61.53	45.01	54.94	44	264
RP3-SP-WM-S	87.39	89.63	89.24	89.29	58.33	64.8	58.33	60.62	15.2	80.8
RP3-SP-FDT	87.59	91.14	90.92	90.91	61.67	71.07	66.67	67.24	67.6	332.5
RP3-SP-FDT-S	**87.95**	**91.43**	**91.11**	**91.12**	**61.67**	**70.68**	**66.66**	**65.19**	**16.3**	**66.9**
RP5-SP-WM	96.85	97	96.85	96.85	6.67	18.34	6.67	9.61	51.5	309
RP5-SP-WM-S	96.85	97	96.85	96.85	41.66	59.21	41.66	49.1	21.2	115.7
RP5-SP-FDT	96.66	96.79	96.66	96.68	51.66	58.82	51.66	52.73	78.4	293.6
RP5-SP-FDT-S	96.66	96.71	96.66	96.67	48.33	58.36	48.33	52.35	23.3	78.9
RP7-SP-WM	99.24	99.28	99.24	99.25	3.34	15	3.34	10	53.5	321
RP7-SP-WM-S	99.24	99.28	99.24	99.25	15	30.84	15	21.84	34.4	192.8
RP7-SP-FDT	98.67	98.74	98.67	98.73	61.67	63.74	61.67	62.17	96.8	319.6
RP7-SP-FDT-S	98.86	98.92	98.86	98.91	53.34	61.23	53.34	55.38	23.8	69

TABLE 11. Results of FRBS56 subsystem compared with machine-learning techniques.

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
Naïve Bayes	58.33	58.4	58.3	58.4
SVM	61.67	63.1	61.7	59.1
Logistic regression	58.33	58.2	58.3	58.3
KNN (k=3)	58.33	58.2	58.3	58.3
Decision tree (C4.5)	50	49.3	50	49.1
ANN	58.33	58.4	58.3	58.4
Random forest	61.33	64.1	63.3	61.8
FRBS 56: RP3-SP-FDT-S (16.3 rules)	**61.67**	**70.68**	**66.66**	**65.19**

Select Symptom button and an ontology is displayed to select the required symptoms, see Fig. 7. The same process is performed for complications. Hence, a physician can select different features for different patients, which supports making customized and personalized decisions.

A critical property of SWH-FRBS is the Query Enrichment. As indicated in Fig. 4, this property supports the auto completion of queries. It enables the system to search other sources (distributed EHR or social media) for any data related to the current patient. Query enrichment can collect other features from the created patient database. The Diagnose button performs two functions. First, it determines the semantic similarity between the queries' collected symptoms and complications and the diabetes-related concepts. After determining the number of patient symptoms and complications, it executes the fuzzy inference process of SWH-FRBS. Based on the GUI proposed in Fig. 7, we offer two case studies that illustrate the comparison between semantic FRBS56 and regular FRBS 5 and 6 in the previous design.

Case Study 1: A male patient; living in Rural; aged 47; has BMI of 23; has symptoms of (*blurred vision*, *conjunctivitis*, *lethargy*, *headache*, *sore throat*, and *polyuria*); and has complications of *atrial fibrillation* and *megacalycosis*. This patient has a real diagnosis of Diabetic. The regular FRBSs 5 and 6 provide a diagnosis as Non-Diabetic. However, the semantic FRBS56 indicates Diabetic as the output class.

Case Study 2: A female patient; living in Urban; aged 53; has BMI of 40; has symptoms of (*blurred vision*, *emphysema*, *polyphagia*, and *polyuria*); and has complications of (*hyperuricemia*, *shrunken kidney*, *keratoconjunctivitis*, and *hyperalphalipoproteinemia*). The domain real diagnosis of this patient is Diabetic, and the semantic FRBS56 diagnoses it as Diabetic. However, the regular FRBSs 5 and 6 provide a diagnosis as Non-Diabetic.

