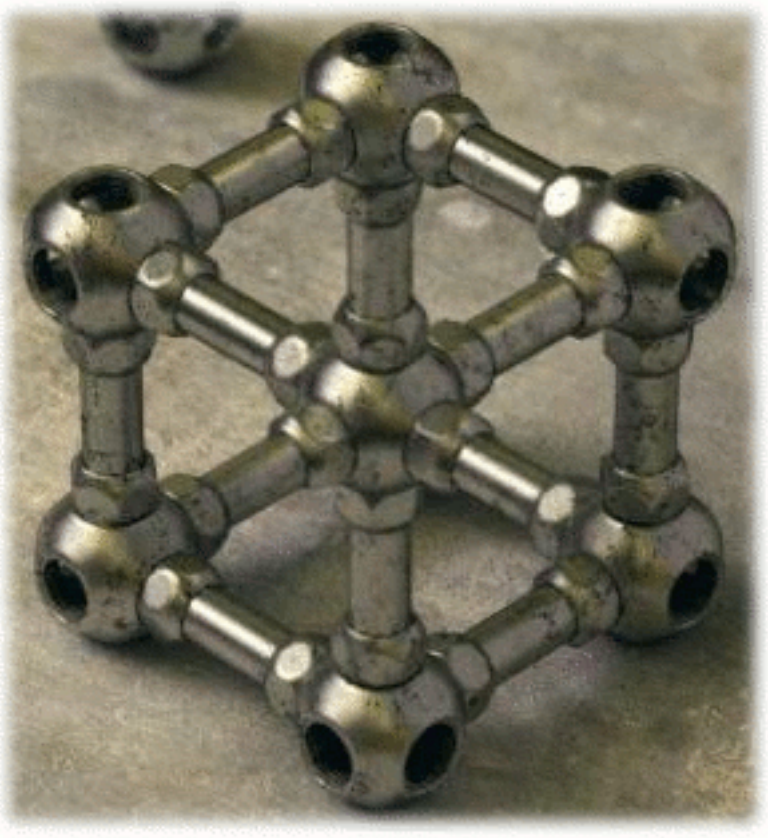

A region-based ontology of the brain ventricular system and its relation to schizophrenia [1]

Author(s):

Paulo Santos
Rodolpho Freire
Danilo N. dos Santos
Carlos Thomaz
Paulo Sallet
Mario Louzã
Anthony Cohn

[1]This work was supported by Fapesp Project LogProb, grant 2008/03995-5, São Paulo, Brazil.

A region-based ontology of the brain ventricular system and its relation to schizophrenia

Paulo Santos Rodolpho Freire Danilo N. dos Santos Carlos Thomaz
Centro Universitário da FEI, S. Paulo, Brazil

Paulo Sallet Mario Louzã
Instituto de Psiquiatria, Universidade de S. Paulo, S. Paulo, Brazil

Anthony Cohn
School of Computing, University of Leeds, Leeds,

Abstract

This paper describes an initial region-based formalisation of some concepts about neuroanatomy into ontological and epistemic terms, as part of a major effort into the formalisation of the knowledge contained in neuroimages of patients with schizophrenia. Our long-term goal is to build an ontology that is a formal basis for the expectations generated from statistical data analysis. To this end we present an example of applying this ontology to interpret the results of image-based analysis of neuroimages from schizophrenic patients.

Keywords: Foundational ontologies, neuroimaging, schizophrenia.

1. Introduction

Schizophrenia is a mental disorder that affects 1% of the population worldwide and whose causes are still unknown[1] . Both genetic (Harrison and Owen, 2003) and environmental factors (including biological - e.g., prenatal infection and obstetric complications - and psychosocial factors) appear to play a role in its etiology. These factors, however, are not sufficient for the emergence of schizophrenia (Nuechterlein and Dawson, 1984). Previous studies (Gattaz et al., 1988, Thomaz et al., 2007b) investigated the possibility that the neuroanatomical changes of this disease are diffuse, therefore suggesting that multidimensional techniques may be the right tools for providing information of the possible influence of structural brain anomalies related to the disease. In another work (Santos et al., 2009), we have proposed an integrated framework for extracting and describing patterns of changes from neuroimages of schizophrenic patients using a combination of linear discriminant analysis and active contour models. As a result, we obtained clusters of the most statistically significant differences of neuroimages between healthy controls and patients (an example of such clusters is shown in Figure 1).

