

The Islamic University–Gaza  
Research and Postgraduate Affairs  
Faculty of Information Technology  
Master of Information Technology



الجامعة الإسلامية – غزة  
شؤون البحث العلمي والدراسات العليا  
كلية تكنولوجيا المعلومات  
ماجستير تكنولوجيا المعلومات

## Diagnosing Heart Diseases Using Ontology and SWRL Rules

تشخيص أمراض القلب باستخدام الأنتولوجيا وقواعد SWRL

Hosam Mohammed Alagha

Supervised by

Dr. Rebhi S. Baraka

Associate Professor of Computer Science

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Science in Information Technology

March/2017



## نتيجة الحكم على أطروحة ماجستير

بناءً على موافقة شئون البحث العلمي والدراسات العليا بالجامعة الإسلامية بغزة على تشكيل لجنة الحكم على أطروحة الباحث/ حسام محمد مصطفى الأغا لنيل درجة الماجستير في كلية تكنولوجيا المعلومات برنامج تكنولوجيا المعلومات وموضوعها:

### تشخيص أمراض القلب باستخدام الأنتولوجيا وقواعد SWRL

### Diagnosing Heart Diseases Using Ontology and SWRL Rules

وبعد المناقشة التي تمت اليوم الأربعاء 15 رجب 1438هـ، الموافق 2017/04/12م الساعة التاسعة صباحاً، اجتمعت لجنة الحكم على الأطروحة والمكونة من:

.....  
.....  
.....

مشرفاً و رئيساً

مناقشاً داخلياً

مناقشاً خارجياً

د. ربحي سليمان بركة

د. إياد محمد الأغا

د. سناء وفا الصايغ

وبعد المداولة أوصت اللجنة بمنح الباحث درجة الماجستير في كلية تكنولوجيا المعلومات / برنامج تكنولوجيا المعلومات.

واللجنة إذ تمنحه هذه الدرجة فإنها توصيه بتقوى الله ولزوم طاعته وأن يسخر علمه في خدمة دينه ووطنه.

والله ولي التوفيق،،،

نائب الرئيس لشئون البحث العلمي والدراسات العليا

أ.د. عبدالرؤوف علي المناصرة



## Abstract

Heart disease is the most common disease worldwide which is considered the main cause of death. According to cardiologists in Palestine, the heart disease was the main cause of death among the Palestinians, at a rate of 27.5% of all deaths.

The differential diagnosis between different types of heart diseases requires the results of several clinical tests. The patient's symptoms alone are not sufficient to give an accurate diagnosis because many types of heart diseases have the same symptoms. Currently there is no specific system in the domain of heart diseases in Palestine. Also, available medical systems do not employ semantic approaches, they are just using database-oriented methodologies. They are not flexible and adaptable to complex requirements and processes and lack intelligence.

This work aims to improve the diagnosis of heart diseases by exploiting Semantic Web technologies. We use ontology and Semantic Web Rule Language (SWRL) to diagnose heart diseases. We have built a domain ontology (HeartOnt) that covers domain knowledge of heart diseases. The ontology contains terms, relationship and properties to be used in the approach of diagnosing heart diseases. SWRL rules are created from valid relationships between ontology concepts to detect heart disease and estimate the risk of heart disease. The rules are used to infer new knowledge from the ontology, knowledge base and patient data.

The proposed system was tested using a sample set of patients with heart diseases provided by a domain expert. Results have shown that the system have correctly diagnosed 27 out of the 30 patients (ratio of correctness is 90%).

**Keywords:** Diagnosis heart diseases, Semantic Web, Ontology, SWRL Rules, OWL, Inference Engine.

## الملخص

تعتبر أمراض القلب من أكثر الأمراض شيوعاً في جميع أنحاء العالم وهي السبب الرئيسي للوفاة. وفقاً لأطباء القلب في فلسطين، تعتبر أمراض القلب السبب الرئيسي للوفاة بين الفلسطينيين، بمعدل 27.5% من مجموع الوفيات.

التشخيص الدقيق لمختلف أنواع أمراض القلب يتطلب نتائج عدة اختبارات سريرية. أعراض المريض وحدها ليست كافية لإعطاء تشخيص دقيق، وذلك لوجود عدد من أمراض القلب المختلفة لديها نفس الأعراض. حالياً لا يوجد أنظمة محوسبة تخدم مجال أمراض القلب في فلسطين. كذلك، فإن الأنظمة الطبية الحالية لا توظف النهج الدلالي، لأنها تستخدم منهجيات التعامل البسيط مع قواعد البيانات، فهي ليست ذكية ومرنة وقابلة للتكيف مع متطلبات العمليات المعقدة.

إن الهدف الأساسي من وراء هذا البحث هو تحسين تشخيص أمراض القلب من خلال تطوير نظام قادر على تشخيص أمراض القلب وتقدير خطر الإصابة واقتراح العلاج المناسب باستخدام تقنيات الويب الدلالي. يعتمد النظام بشكل أساسي على قاعدة المعرفة (الأنطولوجيا) التي أنشئت، بالتعاون مع خبير المجال، لنمذجة مختلف الكيانات في مجال أمراض القلب. تم كتابة القواعد الدلالية SWRL من خلال العلاقات الصحيحة بين مكونات الأنطولوجيا وفقاً لإرشادات الممارسة السريرية لأمراض القلب. محرك الاستدلال سوف يستخدم القواعد الدلالية والأنطولوجيا لاستنتاج التشخيص الصحيح وتقدير خطر الإصابة واقتراح العلاج المناسب. هذا البحث يشرح تصميم وهيكلية النظام وتركز على كيفية الاستدلال على أساس الأنطولوجي وSWRL يمكن أن يؤدي إلى تحسين تشخيص أمراض القلب وتقديم نتائج صحيحة.

تم اختبار النظام المقترح باستخدام عينة من المرضى الذين تم تشخيص اصابتهم بأمراض القلب مسبقاً. وقد أظهرت النتائج أن النظام قام بتشخيص 27 من أصل 30 مريضاً بشكل صحيح (نسبة الدقة 90%).

**كلمات مفتاحية:** تشخيص أمراض القلب، الويب الدلالي، أنطولوجيا، لغة أنطولوجيا الويب، لغة قواعد الويب الدلالي، محرك الاستدلال.

## Dedication

*To my beloved parents*

*To my dear Wife*

*To my brothers and sisters*

## Acknowledgements

At the very outset, all my prayers and thankfulness are to Allah for giving me the strength and ability to complete this thesis.

I would like to thank my supervisor, **Dr. Rebhi Baraka**, for the guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly.

I should thank to domain expert, **Dr. Amal Jamee** the head of cardiology department of Al-Shifa hospital for helping me to obtain the necessary information in the field of heart diseases.

To my father, who taught me the value of hardwork and an education, who encouraged me to be the best I can be. Without his, I may never have gotten to where I am today.

To my mother, without her continuous support and encouragement I never would have been able to achieve my goals.

No acknowledgement would be complete without expressing my gratitude and thankfulness for my wife; without her encouragement, I can't do this work.

I also thank my brothers and sisters, I also want to thank my friends, for their moral support during this study.

**Hosam M. Alagha**  
**March, 2017**

## Table of Contents

|  |             |
|--|-------------|
| <b>Abstract.....</b>                           | <b>I</b>    |
| <b>المخلص.....</b>                             | <b>II</b>   |
| <b>Dedication .....</b>                        | <b>III</b>  |
| <b>Table of Contents .....</b>                 | <b>V</b>    |
| <b>List of Tables .....</b>                    | <b>VIII</b> |
| <b>List of Figures.....</b>                    | <b>IX</b>   |
| <b>List of Abbreviations .....</b>             | <b>X</b>    |
| <b>Chapter 1 Introduction .....</b>            | <b>2</b>    |
| 1.1 Statement of the problem .....             | 4           |
| 1.2 Objectives .....                           | 4           |
| 1.2.1 Main objective .....                     | 5           |
| 1.2.2 Specific objective.....                  | 5           |
| 1.3 Importance of the Research .....           | 5           |
| 1.4 Scope and Limitations .....                | 6           |
| 1.5 Methodology .....                          | 6           |
| 1.5.1 information Acquisition.....             | 6           |
| 1.5.2 Construction of Domain Ontology .....    | 7           |
| 1.5.3 Writing SWRL Rules.....                  | 7           |
| 1.5.4 Develop a Prototype.....                 | 7           |
| 1.5.5 Evaluation .....                         | 8           |
| 1.6 Overview of the Thesis .....               | 8           |
| 1.7 Summary.....                               | 9           |
| <b>Chapter 2 State of the Art.....</b>         | <b>10</b>   |
| 2.1 Heart Diseases.....                        | 10          |
| 2.1.1 Heart Diseases Risk Factors .....        | 11          |
| 2.1.2 Diagnosing Heart Diseases .....          | 12          |
| 2.1.3 Heart Disease Risk Assessment .....      | 12          |
| 2.1.4 Clinical Practice Guidelines (CPGs)..... | 13          |
| 2.2 Semantic Web .....                         | 14          |
| 2.2.1 Predicate Logic .....                    | 14          |
| 2.2.2 Ontology .....                           | 15          |
| 2.2.3 Ontology Building Methodologies .....    | 15          |
| 2.2.4 Web Ontology Language (OWL) .....        | 17          |
| 2.2.5 Semantic Web Rule Language (SWRL) .....  | 18          |

|  |           |
|--|-----------|
| 2.2.6 Semantic Reasoner.....   | 19        |
| 2.3 Related Works.....   | 20        |
| 2.3.1 Ontology and Disease Diagnosis System .....                    | 20        |
| 2.3.2 Ontology and Clinical Decision Support Systems.....            | 23        |
| 2.4 Summary .....  | 26        |
| <b>Chapter 3 Heart Diseases Ontology .....</b>                       | <b>25</b> |
| 3.1 Ontology Overview.....   | 25        |
| 3.2 Classification of Heart Diseases Domain .....                    | 25        |
| 3.2.1 Extraction and classification of the heart diseases.....       | 26        |
| 3.2.2 Extracting the characteristics related to each disease .....   | 29        |
| 3.2.3 Building the hierarchy of concepts .....                       | 30        |
| 3.2.4 Identifying the relations between concepts and properties..... | 31        |
| 3.3 Ontology Development.....  | 32        |
| 3.4 Summary .....  | 41        |
| <b>Chapter 4 Writing SWRL Rules .....</b>                            | <b>43</b> |
| 4.1 Rules for Diagnosis Heart Diseases.....                          | 43        |
| 4.2 Rules for Heart Diseases Risk Estimation .....                   | 46        |
| 4.3 SWRL predicates .....  | 49        |
| 4.3.1 SWRL rules with class expressions .....                        | 49        |
| 4.3.2 SWRL rules with data range restrictions .....                  | 49        |
| 4.3.3 SWRL rules with core built-ins .....                           | 50        |
| 4.3.4 SWRL rules with custom built-ins.....                          | 51        |
| 4.4 Summary.....   | 51        |
| <b>Chapter 5 The DHDOno System.....</b>                              | <b>53</b> |
| 5.1 Knowledge Base Component.....                                    | 55        |
| 5.2 Rule Base Component .....  | 55        |
| 5.3 Inference Engine Component .....                                 | 56        |
| 5.4 Front End UI Component .....                                     | 56        |
| 5.4.1 Patient Profile Interface .....                                | 56        |
| 5.4.2 Diagnosis Result Interface .....                               | 57        |
| 5.4.3 Risk Estimation Result Interface .....                         | 58        |
| 5.4.3 Patient Questionnaire Interface.....                           | 59        |
| 5.5 Process of Heart Diseases Diagnosis .....                        | 60        |
| 5.5.1 Patient Questionnaire .....                                    | 60        |
| 5.5.2 Create Patient Medical Profile .....                           | 61        |
| 5.5.3 Run Inference Engine .....                                     | 61        |



|  |           |
|--|-----------|
| 5.5.4 Patient Risk Factors Analysis .....                  | 61        |
| 5.5.5 Calculate Heart Risk Score .....                     | 62        |
| 5.5.6 Giving the recommended diagnosis.....                | 63        |
| 5.6 Case study .....                                       | 63        |
| 5.7 Implementation Issues .....                            | 68        |
| 5.8 Summary .....  | 70        |
| <b>Chapter 6 Experimental Results and Evaluation .....</b> | <b>71</b> |
| 6.1 Experimental Settings .....                            | 71        |
| 6.2 Evaluation of Diagnosis Results .....                  | 73        |
| 6.3 Summary .....  | 81        |
| <b>Chapter 7 Conclusions and Future Work .....</b>         | <b>83</b> |
| <b>References.....</b>                                     | <b>86</b> |
| <b>Appendix A .....</b>                                    | <b>90</b> |
| Rules for the Diagnosis of Heart Diseases.....             | 90        |
| Rules for Heart Disease Risk Estimation.....               | 94        |

## List of Tables

|   |    |
|---|----|
| Table (3.1): Types of heart diseases and their treatments.....                                  | 28 |
| Table (3.2): Heart diseases and their characteristics.....                                      | 30 |
| Table (3.3): Classes of heart disease ontology .....  | 34 |
| Table (3.4): Object properties in heart disease ontology .....                                  | 36 |
| Table (3.5): Data properties in HeartOnto .....   | 37 |
| Table (4.1): Risk factors to determine the likelihood of Acute Coronary Syndrome diseases. .... | 44 |
| Table (4.2): Heart Disease Risk Points. ....  | 46 |
| Table (4.3): Heart Disease Risk. ....   | 47 |
| Table (4.4): SWRL rules for calculating the HDL level score. ....                               | 48 |
| Table (5.1): Patient data which are entered through the questionnaire interface.....            | 64 |
| Table (6.1): The number of patients and their disease type. ....                                | 71 |
| Table (6.2): An example of a patient data that have been diagnosed by doctor. ....              | 72 |
| Table (6.3): The size of the heart disease ontology .....                                       | 73 |
| Table (6.4): The number of individuals .....  | 73 |
| Table (6.5): The number of patients and their diagnosis results. ....                           | 74 |
| Table (6.6): The patients that have been diagnosed.....   | 74 |

## List of Figures

|  |    |
|--|----|
| Figure (2.1): Relation child of married parents .....  | 19 |
| Figure (3.1): Part of heart diseases and their characteristics .....   | 29 |
| Figure (3.2): Core classes of heart diseases ontology and their relationships .....                                | 31 |
| Figure (3.3): Heart disease ontology classes in Protégé .....  | 35 |
| Figure (3.4): Classes and object properties.....   | 35 |
| Figure (3.5): Data restriction of hasName data property .....  | 39 |
| Figure (3.6): Data cardinality.....  | 39 |
| Figure (3.7): Instances of the HeartOnto .....   | 40 |
| Figure (5.1): Architecture of DHDOnto system .....   | 54 |
| Figure (5.2): A snapshot of patient profile interface .....  | 57 |
| Figure (5.3): A snapshot of the diagnosis result interface .....   | 58 |
| Figure (5.4): A snapshot of risk estimation result interface .....   | 59 |
| Figure (5.5): Part of patient questionnaire interface .....  | 59 |
| Figure (5.6): The process of heart disease diagnosis .....   | 60 |
| Figure (5.7): The SWRL rule to compute the patient risk score.....   | 62 |
| Figure (5.8): Instance of Patient class and its data properties .....  | 64 |
| Figure (5.9): Risk point for every risk factor.....  | 66 |
| Figure (5.10): Diagnosis results contain risk estimation, treatment recommendation<br>and result explanation. .... | 68 |
| Figure (6.1): Diagnosis results for patient with id P08.....   | 77 |

## List of Abbreviations

|                 |                                   |
|-----------------|-----------------------------------|
| <b>ACS</b>      | Acute Coronary Syndrome           |
| <b>CDSS</b>     | Clinical Decision Support System  |
| <b>CPGs</b>     | Clinical Practice Guidelines      |
| <b>CT</b>       | Computerized Tomography           |
| <b>DBP</b>      | Diastolic Blood Pressure          |
| <b>DHDOnto</b>  | Diagnosis Heart Diseases Ontology |
| <b>HDL</b>      | High Density Lipoproteins         |
| <b>ECG</b>      | Electrocardiography               |
| <b>HeartOnt</b> | Heart Ontology                    |
| <b>MPI</b>      | Myocardial Perfusion Imaging      |
| <b>MRI</b>      | Magnetic Resonance Imaging        |
| <b>OWL</b>      | Web Ontology Language             |
| <b>RDF</b>      | Resource Description Framework    |
| <b>SBP</b>      | Systolic Blood Pressure           |
| <b>SW</b>       | Semantic Web                      |
| <b>SWRL</b>     | Semantic Web Rule Language        |

# Chapter 1

## Introduction

# Chapter 1

## Introduction

Heart disease is a group of conditions affecting the structure and functions of the heart. The term 'heart disease' is often used interchangeably with the term 'cardiovascular disease'. Cardiovascular disease generally refers to conditions that involve narrowed or blocked blood vessels. Heart disease is the most common disease worldwide which is considered the main cause of death (Lloyd-Jones et al., 2010).

The heart is a muscular organ, which receives blood through the blood vessels of the cardiovascular system. The purpose of the cardiovascular system (heart and blood vessels) is to provide the cells of the body with oxygen, nutrition, and essential fluids. There are many causes of heart disease. The most common cause is narrowing or blockage of the coronary arteries, the blood vessels that supply blood to the heart itself.

According to cardiologists in Palestine, there is an alarming rise of heart diseases among Palestinian population. The latest figure published by the Palestinian Health Information Center for 2015 indicates that the cardiovascular disease was the main cause of death among the Palestinians, at a rate of 27.5% of all deaths. (Ministry of Health, 2015).

Diagnosis is the result of a decision-making process made by physicians to identify patient's diseases from their signs, symptoms and clinical tests results. The differential diagnosis between different types of heart diseases are often hard to distinguish because symptoms are confused with other possible conditions easily. Accurate diagnosis requires the results of several clinical tests like electrocardiogram (ECG), X-ray and Stress Test. Currently there is no specific system in the domain of heart diseases in Palestine. Also, available medical systems do not employ semantic approaches, they are just using database-oriented methodologies. They are not flexible and adaptable to complex requirements and processes and lack intelligence.

In this research, our intention is to improve the diagnosis heart diseases by exploiting Semantic Web technologies. We use ontology and Semantic Web Rule Language (SWRL) to diagnose heart diseases. Ontology is a technology created for sharing the knowledge in machine understandable format. Ontology is defined as *'a formal explicit description of concepts in a domain of discourse (classes). Properties of each concept describe various features and attributes of the concepts (slots), and restrictions on slots (facets). Ontologies together with a set of individual instances of classes constitutes a knowledge base'* (Noy & McGuinness, 2001).

We build a prototype ontology (HeartOnt) that covers domain of knowledge from heart diseases. The ontology contains terms, relationship and properties to be used in the development of the heart diseases diagnosis approach. HeartOnt is represented by Web Ontology Language (OWL) format, which is an ontology language for building ontologies or knowledge bases. The current version of OWL is 'OWL 2' (Bechhofer, 2009). The OWL 2 language is not able to express all relations. For example, there is no way in OWL 2 to express the relation between individuals with which an individual has relations. In order to cover this limitation, we utilize SWRL to detect heart disease type from the given symptoms and clinical tests results.

SWRL was developed to be the rule language of the semantic web. It allows the users to write first order logic rules required to reason about the specified OWL individuals (Horrocks et al., 2004). SWRL rules are created from valid relationships between ontology concepts to diagnosis heart diseases and estimate the risk of heart diseases. The rules are used to infer new knowledge from existing ontology knowledge bases and patient data. All rules will be expressed in terms of ontology concepts (classes, properties, individuals).

Several works have used ontology and SWRL rules in the field of biomedical, especially in diagnosing heart diseases. Most of these works focused on part of heart diseases. Additionally, the diagnosis is only based on the patient's symptoms, without taking into account the clinical tests results. In this work, we focus on diagnosing all types of coronary artery diseases, valve diseases and cardiomyopathy diseases. The proposed system would be an improvement achievement for the medical sector in

Palestine to assist the medical professionals to diagnose/predict the risk of heart diseases.

The system functions are divided into four tasks which are heart disease diagnosis, heart diseases risk estimation, treatment recommendation and result explanation. Diagnosis process performed by checking the input provided by the users to different questions related to the symptoms and risk factors for heart disease. All diagnosis rules and relevant OWL knowledge are extracted from domain ontology. The Pellet reasoner uses patient data and rules to draw inference and gives recommended diagnosis, treatment recommendation and result explanation.

The proposed system is tested using a set of patients with heart diseases provided by a domain expert. Results have shown that the system can correctly diagnosis 27 out of the 30 patients with ratio of correctness of 90%.

## **1.1 Statement of the problem**

Heart disease is the most common disease worldwide which consider main cause of death. Currently there is no specific system in the heart diseases domain in Palestine. The main problem addressed in this research is how diagnosing all types of coronary artery diseases, valve diseases and cardiomyopathy diseases using ontology and SWRL rules according to clinical practice guidelines followed in Palestine.

This problem can be divided into the following:

- How to build domain ontology containing the various entities, concepts and relationships mentioned in a heart diseases domain.
- How to write SWRL rules according to medical guidelines followed in Palestine.
- How OWL 2 and SWRL rules can be used to improve the diagnosis results of heart diseases.

## **1.2 Objectives**

In this section, we present both main and specific objectives of the research work.



### **1.2.1 Main objective**

Building an ontology-based approach to diagnose heart diseases using SWRL rules.

### **1.2.2 Specific objective**

- Acquisition of information about the domain of heart diseases to build ontology and create valid SWRL rules.
- Building heart diseases ontology, including terms, relationship and properties to be used in the development of the proposed approach.
- Feeding the ontology with instances and information about heart diseases to create knowledge base.
- Writing SWRL rules depending on the valid relationships between diseases, symptoms and clinical test results to detect heart disease types and treatment recommendations.
- Implementing a prototype that uses the ontology and SWRL rules for heart diseases diagnosis.
- Evaluating the prototype for accuracy based on a previously diagnosed heart disease cases.

### **1.3 Importance of the Research**

- Heart diseases are the number one cause of death worldwide, diagnosing the disease in the right time could lower the danger that it may cause.
- Currently there is no specific system in the domain of heart diseases in Palestine. Also, currently available medical systems do not employ semantic approach.
- Building standardized representation (ontological) of heart diseases could lead to develop efficient and intelligent systems for retrieving specific and appropriate information from them.
- Developing an ontology-based system for diagnosing heart diseases assist physicians in diagnosing heart diseases.
- Providing facilities for the scholars and researchers in order to obtain satisfactory and intelligence results with minimal efforts.

