

Towards an hybrid system for annotating brain MRI images

Ammar Mechouche¹, Christine Golbreich², Bernard Gibaud¹

¹ INSERM, U746, Campus de Villejean, F-35043 Rennes, France

¹ INRIA, VisAGeS, Univ. Rennes I, CNRS, UMR 6074, F-35042 Rennes, France
{Ammar.Mechouche, Bernard.Gibaud}@irisa.fr

²LIM, University Rennes I, Av du Pr. Léon Bernard, 35043 Rennes, France
Christine.Golbreich@univ-rennes1.fr

Abstract

This paper describes a method combining symbolic and numerical techniques for annotating brain Magnetic Resonance images. The goal is to assist existing automatic labelling methods which are mostly statistical in nature and do not work very well in certain situations such as the presence of lesions. The system uses existing statistical methods for generating ABox facts that constitute a set of initial information sufficient for fruitful reasoning. The reasoning is supported by an OWL ontology extended with rules, and the facts extracted by the statistical component

1 Introduction

Identifying anatomical structures in brain Magnetic Resonance Images (MRI) is an important aspect of the preparation of a surgical intervention in neurosurgery, especially when the lesion is located in the cerebral cortex. A precise labelling of cortical structures (gyri, sulci) surrounding the lesion is particularly necessary to determine an optimal surgical strategy. Existing automatic approaches for annotating brain images are often statistical, e.g., based on Statistical Probability Anatomy Maps (SPAMs). A SPAM is a 3D probabilistic map associated to a particular anatomical structure. The value at each voxel position represents the probability of belonging to this structure at that location. The statistical data used in our system were derived from a database of 305 normal subjects, after re-alignment of MRI data into a common reference system (called stereotaxic space). SPAMs-like methods have an important drawback. They are not robust against deformations and shifts caused by a lesion in the brain. A symbolic method, using a priori knowledge about topological relations between the cerebral structures may be an alternative or a complement to compensate it, since in contrast topological relations are preserved. This paper describes a new hybrid

method for annotating brain images where SPAMs are used to get a sufficient set of initial facts for reasoning. Reasoning is supported by an OWL [1] ontology about the brain cortex anatomical structures and Horn rules capturing the topological dependencies between the brain structures.

2 Method

The method consists of two main steps. The first step is the segmentation of the brain and the extraction of the sulci tracks from an MRI exam. The second step, is the annotation of a region of interest (ROI) selected from the sulci graph. This paper mainly focuses on the second step.

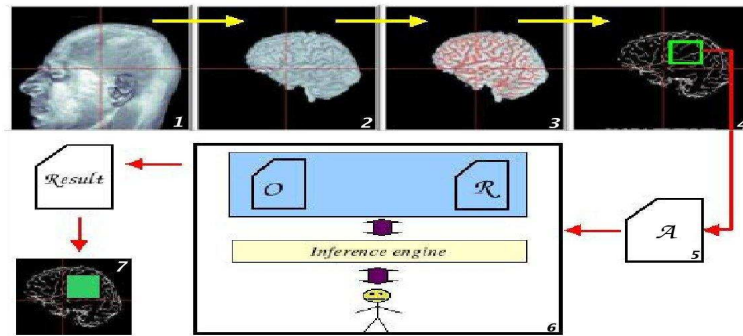


Figure 1: The general diagram

Reasoning is performed from an ontology of the brain structures enriched by rules representing their topological dependencies, and initial facts provided by numerical and statistical tools (SPAMs). Figure 1 illustrates the complete process of the application: (1) acquiring the patient MRI ; (2) brain segmentation; (3) extraction of the external tracks of the sulci; (4) selection by the user of a region of interest and extraction of the corresponding subgraph of sulcus segments delimiting surfaces corresponding to the parts of gyrus (called patches) present in the region; (5) initialization of the ABox \mathcal{A} . The above numerical and statistical treatments lead to the initial facts, OWL individuals and role values representing their topological relations, as explained thereafter; (6) reasoning based on the brain ontology \mathcal{O} , the rule base \mathcal{R} , and the ABox \mathcal{A} (with some user interaction), (7) inferred labels of the structures involved in the ROI.

