POLITECNICO DI MILANO

FACOLTÀ DI INGEGNERIA DEI SISTEMI
Corso di Laurea in Ingegneria Biomedica

Ontological modeling for Neurosurgery: application to automatic classification of temporal and extratemporal lobe epilepsies

Relatore:
Ing. Elena DE MOMI

Correlatori:
Dott. Alberto VACCARELLA
Prof. Andras LORINCZ

Tesi di Laurea di:
Roberta PERRONE
Matricola 765971

Anno Accademico 2011-2012

Ontological modeling for Neurosurgery: application to automatic classification of temporal and extratemporal lobe epilepsies

Roberta Perrone

22 April 2013

Abstract

The development of imaging devices providing multimodal information, in addition to the set of data relative to the specific patient and the surgical procedure has led to a growing amount of information in Computer Assisted Surgery (CAS). Along with the vast and various sources of existing surgical knowledge, e.g. anatomic atlases, medical databases with patient records, documented surgical procedures, etc. and knowledge coming from new devices like systems of image guidance, the quantity of information keeps growing with every new patient. As CAS is becoming more common for various types of surgical interventions (such as neurosurgery, orthopedic surgery, maxillofacial surgery etc.), there is a need for representation, storage and processing of surgical knowledge in a more structured way. In particular, beginning in the latter half of the 20th century, the field of neurosurgery has been transformed by a technological revolution providing increasingly capable and affordable digital computer technology. Moreover, the constant evolving nature of this information has forced the evolution of knowledge representation mechanism as well. It is fundamental and at the same time hard to keep track of all the information and to use it for specific applications.

Medical research can be represented using ontological modeling and used in real time application in Operating Room (OR). Surgical ontologies, or surgical process models, are increasingly used to model clinical practice in terms of discrete steps, thus defining actions that have to be performed by the surgeon, and the patient information that has to be retrieved during the surgical procedure. In this context, surgical ontologies and models are obtained by witnessing a set of interventions, noting each gesture by computer, and fusing these descriptions together. These models can be useful in defining requirments for surgical simulator and helping the surgeons in the choice of the surgical approach for a specific patient. This work is onerous in terms of resource expenditure and has the risk that can be biased by institute-specific practices, and may not characterize general procedures.

On the Web there are many medical ontologies (SNOMED CT , Open-GALEN etc.) that provide taxonomies and relations between concepts that belong to the surgical field.

The general aim of this thesis is to build an epilepsy neurosurgery ontology adopting a more generic approach using a textbook to flesh out concepts included in the neurosurgical domain, which in turn will be complemented by

interviews to surgeons to finalize these descriptions. As CAS for neurosurgery uses computer technology in all the phases (pre-operative, intra-operative, post-operative), in the developed ontology are included actors to each of the three phases, such as instruments that can be used in intra-operative phase in a specific neurosurgical procedure (such as craniotomy), or a list of imaging tests that can be performed in one or more of the three phases. For the specific application of the thesis work, the epilepsy neurosurgery ontology is enlarged with specific concepts to support the diagnosis of temporal lobe epilepsies or extratemporal lobe epilepsies.

The epilepsy diagnosis and neurosurgery ontology is devised in Web Ontology Language (OWL), using Protègè as the ontology editor. The ontology is implemented on a portal (`http://rsp.inf.elte.hu`) developed for the EuRoSurge project (FP7-ICT-2011-7-288233), allowing the registered users to provide their contributions to the ontology (editing, adding or discussing). Because of the possible use of the ontology by anyone (e.g. surgeons with different background and nationality), it is necessary to make clear the concepts that are included and resolve ambiguities since different words of different data resource may refer to the same concept. The Web can be used to connect related concepts that aren't previously connected, with the aim of enlarging and sharing data, information or knowledge with the community. DBpedia (a collection of resources extracted from Wikipedia) is used to connect all the concepts relative to the diagnosis and surgery of epilepsy included in the ontology.

The ontology is used also to model epileptic seizures semiology and to understand the correlation between epileptic symptoms. Temporal lobe epilepsy is a form of focal epilepsy while the extratemporal lobe epilepsies originates outside the temporal lobe. Temporal and extratemporal lobe epileptic seizures induce different symptoms, that can be classified in different ways. A traditional way is to classify objective and subjective symptoms, where subjective symptoms (such as fear and epigastric sensation) differ from others (such as head deviation and eyes deviation) by the inability to physically measure them. In this work, 29 symptoms (14 of which subjective symptoms and 15 of which objective symptoms) are considered and included in the ontology.

Three clinical centers (Besta, Niguarda Ca Granda, San Paolo Hospital) provided the analysis of 109 patients, 60 of which are patients with temporal lobe epilepsies (TLE) while 49 are patients with extratemporal lobe epilepsies (eTLE). Every patient was video registered during epileptic seizure for 60 seconds. In each second, symptoms that appear during the epileptic seizure are set to '1'. Thus a matrix composed of zeros and ones is created for each patient. Starting from the ontological modeling of epileptic seizures and the dataset of TLE and eTLE patients, for each patient a correlation matrix is calculated: this matrix include all the information of correlation between symptoms (i.e. observed symptoms are more correlated than a pair of symptoms in which one is observed and the other one is not observed). Next, a mean correlation matrix (thus a model) for the two set of patients is calculated. A classifier is trained subdividing the dataset in three parts: 70% for training the classifier, 15% for validating and 15% for testing. These models give information relatively to the

occurrence of a symptom and the correlation with other symptoms, in addition to the temporality, thus the time of onset during epileptic seizure. Solution is evaluated with cross-validation method. The performance of classifier, thus the number of patients correctly identified as TLE or eTLE patients, is compared to the performance of 7 clinicians of Besta, Niguarda Ca Granda and San Paolo Hospitals that based their evaluations on the same dataset used to train the classifier. The classifier is tested on a subset of patients randomly extracted from the original dataset. The classifier allows to classify every new epileptic seizure of a patient as TLE or eTLE. With the trained classifier an overall accuracy (mean value of epileptic seizures correctly identified in testing phase) of 74,53 % and a standard deviation of 1,1226 (number of seizures correctly identified that deviate from the average that is 11,18), are reached. Both the algorithm and the clinicians make decisions only on the basis of video registrations of ictal seizures. The accuracy of classification of the algorithm presented in this work is 12,62 % better than the classification made by clinicians (mean).

The ontological modeling, in this sense, help summarizing the information that stands between each symptom with the others, and gives an improvement to the simple analysis made by clinicians on video registrations of epileptic seizures.

A greater evolution of the algorithm could regard the integration of different types of clinical information of patients affected by epilepsies. It is known that for the diagnosis of epilepsy requires careful evaluation of symptoms and clinical history, which should preferably include detailed observations by other persons, since the alteration or loss of consciousness often preclude a description of symptoms by the patient himself. Other diagnostic tests include magnetic resonance imaging or CT and laboratory tests, which can verify or exclude specific causes. With the help of the proposed ontological modeling, different types information could be unified and then processed to obtain new information and more accurate diagnosis.

Keywords: ontological modeling, epilepsy surgery, ictal symptoms, epilepsy classification

Sommario

Lo sviluppo di nuovi dispositivi di imaging ha fornito informazioni multimodali, in aggiunta all'insieme di dati relativi allo specifico paziente e alla procedura chirurgica, ha portato ad una crescente quantità di informazioni nella chirurgia assistita dal calcolatore (CAS). La chirurgia assistita sta rapidamente evolvendo in un dominio a sé stante, con ogni nuovo avanzamento tecnologico nel settore dell'assistenza sanitaria personale. Insieme con le ampie e varie sorgenti di conoscenza chirurgica, come ad esempio atlanti anatomici, banche dati medicali con i dati dei pazienti, interventi chirurgici documentati, ecc. e le conoscenze provenienti da nuovi dispositivi, come i robot e i sistemi di *image guidance*, la quantità di informazioni continua a crescere con ogni nuovo paziente e ogni nuovo intervento.

Dato che la chirurgia assistita da calcolatore è sempre più comune in diversi tipi di interventi chirurgici (ad es. la neurochirurgia, la chirurgia maxillo-facciale, ecc.), vi è la necessità di rappresentare, memorizzare ed elaborare la conoscenza chirurgica in modo più strutturato. In particolare, a partire dalla seconda metà del 20ř secolo, la neurochirurgia è stata trasformata da una rivoluzione che fornisce sempre più una tecnologia efficiente e affidabile grazie all'uso del calcolatore. Inoltre, data la costante evoluzione della natura di queste informazioni è necessaria allo stesso tempo l'evoluzione del meccanismo di rappresentazione della conoscenza. È fondamentale e allo stesso tempo difficile tenere traccia di tutte le informazioni e di usarle per applicazioni specifiche. La ricerca medica, può essere rappresentata utilizzando una modellazione ontologica e essere utilizzata poi, in applicazioni *real time* in sala operatoria (OR). Le ontologie chirurgiche, o i modelli di processi chirurgici, sono sempre più utilizzati per modellizzare la pratica clinica in termini di passi discreti, ovvero definendo gli steps e le azioni che devono essere eseguite dal chirurgo prima, durante e dopo la procedura chirurgica. In questo contesto, le ontologie e modelli chirurgici sono ottenuti assistendo a una serie di interventi chirurgici, registrando ogni gesto con il calcolatore, e fondendo insieme queste descrizioni. I modelli possono essere utili nella definizione di requisiti per un simulatore chirurgico e aiutare poi i chirurghi nella scelta dell'approccio chirurgico per uno specifico caso. Questo lavoro è oneroso in termini di dispendio di risorse e ha il rischio di essere influenzato da pratiche istituto-specifiche, senza caratterizzare le procedure in modo generico.

Sul Web esistono ontologie mediche (SNOMED CT, OpenGALEN, ecc) che

forniscono le tassonomie e le relazioni tra i concetti che appartengono al dominio chirurgico.

L'obiettivo generale di questa tesi è di costruire una ontologia per la neurochirurgia dell'epilessia adottando un generico approccio, ovvero utilizzando un libro di testo per approfondire i concetti inclusi nel dominio neurochirurgico e integrando l'informazione estratta da interviste con i chirurghi per finalizzare le descrizioni dei concetti. Dato la chirurgia assistita utilizza la tecnologia informatica in tutte le fasi (fase pre-operatoria, intra-operatoria, post-operatoria), nell'ontologia sviluppata sono inclusi gli attori che partecipano a ciascuna delle tre fasi, come anche gli strumenti che possono essere utilizzati nella fase intra-operatoria di un intervento neurochirurgico specifico (come una craniotomia), o un elenco di test di imaging che possono essere effettuati in una o più delle tre fasi operatorie. Per l'applicazione specifica di questo lavoro di tesi, l'ontologia per la neurochirurgia dell'epilessia è ampliata con concetti specifici per il supporto alla diagnosi delle epilessie del lobo temporale o epilessie del lobo extratemporale.

L'ontologia per la neurochirurgia dell'epilessia insieme alla sua parte applicativa (di supporto alla diagnosi delle crisi epilettiche) è stato ideata in linguaggio Web Ontology Language (OWL), utilizzando Protègè come editor di ontologie. L'ontologia è implementato inoltre sul portale "http://rsp.inf.elte.hu" (sviluppato per il progetto EuRoSurge, FP7-ICT-2011-7-288233), che consente agli utenti registrati di fornire il loro contributo all'ontologia (apportando modifiche, aggiungendo nuovi concetti o discutendo dell'ontologia con altri utenti). A causa del possibile uso dell'ontologia da parte di chiunque (ad esempio, chirurghi con diverso background e nazionalità), è necessario chiarire univocamente i concetti che sono inclusi e risolvere le ambiguità, dato che parole diverse provenineti da differenti sorgenti di dati possono riferirsi allo stesso concetto. Il Web può essere utilizzato per collegare concetti correlati non ancora connessi, con l'obiettivo di ampliare e condividere dati, informazioni o conoscenze con la comunità. DBpedia, che raccoglie un insieme di risorse estratte da Wikipedia, viene utilizzata per collegare tutti i concetti relativi alla diagnosi e alla chirurgia dell'epilessia inclusi nell'ontologia.

L'ontologia è utilizzata per modellare i sintomi che compaiono durante una crisi epilettica e per comprendere la correlazione esistente tra i sintomi epilettici. L'epilessia del lobo temporale è una forma di epilessia focale, mentre l'epilessia del lobo extratemporale ha origine al di fuori del lobo temporale. Crisi epilettiche del lobo temporale e extratemporale inducono sintomi diversi, che possono essere classificati in diversi modi. Un modo tradizionale è quello di classificare i sintomi oggettivi e soggettivi, in cui i sintomi soggettivi (come la paura e la sensazione epigastrica) differiscono dagli altri (come la deviazione testa e la deviazione degli occhi) per l'impossibilità di essere misurati fisicamente. In questa tesi vengono inclusi nell'ontologia 29 sintomi (14 dei quali sintomi soggettivi e 15 dei quali sintomi oggettivi).

Tre centri clinici (Ospedali Besta, Niguarda Ca Granda, San Paolo di Milano) hanno fornito l'analisi di 109 pazienti, 60 dei quali sono pazienti con epilessia del lobo temporale (TLE), mentre 49 sono i pazienti con epilessia del lobo extratemporale (eTLE). Ogni paziente è stato video-registrato durante una crisi epilettica per 60 secondi. In ogni secondo, sintomi che compaiono durante la crisi epilettica sono impostati a '1 '. Quindi una matrice composta di zero e uno viene creata per ogni paziente. A partire dalla modellazione ontologica di crisi epilettiche e il dataset di TLE e eTLE, viene calcolata per ogni paziente una matrice di correlazione: questa matrice contiene tutte l'informazione di correlazione tra i sintomi (ad esempio i sintomi osservati durante una crisi sono più correlati rispetto ad un paio di sintomi in cui uno compare e l'altro no). Successivamente, una matrice di correlazione media (quindi un modello) per le due popolazioni di pazienti è calcolato. Un classificatore viene addestrato suddividendo il dataset in tre parti: il 70% per la addestramento, il 15% per la validazione e il 15% per il test. Questi modelli forniscono informazioni relativamente alla probabilità che un sintomo compaia durante una crisi e la correlazione con gli altri sintomi, oltre alla temporalità, quindi, il tempo di insorgenza durante la crisi epilettica e sono utilizzati per assegnare ogni nuovo paziente che presenta una certa semiologia ictale ad uno dei due tipi di epilessia. La soluzione viene valutata con il metodo di cross-validazione. L'accuratezza del classificatore, quindi il numero di pazienti correttamente identificati come pazienti TLE o eTLE sul numero di pazienti testati, viene confrontato con l'accuratezza di classificazione di 7 clinici degli Ospedali Besta, Niguarda Ca Granda e San Paolo, i quali hanno fatto le loro valutazioni sugli stessi dati utilizzati per istruire il classificatore. Il classificatore è testato su un sottoinsieme di pazienti estratti casualmente dall'insieme di pazienti originale e permette di classificare ogni nuova crisi epilettica di un paziente come TLE o eTLE. Con il classificatore addestrato viene raggiunta un'accuratezza complessiva (valore medio di crisi epilettiche correttamente identificate in fase di test) del 74,53% e una deviazione standard di 1,1226 (numero di crisi correttamente identificate che si discostano dalla media che è 11,18). Sia l'algoritmo che i clinici hanno tentato di classificare le crisi solo sulla base delle registrazioni video dei pazienti durante le crisi epilettiche. La accuratezza di classificazione dell'algoritmo presentato in questo lavoro è 12,62% in più rispetto alla classificazione fatta da medici (media).

La modellazione ontologica, in questo senso, contribuisce a riassumere le informazioni di correlazione tra ogni sintomo e gli altri, e offre un diagnosi più accurata rispetto alla semplice analisi eseguita da clinici sulle registrazioni video delle crisi epilettiche.

Una maggiore evoluzione dell'algoritmo potrebbe considerare l'integrazione di diversi tipi di informazioni cliniche di pazienti affetti da epilessia. È noto che la diagnosi dell'epilessia richiede un'attenta valutazione dei sintomi e della storia clinica, che dovrebbe preferibilmente includere osservazioni dettagliate da terze persone, poiché l'alterazione o perdita di coscienza spesso preclude una descrizione dei sintomi da parte del paziente stesso. Altri esami diagnostici includono la risonanza magnetica o TC e test di laboratorio, in grado di verificare o escludere cause specifiche. Con l'aiuto della modellazione ontologica proposta, diversi tipi di informazioni possono essere unificate e poi elaborate per ottenere nuove informazioni e diagnosi più accurate.

Parole chiave: modellazione ontologica, chirurgia dell'epilessia, sintomi ictali, classificazione delle epilessie

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Ontological modeling for surgical knowlege

Computer Assisted Surgery (CAS) is becoming more common for various types of surgical interventions (such as neurosurgery, orthopedic surgery, maxillofacial surgery etc.), and it uses computer based technology in all the phases, i.e. *pre-operative* planning, *intra-operative* procedures, *post-operative* analysis. In the pre-operative phase CAS receives data from various sources related to surgical knowledge, which is mostly represented in the form of discrete databases, in particular CAS for neurosurgery involves a guide based on diagnostic images of the patient such as Computed Tomography (CT), Magnetic Resonance Images (MRI), angiography, tractography. This information is merged together with information from brain atlas on which anatomical structures are segmented and labeled, based on the risk of damaging brain functionalities in crossing them.

The domain of CAS also involves various other closely connected areas such as *engineering* of surgical devices, *modeling* of surgical procedures, *assessment* of various methods adopted by different surgeons and the representation of all this knowledge in a computable manner. Along with the vast and various sources of medical knowledge, the information keeps growing with every new patient. Thus, there is the need for representation, storage and processing of surgical knowledge representation system: keeping track of all information and using it for specific application is a very demanding task.

Because of the use of heterogeneous data in CAS application, it is necessary to interface with the latest knowledge representation system, that deals with them in order to integrate and unified them. A relevant open point is the kind of information that CAS applications typically need, and how should that information be represented and processed. Different phases of a neurosurgical procedure require different collection of data. For example, the pre-operative phase involves using medical images, while during the intra-operative phase anatomical and pathological data are used and the post-operative analysis may need need information also from the past diagnostic records of the patient. It is easy to imagine that data are represented in different forms, which makes it difficult for software/hardware tools to understand the meaning of this data and process it accordingly.

In this ambit ontological modeling provides the means to represent conceptual knowledge using expressive formal logics. An ontology is defined as an explicit formal specification of a shared conceptualization, thus a domain of interest can be represented using ontologies by locally defining the concept in a semantically interoperable manner. The great advantage of the ontological modeling is that such a representation is easy to comprehend both for humans and for the machines.

Ontological modeling of a domain of interest requires the definition of actors (i.e. persons or devices), tasks, systems, states and the information or knowledge involved. A possible implementation of this knowledge representation is in a surgical planning software 1.1, allowing the number of surgical steps to be defined for each surgical procedure. In this way, surgeons can access to descriptions of surgical procedures and querying the ontology using patient-specific parameters, they can choose the best surgical approach for the specific case.

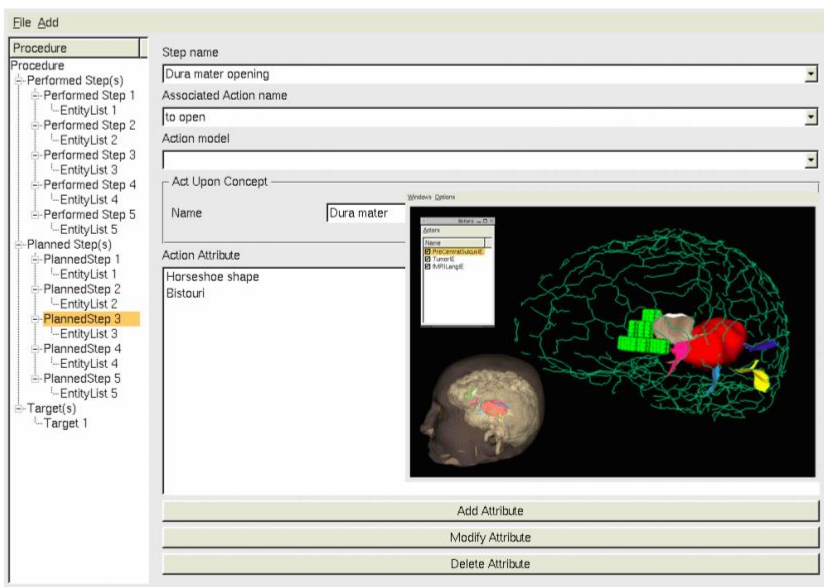

Figure 1.1: Ontology for CAS application. Description of a surgical step (planned) and corresponding 3D scene with relevant images entities.

A formal model can be implemented using ontology languages and also shared with a web application. The aim of creating such a model is to give a meta-level description of the terms and definitions that belong to the chosen surgical application.

1.2 Aim of the work

The aim of this thesis is to design and develop an ontology for epilepsy diagnosis and surgery so that a representation of a domain knowledge will allow to unify all medical records and make them understandable by both humans and machines. The ontological information can be processed to achieve a more accurate diagnosis.

The surgical domain is described from different perspective according to the surgical timeline: pre-operative, intra-operative, post-operative. Each of these aspects are studied at different levels of granularity. Each study focuses on participants to the surgical procedure (i.e. humans and devices), on tasks, actions and other information like the observation of video registrations of epileptic seizures. Ontology is linked to knowledge repositories already present on the Web, in order to make it reliable and uniquely unambiguos: in particular DBpedia was chosen -the ontology behind Wikipedia- to connect resources accepted by researchers communities.

In a second step, then, we will look for a direct application of the proposed ontology relative to the diagnosis of TLE and eTLE : starting from the developed ontology, a classifier that consider only the symptoms that appear during an epileptic seizure is trained.

1.3 Contents

The organization of the thesis is the following:

- In Chapter 2 the overview of the recent knowledge representation language (ontology language) is described; a more detailed description of the medical ontologies and the use of this Knowledge Representation (KR) in surgical model and its application in operating room. In this chapter the exemplary application field of interest (epilepsy diagnosis) is presented;
- In Chapter 3 language Ontology Web Language (OWL), the tools, the basic elements and protocol that helps building a new ontology of domain of interest are presented. The flowchart for the ontology application in epilepsy diagnosis is presented;
- In Chapter 4 results related to the implemented ontology for epilepsy neurosurgery and to the epilepsy diagnosis are presented. Also results related to the application of the part of the ontology that describes ictal semiology (epilepsy diagnosis) to a database of 109 patients are presented;
- In Chapter 5 discussions of results and future works are presented;

Chapter 2

State of the art

2.1 Definition and goal of ontology

Ontologies can be used to model a domain and support reasoning about entities [15]. The term ontology derives from Greek, from "*òntos*" and "*logos*" and it literally means "*speech on being*", thus the study of being, thus of that it is, it exists and it is thinkable.

In computer science and information science, an ontology formally represents knowledge as a set of concepts within a domain, and the relationships between pairs of concepts. It was first introduced in the field of artificial intelligence (AI) as a method for representing and sharing knowledge [45].

Some authoritative definitions state that:

- "Ontology is an explicit and formal specification of a conceptualization", where "explicit" means that typology of concepts and relative restrictions have to be well defined; "formal" means that it has to be machine-readable and "conceptualization" is an abstract and simplified model of the domain of interest that would be represented. Representation of knowledge is based on conceptualization of objects, concepts and other entities that could belong to a certain area of interest and of the relations between them [2];

- "The subject of ontology is the study of the categories of things that exist or may exist in some *domain*. Thus, a rigorous and exhaustive description of a specific domain is inside the ontology, and with the term "domain" an area of knowledge, like for example the healthcare field or genetics area, is meant. An ontology is a catalog of the types of things that are assumed to exist in a domain of interest "D" from the perspective of a person who uses a language "L" for the purpose of talking about "D"" [3].