- The developed ontology can be used as a basis for other medical applications.

## 1.4 Scope and Limitations

- This work is targeted to patients in Palestine, where the collected data are from local hospitals and physicians.
- We have built knowledge based and write the SWRL rules according to 'Guidelines for Diagnosis and Treatment of the Chronic Heart Failure' followed in Palestine.
- The ontology covers the heart diseases in terms of description of diseases, symptoms and treatment method.
- We focus on diagnosing all types of coronary artery diseases, valve diseases and cardiomyopathy diseases.
- We develop a prototype for heart diseases diagnosing and risk estimation not a complete system.
- Diagnosis results are divide into four types: diagnosis result, heart risk estimation, treatment recommendations and results explanation.

## 1.5 Methodology

In order to improve the heart disease diagnosis, we need to build an ontology-based approach using SWRL rules. To accomplish the objectives of the research, the following methodology is followed:

### 1.5.1 information Acquisition

Heart diseases knowledge acquisition is important phase in our work as the correctness of the results of the system relies mainly on the correctness and completeness of the knowledge base and on the reasoning process that will infer diagnosis result. In this phase, we collect information such as detailed description of heart diseases, symptoms, clinical tests and treatments from various resources.

There are many clinical practice guidelines that belong to different medical associations and bodies. Most of these resources have evidence-based guidelines. we

adopted ‘Guidelines for Diagnosis and Treatment of the Chronic Heart Failure’ (McMurray et al., 2012). We also held a variety of meetings and interviews with the heart diseases physicians, who are asked to describe in details the procedures they employ to diagnose and treat patients.

### **1.5.2 Construction of Domain Ontology**

To build the HeartOnt ontology, we follow the ontology building processes (Noy & McGuinness, 2001) using Protégé 5.1. The building of ontology consists of the following steps:

- Determine the domain and scope of the ontology.
- Enumerate the important terms in the ontology.
- Define the classes and the class hierarchy.
- Define the properties of classes.
- Define the facets of the slots.
- Create instances.
- Apply ontology reasoning.

### **1.5.3 Writing SWRL Rules**

In this phase, we write SWRL rules depending on the valid relationships to detect heart disease type from the given symptoms and clinical test results. The rules will be used to infer new knowledge from existing ontology knowledge base. All rules will be expressed in terms of ontology concepts (classes, properties, individuals). The SWRL rules will stored as Web Ontology Language (OWL) syntax in domain ontology (HeartOnto).

### **1.5.4 Develop a Prototype**

In this phase, we develop a prototype of the proposed approach using Java, the development includes:

- Specifying the requirements of the system.
- Identify the required tools.

- Design the system.
- Implement the prototype.
- Test the prototype.

### 1.5.5 Evaluation

The main goal of this evaluation is to assess the effectiveness of the heart diseases diagnosis offered by our system. we will ask the expert in heart diseases to select a sample data of patients with heart diseases. The sample contains the various cases needed to examine the various aspects of the system.

### 1.6 Overview of the Thesis

The thesis consists of seven chapters. They are organized as follows:

**Chapter 1: Introduction:** introduces the research, problem, objectives, importance, scope and limitations, and methodology followed in the research.

**Chapter 2: State of the Art:** focuses on the background that is related the theoretical and technical foundations needed for thesis work, the concepts of heart diseases diagnosis and semantic web technologies. In addition to the related work that used semantic web technologies to diagnose heart diseases.

**Chapter 3: Heart Diseases Ontology:** describes the steps of developing heart disease domain ontology (HeartOnto). The ontology is developed using Protégé and OWL.

**Chapter 4: The DHDOnto System:** presents the architecture of the proposed system that involves HeartOnto ontology and SWRL rules to improve heart disease diagnosis. It describes the system modules to be implemented and how they relate to each other.

**Chapter 5: System Evaluation:** presents a set of diagnosing to test and evaluation the system with discussion of the results.

**Chapter 6: Conclusions and Future Work:** conclude the research and presents possible future works.

## 1.7 Summary

In this chapter, we presented an introduction to the heart disease diagnosis and the terminologies related to it. We presented the research problem and the importance of using ontology and SWRL in the development of the heart disease diagnosis system. We also stated the scope and limitations of the research. The limitation is that the system model depends on specific domain ontology, and we build a system prototype not a complete system. We presented the methodology that will be followed in this research. Finally, the thesis structure has been described.

In the next chapter, we review and investigate the concepts of heart diseases and semantic web technologies that are used to develop heart diseases diagnosis system. In addition, we will present different related works which address diseases diagnosis using ontology and SWRL rules.

# **Chapter 2**

## **State of the Art**

## **Chapter 2**

### **State of the Art**

This chapter presents fundamental concepts and technical foundation related to heart diseases and semantic web technologies including Clinical Practice Guidelines (CPGs), ontology, ontology development, OWL and SWRL. It also various applications and works that used semantic web technologies to develop medical diagnosis systems.

#### **2.1 Heart Diseases**

Cardiovascular disease is the most common cause of death in developed countries (Go et al., 2014). Across the globe, the incidence of death from cardiovascular and circulatory diseases has risen by one third between 1990 and 2010 (Lozano et al., 2013).

Heart disease includes a range of conditions that affect the heart. Diseases under the heart disease umbrella include blood vessel diseases, such as coronary artery disease; heart rhythm problems (arrhythmias); and heart defects you're born with (congenital heart defects), among others.

Cardiovascular disease is caused by narrowed, blocked or stiffened blood vessels that prevent the heart, brain or other parts of the body from receiving enough blood. Cardiovascular disease symptoms may be different for men and women. For instance, men are more likely to have chest pain; women are more likely to have symptoms such as shortness of breath, nausea and extreme fatigue.

The symptoms include:

- Chest pain (angina).
- Shortness of breath.
- Pain, numbness, weakness or coldness in the legs or arms if the blood vessels in those parts of the body are narrowed.
- Pain in the neck, jaw, throat, upper abdomen or back.

### 2.1.1 Heart Diseases Risk Factors

There are many risk factors for heart diseases that lead to developing heart diseases including:

- **Age.** Aging increases the risk of damaged and narrowed arteries and weakened or thickened heart muscle.
- **Sex.** Men are generally at greater risk of heart disease. However, women's risk increases after menopause.
- **Family history.** A family history of heart disease increases the risk of coronary artery disease, especially if a parent developed it at an early age (before age 55 for a male relative, such as a brother or a father, and 65 for a female relative, such as a mother or a sister).
- **Smoking.** Nicotine constricts the blood vessels, and carbon monoxide can damage their inner lining, making them more susceptible to atherosclerosis. Heart attacks are more common in smokers than in nonsmokers.
- **Poor diet.** A diet that's high in fat, salt, sugar and cholesterol can contribute to the development of heart disease.
- **High blood pressure.** Uncontrolled high blood pressure can result in hardening and thickening of the arteries, narrowing the vessels through which blood flows.
- **High blood cholesterol levels.** High levels of cholesterol in the blood can increase the risk of formation of plaques and atherosclerosis.
- **Diabetes.** Diabetes increases the risk of heart disease. Both conditions share similar risk factors, such as obesity and high blood pressure.
- **Obesity.** Excess weight typically worsens other risk factors.
- **Physical inactivity.** Lack of exercise also is associated with many forms of heart disease and some of its other risk factors, as well.
- **Stress.** Unrelieved stress may damage the arteries and worsen other risk factors for heart disease.



### 2.1.2 Diagnosing Heart Diseases

The clinical tests used to diagnose the heart disease depend on patient condition. The doctor performs a physical exam and asks patient about the personal and family medical history before doing any tests. Besides blood tests and chest X-ray, tests to diagnose heart disease can include:

- **Electrocardiogram (ECG).** An ECG records these electrical signals and can help the doctor detect irregularities in heart's rhythm and structure.
- **Holter monitoring.** A Holter monitor is a portable device to record a continuous ECG, usually for 24 to 72 hours. Holter monitoring is used to detect heart rhythm irregularities that aren't found during a regular ECG exam.
- **Echocardiogram.** This noninvasive exam, which includes an ultrasound of the chest, shows detailed images of the heart structure and function.
- **Cardiac catheterization.** In this test, a short tube (sheath) is inserted into a vein or artery in the leg (groin) or arm. A hollow, flexible and longer tube (guide catheter) is then inserted into the sheath. Aided by X-ray images on a monitor, the doctor threads the guide catheter through that artery until it reaches the heart. The pressures in the heart chambers can be measured, and dye can be injected. The dye can be seen on an X-ray, which helps the doctor see the blood flow through the heart, blood vessels and valves to check for abnormalities.
- **Cardiac computerized tomography (CT) scan.** This test is often used to check for heart problems. In a cardiac CT scan, patient lie on a table inside a doughnut-shaped machine. An X-ray tube inside the machine rotates around the body and collects images of the heart and chest.
- **Cardiac magnetic resonance imaging (MRI).** For this test, patient lie on a table inside a long tube-like machine that produces a magnetic field. The magnetic field produces pictures to help the doctor evaluate the heart.

### 2.1.3 Heart Disease Risk Assessment

The Heart Disease Risk Assessment is based on findings from a major research project called the “Framingham Heart Study” in which three generations of men and women

from Framingham, Massachusetts were evaluated to better understand the causes of cardiovascular disease. According to Framingham heart study, the cardiovascular risk is calculated based on factors such as Gender, Age, Total Cholesterol, High-density lipoproteins (HDL) , Systolic Blood Pressure, Smoking habit, and Hypertension.

Total heart disease risk of over 20% over 10 years is defined as high-risk. People with moderate-to-high risk are more likely to be compliant with lifestyle changes and preventative medication if given information about their individual cardiovascular risk (Casebeer et al., 2009).

#### **2.1.4 Clinical Practice Guidelines (CPGs)**

Clinical Practice Guidelines (CPGs) are recommendations for clinicians about the care of patients with specific conditions. They should be based upon the practice experience. The Institute of Medicine defines clinical practice guidelines as '*statements that include recommendations, intended to optimize patient care, that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options*' (Steinberg, Greenfield, Wolman, Mancher, & Graham, 2011). The use of CPGs allows clinicians to improve the quality and enhance the accuracy of their decisions and activities which affects positively the patient care quality and reduce the cost of the treatment.

Producing clinical guidelines is not sufficient and implementation of guidelines presents a significant challenge. Several barriers to implementation of guidelines exist throughout the patient pathway, from problems with delayed referral, limited access to specialists and to specialist tests and rationing of some treatments. It remains difficult to ensure that all health care professionals are aware of new guidelines and implement them (Farooq, Hussain, Leslie, Eckl, & Slack, 2011).

There are many clinical practice guidelines resources that belong to different medical associations and bodies. Most of these resources have evidence-based guidelines. One of these guidelines is 'Guidelines for Diagnosis and Treatment of the Chronic Heart Failure' (McMurray et al., 2012). This guideline covers diagnosing and managing

heart diseases in people aged 18 and above. It aims to improve early diagnosis and treatment to improve the length and quality of life for people with heart diseases.

## **2.2 Semantic Web**

The Semantic Web (SW) is the extension of the World Wide Web that enables people to share content beyond the boundaries of applications and websites (Jain & Singh, 2013a). SW adds meaning to the content. It is a web of data described and linked in ways to establish context and semantics that adhere to defined grammar and language constructs (Hebeler, Fisher, Blace, & Perez-Lopez, 2011).

The main goal of the Semantic Web is to add logic to the current Web, in other words, express the meaning of data, the properties of objects, and the complex relationships existing between them by a series of formal rules, which would make information accessible to machines. Machine accessibility should be understood as representing information in such a way that it is possible to make queries based on the meaning (i.e., semantics) of the data, independent of the form in which the information is presented (Robu, Robu, & Thirion, 2006).

There are many existing applications and research projects in the medical and health sciences based on SW technologies. SW has a crucial role in information retrieval of biomedical vocabularies, terminologies, and taxonomies (Lorence & Spink, 2004). Most of these projects are related to their formalization, namely defining the classes of contained entities and, especially, establishing the relations among them to develop them into ontologies.

### **2.2.1 Predicate Logic**

Predicate logic, also known as first-order predicate calculus, is a collection of formal systems used in mathematics, philosophy, linguistics, and computer science. Predicate logic uses quantified variables over non-logical objects and allows the use of sentences that contain variables, so that rather than propositions such as ‘The heart is a muscular organ’, one can have expressions in the form ‘there exists X such that X is the heart and X is a muscular organ’ where there exists is a quantifier and X is a variable (Andrews, 2002).

Predicate logic is the logic which adds predicates (like  $P(x, y)$ ) which represent relations, i.e. produce true or false for a combination of values of the terms  $x$  and  $y$ . quantifiers: existential  $\exists$  ("there exists") and universal  $\forall$  ("for all"); and terms made of variables and functions, like  $f(x)$ ,  $g(y, z)$ . Thus, predicate logic can form formulas like:  $\forall x \exists y (P(x) \rightarrow Q(f(y)))$ .

First order predicate logic is the logic where the quantifiers can range only over elements of sets. In higher order logics, quantifiers can range over sets, functions and other objects. So, for example, the sentence that "every set of real numbers has a minimum" cannot be expressed in first order logic, because the quantification ranges over sets, not set elements.

Predicate logic is the core base for the semantic web particularly the ontology and semantic rule language which are essentially predicate logic statement of various types and quantifications.

### **2.2.2 Ontology**

Ontology is a formal specification of the concepts within a domain and their relationships. It is a methodology which describes the domain of knowledge structure in the specific area. It promotes various kinds of data processing intended to provide systematic and semantic links among groups of related concepts. Ontologies are used to represent knowledge. Domain knowledge is contained in the form of concepts, individuals belonging to these concepts and relationships between the concepts and between concepts and individuals (Sherimon, Vinu, Krishnan, & Takroni, 2013).

### **2.2.3 Ontology Building Methodologies**

Ontology is developed from information gathered by domain experts and assigned to the ontology expert in the form of a set of concepts, relationships and definitions. There are several well established and defined methodologies to develop an ontology (Saad, Salim, Zainal, & Muda, 2011). For manual construction of ontology, we follow mainly Noy and McGuinness methodology (Noy & McGuinness, 2001). It includes the following steps:

### 1. Determine the domain and scope of the ontology

The first step in ontology development is defining ontology domain and scope in which the ontology will be developed in order try to find an answer to questions such as: what is the domain that the ontology will cover? For what we are going to use the ontology? For what types of questions should the ontology provide answers? Who will use and maintain the ontology?

### 2. Consider reusing existing ontologies

It is almost always worth considering what someone else has done and checking if we can refine and extend existing sources for our particular domain and task. Reusing existing ontologies may be a requirement if our system needs to interact with other applications that have already committed to particular ontologies or controlled vocabularies (Saad et al., 2011).

### 3. Enumerate important terms in the ontology

It is useful to write down a list of all terms we would like either to make statements about or to explain to a user. What are the terms we would like to talk about? What properties do those terms have? What would we like to say about those terms?

### 4. Define the classes and the class hierarchy

This step defines classes used in an ontology domain. There are several possible approaches in developing a class hierarchy: a top down development process, which starts with the most general concepts and subsequent specialization of the concepts. Bottom-up starts with the most specific concepts or classes, the leaves of the hierarchy with subsequent grouping of these classes into more general concepts. Middleout is a combination of the top-down and bottom-up approaches starts with the salient concepts first and then generalize and specialize them appropriately (Noy & McGuinness, 2001).

## 5. Define the properties of classes (slots)

This step used to describe the attributes or properties of the classes. These properties Defined as the “slots” of the models. Once the classes have defined, the next step is to describe the internal structures (properties) of the concepts. Again, these should be readily available from the list produced because of Step 3.

## 6. Define the facets of the slots

Slots can have different facets describing the value type, allowed values, the number of the values (cardinality), and other features of the values the slot can take.

## 7. Create instances

The last step is creating individual instances of classes in the hierarchy. Defining an individual instance of a class requires (1) choosing a class, (2) creating an individual instance of that class, and (3) filling in the slot values.

## 8. Ontology Evaluation

The evaluation of the quality of ontology is an important part of ontology development. An ontology can be evaluated against many criteria: its coverage of a particular domain and the richness, complexity and granularity of that coverage; the specific use cases, scenarios, requirements, applications, data sources it was developed to address, formal properties such as the consistency and completeness of the ontology and the representation language in which it is modeled.

### **2.2.4 Web Ontology Language (OWL)**

Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things (Jain & Singh, 2013b). OWL is a computational logic-based language such that knowledge expressed in OWL can be exploited by computer programs. OWL has three sub-languages (Horridge et al., 2009):

- **OWL-Lite:** This sub-language intended to use in situations where only a simple class hierarchy and simple constraints needed, its cardinality is limited to either 0 or 1.
- **OWL-DL:** To fill the shortage of OWL-Lite, this sub-language comes with features that enrich the use of OWL. Class Boolean combinations and class property restrictions are some of the added features. Other properties in describing a class in term of other disjoint classes are another new feature. With all these features, OWL-DL becomes the most used language since it provides the user with full expressiveness (McGuinness & Van Harmelen, 2004).
- **OWL-Full:** this sub-language offers to its users maximum expressiveness and syntactic freedom of RDF (Roussey, Pinet, Kang, & Corcho, 2011). As instance, OWL-Full treats a class as a set of individuals and as an individual at the same time. Its data type property generalizes to include inverse functional property.

### 2.2.5 Semantic Web Rule Language (SWRL)

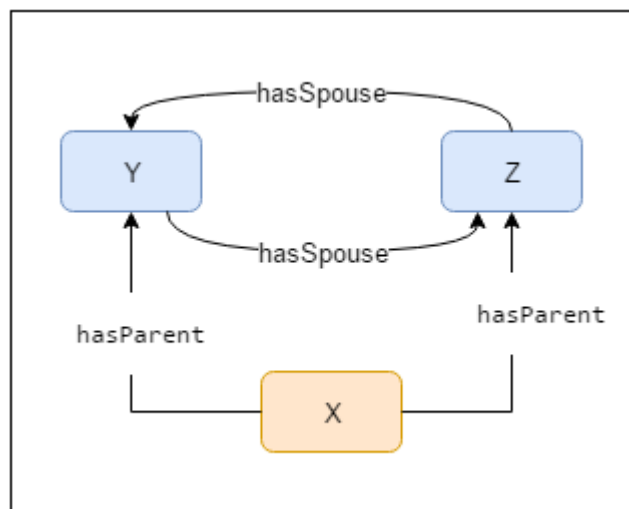
Semantic Web Rule Language (SWRL) was developed to be the rule language of the semantic web. SWRL allows users to write rules that can be expressed in terms of OWL concepts and that can reason about OWL individuals. One of SWRL's most powerful features is its ability to support custom built-ins, user defined, to extend SWARL's core built-ins so the user can achieve extra extensibility. (Horrocks et al., 2004).

There are several atom types that are supported by SWRL, such as class atoms, individual property atoms, data value property atoms, and data range atoms. The most powerful atoms are built-in atoms, where SWRL provides several types of existing built-ins and allow the user to design and use his own built-ins (O'connor et al., 2005).

OWL 2 language is not able to express all relations. For example, SWRL cannot express the relation child of married parents, because there is no way in OWL 2 to express the relation between individuals with which an individual has relations. In order to cover this limitation, SWRL rules can be formed as following:

$Person(?x), hasParent(?x, ?y), hasParent(?x, ?z), hasSpouse(?y, ?z) \rightarrow ChildOfMarriedParents(?x)$

The SWRL rule says that an individual X from the *Person* class, which has parents Y and Z such that Y *hasSpouse* Z, belongs to a new class *ChildOfMarriedParents*. Figure 2.1 depicts the relation child of married parents.



**Figure (2.1):** Relation child of married parents

### 2.2.6 Semantic Reasoner

A semantic reasoner, a reasoning engine, a rules engine, or simply a reasoner is a piece of software able to infer logical consequences from a set of asserted facts/axioms. The capabilities of a reasoner depend on the axioms and inference rules that it knows about, which are related to a particular kind of logic.

Reasoner is a key component for working with OWL ontologies. All querying of an OWL ontology should be done using a reasoner. This is because knowledge in an ontology might not be explicit and a reasoner is required to deduce implicit knowledge so that the correct query results are obtained.

OWL reasoner such as Pellet, FaCT++ and HerMiT would be required for executing SWRL rules and infer new ontology axioms. Pellet has more direct functionality for working with OWL and SWRL rules, it allows to define custom SWRL built-ins.



When applying Pellet to reason over ontology with SWRL rules, it takes these rules into consideration, and returns conclusions based on them.

## **2.3 Related Works**

This section presents several related works that use semantic web techniques in similar situations and domains. They focus on disease diagnosis using ontology and SWRL rules, ontology and clinical decision support system.

### **2.3.1 Ontology and Disease Diagnosis System**

There are various works which have explored heart diseases and related diagnosis systems using different approaches. However, this section focuses only on the approaches that use ontology and SWRL rules.

(Al-Hamadani & Alwan, 2015) presents an expert system to diagnose coronary artery diseases. The design of the system depends on ontology knowledge about the patient's symptoms to build the knowledge base. SWRL rules are used to deduce the suitable medication and the required operation for the patient. The architecture of this system consists of several modules: knowledge base module which consists of the fact base and the rule base, the inference engine module and explanation module. The facts are extracted using the user interface, from the user as the patient's symptoms. The inference engine depends on the facts and the rules to reason the required decision. The final decision results will be introduced to the user through the user interface alongside with the explanation about this decision inferred from the explanation module. This system was tested by several general practitioners using 16 instances to test the validation and the evaluation of the system. Recall and precision factors were calculated 0.83 and 0.87.

This work focuses on diagnosing coronary artery diseases which are considered as part of heart diseases. The diagnosis in this system are based only on the patient's symptoms, without taking into account the diagnosis tests results (i.e. ECG, X-ray, CT scan). Our system diagnoses heart diseases based on patient symptoms, risk factors and diagnosis tests results, it focuses on several types of heart diseases.

(Rawte & Roy, 2015) presents ontology based expert system for diagnosis of thyroid disease. This system uses ontology to construct the domain knowledge and rules to infer the thyroid disease related diagnosis. It consists of ontology, reasoner, rule base and MySQL Database. Based on the symptoms entered through the user interface, data and SWRL rules get generated for the users. Then, a reasoner called Jena uses these data and rules to infer and make the diagnosis accordingly. This system was implemented based on neural networks and ontology. Total number of instances taken is 60. Neural network considers only TSH, T3, T4 levels in the blood to train the neural network whereas there is no training for the ontology based expert system. The results show that expert system based on ontology gives more accurate results with lesser complexity than the one which uses neural network.