[1] Source: National Institute of Mental Health (http://www.nimh.nih.gov/, last access 8/6/2009.

The voxels within these clusters were mapped using the Talairach atlas (Lancaster et al., 2000) to the corresponding labels of neuroanatomic structures. However, that work falls short on placing the obtained descriptions into the context of neuroanatomy, as the relationships between the labels and the structures they describe were not represented explicitly. Besides, the knowledge about the known facts of how anatomical changes are related to schizophrenia was also absent.

The goal of the present work is to describe a spatial ontology for the brain structure that is most affected by schizophrenia: the Ventricular Brain System (VBS). This ontology provides both the neuroanatomical context for parts of the VBS, and also the findings relating brain structures to schizophrenia as published in the literature. For this purpose, in this paper we assume the Basic Inclusion Theory (BIT) (Donnelly et al., 2006), a spatial theory for formalising biomedical ontologies. The main reason for assuming BIT is the distinction that it makes between mereological and location relations. Mereological relations are properties defined in terms of parts and their respective wholes (Casati and Varzi, 1999). For instance, in BIT we could express formally that the fourth ventricle is part of the ventricular system, but it would be a mistake to say that it is also part of the hindbrain, although it is located within it. This distinction is explicit in BIT (as summarised in Section 2). BIT is one of many existing frameworks for building bio-ontologies. Perhaps central to this area are the Foundational Model of Anatomy Ontology (FMA) (Rosse and Mejino, 2003) and the Open Biomedical Ontologies (Smith et al., 2005), among others (Chen et al., 2006). FMA is a knowledge source of classes and relations about observable characteristics of the human body structure; thus, FMA is mainly concerned with representing anatomical information. The OBO Foundry project is a collaborative development which includes a large amount of biological information, such as amphibian gross anatomy, chemical entities of biological interest, as well as human disease information and so on. There are a few attempts to build ontologies of neuroanatomical structures. The work proposed in (Martin et al., 2001) integrates into FMA a terminology whose purpose is to model the structure of knowledge of neuroanatomy. An ontology specifically about the nervous system is proposed in (Bota and Swanson, 2008). However, to the best of our knowledge, the present paper is the first attempt to encode neuroanatomical terms into BIT, providing thus a clear account of the mereological and location relations underlying this field. The main reason for choosing the Basic Inclusion Theory in this work, instead of FMA or OBO (for instance), is that BIT provides a clear formalisation for the structure underlying the elements of bio-ontologies in terms of spatial regions. We believe that the region-based relations of BIT provide a proper basis for describing the main differences between neuroimages of controls and patients, such as those shown in Figure 1.

There are a number of decision support systems for schizophrenia reported in the literature (Spitzer and Endicott, 1968, Bronzino et al., 1989, First et al., 1993), all of them were based

on formalising the symptomatology of the disease (Razzouk, 2001). The challenge of the present work is to develop a knowledge representation and reasoning system that incorporates (as elementary entities) spatial regions in neuroimages that represent structures that are known to be affected in mental disorders. This paper is part of a major effort to formalise the basic concepts for an automated process of image interpretation that would facilitate the medical assessment of the information contained in neuroimages. The pursuit of an ontological basis for this investigation resides in the largely believed fact that the pathogenesis of schizophrenia may rely on a myriad of factors, ranging from neuroanatomical and neurochemical abnormalities to genetic predisposition (Kandel et al., 2000). Defining a common ontology underlying the search of the possible causes of this disease is essential in order to develop computer systems that process findings from distinct analysis of the same data sets. The present paper reports initial steps towards these goals.

<<Fig1.tiff>>

Figure 1: Results of the active contour model segmentation from statistical

In order to provide a region-based formalisation for the ventricular brain system we define consensual (fiat) boundaries using BIT relations. We also present axioms constraining the idea of spatial continuity. In this work, variables are universally quantified, unless explicitly stated otherwise.