The above definitions suggest the main steps to build an ontology. They can be summarized as:

1. Identification of the domain of interest;
2. Evaluation of the possibility of re-using existing ontology (at least partially);
3. Identification of basic terms;
4. Definition of taxonomy (hierarchy of concepts and relations);
5. Definition of properties and rules (specifying domain and range);
6. Individuation of new terms resulting from rules;
7. Definition of instances (population of ontology);
8. Identification of errors and inconsistencies.

One of the main purposes of ontologies in computer science is to share knowledge between users and machines; for example if different web sites containing medical information or hospital services, publish and share the same ontology, a software could extract and merge these data to answer to users' questions, or use them as input for further processing [4].

Considering web application, it would be more useful to use standardized languages with the aim of simplifying use of the application, its diffusion and its interoperability with other applications or tools. Building an ontology requires to know basically two languages: Resource Description Framework (RDF) and Ontology Web Language (OWL).

1. RDF, Resource Description Framework, is the basic language for coding, for re-using of structured metadata and for interoperability with other application that exchange machine-readable information [5]. The basic unit to represent information is the statement, or the triple subject-predicate-object, and each element of the triple is a resource. The RDF model can be represented by a directed graph in which the nodes are resources or primitive types and branches are properties [10] (par. 3.1.3, pg. 43);
2. OWL, is a language adopted as standard by W3C (World Wide Web Consortium) from 2002, for the creation of structured ontology based on RDF model; it adds a larger vocabulary in order to describe properties and classes, combining concepts to logic rules (par. 3.2, pg. 45).

2.2 Ontologies on the Web

2.2.1 DBpedia: the ontology behind Wikipedia

DBpedia is an open and collaborative project for the extraction and reuse of information by semantically structured version of Wikipedia. It aims at making reusable Wikipedia information by software applications. The DBpedia allows the user to perform sophisticated and complex queries on the contents of

Wikipedia and to link other datasets on Wikipedia. The DBpedia dataset uses multiple-domain ontology which was derived from Wikipedia.

Currently the English DBpedia describes 3,770,000 "things". DBpedia uses the RDF language as a flexible data model for representing the extracted information and for publishing it on the web. Everything in the DBpedia dataset is indicated by a de-referenceable type Uniform Resource Identifier (URI) `http://dbpedia.org/resource/Name`, where Name is derived from the Uniform Resource Locator (URL) of the source wiki article `http://en.wikipedia.org/wiki/Name` .

In the architecture of DBpedia there is an extractor manager, that is the core component of the overall structure and controls the passage of the Wiki articles to extractors and brings outputs to destination. 11 extractors process the following types of content in Wikipedia. For example:

Labels: all Wikipedia articles have a title that is used as a label for the corresponding DBpedia resource;

Abstracts: a short abstract (first paragraph, represented as `rdfs:comment`) and a longer abstract (the text before a table of contents, using the property `dbpedia:abstract`) are extracted for each item;

Images: links pointing to images (Wikimedia Commons images) are extracted and represented using the property `foaf:depiction`[1];

Redirects: in order to find synonymous terms, the Wikipedia articles can be redirected to other articles;

External links: the articles contain external links that are represented as properties of `dbpedia:reference`. For example: `dbpedia:Providence_Hospital dbpedia:reference http://www.providencehospital.org/`;

Categories: articles are arranged in categories, represented using the Simple Knowledge Organization System SKOS[2] vocabulary. The categories become `skos:concepts`;

In addition to this DBpedia 2.1 is inserted in a more large project named Linking Open Data[3]. The goal of the W3C Semantic Web Education and Outreach group's is using the Web to connect related data that were not previously linked, or using the Web to lower the barriers to link data currently linked using other methods. More specifically, Wikipedia defines Linked Data as "a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information and knowledge on the Semantic Web using URIs and RDF". The goal of Linked Data is to enable people to share structured data on the Web as easily as they can share documents today.

[1] `foaf:depiction` defines a relation between a thing and an image that depicts it. <`http://xmlns.com/foaf/spec/#term_depiction`>

[2] SKOS is a common data model for sharing and linking knowledge organization systems via the Semantic Web.

[3] <`http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData`>

The figure 2.1 shows datasets that have been published in Linked Data, by contributors to the Linking Open Data community project and other individuals and organisations.

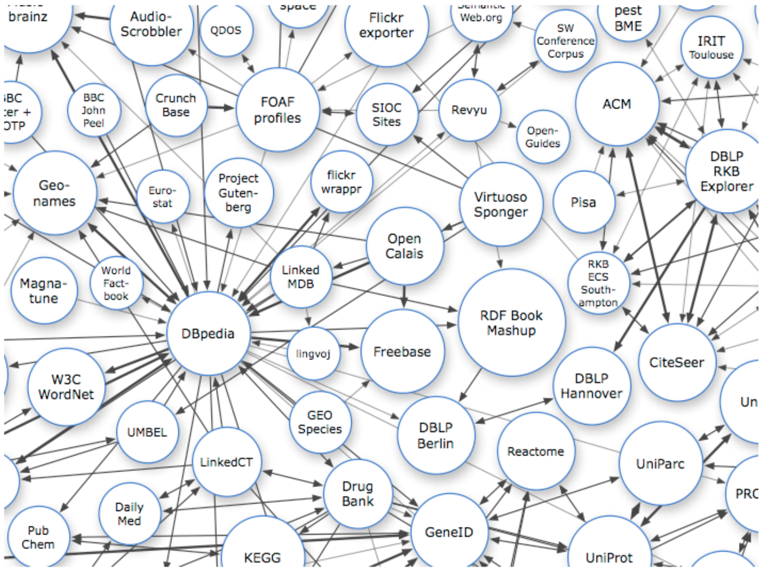

Figure 2.1: Part of the Linking Open Data (LOD) Project Cloud Diagram that shows datasets that have been published by contributors to the Linking Open Data community project and other individuals and organisations.

The basic principles of Linked Data are to:

1. use the RDF data model to publish structured data on the Web;

2. use RDF links to interlink data from different data sources;

The main benefits of using the RDF data model in a Linked Data context are:

- Users can look up every URI in an RDF graph[4] over the Web to retrieve additional information;

- Information from different sources merges naturally;

- The data model enables the user to set RDF links between data from different sources;

- The data model allows the user to represent information that is expressed using different schemata, in a single model;

[4] An RDF graph, or simply a graph, is a set of RDF triples in the form of subject-predicate-object. When the user combine a bunch of triples, one end up with a graph, consisting of subject and object nodes, and predicates that are the connections between them.

- Combined with schema languages such as Resource Description Framework Schema (RDFS) or OWL, the data model allows the user to use as much or as little structure as needed, meaning that one can represent tightly structured data as well as semi-structured data [14];

2.3 Medical ontologies classification

Many ontologies are available on the web in the medical field. They can be divided in three different groups:

1. *Top-level ontologies*: they have categories used in other existing ontologies. Thus they comprise objects, events, part or the whole ontology, instance and classes (for example the General Formal Ontology, GFO [47]);

2. *Domain ontologies*: they apply a top level system to a particular domain: cells, genes, therapies (for example Gene Ontology -GO- [8], the Core Ontology fo CAS -COCAS- [29]);

3. *Application ontologies*: they are focused on a specific task (Functional Endoscopic Sinus Surgery Ontology, FESSOnt [29]);

4. *Reference ontologies*: intermediate layer between domain ontology and top-level ontology. They are independently developed from a particular purpose and serve as modules shared across domains. Technically, reference ontologies are domain ontologies, as they refer to a well-defined context, such as medicine. In spite of everything they can be seen as top-level ontologies for the domain they serve (for example the Surgical Workflow Ontology, SWOnt [29]).

Ontologies allow:

1. Ability to perform advanced software elaborations like automatic reasoning [47];

2. Facilitate the exchange of information between different system.

These advantages are more evident in biomedical ontologies since they allow building more powerful and interoperable health information systems.

In fact all medical research is afflicted by problems of communication and different community of researchers make use of different technologies and most of the time they are incompatible leading to problem of integration of data. By the use of ontologies it would be possible to unify all medical data and make them understandable to anyone.

On the other hand, many people are skeptical about the impact that ontologies can have on the design and maintenance of health information systems: ontologies have some problems. It is often difficult task to design and implement the ontology and, once that it is done, to mantain a consistent logic together with the level of abstraction.

Some examples of application ontologies in medical field are:

1. Electronic health records: they are an electronic version of paper medical records [49];
2. Clinical decision: they support clinicians in decision making tasks, e.g. diagnosis (Computer Aided Diagnosis, CAD) [12, 9];

The next section presents some of the most known medical ontologies available on the Web, together with their structure and basic components.

2.3.1 Medical ontologies examples

Bioportal

As the number of biomedical ontologies increases, so does the number of repositories that index and organize ontologies.

BioPortal[5] is an open repository of biomedical ontologies, developed in 2005 by the National Center for Biomedical Ontology (NCBO), that provides access via Web browsers and through Web services. Users of BioPortal can browse and search the ontologies, submit new versions of ontologies in the repository, comment on any ontology, add a review or make suggestions [50]. The focus on enabling members to contribute actively to BioPortal content and to increase the value of that content to other users distinguishes BioPortal from other ontology repositories.

Available data on BioPortal.com are ontologies for anatomy, phenotype, experimental conditions, imaging, chemistry and health. In the last statistic research, the repository included 329 ontologies (4.849.100 total terms). The full list of ontologies available can be seen at `http://bioportal.bioontology.org/ontologies`. Metadata collected for each ontology include terms, text descriptions, release date, ontology author information and links to documentation. RSS feeds provide alerts of changed content, notes or mappings, new ontology versions and new ontology submissions.

Bioportal supports peer review of ontologies. When a user is evaluating an ontology for a project, a key piece of information is for what other projects have used the ontology and the suitability of the ontology for the tasks of the project. Users can submit descriptions of their ontology-based projects to Bioportal and link these descriptions to BioPortal ontologies.

BioPortal provides also mappings between terms in different ontologies. Users can browse the mappings, create new point-to-point mappings, upload mappings created with other tools, and download the mappings [6].

Developers of Bioportal have added many functionalities to the user interface in order to facilitate the human-computer interaction. Users can filter ontologies by group, such as all the Open Biological and Biomedical Ontologies (OBO) Foundry[6] [7] (e.g. `http://bioportal.bioontology.org/?filter=OBOFoundry`).

[5] <http://bioportal.bioontology.org/>

[6] The OBO Foundry is a collaborative experiment involving developers of science-based ontologies. The foundry is concerned with establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain.

Users can also filter by domain or relevant organism (fig. 2.2).

For what concern the architecture of BioPortal, ontologies may be represented in different languages like OWL, RDF format or the Protègè frame language. BioPortal uses the Mayo Clinic's LexGrid system (`http://informatics.mayo.edu/LexGrid`) to store ontologies in OBO format and to access standard biomedical terminology, such as Unified Medical Language System[7] (UMLS). Protègè[8], an ontology editor, serves as backend for OWL and RDF. BioPortal adopts a layered structure, with the aim of decoupling the logic and domain object models between each layer. The NCBO Bioportal architecture is presented in the figure 2.3.

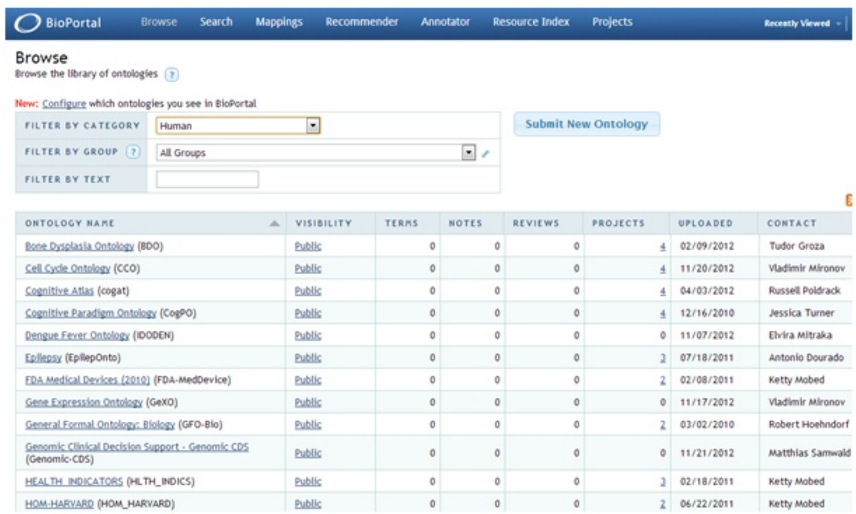

Figure 2.2: BioPortal browser. Users can filter ontologies by category (such as Human), group (collections of ontologies that are designed using similar practices, such as OBO Foundry) or text and directly visualize relsults of filtering. Every result is described with related information like visibility to users, date of uploading and contact to the developer of the ontology.

[7] <`http://www.nlm.nih.gov/research/umls/`>
[8] <`http://protege.stanford.edu/`>

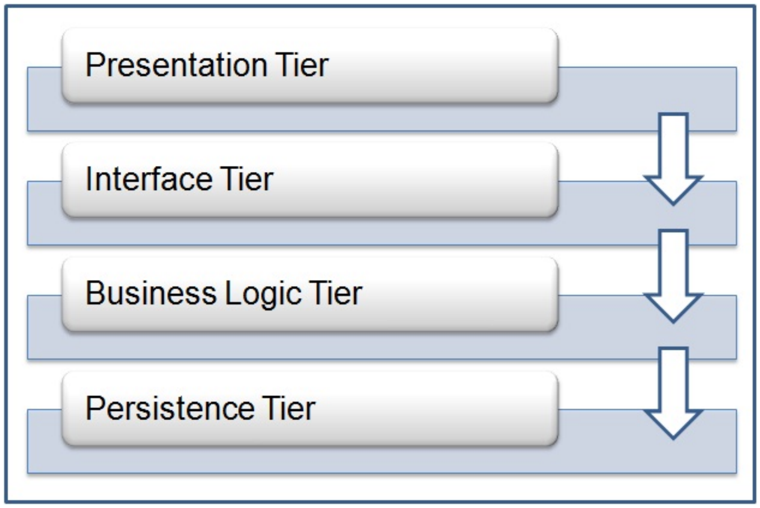

Figure 2.3: NCBO BioPortal architecture. The Presentation Tier delivers the BioPortal user interface. The Interface Tier consists of web services that present all BioPortal functionalities to the upper tiers (e.g. browse, search, visualization, ontology concept display, etc.). The Business Logic Tier provides API access to ontologies and resource index. The Persistence Tier provides a basic object-relational mapping to the back-end relational database.

"The Gene Ontology (GO)[9]project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product (such as ribosome) across species and databases. The project provides a structured, precisely defined, common, controlled vocabulary of terms for describing gene product characteristic and gene product annotation data from GO Consortium members[10], as well as tools to access and process this data". Every concept in the ontology has a precise term, that can be a word or a string of words, a unique alphanumeric identifier, a definition with cited sources, and a namespace which refers to the domain to which it belongs. Terms can also have synonyms, which are classified as exactly equivalent to the name of term (fig. 2.4).

The GO ontology is structured as an oriented acyclic graph, and every term has relations with one or more terms in the same domain, and sometimes to other domains. Three different ontologies are being developed: biological process, molecular function and cellular component.

1. Biological process refers to a biological objective to which the gene or gene product contributes. Examples of broad (high level) biological process terms are 'cell growth and maintenance' or 'signal transduction'. A process is accomplished via one or more ordered assemblies of molecular functions.

[9] <http://www.geneontology.org>

[10] <http://www.geneontology.org/GO.consortiumlist.shtml#full>, the list of current GO Consortium members;

Processes often involve a chemical or physical transformation: something goes into a process and something different comes out of it;

2. Molecular function is defined as the biochemical activity (including specific binding to ligands or structures) of a gene product. This definition also applies to the capability that a gene product (or gene product complex) carries as a potential. It describes only what is done without specifying where or when the event actually occurs. Examples of broad functional terms are ‘enzyme’, ‘transporter’ or ‘ligand’. Examples of narrower functional terms are ‘adenylate cyclase’ or ‘Toll receptor ligand’.

3. Cellular component refers to the place in the cell where a gene product is active. Not all terms are applicable to all organisms; the set of terms is meant to be inclusive. Cellular component includes terms as ‘ribosome’ or ‘proteasome’, specifying where multiple gene products would be found [8].

Goals of Gene Ontology project are:

- creating controlled vocabularies (i.e. terms and definitions);
- producing annotation to terms (i.e. gene products);
- producing GO tools (i.e. browsing, searching and editing);
- making everything publicly available, without any licensing requirements than original source.

The GO is not a static ontology: corrections, updates and suggestions are made directly by community researchers that are involved in GO project.

```
id:         GO:0000016
name:       lactase activity
namespace:  molecular_function
def:        "Catalysis of the reaction: lactose + H2O = D-glucose + D-galactose." [EC:3.2.1.108]
synonym:    "lactase-phlorizin hydrolase activity" BROAD [EC:3.2.1.108]
synonym:    "lactose galactohydrolase activity" EXACT [EC:3.2.1.108]
xref:       EC:3.2.1.108
xref:       MetaCyc:LACTASE-RXN
xref:       Reactome:20536
is_a:       GO:0004553 ! hydrolase activity, hydrolyzing O-glycosyl compounds
```

Figure 2.4: example of GO concept. The lactase activity is represented by a ID number, a definition, a namespace (a container for a set of identifiers (names) that usually group names based on their functionality), synonims and relations with other concepts (i.e. is_a hydrolase activity).

OpenGALEN ontology

The OpenGALEN[11] ontology, is in the field of interest of the present thesis. In 1991 OpenGALEN was formed and non-profit organization for open source

[11] <http://www.opengalen.org>

medical terminology. This terminology is written in a formal language called GALEN Concept Representation Language (GRAIL) and today is possible to translate GRAIL ontologies in OWL-RDF (OWL 2.0) ontologies.

The aim of GALEN is to make as more machine-friendly as possible knowledge and information of biomedical field. It is intended for use by clinical application builders, both when developing clinical applications, and as a run-time resource when those applications are in service. The result is a computerized system of multilingual encoding for use in the medical field to address some key issues such as the problem of having multilingual systems that can be able to preserve and share knowledge underlying to the specific language. Moreover it overcomes the barrier between the high level of detail of information (close to natural language) that is necessary for managing clinical data of patient and aggregate language useful for statistical purpose.

One primary result is the GALEN Common Reference Model, a set of medical concepts re-usable and independent of language used. Common Reference Model is structured in 4 parts (fig. 2.5):

1. High Level Ontology, sometimes called high level schemata, which represents schemes and concepts more abstract and of general use. It contains concepts and broad patterns used for more detailed concepts. The high level ontology of GALEN is independent from the used language used and it is a high-level scheme that allows owner to establish the main rules for collaboration and information sharing;

2. Common Reference Model, that is a collection of easily re-usable definitions and parts of definitions of terms related to anatomy (e.g. Body part), pathologies, symptoms and others. The Common Reference Model itself is broad and shallow and expected to be shared by most applications (although it is highly modular internally and can be subdivided if needed);

3. Subspecialty extensions to manage concepts that need a more detailed level and information for example on specific body regions (e.g. Distal phalanx of the left fourth finger), types of surgery, etc, required in a particular subdomain of surgery, e.g. cardiovascular, respiratory, urologic, etc.;

4. A model of composed concepts that use Common Reference Model and its extensions (e.g. surgical procedure).

There is no sharp line dividing these four levels. There is a continuum between high level concepts such as Body part and highly specific concept such as Distal phalanx of the left fourth finger.

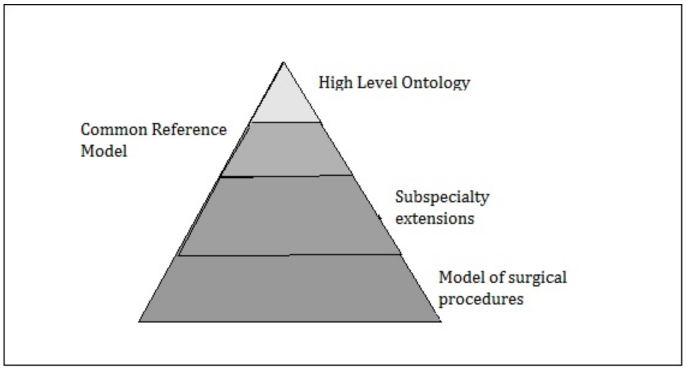

Figure 2.5: Structure of GALEN ontology. High level ontology is on the top of the pyramid and includes abstract concepts that can be used by lower levels. At the base of the pyramid there are the most specific concepts related to surgical procedures and are defined thanks to the composition of concepts belonging to upper level.

GALEN is based on a fundamental principle that is taxonomic separation. In practice, the concepts are described in terms of their components, which are grouped according to pure hierarchies recombined to form all the necessary concepts. The classification of such compositions can be fully automated.

SNOMED CT

Systematized Nomenclature of Medicin-Clinical Terms (SNOMED CT) is the most complete and multilingual clinical sanitary terminology in the world and it is owned, maintained and distributed by the International Health Terminology Standard Development Organisation (IHTSDO). The objective of SNOMED CT is precisely represent clinical information in all fields of application of healthcare: it provides content for the clinical documentation. It is an essential resource for electronic health records, with complete and scientifically validated contents. It is used in about 50 countries but Italy does not appear in this list. SNOMED CT is primary used as:

1. Standard for clinical information: software applications can use concepts, hierarchies and relations in SNOMED CT as common point of reference for data analysis;

2. Base for analysis applications: sanitary assistance can analyze research, evaluate the quality and cost of care and design efficient line guides for a treatment.

In 2011, SNOMED CT includes more than 311,000 concepts. A “concept” has a precise clinical meaning and has defined by a numerical identifier (conceptID). For example the conceptID 556790 identifies the Peribronchial pneumonia which is a disease. The sequence of digits in conceptId does not convey any information about the meaning or nature of the concept. Every concept is represented by:

1. Fully Specified Name (FSN): it provides a human-readable representation;
2. Relations with other concepts: they provide a formal definition to the concept;
3. Collection of terms: they explain the concept in a human-readable way;

Typical approaches to meaningful coding impose limits on both the number of levels of specificity (i.e. the length of the code) and the number of options at each level (i.e. the number of different symbols that can be used in each character position).

Currently the question whether SNOMED CT is an ontology or not is still open: according to official documents released by IHTSDO, SNOMED CT is a "clinical terminology increasingly guided by ontological principles", as subject to continual revisions. SNOMED CT presents the following characteristics, which contribute to its large adoption:

1. Completeness: it covers topics related to health with an unprecedented depth, which allows doctors to record data at the appropriate level of granularity;
2. Scalability: the number of concepts in SNOMED CT continues growing;
3. Multilingualism: although the "International Release" includes language-independent concepts and relations, it incorporates the framework to handle different languages and dialects;
4. Cross maps: provides explicit connections to classifications and coding schema related to health, that are used in all the world (e.g. diagnostic classification as ICD-O3[12])

All SNOMED CT concepts are organized into acyclic taxonomic (is-a) hierarchies (fig. 2.6).