The main reason for this is that the regular FRBSs can only search for specific features with specific values. For example, in Case study 1, it cannot understand that *lethargy* means *fatigue* and *megacalycosis* is a *nephropathy*. The semantic FRBS56 is more flexible, owing to the ontology reasoning. It can recognize different numbers (i.e., 5 and 4 symptoms, 1 and 4 complications) and different types of symptoms and complications. In the first case study, the system detects that *conjunctivitis* and *sore throat* symptoms and *atrial fibrillation* complication are not related to diabetes, hence it does not consider them in the diagnosis process. As can be observed, the semantic system is more dynamic and more intelligent than the previous "regular" one. The problems discussed in the previous case studies can be generalized to all cases in our dataset. If the physician selected other semantic representations for symptoms and complications, the regular FRBSs 5 and 6 would misinterpret these concepts and would make incorrect decisions. Further, FRBSs 5 and 6 cannot consider a different number of symptoms and complications for different patients; consequently, it is impossible to integrate

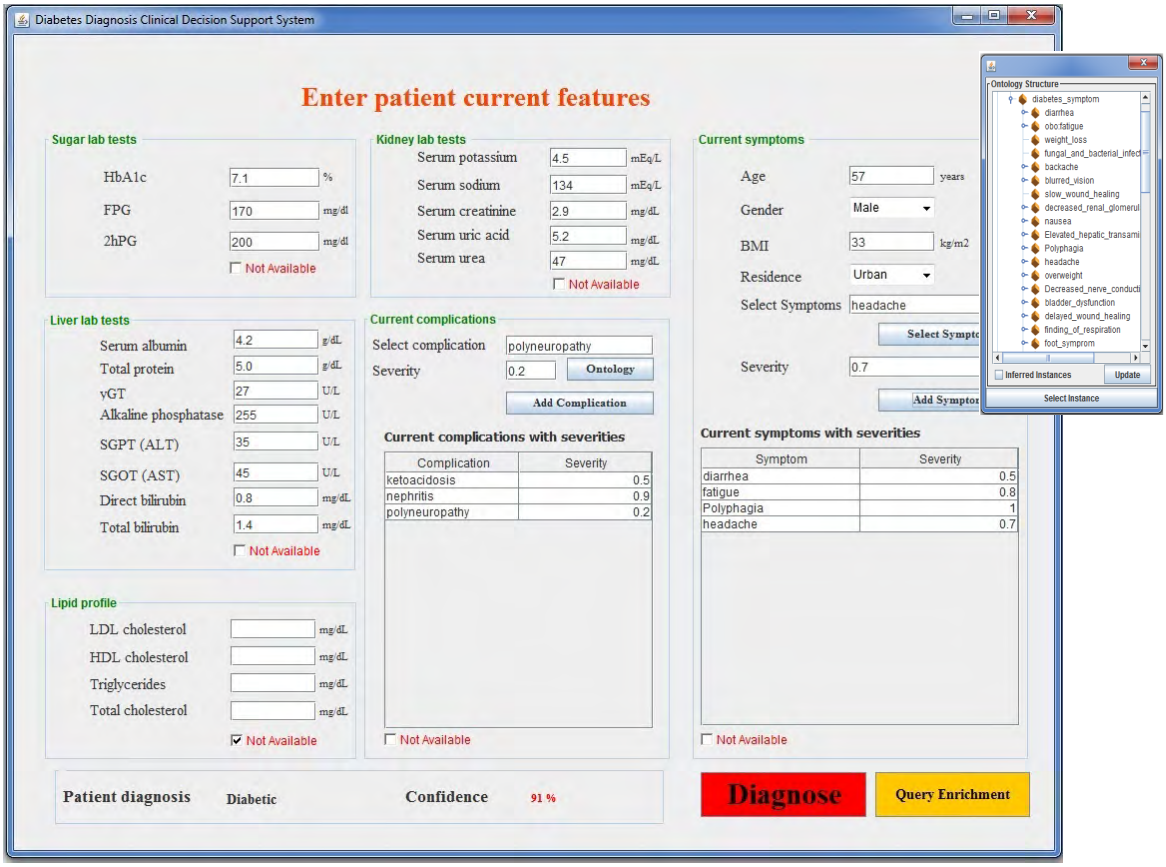

FIGURE 7. Query interface form for combined FRBS.