This system is in line with our research in terms of using ontology to represent knowledge and SWRL rules for disease diagnosis. It is evaluated through its implementation based on neural networks and is compared with ontology-based implementation. Our system is evaluated using a sample set of patients with heart diseases provided by a domain expert.

(Alharbi, Berri, & El-Masri, 2015) presents a diagnosis and treatment recommendation system for diabetes. The implemented system relies on two components: a domain ontology which developed to standardize the domain expertise and a knowledge base system which is includes the necessary rules dedicated to diagnose and propose treatment for the disease. The domain ontology is designed and developed by OWL-DL, the rules are constructed by SWRL and executed by JESS inference engine. The system contains four modules: the graphical user interface, the inference engine, the knowledge base and the ontologies. The user interacts through the graphical user interface to test the system or to request any diagnosis for a specific case. The inference engine is the reasoning component which uses the ontologies and the rules in the knowledge base to infer a diagnosis for the specific case.

(Thirugnanam, 2013) develops an ontology based approach to create disease information system. This system consists of three different components such as knowledge base component, rules component and query processing component.

knowledge base consists of the ontology to model diseases and their relationships with symptoms. Rules component contains the semantic reasoner and SWRL Rules. The query processor checks with SWRL rules for relations between the diseases. It returns the diseases associated with the given symptoms. The query processor then returns the output to the user. The predicted human diseases are done with executing rules which extract disease details with symptoms based on the rules specified, then the inferred axioms are reflected in the ontology.

This approach is similar to our research in terms of building ontology and using SWRL rules to predict diseases using a knowledge base. But it is different in terms of the field of diseases and diagnosis procedures.

(Almutair & El-Masri, 2014) propose general knowledge base ontology framework for patient diagnosis based on clinical practice guidelines (CPG) to help the physicians and medical staff to making the right diagnosis decision for the patient situation. This framework is a general base, which can be used with more specialization for quickly modeling a specific clinical practice guideline. The methodology content on four steps. The first step of this methodology is to choose an appropriate clinical practice guidelines resource as the base of this research. The second step and as the domain of the research, 30 different diseases has been chosen from different human organs and have been visualized in tables based on their symptoms, signs, and diagnosis procedures. The third step is capturing and modeling the common symptoms and signs among these 30 different diseases and with the help of the differential diagnosis that will go out at the end, the patient will be successfully diagnosed. The fourth and the last step in this methodology is the transformation of these models into a knowledge base ontology framework for patient diagnosis based on clinical practice guidelines by using Protégé.

This approach can be considered as a general knowledge base for various diseases, it lacks further information about the characteristics of each disease in terms of laboratory tests and risk factors, particularly that shares the symptoms and diagnostic methods.

(Alfonse, Aref, & Salem, 2014) propose an ontology based diagnostic method for cancer diseases with knowledge management. The proposed method contains three basic modules namely; the diagnostic module, the staging module and the treatment recommendation module. All the three modules interact with a database of cancer ontologies through the query module, which maps from the query of the asking module to the structure of the vocabulary of the ontologies stored in the database of cancer ontologies. The database of cancer ontologies describes the different types of cancer diseases in detail. Each cancer ontology describes the cancer in terms of its structure, signs and symptoms, staging and treatment.

This method can be applied to help patients, medical students and doctors to decide what cancer type the patient has, what is the stage of the cancer and how it can be treated. The system is used to help doctors to decide what cancer type the patient has. Our system diagnoses heart diseases, assesses the heart risk and provides a recommended treatment for a specific case.

From the above works, only one work addressed heart diseases diagnosis using ontology and SWRL rules. This work implements and uses ontology on diagnosing coronary artery diseases which are considered as part of heart diseases. The diagnosis in this system are based only on the patient's symptoms without taking into account the diagnosis test results (i.e. ECG, X-ray, CT scan). Our work focuses on diagnosing all types of coronary artery diseases, valve diseases and cardiomyopathy diseases. In addition, our system diagnoses heart diseases based on patient symptoms, risk factors and diagnosis tests results. Other works that have been mentioned focused on different types of diseases but we focus in this research on heart diseases.

### **2.3.2 Ontology and Clinical Decision Support Systems**

Clinical Decision Support Systems (CDSS) is applications that analyze data to help healthcare providers make clinical decisions. These types of systems require computable medical knowledge, person-specific data, and a reasoning or inferencing mechanism that combines knowledge and data to generate and present helpful information to clinicians (Bau, Chen, & Huang, 2014). This section presents several ontology-based clinical decision support systems.

(Sherimon et al., 2013) presents an intelligent system to predict the risk of hypertension in three main related areas; diabetes, cardiovascular problems, and kidney disorders. The system uses ontologies with knowledge base (medical knowledge base), patient medical profile stored in a semantic way and an inference mechanism to extract data in the decision-making process. Predicting the risk of hypertension is performed through three phases. In the initial phase, the user fills adaptive questionnaire. In the second phase, a semantic profile of the patient is generated automatically by the system based on the input (answers) of the questionnaire. The patient profile generated is semantic in nature and is represented in OWL. In the final phase, this semantic patient profile is analyzed by ontology reasoner with the help of clinical guidelines ontology. The output of the reasoner and the rule engine together generates a risk assessment report of hypertension in three main related cases of diabetes, cardiovascular problems, and kidney disorders.

This work is similar to our system in terms of creating medical patient profile in OWL format and in terms of analyzing patient data using reasoning to predict the risk of hypertension.

(Bau, Chen, & Huang, 2014) presents a clinical decision support system (CDSS) for undergoing surgery based on domain ontology and rules reasoning in the setting of hospitalized diabetic patients. The ontology was created with a modified ontology development method, including specification and conceptualization, formalization, implementation, evaluation and maintenance. Embedded clinical knowledge was elicited to complement the domain ontology with formal concept analysis. The decision rules were translated into JENA format; which JENA used to infer recommendations based on patient clinical situations. The evaluation confirms the correctness of the ontology, acceptance of recommendations, satisfaction with the system, and usefulness of the ontology for blood sugar management of diabetic patients undergoing surgery, especially for domain experts. This work is similar to our system in terms of using decision rules.

This system uses JENA semantic rules to infer recommendations based on patient clinical situations. Our system uses SWRL rules to infer correct diagnosis and recommend appropriate treatment.

(Martínez-Romero et al., 2013) propose an ontology-based decision support system designed to supervise and treat patients affected by acute cardiac disorders. The architecture of the system contains: cardiac intensive care units (CICU) devices to connected to a monitor, communication APIs to enable the system's interaction with the CICU devices, Expert System which has 4 essential components: knowledge base, fact base, inference engine, and explanation facilities, graphical user interface to enables the communication between the doctor and the expert system and database. The system analyzes the patient's condition and provides a recommendation about the treatment that should be administered to achieve the fastest possible recovery. The knowledge base is consisting of an OWL ontology and a set of SWRL rules that represent the expert's knowledge. This approach provides supervision and treatment of critical patients with acute cardiac disorders.

(Abidi et al., 2007) presents an approach to computerize and deploy the CPG in a breast cancer decision support system (BCF-DSS) for use by family physicians in a primary care setting. This approach using semantic web approach to model the CPG knowledge and employ a logic-based proof engine to execute the CPG in order to infer patient-specific recommendations. this work takes a semantic web approach to model the CPG knowledge and to reason over the ontology to provide 'trusted' CPG-driven recommendations. The knowledge base in this system consists of three ontologies: CPG ontology that models the structure of the CPG, breast cancer ontology that represents the medical knowledge encapsulated within the CPG and general breast cancer related concepts; and patient ontology that models the patient's parameters. The ontologies are developed using Protégé and are in OWL format.

This approach uses a logic-based reasoning engine that reasons over the knowledge from these three ontologies. the system allows family physicians to collect patient data and assists them to make CPG mediated decisions, recommendations and referrals for breast cancer survivors. It consists of three stages: computerization of the paper-based

CPG for the care and treatment of breast cancer, development of the ontologies, in particular the breast cancer ontology; and execution of the breast cancer follow-up CPG through a logic-based CPG execution engine. Our work is affected by this work through how the knowledge in clinical practice guidelines is transformed to knowledge base and rule base.

(Sherimon, Vinu, Krishnan, & Saad, 2014) proposes an ontology based system to collect the patient history to assess the patient risk in diabetes due to smoking history, alcohol history, and cardiovascular history. According to the patient history, a total score is calculated for each of the above factors. Based on the score, the ontology performs the risk assessment on a patient profile and predicts the potential risks and complications of the patient. The system instantiates the questionnaire ontology and stores the corresponding answers in it. The system processes this information and automatically generates a patient ontology instance in the server. Patient medical is an OWL file which encapsulates patient details as entered by the patient, nurse and other users in a web/mobile application. The clinical guidelines are hard coded in Java and the values generated are written back to the ontology.

This work is similar to our system in terms of procedures for estimating risk of diabetes. Our system estimates the heart risk according to five factors which are: Age, total cholesterol, HDL, systolic blood pressure and smoking habit.

All clinical decision support systems that have been mentioned in this section focused on the diagnosis of different types of diseases such as diabetes, cardiac disorders, breast cancer and hypertension. These works are very helpful in the development of our proposed ontology. We focus in this research on heart diseases diagnosis and heart risk estimation.

## **2.4 Summary**

This chapter presented required background and related works. It is divided into two sections. In the first section we presented overview of the heart diseases and their risk factors, their diagnosis approaches as well as their related clinical guidelines. The first section also provides overview of semantic web and its associated technologies and

functionalities such as ontology, OWL, SWRL rules, knowledge base and reasoner that used in developing diseases diagnosis systems. In the second section, we studied several related works that use semantic web techniques in similar situations and domains. They focus on disease diagnosis using ontology and SWRL rules, ontology and decision support system for heart disease diagnosing.

In the next chapter, we will discuss the steps of developing heart diseases ontology in order to model the concepts and relationships within the heart diseases domain area, which will form the basis of the system for diagnosing and recommendation treatments for heart diseases.



# **Chapter 3**

## **Heart Diseases Ontology**

## **Chapter 3**

### **Heart Diseases Ontology**

In this chapter, we present the steps to develop the heart disease domain ontology to be used as a basis to provide a proof of concept of how ontology and SWRL rules can improve diagnosing heart diseases based on semantics and reasoning. Section 3.1 presents an overview of HeartOnto, Section 3.2 presents classification of heart diseases domain and Section 3.3 presents the process of HeartOnto development.

#### **3.1 Ontology Overview**

We have developed the ontology contents for heart diseases domain, collected from a number of relevant research papers and documentations of medical domain. In the design of the heart diseases ontology we have started from the terms defined in the ‘Guidelines for Diagnosis and Treatment of the Chronic Heart Failure’ (McMurray et al., 2012), as well as, we held a variety of meetings and interviews with the heart doctors, who were asked to describe in detail the procedures they employ to diagnose and treat patients.

Heart diseases ontology (HeartOnto) is a model of the knowledge from the heart diseases domain. It contains the relevant concepts related to the diagnostics, data of patient (personal information, symptoms, risk factors and clinical tests results) and treatment. The ontology is implemented with Protégé in OWL format. We used a top-down approach to build the ontology, where abstract concepts are identified first, then specialized into more specific concepts. HeartOnto contains 65 well-defined terms (classes) frequently used by experts in the area of heart diseases organized as a taxonomy, 10 object property, 42 datatype properties, 139 Instances and a set of inference rules that guide the diagnosis and risk assessment process.

#### **3.2 Classification of Heart Diseases Domain**

At this stage, we classified the domain knowledge to have a better and easier understanding of heart diseases. We have used the following stages to construct the ontology of heart diseases:

1. Extracting and classifying the main topics (concepts) and subtopics in a heart diseases domain.
2. Extracting the characteristics related to each disease.
3. Building of the hierarchy of concepts.
4. Identifying the relations between concepts and properties.

Each of the above stages is explained next (section 3.2.1 – 3.2.4) with some details.

### 3.2.1 Extraction and classification of the heart diseases

We have classified the content of the heart diseases into the following categories:

1. Types of heart diseases
2. Symptoms of heart diseases
3. Diagnosis of heart diseases
4. Causes and risks of heart diseases
5. Treatment of heart diseases

Each of these categories are explained in what follows:

#### 3.2.1.1 Types of heart diseases

Heart disease refers to various types of conditions that can affect heart function.

These types include:

- Coronary Artery Diseases (أمراض الشريان التاجي)
  - Acute Coronary Syndrome (أمراض الشريان التاجي)
  - Heart Failure (فشل القلب)
  - Unstable Angina (الذبحة الصدرية غير المستقرة)
- Heart valve Diseases (أمراض صمامات القلب)
  - valvular stenosis (تضييق الصمامات)
  - valvular insufficiency (قصور الصمامات)
- Cardiomyopathy Diseases (أمراض اعتلال عضلة القلب)
  - Dilated Cardiomyopathy (تمدد عضلة القلب)
  - Hypertrophic Cardiomyopathy (تضخم عضلة القلب)
  - Restrictive Cardiomyopathy (قصور عضلة القلب)

- Irregular Heart Rhythm (عدم انتظام إيقاع القلب)
- Atrial Fibrillation (الرجفان الأذيني)
- Heart Rhythm Disorders (اضطرابات إيقاع القلب)
- Sudden Cardiac Death (توقف القلب المفاجئ)
- Congenital Heart Disease (مرض قلب خلقي)

### 3.2.1.2 Symptoms of heart diseases

We have listed some symptoms of heart diseases such as:

- Shortness of breath (ضيق في التنفس)
- Palpitations (خفقان)
- Faster heartbeat (تسارع ضربات القلب)
- Extreme Weakness (الضعف الشديد)
- Dizziness (دوخة)
- Pressure (ضغط)
- Heaviness (ثقل)
- Pain in the chest or arm (ألم في الصدر أو الذراع)
- Discomfort radiating to the jaw, or arm (انزعاج "عدم ارتياح" في الفك أو الذراع)
- Fullness (الامتلاء)
- Indigestion (عسر الهضم)
- Sweating (التعرق)
- Nausea (الغثيان)
- Vomiting (التقيؤ)

### 3.2.1.3 Diagnosis of heart diseases

Some standard and simple exam techniques provide with the first clues as to how heart functions and whether patient has heart disease. We list exam techniques of heart diseases such as:

- Checking heart rate (فحص معدل ضربات القلب)
- Checking heartbeat (فحص نبضات القلب)
- Checking blood pressure (قياس ضغط الدم)
- Checking heart by a physical exam (الفحص السريري للقلب)

- Checking heart through blood tests (فحص القلب من خلال فحص الدم)

Heart disease exam techniques are insufficient to decide whether or not the patient has a heart disease. There are several ways to diagnose heart diseases. We list heart disease tests such as:

- Chest X-Ray (تصوير القلب بالأشعة السينية)
- Stress Test (اختبار الإجهاد)
- Echocardiogram (ECG) (مخطط صدى القلب)
- CT Heart Scan (تصوير القلب بالأشعة المقطعية)
- Heart MRI (تصوير القلب بالرنين المغناطيسي)
- Myocardial perfusion imaging test (MPI) (تصوير عضلة القلب)

#### 3.2.1.4 Causes and risks of heart diseases

The risk of heart diseases increases around the age of 55 in women and 45 in men. The risk may be greater if the person has close family members who have a history of heart disease. Other risk factors for heart disease include:

- Obesity
- Diabetes
- High cholesterol and blood pressure
- Family history of heart disease
- Physically inactive
- Smoking

#### 3.2.1.5 Treatment of heart diseases

There are many types and combinations of drugs used to treat heart diseases. We list some diseases and their treatment as shown in Table 3.1

**Table (3.1):** Types of heart diseases and their treatments

| Disease                 | Treatment   |
|-------------------------|---|
| Coronary Artery Disease | Daily aspirin, ACE inhibitors, beta-blockers          |
| Heart Failure           | Diuretics, water pills, beta-blockers, ACE inhibitors |

| Disease                | Treatment                           |
|------------------------|-------------------------------------|
| Heart Valve Disease    | Defibrillators, beta-blockers       |
| Cardiomyopathy Disease | L-carnitine, coenzyme Q10, garlic   |
| Heart Arrhythmias      | Warfarin, dabigatran, beta-blockers |

### 3.2.2 Extracting the characteristics related to each disease

In this step, we have listed the characteristics according to heart disease type. Figure 3.1 shows the characteristics of some of heart diseases, their respective priorities in the diagnostic process and how each one is diagnosed. For example, in Arrhythmia disease (عدم انتظام ضربات القلب), the age factor is the most important characteristic, as most patients diagnosed with this type are 60 years old or more. Therefore, the unique feature of Arrhythmia disease is that patients are within this age range. In the intervention tests, only Electrocardiogram test (ECG) and Myocardial perfusion imaging test (MPI) are determinant for diagnosing Acute Coronary Syndrome (متلازمة الشريان التاجي الحادة). Tables 3.2 shows some of heart diseases and their symptoms, risk factor and tests.

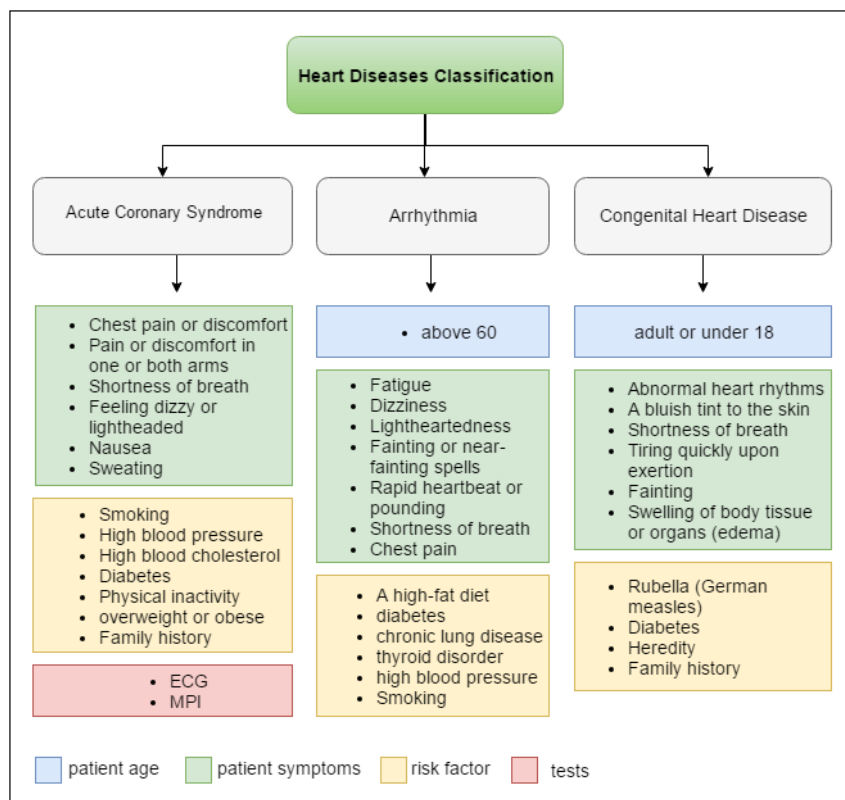


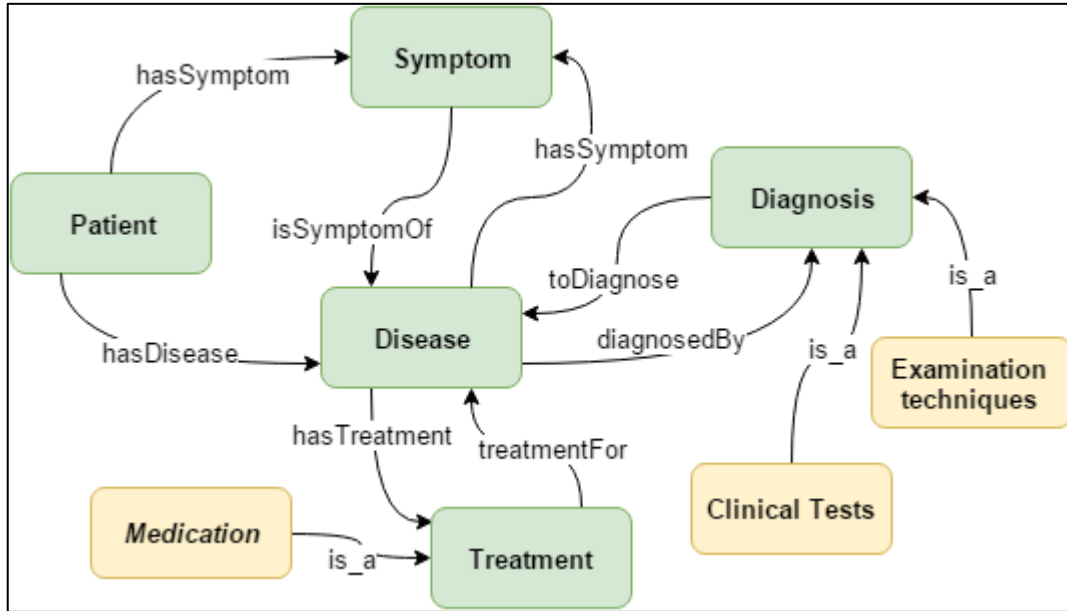
Figure (3.1): Part of heart diseases and their characteristics

**Table (3.2):** Heart diseases and their characteristics.

| Disease                  | Symptoms  | Risk Factor           | Tests                |
|--------------------------|---|-----------------------|----------------------|
| Coronary Artery Disease  | Shortness of breath, Palpitations, Faster heartbeat, Weakness, dizziness, Nausea and Sweating   | - Diabetes            | - EKG                |
| Heart Attack             | Discomfort, Pressure, Heaviness, Fullness, Indigestion, Sweating, Nausea, Vomiting, Dizziness, Extreme weakness and Rapid or irregular heartbeats                                       | - Family History      | - ECG                |
| Arrhythmias              | Palpitations, Dizziness or feeling, Light-headed, Fainting, Shortness of breath, Chest discomfort and Weakness or fatigue (feeling very tired)  | - Smoking             | - Stress test        |
| Atrial Fibrillation      | Heart palpitations (a sudden pounding, fluttering, or racing feeling in the heart), Lack of energy, Dizziness (feeling faint or light-headed), Chest discomfort and Shortness of breath | - High Blood Pressure | - Carotid ultrasound |
| Heart Valve Disease      | Shortness of breath, difficulty catching breath, Weakness or dizziness, Discomfort in chest, Palpitations and Quick weight gain   | - Overweight or Obese | - Chest X-ray        |
| Congenital Heart Defects | Cyanosis (a bluish tint to the skin, fingernails, and lips), Fast breathing and poor feeding, Poor weight gain, Recurrent lung infections, Inability to exercise                        | - Physical inactivity | - Heart MRI          |

### 3.2.3 Building the hierarchy of concepts

Figure 3.2 illustrates the core classes of the heart disease ontology and the relationships between them.