[12] <`http://www.who.int/classifications/icd/adaptations/oncology/en/`>, International Classification of Diseases for Oncology

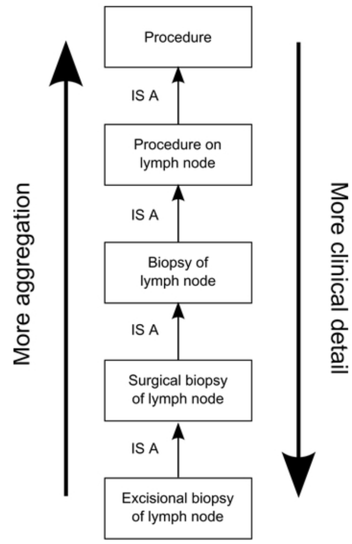

Figure 2.6: Granularity of SNOMED CT concepts. Excisional biopsy of lymph node is more specific than Surgical biopsy of lymph node, Biopsy of lymph node, Procedure on lymph node and Procedure. In taxonomic hierarchies, more specific concepts have finer granularity (more granular) and represent clinical detail, while more general concepts have coarser granularity, represent less clinical detail and aggregate similar concepts;

There are four types of relationships that can be assigned to concepts in SNOMED CT.

1. Defining relationships: used to model concepts and to create their logic definitions. There is a further distinction between:
   (a) Is_a relationship: used to create hiererchies between concepts, thus to crete parent-child relationships;
   (b) Defining attribute: used to link two concepts that belong to different hierarchies and state the type of relation between them (e.g. has finding site, has causative agent);
2. Qualifying relationships;
3. Historical relationships: used to link inactive concepts to active concepts;
4. Additional relationships;

Each concept in SNOMED CT is logically defined through its relationships to other concepts.

Example in fig. 2.7: Bacterial pneumonia is defined as:

- Is_a Infective pneumonia (Clinical finding);

- and has finding site Lung structure of tarsus (Body structure);
- and has causative agent Bacteria (Organism).

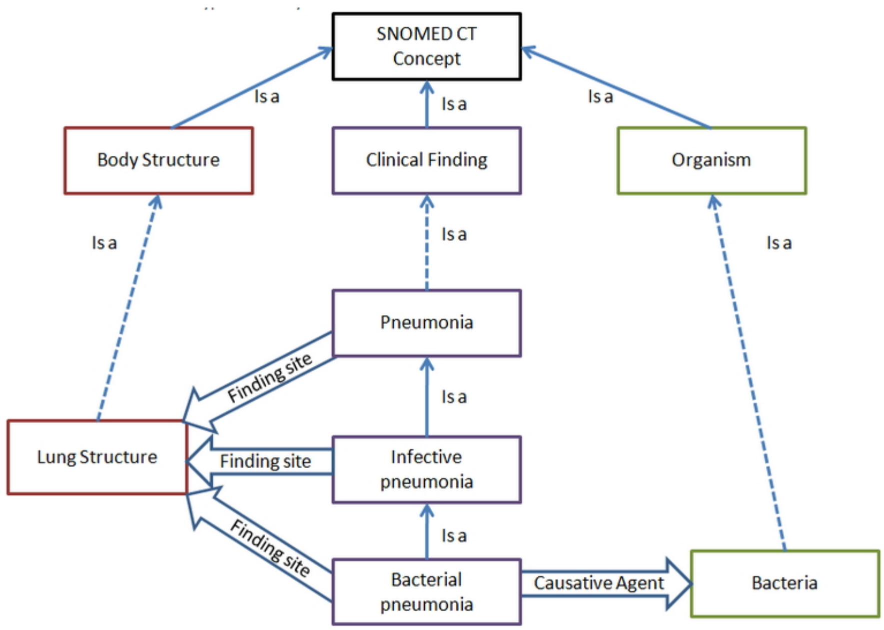

Figure 2.7: Extract of SNOMED CT ontology. Concepts are represented by rectangles, relationships with arrows. Bacterial pneumonia concept is linked to other concepts by defining relationships: in particular, it is linked to the concept Lung structure -that belongs to another hierarchy- by the defining attribute has finding site and it is linked to the more generic concept Infective pneumonia by the parent-child relationship (Is_a).

Another repository of biomedical vocabularies is the Unified Medical Language System (`http://umlsks.nlm.nih.gov`). The UMLS was designed and is maintained by the US National Library of Medicine, is updated quarterly and may be used for free. The project was initiated in 1986 by Donald A. B. Lindberg, M.D., Director of the Library of Medicine. The major component of UMLS is the Methatesaurus, a repository of inter-related biomedical concepts.

The two knowledge sources in the UMLS are the "Semantic Network" providing high-level categories used to categorized every Methatesaurus concept, and lexical resources including the specialist lexicon and programs for generating the lexical variants of biomedical terms.

2.3.2 Surgical workflow ontologies

Computer assisted neurosurgery (CAS) can be improved using surgical models along with patient-specific models built from multimodal images.

P. Jannin and X. Morandi propose a methodological framework for surgical models [46] that includes the definition of a surgical ontology, the development of a planning software for surgical procedure based on this ontology, and the analysis of these descriptions to generate knowledge about surgical practice (i.e. how to perform a procedure for a specific patient).

They defined a surgical workflow as "the automation of a business process in the surgical management of patients, in whole or part, during which documents, information, images or tasks are passed from one participant to another for action, according to a set of procedural rules"[13] [9]. The basic idea was to break down the surgical procedure into steps defining the surgical script. A list of relevant image entities is assigned to each step (i.e. anatomical or functional extracted from pre-operative phase) and the role of each image is annotated. In the figure 2.8 it is presented the Unified Modeling Language (UML) class diagram of the hypothesized ontology.

[13] <http://www.wfmc.org>

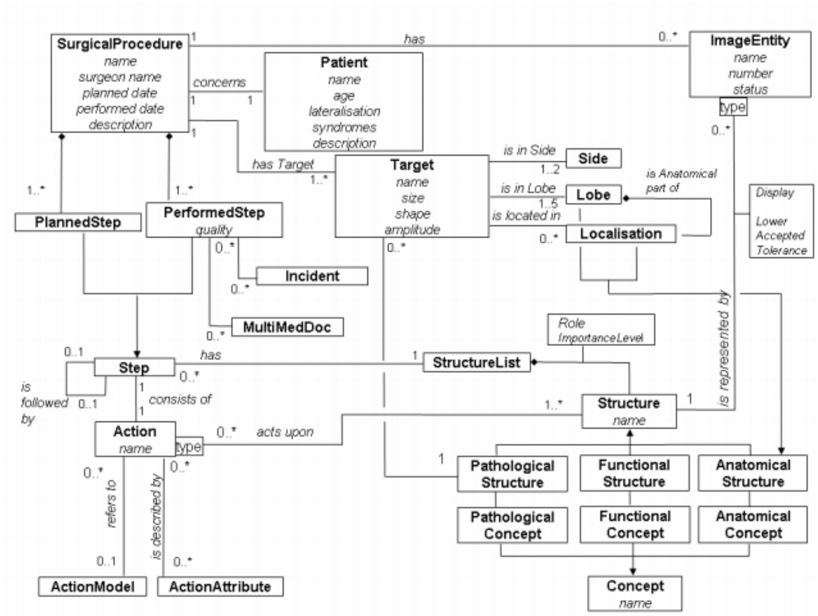

Figure 2.8: UML class diagram of the ontology. Surgical procedure is the basic entity of this ontology. A Surgical procedure is defined with a set of information like the name of the procedure to be performed, the name of the surgeon that operates, the planned date, the performed date and a description. The Surgical procedure is broken down into a sequential list of surgical steps. Each step is described by an action (such as incision) with associated attributes (such as patient positioning). Each structure can be represented by an image extracted from multimodal images of the patient. The role of each image is specified: target area, area to be avoided (such as high-risk functional area or vessels), reference area (such as cortical sulci)[46].

159 patients cases were "described" using the ontology classes derived from post-operative reports. Each class includes a set of parameters. They divide the set of parameters included in the ontology in two subgroups:

1. Predictive, belongs to classes describing the surgical case (e.g. patient information, surgery type, name of pathology, its depth and hemisphere, lobes and gyrus of location);

2. Predicted, belongs to classes describing surgical procedure. These classes include information relative to different surgical steps, the corresponding actions, actions attribute and action models.

The objective of data analysis was to predict the course of surgery, which is described by predicted parameters starting from the description of surgical case (provided by predictive parameters).

This work make use of *Cart* [30], a software for decision tree computation with the aim of understanding the relationship between patient positioning in operating room and the anatomical location of tumors in the subgroup of 159 patients. They randomly divide the dataset in two subgroup: one for training the classification tree and one for testing and evaluating prediction error. Parameters involved in the classification algorithm are:

1. Predictive: values for tumor location (temporal, parietal, occipital and frontal);

2. Predicted: values for patient positioning (dorsal, ventral or lateral decubitus);

They computed a decision tree using a prediction approach giving the probability of patient positioning as a percentage of the learning set and according to tumor location inside one or more lobes. Classification error was 20% when computed on testing set. Results showed the relationship between tumor location inside a lobe and patient positioning in the OR.

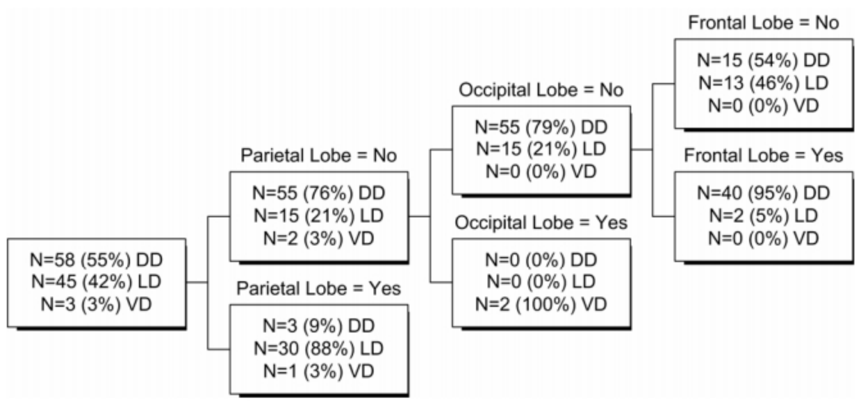

Figure 2.9: Decision tree for patient positioning according to tumor location in one or more cortical lobe for population undergoing brain tumor surgery. DD: Dorsal Decubitus, LD: Lateral Decubitus, VD: Ventral Decubitus.

Having described and stored data relative to workflows of surgical procedure and having implemented this workflow on a planning software, an agent can query the database of surgical cases using patient-specific parameters and access descriptions of surgical procedures performed in similar cases.

Raj Mudunuri *et al.* [29] have presented how CAS knowledge can be represented using ontological modeling, and make use of it in real time applications in Operating Room (e.g. providing information for workflow recording in the OR). This type of work involves description of relevant concepts and roles (attributes) that capture the domain of interest. They divided surgical information into various ontologies at different hierarchical level. Ontologies at the higher level contain concepts that are defined in broader logical expressivity, and ontologies at the lower level are applicable only for a particular surgical discipline for example. The advantage of this top-down and bottom-up approach is that all the concepts that are defined in the whole framework of ontologies follow a logical pattern where concepts of lower level ontologies are subsumed by the concepts of higher level ontologies, and vice versa.

Surgical Ontologies for Computer Assisted Surgery (SOCAS) is an ontological framework that contains various ontologies at different levels of abstraction.

The four ontologies involved in the definition of a ontological modeling for neurosurgical knowledge are of different level of abstraction starting from general to specific:

1. General Formal Ontology (GFO), developed by the Onto-Med research group[14] at the Institute for Medical Informatics, Statistics, and Epidemiology (IMISE) of Leipzig (D). This ontology contains more general concepts, like Process, that can be applied for real world scenarios, for example in the Mudunuri's work the OperativePhase concept -of a surgical procedure- (`OperativePhase isA Process`);

2. The Core Ontology for CAS (COCAS), which contains close to 70 concepts and describe different aspects of surgical knowledge. For example Medical Imaging techniques (e.g. Microscopy-MI, Functional-MI), different types of surgeries (HNO, NeuroSurgery);

3. SWOnt, Surgical Workflow Ontology, that has concepts related to Surgical Workflows;

4. FESSOnt, an ontology with concepts that are related to Functional Endoscopic Sinus Surgery.

Concepts belonging to one of these ontologies are not just linked with simple subsumption relations, but also classified according to different roles between them. For example the Scalpel concept (that belong to the FESSOnt ontology) is linked to the Instrument concept that belong to the SWOnt ontology and the Surgical procedure concept is linked to Scalpel, as in a specific Surgical Procedure, surgeon may need to use a scalpel.

The figure 2.10 shows how the listed four ontologies are connected with each other (hierarchical levels).

[14]OntoMed is a spin-off from Ontonix. Created in mid-2009, the company has the mission of delivering complexity concepts and technology to the medical community. `http://www.onto-med.com/aboutus.html`

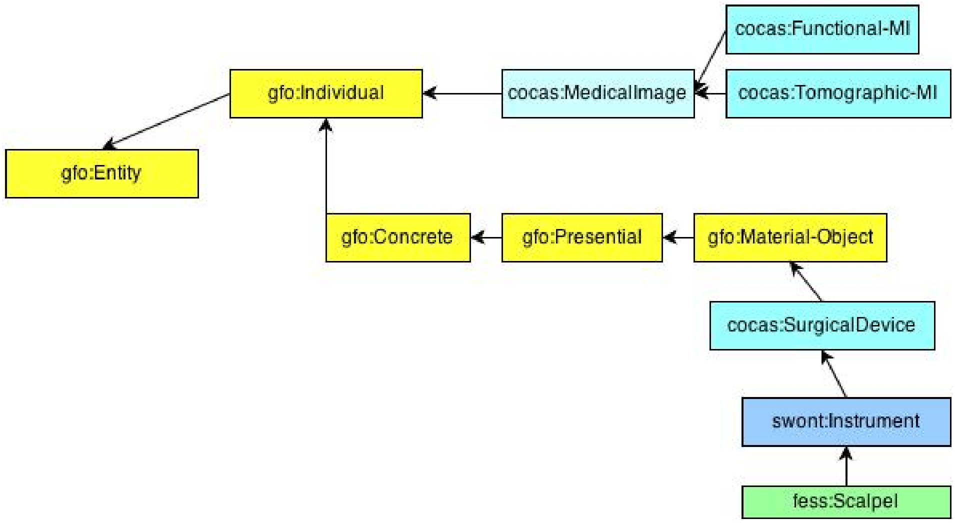

Figure 2.10: SOCAS ontology extract. Each ontology has its own namespace, and each concept is prefixed with this namespace. FESSOnt ontology, the most specific ontology is subsumed to the SWOnt, COCAS ontology and GFO ontology (the most general ontology). SOCAS ontology contains information about the core concepts of CAS (COCAS), and surgical disciplines such as Functional Endoscopic Sinus Surgery (FESSOnt).

An ontology with surgical workflow concepts (SWOnt, 3) acts as a bridge between the core ontology COCAS (2) and the specific discipline ontology (e.g. Functional Endoscopic Sinus Surgery), to propagate the information between them. At the end, these logical connections from top to bottom, and vice versa, which include simple concept subsumptions within the ontology and complex axiomatic relations across the ontologies, form a classified knowledge-based model that can answer, with a particular querying language named SPARQL[15], questions such as:

What are the possible instruments that can be used for a "particular activity" that involves certain "anatomical structures"?

Though it seems that this type of "reasoning" could be answered with a well-built database, the model itself does make a difference. Ontological modeling can be easily updated: facts addition, modification, deletion of concept that became inconsistent with new knowledge can maintain a knowledge base, whereas a database is not meant for revising the model but only to dump the values in the predefined structure tables.

There is an on-going project [41] which has the aim of developing a strategy for elaborating the clinical requirements for a general-purpose neurosurgery simulator, based on surgical ontologies.

[15]SPARQL stands for **S**PARQL Protocol And RDF Query Language, <`http://www.w3.org/TR/rdf-sparql-query/`>

Audette *et al.* used a textbook to flesh out neurosurgical ontologies [41], in particular "*Cranial anatomy and surgical approaches*" by Rhoton A. L. (2007). This group of research involved clinicians in describing surgical processes, ensuring them that their needs are met with the future prototypes. This will lead to a simulation architecture whose organization is based on two nearly parameters:

- The nature of pathology, namely the specific type of tumor as classified by the Central Brain Tumor Registry of the United States[16] (CBTRUS), based on histological types (e.g. glioblastoma, astrocytoma, etc.). In addition, the anatomical site of the pathology, one of 12 CBTRUS categories (e.g. lobes of brain, cerebrum, ventricle, etc.). This correlates strongly with the choice of surgical tools.

- The choice of neurosurgical approach (transnasal, pterional, frontal, occipital and so on), which determines the position and size of the craniotomy, orientation of patient's head, subset of critical tissues that are at risk in the procedure and should be underlined in the anatomical model.

[16] <http://www.cbtrus.org/>

Chapter 3

Materials and Methods

In this chapter language Ontology Web Language (OWL), the tools, the basic elements and protocol that helps building a new ontology of domain of interest are presented. The flowchart for the ontology application in epilepsy diagnosis is also presented.

3.1 Ontologies background

Several ontology languages have been developed during the last few years. Some of them are based on XML syntax, such as Ontology Exchange Language (XOL [37]), SHOE [38] (which was previously based on HTML), and Ontology Markup Language (OML). Resource Description Framework (RDF [39]) and RDF Schema are languages created by the World Wide Web Consortium (W3C). Finally, an additonal language was built on top of the union of RDF and RDF Schema to improve its features (Ontology Web Language) [36].

In the following paragraphs basic components of ontological language will are described.

3.1.1 Uniform Resource Identifiers (URI)

The current Web infrastructure supports a distributed network of web pages that can refer to each other with global links called Uniform Resource Locators (URLs). The special infrastructure of Web provides a data model whereby information about a single entity can be distributed over the Web. Data model for the single application is not held inside the same application but is part of the Web infrastructure. Publishing information for an application, it is not just publishing a human-readable presentation of this information but instead a distributable, machine-readable description of data. The new idea of Semantic Web is to support a distributed Web at the level of data rather than at the level of presentation. In practice, instead of having one web page point to another, one data item can point to another, using global references called Uniform Resource

Identifiers (URIs). A URI provides a global identification for a resource that is common across the Web. If two users on the Web wants to refer to the same resource, it is recommended practice across the Web, to agree to a common URI for the resource. Any two Web application in the world can refer to the same thing by referencing the same URI.

URI can be classified as URL or URN:

URL is an URI that, in addition to identifying a resource, provides instruments to act upon or to obtain a representation describing its primary access mechanism or its "location" in a network;

URN (Uniform Resource Name) that identifies a resource with a "name" in a particular domain of names (i.e. namespace). URN can be used to refer to a resource without using reference to its location or saying how to have access to its representation.

For example in the ontology for epilepsy surgery, "`http://purl.bioontology.org/ontology/SNOMEDCT/42365007`" is used to identify the atonic seizure (disorder), a type of seizure that consist of a brief lapse in muscle tone that are caused by temporary alterations in brain function.

3.1.2 XML, Namespace and XML schema

XML, that stands for eXtensible Markup Language, is a markup language for structured documents. Despite the similarity with HTML that specifies how text and other layout components in a browser should be visualized, XML uses tags only to delimit parts of data, leaving the interpretation to applications that reads them. While HTML has a well-defined and narrow set of tags, XML can define new one according to the requirements of the user.

XML is modular: it allows defining a new format for document combining different formats. In order to avoid confusion on names during union between formats, XML provides a mechanism of namespace. Namespace is a standard, introduced by W3C , that is necessary to make the identification of some elements and attributes in a XML instance. In fact a XML document can import elements or attributes defined in other documents and it can be possible that appear homonyms.

Most of the time an OWL ontology begins with a declaration of the namespaces, such as owl:, rdf: and rdfs: (fig. 3.1).

xmlns:owl ="http://www.w3.org/2002/07/owl#"	It states that the elements in the documents with prefix owl: should be understood as referring to things drawn from the namespace called http://www.w3.org/2002/07/owl#. This is a conventional OWL declaration, used to introduce the OWL vocabulary.
xmlns:rdf ="http://www.w3.org/1999/02/22-rdf-syntax-ns#"	refers to everything that belongs to the namespace at the http://www.w3.org/1999/02/22-rdf-syntax-ns# address.
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"	refers to everything that belongs to the namespace at the http://www.w3.org/TR/rdf-schema/ address.

Figure 3.1: Standard namespaces at the beginning of an ontology.

To these basic declaration, it has been added a list of other namespaces for the epilepsy neurosurgery ontology: foaf:[1], snomed:, dbpedia-owl:, rsp: (fig. 3.2).

xmlns:foaf="http://xmlns.com/foaf/0.1/"	From FOAF ontology we exploit some properties that would be useful to link data of the current ontology to ontology already exsisting and more reliable (i.e. foaf:isPrimaryTopicOf property).
xmlns:snomed="http://purl.bioontology.org/ontology/"	In the epilepsy neurosurgery ontology, some concepts belonging to the SNOMED CT ontology are imported.
xmlns:dbpedia-owl="http://dbpedia.org/ontology#"	In the epilepsy neurosurgery ontology, some concepts belonging to the DBpedia ontology are imported.
xmlns:rsp="http://www.owl-ontologies.com/Neurosurgery.owl#"	It defines the namespace for the epilepsy neurosurgery ontology (i.e. rsp:)

Figure 3.2: Additional namespaces to the epilepsy neurosurgery ontology. The `foaf:`prefix is used to include a property that will allow the linking of two concepts belonging to different ontologies. The `snomed:` and `dbpedia-owl:` namespaces are used to import part of the SNOMED CT ontology and DBpedia repository, recognized by the community of researchers. The last namespace, the `rsp:` prefix, identifies the epilepsy neurosurgery ontology on the Web.

[1] FOAF is a project devoted to linking people and information using the Web (`http://xmlns.com/foaf/spec/`)

3.1.3 Resource Description Framework (RDF)

The Resource Description Framework (RDF) is a family of World Wide Web Consortium specifications originally designed as a metadata[2] data model. Now it is used as a general model for conceptual description or modeling of information that is part of web resources, using different kind of syntax formats. So RDF syntax provides a model for describing resources.

RDF defines a “resource” as any object that is uniquely identifiable by the URI. The properties associated with resources are identified by Property Types, that have specific values. Property types describe the relationships of values associated with resources. A collection of these properties that refers to the same resource is called description.

RDF data model is similar to classic conceptual modeling approaches such as entity-relationship or class diagrams, as it is based upon the idea of making statements about Web resources in the form of subject-predicate-object expressions. These expressions are known as triples in RDF language.

The subject represents the resource, and the predicate denotes traits or aspects of the resources and defines a relationship between the subject and the object. The “surgeon make use of bipolar forceps” in RDF is as the triple: a subject denoting the surgeon (identified by an URI), a predicate denoting “make use of”, and an object denoting “bipolar forceps”.

A collection of RDF statements intrinsically represents a labeled, directed multi-graph. As RDFs and Ontology Web Language (OWL) demonstrate, one can build additional ontology languages upon RDF [32].

Resource may be divided into groups called classes. Members of each class are known as instances of that class. Classes are themselves resource.