This paper is organised as follows: the next subsection describes informally the ventricular brain system; Section 2 summarises Basic Inclusion Theory; Section 3 presents our definitions of fiat boundaries and continuity used to represent the VBS; Section 4 presents the bio-ontology for the VBS; Section 5 presents an example of how the proposed ontology can be used along with neuroimage analysis. Finally, Section 6 concludes this paper.

1.1. Ventricular Brain System (VBS)

The ventricular brain system (Figure 2) is a cavity disposed within the brain, which is composed of the third, fourth and lateral ventricles. The lateral ventricles are subdivided as body, frontal horn, occipital horn and temporal horn. The communication between the lateral ventricles and the third ventricle is done via the Monro foramina (Parent and Carpenter, 1996).

The third ventricle is sub-divided into optical recess, supra-parietal recess and infundibulum. The third ventricle also communicates with the fourth ventricle by the cerebral aqueduct and in the centre of the third ventricle is located the interthalamic connection.

The fourth ventricle is composed of the lateral recess and the Luschka foramina linking up with the third ventricle through the brain aqueduct.

<<Fig2.Tiff>>

Figure 2: Ventricular Brain System. A lithography plate from Gray's Anatomy (Gray, 1918)

As we shall discuss further in this paper, enlarged lateral and third ventricles are the most commonly found anatomical abnormalities in schizophrenic patients.

2. The Basic Inclusion Theory

In this section we describe the Basic Inclusion Theory (BIT) (Donnelly et al., 2006), a region-based formalism for building biomedical ontologies, which makes a clear distinction between mereological and location relations.

In the mereological part of BIT, variables range over domain individuals, in the present case, over distinct neuroanatomical structures.

Axioms (1), (2) and (3) constrain the meaning of the primitive BIT relation part of (P/2).

$$P(x, x). \tag{1}$$

$$P(x, y) \wedge P(y, x) \rightarrow x = y. \tag{2}$$

$$P(x, y) \wedge P(y, z) \rightarrow P(x, z). \tag{3}$$

Axiom (1) states that every individual x is a part of itself; Axiom (2) states that if x is part of y (and y is a part of x) then x and y are the same individual. Finally, Axiom (3) represents the fact that if x is part of y and y is part of z, then x is part of z. With P/2, two important relations can be defined: the proper part relation PP (Formula (4)) and the overlap relation O (Formula (5)).

$$PP(x, y) \equiv P(x, y) \wedge \neg x = y \tag{4}$$

$$O(x, y) \equiv \exists z\,(P(z, x) \wedge P(z, y)) \tag{5}$$

In other words, x is proper part of y if and only if x is part of y and x and y are distinct individuals (cf. Formula (5)); similarly, x overlaps with y if there is at least one z that is part of both x and y.

The distinction between mereological and location relations is accomplished by defining location relations with spatial regions as arguments. In this paper, spatial regions are unique mappings from neuroanatomical structures to the distinct spatial regions they occupy. Thus, location relations can be defined using relations from the mereological part of BIT using a region function r(x) that maps each individual x to its occupancy region.

There are two basic location relations in BIT: the LocIn, standing for "object x is located in object y" which holds if and only if the occupancy region of x, r(x), is part of the occupancy region of y" (cf. Formula (6)) and PCoin, read as "objects x and y partially coincide", which holds if and only if their occupancy regions overlap (cf. Formula (7)).

$$\text{LocIn}(x, y) \equiv P(r(x), r(y)). \quad (6)$$

$$\text{PCoin}(x, y) \equiv O(r(x), r(y)). \quad (7)$$

The qualitative distinctions within the basic part of BIT (Axioms (1) – (7)) are shown in Fig. 3 (Donnelly et al., 2006).

<<Fig3.tiff>>

Figure 3: Basic Inclusion Theory relations.

The BIT axioms imply a number of theorems that in turn can be used to infer many facts about biomedical ontologies (Donnelly et al., 2006). Assertions about classes of individuals, rather than about particular instances, are also feasible in an extension of BIT that includes an instantiation relation, linking a class to its individual instances. Therefore, in this extended version of BIT, it is possible to express (for instance) the spatial relation between the classes Ventricular System and Cerebrospinal Fluid, as well as the spatial relation between the individuals' right and left lateral ventricles. The instantiation relation is represented as Inst(x, A) read as "x is an instance of the class A".

