AN ONTOLOGY FOR DIABETES EARLY WARNING SYSTEM

[1]NOUF ALMUTAIRI, [2]RIYAD ALSHAMMARI, [3]IMRAN RAZZAK

[1,2,3]King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia
E-mail: razzakmu@ngha.med.sa

Abstract- Diabetes mellitus is one of the common diseases with a high prevalence across Saudi society. Diabetes is a chronic disease where human body is not able to produce the insulin or use it properly. There is a wealth of hidden information within Healthcare Information System (HIS) that can be used to support clinical decision-making in addition to useful information that can help healthcare provider to predict diseases before it occurs. This paper is about building the early warning system ontology. A models generated from the data mining algorithms will be used in building an early warning system that can benefit healthcare sector in providing prevention and control programs for community health improvement.

Keywords- Diabetes, Early Warning, Chronic Disease Ontology

I. INTRODUCTION

Nowadays, Semantic web is the solution for efficient web content and sharing knowledge among different systems. Semantic web works by define all related concepts and properties about specific domain, which is called ontology. Ontology is the representation of domain terms, concepts and the relationship between them. It is also contains the properties that describe each concepts in the domain. Ontology used to facilitate sharing and representing domain knowledge among people or software.

Ontology has been used in different applications and software to support the knowledge representation in a proper way. Different studies used ontology for diabetes but in different concept like diet planner and CDSS for diabetic patient where different classes and relationships were used. Different SPARQL queries were applied to answer different questions.Chen and Bau (2013), apply ontology approach to build the clinical guidelines in the clinical decision support system CDSS. In this research theyproposed a management guideline for diabetic patient by designing an ontology-driven CDSS. This system will address the issue and the guideline that similar to what clinicians have during treatment of the patient. Kafalı, et al., (2013), buildan ontology reasoning component ORC, a tool that embedded on a personal health setting for diabetic patients and the clinician that treat them. On other hand Cantais et al. (2005), apply ontology concept to assist diabetic patient in their planning for a good diet plan, which help them to control that disease.

Early Warning System will be build based on knowledge that extracted from applying data mining in healthcare data. This data contains different concepts and attributes that describe the domain of the information that gathered from patient data. The ontology used to represent the domain concept as classes and subclasses.Classes and subclasses are described by defining the properties for each class. However, all the relationship between these classes should be define to retrieve needed information. The rest of the paper is organized as: the methodology that has been used will be described in next section. The classes, properties, and the relation are presented in the implementation section followed by discussion.

II. METHODOLOGY

To build the system ontology, we have used theProtégé4.3 that support OWL DL [4-5].Noy & McGuinness (2001), definedthe following processes to define the domain ontology. These processes were followedin this project as below:

1. Define ontology domain and scope: by list of the questions that related to the domain, which called competency questions. These questions help to know all the terms and concepts that would be used during building the diabetes ontology.

2. Reuse existing ontology: in the Web there are many libraries of ontology that can be use and modify based on our needs.To find a sufficient ontology that related or similar to the proposed system,related literatures was explored. Up to our knowledge, other literatures were considering diabetes from different perspective like CDSS for diabetic patients, Personal health record for diabetic patient and diet plan. As a result,webuild the ontology for diabetes warning system instead of existing ontology.

3. Define the most important ontology terms: all terms that can be related to the system and the domain of diabetes should be defined. This will facilitate to define the classes, properties and relation within the system.

4. Define the classes and the class hierarchy:
Three approaches to define classes:

- Top-down: create the general concept then define the subclasses for each one on general class.
- Bottom-up: define of the most specific classes then move to the general one.
- A combination of the above approaches.

5. Define the properties of classes–slots: for each classes a slots or properties that describe the class

relationship have to be defined for example for patient class there are many slots like has diagnosis and has medication. In protégé it is called objects properties.

6. Define the facets of the slots: for each slot there is a fact that contains the value of that slot like patient name in patient class. Also each value described by slot-value type (string, number and etc) and the slot cardinality (define how many values slot can have).

7. Create instances:for each class create individual instances.This instance is like an object for that class and have the same properties and data type of that class.

III. IMPLEMENTATION

This section discusses the building of ontology for diabetes early warning system that build based on knowledge learned from applying data mining algorithms. Domain was defined to determine concepts, terms, relationship and objects that used in the diabetes ontology. We have defined different classes, properties, relations, individuals for diabetes ontology and finally sparql query was performed to answer and retrieve information from the ontology.

1. Classes:Main domain concepts are represented as class as shown in Table 1.The list of the classes and description of each one is presented in table 1 whereas the figure 1shows the classes in protégé tool.