The rdf:type property may be used to state that a resource is an instance of a class. In the following list are presented part of the vocabulary useful in the definition of the ontology:
rdf

- **rdf:Property** - the class of RDF properties;
- **rdf:Statement** - the class of RDF statements;
- **rdf:type** – is an instance of rdf:Property used to state that a resource is an instance of a class;

rdfs

- **rdfs:Resource** - the class resource, everything;
- **rdfs:Class** - the class of classes;

[2]The term metadata refers to "data about data". The term is ambiguous, as it is used for two fundamentally different concepts (types). Structural metadata is about the design and specification of data structures and is more properly called "data about the containers of data". Descriptive metadata, on the other hand, is about individual instances of application data, the data content. <`http://en.wikipedia.org/wiki/Metadata`>

- **rdfs:Datatype** - the class of RDF datatypes;
- **rdfs:domain** – is an instance of rdf:Property and it is used to state that any resource that has a given property is an instance of one or more classes;
- **rdfs:range** – is an instance of the class rdf:Property, used to state that the values of a property are instances of one or more classes;
- **rdfs:subClassOf** – is an instance of rdf:Property that is used to state that all the instances of one class are instances of another;
- **rdfs:label** – is an instance of rdf:Property that may be used to provide a human-readable version of a resource's name;
- **rdfs:comment** – is an instance of rdf:Property and it may be used to readable description of a resource. A triple of the form: `A rdfs:comment B`, states that B is a human readable description of A;
- **rdfs:seeAlso** – is an instance of `rdf:Property` that is used to indicate a resource that might provide additional information about the subject resource;
- **rdfs:isDefinedBy** – is an instance of rdf:Property that is used to indicate a resource defining the subject resource [32].

On the other hand RDF Schemas have some limitations. It is not possible to define two classes as equivalent, for example if two classes are defined by two different authors that might used different identifiers in order to define the same concept. Moreover it is not possible to limit the number of properties that can be used (for example every diagnostic image refers to one patient).

3.2 Ontology Web Language (OWL)

Ontology Web Language is currently the most expressive ontology language defined for Semantic Web. There's a long history of ontological language development in computer science. These included language based on HTML, based on XML (called Ontology Inference Layer -OIL-), and various frame-base Knwledge Representation languages.

Developers of OWL choose the DAML[3]+OIL language and try to understand if its semantic expressivity was sufficient to develop usable ontology on the web. A new language emerged as the revision of the DAML+OIL language. OWL can be subdivided in tree subgroups with incremental expressivity, built with the aim of making them usable by specific communities of developers and users: OWL-Lite, OWL-DL and OWL-Full with incrasing level of expressiveness.

In particular:

[3]DARPA Agent Markup Language developed by DARPA in 1999 and it is an agent markup language based on RDF

1. **OWL Lite** is the syntactically simplest sub-language. It is intended to be used in situations where only a simple class hierarchy and simple constraints are needed;

2. **OWL DL** is based on Description Logic (hence the suffix DL). Description Logics are a fragment of First Order Logic and are therefore amenable to automated reasoning. It is also possible to automatically compute the classifcation hierarchy and check for inconsistencies in an ontology that conforms to OWL-DL;

3. **OWL-Full** is intended to be used in situations where very high expressiveness is more important than being able to guarantee the computational completeness of the language. It is not possible to perform automated reasoning on OWL-Full ontologies;

As it can be seen in the figure 3.3,OWL-DL may be considered as an extension of OWL-Lite and OWL-Full an extension of OWL-DL.

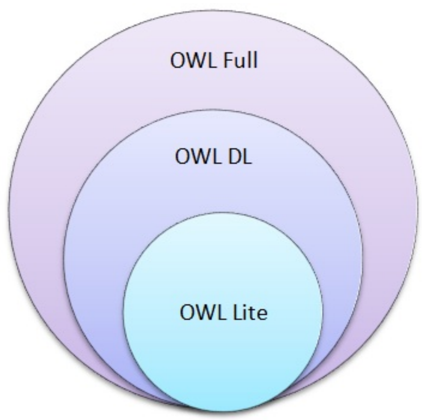

Figure 3.3: OWL Full, DL and Lite

The differences between the three sublanguages are due to differences in the definition of `owl:class` and `owl:ObjectPropertyType` in the three sublanguages of OWL.

1. `owl:class` vs `rdfs:class.` In OWL full, `owl:class` is defined as equivalent to `rdfs:class`. Thus any class that is a subclass of rdfs:class is also a subclass of `owl:class`. As a result, any valid rdf document can be considered as a valid OWL full document. In constrast to owl full, owl:class is defined as subclass of rdfs:class in Lite and DL. Thus not all classes of RDF document (that are subclass of `rdfs:class`) can be an instances or subclasses of `owl:class`. As a result, a valid rdf document cannot be considered as a valid OWL Lite or DL document.

2. `Owl:ObjectTypeProperty` and `rdf:Property` In OWL full,

`owl:ObjectTypeProperty` is considered equivalent to rdf:Property. As a result, `owl:DataTypeProperty`, which is a subclass of `rdf:property`, is also a subclass of `owl:ObjectTypeProperty`. Thus, any property in OWL that is defined as datatype can also be interpreted as objectype property. This provides much of the expressiveness in OWL full. Again in constrast to OWL full, `owl:DataTypeProperty` and `owl:ObjectTypeProperty` are defined as disjoint subclasses of rdf:property. As a result, built-in relations such as "InverseOf", "InverseFunctional", "Symmertric", and "Transitive" cannot be specified as datatype property. This is because a dataType property defines a relation from one individual to a literal (integer, number, date). Thus, inverse of dataType Property does not hold much meaning.

3.2.1 Header ontology

Once that namespaces have been defined, the ontology is filled with a set of statements that refers to the ontology and grouped under a tag called `owl:Ontology,` that starts the ontology file. These tags support such critical and ordinary tasks as comments, version control and inclusion of other ontologies. For example, in epilepsy neurosurgery ontology:

```
<owl:Ontology rdf:about="">

<owl:versionInfo
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
version 1.3</owl:versionInfo>
<rdfs:comment xml:lang="en">This is a seed ontology for...
neurosurgery for epilepsy domain.  It contains classes...
for instruments, surgical procedure and robot used in these.
</rdfs:comment>
<owl:imports rdf:resource=...
..."http://protege.stanford.edu/plugins/owl/protege"/>
<protege:defaultLanguage
rdf:datatype="http://www.w3.org/2001/XMLSchema#string">...
...en</protege:defaultLanguage>
</owl:Ontology>
```

The `owl:Ontology` element is the place where most of the meta-information (i.e. information about the ontology file) can be preserved for the document. It doesn't guarantee that the document describes an ontology in the traditional sense.

When using OWL to describe a collection of instance data the `owl:Ontology` tag may be needed in order to record version information and to import the definitions that the document depends on. Thus, in OWL the term ontology has been broadened to include instance data. The rdf:about attribute provides a name or a reference for the ontology.

When the value of this attribute is null (""), that is the standard case, the ontology name is the basic identifier of the owl:Ontology element. Typically it represents the URI identifier of the document that contains the ontology. Then, rdfs:comment provides the needed capability to comment an ontology. `Owl:imports` gives the possibility to include other document. It provides a mechanism of inclusion and it takes as input a single argument, identified by the rdf:resource attribute. Importing an ontology implies importing all the assertions provided in the current ontology.

3.2.2 Methodology for building an ontology

Building an ontology is a crucial and required step in modeling a work domain [15]. There is no standard methodology for builing an ontology. Individual researchers have taken different approaches. A common methodology can be summarizes in 9 sequential activities:

1. *Planning*. Before starting in building an ontology, a plan should be made about what should be done, who will do what, and what resources are needed. This is a common step for starting any project, though some authors did not mention it in their methodology descriptions.

2. *Specification*. In this activity, the scope and purpose of an ontology should be clarified. Why is the ontology being built? What is its intended usage?

3. *Knowledge acquisition*. Most of the time, developer of ontology has not enough knowledge about the domain of interest to know all terms and their meanings. Thus, knowledge acquisition is fundamental. In this step, it is necessary to locate a set of knowledge sources, such as domain experts, textbooks, and other publications.

4. *Conceptualization*. This activity requires a lot of intellectual thought to structure the domain knowledge. What are the concepts existing in the domain? What are the relations between two concepts? How should concepts and relations be defined? What are the attributes of a concept that help us identify the concept? Once we have responded to these questions we will have a conceptualization of the domain. The final ontology is an explicit specification of the conceptualization.

5. *Integration*. Existing ontology that may help in enlarging and developing the current ontology should be taken in consideration, to avoid duplicating efforts.

6. *Implementation*. In this step we have to choose a computer language and an ontology editor with the aim of implementing the ontology.

7. *Evaluation*. Unlike a KBS (Knowledge Based System), whose evaluation can be done by comparing its performance with an expert's performance, there are no specific methods to evaluate ontologies. Gruber [16] proposed a set of design criteria for ontology development. Some criteria offer a primitive guidelines for evaluating goodness of the current ontology:

(a) Clarity: definition in an ontology should be clear and documented with natural language. Logical axioms should be preferred when possible.

(b) Coherence: both the formal definitions and the informal definitions should be consistent.

(c) Extensibility: an ontology should be designed to be a conceptual framework for a range of anticipated tasks. One should be able to add new terms starting from existing vocabulary. Most of the element of an OWL ontology concerns classes, properties instances of classes, and relationships between these instances.

3.2.3 Classes

Classes – concepts that are also called type, sort, category, and kind – can be defined as abstract objects that are defined by values of aspects that are constraints for being member of the class. Thus a mechanism to describe classes to which individuals belongs to is necessary. At the same time it is important the heritability of properties due to belonging to classes.

It is always possible to define new specific properties that concern individual, but the power of ontological modeling comes from the capability of reasoning based on classes hierarchy. Every individual in the OWL ontology is a member of the `owl:Thing` class.

Thus every new class defined by users will be implicitly a subclass of the default `owl:Thing` class. A class identifier (syntactically represented as a URI reference) describes a class through a class name. In diagnosis and surgery for epilepsy ontology, we define for example the "instrument" class as follows:

```
<owl:Class rdf:about="#instrument"/>
```

Another way to define a class is:

```
<owl:Class rdf:ID="instrument"/>
```

The conceptual difference between `rdf:ID` and rdf:about is that the former declares the concept, while the latter references it. In an RDF/XML file, you can only have one declaration per concept, but you can reference any concept as many times as you like (even if it hasn't been declared).

At this stage we know only about the existence of classes, and once they only exist they could not have members.

The most important construct that is used to define classes taxonomy is `rdfs:subClassOf`. It allows to couple a specific class with a more general class. If X is a subclass of Y, then every instances that belongs to X belong also to Y. The `rdfs:subClassOf` is a transitive relation.

If X is a subclass of Y and Y is a subclass of Z, then X is a subclass of Z.

The definition of a class is divided in two step: an introductive name or a reference and a set of restrictions. Every expression that is in the definition of the class, restricts properties that can be applied to instances that belong to the defined class.

There's also a special class named `protégé:ExternalResource`, which allows to add some external resources (untyped URIs) from repositories available on the Web (e.g. DBpedia).

3.2.4 Properties

Properties allows to assert general facts about classes' members and to assert specific facts about individuals. In OWL language, there are two types of property:

1. Datatype property. Relationships that link individuals to data values;
2. Object property. Relationships that link individuals to individuals.

An object property is defined as an instance of the built in OWL class `owl:ObjectProperty.` A datatype property is define as instance of the built-in class `owl:datatypeProperty`.

Both `owl:ObjectProperty` and `owl:datatypeProperty` are subclasses of the RDF class rdf:Property.

A property axioms defines characteristics of a property. In its simplest form, a property axiom just defines the existence of a property. For example in the *Neurosurgery* for epilepsy ontology: `<owl:ObjectProperty rdf:ID="hasTarget"/>` This axioms defines a property with the restriction that its values should be individuals.

Moreover, property axioms define additional characteristics of properties. OWL supports the following constructs for property axioms:

- RDF Schema constructs like: `rdfs:subPropertyOf`, `rdfs:domain` and `rdfs:range`;
- Relations to: `owl:equivalentProperty` and `owl:inverseOf`;
- Global cardinalities constraints: `owl:FunctionalProperty` and `owl:InverseFunctionalProperty`;
- Logical property characateristics: `owl:SymmetricProperty` and `owl:TransitiveProperty`;

rdfs:subPropertyOf: this kind of property defines that the property is a subproperty of some other property.

Formally this means that if P1 is a subproperty of P2, then the property extension of P1 (a set of pairs) should be a subset of the property extension of P2 (also a set of pairs).

An example extracted from epilepsy neurosurgery ontology:

```
</owl:ObjectProperty>
```

```
<owl:ObjectProperty rdf:about="#hasPrecedingProcedure">
<rdfs:domain rdf:resource="#surgicalStep"/>
<rdfs:subPropertyOf rdf:resource="#hasTemporalProcedure"/>
</owl:ObjectProperty>
```

These code lines state that all instances (pairs) contained in the property extensions of the "hasPrecedingProcedure" are also members of the property "hasTemporalProcedure".

Subproperty axioms can be applied to both datatype properties and object properties. In OWL DL the subject and the object of a subproperty statement must be either both datatype or both object properties.

rdfs:domain. For a property the user can define rdfs:domain axioms. The rdfs:domain is a built-in property that links a property (instance of the class `rdf:Property`) to a class description. This kind of axiom asserts that the subject of such property statements must belong to the class extension of the indicated class description. For example:

```
<owl:DatatypeProperty rdf:ID="size">

<rdfs:domain rdf:resource="#instrument"/>
</owl:DatatypeProperty>
```

If we are interested in saying that multiple classes can act as domain, we should use a class description of the owl:unionOf form.

```
<owl:DatatypeProperty rdf:ID="name">
<rdfs:domain>
<owl:Class>
<owl:unionOf rdf:parseType="Collection">
<owl:Class rdf:about="#Surgeon"/>
<owl:Class rdf:about="#Assistent"/>
<owl:Class rdf:about="#Student"/>
<owl:Class rdf:about="#Patient"/>
</owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:DatatypeProperty>
```

rdfs:range. For a property the user can define (multiple) rdfs:range axioms. It is a built-in property that links a property (some instance of class `rdf:Property`) to either a class description or a data range. An `rdfs:range` axiom assert that the values of this property must belong to the class extension of the class description or to data values in the specified data range.

Multiple range restrictions are interpreted as stating that the range of a property is the intersection of all ranges.

Similarly to `rdfs:domain`, multiple alternative ranges can be specified using the `owl:unionOf` construct.

```
<owl:ObjectProperty rdf:about="#actsUpon">

<rdfs:domain rdf:resource="#action"/>
```

```
<rdfs:range rdf:resource="#bodyPart"/>
</owl:ObjectProperty>
```

owl:equivalentProperty. In order to link a set of ontologies with the aim of building a third ontology, usually it is useful the fact that we can say that a property (or a class, `owl:equivalentClass`) in one ontology is equivalent to a second property (or a second class) in a second ontology.

We report an example of the use of `owl:equivalentClass`, used to unify two definitions that belong to two different ontology:

```
<owl:Class rdf:ID=”Y”>

<owl:equivalentClass rdf:resource=”&01;X”/>
</owl:Class>
```

It is important to note that property/class equivalence is not the same as property/class equality. Equivalent properties/classes have the same “values” (i.e., the same property/class extension), but may have different intensional meaning, i.e. denote different concept.

Property equality should be expressed with the `owl:sameAs` construct. As this requires that properties are treated as individuals, such axioms are only allowed in OWL Full. On the other hand in the epilepsy neurosurgery ontology we make use of the `owl:sameAs` construct to link equal individuals that belong to different ontology and different classes. owl:inverseOf. Properties have direction from domain to range.

Most of the time user find useful to define relations in both direction (surgical procedures have complications, some complications are complications of surgical procedures). If a property P1 is defined as the inverse of the P2 property, for every x,y, P1(x,y) if and only if P2(y,x).

```
<owl:ObjectProperty rdf:ID="isComplicationOf">

<owl:inverseOf>
<owl:ObjectProperty rdf:ID="hasComplication"/>
</owl:inverseOf>
</owl:ObjectProperty>
```

owl:FunctionalProperty. A functional property is a property that can have only one (unique) value y for each instance x, i.e. there cannot be two distinct values y1 and y2 such that the pairs (x,y1) and (x,y2) are both instances of this property. Both object properties and datatype properties can be declared as “functional”. OWL defines the built-in class owl:FunctionalProperty as a special subclass of the RDF class rdf:Property.

3.2.5 Restrictions on proprieties

In addition to characteristics specified in the previous paragraph,it is possible to limit the interval of values that can assume a property in specific contexts, in different ways. `owl:allValuesFrom` is a built-in property that links a restriction class to either a class description or a data range.

The constraint is used to describe a class of all individuals for which all values of the property under consideration are either members of the class extension of the class description or are data values within the specified data range.

```
<owl:Class rdf:ID="CT">

<rdfs:subClassOf>
<owl:Class rdf:ID="Imaging_test"/>
</rdfs:subClassOf>
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty>
<owl:ObjectProperty rdf:ID="hasInterventionPhase"/>
</owl:onProperty>
<owl:allValuesFrom>
<owl:Class>
<owl:unionOf rdf:parseType="Collection">
<owl:Class rdf:ID="preOperativePhase"/>
<owl:Class rdf:ID="intraOperativePhase"/>
</owl:unionOf> </owl:Class>
</owl:allValuesFrom>
</owl:Restriction>
</rdfs:subClassOf>
</owl:Class>
```

This constraint is analogous to the universal (for-all,) quantifier of predicate logic, for each instance of the class that is being described, every value for P must fullfil the constraint. `owl:someValuesFrom`. It is a built-in property that links a restriction class to a class description or a data range.

A restriction of this kind describes a class of all individuals for which at least one value of the property concerned is an instance of the class description or a data value in the data range. It defines a class of individuals x for which there is at least one y such that the pair (x,y) is an instance of P.

This does not exclude that there are other instances (x,y') of P for which y' does not belong to the class description or data range.

```
<owl:Class rdf:about="#bodyPart">

<rdfs:subClassOf>
<owl:Restriction>
<owl:someValuesFrom>
<owl:Class rdf:about="#Imaging_test"/>
</owl:someValuesFrom>
<owl:onProperty>
<owl:ObjectProperty rdf:ID="isRepresentedBy"/>
</owl:onProperty>
</owl:Restriction>
</rdfs:subClassOf>
```

```
</owl:Class>
```

The `owl:someValuesFrom` constraints is analogous to the existential quantifier of Predicate logic, that states that for each instance of the class that is being defined, there exists at least one value for P that fulfills the constraint.

owl:cardinality. This property allows to exactly specify the number of elements in a relationship. Positive integer numbers (different from 0 and 1) are allowed in OWL DL. `owl:maxCardinality` can be used to define a superior limit, on the other side `owl:minCardinality` is used to define a inferior limit. They can be combined to state an interval of values.

owl:hasValue. This construct allows to define classes on the basis of the existence of particular values of the property. However, an individual will be a member of a class if at least one of the values of its property is equal to the one defined by the `owl:hasValue` construct.

owl:disjointWith. This construct is used to assert that the class extensions of the two class descriptions involved have no individuals in common. Like axioms with `rdfs:subClassOf`, declaring two classes to be disjoint is a partial definition: it imposes a necessary but not sufficient condition on the class.

In order to say that a set of classes is mutually disjoint, there must be an `owl:disjointWith` assertion for every pair.

3.2.6 Individuals

In addition to classes, we described also their members. Individuals, also named instances, are the basic, "ground level" components of an ontology.

The individuals in an ontology may include concrete objects such as people, instruments and body part as well as abstract individuals such as action, pathology and images.

Strictly speaking, an ontology need not to include any individuals, but one of the general purposes of an ontology is to provide a means of classifying individuals, even if those individuals are not explicitly part of the ontology.

For defining an individual we have to use the `rdf:type` that links an individual to a class to which it belongs.

3.2.7 Equivalence

It would be useful to link ontology, maybe two predefined ontologies, or a new small ontology with others that are large and accepted by the community of researchers, with the aim of creating a third one.

It is frequent the fact that it has to be said that a particular class or a property in one ontology is equivalent to a class or a property in a second ontology. An example could be the one in which we depends on the useage of two ontologies separately developed and in which they use a different URI identifiers, o1:X and o2:Y, to refer to the same class. The owl:equivalentClass property can be used to collapse these definitions so that derivations obtained from the two ontology are effectively combined with each other. The lines of code to express these property are:

```
<owl:Class rdf:ID=”Y”>

<owl:equivalentClass rdf:resource=”&O1;X”/>
</owl:Class>
```

In order to link properties in a similar way, it might be used the `owl:equivalentProperty`. For what concern the equivalence between individuals, it might be useful to declare that two individuals (also belonging to different ontologies) are the same. It is frequently used to support Linked Data integration via declaratively interconnecting "equivalent" resources across distributed datasets. Prior to rise of owl:sameAs, the `rdf:seeAlso` property was heavily used in linking Friend of a Friend (FOAF) data: it links from one FOAF to another in which additional property can be found. More recently, `owl:sameAs` has been widely used in linked data datasets, such as DBpedia , and it provides an alternative way to refer to an external equivalent resource.

In the epilepsy neurosurgery ontology there are many of these declarations, because of the need to link the ontology created (seed ontology) to bigger ones that have been edited, integrated and modified by the contributions of worldwide community (including also experts). An example that can be find in the ontology is:

```
<rsp:suctionTube rdf:ID="YankauerSuctionTip">

<rdfs:label ...
...rdf:datatype=...
..."http://www.w3.org/2001/XMLSchema#string">...
...Yankauer suction tip</rdfs:label>
<owl:sameAs>
<rdf:Description...
...rdf:about="http://dbpedia.org/resource/Yankauer_suction_tip">
<rdfs:subClassOf ...
...rdf:resource=...
..."http://protege.stanford.edu/plugins/owl/protege#ExternalResource"/>
</rdf:Description>
</owl:sameAs>
```

In this example it is said that "YankauerSuctionTip" defined in neurosurgery ontology (rsp:) and "Yankauer_suction_tip" resource of DBpedia are the same thing. This is an important point: OWL has not unique name assumption and this means that if two names are different, it does not mean that have to refer to different individuals.

It has to be noticed that the `owl:sameAs` property must be used if and only if anything that can be said for the individual A is true for B and viceversa . In OWL Full the owl:sameAs property can be used to equal everything: a class and an instance, a property and a class, and it consider both the elements as instances.

On the contrary in OWL DL both the elements must be instances.

3.2.8 Set operators

The extensions of OWL class are some sets composed of instances that are members of that class. OWL provides instruments to manage the extensions of a class using base operators, that can be seen as the logic operators AND, OR, NOT. These tree class descriptors are the most advanced constructors used in Description Logic (thus they will be supported only by OWL DL and OWL Full).

Union. The following example shows the use of the owl:unionOf property.

```
<owl:Class rdf:ID=''Manifestation''

<owl:unionOf rdf:parseType=''Collection''>
<owl:Class rdf:about=''#subjectiveManifestation'' />
<owl:Class rdf:about=''#objectiveManifestation'' />
</ owl:unionOf>
</ owl:Class>
```

The Manifestation class include both the extensions of subjectiveManifestation and objectiveManifestation (subjective and objective ictal symptoms).

Intersection. The following example demonstrate the use of the `owl:intersectionOf` property.

```
<owl:Class rdf:ID=''A''>

<owl:intersectionOf rdf:parseType=''Collection''>
<owl:Class rdf:about=''B'' />
<owl:Restriction>
<owl:onProperty rdf:resource=''#C'' />
<owl:hasValue rdf:resource=''#D'' />
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
```

The presented construction states that the class A is exactly the intersection of the B class with the set of things that have value D. This means that if something has value D on the property C, then it will necessarily be instance of the A class.

3.3 Building the ontology

3.3.1 Protégé

Protègè [4] is an application developed by the Stanford University (Palo Alto, California, USA) in collaboration with the Manchester University (United Kingdom) in 1987.

It is 'free and open-source platform', that supports an always growing community of users (in 2013, 219155) for building domain models and application

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

based on knowledge using ontologies[5] such as models building, data inclusion, visualization and modification of the already exsisting models.