This class instantiation relation is defined in (Donnelly et al., 2006) as the relation Is_a(A, B) (Formula 8) which represents that the class A is a subclass of B and that if x is an instance of A, it is also an instance of B.

$$\text{Is\_a}(A, B) \equiv \forall x(\text{Inst}(x, A) \rightarrow \text{Inst}(x, B)) \quad (8)$$

Also relevant to this work is the relation PP12 (A, B) which represents that "every instance of A has an instance of B as proper part, and every instance of B has an instance of A as proper part" (Donnelly et al., 2006) . With this relation it is possible to state, for instance, that every lateral ventricle is a proper part of the ventricular brain system and every ventricular brain system has a lateral ventricle as proper part.

3. Continuity and Fiat Boundaries

In order to represent the ventricular brain system, we have to characterise regions that do not have physical boundaries, i.e., they are consensually defined by the experts (such as the body, the anterior or posterior horns of the lateral ventricles, Figure 2). Besides, we also want to be able to represent continuous portions of the ventricular system (the foramina for

instance). In order to accomplish these tasks we define the notion of fiat boundaries and maximal continuous part, as introduced below.

Continuity

In order to define the notion of maximum continuous part, we introduce a relation about discontinuity (Disc/2), following some ideas in the "hole ontology" presented in (Varzi, 1996). A discontinuity can be thought of as a special case of a hole: a hole is normally thought of something which is part of the convex hull of an object but not part of the object itself, normally manifesting itself as an interior void or a surface concavity. We can think of discontinuities as a third case where the discontinuity is formed by two (or more) separated parts, and the discontinuity again is part of the convex hull of the sum of these parts but not overlapping any of them.

This notion is used to represent, for instance, the material discontinuity in the third ventricle called the interthalamic connection.

We represent a discontinuity using the relation Disc(x, y) ("x is a discontinuity in y") which is constrained by the following axioms:

$$\exists y \, Disc(x, y) \tag{9}$$

"Every discontinuity x has a host y" ;

$$Disc(x, y) \rightarrow LocIn(x, y) \tag{10}$$

"Every discontinuity x is located in its host y";

$$Disc(x, y) \rightarrow \neg \exists z Disc(z, x) \tag{11}$$

"no discontinuity hosts a discontinuity itself".

Using Disc/2 we can define a relation for continuous part: PCont(x, y), meaning that "x is a continuous part of y", (Formula (12)). This relation was first introduced in (Santos and Cabalar, 2008).

$$PCont(x, y) \equiv P(x, y) \wedge \forall z \neg Disc(z, x). \tag{12}$$

Then, we define a segment x of an object y (Segm(x, y)) as the maximal continuous part of y, according to Formula (13).

$$Segm(x, y) \equiv PCont(x, y) \wedge \neg \exists z (PP(y, z) \wedge PP(x, z) \wedge PCont(z, y)). \tag{13}$$

Fiat boundaries

Fiat boundaries are limits that are commonly accepted but which do not have a concrete existence (Varzi, 1997, Smith and Varzi, 2000). In this paper we define the relation Fiat(x, y), read as "x is a fiat boundary in y", constrained by axioms (14) to (17), below. It is worth noting that, in order to provide a region-based definition for boundaries, we assume in this work that boundaries are of the same dimensionality as the domain objects.

$$\exists\, y\ \mathrm{Fiat}(x, y), \tag{14}$$

"Every fiat boundary x has a host y".

$$\mathrm{Fiat}(x, y) \rightarrow \mathrm{PCoin}(x, y), \tag{15}$$

"Every fiat boundary x partially coincides with its host y".

$$\mathrm{Fiat}(x, y) \rightarrow \neg\, \exists\, z\ \mathrm{Fiat}(z, x), \tag{16}$$

"a fiat boundary cannot host a fiat boundary itself".

$$\mathrm{Fiat}(x, y) \rightarrow \neg\, \exists\, z\ \mathrm{Disc}(x, z). \tag{17}$$

"no fiat boundary is a discontinuity".