2.1 Populating the Abox

First, the numerical tools extract the sulci of a ROI and delimit the surfaces limited by the sulci (patches). They also provide the topological relations and the orientations between the different patches and sulcus segments.

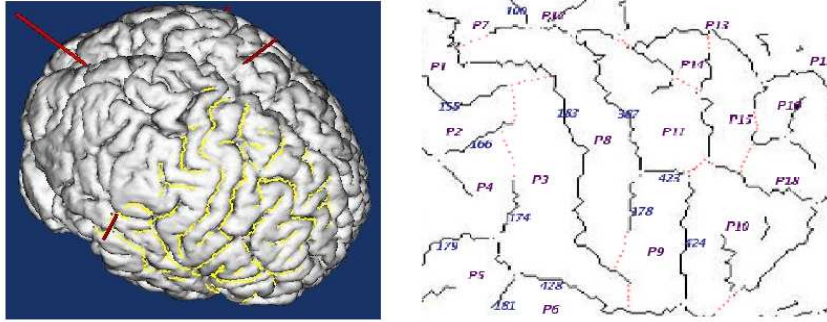


Figure 2: Extraction of the ROI graph

The yellow segments figure 2 (left) show the sulci of the ROI. The patches (e.g.; $P7$, $P8$, $P9$ etc.) are delimited by the sulci (e.g.; 178, 124 etc.) (right). The facts extracted by the numerical tools from this graph are represented in OWL DL (figure 3 left). For example $P9$ is an individual of the class Patch while 178 is a SulcusSegment. The property *isMAEBoundedBy* has a value (individual of the class *AttributedEntity*) expressing that $P9$ is bounded by the segment 178 with a posterior orientation, and other segments 423, 424 etc.

```

<Patch rdf:ID="p9">
...
  <isMAEBoundedBy>
    <AttributedEntity rdf:ID="AttEntity1">
      <entity>
        <SulcusSegment rdf:ID="178"/>
      </entity>
      <orientation>
        <Posterior rdf:ID="posteriorTo"/>
      </orientation>
      <orientation>
        <Unknown rdf:ID="unknown1"/>
      </orientation>
      <orientation>
        <Unknown rdf:ID="unknown2"/>
      </orientation>
      <MAEBounds rdf:resource="#p9"/>
    </AttributedEntity>
  </isMAEBoundedBy>
...
</Patch>

```

```

<owl:Class rdf:ID="Orientation">
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Posterior"/>
  </rdfs:subClassOf>
</owl:Class>

<owl:ObjectProperty rdf:ID="isMAEBoundedBy">
  <inverseOf>
    <owl:ObjectProperty rdf:about="#MAEBounds">
      </inverseOf>
    </owl:ObjectProperty>
  </inverseOf>
</owl:ObjectProperty>

```

Figure 3: Facts in the ABox (left), OWL class and ObjectProperty (right)

These facts are then completed by data computed from the SPAMs. A SPAM is a 3D image file associated to a particular anatomical structure, for instance, a particular gyrus. The information at each point of this 3D image $pt(x, y, z)$ represents an estimate of the probability to belong to this particular structure. Each segment s_i of the ROI is a set of points. We first transform the points coordinates into coordinates of the reference space, i.e. the stereotaxic space.

SulcusSegment \ Gyrus	Precentral	Postcentral	Angular	SupTemporal	...
ID = 183	0.384	0.186	0	0	.
ID = 178	0.218	0	0	0	.
ID = 155	0	0.477	0.105	0	.
ID = 298	0	0.038	0	0.076	.

Figure 4: Example of probabilities

Then we calculate the probability p_{ij} of the segment s_i to belong to a SPAM sp_j by calculating the average of the probabilities of all its (transformed) points. For each segment we store the two highest probabilities that have been computed (figure 4). These values computed from the SPAMs, and the abstractions rules presented below enable to automatically acquire the initial facts of the Abox \mathcal{A} .