TABLE 12. SWH-FRBS performance after isolating specific subsystems.

# of utilized subsystems	Combinations	Type of subsystems	Acc (%)	AG (%)	P (%)	PG (%)	R (%)	RG (%)	FM (%)	FMG (%)
5	C1	With All subsystems	95	5	91.43	3.19	100	6.25	95.52	4.61
4	C2	Without glucose	86.67	15.07	81.58	14.14	96.88	6.25	88.57	11.24
	C3	Without kidney	100	10	100	3.57	100	15.62	100	10.01
	C4	Without liver	95	5	91.43	3.19	100	6.25	95.52	4.62
	C5	Without symptoms	85	-	79.48	-	96.88	6.25	87.32	-
	C6	Without lipid	95	5	91.43	3.19	100	6.25	95.52	4.62
3	C7	Without glucose and kidney	88.33	10	87.87	-	90.63	21.88	89.23	12.04
	C8	Without glucose and liver	86.67	15	81.58	14.14	96.87	6.24	88.57	11.24
	C9	Without glucose and lipid	86.67	15	81.58	14.14	96.87	6.24	88.57	11.24
	C10	Without glucose and symptoms	55	-	54.54	-	93.75	12.5	68.96	-
	C11	Without symptoms and lipid	85	-	79.48	-	96.88	6.26	87.32	-
	C12	Without symptoms and liver	86.66	-	80	-	100	9.38	88.88	1.01
	C13	Without symptoms and kidney	93.33	-	100	-	87.5	-	93.33	-
2	C14	Without glucose and kidney and liver	91.66	13.33	88.57	0.57	96.87	28.12	92.53	15.34
	C15	Without glucose and kidney and lipid	86.66	8.33	87.5	-0.5	87.5	18.75	87.5	10.31
	C16	Without glucose and kidney and symptoms	68.33	-	78.26	-	56.25	-	65.45	-
	C17	Without glucose and liver and lipid	86.66	14.99	81.57	14.13	96.87	6.24	88.57	11.24
	C18	Without glucose and liver and symptoms	60	-	57.14	-	100	18.75	72.72	1.49
	C19	Without glucose and lipid and symptoms	55	-	54.54	-	93.75	12.5	68.96	-

regular FRBSs in an EHR distributed environment or use these systems in mobile CDSS. Semantic FRBS addressed all of these challenges.

Table 12 presents the results of the combined SWH-FRBS and the gains or improvements in comparison with WH-FRBS. The system has been validated as a whole and by isolating specific subsystems. The purpose of this isolation is to determine the effect of each subsystem. From a medical point of view, the unavailability of information can affect the level of confidence of the resulting decision, even if the decision has high accuracy. The system demonstrated an overall accuracy of 95% (C1) when all the data were available. In the case of isolating one subsystem, the poorest performance was achieved in C5, where symptoms were not available. As glucose level and symptoms are the most critical features, their absence in C10 and C19 provided an accuracy of 55%. As it can be observed, there are accuracy gains (AG), precision gains (PG), recall gains (RG), and F-measure gains in the majority of combinations.

The system's overall accuracy was improved by 5%. Because the symptom subsystem has the highest weight (0.3184) in the SWH-FRBS, its isolation causes no

improvements, as in C5, C10, C11, C12, C13, C16, C18, and C19.

