Towards a Sharable Numeric and Symbolic Knowledge Base on Cerebral Cortex Anatomy: Learnings from a Prototype

Olivier Dameron[1], Bernard Gibaud[1], Anita Burgun[2], Xavier Morandi[1,3]
(1) Laboratoire IDM UPRES-EA 3192, Faculté de Médecine, F35000 Rennes, France
(2) LIM, Faculté de Médecine, F35000 Rennes, France
(3) Service de neurochirurgie, CHRU de Rennes, F35000 Rennes, France

Abstract

We propose a numeric and symbolic knowledge base about sulco-gyral brain cortex anatomy implemented using web technologies. This knowledge base is supposed to be easily reusable in various application contexts such as teaching, decision support in neurosurgery and sharing of neuroimaging data for research purposes.

Our analysis shows that those various application contexts call for a more formalised representation of taxonomical and mereotopological relations and for the expression of identity criteria for the symbolic concepts.

Keywords : cortex anatomy, knowledge representation

1 Introduction

The evolution of medical practice leads to a more complete documentation of care, that has to be shared between health care professionals and can be reused in various contexts such as research and teaching. Computerization of medical information facilitates these reuse processes. However, applications such as intelligent data retrieval and decision support require that medical data be structured, with explicit references to the related medical knowledge and concepts. Particularly, anatomy is a fundamental discipline that underlies most medical fields [1]. Consequently, it is of major interest to be able to share a representation of anatomical knowledge [2]. However, the mere transposition of the classical paper-based anatomical atlases [3-5] into electronic systems with only textual descriptions and bidimensional illustrations is inadequate [6]. The need for a context-independent and formalised representation of anatomy triggered the creation of computerized models such as Digital Anatomist [7] or GALEN [8].

Neuroanatomy teaching, surgical planning in neurosurgery and sharing of clinical or experimental neuroimaging data for research in neurosciences are application fields that make reference either implicitly or explicitly to sulco-gyral anatomy of the cerebral cortex. Previous works by Nowinski [9], Höhne [10] or Kikinis [11] rely on the segmentation and the labelling of anatomical structures in images. It is based on a limited set of individuals so anatomical variability is poorly represented, and knowledge about neighborhood relationships lays within the images in an implicit rather than explicit way. Although this may be relevant for teaching applications, we believe this approach to be irrelevant when inferences may occur, such as in decision support or in data retrieval applications.

This paper briefly describes and discusses a prototype system implementing a model of brain cortex anatomy. It represents both numeric knowledge (i.e. derived from direct measures such as images) and symbolic knowledge (i.e. derived from conceptual descriptions). This model provides a semantic reference that can be used for the description and the indexing of neuroimaging data. Our study focuses on sulco-gyral anatomy of normal cerebral cortex in man. The model was implemented using web technologies in order to make it easily sharable and therefore easily reusable in various application contexts. We analyse the limitations of this prototype according to the requirements of these application contexts. This discussion hightlights needs for a more thorough specification of symbolic relations and identity criteria associated to anatomical features.

2 Development of a prototype

2.1 Knowledge Base

We first defined a higher-level model of cortex anatomy. For the symbolic knowledge, it describes the classes of concepts such as sulci or gyri and the relations between them. This modelling was performed in UML[1] with associated textual definitions of the entities [12]. For the numeric knowledge, it consists of information about the positions and shapes of the anatomical structures. It was extracted from MRI datasets.

The symbolic entities of the model are represented in an XML[2] file. The grammar of this file is defined in a Document Type Definition (DTD), derived from entities and relations defined at the semantic level. The symbolic elements are entities such as ‘Central Sulcus‘ or ‘PreCentral Gyrus‘, and relations such as ‘is-anatomical-part-of‘ or ‘delimits‘. Numeric