**Figure (3.2):** Core classes of heart diseases ontology and their relationships

### 3.2.4 Identifying the relations between concepts and properties

Based on the heart diseases domain and as shown in Figure 3.2, we can identify the following relations between classes and properties.

- Coronary Artery *diagnosedBy* ECG test.
- Enlarged Heart *diagnosedBy* MPI test.
- Electrophysiology Test *toDiagnose* Atrial Fibrillation.
- Congenital Heart *toDiagnose* Heart MRI.
- Heart Attack *hasSymptom* shortness of breath.
- Sweating *isSymptomOf* Heart Valve.
- Coronary Artery *hasTreatment* beta blockers.
- Nitroglycerin *treatmentFor* Irregular Heart Rhythm.
- Patient *hasDiseases* Sudden Cardiac Death.

After we have presented a classification of the heart diseases domain knowledge as essential part of the heart disease ontology and presented the four stages of constructing the ontology. Next, based on the classification stage, we present the ontology development emphasizing the steps presented in Section 2.2.2 for ontology



engineering methodology. We also emphasize its representation in OWL and the use of Protégé OWL in the development.

### **3.3 Ontology Development**

Heart diseases ontology is a model of the knowledge from the medical domain. It contains all of the relevant concepts related to the diagnostics, treatments, clinical procedures and patient data. It is designed in a way that allows knowledge inference and reasoning.

Based on heart diseases classification presented in section 3.2, we present the development of HeartOnto following a common ontology engineering method (Noy & McGuinness, 2001).

There is no single method on how to construct a medical ontology or any other type of ontology. An ontology can be constructed manually or (semi)automatically. Manual extraction has been done for the heart diseases ontology, domain experts are consulted to explain the meaning of domain-specific concepts and relationships.

There are many different tools available for developing ontologies. We use Protégé which is one of the most widely used ontology development editors that help to define ontology concepts (classes), properties, taxonomies, various restrictions and class instances. The process of ontology construction can be divided into the following steps:

#### **Step 1: Determine the Domain and Scope of the Ontology**

The definition of ontology domain and scope is considered the first step in ontology development. The ontology is developed to answer some basic questions:

1. What is the domain that the ontology will cover?

The domain of the ontology is heart diseases.

2. What is the use of the ontology?

The ontology will be used to provide a model-base for heart diseases domain, which is used in diagnosing heart diseases.

3. What types of questions the information in the ontology should provide answers?

The ontology would provide answers to questions related to heart diseases domain such as:

- What is the recommended diagnosis for a particular patient?
- What is the risk assessment of heart diseases for a particular patient?
- What is the recommended treatment for a particular patient?
- What are the symptoms of a particular disease?
- Who are the patients with a particular disease?
- What is the appropriate treatment for a particular disease?
- What are the symptoms of a particular patient?
- Who are the patients who have a high risk of heart disease?

4. Who will use the ontology?

The ontology will be used by doctors and medical students to access the appropriate diagnosis for a patient.

## **Step 2: Reuse Existing Ontologies**

With the enormous applications of the semantic web, ontologies are becoming more widely available. There is no single standard way to develop ontology. It is not necessary to start from scratch always. We use Heart Failure Ontology (Jović, Gamberger, & Krstajić, 2011) as a basis for developing HeartOnto ontology. We have taken advantage of the ontology structure and relationships between classes and used it to determine the method of building ontology, as well as to identify some relationships and properties in the diagnosis of heart diseases.

### Step 3: Define Classes and Class Hierarchy of the Ontology

There are many approaches for developing a class hierarchy, we name a three of them. The first is top-down approach, which starts with creating definition of the most general concepts and then their specializations (creating subclasses); the process is recursive for every class until we reach the most specific definitions. The second is the bottom-up approach, which goes the other way first we define the most specific concepts and then group them into more general concepts by creating common superclass for them. Combination of both when we start with few general (or 'top-level') concepts and few specific (or 'bottom-level') concepts and fill the middle levels consequently

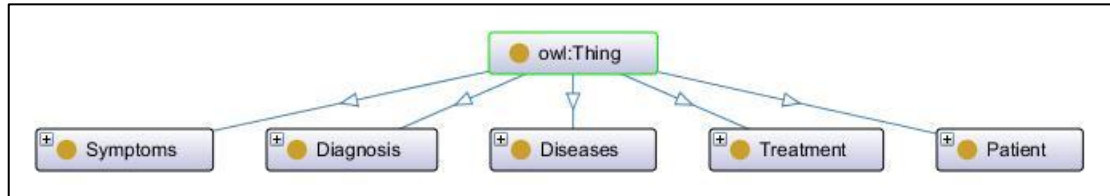
This step starts by defining classes which are selected in earlier stages as mentioned in the above procedure. We used a bottom-up approach to develop the class hierarchy in the ontology, through representing the core concepts (main classes) and subclasses as classes in the ontology.

Classes (concepts) have direct relation with user requirements used in diagnosing heart diseases. We define five main classes: Disease class, Symptom class, Diagnosis class, Treatment class and Patient class. Table 3.3 shows descriptions of these classes (Noy & McGuinness, 2001).

**Table (3.3):** Classes of heart disease ontology

| No. | Class      | Description  |
|-----|------------|--|
| 1   | Disease    | It is a medical condition associated with specific symptoms and signs.               |
| 2   | Symptom    | It is a departure from normal function or feeling which is noticed by a patient.     |
| 3   | Diagnosing | It is the identification of the nature and cause of a certain phenomenon.            |
| 4   | Treatment  | It is the end of a medical condition; the procedure that ends the medical condition. |
| 5   | Patient    | A person who is under medical care or treatment.                                     |

Figure 3.3 illustrates heart disease ontology classes in Protégé visual interface.

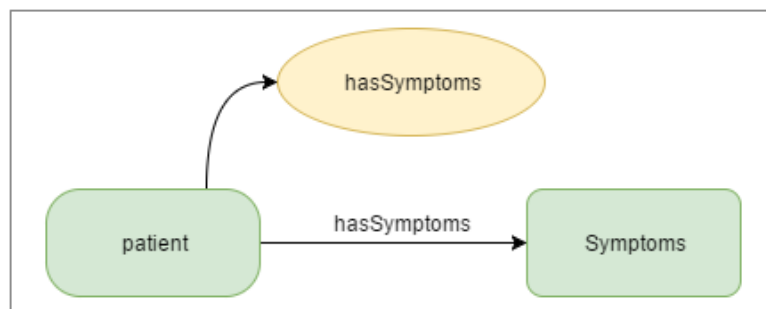


**Figure (3.3):** Heart disease ontology classes in Protégé

#### Step 4: Define the Properties of Classes (Slots)

Defining object properties (relations) among classes is a requirement to come up with the ontology. Classes alone will not provide enough information. Creating object properties plays important role in connecting classes (concepts) of the ontology.

There is a close relation between the properties of classes and relationships between classes. For example, the property of the class *Patients* is *hasSymptoms*. It is also a relation from the class of *patients* to the class of *symptoms* as shown in Figure 3.4. Other examples of implicit relations are *is-a* and *part-of* relations. They always exist between classes and their subclasses.



**Figure (3.4):** Classes and object properties

Ontology classes in a hierarchy inherit properties from their upper classes. For example, if the subclass *Drugs* has properties *specificDiagnosis* and *maxDailyDose*, then *Anti thrombotic Medication* will also have these properties and so will its individual *Aspirin*. Class *Treatment* does not have property *maxDailyDose*, but it does have property *specificDiagnosis*. However, the subclass *Anti thrombotic Medication* might also have a property *targetDose* which subclass *Medication* does not have.

Individuals such as *Aspirin* cannot have additional properties other than those provided by their class.

There are two kinds of properties: object properties and datatype properties. Object properties point at a class of the ontology. The datatype properties do not point at any class, but they have a specific value (a String, an Integer, a Boolean, etc.). The object properties we defined are illustrated in Table 3.4

**Table (3.4):** Object properties in heart disease ontology

| No. | Object property      | Domain     | Range           |
|-----|----------------------|------------|-----------------|
| 1   | toDiagnose           | Diagnosing | Diseases        |
| 2   | diagnosedBy          | Diseases   | Diagnosing      |
| 3   | treatmentFor         | Treatment  | Diseases        |
| 4   | hasTreatment         | Diseases   | Treatment       |
| 5   | hasDisease           | Patient    | Diseases        |
| 6   | hasSymptom           | Diseases   | Symptoms        |
| 7   | isSymptomOf          | Symptoms   | Diseases        |
| 8   | specificDiagnosis    | Patient    | Diseases        |
| 9   | RelatedToHD          | Diseases   | RelatedDiseases |
| 10  | recommendedDiagnosis | Patient    | Diseases        |
| 11  | recommendedTreatment | Patient    | Treatment       |

Data properties describe the relationships between instances (individuals) and data values. Data properties such as *bloodType*, *hasDiabetes*, *isSmoker* and *hasTotalCholesterol* are added to the ontology to link instances and classes and give values to instances in classes. Table 3.5 illustrates the data properties of heart disease ontology. These properties are explained as follows:

- *familyHistory*: refers to health information about a patient and his close relatives. Family health history is one of the most important risk factors for health problems like heart disease, stroke, diabetes, and cancer.

- *hasBloodPressureMedication*: refers to whether the patient takes high blood pressure drugs.
- *hasDiabetes*: refers to whether the patient has diabetes disease.
- *hasDiastolicBloodPressure* and *hasSystolicBloodPressure*: refers to the measurements of systolic and diastolic blood pressure of the patient.
- *bloodType*: refers to the blood type of the patient.
- *isSmoker*: refers to whether the patient is smoker.
- *maxDailyDose*: refers to the amount of daily medicine needed for the patient.
- *hasTotalCholesterol*: measure of the total amount of cholesterol in the patient blood.
- *hasHDLCholesterol*: measure of the total amount of good cholesterol in the patient blood.
- *hasLDLCholesterol*: measure of the total amount of bad cholesterol in the patient blood.
- *hasHeartRate*: refers to the number of the patient heartbeats per unit of time.
- *doExercise*: refers to whether the patient exercises physical activity.
- *hasBloodPressureScore*, *hasCholesterolScore*, *hasHDLLevelScore*, *hasSmokingStatusScore* and *hasAgeScore*: These data properties store the values which are used in calculation of the risk estimation of the heart disease.

**Table (3.5):** Data properties in HeartOnto

| No. | Data property              | Domain              | Range   |
|-----|----------------------------|---------------------|---------|
| 1   | has_name                   | Diseases, Patient   | String  |
| 2   | has_description            | Diseases, Treatment | String  |
| 3   | hasAge                     | Patient             | Integer |
| 4   | familyHistory              | Patient             | Boolean |
| 5   | hasBloodPressureMedication | Patient             | Integer |
| 6   | hasDiabetes                | Patient             | Boolean |
| 7   | hasSympValue               | Symptoms            | String  |
| 8   | hasDiastolicBloodPressure  | Patient             | Integer |
| 9   | hasSystolicBloodPressure   | Patient             | Integer |

| No. | Data property         | Domain  | Range   |
|-----|-----------------------|---------|---------|
| 10  | bloodType             | Patient | String  |
| 11  | isSmoker              | Patient | Boolean |
| 12  | maxDailyDose          | Drugs   | Integer |
| 13  | hasTotalCholesterol   | Patient | Integer |
| 14  | hasHDLCholesterol     | Patient | Integer |
| 15  | hasLDLCholesterol     | Patient | Integer |
| 16  | hasHeartRate          | Patient | Integer |
| 17  | doExcercise           | Patient | Boolean |
| 18  | bodyMassIndex         | Patient | Integer |
| 19  | hasBloodPressure      | Patient | Boolean |
| 20  | hasBloodPressureScore | Patient | Integer |
| 21  | hasCholesterolScore   | Patient | Integer |
| 22  | hasHDLLevelScore      | Patient | Integer |
| 23  | hasSmokingStatusScore | Patient | Integer |
| 24  | hasAgeScore           | Patient | Integer |

### Step 5: Define Facets of the Slots

Slots have different facets that describe the value type, allowed values, the number of the values cardinalities, and other features of the values the slot can take. In our ontology, most of the slot values are integer and boolean. For example, the value type of data property *hasName* for *Patient* class is string and the number of the values cardinality has exactly cardinality of 1: which means that each patient has exactly one name. And value type of *hasTotalCholesterol* is integer. The Figure 3.5 illustrates the data restriction for has name property.

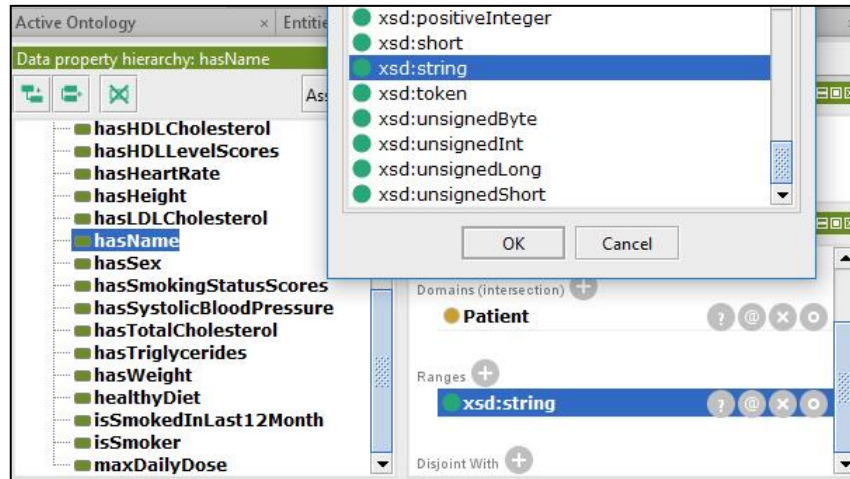


Figure (3.5): Data restriction of hasName data property

Figure 3.6 shows restrictions on the cardinalities of object property *hasSymptom*, data property *hasName* and data property *maxDailyDose*.

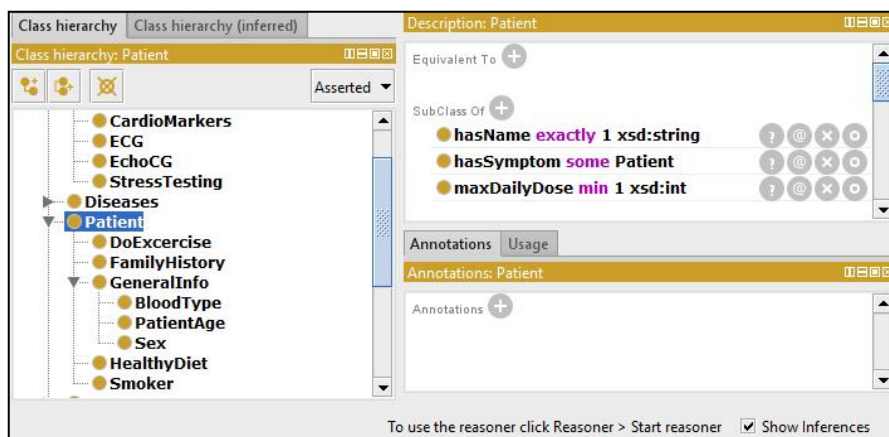
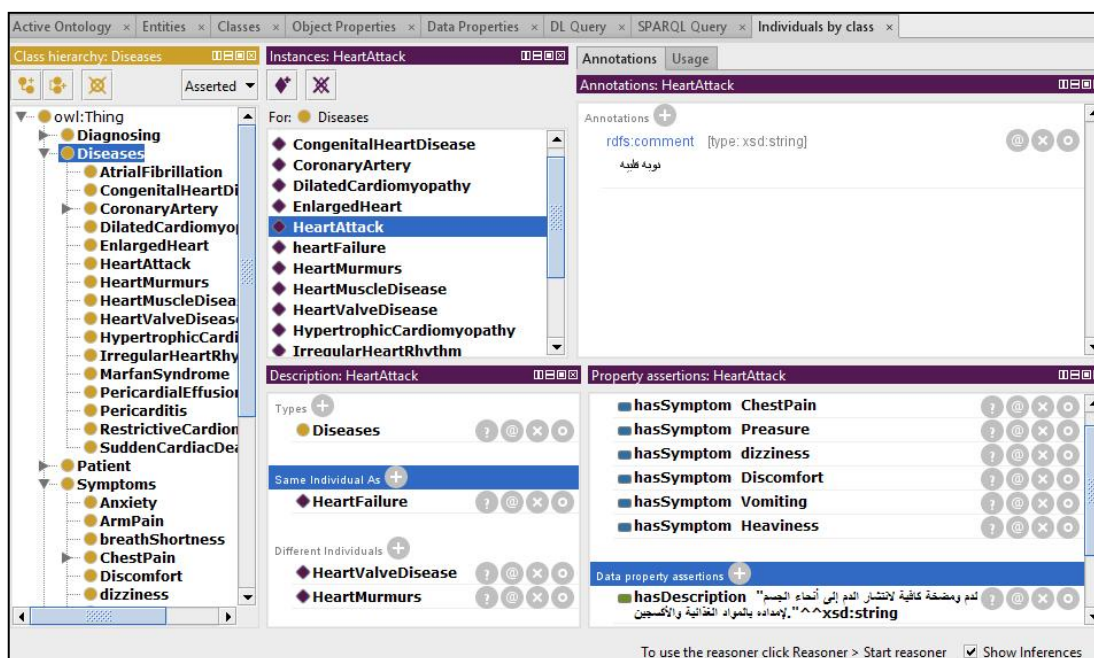


Figure (3.6): Data cardinality

### Step 6: Create Instances

OWL classes are interpreted as sets that contain individuals. Individuals are also known as instances and can be referred to as being ‘instances of classes’. Individuals represent objects in the domain in which we are interested. The information of individuals is taken from a number of relevant research papers and documentations of heart diseases domain.





**Figure (3.7):** Instances of the HeartOnto

Figure 3.7 shows part of the instances used in the heart disease ontology domain. For example, the class *Disease* has instances ‘*Heart Attack*’, description ‘*Heart failure occurs when your heart muscle doesn't pump blood as well as it should*’ and their symptoms ‘*ChestPain, Pressure, Dizziness, Discomfort, Vomiting, Heaviness*’ Also, the class *Treatment* has instances for its name ‘*Aspirin*’ and description ‘*white drug used to relieve pain and fever*’.

### Step 7: Create SWRL Rules

SWRL is used to add rules to OWL to provide an additional layer of expressivity. SWRL allows users to write rules that can be expressed in terms of OWL concepts and that can reason about OWL individuals. We write SWRL rules from valid relationships between HeartOnto concepts to detect heart diseases and heart risk estimation. The rules used to infer new knowledge from existing ontology knowledge bases and user input. All rules are expressed in terms of ontology concepts (classes, properties and individuals). We write around 75 SWRL rules under two HeartOnto concepts ‘Heart Disease Diagnosis’ and ‘Heart Risk Estimation’. More details and examples about writing SWRL rules are illustrated in Chapter 4.

### **Step 8: Apply Ontology Reasoner**

The reasoner play a vital role in developing and using an ontology written in OWL. After creating instances and SWRL rules, we apply an ontology reasoner (e.g. Pellet reasoner) on the HeartOnto to check the consistency of the relationship among the classes and their properties. The reasoner is used also to identify new relations from existing ones. The reasoner is able to identify the different types of ontological relations such as transitive, symmetric, inverse and functional properties and use them to add new facts. Note that the reasoner in this step is used to check the consistency of HeartOnto, later we will use the reasoner to infer heart diseases diagnosis using SWRL rules and patient data.

### **3.4 Summary**

In this chapter, we have explained the development of the HeartOnto. We have presented the steps of developing the ontology. The procedure used to construct the HeartOnto included extracting and classifying the main topics and subtopics in the heart diseases domain, extracting the characteristics related to each disease, building of the hierarchy of concepts, and identifying the relations between concepts and properties. Finally, the ontology was designed depending on heart diseases clinical guidelines and expert consultation. The HeartOnto domain was constructed using Protégé and OWL.

In the next chapter, we will discuss the steps of writing the SWRL rules which are used to infer diagnosis from the domain ontology. Also, we will present the steps of formulating the diagnostic rules which are defined by the experts to SWRL format. Finally, we well explained different uses of SWRL predicates.

# Chapter 4

## Writing SWRL Rules

## Chapter 4

### Writing SWRL Rules

SWRL rules are created from valid relationships between ontology concepts to detect and estimate the risk of heart disease. The rules are used to infer new knowledge from existing ontology knowledge and user input. All rules will be expressed in terms of ontology concepts (classes, properties, individuals) presented in Chapter 3.

The writing of the semantic rules starts with the concept in which the property belongs, and then chains the concept to other facts in a step-by-step manner until the objective is achieved. Each step is expressed as an atom and the rule is expressed in the form of “(atom\_1^...^ atom\_n) → consequence” to express the cause-effect relationship (Lezcano, Sicilia, & Rodríguez-Solano, 2011). OWL 2 language is not able to express all relations, for example, there is no way in OWL 2 to express the relation between individuals with which an individual has relations, the expressivity of OWL 2 can be extended by adding SWRL rules to our ontology. The semantic web rule language (SWRL) is used to express the rules as explained in Section 2.2.4.

This chapter presents SWRL rules which are used to detect and estimate the risk of heart disease. We present details about writing SWRL rule for heart disease diagnosis and heart disease risk estimation. Subsequently, SWRL rules predicates.

#### 4.1 Rules for Diagnosis Heart Diseases

The diagnostic rules are specialized for each type of heart diseases according to the existing cardiac guidelines followed in Palestine and the expert physician. For example, Table 4.1 shows the risk factors needed to determine the likelihood of Acute Coronary Syndrome diseases (ASC), which are considered as part of coronary artery diseases (Achar, Kundu, & Norcross, 2005). The likelihood of ACS diseases either high, intermediate or low. The assessment requires a focused history, risk factor, a physical examination, an electrocardiogram (ECG), frequently and cardiac marker determinations.