Table 1: Table 1: Domain class

Class Name	Description
Patient	The main class that contains patient information.
Proc	Procedure class is contain all procedure that patient received it during the hospital visit like lab, radiology and etc.
Disease	List of diseases that patient come to the hospital to get treatment.
Medication	List of the prescribed patient's medication during each visit with the information about the dosage and frequencies.
Behavior	Patient's behavior about the diet and smoking.

Classes Disjointness: when two classes disjoint, it is mean no object can be related to these classes at the same time like no object can be proc and medication. Disjoint classes are described in Table 2 and figure 2.

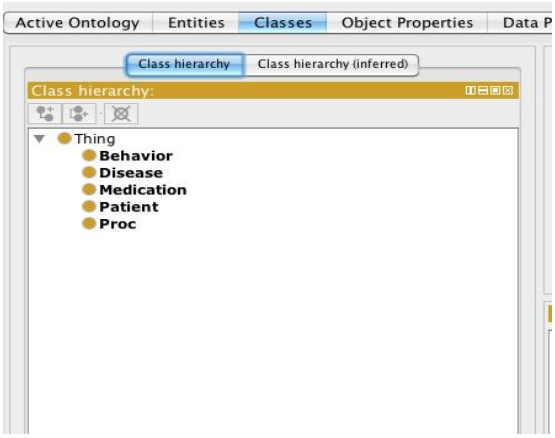


Figure 1: Class in Protégé

Table 2: Disjoint Classes

Class Name	Disjoint With
Proc	Medication
Disease	Proc
Proc	Disease

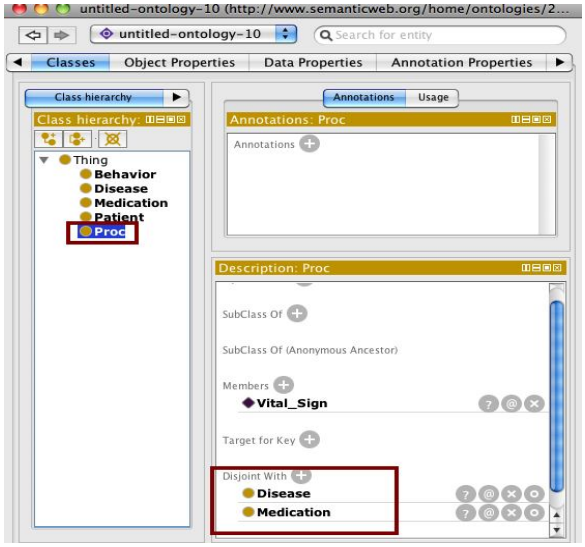


Figure 2: Disjoint Classes in protégé

2. Properties

Object Properties:objectproperties usedto describe the relationship between the classes, where for each objectproperties, domain and range should be define. Objectproperties defined as in Table 3. And in figure3 the object properties in protégé.

Inverse Properties: some object properties have inverse properties, which represent the inverse relationship of those properties. Table 4 contain inverse Properties of this ontology and in figure 4 the inverse Properties in protégé

Table 3 Object Properties

Object Properties	Domain	Range
Diagnose_By	Disease	Proc
Diagnosis_with	Patient	Disease
Has_Behavior	Patient	Behavior
Has_Medication	Patient	Medication
Has_Proc	Patient	Proc
Treated_By	Disease	Medication
Belong_to	Behavior	Patient
Related_to	Medication	Disease
Medication_of	Medication	Patient

Table 4: Inverse Properties

Object Properties	Inverse To
Belong_to	Has_Behavior
Related_to	Treated_By
Medication_of	Has_Mediaction

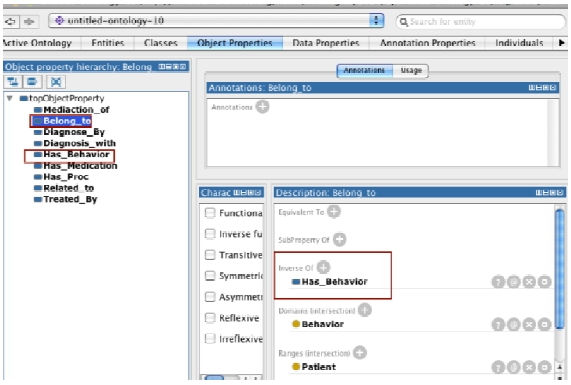


Figure 4: Inverse Properties in Protégé

Property Restrictions: used to apply restrictions to the relationship between the classes. There is different type of these restrictions like somevalue, allvalue and hasvalue.Property restriction of this ontology listed inTable 5andFigure 5.