The development of Protègè has historically been mainly driven by biomedical applications, but the system preserves its domain-independent characteristic. Protègè is entirely written in Java and it is structured in plug-in: they are grafted to the central module of the application that developers call Protègè core. Thus, developers can extend the core with the addition of new functionalities, without altering the original source code. In the figure 3.4 the homescreen of Protègè is presented: every new project has an URI that uniquely identifies the ontology over the Web in addition to a list of basic characteristic like the default language.

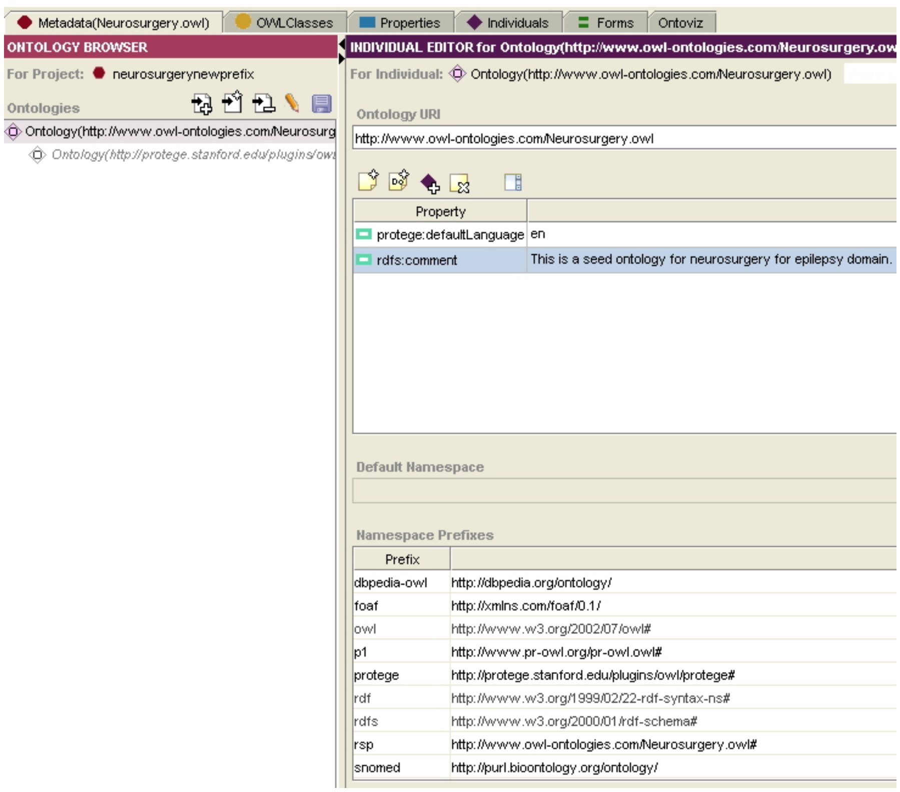

Figure 3.4: Screen of Protègè 3.4.8. On the top of the figure there are tabs for inserting classes, properties and individuals. Metadata tab: on the left ontologies included in the project are visualized (e.g. epilepsy neurosurgery ontology). In the center (from top to bottom): the ontology URI, that uniquely defines the epilepsy neurosurgery ontology, the definition of the default language of the ontology and a comment that introduces the field of interest of the ontology and a list of namespaces of ontologies cited in the project.

[5] <(`http://protege.stanford.edu/overview/index.html`>

Protégé is released under Mozilla Public License (MPL), which is a free license. This license is extended to plug-ins developed by Protègè team (such as Protègè-OWL). Protégé 3 is structured in a modular way. The base module is named Protègè core and provides main functionalities like creation, visualization and manipulation of ontology. A support for OWL has been included in a plug-in that can be separately installed.

Protègè core provides the main graphic interface (GUI) and modelization APIs. The plug-in of Protègè-OWL includes a set of APIs that exploits reusable interfaces (for example classes, proprieties, instances or individuals). At the same time Protègè-OWL does not have a graphic interface but it uses the same proposed by the core module (in fig. 3.5 Protègè Core System) adding some elements, in the form of tab or widget. Protégé-OWL make use of Jena APIs[6] to parse .owl file and as instrument of reasoning and supporting queries in SPARQL language. In figure 3.5 the structure of Protègè-OWL is described.

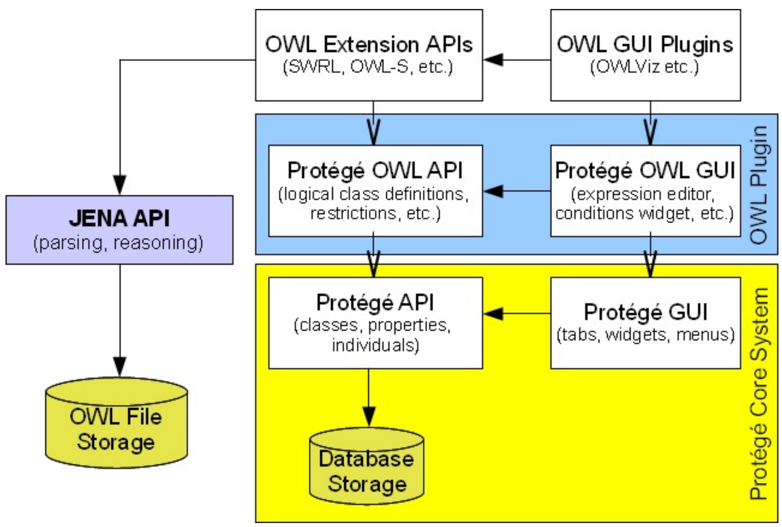

Figure 3.5: Protègè-OWL architecture

At this time Protègè developers are moving in two directions, thus they are working on two different release. Protègè 3 is the more mature release and it sometimes undergoes some changes like improvements, optimizations and bugs corrections, and rarely addition of functionalities that bring to releasing of new version. The most recent version is the 3.5, released in May 2012.

The other section on which Protègè developers are working is named Protègè 4. It appears as a completely new software that is developed to answer to the same needs that have encouraged the development of the predecessor, with the

[6] Jena™ is a Java framework for building Semantic Web applications. Jena provides a collection of tools and Java libraries to help you to develop semantic web and linked-data apps, tools and servers. The Jena Framework includes an API for reading, processing and writing RDF data in XML.

aim of supporting new standards. However it has many gaps and for this reason most of the time the stable Protègè 3 is preferred. Protégé team encourage anybody who have developed and or want to extend application's functionalities, both for the 3. and for the 4. versions, allowing designer to access to database that collects all plug-in that have been released since now and to allow them to upload their work on the database. At the `http://protegewiki.stanford.edu/wiki/Protege_Plugin_Library` address it is possible to see the list of all the available plug-in.

3.3.2 Protégé client-server

Existing applications of Protègè server are essentially two: Collaborative Protègè and Webprotege. The first one is an extension of Protègè for collaborative editing of shared ontologies, that allows users to annotate changes that they have made (thus they write in natural language with the aim of helping other users to understand changes introduced)[51]. Colaborately protégé application is allowed to interact with one at a time.

Users can collaboratively work to the same ontology stored on a Protègè server. The interaction between client and server is based on Remote Method Invocation (RMI) of Java and it is centrally controlled. A special ontology, named metaproject (fig. 3.6), controls the politic of accessing, registering users that can have access to the system, objects -usually ontologies- that have be accessible (i.e. sharable) and finally access policies associated to every shared object. An example of a metaproject can be viewed in the following figure.

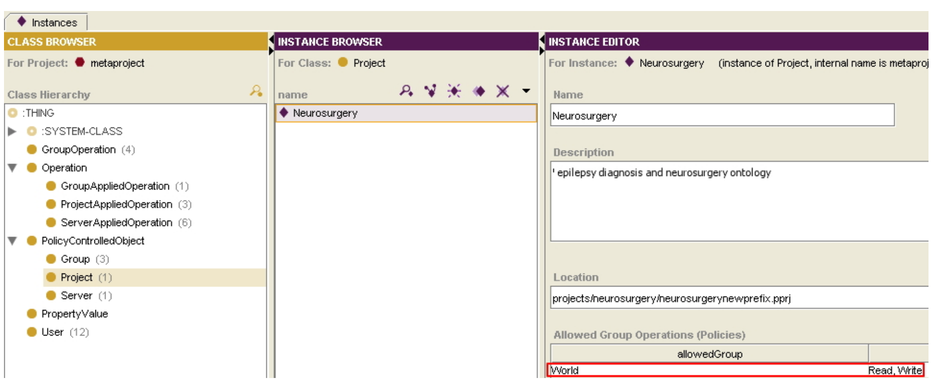

Figure 3.6: Metaproject including Neurosurgery epilepsy ontology. Metaproject, included in the Protègè server side, contains information about which Protege projects are exported, i.e. available to Protege clients, and which users have access to these projects. It also stores policy configurations (in red rectangle). In metaproject in figure, the epilepsy neurosurgery ontology is shared and accessible to users: it is defined by a name, a short a description, the location of the project and the access policies.

Webprotege is a web-based lightweight ontology editor for the web that uses Protègè as backend. For developers the goal of this application is to better support the collaborative development process in a web environment. For example it provides support for the simultaneous editing, thus a change by a user is immediately seen by the other users.

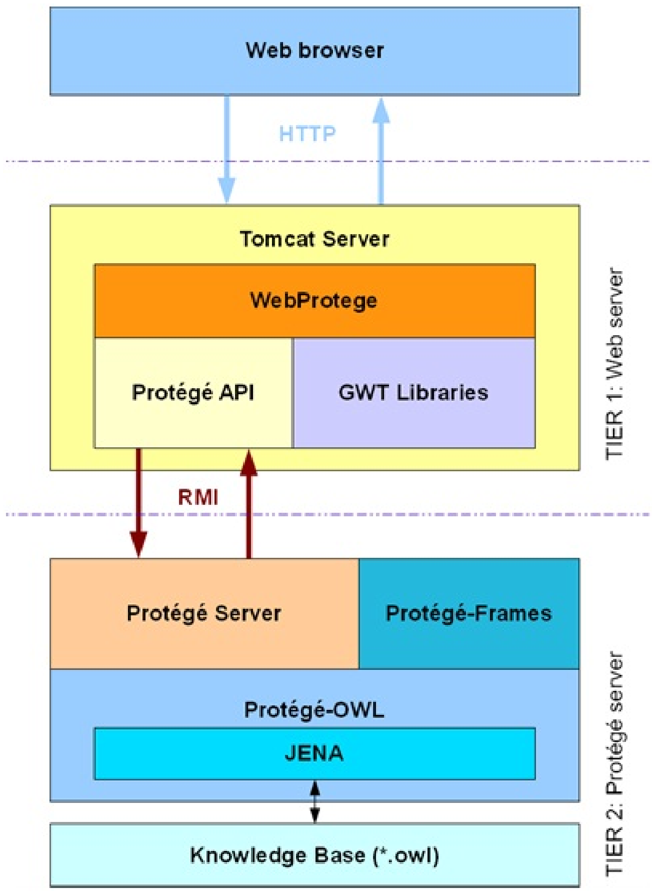

Figure 3.7: Architecture of Webprotege.

Webprotege implements a three levels client-server architecture 3.7. The inner level (Protègè server) provides ontology-access services through the Ontology Jena API (currently Protègè-OWL 3.x API). This Java API contains methods for reading and writing OWL ontologies. Most of the business logic is implements in the intermediate level (Web server). Webprotege uses APIs of Protègè 3 to access to the ontology and to interact with it, interacting with the server through Remote Method Invocation (RMI), while it uses the Google Web Toolkit (GWT) to implement the user interface which is a development toolkit for building and optimizing complex browser-based applications. The outer level is simply made of the web browser of the user.

Once that the user is registered to webprotege and he/she has logged in, he/she can see the list of the available ontologies on which he/she can to make changes (these authorization are inserted in the metaproject side server). With Webprotege is possible to interact with multiple ontologies at the same time.

3.4 Ontology on Robotic SurgePedia (RSP)

Robotic SurgePedia [7] (RSP) is a community platform developed by the team of ELTE University of Budapest, to identify available competences and potential missing topics in the field of robotics assisted surgery with the help of collective intelligence. Robotic SurgePedia propose an intelligent encyclopedia, or repository of concepts that:

1. Is kept up-to-date by the community and natural language processing (NPL) methods;
2. Provides a bridge between different disciplines, between the experts themselves, and the public.

Robotic SurgePedia takes inspiration from Wikipedia and Scholarpedia. Drafts (i.e. new pages in RSP) are written in wiki text and everyone that is registered to the RSP portal can contribute to these pages like in Wikipedia while every draft has to be promoted passing a voting system, in which all contributors have to give their consense to the completeness and clearness of the draft and a peer review like in Scholarpedia. If the draft pass the voting system and the peer review it becomes an article.

RSP is fostered by machine recommendation systems and disseminates knowledge for non-experts, also by translating and simplifying expert texts. Potential patients get qualified, cost-effective and easy-to-understand solutions and answers without personal contact, can share information with others with the same problem, and can participate actively in searching for the solution. Medical experts/surgeons receive new high quality information quickly to update their knowledge, gain recognition, visibility, and patients. They can realize underlying hidden problems of their patients as well. Pharmaceuticals, business-oriented partners in the fields of surgery, robotic developments, etc. can prove and advertise methods, discover new directions and ideas for developments, get potential clients and business partners. One of the goal is to develop a Robotic SurgePedia ontology. Using Webrotege (3.3.2), the *diagnosis and surgery of epilepsy ontology* is uploaded on the RSP portal.

Every concepts included in the ontology is checked also on Wikipedia and DBpedia repository and if a concept is already present and described in them, then on RSP it will be simply connected with the OWL properties owl:sameAs (for DBpedia concept) and foaf:isPrimaryTopicOf (for a Wikipedia page). While for the other concepts, for which there is no dedicated page on Wikipedia, a

[7] http://rsp.inf.elte.hu/

new page (draft) on RSP is created and subject to the voting system in order to become an article; finally the concept in the ontology will be connected to the relative RSP page.

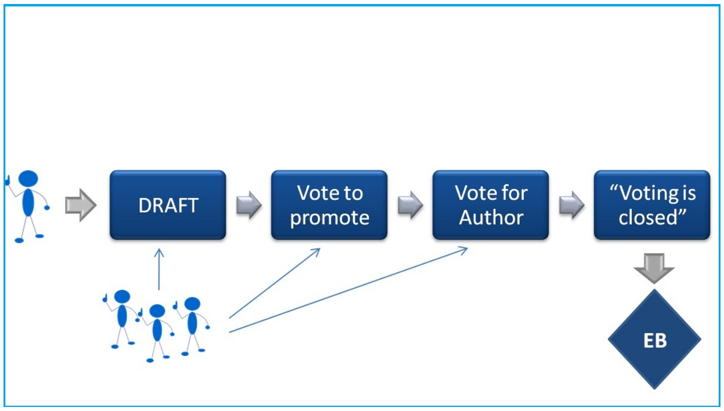

Figure 3.8: If a contributor of a draft thinks that it is ready to become an article, he can start the voting system. the contributors(of the draft only) can 'vote' on the promotion of the draft. Each vote of each contributor has a weight: the more the contributor edited the draft, the more the weight of his vote is. If (and when) the votes exceed 70% of contributors, the draft is promoted to become an article, otherwise it will remain a draft. Members have also to decide on who will be asked to write the article; when "*voting is closed*" appears on the page, a notification is sent to a Special Page and members of the Editorial Board of Robotic SurgePedia(people who have rights to decide if a draft is correct, complete and consistent) will accept or reject a draft which can be the basis of an article.

Every user that is registered at RSP can log in and make their contribution to both the SurgePedia pages and the ontology.

3.4.1 Ontology browser

As already said, Protègè is easily extensible and many people have developed plug-ins and applications for the Protègè-OWL and/or protégé-Frames[8] editors. Once downloaded Protègè, it is recommended to visit the page that list all the library plug-ins and applications (`http://protegewiki.stanford.edu/wiki/Protege_Plugin_Library`). In the epilepsy neurosurgery ontology the Ontoviz plug-in is used to browse the ontology.

Techniques to visualize ontologies can be grouped in 4 modalities: network, tree, neighborhood and hyperbolic view. The network view is the most simple

[8]The Protégé-Frames editor provides a user interface and knowledge server to support users in constructing and storing frame-based domain ontologies, customizing data entry forms, and entering instance data.

and it is applicable to every ontology, because it consists of a graph in which nodes represent concepts, while arcs represent connections between them. If hierarchical relations exist between members in an ontology (i.e. taxonomies) it would be useful a tree view. Three different variants have been implemented (fig. 3.9):

1. the indentation, which is used in many ontology editors to show the hierarchy of concepts (and sometimes also the properties);

2. the tree made of nodes and arcs is used many times, especially in graphical viewers associated to ontologies editors;

3. the treemap (i.e. "*Chinese box*") which is useful to show hierarchical relations in a limited space making use of geometrical structures represented one inside the other.

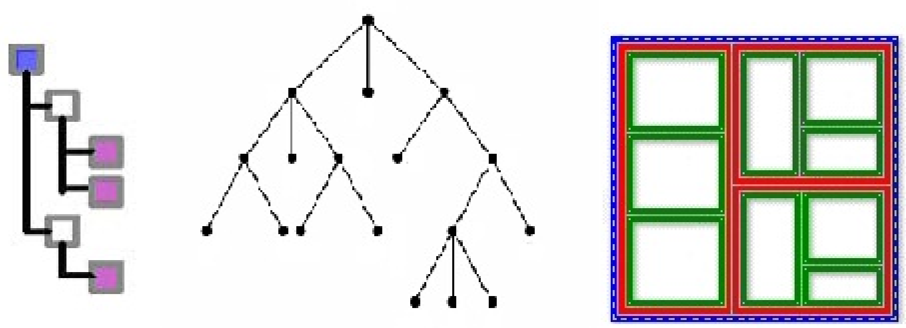

Figure 3.9: Tree views: indented, nodes and arcs, treemap. The first one is a style of representation in which data is displayed with indented lines. The second is the classic tree map composed of nodes and arcs. The third is a display mode in which child objects arerectangles, nested within rectangles parents.

The other two schemes (neighborhood and the hyperbolic views, fig. 3.10) focus the attention on a chosen node and its nearest neighbors. The neighborhood view uses the nodes and arcs representation and the selected node becomes the new root. The hyperbolic view projects into a hyperbolic plane the ontology and it allows to show hundreds of nodes in a small space.

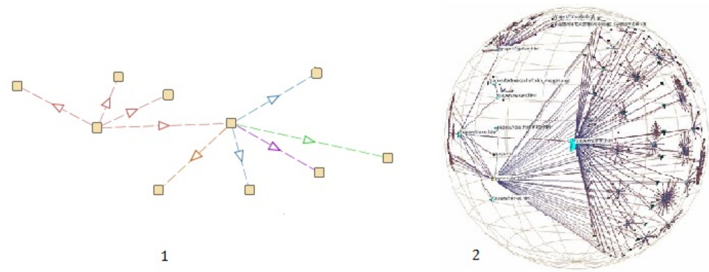

Figure 3.10: Neighborhood (1) and hyperbolic views (2). Both the neighborhood and the hyperbolic views focus the attention on a chosen node and its nearest neighbors. In the former case only the semantically nearest nodes are displayed, whereas in the latter case the nodes are displaced onto a hemi-spherical surface, projected onto the visual window, therefore magnifying the central nodes while shrinking the peripheral nodes.

The OntoViz Tab allows the user to visualize Protègè ontologies with the help of a highly sophisticated graph visualization called GraphViz from AT&T. The visualization of the ontology is highly configurable by the user and include:

- Picking a set of classes or instances to visualize part on an ontology;
- Displaying relationships between classes;
- Specifying colors for nodes and arcs;
- When picking only a few classes or instances, various closure operators can be applied (e.g. subclasses, superclasses) to visualize their vicinity;

3.5 Ontology based classification example: diagnosis of temporal (TLE) and extratemporal (eTLE) lobe epilepsy

Temporal lobe epilepsy (TLE) is a form of focal epilepsy, a chronic neurological condition characterized by recurrent seizures while the extratemporal lobe epilepsies (eTLE) term describes specific form of seizures which have origin outside the temporal lobe. Temporal and extratemporal lobe epileptic seizures induce different symptoms. TLE are the most studied form of focal epilepsy and the most subject to surgical treatment. The presurgical iter includes videoEEG of seizures, to be correlated with clinical and neuroradiological data to define the epileptogenic zone [9] (EZ).

[9] Epilelptogenic Zone (EZ) was originally defined in the '50s by Jean Bancaud and Jean Talairach as "the site of the beginning and of primary organization of the epileptic seizures" (La règion <<d'origine des dècharges critiques et leurs voies de propagation primarie>>). http://rsp.inf.elte.hu/mediawiki/index.php/Epileptogenic_zone

In literature there are different works that try to classify epileptic seizures considering only ictal semiology.

In particular in [19] the International League Against Epilepsy (ILAE) introduces a seizure classification based on clinical semiology, interictal EEG findings, and ictal EEG patterns. In [20] the second work try to classify epileptic seizure considering exclusively ictal semiology: this classification has been used in daily clinical practice for more than 10 years in selected epilepsy centers. It is believed that all additional clinical information, such as EEG, anatomic neuroimaging, clinical history, neurological examination and seizure evolution should be analyzed separately and then integrated to define the epileptic syndrome. The advantage of classifying the epileptic seizures by semiology are:

1. It provides a terminology that permits clear identification of ictal semiology features independent of any other results;

2. A semiological seizure classification focuses the attention of the observer on clinical semiology;

3. A semiological seizure classification, can be applied to any age group. However, certain types of seizures will not occur in newborn and infants because of their incompletely developed nervous system.

In this thesis work, each patient was video registered during epileptic seizure and interviewed in the post-ictal seizure. Video registration has a variable duration, but storing of ictal manifestations has a duration of one minute, starting from the first symptoms that patient show. Storing information of only one minute of epileptic seizure is an approximation, but these firsts 60 seconds include the information of the area of starting of epileptogenic zone.

As time passes, the seizures spread to other brain structures, giving no more indications on the epileptogenic zone. Thanks to the contributions of 7 clinicians from Besta, Niguarda and San Paolo Hospital (Milan), a list of manifestations that we will take into account to classify the epileptic seizure as temporal or extratemporal is defined. The list of symptoms have 29 elements, which are subdivided in two subgroup: subjective symptoms (14) and objective symptoms (15) 3.1. The ontological representation is used also to model the epileptic seizures semiology as support to diagnosis of TLE and eTLE. From the ontology an occurrence matrix of symptoms, in 60 seconds of seizures, is created. Then, a classifier of TLE and eTLE is trained (par.).

1	epigastic
2	auditory illusion
3	auditory hallucinations
4	visual illusion simple
5	visual illusion complex
6	visual hallucination simple
7	visual hallucinations complex
8	olfactive hallucinations
9	gustatory hallucinations
10	psychic manifestation
11	fear
12	autonomic (tachicardia, dispnoea, etc)
13	motor-sensitive manifestation monolateral
14	motor-sensitive manifestation bilateral
15	oroalimentary automatisms
16	gestural automatisms monolateral
17	gestural automatisms bilateral
18	head deviation
19	eyes deviation
20	motor clonic monolateral arm
21	motor clonic bilateral arms
22	motor clonic monolateral leg
23	motor clonic bilateral legs
24	motor clonic monolateral mouth
25	motor dystonic monolateral arm
26	motor dystonic bilateral arms
27	motor dystonic monolateral leg
28	motor dystonic bilateral legs
29	autonomic (flushing, shialorrhea, vomiting)

Table 3.1: List of ictal symptoms: (1 : 14) Description of subjective symptoms was obtained at the time they occurred or during the post-ictal interview. (18 : 28) describe motor signs and they can be defined as 'single type of contraction of a muscle or a group of muscles that is usually stereotyped and not decomposable into phases'; they could involve the head or the eyes or body segments and they were differentiated in 'clonic' motor signs, which are myoclonic contractions that regularly recur at a rate of 0.2-5/s -during the seizure- , and 'dystonic' motor signs, which are unnatural posturing of one upper extremity with rotational component. (15,16,17) Describe complex behaviors that looked like natural behaviors; gestural automatism essentially affected the distal body segments and they were characterized by 'more or less coordinated, repetitive, motor activity'.