**Table (4.1):** Risk factors to determine the likelihood of Acute Coronary Syndrome diseases.

| Assessment           | High likelihood of ACS   | Intermediate likelihood of ACS   | Low likelihood of ACS  |
|----------------------|--|--|--|
| History              | <ul style="list-style-type: none"> <li>- Chest or left arm pain or discomfort as chief symptom</li> <li>- Reproduction of previously documented angina</li> <li>- Known history of coronary artery disease, including myocardial infarction</li> </ul> | <ul style="list-style-type: none"> <li>- Chest or left arm pain or discomfort as chief symptom</li> <li>- Age &gt; 50 years</li> </ul> | <ul style="list-style-type: none"> <li>- Probable ischemic symptoms</li> </ul>   |
| Physical examination | <ul style="list-style-type: none"> <li>- hypotension</li> <li>- diaphoresis</li> </ul>   | Extracardiac vascular disease  | Chest discomfort reproduced by palpation   |
| ECG                  | ST-segment deviation (> 0.05 mV) or T-wave inversion (> 0.2 mV) with symptoms  | Fixed Q waves<br>Abnormal ST segments or T waves not documented to be new  | <ul style="list-style-type: none"> <li>- T-wave flattening or inversion of T waves in leads with dominant R waves</li> <li>- Normal ECG</li> </ul> |
| Cardiac Markers      | Elevated cardiac troponin T or I, or elevated CK-MB  | Normal   | Normal   |

We formulate diagnostic rules in SWRL format according to the rules followed by experts to determine the likelihood of ACS. In order to illustrate how rules are codified, the following rule is resulting from Table 4.1 to diagnose the high likelihood of ACS:

*IF*        *patient has any symptom or any sign of ACS disease*  
*AND*      *patient has history of Angina*  
*AND*      *patient has Hypotension*  
*AND*      *patient has Diaphoresis*  
*AND*      *patient has ECG test with ST-segment (> 0.05) and T-wave (> 0.2)*  
*AND*      *patient has cardiac troponin test T or I*  
*THEN*     *recommend diagnostic: Acute Coronary Syndrome*  
*AND*      *likelihood of Acute Coronary Syndrome: high*

This rule is codified in predicate logic presented in Section 2.2.1 as shown in the following rule. P, T and S are variables where P is universally quantified, T and S are existentially quantified. This rule example concludes on “diagnosing the high likelihood of ACS disease” based on assessment risk factors.

$\forall P, \exists T, S$   
*Patient (P), hasT-WaveValue(T), hasSystolicBloodPressure(S)*  
*hasPatientSymptoms (P, “Acute Coronary Syndrome”)*  
 $\wedge$     *familyDiseaseHistory (P, “Angina”)*  
 $\wedge$     *hasSystolicBloodPressure (P, S)*  
 $\wedge$     *hasPatientSymptoms (P, “diaphoresis”)*  
 $\wedge$     *hasST-SegmentValue (P, T)*  
 $\wedge$     *hasT-WaveValue (P, W)*  
 $\wedge$     *value(S) <= 70*  
 $\wedge$     *value(T) < 0.05*  
 $\wedge$     *value(W) < 0.02*  
 $\rightarrow$     *diagnosticPatient (P, “Acute Coronary Syndrome”)*  
 $\wedge$     *likelihood (P, “high”)*

The above rule codified in predicate logic is easily converted to SWRL based on the syntax presented in Section 2.2.5 as explained in the following rule:

*Patient(?p)*  
*A hasPatientSymptoms (?p, ?s)*  
*A hasSympValue (?s, "AcuteCoronarySyndrome")*  
*A familyDiseaseHistory (?p, "Angina")*  
*A hasPatientSymptoms (?p, "Diaphoresis")*  
*A hasSystolicBloodPressure (?p, ?spd)*  
*A hasST-SegmentValue (?p, ?t)*  
*A hasT-WaveValue (?p, ?w)*  
*A swrlb:lessThanOrEqual(?spd, 70)*  
*A swrlb:lessThan (?t, 0.05)*  
*A swrlb:lessThan (?w, 0.02)*  
 → *recommendedDiagnosisResult (?p, "Acute Coronary Syndrome")*  
*A likelihood (P, "high")*

We developed 25 rules to diagnose all types of coronary artery diseases, valve diseases and cardiomyopathy diseases (see Appendix A for the full list of SWRL rules). Next, we present the set of rules written for heart diseases risk estimation.

#### 4.2 Rules for Heart Diseases Risk Estimation

We are describing the estimation of cardiac score and risk according to Framingham Heart Study, a risk assessment tool to predict a person's chance of having a heart attack in the next 10 years (Sherimon et al., 2014). According to Framingham Heart Study, the heart disease risk is calculated based on factors such as Gender, Age, Total Cholesterol, HDL, Systolic Blood Pressure, Smoking habit, and Hypertension. Table 4.2 represents the heart disease risk points associated with each factor. The actual risk estimation is obtained by calculating the points of risk factor according to Table 4.3.

**Table (4.2):** Heart Disease Risk Points.

| Points | Age, years | HDL       | Total Cholesterol | SBP (Not Treated) | SBP (Treated) | Smoker | Diabetic |
|--------|------------|-----------|-------------------|-------------------|---------------|--------|----------|
| -3     |            |           |                   | <120              |               |        |          |
| -2     |            | >1.56     |                   |                   |               |        |          |
| -1     |            | 1.30–1.55 |                   |                   | <120          |        |          |
| 0      | 30–34      | 1.17–1.29 | <4.14             | 120–129           |               | No     | No       |

| Points | Age, years | HDL      | Total Cholesterol | SBP (Not Treated) | SBP (Treated) | Smoker | Diabetic |
|--------|------------|----------|-------------------|-------------------|---------------|--------|----------|
| 1      |            | 0.9–1.16 | 4.14–5.15         | 130–139           |               |        |          |
| 2      | 35–39      | <0.9     |                   |                   | 120–129       |        |          |
| 3      |            |          | 5.16–6.19         | 140–149           | 130–139       | Yes    |          |
| 4      | 40–44      |          | 6.20–7.24         | 150–159           |               |        | Yes      |
| 5      | 45–49      |          | >7.25             | 160+              | 140–149       |        |          |
| 6      |            |          |                   |                   | 150–159       |        |          |
| 7      | 50–54      |          |                   |                   | 160+          |        |          |
| 8      | 55–59      |          |                   |                   |               |        |          |
| 9      | 60–64      |          |                   |                   |               |        |          |
| 10     | 65–69      |          |                   |                   |               |        |          |
| 11     | 70–74      |          |                   |                   |               |        |          |
| 12     | 75+        |          |                   |                   |               |        |          |

**Table (4.3):** Heart Disease Risk.

| Points | <0 | 0  | 1  | 2   | 3   | 4   | 5   | 6   | 7   | 8     |
|--------|----|----|----|-----|-----|-----|-----|-----|-----|-------|
| Risk % | 1% | 1% | 1% | 1%  | 1%  | 1%  | 2%  | 2%  | 3%  | 4%    |
| Points | 9  | 10 | 11 | 12  | 13  | 14  | 15  | 16  | 17  | > 17  |
| Risk % | 5% | 6% | 8% | 10% | 12% | 16% | 20% | 25% | 30% | > 30% |

In the data properties detailed in Chapter 3, Table (3.5), we have 7 inferred data properties which are *hasBloodPressureScore*, *hasAgeScore*, *hasCholesterolScore*, *hasHDLLevelScore*, *hasSmokingStatusScore*, *hasBloodPressure* and *BodyMassIndex*. These data properties are designed as sub-tasks in term of require semantic rules to combine related facts for inference. For example, the steps for evaluating patient risk (*hasRiskLevel*) require the steps from obtaining the blood pressure score value, cholesterol score value, smoking status score value, HDL level score value and age score value to acquire proper risk level. The steps can be expressed as SWRL rule as follows:



*Patient(?p)*  
*A hasBloodPressureScore (?p, ?BP)*  
*A xsd:integer[>= 140 , <= 180](? BP)*  
*A hasTotalCholesterolScore (?p, ?TotalCholesterol)*  
*A xsd:integer[>= 40, <= 49](? BP)*  
*A hasAgeScore (?p, ?age)*  
*A xsd:integer[>= 50 , <= 54](?age)*  
*A hasHDLCholesterolScore (?p, ?HDL)*  
*A xsd:integer[< 40](?HDL)*  
 → *hasRiskLevel(?p, 4.5)*

In total, 38 rules are developed to calculate heart diseases risk level, 17 rules to calculate total Cholesterol score, 4 rules to calculate smoking score, 7 rules to calculate Blood Pressure score and 10 rules to calculate Age score (See Appendix A for the full list of SWRL rules).

For example, according to Framingham study, HDL level score for patient with HDL 60 or more is -1, if the HDL between 50 and 59 the HDL level score is 0, if the HDL between 50 and 59 the HDL level score is 0, while if the HDL less than 40 the HDL level score is 2. Table 4.4 illustrates the rules for calculating the HDL level score.

**Table (4.4):** SWRL rules for calculating the HDL level score.

| Rule Number | Rule  | Rule Description   |
|-------------|---|--|
| Rule-2      | Patient(?p),<br>hasHDLCholesterol (?p, ? HDL),<br>xsd:integer[>= 60](? HDL)<br>-> hasHDLLevelScore(?p, -1)        | To set data property <i>hasHDLLevelScore</i> value (-1) when the patient has <i>hasHDLLevelScore</i> value more than 60  |
| Rule-3      | Patient(?p),<br>hasHDLCholesterol (?p, ? HDL),<br>xsd:integer[>= 50 , <= 59](? HDL)<br>-> hasHDLLevelScore(?p, 0) | To set data property <i>hasHDLLevelScore</i> value (0) when the patient <i>hasHDLLevelScore</i> value between 50 and 59. |

| Rule Number | Rule  | Rule Description   |
|-------------|---|--|
| Rule-4      | Patient(?p),<br>hasHDLCholesterol (?p, ? HDL),<br>xsd:integer[>= 40 , <= 49](? HDL)<br>-> hasHDLLevelScore(?p, 1) | To set data property<br><i>hasHDLLevelScore</i> value<br>(1) when the patient has<br><i>hasHDLLevelScore</i> value<br>between 40 and 49. |
| Rule-5      | Patient(?p),<br>hasHDLCholesterol (?p, ? HDL),<br>xsd:integer[<40](? HDL) -><br>hasHDLLevelScors(?p, 2)           | To set data property<br><i>hasHDLLevelScore</i> value<br>(2) when the patient has<br><i>hasHDLLevelScore</i> value<br>less than 40,      |

### 4.3 SWRL predicates

We wrote SWRL rules for other HeartOnto predicates in addition to class and property names, such as class expressions, property expressions, data range restrictions, core SWRL built-ins and custom SWRL built-ins.

#### 4.3.1 SWRL rules with class expressions

We used SWRL rules with class expressions for many cases. For example, to determine if the patient has a disease related to heart disease, we wrote the following rule:

$$\begin{aligned}
 & Patient(?p) \\
 & A \quad (hasRelatedDisease \text{ min } 1 \text{ RelatedDiseases}) (?p) \\
 & \rightarrow RelatedDiseaseHistory (?p, true)
 \end{aligned}$$

This rule indicates that individuals who are in the *Patient* class and have at least one object property *hasRelatedDisease* with an individual from the class *RelatedDiseases*, then the data property *RelatedDiseaseHistory* has Boolean value true.

#### 4.3.2 SWRL rules with data range restrictions

Datatypes, such as *xsd:string*, *xsd:integer* and literals such as "1"^^*xsd:integer*, can be used to express data ranges. We wrote SWRL rules with data range restrictions to

calculate cardio risk score. For example, the following rule is created for calculate total cholesterol score, this example shows a complex data range restriction with facets:

```

Patient(?p)
A hasAge(?p, ?age)
A hasTotalCholesterol(?p, ?TotalCholesterol)
A swrlb:graeterThanOrEqual(?age, 20)
A swrlb:lessThanOrEqual(?age, 39)
A swrlb: graeterThanOrEqual (?TotalCholesterol, 200)
A swrlb:lessThanOrEqual(?TotalCholesterol, 239)
→ hasCholesterolScore (?p, 7)

```

The data range restriction is satisfied when *TotalCholesterol* variable has an integer value between 200 and 239 and *age* variable has an integer value between 20 and 39.

### 4.3.3 SWRL rules with core built-ins

SWRL defines some special predicates which can be used to manipulate data values. These predicates are called SWRL core built-ins. This core set includes basic mathematical operators and built-ins for string and date manipulations. These built-ins can be used directly in SWRL rules. We used SWRL rules with *swrlb:multiply* and *swrlb:divide* math core built-ins to calculate the body mass index. The Body Mass Index (BMI) is calculated based on the following formula:

$$\text{BMI} = x \text{ KG} / (y \text{ M} * y \text{ M})$$

Where x is Body weight in kilograms and y is height in meters. The formula is used in the following rule to infer the BMI:

```

Patient(?p)
A hasHeight(?p, ?h)
A hasWeight(?p, ?w)
A swrlb:multiply (?HSquare, ?h, ?h)
A swrlb:divide(?BMI, ?w, ?HSquare)
→ bodyMassIndex (?p, ?BMI)

```

The rule takes the data value of the *hasHeight* and *hasWeight* property, which must be of *xsd:integer* type, divide value from *hasWeight* (in kilograms) by the square value from *hasHeight* (in meters), then add the *?BMI* part as the value of the *bodyMassIndex* property for the patient.

#### 4.3.4 SWRL rules with custom built-ins

SWRL provides an extension mechanism that allows user-defined methods to be used directly in SWRL rules (O'Connor and Das, 2006). We define our SWRL custom built-ins which are used in risk assessment of heart diseases. This built-ins map total Framingham point score to risk level. We define *heartOnto: RiskLevel* custom built-in to get the percentage of risk, where *heartOnto* prefix refers to our custom built-in which is used in HeartOnto (more implementation issues are detailed in Chapter 5 Section 5.7). The following rule illustrate how we use *RiskLevel* custom built-ins to infer risk level:

```
Patient(?p)
∧ hasBloodPressureScore (?p, ?BPS)
∧ hasTotalCholesterolScore (?p, ?TotalCholesterolS)
∧ hasAgeScore (?p, ?ageS)
∧ hasHDLCholesterolScore (?p, ?HDLS)
∧ swrlb:add(?subTotal,?BPS,?TotalCholesterolS)
∧ swrlb:add(?subTotal2, ?ageS, ?HDLS)
∧ swrlb:add(?TotalScors,?subTotal, ?subTotal2)
∧ swrlb:add(?TotalScors, ?subTotal, ?subTotal2)
∧ heartOnto: RiskLevel (?TotalScors)
→ hasRiskLevel(?p, ?TotalScors)
```

we used the *swrlb:add* core built-in to compute the patient risk score, then uses the *heartOnto:RiskLevel* custom built-in to bind the integer value of the percentage of risk to the *TotalScors* variable, then it sets the result as the value of the *hasRiskLevel* property.

#### 4.4 Summary

This chapter presented the steps of writing the SWRL rules which used to infer diagnosis from the domain ontology, we have presented the steps of formulated the diagnostic rules which defined by the experts to SWRL format. We explained different uses of SWRL predicates. The SWRL predicates used in developed rules included class expressions, data range restrictions, core built-ins and custom built-ins. The SWRL rules was written using Protégé editor. In the next chapter, we present the design of DHDOnto system and explain how the system works using the ontology, the SWRL rules and reasoning

# Chapter 5

## The DHDOnto System

## Chapter 5

### The DHDOnto System

This chapter discusses the design of DHDOnto system. The proposed system is based on HeartOnto and SWRL rules to diagnosis heart diseases. The core of DHDOnto system is the HeartOnto ontology that we built in Chapter 3 to model the different entities included in the domain of heart diseases. We present the architecture of the approach, explain how the system works, and discuss how HeartOnto, SWRL rules and inference engine (reasoner) have the potential to improve heart disease diagnosis.

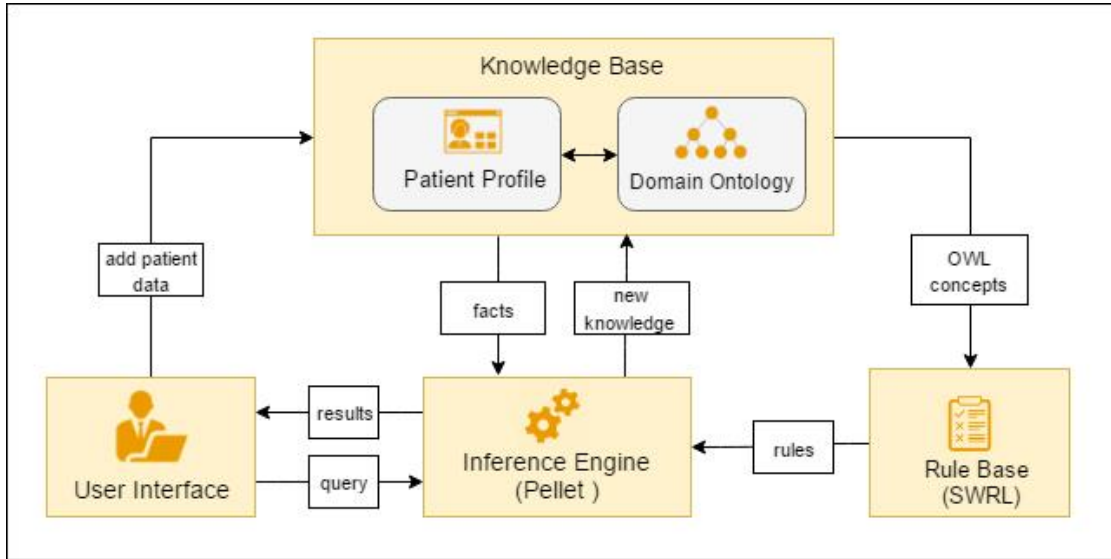
The DHDOnto system can be used by doctors and medical students to diagnosis heart diseases. We divide the system's functionality into four main tasks as follows:

- Heart disease diagnosing: it is performed by checking the input provided by the user to different questions related to the symptoms and risk factors for heart disease. All diagnosing rules and relevant knowledge is extracted from the HeartOnto ontology. Then, the reasoner uses patient data and diagnostic rules to draw conclusion and to give recommended diagnosis for the selected patient.
- Heart disease risk estimation: it is used to predict a person's chance of having a heart disease in the next 10 years. the heart disease risk is calculated based on risk factors which are entered by the user.
- Treatment recommendation: based on the diagnosis results, the treatment recommendation component can recommend a specific treatment for the case being diagnosed.
- Result explanation: permits the user to ask the system about the reasons behind the returned results.

Later in Section 5.3, we present a complete case study to illustrate how these four tasks are performed by the system.

To perform these main tasks, the DHDOnto system is architected to have four main components as depicted in Figure 5.1. The four components: Knowledge Base

component, Rule Base component, Inference Engine component and User Interface component.



**Figure (5.1):** Architecture of DHDOnTo system

The user interacts through the graphical user interface to request any diagnostic for a specific case. The inference engine is the reasoning component which uses the HeartOnto, patient data and the SWRL rules in the Rule Base to infer a diagnostic for the specific case. The Rule Base includes all the rules used to infer diagnosis from the knowledge base which are acquired from the expertise. These rules are mapped with the ontology concepts which allows them to be expressed in SWRL. The HeartOnto feeds the reasoning process with the necessary concepts and their relationships which allows the inference engine to combine rules with concept instances during inferences.

The process of diagnosing heart diseases consists of several steps, the system display the questionnaire to the user who is required to answer questions related to his medical profile. When the user completes filling the questionnaire, the system stores the values entered by user in his own medical profile which is automatically generated. After medical patient profile is created, the inference engine uses the ontology concepts, their relationships and the rules in the knowledge base to draw diagnosis results for the specific case. All inferred information is stored as values for inferred data properties which are used for patient risk analysis and disease diagnosis. The potential risks of a patient are predicted based on some factors by calculating the points

of risk factors according to Framingham Heart Study. The total heart risk points are calculated and then the heart risk ratio is calculated. Finally, The DHDOnto system gives the recommended diagnosis of heart disease, heart disease risk estimation, treatment recommendation and result explanation. All of this is based on patient symptoms, risk factors and clinical test results according to cardiac guidelines followed in Palestine. More details and examples about the process of diagnosing heart diseases are presented in Section 5.5.

The components of the system are explained as follows:

### **5.1 Knowledge Base Component**

Knowledge Base component contains the domain ontology and patient profile. The ontology together with patient profile acts as the knowledge base for the DHDOnto system. The ontology represents concepts in the heart disease domain which is collected from a number of relevant research papers and documentations in the medical domain. The HeartOnto ontology is created using an ontology editor called Protégé in OWL format. It contains the relevant concepts related to the diagnostics, to patient (personal information, symptoms, and risk factors) and to treatment. Patient profile contains patient information such as symptoms, gender, age, total cholesterol, HDL, systolic blood pressure, smoking habit, hypertension, diabetic history, family history, smoking history, physical-activity history, etc. which is stored into ontology by the user through the questionnaire interface as presented in Section 5.4.3.

### **5.2 Rule Base Component**

This component includes a collection of SWRL rules which are used to infer diagnosis from the knowledge base. SWRL rules are developed under the two concepts of ‘Heart Disease Diagnosis’ and ‘Heart Disease Risk Estimation’. All rules are expressed in terms of ontology concepts (classes, properties, individuals).

We have written around 62 SWRL rules (see Chapter 4). These rules are mapped to the domain ontology concepts which allow them to be expressed in SWRL. The ontology feeds the reasoning process with the necessary concepts and their



relationships which allow the inference engine to combine rules with concept instances during inferences. SWRL rules was created using Protégé editor. After the SWRL rules are created, we used a reasoner to test and check for inconsistencies.

### **5.3 Inference Engine Component**

OWL inference engine is a tool which has the capability to infer logical consequences from a set of facts. The inference engine would be required for executing the SWRL rules and infer new ontology axioms. We implemented the Inference Engine component in Java and used the Pellet reasoner. Pellet has more direct functionality for working with OWL and SWRL rules, it allows us to define custom SWRL built-ins. When we apply Pellet to reason over ontology with SWRL rules, it takes these rules into consideration, and returns conclusion based on these rules.