The resulting SWH-FRBS achieved both the interpretability of the resulting fuzzy system and semantic interoperability with a distributed EHR environment. The main objective was accomplished in developing an FRBS that helps physicians diagnose DM in an accurate and medically intuitive manner. This FRBS can assess patient conditions based on the complete historical profile, helping even non-specialists offer patients an opportune, adequate, and on-time treatment. The proposed system considered the uncertain and vague nature of the medical domain. Further, it considered the semantic relationships between medical concepts. Consequently, this framework has an excellent opportunity to be utilized in the implementation of mobile health systems for monitoring patients transparently based on their EHRs, social media data, and sensor data. This remote patient monitoring can discover probable diabetics as early as possible to initiate treatment and prevent possible complications.

V. CONCLUSION

In this paper, the authors proposed a major advancement toward the improvement of medically acceptable CDSSs. The proposed system integrates ontology reasoning with fuzzy reasoning in a novel manner. The resulting system is more accurate, interpretable, dynamic, and interoperable. We first developed a hierarchical FRBS based on fuzzy sets and fuzzy rules learned from a real dataset. We used advanced techniques and APIs to achieve this step. Second, we extended this framework by integrating ontology reasoning in the fuzzy inference process in a novel manner. This extension improved the semantic intelligence of the system and its semantic interoperability with other CDSSs and distributed EHR systems. We used one of the most popular MCDM techniques named fuzzy AHP to determine the weight of every subsystem in the hierarchy. This step supported the ability to isolate any number of subsystems and continue making decisions. Further, the present study is a new and innovative proposal because it uses a complete list of diabetes features, which have not yet been used, even in systems that are similar to the proposed. The majority of the literature research discusses diabetes diagnosis based on the freely available Pima Indians Dataset. The proposed framework was implemented and tested using real cases and produced accurate results.

We expect that the developed framework is highly intelligent and can provide more accurate results than existing studies. It has an open architecture where other enhancements can be appended to extend its functionality. The framework can be applied in any other medical domain in a straightforward way.

The current limitations of this system are that it does not yet support full interoperability with the EHR system, and does not consider data regarding patients from social media. Although these two requirements are required for the success of an FRBS in a hospital, they did not affect the implementation of our interpretable fuzzy system. Full interoperability of CDSSs and EHR systems requires the utilization of standards in medical knowledge representation such as vMR, medical terminology representation such as SNOMED CT and LOINC, and standard data modeling such as HL7 FHIR and openEHR. In addition, using sentiment analysis techniques to extract patient symptoms and complications from social media (e.g. tweets and comments) add to the available data that CDSS uses to make decision. We consider these requirements as enhancements to the proposed system, and we they will be addressed in our future work. In addition, we will apply the resulting framework in the mobile health environment to build a cloud based distributed CDSS system to remotely monitor patients.

ACKNOWLEDGEMENT

This work was supported by National Research Foundation of Korea-Grant funded by the Korean Government (Ministry of Science, ICT and Future Planning)-NRF-2017R1A2B2012337).