[1] Unified Modelling Language
[2] eXtensible Markup Language

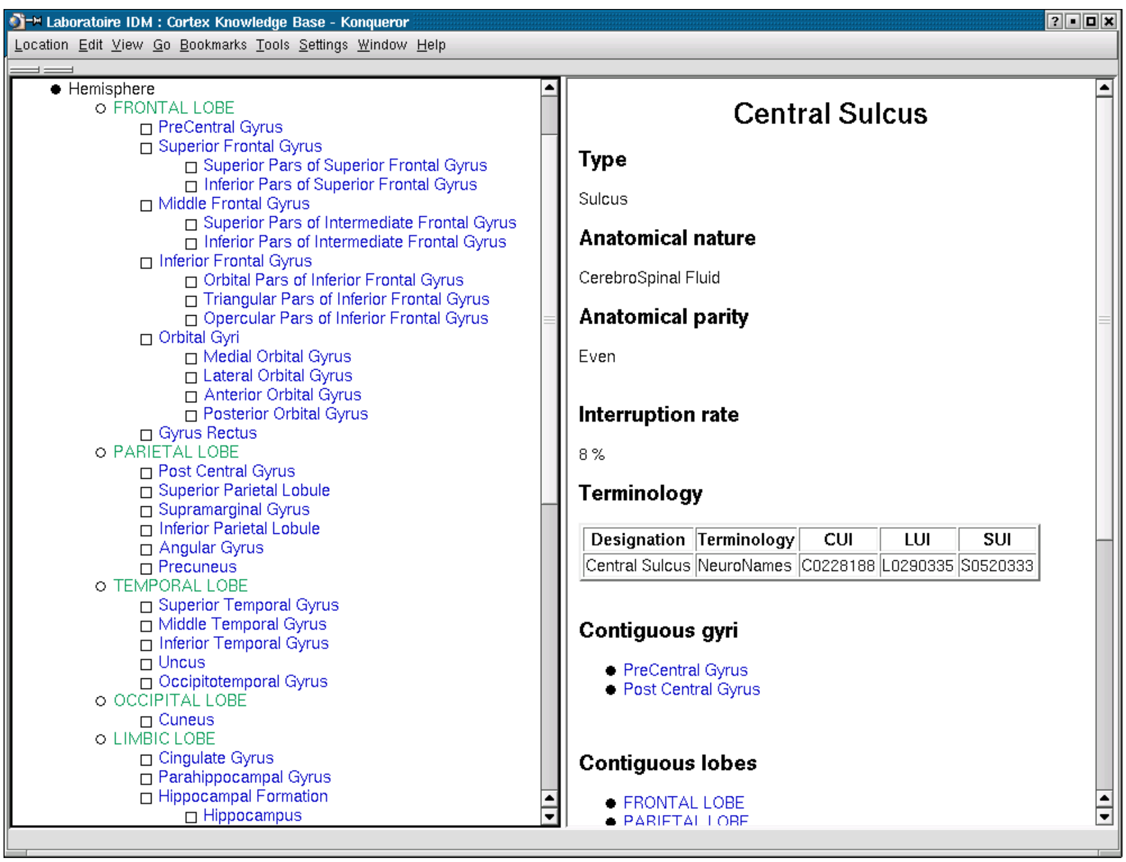
elements, among which 3D scenes, are represented in VRML[3]. The links between symbolic and numeric elements are represented using XLink compliant companion XML files for modularity.

When dealing with actual patients' data (e.g. in a decision support application), it is necessary to refer to a symbolic model of anatomy taking into account the constitution of the brain by two symmetrical hemispheres. Such a "lateralised" model of cortex anatomy was generated from the abstract one (i.e. non lateralised) through XSLT stylesheet transformation. Every concept in the abstract model is duplicated twice (one for each hemisphere). Each copy is tagged according to its side, renamed (prefixed with 'Right' or 'Left'), and linked to the original abstract concept by the 'is-lateral-instance-of' relationship. Relations among lateralised concepts of the same hemisphere are derived from the relations between abstract concepts.

The symbolic model provides a description of the most significant structures of the human cortex. Anatomical concepts are 'Hemisphere', 'Lobe', 'Gyrus', 'Pars', 'Operculum', 'Pli de passage' and 'Sulcus'. Sulci can be bound by 'is-segment-of', 'is-side-branch-of' or 'is-connected-to' relations. The other structures (cortical structures) are bound by 'is–anatomical–part-of' and 'is–anatomically–continuous-to' relations. Moreover, a 'delimits' relation involves a sulcus and two other cortical structures. Concepts and relations may have properties such as 'existence probability', as a simple representation of variability: for example the Intermediate PreCentral Sulcus is present in 18% of individuals, or the Superior Frontal Sulcus and the PreCentral Sulcus are connected in 92% of the right hemispheres [3].

The numeric knowledge is currently limited to 3D models representing the shapes of cortical sulci. The latter were generated from MRI images by representing the median surface of the sulcal folds.