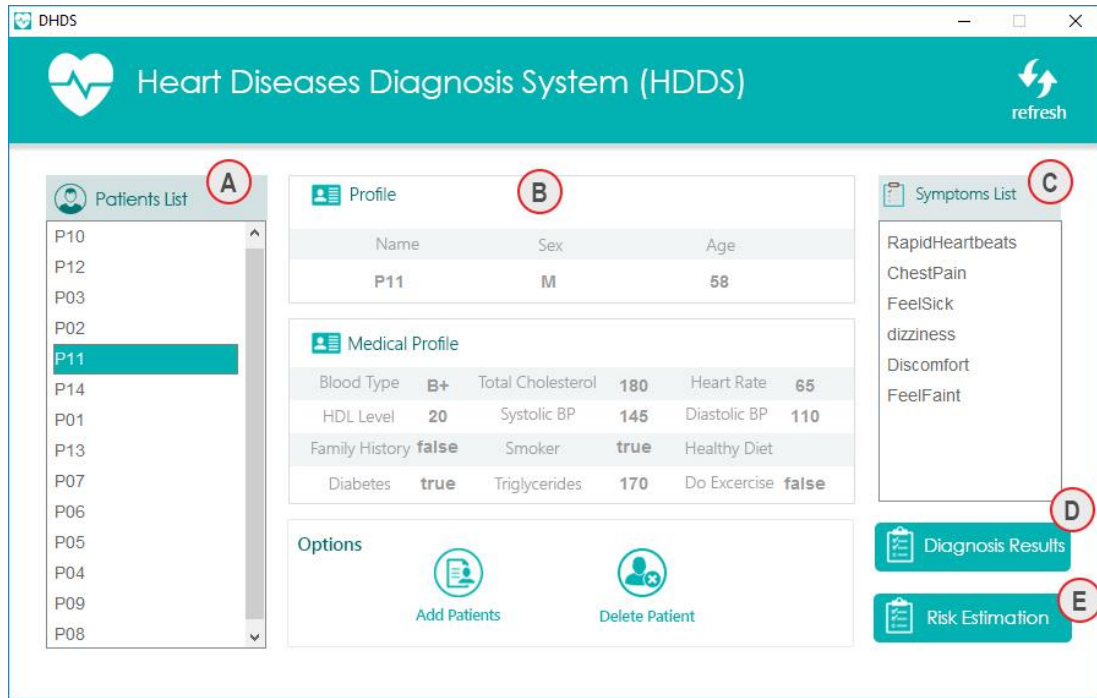
### **5.4 Front End UI Component**

We built DHDOnto system in such a way that the user can easily interact with the system and get the appropriate diagnosis results for specific patient according to the patient symptoms and patient information. We devolved front end Java application using OWL APIs to provide different interfaces for clinicians.

DHDOnto system provide four different interfaces for clinicians which are: patient profile, patient questionnaire, diagnosis results and risk estimation results.

#### **5.4.1 Patient Profile Interface**

Patient profile interface contains patient information such as symptoms, gender, age, total cholesterol, HDL, systolic blood pressure, smoking habit, hypertension, diabetic history, family history, smoking history, physical-activity history. Figure 5.2 shows a snapshot of patient profile interface which contains four parts: list of patients (see Figure 5.2.A), medical information for selected patient (see figure 5.2.B), list of symptoms for selected patient (see Figure 5.2.C), and finally the user interface provides heart disease diagnosis and heart risk estimation services for the selected patient through the clicking on the "Diagnosis Results" button (see Figure 5.2.D) and "Risk Estimation" button (see Figure 5.2.E).



**Figure (5.2):** A snapshot of patient profile interface

### 5.4.2 Diagnosis Result Interface

Diagnosis result interface contains four panels which are: recommended diagnosis which is displays recommended diagnosis based on patient symptoms and clinical test results (see Figure 5.3.A), risk estimation which is display ratio of chance a patient has heart diseases in the next 10 years, calculated based on risk factors which entered by user (see Figure 5.3.B), treatment recommendation which is display recommend a specific treatment for the case being diagnosed (see Figure 5.3.C) and result explanation which is display the reasons behind the returned results (see Figure 5.3.D). Figure 5.3 shows a snapshot of the diagnosis result interface.



**Figure (5.3):** A snapshot of the diagnosis result interface

### 5.4.3 Risk Estimation Result Interface

Risk estimation result interface provide more details about the estimated risk of heart disease in terms of total heart risk ratio and risk points for each factor such as Gender, Age, Total Cholesterol, HDL, Systolic Blood Pressure, Smoking habit, and Hypertension. The risk points are calculated according to Framingham heart study.



Figure (5.4): A snapshot of risk estimation result interface

### 5.4.3 Patient Questionnaire Interface

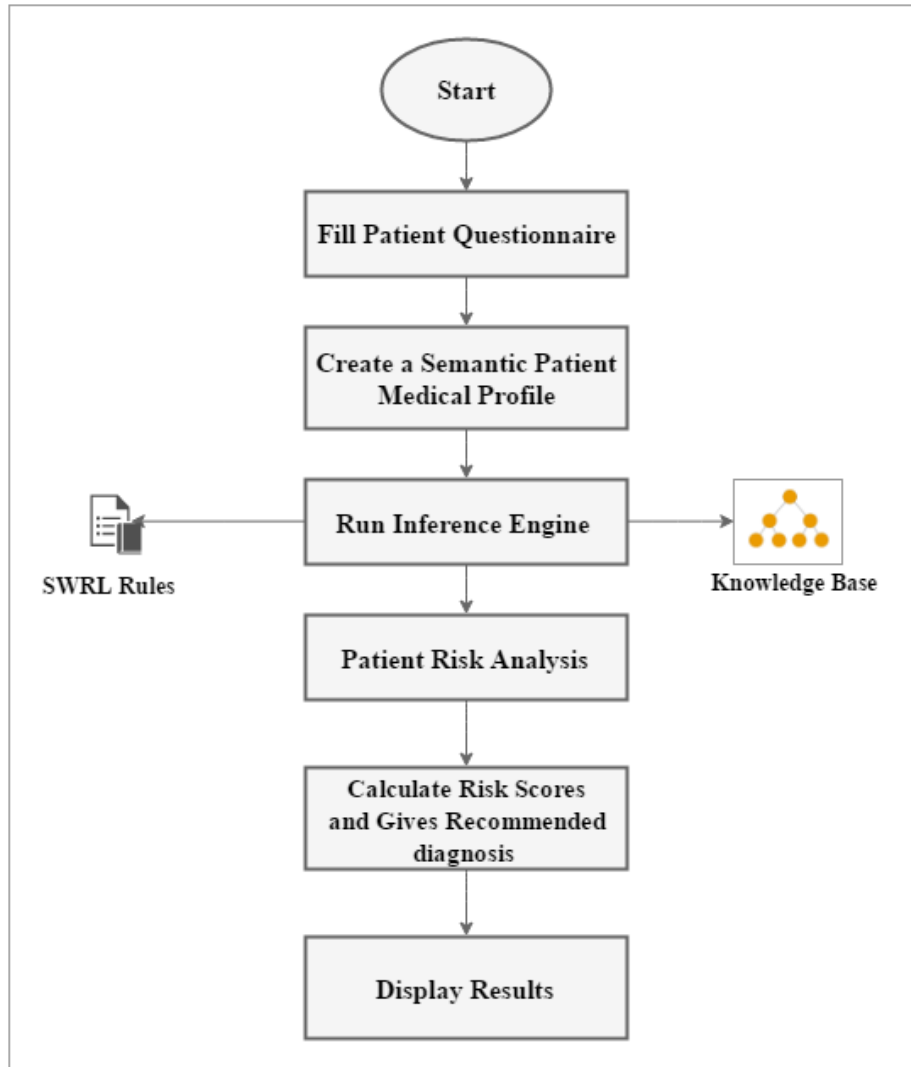
Through questionnaire interface, the patient is required to answer a set of different questions related to the patient factors, which are included as ontological classes, data properties and object properties. Figure 5.5 illustrate part of patient questionnaire interface.



Figure (5.5): Part of patient questionnaire interface

## 5.5 Process of Heart Diseases Diagnosis

Based on the architecture of the system presented on Figure 5.1 and its components presented on Sections 5.1, 5.4, 5.3, and 5.4, we present the process of diagnosis heart diseases. The flow diagram in Figure 5.6 shows the process of heart disease diagnosis.



**Figure (5.6):** The process of heart disease diagnosis

### 5.5.1 Patient Questionnaire

The system displays the questionnaire to the user who is required to answer questions related to personal information, symptoms and signs, diabetic history, family history, smoking history, lab tests, vital information about the patient such as body

temperature, blood pressure, height, etc. For every new patient, the system stores the values entered by the users in his medical profile.

### 5.5.2 Create Patient Medical Profile

Patient medical profile contains all patient medical data which stored in different data properties using answers acquired through questionnaire. Patient medical data contains personal and family medical history, symptoms, risk factors and clinical test results. When the users complete filling the questionnaire, a new instance of *Patient* class is automatically generated with patient name as the instance name. For each instance, the values entered by the users are asserted into different data properties. For example, *has\_Systolic\_Blood\_Pressure* is a data property to represent the systolic blood pressure measurement. The domain of this property is *Patient* class and the range is {60 or more}. If a patient has chosen that the systolic blood pressure to be 130, then the value of the property *has\_Systolic\_Blood\_Pressure* is '130' for that particular instance.

### 5.5.3 Run Inference Engine

After medical user profile is created, the inference engine (i.e. Pellet Reasoner) uses the ontology concepts and their relationships and the rules in the knowledge base to draw diagnosis results for the specific case. After running the inference engine over the ontology and SWRL rules, all inferred information is stored as values for inferred data properties which used for patient risk analysis and disease diagnosis. For example, *bodyMassIndex* is inferred data property used to store patient Body Mass Index which is calculated using SWRL rules based on patient bodyweight and height entered by the user.

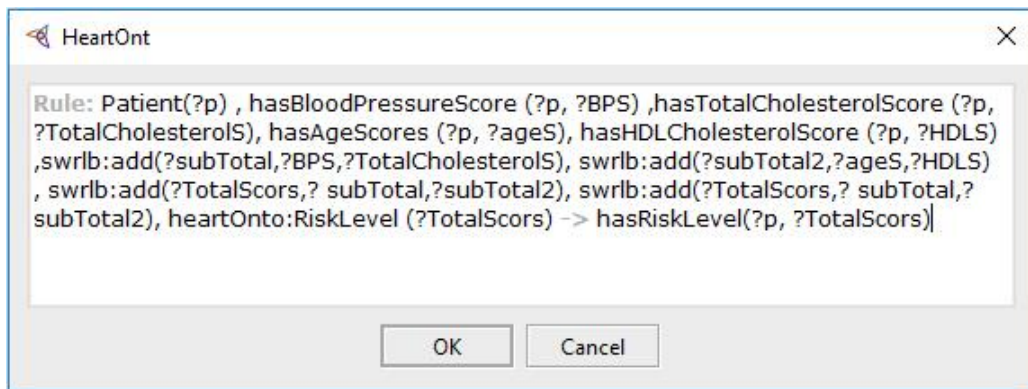
### 5.5.4 Patient Risk Factors Analysis

By performing the risk estimation, the potential risks of a patient are predicted. The heart risk for a patient is assessed according to five factors which are: Age, Total Cholesterol, HDL, Systolic Blood Pressure and Smoking habit. It is done by checking the input answers provided by the user to different questions related to the above factors. A score is associated with each answer choice. The actual risk estimation is

obtained by calculating the points of risk factors according to Framingham Heart Study (see Chapter 4, Table 4.2).

### 5.5.5 Calculate Heart Risk Score

Based on the user input to the questionnaire, the total heart risk points are calculated and then the heart risk ratio is calculated. Data properties are used to store the score of each factor. For example, we have 5 inferred data properties which are *hasBloodPressureScore*, *hasAgeScore*, *hasCholesterolScore*, *hasHDLLevelScore* and *hasSmokingStatusScore*. These properties require SWRL rules to store score values. Every answer to a question carries a particular point. So according to the user input, the points are allocated and the total score is calculated for every factor. We assessed the risk level of heart diseases by using custom built-ins SWRL rule which map total score to risk level. For example, suppose that a patient has age score 5, blood pressure score 3, cholesterol score 6, HDL level score 4 and smoking status score 3, These risk scores are calculated using the *swrlb:add* core built-in, the total risk score is 21. *HeartOnto:RiskLevel* SWRL custom built-in is used to bind the integer value of the total risk score to the *TotalScores* variable, then it sets the result ( which is 14%) as the value of the *hasRiskLevel* property as shown in Figure 5.6.



**Figure (5.7):** The SWRL rule to compute the patient risk score

SWRL custom built-in is user-defined methods are used directly in SWRL rules. We define RiskLevel custom built-in with OWL API 3 using Java programming language. Then we register our built-in to the Pellet reasoner to be available to use with SWRL rules (more implementation issues are detailed in Section 5.7).

### 5.5.6 Giving the recommended diagnosis

The DHDOnto system gives the recommended diagnosis of heart disease based on patient symptoms, risk factors and lab tests according to cardiac guidelines followed in Palestine. The diagnostic process depends primarily on the clinical tests and risk factors. The patient's symptoms alone are not sufficient to give an accurate diagnosis, because many types of heart diseases have the same symptoms. Distinguish type of heart diseases is mainly dependent on the results of the necessary clinical tests (i.e. ECG, Chest X-Ray, CT scan, and Stress Test).

We define the diagnostic rules (see chapter 4 Section 4.2) which are specialized for each type of heart diseases according to the clinical guidelines and to the domain expert. When the user completes filling the questionnaire and his medical profile is created, the inference engine extracts diagnostic rules and relevant OWL knowledge from HeartOnt ontology to draws inference and gives the diagnosis results for the patient. We used *recommendedDiagnosisResult* data property to store the type of disease the patient has. The diagnostic results contain: type of heart diseases, heart disease risk estimation which are obtained from calculating the heart risk score phase, treatment recommendation and result explanation.

Next, we present a case study covering all the steps of the heart diseases diagnosis process including creating medical profile, analysis of patient risk factors, calculate heart risk score and giving the recommended diagnosis, treatment recommendation and result explanation.

## 5.6 Case study

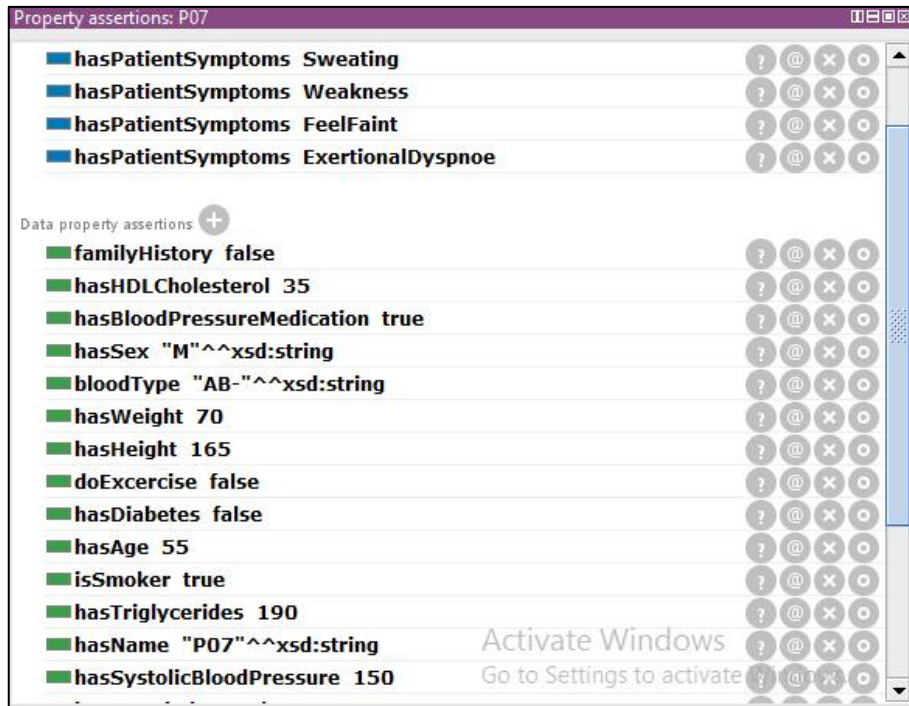
For understanding the DHDOnto system, we illustrate a usage scenario of the system showing how the user interacts through the graphical user interface to get the appropriate diagnosis for a specific patient. Suppose that the user entered the following values through the questionnaire interface for a specific patient as shown in Table 5.1.



**Table (5.1):** Patient data which are entered through the questionnaire interface.

| name  | Age      | Gender  | F. History | Diabetes   | Total.Ch        |
|---|----------|---------|------------|------------|-----------------|
| P07   | 55       | M       | false      | false      | 200             |
| Smoker  | HDL.Ch   | LDL.Ch  | HeartRate  | doExercise | P.B. medication |
| true  | 35       | 60      | 78         | false      | true            |
| SBP   | DBP      | ECGTest | SCBTest    | Q_waves    | R_waves         |
| 150   | 100      | true    | false      | 160 mm     | 35 mm           |
| T_waves   | ST_value | XRyTest | XRyValue   | S_waves    |                 |
| inversion   | abnormal | true    | abnormal   | 39 mm      |                 |
| Symptoms  |          |         |            |            |                 |
| shortness of breath, headache, weakness, palpitations, weakness |          |         |            |            |                 |

To create patient medical profile, the user fills the questionnaire, the system generates a new instance of *Patient* class with patient name as the instance name and asserts entered values into different data properties. Figure 5.8 shows instance of *Patient* class and its data properties in protégé editor.



**Figure (5.8):** Instance of Patient class and its data properties

After medical user profile is created, the DHDOnTo system run Pellet inference engine over the ontology and SWRL rules. In our case, the risk estimation and diagnosis rules are executed to combine related facts for the inference process. The necessary information for the diagnostic process and risk assessment are stored as values for inferred data properties.

- **Heart risk assessment**

The risk assessment process requires the values obtained from the inferred data properties which are used to store the risk factor score obtained by executing risk assessment rules. In our case, according to the user input, total cholesterol score is 3, blood pressure score is 2, age score is 8, HDL level score is 2 and smoking status score is 3 according to following SWRL rules:

*Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age),  
hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=200, <=  
239](?TotalCholesterol) -> hasCholesterolScores(?p, 3)*

*Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 130 , <=  
159](? SBP), hasBloodPressureMedication(?p, true) ->  
hasBloodPressureScore (?p, 2)*

*Patient(?p), hasAge (?p, ?age), xsd:integer[>= 55 , <= 59](?age) ->  
hasAgeScores (?p, 8)*

*Patient(?p), hasHDLCholesterol (?p, ? HDL), xsd:integer[<40](? HDL) ->  
hasHDLLevelScores(?p, 2)*

*Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age), isSmoker  
(?p, true) -> hasSmokingStatusScores (?p, 3)*

The total score is calculated for each factor. We assessed the risk level of heart diseases by using custom built-ins SWRL rule which map total score to risk level. The total score is 18 and the risk level is ‘30% ≥’ according to the following SWRL rule

*Patient(?p)*  
*A hasBloodPressureScore (?p, ?BPS)*  
*A hasTotalCholesterolScore (?p, ?TotalCholesterolS)*  
*A hasAgeScore (?p, ?ageS)*  
*A hasHDLCholesterolScore (?p, ?HDLS)*  
*A swrlb:add(?subTotal,?BPS,?TotalCholesterolS)*  
*A swrlb:add(?subTotal2,?ageS,?HDLS),*  
*A swrlb:add(?TotalScors,? subTotal,?subTotal2)*  
*A swrlb:add(?TotalScors,? subTotal,?subTotal2)*  
*A heartOnto: RiskLevel (?TotalScors)*  
 → *hasRiskLevel(?p, ?TotalScors)*

Figure 5.9 shows a snapshot of risk estimation interface which displays the risk point for every risk factor and the total risk points for the patient.



**Figure (5.9):** Risk point for every risk factor

- **Heart Disease diagnosis**

We utilize SWRL to detect heart disease type based on the given symptoms, the risk factor and clinical test. In this case, the patient has a symptom shortness of breath, headache, weakness, palpitations and weakness. He has ECG and X-Ray tests. Based on the patient data, the recommend diagnostic is ‘Valvular Stenosis’ according to the following rule:

*Patient(?p)*  
*∧ hasPatientSymptoms (?p, ?s)*  
*∧ hasSympValue (?s, "Valvular Stenosis")*  
*∧ hasSystolicBloodPressure (?p, ?spd)*  
*∧ hasTriglycerides (?p, ?tg)*  
*∧ swrlb: greaterThanOrEqual(?spd, 140)*  
*∧ swrlb: greaterThan (?tg, 200)*  
*∧ hasECGTest (?p, true)*  
*∧ ECG\_T\_Waves (?p, "inversion")*  
*∧ ECG\_R\_Waves (?p, ?rW)*  
*∧ swrlb: greaterThanOrEqual(?rW, 35)*  
*∧ ECG\_Q\_Waves (?p, ?qW)*  
*∧ swrlb: greaterThanOrEqual(?qW, 30)*  
 → *recommendedDiagnosisResult (?p, ValvularStenosis)*

- **Treatment recommendation and result explanation**

According to the diagnosis result and the total risk score, the system suggests the drugs and/or other required operations to be taken with the explanation about the given results. In the case of patient having ‘Valvular Stenosis’, so the system suggests appropriate medications and procedures to be followed to reduce the risk as much as possible. In this case, no medications can reverse valve stenosis. However, the system suggests required procedures to repair or replace the Aortic Valve Stenosis, Also, it suggests medications to help patient symptoms, such as ones to reduce fluid accumulation, to slow heart rate associated with valve stenosis. The following rule illustrates the appropriate drugs for ‘Valvular Stenosis’.

*Patient(?p)*  
*∧ recommendedDiagnosis (?p, ValvularStenosis)*  
 → *recommendedTreatment (?p, BetaBlockers)*  
*∧ recommendedTreatment (?p, CalciumChannelBlockers)*  
*∧ requiredProcedure (?p, 'Aortic valve replacement (استبدال صمامات القلب)')*  
*∧ requiredProcedure (?p, 'Balloon valvuloplasty (قسطرة صمامات القلب)')*

The system displays risk factors and clinical test results as reasons behind the diagnosis results. Figure 5.11 displays the diagnosis results containing risk estimation, treatment recommendation and result explanation.