REFERENCES

[1] K. Zarkogianni *et al.*, "A review of emerging technologies for the management of diabetes mellitus," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 12, pp. 2735–2749, Dec. 2015.
[2] F. Mansourypoor and S. Asadi, "Development of a reinforcement learning-based evolutionary fuzzy rule-based system for diabetes diagnosis," *Comput. Biol. Med.*, vol. 91, pp. 337–352, Dec. 2017.
[3] L. Bickley, *Bates' Guide to Physical Examination and History-Taking*, 11th ed. Philadelphia, PA, USA: LWW, 2013.
[4] A. Minutolo, M. Esposito, and G. De Pietro, "A fuzzy framework for encoding uncertainty in clinical decision-making," *Knowl.-Based Syst.*, vol. 98, pp. 95–116, Apr. 2015.
[5] M. Maniruzzaman *et al.*, "Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm," *Comput. Methods Programs Biomed.*, vol. 152, pp. 23–34, Dec. 2017.
[6] O. Biran and C. Cotton, "Explanation and justification in machine learning: A survey," in *Proc. Int. Justification Conf. Artif. Intell. Workshop Explainable Artif. Intell.*, 2017, pp. 8–13.
[7] J. M. Alonso, C. Castiello, M. Lucarelli, and C. Mencar, "Modeling interpretable fuzzy rule-based classifiers for medical decision support," in *Medical Applications of Intelligent Data Analysis: Research Advancements*. Hershey, PA, USA: IGI Global, 2013, pp. 1064–1081.
[8] S. Guillaume and B. Charnomoridc, "Learning interpretable fuzzy inference systems with FisPro," *Inf. Sci.*, vol. 181, no. 20, pp. 4409–4427, 2011.
[9] L. A. Zadeh, "From computing with numbers to computing with words. From manipulation of measurements to manipulation of perceptions," *IEEE Trans. Circuits Syst. I. Fundam. Theory Appl.*, vol. 46, no. 1, pp. 105–119, Jan. 1999.
[10] J. M. Alonso, A. Ramos-Soto, E. Reiter, and K. van Deemter, "An exploratory study on the benefits of using natural language for explaining fuzzy rule-based systems," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jul. 2017, pp. 1–6.
[11] J. M. Alonso, C. Castiello, and C. Mencar, "Interpretability of fuzzy systems: Current research trends and prospects," in *Springer Handbook of Computational Intelligence*. Berlin, Germany: Springer, 2015, pp. 219–237.
[12] R. Meza-Palacios, A. A. Aguilar-Lasserre, E. L. Ureña-Bogarín, C. F. Vázquez-RodrÃguez, R. Posada-GÃşmez, and A. Trujillo-Mata, "Development of a fuzzy expert system for the nephropathy control assessment in patients with type 2 diabetes mellitus," *Expert Syst. Appl.*, vol. 72, no. 1, pp. 335–343, Apr. 2017.
[13] A. D. Torshizi, M. H. F. Zarandi, G. D. Torshizi, and K. Eghbali, "A hybrid fuzzy-ontology based intelligent system to determine level of severity and treatment recommendation for benign prostatic hyperplasia," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 301–313, 2014.
[14] S. El-Sappagh, M. Elmogy, and A. M. Riad, "A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis," *Artif. Intell. Med.*, vol. 65, no. 3, pp. 179–208, 2015.

[15] S. Borgwardt, F. Distel, and R. Peñaloza, "The limits of decidability in fuzzy description logics with general concept inclusions," *Artif. Intell.*, vol. 218, pp. 23–55, Jan. 2015.

[16] F. Bobillo and U. Straccia, "An OWL ontology for fuzzy OWL 2," in *Foundations of Intelligent Systems* (Lecture Notes in Computer Science and Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 5722. Berlin, Germany: Springer, 2009, pp. 151–160.

[17] M. Buzzi *et al.*, "Facebook: A new tool for collecting health data?" *Multimedia Tools Appl.*, vol. 76, no. 8, pp. 10677–10700, 2017.

[18] M. del Pilar Salas-Zárate *et al.*, "Sentiment analysis on tweets about diabetes: An aspect-level approach," *Comput. Math. Methods Med.*, vol. 2017, Feb. 2017, Art. no. 5140631.

[19] E. H. Wagner *et al.*, "Quality improvement in chronic illness care: A collaborative approach," *Joint Commission J. Qual. Improvement*, vol. 27, no. 2, pp. 63–80, 2001.

[20] A. Riid and E. Rüstern, "Transparent fuzzy systems in modelling and control," in *Proc. Interpretability Issues Fuzzy Modeling*, 2003, pp. 452–476.

[21] A. C. Adoko, C. Gokceoglu, L. Wu, and Q. J. Zuo, "Knowledge-based and data-driven fuzzy modeling for rockburst prediction," *Int. J. Rock Mech. Mining Sci.*, vol. 61, pp. 86–95, Jul. 2013.

[22] S. Guillaume and B. Charnomordic, "Generating an interpretable family of fuzzy partitions from data," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 3, pp. 324–335, Jun. 2004.