The representation of links between symbolic concepts and numeric elements, particularly 3D models, is generic. It does not depend on the nature of the link nor on the nature of the elements linked. They can be used to associate any concept with one or more prototypical shape representations, or with some statistical model of their shape variations.



Picture 1 : HTML representation of the symbolic knowledge base

[3] Virtual Reality Modelling Language

2.2 Knowledge base browsing applications

These applications were implemented for demonstration and browsing purposes, rather than for a specific goal such as teaching or decision support. Both applications display symbolic concepts and 3D data and manage bidirectional links between them.

We implemented a standalone Java application first, and then a set of XSLT[4] stylesheets generating HTML pages from the XML files, enabling an user-friendly access to the knowledge base using any web browser. The knowledge can be consulted on the web[5], in both the XML and the HTML forms.

Picture 1 is a snapshot of the symbolic knowledge base. The left frame displays concepts sorted according to the 'is-anatomical-part-of' relation. The right frame displays information available about a concept, including hyperlinks to other related concepts.

3 Key points for reusability

This prototype shows that an anatomical model combining explicit and well-formalised symbolic knowledge with numeric knowledge can be achieved using web technologies. The use of web standards provides independence from application and operating system constraints.

Based on this model, we tried to assess its suitability to three main kinds of contexts: teaching, decision support, and sharing of neuroimaging data. We identified some specific needs from each of these domains, some of them highlighting limitations of our model. The following paragraphs detail the results of this assessment. We first lay the stress on taxonomic aspects. We then analyse the mereological needs and their relation with topology. Finally, we synthesize these needs and mention some methodological key points for the successful creation and maintenance of an adequate anatomical model.

3.1 Taxonomy

Generalization In a given application context, cortex anatomy may occur as a sub-domain of anatomy. For example, a teaching application would present the skull bones, the various brain envelopes as well as the cortex. Similarly, a clinical data management application would need to manage references to cortex anatomy just like any other anatomical references. Consequently, we think it is advisable to make our model compatible with more general models of anatomical structures such as Digital Anatomist or GALEN. Rather than developing a specific extension to either of them, we prefer to develop an autonomous yet compatible model. The next step will consist in establishing a semantic correspondence between our model and more general models. Our model's anatomical structures have to be restrictions of more general ones. This is possible for Digital Anatomist, of which classes may be superclasses of the ones we defined. For example, Sulcus would inherit from AnatomicalSpatialEntity. Other cortical structure such as Hemisphere, Lobe, or Gyrus would be subclasses of OrganSubdivision [12]. As for GALEN, this has yet to be studied. This relation to more general models may raise problems regarding compatibility of relations. The Digital Anatomist 'is-boundary-of' relation is a binary one [13]. Such a relation is incompatible with relations existing in our model and with existing inference capabilities. It is not because a particular sulcus is involved in two binary delimitation relationships with two gyri that the two gyri are contiguous. Actually, we use a ternary delimitation relation between a sulcus and two cortical structures to represent contiguity. This relation can be used to infer contiguity of compound structures such as lobes. For example, we can deduce from the delimitation of the PreCentral gyrus and the PostCentral gyrus by the CentralSulcus and from the fact that these two gyri are anatomical parts of the Frontal Lobe and of the Parietal Lobe, respectively, that the two lobes are delimited by the Central Sulcus.

Specialization A lateralised model of anatomy may be referred to in image data descriptions as well as in knowledge representations about asymmetries of the two brain hemispheres. This may be relevant in both teaching and decision support contexts. A generic concept such as PreCentral Sulcus may have different relations or different shapes or components (with different existence probabilities), depending on location within the left or the right hemisphere. This problem cannot be solved by simply adding an attribute to specify the side we are referring to. This has two implications. On the one hand, the lateralised model should specialise the abstract model. On the other hand, the lateralised model should be able to associate different knowledge to anatomical features of either sides. For example, this is relevant for representing the language areas (notably the inferior frontal area and the planum temporale), which are particularly important in the context of surgical planning. The first point reinforces generation of the lateralised model from the abstract model, in order to ensure consistency between them. The specialization of two lateralised concepts instances from an abstract concept may become an issue during integration with more general models. Neither the UMLS[6] nor NeuroNames [14] have an explicit representation of laterality. When organs such as lungs or parts of an

[4] eXtensible Stylesheet Language Transformation

[5] http://idm.univ-rennes1.fr/users/dameron/anatomy

[6] Unified Medical Language System

organ such as cardiac chambers are composed of two instances, the two instances are defined as two different organs or organ parts, with two different CUI[7]. However, the fact that they are instances of the same concept (lung or cardiac chamber) is not represented. Conversely, NeuroNames represents the generic concept, but not its lateralised instances. During the realization of our lateralised model, the reference to an UMLS identifier becomes indirect: one must go from the lateralised concept to the abstract one in order to access it. The second point, i.e. representation of knowledge about aymmetries, is not yet taken into account in our model. We consider overloading the abstract model's values with the lateralised concepts' specific properties and relations.