**Figure (5.10):** Diagnosis results contain risk estimation, treatment recommendation and result explanation.

## 5.7 Implementation Issues

To implement DHDOnto system, we use several tools and methods. The ontology is built using the ontology editor Protégé 5.1<sup>1</sup>. The ontology was formalized in OWL DL, a description logics-based sublanguage of the OWL. It was chosen because it is highly expressive. In addition, several well-known reasoning systems are available for OWL DL, such as Pellet<sup>2</sup> which supports custom built-ins

We implement the DHDOnto system using Java programming language and OWL API 3<sup>3</sup> for manipulating HeartOnto. The OWL API is a Java API and reference implementation for creating, manipulating and serialising OWL ontologies. We used NetBeans IDE<sup>4</sup> to help us to implement the system.

<sup>1</sup> <http://protege.stanford.edu/>

<sup>2</sup> <https://github.com/stardog-union/pellet>

<sup>3</sup> <https://owlapi.sourceforge.net/>

<sup>4</sup> <https://netbeans.org>

The diagnostic rules were written in SWRL. There are many tools for managing SWRL rules. One of these tool is SWRLTab, The SWRLTab provides a set of standalone graphical interfaces for managing SWRL rules. The SWRLTab available only in the old Protege 3 which does not support OWL2, and is not usable in the newer Protégé 5 or in Java programs using the newer OWL API 3. The rules were written using the Protégé 5.1 Editor. When editing rules in this environment, we directly refer to OWL classes, properties, and individuals within OWL HeartOnto ontology. The rules are stored as OWL individuals in the HeartOnto ontology (see Appendix A for the full list of SWRL rules).

To make Pellet 2.2 reasoner compatible with custom built-in using OWL2 and OWL API 3, we extend Pellet reasoner to support built-in through three steps:

### 1- Create custom built-in class

We have created custom built-in class in Java called GetHeartRiskLevel. Then we have created getRiskLevel function which returns the heart risk level based on a patient total score as shown below.

```

..
...
import com.clarkparsia.pellet.rules.builtins.GeneralFunction;
..
..
public class GetHeartRiskLevel {
    private static class RiskLevel implements GeneralFunction {
        public String getRiskLevel (int totalPonits) {
            String riskLevel;
            if (totalPonits < 0) {
                riskLevel = "1<";
            } else if (totalPonits == 0) {
                riskLevel = "1";
            }
            ..
            ..
            Return riskLevel;
        }
    }
}

```

## 2- Register built-in to Pellet

We have inserted the code below before loading Pellet in the main class of DHDOnto system.

```
BuiltInRegistry.instance.registerBuiltIn( "urn:HeartOnto:builtIn# RiskLevel",  
new GeneralFunctionBuiltIn( GetHeartRiskLevel. RiskLevel))
```

## 3- Use custom built-in in SWRL rules

Finally, we use our built-in in HeartOnto to be available to SWRL rules as shown below (Manchester Syntax)

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
PREFIX owl: <http://www.w3.org/2002/07/owl#>  
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>  
PREFIX: <urn:HeartOnto:RiskLevel #>  
Ontology: <http://www.semanticweb.org/hosam/ontologies/2016/7/CardioOnt>
```

## 5.8 Summary

This chapter has presented the architecture of the DHDOnto system which depends on the HeartOnto ontology which we have built to model the different entities in heart diseases domain. The functions of the system consist of four tasks: heart disease diagnosing, heart disease risk estimation, treatment recommendation and result explanation. We have presented the steps of heart diseases diagnosis process which contains: create patient medical profile, run inference engine, patient risk factors analysis, calculate heart risk score and give the recommended diagnosis. We presented a case study covering all the steps of the heart diseases diagnosis process. Finally, we mentioned some implementation issues related to realizing the system.

# **Chapter 6**

## **Experimental Results and Evaluation**



## Chapter 6

### Experimental Results and Evaluation

This chapter presents the experiments performed for the evaluation of the proposed DHDOnto system. The experiments are conducted to assess the effectiveness of the heart diseases diagnosis offered by the system.

The system is tested using a set of patients with heart diseases provided by a domain expert. Diagnosis results are compared with the results obtained by the domain expert. In the following sections, we present details about the experimental settings including the knowledge base and SWRL rules used. Subsequently, the experimental procedures and results are discussed.

#### 6.1 Experimental Settings

To evaluate the system, we have asked the expert in heart diseases to select a sample set of patients with heart disease. The sample contains the various heart diseases cases to examine through the system. The patient data we used for evaluation were obtained from the cardiology unit in Alshifa hospital in Gaza strip. The patient dataset was restricted to 30 patients, it sufficient to cover all diagnosis cases that the system provides. Table 6.1 contains the disease category, disease types in each category and the number of patients and their disease type.

**Table (6.1):** The number of patients and their disease type.

| Disease Category          | Disease Type                | Number of Patient |
|---------------------------|-----------------------------|-------------------|
| Coronary Artery Disease   | Acute Coronary Syndrome     | 4                 |
|                           | Stable Angina               | 3                 |
|                           | Heart Failure               | 6                 |
| Valve Disease             | Valvular stenosis           | 4                 |
|                           | Valvular insufficiency      | 3                 |
| Cardiomyopathy Disease    | Hypertrophic Cardiomyopathy | 5                 |
|                           | Dilated Cardiomyopathy      | 2                 |
|                           | Restrictive Cardiomyopathy  | 3                 |
| <b>Total: 30 patients</b> |                             |                   |

Diagnosing all cases of heart disease requires patient profile data such as symptoms, age, diabetes, hypertension, cholesterol, etc. in addition of clinical investigations and tests. Note that the accuracy of diagnostic results is mainly dependent on the results of the necessary tests (i.e. ECG, Chest X-Ray and Stress Test). We asked the expert for the results of diagnosis for each case to compare them with the results of our system. Table 6.2 shows the sample patient data includes diagnosis results and treatments.

**Table (6.2):** An example of a patient data that have been diagnosed by doctor.

| id   | Age      | Gender  | F. History   | Diabetes     | Total.Ch        |
|--|----------|---------|--|--------------|-----------------|
| P27  | 66       | M       | false  | true         | 220             |
| Smoker   | HDL.Ch   | LDL.Ch  | HeartRate  | doExcercise  | P.B. medication |
| true   | 38       | 60      | 80   | false        | true            |
| SBP  | DBP      | ECGTest | SCBTest  | Q_waves      | R_waves         |
| 160  | 100      | true    | true   | 35 mm        | abnormal        |
| T_waves  | ST_value | XRyTest | XRyValue   | LV Thickness | S. Thickness    |
| abnormal   | normal   | true    | normal   | 15 mm        | 1.5             |
| Symptoms   |          |         |  |              |                 |
| shortness of breath, fainting, fluttering or pounding heartbeats, rapid heartbeats |          |         |  |              |                 |
| Diagnosis Result   |          |         | Treatment  |              |                 |
| Hypertrophic cardiomyopathy<br>(تضخم عضلة القلب)                                   |          |         | Beta blockers, propranolol, atenolol<br>(Tenormin), calcium channel blockers |              |                 |

In general, we tested the system with 30 different cases of patients with heart diseases. These 30 patients were carefully chosen to represent the different SWRL rules used in the diagnoses of different types of heart diseases.

SWRL rules depends mainly on the ontology for heart diseases domain and its richness. Table 6.3 depicts the size of the ontology, the number of classes, the number of object properties, the number of data properties and the number of SWRL rules in HeartOnto.

**Table (6.3):** The size of the heart disease ontology

| Ontology Components         | Number |
|-----------------------------|--------|
| Number of Classes           | 65     |
| Number of Object Properties | 10     |
| Number of Data Properties   | 42     |
| Number of Instances         | 139    |
| Number of SWRL Rules        | 63     |

We have defined around 139 instances (individuals) representing all ontology concepts. Table 6.4 depicts the number of instances. The ontology together with these instances form the knowledge base, respectively, the main component of DHDOno system.

**Table (6.4):** The number of individuals

| Name of Classes | Number of individuals |
|-----------------|-----------------------|
| Diseases        | 20                    |
| Symptoms        | 45                    |
| Patient         | 30                    |
| Treatment       | 28                    |
| Diagnosing      | 16                    |

## 6.2 Evaluation of Diagnosis Results

In this section, we present the evaluation of the DHDOno system to determine whether it behaves exactly as we expect. The main goal of this evaluation is to assess the system's ability to diagnosing heart disease based on patient symptoms and tests results.

System evaluation depends on SWRL rules which are used in our approach to find correct diagnosis for patients according to patient symptoms and diagnosis test results. We have tested 30 patients with heart diseases. We verify the system ability to diagnosis most of these cases. Table 7.5 depicts the number of patients who were diagnosed and the number of correct/incorrect diagnostic results.

**Table (6.5):** The number of patients and their diagnosis results.

| Total Patients | Correct Diagnosis Results | Incorrect Diagnosis Results |
|----------------|---------------------------|-----------------------------|
| 30             | 27                        | 3                           |

Results show that the system can correctly diagnoses 27 out of the 30 patients (ratio of correctness is 90%). Table 6.6. shows the patients that have been diagnosed using traditional diagnosis, and using DHDOnto system, diagnosis results and the rules that have been applied.

**Table (6.6):** The patients' information that have been diagnosed

| Patient ID | Correct Diagnosis           | DHDOnto System Diagnosis   | results   | Rules No. |
|------------|-----------------------------|----------------------------|-----------|-----------|
| P01        | Acute Coronary Syndrome     | Acute Coronary Syndrome    | correct   | Rule-3    |
| P02        | Restrictive Cardiomyopathy  | Restrictive Cardiomyopathy | correct   | Rule-11   |
| P03        | Heart Failure               | Heart Failure              | correct   | Rule-05   |
| P04        | Restrictive Cardiomyopathy  | Restrictive Cardiomyopathy | correct   | Rule-11   |
| P05        | Stable Angina               | Stable Angina              | correct   | Rule-04   |
| P06        | Heart Failure               | Heart Failure              | correct   | Rule-05   |
| P07        | Valvular Stenosis           | Valvular Stenosis          | correct   | Rule-06   |
| P08        | Stable Angina               | Stable Angina              | correct   | Rule-04   |
| P09        | Hypertrophic Cardiomyopathy | No Answer                  | incorrect | Rule-08   |
| P10        | Dilated Cardiomyopathy      | Dilated Cardiomyopathy     | correct   | Rule-10   |
| P11        | Acute Coronary Syndrome     | Acute Coronary Syndrome    | correct   | Rule-2    |
| P12        | Restrictive Cardiomyopathy  | Restrictive Cardiomyopathy | correct   | Rule-11   |

| Patient ID | Correct Diagnosis           | DHDOnto System Diagnosis    | results   | Rules No. |
|------------|-----------------------------|-----------------------------|-----------|-----------|
| P13        | Dilated Cardiomyopathy      | Dilated Cardiomyopathy      | correct   | Rule-10   |
| P14        | Hypertrophic Cardiomyopathy | Hypertrophic Cardiomyopathy | correct   | Rule-8    |
| P15        | Heart Failure               | Heart Failure               | correct   | Rule-05   |
| P16        | Acute Coronary Syndrome     | Acute Coronary Syndrome     | correct   | Rule-02   |
| P17        | Hypertrophic Cardiomyopathy | Hypertrophic Cardiomyopathy | correct   | Rule-09   |
| P18        | Valvular Stenosis           | Valvular stenosis           | correct   | Rule-06   |
| P19        | Valvular Stenosis           | No Answer                   | incorrect | Rule-06   |
| P20        | Valvular Insufficiency      | Valvular Insufficiency      | correct   | Rule-07   |
| P21        | Heart Failure               | Heart Failure               | correct   | Rule-05   |
| P22        | Heart Failure               | Heart Failure               | correct   | Rule-05   |
| P23        | Valvular Insufficiency      | valvular Insufficiency      | correct   | Rule-07   |
| P24        | Stable Angina               | Stable Angina               | correct   | Rule-04   |
| P25        | Valvular Stenosis           | Valvular Stenosis           | correct   | Rule-06   |
| P26        | Valvular Insufficiency      | Valvular Insufficiency      | correct   | Rule-07   |
| P27        | Hypertrophic Cardiomyopathy | Hypertrophic Cardiomyopathy | correct   | Rule-08   |
| P28        | Acute Coronary Syndrome     | No Answer                   | incorrect | Rule-01   |
| P29        | Heart Failure               | Heart Failure               | correct   | Rule-05   |
| P30        | Hypertrophic Cardiomyopathy | Hypertrophic Cardiomyopathy | correct   | Rule-9    |

In the following example, we select a patient who was correctly diagnosed and show how the system obtained the correct diagnosis. Subsequently, we explain why the system cannot give the correct diagnosis.

**Example:** The patient with P08

The patient with id P08 has symptoms shortness of breath (ضيق في التنفس), nausea (غثيان), fatigue (تعب), dizziness (دوخة), profuse sweating (تعرق غزير), anxiety (قلق), Pain in arms, neck, jaw (آلام في الذراعين، الفك والرقبة), and he has history of heart disease, diabetes, smoking, not exercising, high blood (systolic > 140), high cholesterol (HDL < 40) and age 57.

**Step 1:** the system generates a new instance of *Patient* class with patient name as the instance name and assert patient data into different data properties.

**Step 2:** The DHDOnTo system run the Pellet inference engine over the ontology and SWRL rules. The diagnosis rules (rule-1 to rule-20) and risk estimation rules (rule-21 to rule-63) are executed to combine related facts for inference process. The necessary information for diagnostic process and risk assessment are stored as values for inferred data property.

**Step 3:** In this specific case, according to patient symptoms and values of different risk factors; the data property *recommendedDiagnosisResult* yields the diagnosis results based the following rule:

```
Patient(?p)
A hasPatientSymptoms (?p, ?s)
A hasSympValue (?s, "Stable Angina")
A familyHistory (?p, true)
A isSmoker (?p, true)
A doExcercise (?p, false)
A hasSystolicBloodPressure (?p, ?spd)
A hasHDLCholesterol (?p, ?hdl)
A swrlb: greaterThanOrEqual(?spd, 140)
A swrlb: lessThan (?hdl, 40)
→ recommendedDiagnosisResult (?p, Stable Angina)
```

Based on the above rule, patient with id P108 has Stable Angina (ذبحة صدرية). According to diagnosis results, the system suggests the drugs and required measures to be taken with the explanation about the given results. Figure 6.1 displays the

diagnosis results for patient with id P08 with risk estimation, treatment recommendation and result explanation.



**Figure (6.1):** Diagnosis results for patient with id P08

Three of the given thirty patients were not correctly diagnosed. In the following part, we explain why the system could not give the correct diagnosis.

### The patient with P09

The patient with id P09 has the following symptoms: dyspnea (زلة تنفسية), syncope (غشيان), dizziness (دوخة), orthopnea and paroxysmal nocturnal dyspnea (ضيق تنفس ليلي) and patient has family history of heart disease, diabetes, non-smoker, not exercising, high blood (systolic > 140), high triglycerides (> 200), abnormal Chest X-Ray test results and no ECG test.

According to the domain expert (doctor), the final diagnosis result for this case is Hypertrophic Cardiomyopathy (تضخم عضلة القلب). This diagnosis result was deduced based on patient symptoms, lab tests, family history and the X-Ray test without need

to ECG test which has a crucial role in the Cardiomyopathy (اعتلال عضلة القلب) disease diagnosis. Based on the doctor's experience, there are previous cases which were diagnosed for hypertrophic cardiomyopathy similar to the given case.

In the DHDOnTo system, diagnosis of the Cardiomyopathy disease takes into account ECG test results, the main factor for Hypertrophic Cardiomyopathy diagnosis is the thickness of left ventricular wall (سمك جدار البطين الأيسر) which is 30 mm or more ,X-Ray test alone is not enough. The following SWRL rule illustrates Hypertrophic Cardiomyopathy diagnosis.

```

Patient(?p)
A hasPatientSymptoms (?p, ?s)
A hasSympValue (?s, "Hypertrophic Cardiomyopathy")
A familyHistory (?p, true)
A hasSystolicBloodPressure (?p, ?spd)
A hasTriglycerides (?p, ?tg)
A swrlb: greaterThanOrEqual(?spd, 140)
A swrlb: greaterThan (?tg, 200)
A hasECGTest (?p, true)
A ECG_LV_wall_thickness (?p, ?lvTH)
A swrlb: greaterThanOrEqual(?lvTH, 30)
→ recommendedDiagnosisResult (?p, HypertrophicCardiomyopathy)

```

### The patient with P28

The patient with id P28 has constricting (تشنج), crushing (الشعور بالسحق), pressure (الشعور بالضغط), tightness (ضيق), pain in shoulder or left arm (ألم في الكتف أو الذراع الأيسر), age of 62, smoker, high cholesterol, high blood pressure and has ECG test with ST-segment deviation > 0.05 mV and Q-wave is Fixed.

Acute Coronary Syndrome is the expected diagnosis result by the domain expert, while our system could not distinguish the type of disease due to lack of full tests results. Acute Coronary Disease diagnosis requires an electrocardiogram test (ECG), the common ECG abnormalities include T-wave tenting or inversion, ST-segment elevation, and Q-waves. In this case, the patient data does not contain T-wave value to



be taken into account according to SWRL rule responsible for the Acute Coronary disease diagnosis as follows:

```

Patient(?p)
A hasPatientSymptoms (?p, ?s)
A hasSympValue (?s, "Acute Coronary Syndrome")
A hasPatientSymptoms (?p, "diaphoresis")
A hasSystolicBloodPressure (?p, ?spd)
A hasST-SegmentValue (?p, ?t)
A hasT-WaveValue (?p, ?w)
A swrlb:lessThanOrEqual(?spd, 70)
A swrlb:lessThan (?t, 0.05)
A swrlb:lessThan (?w, 0.02)
→ recommendedDiagnosisResult (?p, AcuteCoronarySyndrome),
likelihood(?p, "high")

```

Patient data does not achieve the required terms of the Acute Coronary Syndrome diagnosis rule (as described in the previous rule). This result indicates that the full tests result is very important to infer the correct diagnosis. Figure 7.2 displays incorrect diagnosis results for patient with id P28.



Figure (7.2): Diagnosis results for patient with id P28

## The patient with P19

The patient with id P19 has heart palpitations (خفقان القلب), feeling faint (الشعور بالإغماء), pounding in chest (طرق في الصدر), age of 55, type 2 diabetes, smoker, not exercising, high blood pressure, high cholesterol and has X-ray test, computerized tomography (CT) scan, no ECG test.

According to the domain expert, the final diagnosis result for this case is Aortic Valve Stenosis (تضييق صمام الأبهر). Diagnosis result was deduced based on CT scans test. There are many diagnostic tests to detect Aortic Valve Stenosis disease such as ECG, Chest X-ray, CT and Cardiac catheterization. In our system, we depend on the ECG test in the diagnosis of AVS.

The SWRL rule responsible for the AVS diagnosis does not take into account the results of CT test. This result indicates that the other tests results may be useful to infer the correct diagnosis.

In light of the above results we can summarize the strengths and limitations of the system as follows:

- The system has the ability to diagnose all types of coronary artery diseases, valve disease and cardiomyopathy disease.
- The system relies heavily on SWRL rules to diagnose heart diseases. the ability of the system to provide accurate diagnosis depends on the coverage of the SWRL rules.
- The patient's symptoms alone are not sufficient to give an accurate diagnosis, because many types of heart diseases have the same symptoms. Distinguishing the types of heart diseases is mainly dependent on the results of the necessary tests (i.e. ECG, Chest X-Ray, CT scan, and Stress Test). This means that diagnose patient without the necessary tests results does not give accurate results.
- The diagnostic rules use the most important test in the diagnosis, according to the type of disease. For example, in the Aortic Valve Stenosis diagnosis rule, the ECG test is determinant for diagnosis. However, other tests can diagnose

the AVS disease (i.e. CT scan). In DHDOnto system, the CT scan test result is unsupported in the AVS diagnosis. The more complete the SWRL rules in coverage of the different types of diagnosis tests, the more the system becomes capable of providing accurate diagnosis.

As for comparison with other medical diagnosis systems, we are not aware of any specific system in the domain of heart diseases diagnosis in Palestine that uses ontologies. Therefore, we cannot compare our work with other researches.

### **6.3 Summary**

In this chapter, we presented the experimental results and provided an evaluation of the proposed DHDOnto system. The experimental settings are presented in terms of the ontology size and number of patients that have been diagnosed. The evaluation for the approach shows that the system can correctly diagnosis 27 out of the 30 patients. We have discussed the diagnostic results, explained how the system obtain the correct diagnosis and explained why the system could not give the correct diagnosis for some cases.

# **Chapter 7**

## **Conclusions and Future Work**

## Chapter 7

### Conclusions and Future Work

In this research, we have developed an approach for diagnosing heart diseases based on domain ontology and SWRL rules. The ontology contains the relevant concepts related to the diagnostics, data of patient (personal information, symptoms, risk factors and clinical tests results) and treatment. SWRL rules are created from valid relationships between ontology concepts to diagnose heart disease and estimate the risk of heart diseases. Based on ontology and these rules we proposed a system that can be used by doctors and medical students to diagnosis heart diseases. DHDOnto system consists of several components which are Knowledge Base, Rule Base, Inference Engine and User Interface.

The system functions are divided into four tasks which are heart disease diagnosis, heart diseases risk estimation, treatment recommendation and results explanation. Diagnosis process is performed by checking the input provided by the user to different questions related to the symptoms and risk factors of heart disease. All diagnosis rules and relevant knowledge is extracted from the domain ontology (HeartOnt). The reasoner use patient data and rules to draw inference and gives recommended diagnosis, treatment recommendation and result explanation.

Experiments were performed to test the system ability for diagnosing different types of heart diseases. System evaluation depended on SWRL rules which are used to find correct diagnosis for patients according to patient symptoms and diagnosis test results. We evaluate the system based on pre-tested 30 patients with heart diseases. In the evaluation process, the results that generated from the system showed that the system can correctly diagnose 27 out of the 30 patients with accuracy of 90%.

The main contribution of this research is using the ontology and SWRL rules to diagnosis all types of coronary artery diseases, valve disease and cardiomyopathy disease.

There are still several improvements that need to be addressed in a future work:

- We look to extend the system to cover other types of heart diseases and extend SWRL rules to cover the different types of diagnosis tests like CT scan test.
- Integration of machine learning in the system to enhance its diagnosis capabilities through learning from the diagnosis results.
- Increase the number of patients for complete and more accurate evaluation of the system.
- Finally, we look forward to develop a web and mobile application versions of the system.

# References

## References

- Abidi, S. R., Hussain, S., Shepherd, M., & Abidi, S. S. R. (2007). Ontology-based modeling of clinical practice guidelines: a clinical decision support system for breast cancer follow-up interventions at primary care settings. *In Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems* (p. 845). IOS Press.
- Achar, S. A., Kundu, S., & Norcross, W. A. (2005). Diagnosis of acute coronary syndrome. *chest*, 100, 9.
- Al-Hamadani, B. T., & Alwan, R. F. (2015). An Ontology-Based Expert System for General Practitioners to Diagnose Cardiovascular Diseases. *Advances in Computational Sciences and Technology*, 8(1), 53-65.
- Alfonse, M., Aref, M. M., & Salem, A.-B. M. (2014). An ontology-based system for cancer diseases knowledge management. *International Journal of Information Engineering and Electronic Business*, 6(6), 55.
- Alharbi, R. F., Berri, J., & El-Masri, S. (2015). *Ontology based clinical decision support system for diabetes diagnostic*. Paper presented at the Science and Information Conference (SAI), 2015.
- Almutair, H., & El-Masri, S. (2014, January). *Model of Diagnosis Knowledge Base Ontology Framework based on Clinical Practice Guidelines*. In Proceedings of the International Conference on Bioinformatics & Computational Biology (BIOCOMP) (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
- Andrews, P. B. (2002). *An introduction to mathematical logic and type theory* (Vol. 27): Springer Science & Business Media.
- Bau, C.-T., Chen, R.-C., & Huang, C.-Y. (2014). Construction of a clinical decision support system for undergoing surgery based on domain ontology and rules reasoning. *TELEMEDICINE and e-HEALTH*, 20(5), 460-472.
- Casebeer, L., Huber, C., Bennett, N., Shillman, R., Abdolrasulnia, M., Salinas, G. D., & Zhang, S. (2009). Improving the physician-patient cardiovascular risk dialogue to improve statin adherence. *BMC family practice*, 10(1), 48.



- Farooq, K., Hussain, A., Leslie, S., Eckl, C., & Slack, W. (2011). *Ontology-driven cardiovascular decision support system*. Paper presented at the Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on.
- Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Blaha, M. J., . . . Franco, S. (2014). Executive summary: heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation, 129*(3), 399.
- Hebeler, J., Fisher, M., Blace, R., & Perez-Lopez, A. (2011). *Semantic web programming*: John Wiley & Sons.
- Horridge, M., Jupp, S., Moulton, G., Rector, A., Stevens, R., & Wroe, C. (2009). A practical guide to building owl ontologies using protégé 4 and co-ode tools edition 1. 2. *The University of Manchester, 107*.
- Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosz, B., & Dean, M. (2004). SWRL: A semantic web rule language combining OWL and RuleML. *W3C Member submission, 21*, 79.
- Jain, V., & Singh, M. (2013a). Ontology based information retrieval in semantic web: A survey. *International Journal of Information Technology and Computer Science (IJITCS), 5*(10), 62.
- Jain, V., & Singh, M. (2013b). Ontology Development and Query Retrieval using Protégé-8. *International Journal of Intelligent Systems and Applications, 5*(9), 67.
- Jović, A., Gamberger, D., & Krstajić, G. (2011). Heart failure ontology. *Bio-algorithms and med-systems, 7*(2), 101-110.
- Lezcano, L., Sicilia, M.-A., & Rodríguez-Solano, C. (2011). Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules. *Journal of biomedical informatics, 44*(2), 343-353.
- Lorenz, D. P., & Spink, A. (2004). Semantics and the medical web: a review of barriers and breakthroughs in effective healthcare query. *Health Information & Libraries Journal, 21*(2), 109-116.
- Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., . . . Ahn, S. Y. (2013). Global and regional mortality from 235 causes of death for 20

- age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), 2095-2128.
- Martínez-Romero, M., Vázquez-Naya, J. M., Pereira, J., Pereira, M., Pazos, A., & Baños, G. (2013). The iOSC3 system: using ontologies and SWRL rules for intelligent supervision and care of patients with acute cardiac disorders. *Computational and mathematical methods in medicine*, 2013.
- McGuinness, D. L., & Van Harmelen, F. (2004). OWL web ontology language overview. *W3C recommendation*, 10(10), 2004.
- McMurray, J. J., Adamopoulos, S., Anker, S. D., Auricchio, A., Böhm, M., Dickstein, K., . . . Gomez - Sanchez, M. A. (2012). ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012. *European journal of heart failure*, 14(8), 803-869.
- Ministry of Health Annual Report (2015): Health Status in Palestine. Section 3: Trends in Health Status.
- Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology: Stanford knowledge systems laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880, Stanford, CA.
- O'connor, M., Knublauch, H., Tu, S., Grosz, B., Dean, M., Grosso, W., & Musen, M. (2005). *Supporting rule system interoperability on the semantic web with SWRL*. Paper presented at the International Semantic Web Conference.
- Rawte, V., & Roy, B. (2015). *OBESTDD: Ontology Based Expert System for Thyroid Disease Diagnosis*. Paper presented at the Nascent Technologies in the Engineering Field (ICNTE), 2015 International Conference on.
- Robu, I., Robu, V., & Thirion, B. (2006). An introduction to the Semantic Web for health sciences librarians. *Journal of the Medical Library Association*, 94(2), 198.
- Roussey, C., Pinet, F., Kang, M. A., & Corcho, O. (2011). An introduction to ontologies and ontology engineering *Ontologies in Urban Development Projects* (pp. 9-38): Springer.
- Saad, S., Salim, N., Zainal, H., & Muda, Z. (2011). *A process for building domain ontology: An experience in developing Solat ontology*. Paper presented at the

Electrical Engineering and Informatics (ICEEI), 2011 International Conference on.

Sherimon, P., Vinu, P., Krishnan, R., & Saad, Y. (2014). Ontology driven analysis and prediction of patient risk in diabetes. *Can. J. Pure Appl. Sci*, 8(3), 3043-3050.

Sherimon, P., Vinu, P., Krishnan, R., & Takroni, Y. (2013). Ontology based system architecture to predict the risk of hypertension in related diseases.

*International Journal of Information Processing and Management*, 4(4), 44.

Steinberg, E., Greenfield, S., Wolman, D. M., Mancher, M., & Graham, R. (2011).

*Clinical practice guidelines we can trust*: National Academies Press.

Thirugnanam, M. (2013). An ontology based system for predicting disease using SWRL rules. *International Journal of Computer Science and Business Informatics*, 7(1).

## Appendix A

### Rules for the Diagnosis of Heart Diseases

| <b>Rules for diagnosis Acute Coronary Syndrome</b> |  |
|--|--|
| Rule:01  | <p>Patient(?p), hasPatientSymptoms (?p, ?s) , hasSympValue (?s, “AcuteCoronarySyndrome”),familyDiseaseHistory (?p, “Angina”) , hasPatientSymptoms (?p, “diaphoresis”) ,hasSystolicBloodPressure (?p, ?spd), hasST-SegmentValue (?p, ?t), hasT-WaveValue (?p, ?w), swrlb:lessThanOrEqual(?spd, 70) , swrlb:lessThan (?t, 0.05) , swrlb:lessThan (?w, 0.02) -&gt; recommendedDiagnosisResult (?p, AcuteCoronarySyndrome),likelihood(?p , “high”)</p> |
| Rule:02  | <p>hasPatientSymptoms (?p, ?s), hasSympValue(?s, "Acute Coronary Syndrome"), hasAge(?p, ?age) , xsd:integer[&gt;= 50](?age) ,Patient(?p) , hasCardiacMarkersTest(?p , true) , hasCardiacMarkersTestValue(?p , "normal" ) , ECG_Q_Waves (?p , "Fixed" ) , ECG_ST_value (?p , "Abnormal" ) -&gt; recommendedDiagnosisResult (?p, AcuteCoronarySyndrome),likelihood(?p , “intermediate”)</p>  |
| Rule:03  | <p>Patient(?p),hasPatientSymptoms (?p, ?s), hasSympValue(?s, "Acute Coronary Syndrome"), Patient(?p) , hasCardiacMarkersTest(?p , true) , hasCardiacMarkersTestValue(?p , "normal" ) , ECG_T_Waves (?p , "flattening" ) -&gt; recommendedDiagnosisResult (?p, AcuteCoronarySyndrome),likelihood(?p , “low”)</p>  |
| <b>Rule for diagnosis Stable Angina</b>            |  |
| Rule:04  | <p>Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, “Stable Angina”), familyHistory (?p, true), doExcercise (?p, false), hasSystolicBloodPressure (?p, ?spd), hasHDLCholesterol (?p, ?hdl), swrlb:greaterThanOrEqual(?spd, 140), swrlb:lessThan (?hdl, 40) -&gt; recommendedDiagnosisResult (?p, StableAngina)</p>  |

| <b>Rule for diagnosis Heart Failure</b>               |   |
|---|---|
| Rule:05   | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Heart Failure"), hasECGTest (?p, true), ECG_ST_value (?p, "Abnormal"), ECG_T_Waves (?p, "flattening"), hasSystolicBloodPressure (?p, ?spd), hasHDLCholesterol (?p, ?hdl), swrlb: greaterThanOrEqual(?spd, 140), swrlb:lessThan (?hdl, 40) -> recommendedDiagnosisResult (?p, HeartFailure)   |
| <b>Rule for diagnosis Valvular Stenosis</b>           |   |
| Rule:06   | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Valvular Stenosis"), hasECGTest (?p, true), ECG_T_waves (?p, "ubnormal"), ECG_R_waves (?p, "ubnormal"), ECG_S_waves (?p, "normal"), hasSystolicBloodPressure (?p, ?spd), hasHDLCholesterol (?p, ?hdl), swrlb: greaterThanOrEqual(?spd, 140), swrlb:lessThan (?hdl, 50) -> recommendedDiagnosisResult (?p, ValvularStenosis)                              |
| <b>Rule for diagnosis Valvular Insufficiency</b>      |   |
| Rule:07   | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Valvular Insufficiency"), hasECGTest (?p, true), ECG_Q_waves (?p, ?qW), swrlb:lessThan (?qW, 40), ECG_R_waves (?p, "ubnormal"), ECG_S_waves (?p, "normal"), hasSystolicBloodPressure (?p, ?spd), hasHDLCholesterol (?p, ?hdl), swrlb: greaterThanOrEqual(?spd, 140), swrlb:lessThan (?hdl, 40) -> recommendedDiagnosisResult (?p, ValvularInsufficiency) |
| <b>Rule for diagnosis Hypertrophic Cardiomyopathy</b> |   |
| Rule:08   | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Hypertrophic Cardiomyopathy"), hasSystolicBloodPressure (?p, ?spd), hasTriglycerides (?p, ?tg), swrlb: greaterThanOrEqual(?spd, 140), swrlb: greaterThan (?tg, 200), hasECGTest (?p, true), ECG_LV_wall_thickness (?p, ?lvTH), swrlb: greaterThanOrEqual(?lvTH, 30) -> recommendedDiagnosisResult (?p, HypertrophicCardiomyopathy)                       |

|  |  |
|--|--|
| <b>Rule for diagnosis Hypertrophic Cardiomyopathy (another case)</b> |  |
| Rule:09  | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Hypertrophic Cardiomyopathy"), hasSystolicBloodPressure (?p, ?spd), swrlb: greaterThanOrEqual(?spd, 140), hasECGTest (?p, true), hasXRayTest(?p,true), XRayResult(?, "ubnormal" -> recommendedDiagnosisResult (?p, HypertrophicCardiomyopathy)  |
| <b>Rule for diagnosis Dilated Cardiomyopathy</b>                     |  |
| Rule:10  | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Dilated Cardiomyopathy"), hasSystolicBloodPressure (?p, ?spd), swrlb: greaterThanOrEqual(?spd, 140), hasECGTest (?p, true), ECG_LV_wall_thickness (?p, ?lvTH), swrlb: greaterThanOrEqual(?lvTH, 25), ECG_S_Thickness(?p,?sTh), swrlb: greaterThanOrEqual(?sTh, 1.8) -> recommendedDiagnosisResult (?p, DilatedCardiomyopathy) |
| <b>Rule for diagnosis Restrictive Cardiomyopathy</b>                 |  |
| Rule:11  | Patient(?p), hasPatientSymptoms (?p, ?s), hasSympValue (?s, "Restrictive Cardiomyopathy"), hasSystolicBloodPressure (?p, ?spd), swrlb: greaterThanOrEqual(?spd, 140), hasECGTest (?p, true), ECG_LV_wall_thickness (?p, ?lvTH), greaterThanOrEqual(?lvTH, 50), ECG_S_Thickness(?p,?sTh), lessThanOrEqual(?sTh, 0.5) -> recommendedDiagnosis(?p, Restrictive Cardiomyopathy)                |
| <b>Rule for treatment of Acute Coronary Syndrome</b>                 |  |
| Rule:12  | Patient(?p), recommendedDiagnosis (?p, AcuteCoronarySyndrome) -> recommendedTreatment (?p, DailyAspirin), recommendedTreatment (?p, ACE_Inhibitors), requiredProcedure (?p, 'Stenting (تركيب دعامات)'), requiredProcedure (?p, 'Coronary bypass surgery (جراحة لتغيير الشرايين التاجية)')  |
| <b>Rule for treatment of Stable Angina</b>                           |  |
| Rule:13  | Patient(?p), recommendedDiagnosis (?p, StableAngina) -> recommendedTreatment (?p, Nitrates), recommendedTreatment (?p, BetaBlockers), recommendedTreatment (?p, CalciumChannelBlockers)  |

|  |  |
|--|--|
| <b>Rule for treatment of Heart Failure</b>               |  |
| Rule:14  | Patient(?p), recommendedDiagnosis (?p, HeartFailure)<br>-> recommendedTreatment (?p, Diuretics), recommendedTreatment (?p, WaterPills), recommendedTreatment (?p, ACE_Inhibitors ),<br>recommendedTreatment (?p, Pacemakers)   |
| <b>Rule for treatment of valvular stenosis diseases</b>  |  |
| Rule:15  | Patient(?p), recommendedDiagnosis (?p, ValvularStenosis)<br>-> recommendedTreatment (?p, BetaBlockers), recommendedTreatment (?p, CalciumChannelBlockers), requiredProcedure (?p, ‘Aortic valve replacement (استبدال صمامات القلب)’), requiredProcedure (?p, ‘Balloon valvuloplasty (قسطرة صمامات القلب)’)     |
| <b>Rule for treatment of Valvular Insufficiency</b>      |  |
| Rule:16  | Patient(?p), recommendedDiagnosis (?p, ValvularInsufficiency)<br>-> recommendedTreatment (?p, BetaBlockers), recommendedTreatment (?p, anticoagulants), requiredProcedure (?p, ‘Aortic valve replacement (استبدال صمامات القلب)’), requiredProcedure (?p, ‘Adding tissue to patch holes (إضافة أنسجة للثقوب)’) |
| <b>Rule for treatment of Hypertrophic Cardiomyopathy</b> |  |
| Rule:17  | Patient(?p), recommendedDiagnosis (?p, HypertrophicCardiomyopathy)<br>-> recommendedTreatment (?p, Diltiazem), recommendedTreatment (?p, Amiodarone), recommendedTreatment (?p, Disopyramide)  |
| <b>Rule for treatment of Dilated Cardiomyopathy</b>      |  |
| Rule:18  | Patient(?p), recommendedDiagnosis (?p, DilatedCardiomyopathy)<br>-> recommendedTreatment (?p, L-carnitine), recommendedTreatment (?p, Antibiotics), recommendedTreatment (?p, coenzyme_Q10).   |
| <b>Rule for treatment of Restrictive Cardiomyopathy</b>  |  |
| Rule:19  | Patient(?p), recommendedDiagnosis (?p, RestrictiveCardiomyopathy)<br>-> recommendedTreatment (?p, BloodThinning), recommendedTreatment (?p, Chemotherapy), recommendedTreatment (?p, Amiodarone).  |
| Rule:20  | (has_related_diseases min 1 related_diseases) (?x), Patient(?x) -> Candidate(?x)   |

## Rules for Heart Disease Risk Estimation

| Rules for calculate total cholesterol scores |   |
|--|---|
| Rule:21                                      | Patient(?p), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[< 160](?TotalCholesterol) -> hasCholesterolScores(?p, 0)   |
| Rule:22                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 60 , <= 69](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[<= 199 , >= 160](?TotalCholesterol) -> hasCholesterolScores(?p, 1) |
| Rule:23                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>= 160 , <= 199](?TotalCholesterol) -> hasCholesterolScores(?p, 2) |
| Rule:24                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 40 , <= 49](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>= 160 , <= 199](?TotalCholesterol) -> hasCholesterolScores(?p, 3) |
| Rule:25                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 20 , <= 39](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>= 160 , <= 199](?TotalCholesterol) -> hasCholesterolScores(?p, 4) |
| Rule:26                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 60 , <= 69](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=200, <= 239](?TotalCholesterol) -> hasCholesterolScores(?p, 1)   |
| Rule:27                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=200, <= 239](?TotalCholesterol) -> hasCholesterolScores(?p, 3)   |
| Rule:28                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 40 , <= 49](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=200, <= 239](?TotalCholesterol) -> hasCholesterolScores(?p, 5)   |
| Rule:29                                      | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 20 , <= 39](?age), hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>= 200 , <= 239](?TotalCholesterol) -> hasCholesterolScores(?p, 7) |



|  |  |
|--|--|
| Rule:30  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 60 , <= 69](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=240, <= 279](?TotalCholesterol) -> hasCholesterolScores(?p, 2) |
| Rule:31  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=240, <= 279](?TotalCholesterol) -> hasCholesterolScores(?p, 4) |
| Rule:32  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 40 , <= 49](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=240, <= 279](?TotalCholesterol) -> hasCholesterolScores(?p, 6) |
| Rule:33  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 20 , <= 39](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=240, <= 279](?TotalCholesterol) -> hasCholesterolScores(?p, 9) |
| Rule:34  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 60 , <= 69](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol),<br>xsd:integer[>=280](?TotalCholesterol) -> hasCholesterolScores(?p, 3)      |
| Rule:35  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol), xsd:integer[>=280]<br>(?TotalCholesterol) -> hasCholesterolScores(?p, 5)     |
| Rule:36  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 40 , <= 49](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol),<br>xsd:integer[>=280](?TotalCholesterol) -> hasCholesterolScores(?p, 8)      |
| Rule:37  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 20 , <= 39](?age),<br>hasTotalCholesterol(?p, ?TotalCholesterol),<br>xsd:integer[>=280](?TotalCholesterol) -> hasCholesterolScores(?p, 11)     |
| <b>Rules for calculate smoking status scores</b> |  |
| Rule:38  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 60 , <= 69](?age), isSmoker<br>(?p, true) -> hasSmokingStatusScores (?p, 1)  |

|  |  |
|--|--|
| Rule:39  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 50 , <= 59](?age), isSmoker (?p, true) -> hasSmokingStatusScores (?p, 3)   |
| Rule:40  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 40 , <= 49](?age), isSmoker (?p, true) -> hasSmokingStatusScores (?p, 5)   |
| Rule:41  | Patient(?p), hasAge(?p, ?age), xsd:integer[>= 20 , <= 39](?age), isSmoker (?p, true) -> hasSmokingStatusScores (?p, 8)   |
| <b>Rules for calculate HDL level scores</b>              |  |
| Rule:42  | Patient(?p), hasHDLCholesterol (?p, ? HDL), xsd:integer[>= 60](? HDL) -> hasHDLLevelScores(?p, -1)   |
| Rule:43  | Patient(?p), hasHDLCholesterol (?p, ? HDL), xsd:integer[>= 50 , <= 59](? HDL) -> hasHDLLevelScores(?p, 0)  |
| Rule:44  | Patient(?p), hasHDLCholesterol (?p, ? HDL), xsd:integer[>= 40 , <= 49](? HDL) -> hasHDLLevelScores(?p, 1)  |
| Rule:45  | Patient(?p), hasHDLCholesterol (?p, ? HDL), xsd:integer[<40](? HDL) -> hasHDLLevelScores(?p, 2)  |
| <b>Rules for calculate systolic blood pressur scores</b> |  |
| Rule:46  | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[<= 120](? SBP) -> hasBloodPressureScore (?p, 0)   |
| Rule:47  | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 120 , <= 129](? SBP), hasBloodPressureMedication(?p, false) -> hasBloodPressureScore (?p, 0) |
| Rule:48  | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 120 , <= 129](? SBP), hasBloodPressureMedication(?p, true) -> hasBloodPressureScore (?p, 1)  |
| Rule:49  | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 130 , <= 159](? SBP), hasBloodPressureMedication(?p, false) -> hasBloodPressureScore (?p, 1) |

|                                       |   |
|---------------------------------------|---|
| Rule:50                               | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 130 , <= 159](? SBP), hasBloodPressureMedication(?p, true) -> hasBloodPressureScore (?p, 2) |
| Rule:51                               | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 160 ](? SBP), hasBloodPressureMedication(?p, false) -> hasBloodPressureScore (?p, 2)        |
| Rule:52                               | Patient(?p), hasSystolicBloodPressure (?p, ? SBP), xsd:integer[>= 160](? SBP), hasBloodPressureMedication(?p, true) -> hasBloodPressureScore (?p, 3)          |
| <b>Rules for calculate age scores</b> |   |
| Rule:53                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 20 , <= 34](?age) -> hasAgeScores (?p, -9)   |
| Rule:54                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 35 , <= 39](?age) -> hasAgeScores (?p, -4)   |
| Rule:55                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 40 , <= 44](?age) -> hasAgeScores (?p, 0)  |
| Rule:56                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>=45, <= 49](?age) -> hasAgeScores (?p, 3)  |
| Rule:57                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 50 , <= 54](?age) -> hasAgeScores (?p, 6)  |
| Rule:58                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 55 , <= 59](?age) -> hasAgeScores (?p, 8)  |
| Rule:59                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 60 , <= 64](?age) -> hasAgeScores (?p, 10)   |
| Rule:60                               | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 65 , <= 69](?age) -> hasAgeScores (?p, 11)   |

|         |  |
|---------|--|
| Rule:61 | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 70 , <= 74](?age) -><br>hasAgeScores (?p, 12) |
| Rule:62 | Patient(?p), hasAge (?p, ?age), xsd:integer[>= 75 , <= 79](?age) -><br>hasAgeScores (?p, 13) |