Computing the boundaries and separations. Some heuristics have been defined to determine whether a sulcus 'bounds' a SPAM or 'separates' two SPAMs: if the two highest probabilities are small, over a given threshold MIN , then the segment is asserted to separate or to bound the corresponding SPAMs, else if they are very big, over a given threshold MAX , then the segment is asserted to be inside the corresponding SPAM. The thresholds MIN and MAX are defined empirically. More precisely, the rules that abstract the topological relations regarding the boundaries and separations are:

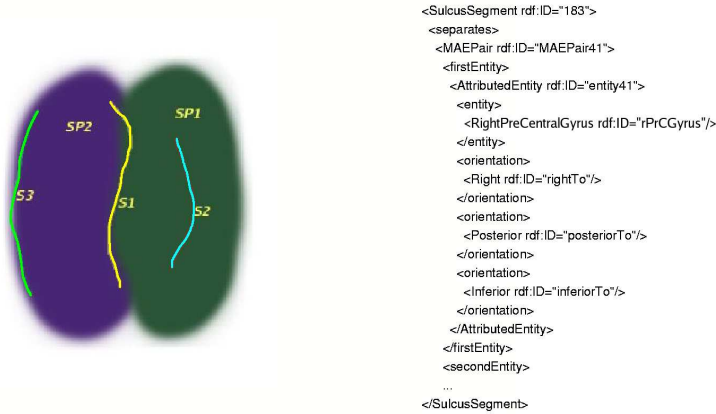


Figure 5: Computing facts about separation and boundary from SPAMS

- if $(p_{i1} \in [MIN, MAX] \text{ and } p_{i2} \in [MIN, MAX])$ then s_i separates sp_1 and sp_2 . Indeed it means that s_i is located between sp_1 and sp_2 that is s_i separates them, e.g.; S_1 figure 5).

- *if*($p_{i1} \in [MIN, MAX]$ and $p_{i2} < MIN$) then s_i bounds sp_1 , indeed low values mean that s_i is located at the extremity of the SPAM thus it is a boundary (e.g.; S_3 figure 5)
- *if*($p_{i1} > MAX$ and $p_{i2} < MIN$) then s_i is *Inside* an instance sp_1 , indeed these values indicate that s_i is within the SPAM (e.g.; S_2 figure 5)

Computing the orientations. Each entity has three orientations: (Right or Left), (Posterior or Anterior) and (Superior or Inferior). To determinate the orientations of the segments w.r.t SPAMs, for example that a segment s_i bounds a SPAM sp_j with an anterior orientation, we compare the coordinates (x, y, z) of the centre of the segment, transformed into the reference space, to the coordinates (x', y', z') of the SPAM centre.

The heuristic rules below abstract the orientations:

- *if* $x > x'$ then s_i is *RightTo* sp_j else s_i is *LeftTo* sp_j
- *if* $y > y'$ then s_i is *AnteriorTo* sp_j else s_i is *PosteriorTo* sp_j
- *if* $z > z'$ then s_i is *SuperiorTo* sp_j else s_i is *InferiorTo* sp_j

Since such rules could not be used by KAON2, they were applied using a C++ procedural program. The resulting role values, e.g.; *separates*($s_0, (prcgr, anteriorTo, rightTo, superiorTo)$), (*pcgr, posteriorTo, leftTo, inferiorTo*) are represented in OWL (figure 5).

2.2 Brain ontology and rules

The knowledge base consists of the brain ontology enriched with rules. For the moment, the ontology about the sulci and gyri is represented in OWL DL, the rules in SWRL. They have been edited using Protégé OWL and the SWRL plugin [2](figure 6).

- **Tbox:** the *Tbox* provides the logical definitions of concepts (classes), roles (properties) and the asserted axioms. For example, the necessary and sufficient condition to be a segment of the right central sulcus is¹ (figure 6): $RightCentralSulcusSegment \equiv ((\exists MAEBounds ((\exists entity (\exists partOf RightPostCentralGyrus)) \sqcap (\exists orientation Anterior))) \sqcap ((\exists MAEBounds ((\exists entity (\exists partOf RightPreCentralGyrus)) \sqcap (\exists orientation Posterior))))$. This OWL definition expresses that a segment of central sulcus is bounded by a part of postcentral gyrus with an orientation which is an instance of Anterior, and is bounded by a part of precentral gyrus with an orientation which is an instance of Posterior.
- **Rule-box:** The *Rule-box* contains all the rules extending the ontology, for example the rule R1 (figure 6) expresses that a boundary is propagated