Instanciation Rosse defines the symbolic model as 'canonical knowledge', i.e. general knowledge about a domain, as opposed to 'instanciated knowledge' for the concretization of canonical concepts in a real individual [1]. The lateralised model provides the canonical knowledge to refer to when describing particular subjects' data.

3.2 Mereotopology

Visible vs Buried Cortex anatomy represented within the atlases or used in a neurosurgical context primarily refers to visible cortical structures such as gyri or sulci and to their relations. However, 2/3 of the cortical surface are buried into the sulcal folds. It follows that an external description of the cortex cannot properly describe tissue continuity within the cortex, nor be well-adapted for the localization of functional data. However anatomy as a scientific field evolves to take into account functional aspects beyond strict morphology. Research on the anatomo-functional organization of the cortex tries to overcome the limitations of the classical representations by adopting unfolded (2D) [15], inflated (3D) [15] or spherical projection [16] models of the cortex. The symbolic model we developed puts emphasis on visible structures and their relations. Applications such as sharing of neuroimaging data for research on brain functional mapping call upon a description of the buried cortical structures as well. Of course models of visible and buried structures have to be consistent.

Identity Criteria The formal definition of concepts and relations requires to define identity criteria. This is not trivial due to the inter-hemispheric and inter-individual variability. However it is a key requirement in all the considered application contexts. In order to deal with the variability issue, our model of anatomy should explicitly distinguish between what is necessary and what is possible. The former constitutes the identity criteria of a structure or a relation and the latter copes with the inconstant character of some structures.

The identity of a structure is partly defined by taxonomic relations: for example, the Ascending Ramus is a kind of Sulcus. Besides, among all the structures of the same kind, a structure can be defined according to its parts as well as to the structures it is part of. This shows that mereological relations are involved as well. Moreover, topological relations contribute to the identification of the parts of a structure. For example, it is because it is adjacent to the Superior PreCentral Sulcus that the Superior Pars of the PreCentral Gyrus can be distinguished from the Inferior Pars of the same Gyrus.

3.3 Additional design constraints

All the needs we identified have to be further investigated. Based on those requirements, we will extend the specifications of our model.

Identity Criteria Modal logics is a sound logical ground to represent what is and what can be. Our model should specify modal expressions on the existence of anatomical concepts and relations. These modal expressions should be applied to taxonomic, mereological and topological relations.

Model Representation In order to be usable by a program (and by a human) for intelligent data query or for decision support, logical operations such as inferences should be possible, requiring first order logics. It implies constraints of completeness (any assertion can be proved or disproved) and consistency (a false assertion cannot be proved true). Furthermore, some properties such as transitivity, involved in mereological relations have to be explicitly represented. Modelling of the symbolic knowledge using UML and textual definitions fails to address these constraints. We begun defining concepts and relation properties with description logics, which provide a more powerful framework to perform inferences involving canonical anatomical knowledge, instanciated anatomical data or both. Furthermore, this formalism should ease the study of semantic compatibility with more general models, since description logics are used by both Digital Anatomist (Protégé) and GALEN (Grail).

Model Implementation The implementation of the numeric and the symbolic models with VRML and XML was satisfactory. They could easily be used and processed locally by the applications. Furthermore, the possibilities offered by Xlinks and XSL are promising for a use in a web context. This trend is strengthened by X3D, a project aiming at implementing VRML using XML. This would go a step further in unifying numeric and symbolic information.

[7] Concept Unique Identifier

Maintenance Some relations can be defined by the composition of other relations. We have already presented the delimitation of the Frontal Lobe and the Parietal Lobe by the Central sulcus, which can be inferred from the delimitation of the PreCentral Gyrus and the PostCentral Gyrus by the same Sulcus. Equally, we state that if a segment of a sulcus delimits two cortical structures, then this sulcus also delimits these structures. Furthermore, transitive relations such as is-anatomical-part-of imply the representation of many relations. Since some relations are composed of other relations, dependency between them has to be explicitly managed.