Table 5Property Restrictions

Object Properties	Restrictions	Class
Has_Medication	Some	Medication
Has_Proc	Some	Proc
Diagnose_By	Some	Proc
Belong_to	Some	Patient

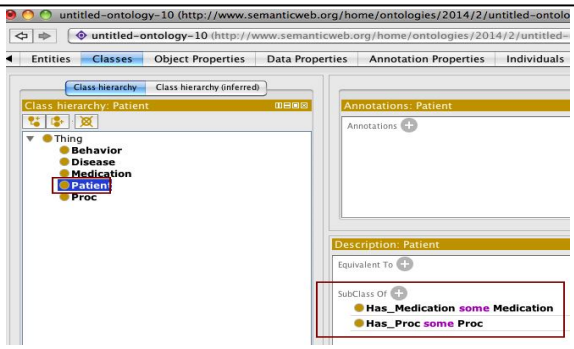


Figure 5:Property Restrictions in Protégé

Data Properties: For each class dataproperties describeobjects to the data type. Table 6 describesdata propertiesof this ontology,and in figure6the data properties in protégé.

Table 6: Data Properties

Data Properties	Domain	Range
Age	Patient	Integer
BMI	Patient	Integer
Diet	Behavior	Sting
Dosage	Medication	String
Frequency	Medication	String
Gender	Patient	String
History	Disease	String
Medication_Name	Medication	String
Proc_Name	Proc	String
Proc_Value	Proc	String
Smoking	Behavior	String

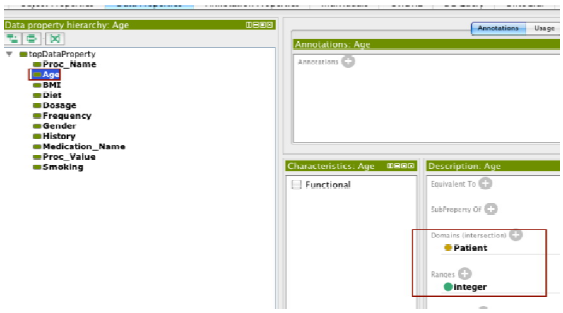


Figure 6: Data Properties in Protégé

3. Individuals

Classes object is define in this ontology as define in Figure 7.

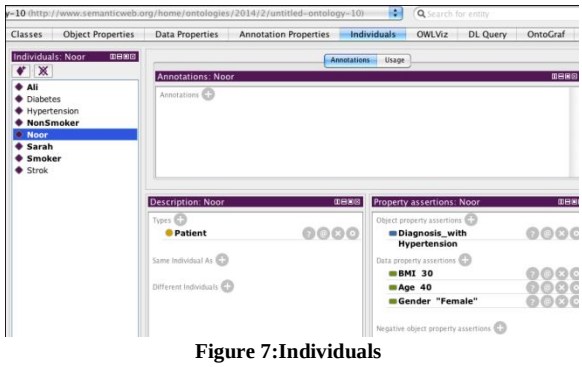


Figure 7:Individuals

4. SPARQL Query: Different sparql query can be apply to answer different question like:

- What medication affect on diabetes treatment?
- List all patient name?
- List of Patient with age grate than 20?
- What factor can control diabetes disease?
- What kinds of diseases usually come with disease?

Query 1: List Of all Patient

```
PREFIX rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl:
<http://www.w3.org/2002/07/owl#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs:
<http://www.w3.org/2000/01/rdf-schema#>
PREFIX
NN:<http://www.semanticweb.org/home/ontologies/2014/2
/untitled-ontology-10#>
SELECT *
WHERE {
?Patient NN:Age ?Age.
?Patient NN:Gender ?Gender.
 What kinds of diseases usually come with disease?
?Patient NN:BMI ?BMI}
```

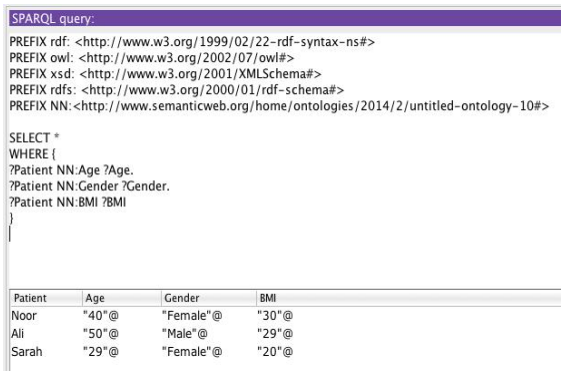


Figure: Result: Query list of all patient

Query 2: List Of all Female Patient

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX
NN:<http://www.semanticweb.org/home/ontologies/2014/2/untitled-ontology-10#>