¹this is not the exact definition but a simplification for the example

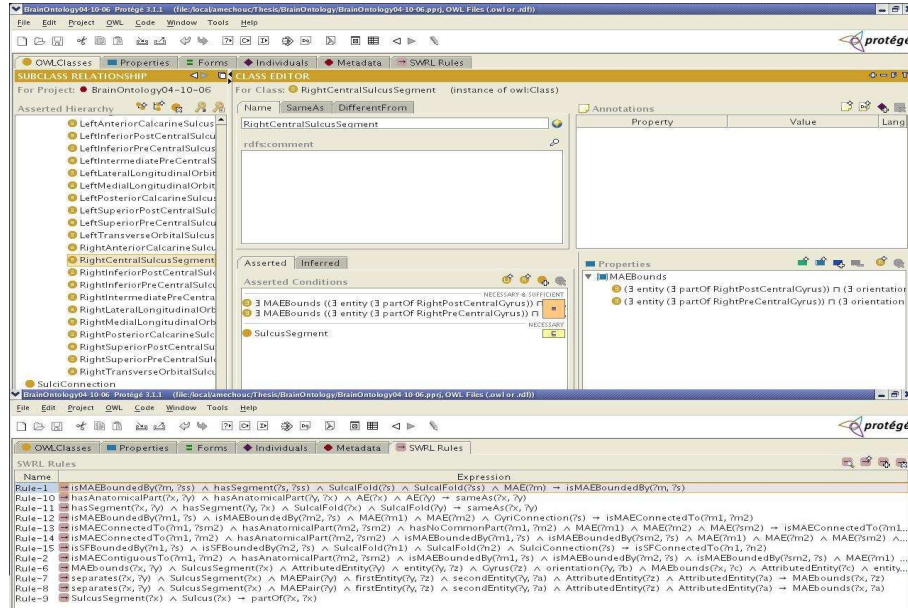


Figure 6: Brain ontology and rules

from parts to whole: $isMAEBoundedBy(?x, ?y) \wedge hasSegment(?z, ?y) \wedge SulcalFold(?z) \wedge SulcalFold(?y) \wedge MAE(?x) \rightarrow isMAEBoundedBy(?x, ?z)$. If a material anatomical entity x is bounded by a sulcal fold y , and y is a segment of z , then x is bounded by z . Such rules are needed to infer the missing knowledge of the classes definitions for instance retrieval. Rules are also useful to express queries. For example, to find all possible instances of gyri of which patches p_i of a ROI are part: $Q(?x_i, \dots, ?x_n) \leftarrow \bigwedge_{i=1 \text{ to } n} (AE(?x_i)) \wedge partOf(p_i, ?x_i)$.

- **Abox:** The *Abox* contains the individuals (instances of classes) and the instances of relations between them as defined section 2.1.

All the knowledge, the ontology in OWL DL (Tbox), the Horn rules (Rulebox), and the facts (Abox), are gathered within a single file provided as input to the reasoner.

3 Reasoning for brain labelling

Figure 7 shows the overall process of reasoning: (1) From the list of sulci (segments) of the ROI and the list of SPAMs we get a table of probabilities (such as Figure 4). This table is first created as an XML file. The heuristics presented above derive the topological relations between the anatomical entities.

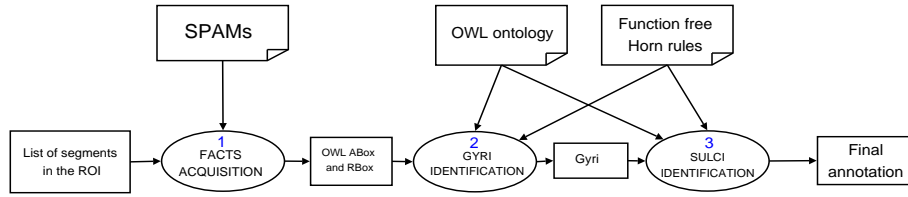


Figure 7: Labelling process

The resulting facts are stored in an OWL file. This file is merged with the ontology, the rules, and the other facts coming from the numerical tools. (2) From this file, the inference engine labels the patches as described below. The user validates the result. (3) Next, the reasoner labels the sulci according to the ontology definitions enriched by rules. The user validates this step, and finally the labelled image is obtained. The reasoning is performed as follows :