Axiom 17 guarantees that a fiat boundary is not a physical discontinuity. The next section presents how these notions are used to formalise the ventricular brain system.

4. A spatial bio-ontology for the ventricular brain system

In the top level of the ontology for the ventricular brain system resides statements about its constituent parts. We assume three classes: VBS (ventricular brain system), VENTRICLE and FORAMINA. The statements are expressed in Formulae 18 to 27, where the following abbreviations are used: LV, lateral ventricle; LLV, left lateral ventricle; RLV, right lateral ventricle; TV, third ventricle; FV, fourth ventricle; LIF, RIF: left and right interventricular foramina (Monro foramina); and CA, Cerebral aqueduct. For the sake of brevity, we omit in the description below more specific parts of the ventricular brain system, such as the anterior or the posterior horns of the lateral ventricles. In the formulae below we use the relation PP12 (A, B) (every instance of A has an instance of B as proper part, and every instance of B has an instance of A as proper part" (Donnelly et al., 2006)) as described in Section 2.

$$\mathrm{PP12}(\mathrm{VENTRICLE}, \mathrm{VBS}) \tag{18}$$

$$\mathrm{Is\_a}(\mathrm{LV}, \mathrm{VENTRICLE}) \tag{19}$$

PP12(LLV, LV) (20)

PP12 (RLV, LV) (21)

Is_a(TV, VENTRICLE) (22)

Is_a(V, VENTRICLE) (23)

PP12(FORAMINA, VBS) (24)

Is_a(LIF, FORAMINA) (25)

Is_a(RIF, FORAMINA) (26)

Is_a(CA, FORAMINA) (27)

We enumerated 21 fiat boundaries limiting all ventricular anatomical elements. We define each fiat boundary according to the structure that hosts it.

Formula 28 states this fact for each fiat boundary Zi .

Fiat(Zi , Ai) (28)

"Zi is a fiat boundary on a structure Ai ."

In order to distinguish individual ventricles (or any neuroanatomical region) within images, the representation and autonomous recognition of shapes are key issues. However, the efficient processing of shapes is still a difficult problem (Costa and Cesar-Jr., 2006, Cohn and Hazarika, 2001, Cohn and Renz, 2008).

In the present paper, we single out the ventricles using their relative volumes. In order to do that, we use a function Vol : VENTRICLE $\rightarrow \Re$, mapping an instance of VENTRICLE to a real value. Formulae (29)–(31) formalise this idea.

Formula 29 states that x is an instance of the lateral ventricle (LV) if its volume (ϕ) is greater than both the volume of the third ventricle (TV) and the volume of the Fourth ventricle (F V). The fourth conjunct on the right hand side of Formula 29 states that x is a maximal continuous part of the ventricular brain system (cf. Section 3). Formulae 30 and 31 are analogous.

$$\text{Inst}(x, \text{LV}) \leftarrow \text{LV}(x) \wedge \text{TV}(y) \wedge \text{FV}(z) \wedge (\text{Vol}(x) > \text{Vol}(y)) \wedge (\text{Vol}(x) > \text{Vol}(z)) \wedge \text{Segment}(x, \text{VBS}) \quad (29)$$

$$\text{Inst}(x, T\ V) \leftarrow \text{LV}(x) \wedge \text{TV}(y) \wedge \text{FV}(z) \wedge (\text{Vol}(x) > \text{Vol}(y)) \wedge (\text{Vol}(y) > \text{Vol}(z)) \wedge$$

$$\wedge \text{Segment}(x, \text{VBS}) \quad (30)$$

$$\text{Inst}(x, F\ V) \leftarrow \text{LV}(x) \wedge \text{TV}(y) \wedge \text{FV}(z) \wedge (\text{Vol}(x) > \text{Vol}(z)) \wedge (\text{Vol}(y) > \text{Vol}(z)) \wedge$$

$$\wedge \text{Segment}(x, \text{VBS}) \quad (31)$$

We can write analogous formulae for the remainder structures of the ventricular brain system, omitted here for brevity.