Therefore, during updates of the model, such composed relations may jeopardize consistency. For example, adding an anatomical part to a structure may require generating many other is-anatomical-part-of relations. Similarly, deletion of an is-anatomical-part-of relation is even more tricky, since the entities involved in this relation may also be involved in other relations such as delimitation by sulcus.

Due to the dependences between relations, we only maintain a minimal model, of which consistency is easier to manage. In practice, the symbolic model is generated by applying a stylesheet that recreates all the derived relations. It requires a tool to compare two successive expanded symbolic models, (before and after an update), in order to make sure that all the expected relations (and only them) have been modified.

Of course, identity criteria have to be re-assessed after any update of the model.

4 Perspectives

The following of this work will consist in taking into account the previous requirements and constraints in an extended version of our model. It will try to achieve a coherent integration of taxonomic, mereological and topological aspects as well as compatibility with Digital Anatomist and GALEN.

The second step will be to assess its relevance in the various application contexts discussed above, particularly in web-distributed environments.

References

[1] Rosse C, Mejino JL, Modayur BR, Jakobovits R, Hinshaw KP, Brinkley JF. Motivation and Organizational Principles for Anatomical Knowledge Representation. JAMIA. Jan-Feb 1998 volume 5 number 1. 17-40.

[2] Brinkley JF, Wong BA, Hinshaw KP, Rosse C. Design of an Anatomy Information System. Medical Graphics. May-June 1999. 38-48.

[3] Ono M, Kubik S, Abernathey CD. Atlas of the cerebral sulci. Georg Thieme Verlag ed. Thieme Medical Publishers Inc, New-York. 1990.

[4] Talairach J, Tournoux P. In: Co-planar stereotactic atlas of the human brain. Georg Thieme Verlag ed. Thieme Medical Publishers Inc, New-York. 1988.

[5] Duvernoy HM. In: The Human Brain: Surface, Three-Dimensional Sectional Anatomy With MRI, and Blood Supply, 2nd ed. Springer-Verlag. 1999.

[6] Gibaud B, Garlatti, Barillot C, Faure E. Computerized brain atlases as decision support systems : a methodological approach. Artificial Intelligence in Medicine. 1998, 14. 83-100.

[7] Brinkley JF, Wong BA, Hinshaw KP, Rosse C. Design of an Anatomy Information System. Computer Graphics and Applications. 1999. Volume 19. 38-48.

[8] Rector AL, Nowlan WA, GALEN Consortium. The GALEN Project. Computer Methods and Programs in Biomedicine. 1993. 45 : 75-8.

[9] Nowinski WL, Fang A, Nguyen BT, Raphel JK, Jagannathan L, Raghavan R, Brian RN, Miller GA. Multiple Brain Atlas Database and Atlas-Based Neuroimaging system. Computer Aided Surgery. 1997, 2. 42-66.

[10] Höhne KH, Pflesser B, Pommert A, Riemer M, Schiemann T, Schubert R, Tiede U. A 'Virtual Body' Model for Surgical Education and Rehearsal. Computer. Jan 1996. 25-31.

[11] Kikinis R, Shenton ME, Iosifescu DV et al. A Digital Brain Atlas for Surgical Planning, Model-Driven Segmentation and Teaching. IEEE Transactions on Visualization and Computer Graphics. Sept 1996 volume 2 number 3. 232-41.

[12] Dameron O, Gibaud B, Morandi X. Numeric and Symbolic Knowledge Representation of Cortex Anatomy using Web Technologies. AIME 2001 proceedings LNAI 2101. 2001. 359-68.

[13] Neal PJ, Shapiro LG, Rosse C. The Digital Anatomist Structural Abstraction : a Scheme for the Spatial Description of Anatomical Entities. AMIA Proceedings. 1998. 423-7.

[14] Bowden DM, Martin RF. NeuroNames Brain Hierarchy. Neuroimage. 1995, 2. 68-83.

[15] Van Essen D, Drury H. Structural and Functional Analyses of Human Cerebral Cortex Using a Surface-Based Atlas. Journal of Neuroscience. Sept 1997. 17(18):7079-7102.

[16] Tao X, Han X, Rettmann ME, Prince JL, Davatzikos C. Statistical Study on Cortical Sulci of Human Brains. IPMI 2001 Proceedings. LNCS 2082. 2001. 475-87.