In TLE the subjective phenomena are more frequent, the loss of the contact more gradual, seizures are longer, the post-ictal confusion is more important and the oro-alimentary automatism are predominant in TLE [31]. The "ideal sequence" of ictal modifications in TLE is a subjective manifestation (like ascendant epigastric feeling, but also psychic symptoms) followed by a loss of contact, oroalimentary automatism and homolateral head deviation and a contralateral arms dystonic posturing with homolateral gestural automatism. The chronological sequence of these events and their relative duration might indicate the seizure onset area and the progressive imvolvement of adjacent brain structures. In addition we know that one particular symptom may originate from different cortical areas, consequently the chronology can help to localize the EZ.

However, it is found a significant association of gestural automatism with TLE [32]. It is also more common to find temporal lobe epileptic seizures that present concomitantly ipsilateral gestural automatisms and contralateral dystonic motor signs [33][34] [22], as in the patient reported in figure 3.11.

On the contrary seizures in eTLE are frequently abrupt, brief, the loss of contact is precocious, rarely oroalimentary automatism are present, and motor modification can be bilateral. Gustatory, vestibular and auditory symptoms, the localizing significance of which remains uncertain were more frequently associated with extratemporal epilepsies. Some signs are common in both the TLE and eTLE. Versive manifestations (*head and eyes deviations*), were more frequently associated to extratemporal seizures, although there's significant occurrence in temporal patients at the same time. Elementary motor signs, for example, as well as complex behaviors were commonly found in temporal patients and extratemporal patients.

However, not all the seizures are so easily recognizable and, particularly the hypermotor gestural automatisms, can occur in bot TLE and eTLE. The analysis of video information is time-consuming and sometime one recorded seizure is not sufficient and every patient has particular chronological and clinical picture of the ictal events.

Figure 3.11: Right mesial frontal lobe epilepsy (eTLE). Dystonic posturing of the left upper limb (Cortesy of A. Bleasel)

3.5.1 Epilepsy diagnosis ontology

Our purpose is to exploit the information stored in the ontology to classify individuals.

The basic idea was to assign, for every patients, a value of correlation between two features, which in this case are ictal symptoms.

According to the domain ontology one feature is treated as an instance of a general class (which in our case is `rsp:ictalSymptom`) and every seizure of a patient is described by a set of these instances according to the set that appears during seizure. The manifestation class is a class union of some instances belonging to `rsp:ictalSymptom` class, which groups also other types of symptom that are not used in our algorithm.

Figure 3.12 shows the ontological modeling for a patient with epileptic seizures, analyzed after video registration.

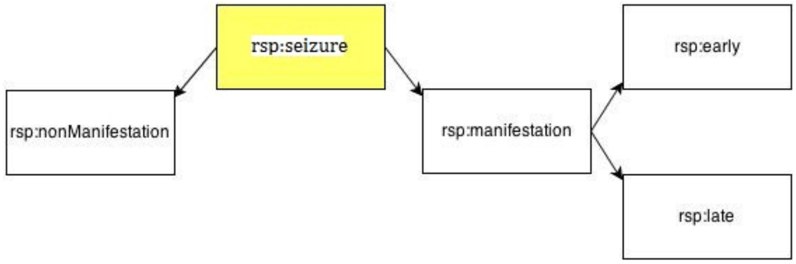

Figure 3.12: Ontological map of patient affected by epileptic seizure. Seizure class has relations with two other classes named Manifestation and nonManifestation. Manifestation class is subdivided in two other classes named early and late.

First, domain ontology tree shows the relationships between concepts and features, including all nodes of concepts [35]. For each patient of the dataset a graph can be drawn: starting from these graphs that represent the ontological information, a correlation matrix of symptoms can be calculated.

The following figure (3.13) describes the ontological modeling for one of the patients of the eTLE dataset.

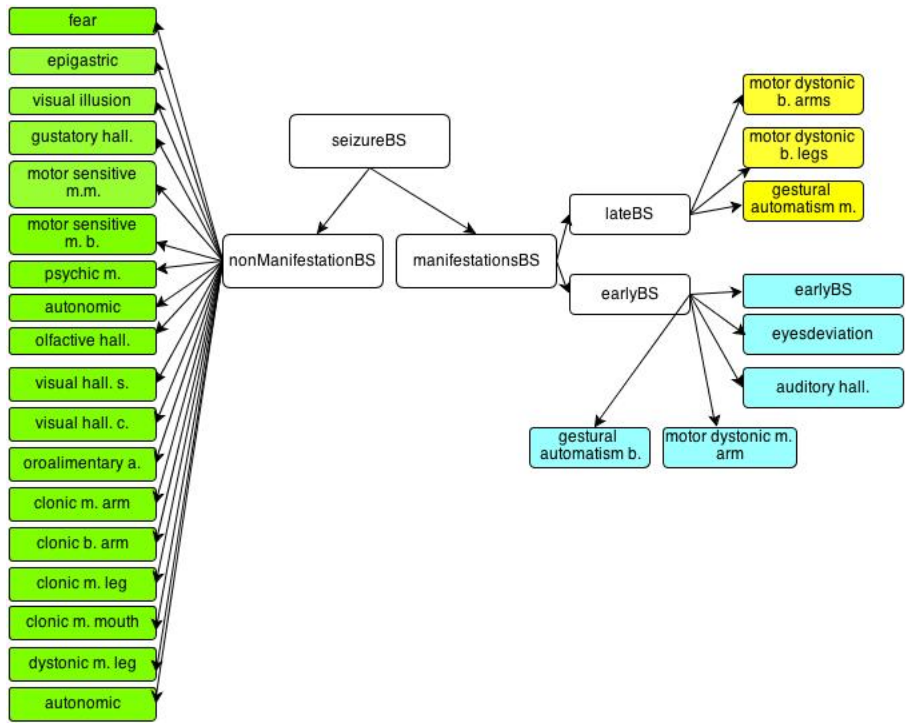

Figure 3.13: Instance of the ontology map with its ictal symptoms (patient BS from Besta Hospital)

3.5.2 Dataset

Dataset of patients that we used in our work is given by three clinical centers: Besta, Niguarda Ca Granda, San Paolo Hospitals. It is composed of 109 patients, of which: 60 patients with temporal lobe epilepsies (TLE) and 49 patients with extratemporal lobe epilepsies (eTLE);

Once clinicians have a collection of video registrations, they created an excel file for each patient in which there are 29 (15+14) rows, one row for each of the 29 symptoms that can be observed and 60 columns, that represent recording time (in seconds). For each patient, clinicians fill in the matrix, assigning to the manifestation observed a ‘1’, while ‘0’ to manifestation not observed.

epigastric	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
auditory illusion	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
auditory hallucinations	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
visual illusion simple	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
visual illusion complex	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
visual hallucination simple	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
visual hallucinations complex	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
olfactive hallucinations	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gustatory hallucinations	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
psychic manifestation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
fear	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
autonomic (tachicardia, dispnoea, etc)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor-sensitive manifestation monolateral	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor-sensitive manifestation bilateral	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
oroalimentary automatisms	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gestural automatisms monolateral	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gestural automatisms bilateral	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
head deviation	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	1	1	0
eyes deviation	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	1	1	0
motor clonic monolateral arm	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor clonic bilateral arms	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor clonic monolateral leg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor clonic bilateral legs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor clonic monolaterale mouth	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor dystonic monolateral arm	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor dystonic bilateral arms	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor dystonic monolateral leg	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
motor dystonic bilateral legs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
autonomic (flushing, shialorrhea, vomiting)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 3.14: Example of patient excel file. Symptom that appear during seizure are underlined with rectangle. In this case all the symptoms starts and ends in the first thirty seconds of the registration.

Fig. 3.14 resumes the information included in a video registration of 1 minute of an epileptic seizure: 29 symptoms are listed and for each symptom a '0' is assigned if not present during the seizure, while a '1' is assigned if present during the seizure. The patient presents four different symptoms (head deviation, eyes deviation, motor dystonic monolateral arm, motor dystonic monolateral leg) that terminate in the first thirty seconds.

Clinicians said that a symptoms can be considered “early”, if it starts in the firsts ten seconds of the registration, while it is considered “late” if the manifestation starts in the last fifty seconds.

The occurrence of manifestations in patients dataset used for this work is shown in fig. 3.15.

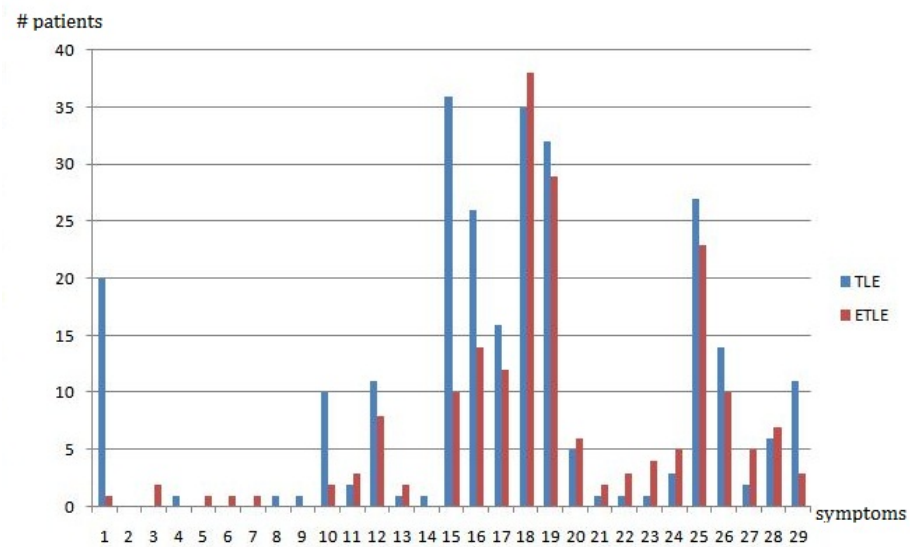

Figure 3.15: Patients dataset's occurrence istogram. On the x-axis there is the list of subjective and objective symptoms that are used to describe every seizure of patients dataset. On the y-axis there is the number of patients that show a specific symptom.

3.5.3 Classification with features correlation matrices

First of all we translate the dataset that has been described in paragraph 3.5.2 in a matrix. A patient is described by a (5*29) A matrix, where 29 refers to the set of possible manifestations that can be observed during an epileptic seizure.

- The first row is used to assign a index to every symptom: A(1,:);
- The second row is used to define if the *i-th* symtpom is objective or subjective. 1 stands for 'subjective' and 2 for 'objective';
- The third raw, is the one that gives information about the presence or the absence of a specific symptom during a seizure. For this information 1 stands for the presence and 0 for the absence of a symptom;
- The fourth row represents the temporality of the symptoms: 1 stands for the early occurrence, 2 for the late occurrence and 0 for the absence of a symptom;
- The fifth row is used to give information relative to the position of each symtpom in the ontological tree. It is calculated as the distance of a concept starting from the 'root' concept, which in our case is the seizure class (fig. 3.16).

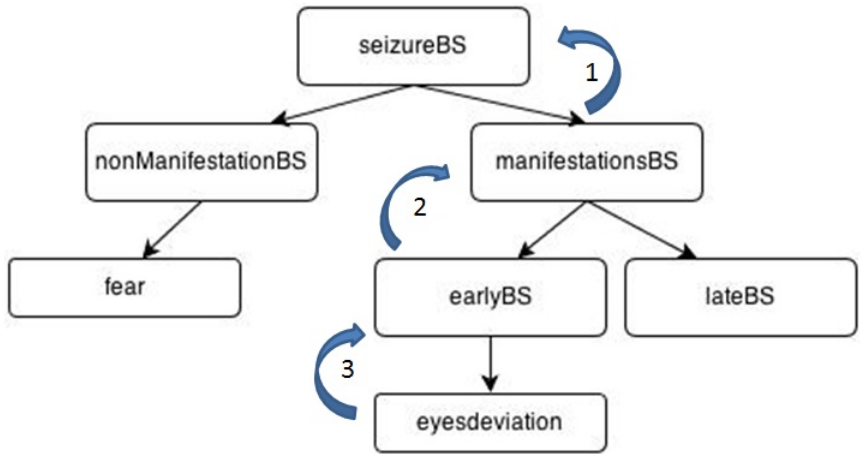

Figure 3.16: Information relative to the position of a symptom in the ontological tree. Between the seizure root class (BS) and the symptom eyes deviation there are three arcs (properties).

The matrix become:

epi.	a.ill.	a.hall.	vis.ill.s	vis.ill.c	vis.h.s.	vis.h.c.	olf.h.	g.h.	psychic	fear	auto.	m.-s. m.	m.-s. b.	oro.a.	g.a.m.	g.a.b.	h.d.	e.d.	m.c.m.a.	m.c.b.a.	m.c.m.l.	m.c.b.l.	m.c.m.m.	m.d.m.a.	m.d.b.a.	m.d.m.l.	m.d.b.l.	a.
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	2	2	2	2	2	3	2	3	2	2

Figure 3.17: Patient Matrix. First row: symptoms list; second row: index of each symptom (1 for subjective, 2 for objective); third row: index of each symptom (1 for presence, 0 for absence); fourth row: index of each symptom, whether they are early (1), late (2) or not present (0); fifth row: index for each symptom (the distance in arcs from the seizure root concept);

Once the dataset has been translated in a set of matrices the correlation matrix was computed as:

$$C(sign.i, sign.j) =$$

$$= \frac{(Depth(sign.i) + Depth(sign.j)) * \alpha + \beta}{(Distance(sign.i, sign.j) + \alpha) + 2 * Max(Depth(sign.i), Depth(sign.j)) + \beta} \quad (3.1)$$

where sign.i and sign.j are the considered symtpoms (e.g. head daviation and eyes deviation), the function Depth*(sign.i)*, Depth(*sign.j*), indicate the distance of the *sign.i* and *sign.j* from the class seizure in the ontology tree. Distance (*sign.i,sign.j*) represents lenght of the shortest path between *sign.i* and *sign.j* in the ontology tree (number of arcs between *sign.i* and *sign.j*). Max(Depth(*sign.i*), Depth(*sign.j*)), represents the bigger distance from the seizure class in the domain ontology tree. *Alpha*, *beta* are variable parameters ranging from 0 to 1. *Alpha* is used to balance the proportion between Depth and Distance.

Most of the information that can distinguish two population -in our case the temporal and extratemporal patients- is concentrated in the correlation matrices, which expresses the relation between each pairs of manifestation (i.e. the probability that if during an epileptic seizure is present the symptom *a*, then will be present also the symtpom *b*).

In order to automatically classify patients (wheter they are TLE or eTLE) correlation matrices for every patient and model of correlation matrices of each population (TLE and eTLE) were computed. Starting from 109 patients -of which 60 patients with temporal lobe epilepsies and 49 patients with extratemporal lobe epilepsies-, we subdivided this dataset in three sets [44]:

- 70% for the training the classifier;
- 15% for validation of the classifier;
- 15% for testing the classifier;

Algorithm description:

From the original dataset, the testing set is randomly extracted (15 patients). Then changing every time the number of patients (from 2 to 27 for each population) that trains the classifier (i.e. the mean correlation matrix), we randomly extract validation set and training set. Validation set is composed of 14 patients that are characterized by index '1' if TLE patient and index '2' if eTLE patient.

For every iter we calculate the mean correlation matrix for TLE and eTLE and we test these models on the validation set: for every patient belonging to the validation set, we calculate the correlation matrix and we compare it with the models. The patient is assigned to the closest model (i.e. euclidean distance): '1' for TLE, '2' for eTLE.

The calculation of the mean correlation matrix is repeated 100 times for every value that the variable 'number of patients for training the classifier' assumes and we calulate the mean value and the standard deviation of the cost function 3.2 on validation set.

$$costfunction = \frac{\#insuccess}{\#success} \tag{3.2}$$

where success stands for the number of patients of testing set correctly identified as TLE or eTLE by the classification algorithm and insuccess stands for the number of patients of testing set incorrectly identified.

The algorithm is reported:

where the mean correlation, thus the two models for TLE and ETLE, are:

$$correlatiomeanT = (\sum_{n=1}^{N} correlation_n)/N \tag{3.3}$$

$$correlatiomeanET = (\sum_{n=1}^{N} correlation_n)/N \tag{3.4}$$

where N is the number of patients used to calculate the model of correlation matrices.

The optim in the cost function 3.2 belongs to the a specific number of patients that has to be used for calculating the matrices correlation models for TLE and eTLE patients.

Once we have calculate the optim value, we randomly extract a new set of training patient, and we test the models of correlation matrices on testing set. We repeat this step 100 times and we calculate the mean value of patients correctly identified. The total accuracy of the algorithm is:

$$accuracy = \frac{\#meansuccess}{\text{totPatients}} \tag{3.5}$$

where 'totPatient' stands for the number of patients tested (i.e. 15).

3.5.4 Validation protocol for clinicians

For this experiment, startig from the dataset presented in previous paragraph, clinicians make a diagnosis for a patient simply on the basis of the excel file that gives information on the occurence of ictal symptoms (subjective or objective) and the timing of occurrence of symptoms. They independently evaluate the specific case and then they emit the diagnosis. In particular, they have to classify epileptic seizures as TLE or eTLE. In this way, we can compare the results of the algorithm -that will be presented in the paragraph 3.4.3- with the results of classification made by clinicians.

Chapter 4

Results

4.1 Epilepsy diagnosis and neurosurgery ontology

Following the methodology reported in paragraph 3.2.2, we build the ontology:

1. *Planning.* The plan is to build a seed ontology that could represent knowledge of diagnosis and neurosurgery epilepsy. For the ontology creation,

2. *Specification.* It is important to build an ontology that models the field of epilepsy diagnosis and treatment because there is the need of representation in a structured way, store surgical case and process information stored, so that every surgeon and clinician can access to it and evaluate every new case compared to the stored cases and makes some considerations on the best surgical approach to be adopted.

3. *Knowledge acquisition.* Most of the knowledge included in the epilepsy neurosurgery ontology comes from textbook [52], interviews to domain experts, i.e. Dr. F. Cardinale (Niguarda Ca' Granda Hosp.) and Dr. L. Bello (Humanitas Hosp.), publications on journals related to neurosurgery and epilepsy [53, 55, 54] and measurments of dimensions of surgical instruments usually used in neurosurgical procedures, while information relative to the diagnosis of epilepsy diagnosis come from interviews to Dr. L. Tassi (Niguarda Ca' Granda Hosp.) and Dr. F. Gozzo (Niguarda Ca' Granda Hosp.) and publications [19, 20]. In the fig. 4.1 is reported a table implemented with contribution of Dr. F. Cardinale: it resumes all neurosurgical procedures that are executed at Niguarda Ca Granda Hospital.

Surgical procedure	Procedure category	Procedure description	# treatments per year at Niguarda	Complications	Patient position (head)
Resection	Craniotomy	The resection consists in the removal of the epileptogenic zone's gray matter, together with the underlying white matter when necessary. All vesselssupporting the resection area can be coagulated and cut.	100 (sum with disconnections)	Bleedings and neurological dysfunctions (gentle or severe, transient or permanent)	Supine (prone or on one side if the surgical field is in posterior region of the brain)
Disconnection	Craniotomy	When the epileptogenic zone is very large, it is better to perform a disconnection instead of a resection, in order to avoid very big post-surgical cavities and the subsequent sequelae (e.g. severe brain shift, hemosiderosis, etc). The disconnected tissue remains in site, and major vessels must be saved.	100 (sum with resections)	Bleedings and neurological dysfunctions (gentle or severe, transient or permanent)	Supine (prone or on one side if the surgical field is in posterior region of the brain)
Subduralgridsi mplantation	Craniotomy	Invasive intracranial monitoring: diagnostic procedure to identify the epileptogenic zone.	No	Bleedings and infections	Supine (prone or on one side if the surgical field is in posterior region of the brain)
Callosotomy	Craniotomy	Corpus callosumDissection (palliative proceduremainlyindicated for suddenictalfallings due to contralateralhemisphericseizurepropagation)	No	Bleedings and infections. Disconnectionsyndrom e.	Supine
Vagus nerve stimulation	Neck surgery	Chronic vagus nerve stimulation, obtained with a coil electrode surrounding the nerve in the cervical region. The electric generator is placed in a subclavearsubcutanous pouch. This palliative procedure is aimed to reduce the seizure's number, via chronic transmission of the stimuli to the brain and subsequent interference with seizure's genesis.	No	Carotid dissection.	Supine
Multiple Subpial Transection	Craniotomy	Palliative procedure based on interruption of inter-columnar connections between cortical neurons. It is performed when epileptogenic zone includes non-resectable areas (because of function).	No	Bleedings and infections	
DBS (Deep Brain Stimulation)	Stereotactic / mini-invasive	Two symmetric electrodes, are implanted in a deep central target (STN, AN,ZonaIncerta or CMN) with the aid of a micro driver, 1 DOF, robot. The effective mechanism of this palliative procedure should be similar to the VNS one.	1	Bleedings and infections	Supine

Figure 4.1: Table for neurosurgical procedure executed at Niguarda Hospital. Every type of procedure is charcaterized by some specifications: a category (every procedure is categorized as craniotomy, stereotactic, mini-invasive, neck surgery), a short description, the number of treatments per year at the Niguarda Ca Granda Hospital, complications of the surgical procedure and the patient position.

The figures 4.2 and 4.3 show two surgical instruments that surgeons usually use in epilepsy neurosurgical procedures. For every instrument we stored photos, operating characteristics and dimensions.

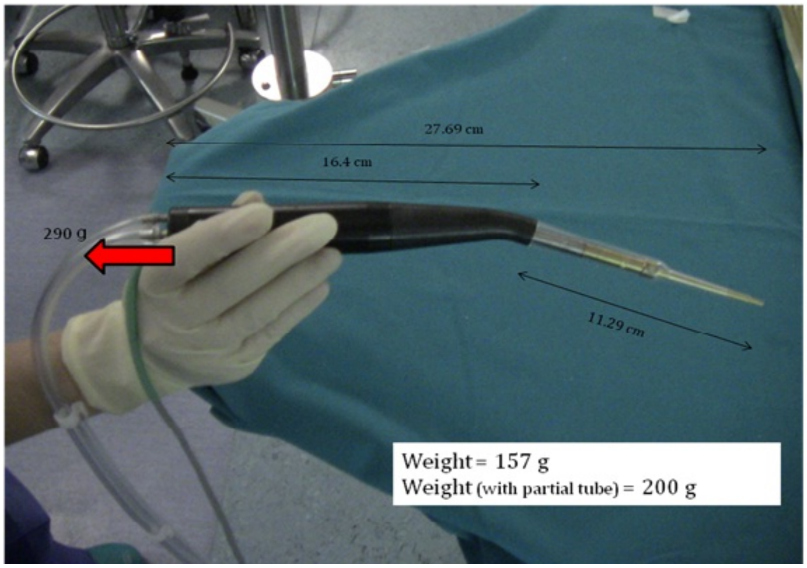

Figure 4.2: SELECTOR US aspirator (25 KHz). It is used for fragmentation, irrigation and aspiration of the area of the brain to be treated.