Labelling the patches. The patches are first labelled using the rules below. The main rule used is a rule ($\mathcal{R}_{\mathcal{M}}$) that makes a matching between the facts extracted from the images by the numerical tools and the facts computed about their boundaries and orientations w.r.t the SPAMs:

- $\text{MAEBounds}(?x, ?c) \wedge \text{AttributedEntity}(?c) \wedge \text{entity}(?c, ?d) \wedge \text{Patch}(?d) \wedge \text{orientation}(?c, ?b) \wedge \text{MAEBounds}(?x, ?y) \wedge \text{SulcusSegment}(?x) \wedge \text{AttributedEntity}(?y) \wedge \text{entity}(?y, ?z) \wedge \text{Gyrus}(?z) \wedge \text{orientation}(?y, ?b) \wedge \text{Orientation}(?b) \rightarrow \text{partOf}(?d, ?z)$

This rule expresses that if it is known from the extracted facts that a segment x bounds a given patch d with a given orientation b and it comes out from the computed orientation that x bounds a SPAM z with the same orientation, then this patch belongs to the gyrus corresponding to that SPAM. The probability associated to it is the probability p_i calculated as explained section 2.1.

In fact, the above rule is simplified and the real rule used involves a *sameAs* construct. Indeed, as the Abox \mathcal{A} is issued from two generated files that are merged into a single OWL file, a same individual has got different ID depending on the file it comes from:

- $\text{MAEBounds}(?x, ?y) \wedge \text{SulcusSegment}(?x) \wedge \text{AttributedEntity}(?y) \wedge \text{entity}(?y, ?z) \wedge \text{Gyrus}(?z) \wedge \text{orientation}(?y, ?b) \wedge \text{MAEBounds}(?x, ?c) \wedge \text{AttributedEntity}(?c) \wedge \text{entity}(?c, ?d) \wedge \text{Patch}(?d) \wedge \text{orientation}(?c, ?b') \wedge \text{Orientation}(?b) \wedge \text{Orientation}(?b') \wedge \text{sameAs}(?b, ?b') \rightarrow \text{partOf}(?d, ?z)$

As there is a possible incertitude in the computed orientations mainly due to the approximations caused by the SPAMs, it may occur that a segment bounds two SPAMs with the same orientation, hence a patch d is inferred to belong to several gyri z_i with probabilities p_i . To decide to which gyrus d finally belongs,

we calculate $\sum(p_i)$ for each gyrus and keep the gyrus with the highest result.

The second rule (\mathcal{R}_S) below infers boundaries from a separation: if a given sulcus separates two gyri then it bounds each of them. This rule is used to infer boundaries from the known separations, information which is needed to fire the first rule above.

- $\text{separates}(\text{?x}, \text{?y}) \wedge \text{SulcusSegment}(\text{?x}) \wedge \text{MAEPair}(\text{?y}) \wedge \text{firstEntity}(\text{?y}, \text{?z}) \wedge \text{secondEntity}(\text{?y}, \text{?a}) \wedge \text{AttributedEntity}(\text{?z}) \wedge \text{AttributedEntity}(\text{?a}) \rightarrow \text{MAEBounds}(\text{?x}, \text{?z})$

Labelling the sulci. After the patches, the sulci are next labelled thanks to the ontology definitions and the rules. If a given segment s_i satisfies a definition of a given sulcus su_j in the ontology, i.e. if it meets the necessary and sufficient condition of su_j , then s_i is classified as an instance of su_j .

Simplified example.