SELECT *
WHERE {
?Patient NN:Age ?Age.
?Patient NN:Gender ?Gender
FILTER (str(?Gender) ="Female").
?Patient NN:BMI ?BMI
```

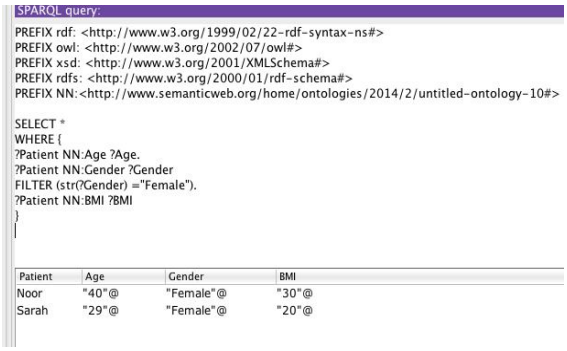


Figure: List Of all Female Patient

Query 3: List of the patient with the diagnoses:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX
NN:<http://www.semanticweb.org/home/ontologies/2014/2/untitled-ontology-10#>
SELECT *
WHERE {
?Patient NN:Age ?Age.
?Patient NN:Gender ?Gender.
?Patient NN:BMI ?BMI.
?Patient NN:Diagnosis_with ?Disease
}
```

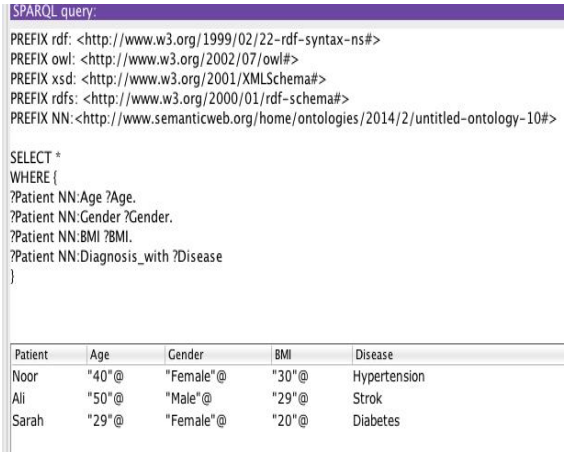


Figure: Result: List of the patient with the diagnoses:

Query 4: List Of all Patient with Age > 30 with diagnoses:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX
NN:<http://www.semanticweb.org/home/ontologies/2
014/2/untitled-ontology-10#>
SELECT *
WHERE {
?Patient NN:Age ?Age
FILTER (?Age > 30).
?Patient NN:Gender ?Gender.
?Patient NN:BMI ?BMI.
?Patient NN:Diagnosis_with ?Disease
}
```

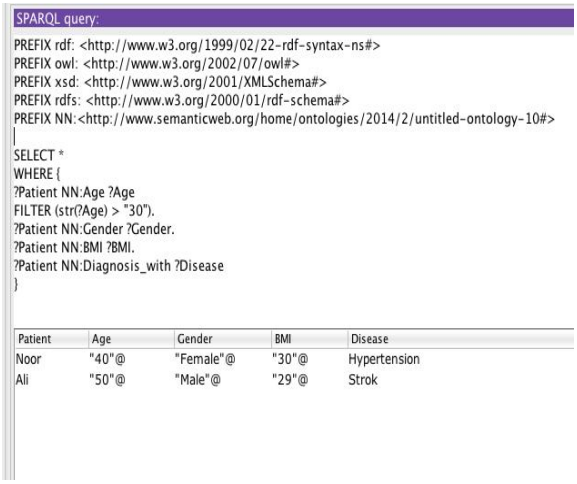


Figure :Result: Query 4: List of all Patient with Age > 30 with diagnoses:

5. Reasoner

Reasoner used to check the consistency. FaCT++ was applied and the ontology was consistent.

CONCLUSIONS

Healthcare data is going to be analyzed to discover hidden information and extract knowledge to improve the health status of the diabetic patients. Diabetes early warning system that based on the learning model generated from applying data mining algorithms will be implemented. Such a system can serve as an assistant tool for physicians and nurses to make better clinical decisions and also it can be utilized for patient protection and control plan. To share the system knowledge, diabetes ontology was build. Different SPARQL queries were applied to answer different questions. Protégé tool was used to build the diabetes ontology.

REFERENCES

[1] Chen, R. C., & Bau, C. T. An Ontological Approach for Guideline-based Decision Support System.

[2] Kafalı, Ö., Sindlar, M., van der Weide, T., & Stathis, K. (2013, August). ORC: an Ontology Reasoning Component for Diabetes. In International Workshop on Artificial Intelligence and NetMedicine (p. 71).

[3] Cantais, J., Dominguez, D., Gigante, V., Laera, L., & Tamma, V. (2005). An example of food ontology for diabetes control. In Proceedings of the International Semantic Web Conference 2005 workshop on Ontology Patterns for the Semantic Web.

[4] http://protege.stanford.edu/

[5] http://protegewiki.stanford.edu/wiki/DLQueryTab

[6] Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology.

★★★