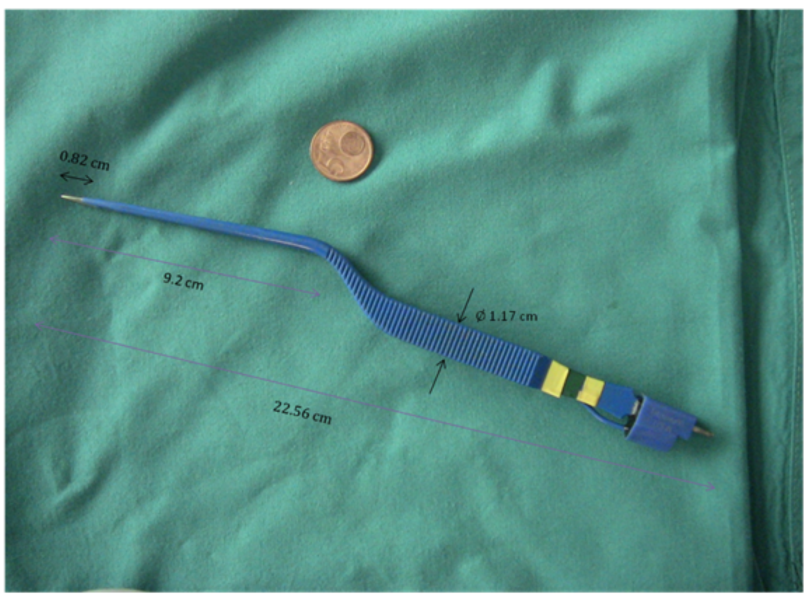

Figure 4.3: CODMAN Malis (CMC III), Bipolar forceps used in surgical procedures for coagulation.

1. *Conceptualization.* In this step all relevant concepts of epilepsy diagnosis

and neurosurgery are identified in addition to the relationships that stand between concepts and the attributes that help the user in identifying each individual.

2. *Integration.* Two existing ontology help in enlarging and developing the seed ontology: DBpedia (for the part of modelization of epilepsy neurosurgical procedures) and SNOMED CT (for the part of modelization of the pre-operative phase, thus clinical semiology of seizures and classification of epilepsy);

3. *Implementation.* In this step we choose the computer language and the ontology editor for implementing the ontology. We decided to use Ontology Web Language (OWL) because of its popularity among the ontology development community, and for the number of tools and reasoning engines that are available. Besides, it provides additional vocabulary, and facilitates greater machine interoperability of Web content. Moreover, we choose the OWL DL language because it allows more expressiveness and decidabilty that OWL Lite and Full. In fact it supports users that want maximum expressiveness without lacking computational completeness. Moreover the OWL Lite language could not manage the ontology because it is not simply made of taxonomies. The ontology editor adopted is Protègè 3.4.8 because of it is a free and open-source software and it provides many extensions for collaborative ontology editing. In the figure 4.4 the information relative to a surgical instrument in the epilepsy neurosurgery ontology is shown.

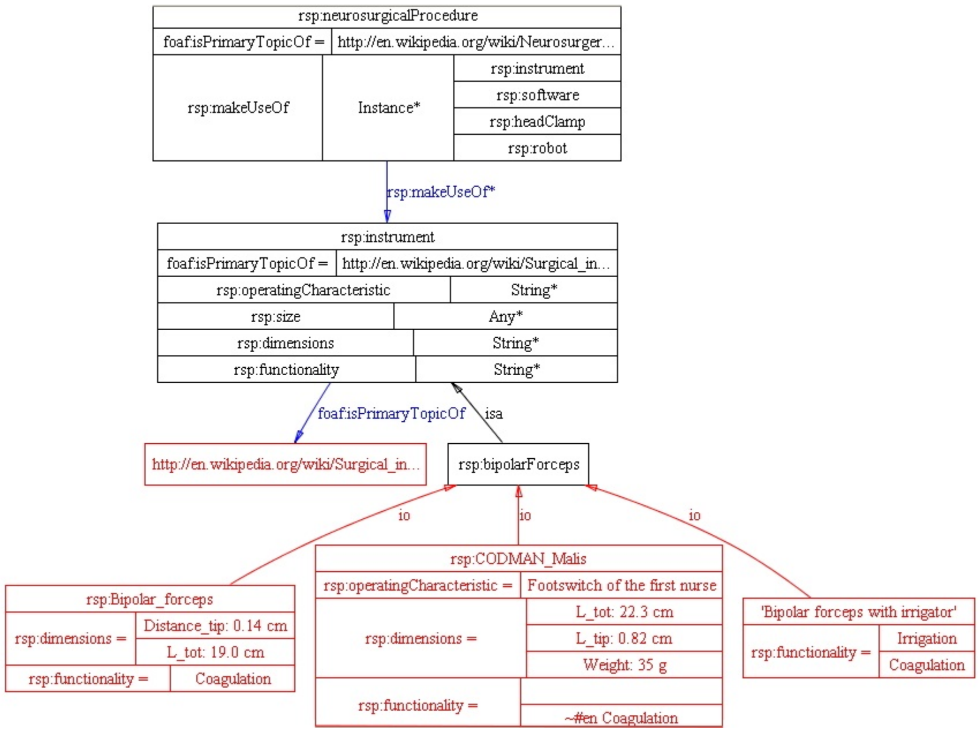

Figure 4.4: Bipolar forceps instances, properties and hierarchy. The surgical procedure class makes use of some surgical instrument. The surgcical instrument concepts is explained in the wikipedia page http://en.wikipedia.org/wiki/Surgical_instrument. Bipolar forceps class is a subclass of the more general instrument class. The bipolar forceps class is populated with three different model of forceps. Every individual is defined with a set of characteristics (i.e. functionality and dimensions).

4. *Evaluation.* There are no specific methods to evaluate ontologies, but the clarity, coherence and extensibility will be evaluated by the community that will contribute in enlarging the ontology (such as on the RSP portal).

The epilepsy diagnosis and neurosurgery ontology is composed of:

- 452 classes;
- 30 object properties;
- 11 datatype properties;
- 180 individuals, of which:
  - 34 individuals are external resources;

- 43 individuals are patients of the dataset provided by three medical centers (San Paolo, Niguarda Ca Granda, Besta);
- 16 individuals are surgical instruments;
- 54 individuals are epileptic symptoms of seizure onset;
- the other individuals remaining are distributed in classes of complications, surgical procedure, surgeon and hospitals;

The figure 4.5 shows the taxonomy of the epilepsy diagnosis and neurosurgery ontology :

```
owl:Thing
    protege:ExternalResource
    rsp:action
    rsp:armAttachment
    rsp:bodyPart
    |   rsp:arm
    |   |   rsp:hand
    |   rsp:brain
    |   |   rsp:AN
    |   |   rsp:CMN
    |   |   rsp:corpusCallosum
    |   |   rsp:hippocampus
    |   |   rsp:lobe
    |   |   rsp:STN
    |   |   rsp:Temporal_horn
    |   |   rsp:Vagus_nerve
    |   |   {'Brain region (body structure)', 'Brain region'}
    |   |   |   {'Brodmann area', 'Brodmann area (body structure)'}
    |   rsp:leg
    |   rsp:skin
    |   rsp:skull
    rsp:capability
    |   rsp:locomotionCapability
    |   rsp:operationalCapability
    |   rsp:sensoryCapability
    rsp:complication
    rsp:coordinates
    rsp:data
    rsp:disease
    rsp:headClamp
    |   rsp:Mayfield_head_clamp
    |   rsp:Three_point_pin_fixation
    rsp:hospital
    rsp:ictalSymptom
```

Figure 4.5: Extract of the taxonomy of epilepsy diagnosis and neurosurgery ontology. Every concept with rsp: prefix identifies a class. Indented classes are subclasses (e.g. rsp: corpusCallosum and rsp:hippocampus are subclasses of the higher rsp:brain class that groups the structures of the brain).

4.2 Automatic patient classification

In this section results of classification algorithm based on correlation matrices of ictal manifestations are presented.

Figures 4.6 and 4.7 show the correlation matrix for a patient with temporal and extratemporal lobe epilepsies respectively. They show different values for different pairs of manifestations. In the equation 3.1, *alpha* assumes the value 0.7 (greater than *beta,* that is 0.3) because it weights the *Depth* and *Distance* functions that include relevant information for calculating correlation between two symptoms.

On the diagonal, there are highest values: darker red is for manifestations observed in video registration of seizure, red more light is for manifestation not observed. Values in correspondence of light blue is for pairs of manifestations that don't appear during the seizure (such as 1-epileptic seizure and 11-fear).

The value of correlation between these two feature is higher than two symptoms of which one does not appear and the other is observed (blue) during the epileptic seizure.

If a couple of symptoms assume the yellow value, this means that these manifestations are both observed in the video registration, they have an higher value than two manifestation both not present.

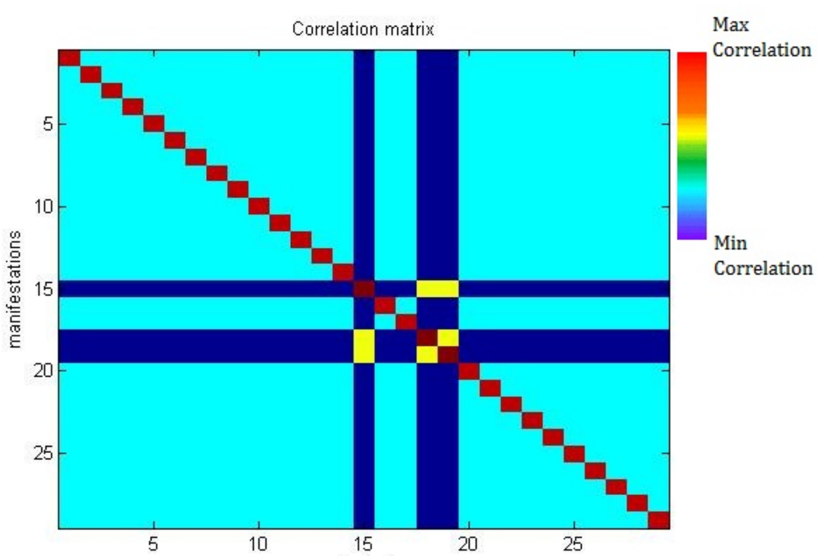

Figure 4.6: Correlation matrix for a temporal patient that presents *oroalimentary automatism* (15), *head deviation* (18) *and eyes deviation* (19). On x and y axis there are the list of ictal observable symptoms.

Figure 4.7 shows a particular case of eTLE patient:

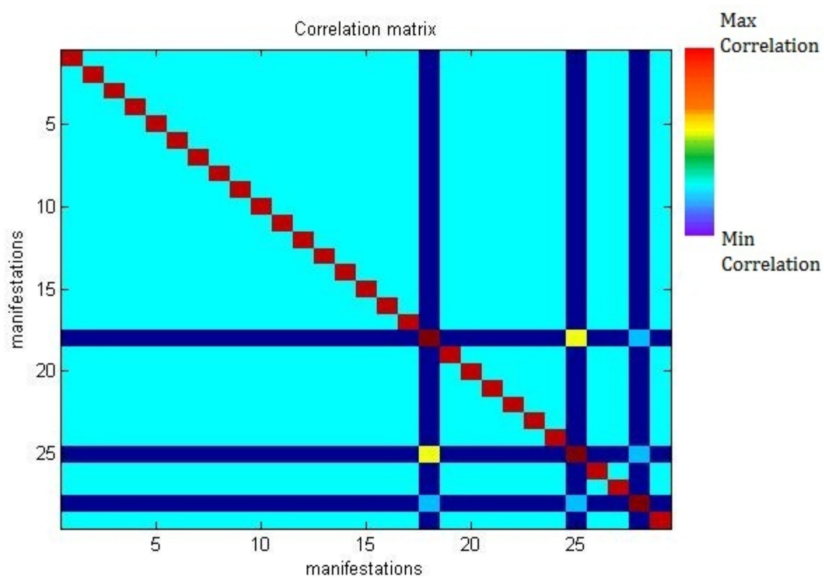

Figure 4.7: Correlation matrix for an extratemporal patient that present head deviation (18), motor dystonic monolateral arm (25), motor dystonic monolateral leg (28). The light blue value in correspondence of the intersection between two dark blue lines is for a couple of symptoms in which one starts in the first ten seconds of the registration, while the other starts later than 10 seconds. They have a lower correlation than a pair of symptoms that appear both in the first ten seconds (yellow).

Figure 4.8 describes the cost function 3.2: on x-axis there is the number of patients for each population (TLE and eTLE) used for calcultating the classifier of temporal and extratemporal lobe epilepsies.

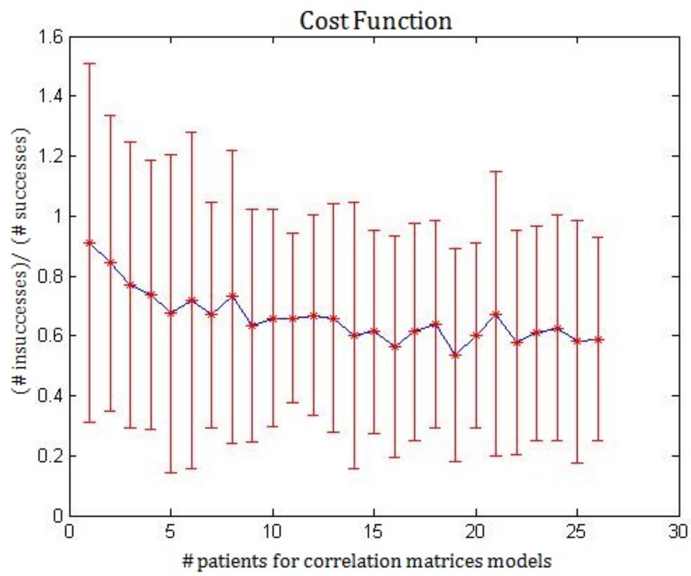

Figure 4.8: Mean Error Performances on validation set. Performance of the classifier is calculated as the fraction of the number of insuccesses on the number of successess. In the graph are underlined also means and standard deviations.

Once the best number of patients to be used for training the classifier has been chosen, the two models are computed (fig. 4.9, 4.10).

In Figures 4.9 and 4.10 x and y axis represent the 29 symptoms that can occure during a seizure.

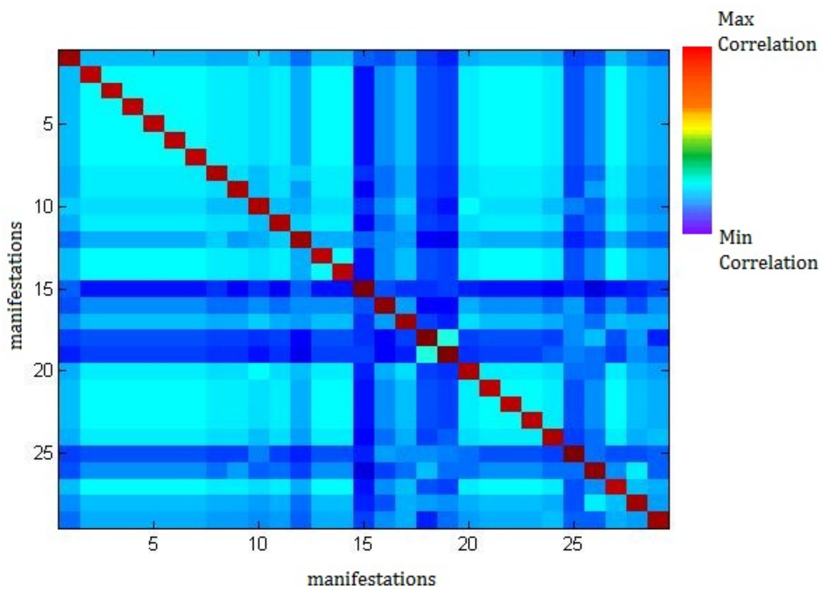

Figure 4.9: Model of correlation matrix TLE calculated with 19 patients. It is the mean of correlation matrices.

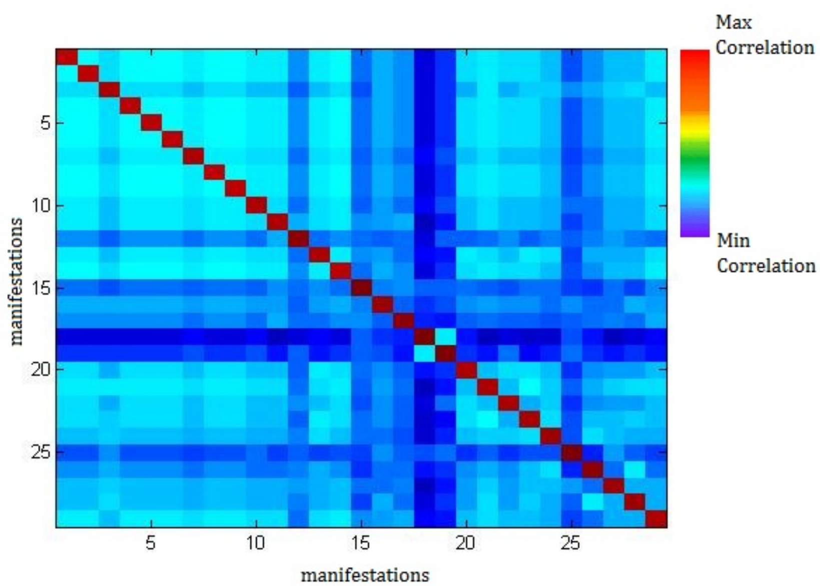

Figure 4.10: Model of correlation matrix eTLE with 19 patients

Results on testing set is presented in the following table for 100 iterations.

	Iter #1	Iter #2	Iter #3	Iter #4	Iter #5	Iter #6	Iter #7	Iter #8	Iter #9	Iter #10	Original
Test #1	1	2	1	1	2	1	2	1	1	1	1
Test #2	2	2	2	1	2	2	2	1	2	2	2
Test #3	2	2	2	2	2	2	2	2	2	2	2
Test #4	1	1	1	1	1	1	1	1	1	1	1
Test #5	2	2	2	2	2	2	2	2	2	2	2
Test #6	1	2	1	1	1	2	1	1	1	1	1
Test #7	2	2	1	1	1	2	2	2	1	1	2
Test #8	1	1	1	1	1	1	1	1	1	1	1
Test #9	2	2	2	2	2	2	2	2	2	2	1
Test #10	2	2	2	2	2	2	2	2	2	2	2
Test #11	2	2	1	1	1	1	2	1	1	1	1
Test #12	1	1	2	2	1	2	2	2	2	2	1
Test #13	1	1	1	1	1	1	1	1	1	1	1
Test #14	1	1	1	1	1	2	1	2	1	1	1
Test #15	2	2	2	2	2	2	1	2	2	2	2
Tot success	13	11	12	11	12	11	10	11	12	12	

Figure 4.11: Testing set is fixed and true values are reported in the last column: 1 stands for TLE patient, 2 stands for eTLE patient. We executed 100 times the test of the algorithm, changing the patients that we use to train the classifier (i.e. to calculate the models of correlation matrices for TLE and eTLE). 'Tot success' indicates the number of patient correctly classified as temporal or extratemporal.

Mean successes	Standard deviation	Accuracy
11,18	1,226	74,53 %

Table 4.1: Accuracy -number of correctly identified seizures - and Standard deviation of classification algorithm

Results are compared to classification results obtained asking to seven clinicians (of different hospitals) to define if a patient has temporal or extratemporal lobe epilepsies having in mind only the manifestations seen during video registration of seizure onset.

Clinician #1	Clinician #2	Clinician #3	Clinician #4
50%	63,33 %	53,33%	63,33%

Clinician #5	Clinician #6	Clinician #7	Classification Algorithm
63,33%	73,33%	66,67%	74,53%

Table 4.2: Seven clinicians comparison. Here are presented results obtained with the classification algorithm and the results of classification made by seven clinicians from Besta, Niguarda and San Paolo Hospitals.

Chapter 5

Discussion

In the context of artificial intelligence, the 'ontology' term means a specification of a conceptualization that is a formal description of a set of concepts and relations between them of a domain of interest.

By means of an ontology, a semantic map of neurosurgery of epilepsy domain is created, with the result of a map that can be interpreted both by humans and machines.

The approach for building the ontology is bottom-up: it is not a huge map imposed from above, but many small maps interconnected that can be built by single individuals who are interested in formalizing a domain.

Medical field is afflicted by a problem of integration of data and information between different medical realities that make use of different technologies and terminology. In particular, since Computer Assisted Surgery seeks data from various sources related to surgical knowledge (most of the time represented in the form of discrete databases), there is the need for representation, storage and processing of surgical knowledge in a more structured way.

Through the developed ontology about epilepsy diagnosis and neurosurgery, it would be possible to unify all medical records related and make them understandable to everyone who has access. Knowledge expressed as ontology can take part in the pre-operative phase and in real time applications in the Operating Room. The developed ontology has a taxonomy composed of 452 classes, 30 object properties, 11 datatype properties and 180 individuals that refer to concepts that describe the domain of diagnosis and neurosurgery of epilepsy. For example, the ontology contains different kinds of Medical Imaging (MI) techniques, different operation phases that are used to scan the time of a procedure (i.e. different operative phases uses different Medical Imaging techniques), different surgical procedures linked to the treatment of the epilepsy and also a list of instruments and robots usually used in neurosurgery for epilepsy procedures (e.g. ultrasonic aspirator, curette, bipolar forceps or NeuroMate robot). This seed ontology would collect also discrete steps (in the action class) that could describe a surgical procedure, thus surgical steps (such as incision of dura mater or insertion of depth electrode), in such a way to guide a surgeon in its work.

Then, every step could be expanded with annotated medical images -of the patient- or additional information to make the step clearer. With the developed ontology every patient can be registered, and each new specific procedure can be used to create -along with many others- a surgical model -made of concepts and relations- that, in future, can guide every surgeon that needs to execute that type of procedure or suggest the best surgical approach for the specific surgical case.

As it is already been said, ontology is composed of many small interconnected maps: a clear example in the presented ontology is the map that model ictal symptoms of a patient. In fact, ontology can be used to make reasoning and help in decisions making and diagnosis.

From the dataset that we have at disposition and that we have included in the implemented ontology (first goal of this work), some considerations can be made. Reminding that symptoms during epileptic seizures may be subjective only (epileptic aura, with clear consciousness) or may progress to seizure signs that can be observed and analysed when recorded during EEG recordings, often associated with impairment of awarness, temporal patients usually present ascendent epigastric sensations and accompained -or isolated- to this manifestation, ipsilateral gestural automatism and contralateral dystonic motor sign. Symptoms like the deviation of eyes or head are present most of the time equivalently in the two form of epilepsies, conversely, gustatory, vestibular and auditory symptoms, the localizing significance of which remains uncertain, were found more frequently associated to eTLE patients. However, not all the seizures are so easily recognizable and, particularly the hypermotor gestural automatisms, can occur in bot TLE and eTLE. For this reason cases with such a symtpom have to be well studied, before ascertaining that seizures arise from the temporal lobe only, especially when the movement is contralateral to the side of the seizure onset, and also when associated with anxiety. Looking to the tab. 4.2, it is clear that classification of the two groups (TLE and eTLE) for clinicians (expert and not-expert) on the basis of general clinical features is very hard work. With the help of the ontological modeling of an epilepsy seizure, this work try to improve the actual accuracy achieved by clinicians.