[23] R. Alcalá, J. Casillas, O. Cordón, A. González, and F. Herrera, "A genetic rule weighting and selection process for fuzzy control of heating, ventilating and air conditioning systems," *Eng. Appl. Artif. Intell.*, vol. 18, no. 3, pp. 279–296, 2005.

[24] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena, "Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: An overview," in *Proc. Interpretability Issues Fuzzy Modeling*, vol. 128, 2003, pp. 3–22.

[25] A. E. Anderson, W. T. Kerr, A. Thames, T. Li, J. Xiao, and M. S. Cohen, "Electronic health record phenotyping improves detection and screening of type 2 diabetes in the general United States population: A cross-sectional, unselected, retrospective study," *J. Biomed. Inform.*, vol. 60, pp. 162–168, Apr. 2016.

[26] S. T. Liaw, J. Taggart, H. Yu, S. de Lusignan, C. Kuziemsky, and A. Hayen, "Integrating electronic health record information to support integrated care: Practical application of ontologies to improve the accuracy of diabetes disease registers," *J. Biomed. Inform.*, vol. 52, pp. 364–372, Dec. 2014.

[27] T.-M. Chang, H.-Y. Kao, J.-H. Wu, and Y.-F. Su, "Improving physicians' performance with a stroke CDSS: A cognitive fit design approach," *Comput. Hum. Behav.*, vol. 54, pp. 577–586, Jan. 2016.

[28] S.-M. Chen and M.-H. Huang, "Automatically generating the weather news summary based on fuzzy reasoning and ontology techniques," *Inf. Sci.*, vol. 279, pp. 746–763, Sep. 2014.

[29] L. Ahmadian, R. Cornet, and N. F. de Keizer, "Facilitating pre-operative assessment guidelines representation using SNOMED CT," *J. Biomed. Inform.*, vol. 43, no. 6, pp. 883–890, 2010.

[30] C. Sáez, A. Bresó, J. Vicente, M. Robles, and J. M. García-Gómez, "An HL7-CDA wrapper for facilitating semantic interoperability to rule-based clinical decision support systems," *Comput. Methods Programs Biomed.*, vol. 109, no. 3, pp. 239–249, 2013.

[31] S.-M. Chen, Y.-H. Huang, R.-C. Chen, S.-W. Yang, and T.-W. Sheu, "Using fuzzy reasoning techniques and the domain ontology for anti-diabetic drugs recommendation," in *Proc. Asian Conf. Intell. Inf. Database Syst., Intell. Inf. Database Syst. (ACIIDS)*, 2012, pp. 125–135.

[32] M. G. Tsipouras *et al.*, "Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 4, pp. 447–458, Jul. 2008.

[33] D. Pal, K. M. Mandana, S. Pal, D. Sarkar, and C. Chakraborty, "Fuzzy expert system approach for coronary artery disease screening using clinical parameters," *Knowl.-Based Syst.*, vol. 36, pp. 162–174, Dec. 2012.

[34] J. M. Alonso, C. Castiello, and C. Mencar, "A bibliometric analysis of the explainable artificial intelligence research field," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*. Cham, Switzerland: Springer, 2018, pp. 1–13.

[35] D. Gunning, "Explainable artificial intelligence (XAI)," Defense Advanced Research Projects Agency, Arlington, VA, USA, Tech. Rep. DARPA-BAA-16-53, 2016.

[36] G. Acampora, B. Di Stefano, A. Vitiello, W. van der Aalst, C. Gunther, and E. Verbeek, "IEEE 1855: The XES standard: The Second IEEE standard sponsored by IEEE computational intelligence society [society briefs]," *IEEE Comput. Int. Mag.*, vol. 12, no. 2, pp. 4–8, May 2016.

[37] *IEEE Standard for Fuzzy Markup Language*, IEEE Standard 1855-2016, 2016.

[38] T. L. Saaty, "How to make a decision: The analytic hierarchy process," *Eur. J. Oper. Res.*, vol. 48, no. 1, pp. 9–26, Sep. 1990.