- Let be a segment s_0 and two patches p_1, p_2 of the ROI.
- the facts provided by the numerical tools include the individuals p_1 and p_2 and the relation $\text{separates}(s_0, (p_1, \text{anteriorTo}, \text{rightTo}, \text{superiorTo}), (p_2, \text{posteriorTo}, \text{leftTo}, \text{inferiorTo}))$ where *anteriorTo*, *rightTo*, *superiorTo*, *posteriorTo*, *leftTo*, and *inferiorTo* are respective individuals of the classes Anterior, Right, Superior, Posterior, Left, and Inferior.
- the facts computed from the SPAMs include the relation $\text{separates}(s_0, (\text{pcgr}, \text{anteriorTo}, \text{rightTo}, \text{superiorTo}), (\text{pcgr}, \text{posteriorTo}, \text{leftTo}, \text{inferiorTo}))$ where RightPreCentralGyrus: *pcgr* and RightPostCentralGyrus: *pcgr*.

The labels of the patches are obtained by answering the query $Q(\text{?}x_i, \dots, \text{?}x_n) \leftarrow \bigwedge_{i=1 \text{ to } n} (AE(\text{?}x_i)) \wedge \text{partOf}(p_i, \text{?}x_i)$. Applying the rule \mathcal{R}_S , facts about boundaries are derived from the initial facts about separations, then as the body of the matching rule \mathcal{R}_M can be satisfied by bindings its variable to known individuals, the reasoner infers: $\text{partOf}(p_1, \text{pcgr})$ and $\text{partOf}(p_2, \text{pcgr})$. Next, at a second step, the labels of the segments are obtained from the class definitions of the ontology. As s_0 satisfies the N&S condition of RightCentralSulcusSegment, the reasoner infers that s_0 is an instance of the RightCentralSulcusSegment.

4 Results and discussion

This section presents some results obtained for real data with the method presented above. The experiments are achieved with the reasoner KAON2 [3], which accepts ontologies extended with rules [4]. The region of interest is the

x	y
a:p6	a:RightSuperiorTemporalGyrus
a:p11	a:RightMiddleFrontalGyrus
a:p9	a:RightPrecentralGyrus
a:p13	a:RightSuperiorFrontalGyrus
a:p7	a:RightPrecentralGyrus
a:p14	a:RightMiddleFrontalGyrus
a:p15	a:RightMiddleFrontalGyrus
a:p1	a:RightPostCentralGyrus
a:p1	a:RightPostcentralGyrus
a:p4	a:RightSupramarginalGyrus
a:p2	a:RightPostcentralGyrus
a:p5	a:RightPostcentralGyrus
a:p8	a:RightPrecentralGyrus
a:p3	a:RightSupramarginalGyrus
a:p10	a:RightInferiorFrontalGyrus

Figure 8: Results obtained with KAON2

automatically extracted region displayed figure 2. We used 45 SPAMs corresponding to the most important gyri of the brain. The *MIN* value approximated from the computations is 0.05 and the *MAX* value 0.75.

Labels of patches. The *SPARQL* query² *SELECT ?x ?y WHERE { ?x rdf:type a:Patch ; a:partOf ?y }* asks for each patch all the possible entities it is part of. The answers of KAON2 to that query provide the labels of the patches, for example patch *P6* is a part of the right superior temporal gyrus (figure 8).

Labels of sulcus segments. The query *SELECT ?x WHERE { ?x rdf:type a:centeral.sulcus.right }* asks for the right central sulcus segments. KAON2 returns the segment 183, which is the single segment of the right central sulcus for this ROI.

For the two experiments reported in this paper, more than 80% of the results given by the reasoner are correct regarding the labels of the patches and over 70% for the labels of the sulcus segments (according to our expertise). For example the patch *P8* is inferred to be a part of the right precentral gyrus, *P1* is inferred to be a part of the right postcentral gyrus, and the segment 183 is inferred to be an instance of the right central sulcus segment, which is correct since it separates parts of the right precentral gyrus from parts of the right postcentral gyrus. This is an ongoing work. It will be interesting in the future

²The query language of KAON2

to assess how this percentage is affected by various aspects of the ontology and rules, and the respective effect of the SPAMs and of the reasoning on the results. The proposed method was adapted to comply with some language and tools limitations, in particular with the version of the KAON2 reasoner available online and the Protégé SWRL editor, for example:

- For the moment the ontology was simplified using OWL DL instead of OWL1.1 [5]. We used existential restriction instead of qualified cardinality restrictions (QCR). But, it should be noted that the real Tbox requires both QCR, disjunctions, inverse. For example, the ontology should express that each right PostCentralGyrus is bounded exactly by one right CentralSulcus. Besides, the rules cannot be expressed as role inclusion axioms (cf. rule ($\mathcal{R}_{\mathcal{M}}$) or the rules of the online Annex at <http://idm.univ-rennes1.fr/~obierlai/>). Thus the required knowledge is not expressible in the \mathcal{EL}^{++} fragment of OWL1.1.
- we defined subclasses of Orientation, e.g.; Posterior, Anterior etc. with individual e.g., posteriorTo and used existential restrictions instead of enumeration or hasValue restrictions, because KAON2 does not support nominals.
- KAON2 reasoner is based on the DL-safe rules assumption [4]. Although the rules used for our system are not DL-safe, KAON2 provides the expected answers for the reported experiments. Indeed, in these cases the rules were fired, because given the initial facts asserted, their body was satisfied by bindings their variable to known individuals. However, this approach is not always relevant and situations may occur where solutions are missed because of the existential construct. For example, a patch is defined with an existential in the equivalent class expression (rhs). Hence, it may happen in some cases that a rule expressing the propagation of a property from parts to whole cannot be not fired, because an instance of Patch is defined without being connected to a known instance of gyrus by the relation partOf [6]. KAON2 does not draw all the consequences according to the first order semantics of SWRL, but only consequences under the "the DL-safe semantics".
- all n-ary relations were transformed into binary relations, using reification for example we defined an artificial class AttributedEntity for it. The ontology was edited using Protégé rules editor which allows to edit only SWRL rules and does not support ordinary predicate that are not DL predicates. KAON2 extends the standard SWRL syntax and offers a swrl:PredicateAtom that allows ordinary predicates, but according to the authors their SWRL extensions were still experimental at the time of these experiments. It would be preferred to have a language extension and tools allowing n-ary relations. N-ary relations is a general needs for example also encountered with the Foundational Model of Anatomy ontology which exhibits more than 30 attributed relationships and where more than 2300 nested classes were generated for their values [7].

- The heuristic rules section 2.1 were implemented in *C++*. A declarative approach was not possible with KAON2 since at the moment it does not handle OWL DL datatypes or OWL1.1 user-defined datatypes and restrictions involving datatype predicates.

5 Conclusion

This paper reports the current stage of development of an hybrid system combining numerical and symbolic techniques for brain MRI images description, and its present limitations. The method is based on an OWL DL ontology extended with rules, and facts coming from numerical tools and SPAMs. Future work will investigate how to overcome some of the work-arounds employed to circumvent the limitations encountered with the representation and tools used. At the moment the method was only tested over a limited set of brain images that did not exhibit a lesion. The experiments will be extended to more cases and to brain images exhibiting a lesion in order to assess its robustness. Automating the annotation of the semantic content of digital images presents promising perspectives for new applications such as retrieval of similar cases for decision support, or statistical medical studies in large populations.

References

- [1] OWL Web Ontology Language : Overview, <http://www.w3.org/tr/owl-features/>.
- [2] Protégé, <http://protege.stanford.edu>.
- [3] KAON2, <http://kaon2.semanticweb.org/>.
- [4] B Motik, U Sattler, and R Studer. Query answering for OWL DL with rules. In *ISWC 2004, LNCS 3298 Springer*. Springer-Verlag, 2004.
- [5] OWL 1.1 Web Ontology Language Syntax: <http://www-db.research.bell-labs.com/user/pfps/owl/syntax.html>.
- [6] C Golbreich. Web rules for Health Care and Life Sciences: use cases and requirements. In *RoW Workshop at WWW2006*, 2006.
- [7] C Golbreich, S Zhang, and O Bodenreider. The Foundational Model of Anatomy in OWL: Experience and Perspectives. In *Journal of Web Semantics, (4)3*, 2006.

6 Acknowledgement

We are grateful to Louis Collins of the Montreal Neurological Institute for providing us with the SPAMs database and to the Regional Council of Brittany for supporting this project.