All the information stored in the ontology, is then compressed in a matrix with dimensions 29*29 (i.e. the number of checked manifestations for each patient): this matrix assumes different values for each patient, and has maximum value in correspondence of symptoms observed on the diagonal, while the minimum value for the correlation between a symptom which is present in the considered video registration and the other symptoms not present. The ontological modeling help us in finding a mean correlation matrix for the two populations of patients and, in order to have statistical significance, the alghortim has been perfomed 100 times to calculate the mean value of the performances.

The variable that has to be optimized is the number of patients thas will be used to calculate the model of correlation matrices for TLE patients and eTLE patients: these models will help in assigning new patient, with its correlation matrix, to one of the two types of epilepsies. Calculated models give information relatively to the occurrence of a symptom and the correlation with other symptoms, in addition to the temporality, thus the time of onset during epileptic seizure.

In order to choose the best value for the variable to be optimized it is considered the cost function 3.2 (i.e. a fraction that expresses the number of insuccess on the number of successes) that has an exponential trend that goes to zero with increasing the number of patients used for training (fig. 4.8).

The function has an minimum (optimum) in correspondence of the value 19. It has been considered the mean value and also the standard deviation of the cost function, respectively 0.535 and 0.375.

A new set of 19 patients with TLE and 19 patients with eTLE patients have been randomly extracted and used to train the classifier. The results, shown in the fig. 4.11, present an overall accuracy (mean successes) of 74,53 % and a standard deviation of 1,1226 (number of successes that deviate from the average that is 11,18 patients correctly identified). This result gives, compared to the tests executed on 7 clinicians from Besta, Niguarda and San Paolo Hospitals, definitely a good outcome. The accuracy of classification of the alghoritm presented in this work is 12,62 % better than the classification made by clinicians (mean). Both the algorithm and the clinicians make decisions only on the basis of video registrations of ictal seizures.

It is clear that there are many cases of epileptic seizure that do not present many and distinct symptoms and this lead to problem of classiffication: for example there are patients of both TLE and eTLE groups that can present simply deviation symtpoms (head or eyes).

In this sense we can think of an evolution of the ontology and also the integration of additional information of the patient:

Ontology evolution. As we already know, ontological model differ from database model in the fundamental aspect of what is defined and what is not defined. In this sense, ontological model takes the approach of Open World Assumption, which means, the information provided to the model is assumed to be partial and open for any new additional piece of information.

The developed ontology can be enlarged, adding part of others available ontologies: for example adding concept that will complete the workflow of surgical procedures executed by surgeons that have been interviewed, and then thinking of implementing it in a surgical planner. In this context the EuRoSurge project play a key role in the sense that, since the ontology is published and shared online at the `http://www.rsp.inf.elte.hu` portal, registered users, that are experts of different fields (robotics, surgery, engineering...), can provide their contributions to the ontology, making it larger and above all it will be checked, avoiding wrong statements (par.).

Algorithm evolution. There are some suggestions that clinicians have reported to us: first of all, in order to add new clinical information of a patient, it would be helpful to insert a new sign, that clinicians named 'staring', a symptom in which patient stare at seizure onset, and which is more frequent during temporal lobe epilepsies. An additonal variable could improve the classifier, even more if that variable has higher occurrence in a set of the population.

A greater evolution of the algorithm could regard the integration of different types of clinical information of patients affected by epilepsies. It is known that for the diagnosis of epilepsy requires careful evaluation of symptoms and clinical history, which should be preferably include detailed observations by other persons, since the alteration or loss of consciousness often preclude a description of symptoms by the patient himself. The electroencephalogram (EEG) detects electrical activity in the brain and it is a fundamental analysis in the diagnosis of epilepsy, because elctrical alterations, often very indicative, could be present also in the absence of other symptoms.

Other diagnostic tests include magnetic resonance imaging or CT (to the brain) and laboratory tests, which can verify or exclude specific causes. In this sense, with the help of the proposed ontological modeling, different types of medical images could be unified (for example medical image could be annotated to extract particular information) and then processed to obtain new information and more accurate diagnosis.

This section takes into account some points of the work just presented in order to suggest some modification or evolution that would improve the algorithm, thus the results.

Appendix

Epilepsy overview

From Wikipedia[36], epilepsy is defined as "*a common and diverse set of chronic neurological disorders characterized by seizures*". Some definitions require that seizures be recurrent and unprovoked, but others require only a single seizure combined with brain alterations which increase the chance of future seizures [34].

Sometimes a cause cannot be defined; however, factors that are associated with epileptical seizures include brain trauma, brain cancer and drug and alcohol misuse among others. Thus epilepsy is not a singular disease and it reflects underlying brain dysfunction that may results from many different causes.

A definition is a useful tool for communication, diagnosis and differential diagnosis. One of the accepted definitio for the *epilepsy* term (by ILAE) is: *epilepsy is a disorder of the brain by an enduring predisposition to generate epileptic seizures and by the neurobiologic, cognitive, psychological, and social consequences of this conditions. The deifnition of epilepsy requires the occurrence of at least one epilepstic seizure [34].*

An epileptic seizuzre is "*transient*", demarcated in time, with a clear start and finish. Termination of an epileptic seizure often is less evident than the onset, because symptoms of the postictal state can blur the end of the seizure.

The start and finish of an epileptic seizure can be determined on a behavioural or EEG grounds, but we must recognized that these two operational criteria do not always coincide. An epileptic seizure is a clinical event; therefore sign and symptoms must feature prominently in the definition.

Detailed specification of subjective and objective clinical phenomena during an epileptic seizure is difficult, because of the wide range of possible manifestations.

Seizure presentation depends on location of onset in the brain, confounding disease processes, sleep-wake cycle, medications, and a variety of other factors. They can affect sensory, motor, and autonomic functions; consciousness; emotional state; memory; cognition; or behavior.

Epilepsies are classified in five ways:

1. by their first cause;

2. by the observable manifestations of the seizure, known as semiology;
3. by the location in the brain where the seizure originate;
4. as a part of discrete, identifiable medical syndromes;
5. by the event that triggers the seizures, such as reading or music;

There are many types of epilepsies syndromes. Epilepsy classification includes more information about the person and the episodes than seizure type alone, such as clinical features (e.g. behavior during seizure) and expected causes [34].

Syndromes can be divided into different groups:

- benign Rolandic epilepsy;
- frontal lobe epilepsy;
- infantile spasms;
- juvenile myoclonic epilepsy;
- juvenile absence epilepsy;
- childhood absence epilepsy;
- Lennix-Gastaut syndrome;
- Landau-Kleffner syndrome;
- progressive myoclonus;
- catamenial epilepsy,
- Lafora disease;
- limbic epilepsy;
- temporal lobe epilepsy;

Surgical treatment of epilepsy concerns any neurosurgical procedure that have as objective the treatment of focal epilepsies not controlled by drugs therapy. Before the surgical intervention, it is necessary to evaluate the patient in a multidisciplinar way, with the aim of defining the *epileptogenic zone.*

Epilelptogenic Zone (**EZ**) was originally defined in the '50s by Jean Bancaud and Jean Talairach as "*the site of the beginning and of primary organization of the epileptic seizures*" (La règion <<d'origine des dècharges critiques et leurs voies de propagation primarie>>). This definition has been slightly updated throughout the years by some other authors: Luders H.O., Engel J., Munari C. defined it in '93 "as the area of cortex that is necessary and sufficient for initiating seizures and whose removal (or disconnection) is necessary for complete abolition of seizures" [36].

EZ must be defined using non invasive and invasive investigations able to provide information about the spatio-temporal structure of the ictal discharge. In fact, the seizure is originated by a paroxysmal dynamic activity that spreads in many directions ("*crise comme un ensemble de signes dètermin*è*s essentiellement par une activit*è *paroxystique dynamique à trajectoire le plus souvent multidirectionelle*").

Non invasive investigations should include clinical history, neurological exam, MRI and video-EEG monitoring. Moreover, other neuro-radiological and neurophysiological modalities can be adopted, such as CT, SPECT, PET, EEG-fMRI, MEG. In about one third of the subjects an invasive recording methodology, such as video-SEEG monitoring, must be performed in order to define tridimensionally the EZ. The definition of the epileptic zone is a crucial step in presurgical evaluation of patient with drug-resistant focal epilepsy.

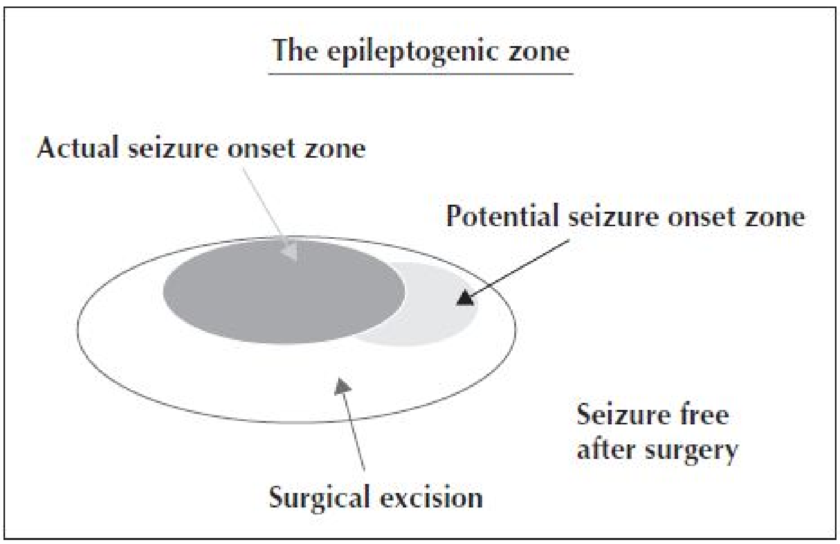

Figure 5.1: A diagram showing the actual seizure-onset zone, the potential seizure-onset zone, and a surgical resection that includes both seizure-onset zones. (provided by [27])

Temporal lobe epilepsies is the most common, and often refractory to medical therapy. In the USA, there are 2.500.000 persorns with epilepsy and 30% of them presents drug-resistant epilepsies and half of this percentage are temporal epilepsies.

Temporal Lobe Epilepsy (**TLE**) means that the seizures arise in the temporal lobe of the brain. Experiences during temporal lobe seizures vary in intensity and quality.

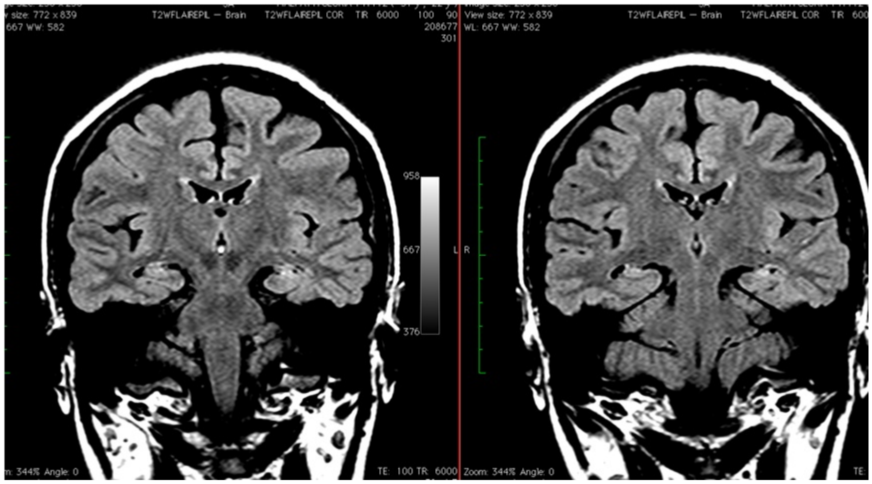

Figure 5.2: MRI (coronal section) of TLE patient with mesial temporal sclerosis (sx).

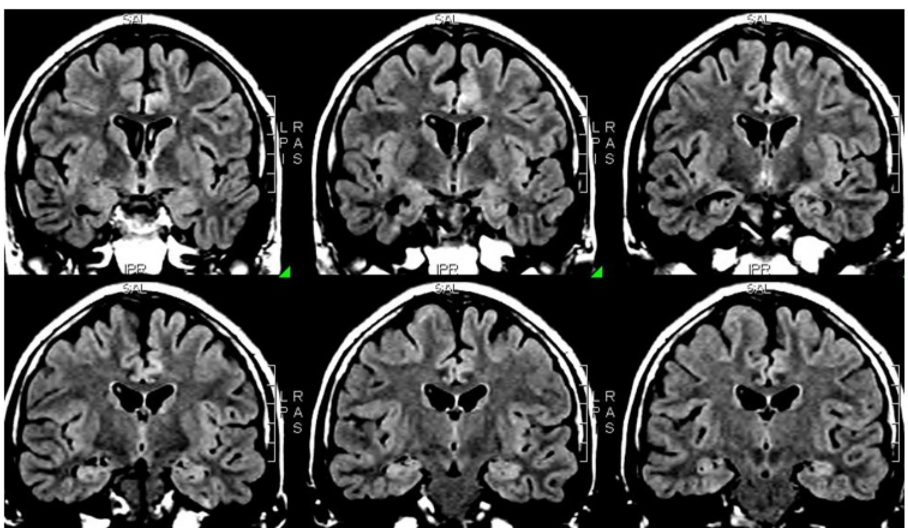

Figure 5.3: MRI (coronal section) of eTLE patient.

Sometimes the seizures are so mild that the person barely notices. In other cases, the person may be consumed with feelings of fear, pleasure, or unreality. A patient may also report an abdominal sensation that rises up through the chest into the throat, an old memory or familiar feeling, or a feeling that is impossible to describe.

The most common seizure type in TLE is a complex partial seizure. During complex partial seizures people with TLE tend to perform repetitive, automatic movements (called automatisms), such as lip smacking and rubbing their hands together [28].

Three-quarters of people with TLE also have simple partial seizures, and about half have tonic-clonic seizures at some time. Temporal lobe seizures usually begin in the deeper portions of the temporal lobe.

This area is part of the limbic system, which controls emotions and memory. This is why the seizures can include a feeling of déjà vu, fear, or anxiety, and why some people with TLE may have problems with memory and depression.

In literature there are different works in which discriminant functions for TLE an eTLE are studied. In the work of J. I. Breier *et al.* the ability of pre-operative memory performance to distinguish between patients who had been diagnosed as TLE and eTLE is examined. Analyses indicated that the use of a combination ofmeasures that assess different aspects of memory were of significant value in distinguishing between TLE and eTLE patients. This approach, as compared the use of single measure, improved classification rates [42].

Bibliography

[1] J. Gennari, *The evolution of Protègè: An environment for knowledge-based systems development,* International Journal of Human-Computer Studies, Volume 58, pages 89 - 123, 2003;

[2] T.R.Gruber, *A translation approach to portable ontologies,* Knoledge Acquisition, Volume 5, pages199-220, 1993;

[3] J.F. Sowa, *Knowledge Representation. Logical, Philosophical,and Computational Foundations*, Brooks Cole Publishing, Pacific Grove, CA, 2000;

[4] A. Venturi, *Sviluppo di un applicativo per il popolamento automatico di una ontologia tramite l'analisi del contenuto di pagine web,* 2006 (thesis, University of Bologna);

[5] <http://www.w3c.it/papers/wsb08.pdf>. Date of last consultation: 15th September 2012;

[6] Natalya F. Noy, Nigam H. Shah, *BioPortal: ontologies and integrated data resources at the click of a mouse*, Nucleic Acids Research, Volume 37, pages 170-173 2009;

[7] B. Smith, M. Ashburner, C. Rosse, J. Bard, et al., *The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration,* Nat Biotechnology, Volume 25, pages 1251-1255, 2007;

[8] The Gene Ontology Consortium, Micheal Ashburner, Catherine A. Ball, et al., *Gene Ontology:tool for the unification of biology*, Nature Genetics, Volume 5, pages 25–29, 2000;

[9] P. Jannin*,* X. Morandi, *Surgical models for computer-assisted neurosurgery*, Neuroimage, Volume 37, 783-791, 2007;

[10] V. Sandrucci, Ontological Software Architecture: Priciples and Applications*,* 2011 (thesis, University of Florence)*;*

[11] M. Peleg, N. Asbeh et al., *Onto-clust: A methodology for combining clustering analysis and ontological methods for identifying groups of comorbidities for developmental disorders,* Volume 42, pages 165-175, 2008;

[12] M. Wang, C. Lee, *Ontology-based Fuzzy Inference Agent for Diabetes Classification*, Fuzzy Information Processing Society, pages 79 - 83, 2007;

[13] D. Allemang, J. Hendler , *Semantic Web for the Working Ontologist*, M. Kauffmann, M. E. James, 2011;

[14] <http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/#intro>. Date of last consultation: 6th November 2012;

[15] B. Chandrasekaran, R. Josephson, V. Richard Benjamins, *What Are Ontologies, and Why Do We Need Them?,* IEEE intelligent systems, 1999;

[16] Thomas R. Gruber, *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*, International Journal Human-Computer Studies, 1993;

[17] <http://en.wikipedia.org/wiki/Epilepsy>. Date of last consultation: 2nd February 2013;

[18] Fisher R., van Emde Boas W., Blume W., Elger C., Genton P., Lee P., Engel J., *Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy* (IBE), Epilepsia, Volume 46, 470-472, 2005;

[19] International League Against Epilepsy, *Proposal for a revised clinical and electroencephalographic classification of epileptic seizures*, Volume 22, pages 489–501, 1981;

[20] H. Luders, J. Acharya, C. Baumgartner, *Semiological Seizure Classification*, Epilepsia, Volume 39, pages 1006-1013, 1998;

[21] Blume et al., *Glossary of descriptive terminology for ictal for ictal semiology,* Epilepsia, Volume 42, pages 1212–1218, 2001*;*

[22] C. Barba, G. Barbati, L. Minotti, D. Hoffmann, P. Kahane, *Ictal clinical and scalp-EEG findings differentiating temporal lobe epilepsies from temporal 'plus' epilepsies,* Brain, Volume 130, pages 1957-1967, 2007;

[23] Henkel et al., *The localizing value of the abdominal aura and its evolution: a study in focal epilepsy*, Neurology, Volume 58, pages 271-276, 2002;

[24] Kotagal P., Arunkumar G., Hammel J., Mascha E., *Complex partial seizures of frontal lobe onset: analysis of ictal semiology*, Seizure, Volume 12, pages 268-281, 2003;

[25] Bleasel A., Kotagal P., Kankirawatana P., Rybicki L., *Lateralizing value and semiology of ictal limb posturing and version in temporal lobe and extratemporal epilepsy*, Epilepsia, Volume 38, pages 168-174, 1997;

[26] S. Shorvon, *The treatment of epilepsy*: *Forms, Causes and Therapy in Children and Adults*, Emilio Perucca, Oxford: Blackwell Science, 1996.;

[27] Luders H.O., Engel J., Munari C., *Surgical Treatment of the epilepsies*, New York: Raven Press, Volume 8, pages S1-S9, 1993;

[28] <http://epilepsy.med.nyu.edu/epilepsy/types-epilepsy/temporal-lobe-epilepsy>. Date of last consultation: 25 February 2013;

[29] Mudunuri Raj, Oliver Burgert, Thomas Neumuth, *Ontological modeling for surgical knowledge*, pages 1044-1054, 2011;

[30] Roman Timofeev, Classication and Regression Trees (CART) Theory and Applications, (thesis: Humboldt University, Berlin), 2004;

[31] Eric Miller, *An introduction to the Resource Description Framework*, Bulletin of the American Society for Information Science and Technology, 1998;

[32] <http://www.w3.org/TR/rdf-schema/#ch_range>. Date of last consultation: 6th November 2012;

[33] <http://www.w3.org/TR/owl-ref/#Ontology-def>. Date of last consultation: 15th September 2012;

[34] W. T. Blume, *Complex partial seizures - clinical description and diagnosis*, Advances in Experimental Medicine and Biology, Volume 497, pages 9-18, 2002;

[35] L. Zhang, Z. Wang, *Ontology-based Clustering Algorithm with Feature Weights,* Journal of Computational Information Systems, Volume 25, pages 15-19, 2010*;*

[36] A. Gómez-Pérez, O. Corcho, Ontology Languages for the Semantic Web*,* (thesis: Universidad Politécnica de Madrid), 2002;

[37] R. Karp, V. Chaudhri, and J. Thomere, *XOL: An XML-Based Ontology Exchange Language* (version 0.4), Volume 17, pages 54-60, 2002;

[38] <http://www.cs.umd.edu/projects/plus/SHOE/spec1>, Date of last consultation: 30th September 2013;

[39] O. Lassila, R. Webick, *Resource Description Framework (RDF) Model and Syntax Specification*, Volume 17, pages 54-60, 2002;

[40] G. Jackson, R. Briellmann, R. Kuzniecky, *Temporal lobe epilepsies,* Magnetic Resonance In Epilepsy: Neuroimaging Techniques, 2005;

[41] M. Audette, H. Yang, A. Enquobahrie*, The application of textbook-based surgical ontologies to neurosurgery simulation requirements*, International Journal of Computer Assisted Radiology and Surgery, Volume 6, pages 138-140, 2011;

[42] J. I. Breier, P. M. Plenger, J. W. Wheless, A. B. Thomas, B. L. Brookshire, V. L. Curtis, A. Papanicolaou, L. J. Willmore, G. L. Clifton, *Memory Tests Distinguish Between Patients with Focal Temporal and Extratemporal Lobe Epilepsy,* Epilepsia, Volume 37, pages:165-170, 1996;

[43] <http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html>. Date of last consultation: 30th January 2013;

[44] <http://www.mathworks.it/it/help/nnet/gs/recognizing-patterns.html>. Date of last consultation: 25th February 2013;

[45] <http://en.wikipedia.org/wiki/Ontology_(information_science)>. Date of last consultation: 15th September 2012;

[46] P. Jannin, M. Raimbault, X. Morandi, B. Gibaud, *Modeling Surgical Procedures for Multimodal Image-Guided Neurosurgery,* Medical Image Computing and Computer-Assisted Intervention, Volume 2208, pages 565-572, 2001;

[47] HeinrichăHerre*, General Formal Ontology (GFO): A Foundational Ontology for Conceptual Modelling*, Theory and Applications of Ontology: Computer Applications. Springer Netherlands, 297-345, 2010;

[48] R. Hoehndorf, M. Dumontier, A. Oellrich, *Interoperability between Biomedical Ontologies through Relation Expansion, Upper-Level Ontologies and Automatic Reasoning*, I. P. Androulakis, Ontologies and Automatic Reasoning, 2011;

[49] <http://trajano.us.es/~isabel/EHR/>. Date of last consultation: 10th September 2012;

[50] P. L. Whetzel, N. F. Noy*, BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications,* Nucleic Acids Research, Volume 39, pages 541–545, 2011;

[51] T. Tudorache, N. Noy, S. Tu, A. Musen, *Supporting Collaborative Ontology Developments in Protègè*, ISWC '08 Proceedings of the 7th International Conference on The Semantic Web, 2008;

[52] A. L. Rhoton, *Operative techniques and instrumentation for neurosurgery,* Neurosurgery, 2003, Volume 5, pages 907-934;

[53] G. D. Cascino*, Epilepsy and Brain Tumors: Implications for Treatment*, Epilepsia, Volume 31, pages 37–44, 1990;

[54] S. Lavallee, J. Troccaz, L. Gaborit, P. Cinquin; A. L. Benabid, D. Hoffmann*, Image guided operating robot: a clinical application in stereotactic neurosurgery,* Robotics and Automation, Volume 1, pages 618-624, 1992*;*

[55] M. S. Eljamel, *Robotic application in epilepsy surgery,* Issue The International Journal of Medical Robotics and Computer Assisted Surgery, Volume 2, pages 233–237, 2006*;*