[39] J. M. Alonso and L. Magdalena, "HILK++: An interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers," *Soft Comput.*, vol. 15, no. 10, pp. 1959–1980, 2011.

[40] A. D'Acierno, M. Esposito, G. De Pietro, "An extensible six-step methodology to automatically generate fuzzy DSSs for diagnostic applications," *BMC Bioinf.*, vol. 14, pp. 1–19, Jan. 2013.

[41] D. P. Pancho, J. M. Alonso, and L. Magdalena, "Quest for interpretability-accuracy trade-off supported by Fingrams into the fuzzy modeling tool GUAJE," *Int. J. Comput. Intell. Syst.*, vol. 6, pp. 46–60, Jun. 2013.

[42] C. Mencar and A. M. Fanelli, "Interpretability constraints for fuzzy information granulation," *Inf. Sci.*, vol. 178, no. 24, pp. 4585–4618, 2008.

[43] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychol. Rev.*, vol. 63, no. 2, pp. 81–97, 1956.

[44] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*, vol. 100. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[45] J. C. Bezdek, "Modified objective function algorithms," in *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York, NY, USA: Plenum, 1981, pp. 155–201.

[46] M.-Y. Chen, "Establishing interpretable fuzzy models from numeric data," in *Proc. 4th World Congr. Intell. Control Autom.*, vol. 3, 2002, pp. 1857–1861.

[47] S. La-Ongsri and J. F. Roddick, "Incorporating ontology-based semantics into conceptual modelling," *Inf. Syst.*, vol. 52, pp. 1–20, Aug./Sep. 2015.

[48] P. C. Sherimon and R. Krishnan, "OntoDiabetic: An ontology-based clinical decision support system for diabetic patients," *Arabian J. Sci. Eng.*, vol. 41, no. 3, pp. 1145–1160, 2016.

[49] S. El-Sappagh and F. Ali, "DDO: A diabetes mellitus diagnosis ontology," *Appl. Inform.*, vol. 3, no. 1, p. 5, 2016.

[50] S. Harispe, D. Sánchez, S. Ranwez, S. Janaqi, and J. Montmain, "A framework for unifying ontology-based semantic similarity measures: A study in the biomedical domain," *J. Biomed. Inform.*, vol. 48, pp. 38–53, Apr. 2014.

[51] E. Adel, S. El-Sappagh, S. Barakat, and M. Elmogy, "Distributed electronic health record based on semantic interoperability using fuzzy ontology: A survey," *Int. J. Comput. Appl.*, vol. 39, pp. 1–19, Dec. 2017.

[52] D. K. Choubey, S. Paul, and V. K. Dhandhenia, "Rule based diagnosis system for diabetes," *Biomed. Res.*, vol. 28, no. 12, pp. 5196–5209, 2017.

SHAKER EL-SAPPAGH received the bachelor's degree in computer science and the master's degree from the Information Systems Department, Faculty of Computers and Information, Cairo University, Egypt, in 1997 and 2007, respectively, and the Ph.D. degree in computer science from the Information Systems Department, Faculty of Computers and Information, Mansura University, Mansura, Egypt, in 2015. In 2003, he joined the Department of Information Systems, Faculty of Computers and Information, Minia University, Egypt, as a Teaching Assistant. Since 2016, he has been with the Department of Information Systems, Faculty of Computers and Information, Benha University, as an Assistant Professor. He is currently a Post-Doctoral Fellow with the UWB Wireless Communications Research Center, Department of Information and Communication Engineering, Inha University, South Korea. He has publications in clinical decision support systems and semantic intelligence. His current research interests include machine learning, medical informatics, (fuzzy) ontology engineering, distributed and hybrid clinical decision support systems, semantic data modeling, fuzzy expert systems, and cloud computing. He is a reviewer for many journals, and he is very interested in disease diagnosis and treatment research.

JOSÉ M. ALONSO received the M.S. and Ph.D. degrees in telecommunication engineering from the Technical University of Madrid, Spain, in 2003 and 2007, respectively. He is currently a Ramon y Cajal Researcher (RYC-2016-19802) with the Research Centre in Information Technologies, Universidade de Santiago de Compostela. He has published over 100 papers in international journals, book chapters, and in peer-review conferences. His research interests include interpretable fuzzy modeling, explainable artificial intelligence, natural language generation, knowledge extraction and representation, integration of expert and induced knowledge, data science and big data, and development of free software tools. He is a Chair of the IEEE-CIS Task Force on Fuzzy Systems Software, and an Associate Editor of the *IEEE Computational Intelligence Magazine*. He is also a Secretary of the European Society for Fuzzy Logic and Technology.

FARMAN ALI received the B.S degree in computer science from the University of Peshawar, Pakistan, in 2011, the M.S. degree in computer science from Gyeongsang National University, South Korea, in 2015, and the Ph.D. degree in information and communication engineering from Inha University, South Korea, in 2018. He is currently with Inha University, where he holds a postdoctoral position with the UWB Wireless Communications Research Center. His current research interests include fuzzy ontology, ontology-based semantic knowledge, fuzzy logic, type-2 fuzzy logic, healthcare monitoring system, sentiment analysis/opinion mining, information extraction, artificial intelligence in text mining, and machine learning-based text classification.

AMJAD ALI received the B.S. and M.S. degrees in computer science from the COMSATS Institute of Information Technology, Lahore, Pakistan, in 2006 and 2008, respectively, and the Ph.D. degree in electronics and radio engineering from Kyung Hee University, South Korea, in 2015. Since 2015, he has been serving as an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad Lahore Campus, Pakistan. He is currently a Post-Doctoral Fellow with the UWB Wireless Communications Research Center, Department of Information and Communication Engineering, Inha University, South Korea. He has published a number of research papers in refereed international journals and conferences. His main research interests include multimedia transmissions over cognitive radio networks, 5G cellular networks, multimedia cloud computing, Internet of Things, and intelligent transportation systems.

JUN-HYEOG JANG received the bachelor's degree from Kyungpook National University and the Ph.D. degree from Texas A&M University. He held fellowships in the Seoul National University Institute of Cancer Research and IBEC. He is currently a Professor and the Chair with the Department of Biochemistry, School of Medicine, Inha University, Incheon, South Korea. Among his many honors are the Bumsuk Academic Award in Medical Biochemistry and the Best Lecture Award from the Korean Society of Biomaterials. He is currently on the Board of Editors, the *Journal of Tissue Engineering* and the *World Journal of Gastrointestinal Oncology*.

KYUNG-SUP KWAK (M'81) received the Ph.D. degree from the University of California at San Diego in 1988. From 1988 to 1989, he was with Hughes Network Systems LLC, San Diego, CA, USA. From 1989 to 1990, he was with the IBM Network Analysis Center, Research Triangle Park, NC, USA. Since 1990, he has been with the School of Information and Communication Engineering, Inha University, South Korea, as a Professor, where he had been the Dean of the Graduate School of Information Technology and Telecommunications from 2001 to 2002. He has been the Director of the UWB Wireless Communications Research Center (formerly the Key National IT Research Center), South Korea, since 2003. In 2006, he served as the President of the Korean Institute of Communication Sciences, and in 2009, the President of the Korea Institute of Intelligent Transport Systems. In 2008, he was selected for an Inha Fellow Professor and currently for an Inha Hanlim Fellow Professor. He published over 200 peer-reviewed journal papers. His research interests include wireless communications, UWB systems, sensor networks, WBAN, and nano communications. He has received a number of awards, including the Engineering College Achievement Award from Inha University, the LG Paper Award, the Motorola Paper Award, the Haedong Prize of research, and various government awards from the Ministry of ICT, the President, and the Prime Minister of Korea, for his excellent research performances. He served as TPC/Track chair/organizing chair for several IEEE related conferences.