The statements presented in this section can be summarised in the diagrams 4 and 5.

<<Fig4.tiff>>

Figure 4: Classes of the VBS formalisation

<<Fig4.tiff>>

Figure 5: Partonomy of the VBS

In order to define common characteristics among distinct groups, the medical specialist relies on the relative literature (using information from meta-analysis), image or statistical analysis. The information available in these sources is not part of the domain *per se* (so it cannot be captured by an ontology) but it is knowledge about it. The next section defines epistemic classes that are used to include knowledge about the domain in the ontology discussed above.

Epistemic Classes

The knowledge about things are not the things themselves, therefore, including it in the ontology would lead to a Kantian confusion (Smith and Ceusters, 2007). In this work we avoid this confusion by assuming epistemic classes, which are related to the ontological classes by a modified Is_a relation (Is_a_{EPIST}). Given an epistemic class E, an ontological class O and a binary primitive relation κ(x, y) (representing that x is the knowledge about a domain y), we define Is_a_{EPIST} in BIT in the following way:

$$\text{Is\_a}_{EPIST}(E, O) \equiv \forall x\ \text{Inst}(x, E) \rightarrow \neg\text{Inst}(x, O) \wedge \kappa(x, O). \quad (32)$$

Informally, E is an epistemic class within the ontology O if and only if every instance of E is not an instance of O but is knowledge about it. Therefore, we can include both ontological and epistemic individuals in the same formalism. In this work, an epistemic individual is a piece of knowledge about anatomical changes in the VBS (related to schizophrenia) that comes from the medical literature (meta-analysis for instance) or from image data analysis procedures.

In order to be absolutely precise, it would be desirable to provide a complete axiomatisation of κ(x, y) according to epistemic logics (Rescher, 2005, Baader et al., 2003). However, for brevity, we introduce it here as a primitive relation, leaving its formal treatment for a future work.

The epistemic classes are described by Formulae 33 to 40, whereby the basic epistemic classes are CONTROL and PATIENT.

$$Is\_a_{EPIST}\,(CONTROL, RLV) \quad (33)$$

$$Is\_a_{EPIST}\,(PATIENT, RLV) \quad (34)$$

$$Is\_a_{EPIST}\,(CONTROL, LLV) \quad (35)$$

$$Is\_a_{EPIST}\,(PATIENT, LLV) \quad (36)$$

$$Is\_a_{EPIST}\,(CONTROL, TV) \quad (37)$$

$$Is\_a_{EPIST}\,(PATIENT, TV) \quad (38)$$

$$Is\_a_{EPIST}\,(CONTROL, FV) \quad (39)$$

$$Is\_a_{EPIST}\,(PATIENT, FV) \quad (40)$$

Subclasses of these epistemic classes would be particular pieces of findings reported in the literature. Thus, we include formulae representing these findings according to the following scheme:

$$Is\_a(< author >, < espist.class >) \wedge Is\_a_{EPIST}\,(< espist.class >, < anatomical.struct. >) \leftrightarrow$$

$$Vol(< anatomical.struct >) \in (\Theta + \sigma, \Theta - \sigma) \quad (41)$$

Formula 41 states that an author (< author >) assigned an epistemic class < espist.class > to a particular anatomical structure < anatomical.struct. > iff this structure has a volume belonging to the interval (Θ + σ, Θ − σ) (where Θ is the mean value found by < author > and σ is the standard variation).

For instance, Formula 42 states that Barr et al. (Barr et al., 1997) reports a mean volume of the control right lateral ventricle as 6.52 mm. Table 1 shows the findings cited from various reference sources.

$$\text{Is\_a(Barr, Control)} \wedge \text{Is\_a}_{EPIST}\text{ (Control, RLV)} \leftrightarrow \text{Vol(RLV)} = 6.52\ (\pm 2.69) \qquad (42)$$

We encoded the classes and the partonomy (Figures 4 and 5) of the ontology described above in *Protégé* (Knublauch et al.2004). *Protégé* is an open source tool that combines a number of knowledge-modelling structures for the creation and usage of ontologies. It is worth pointing out that not all of the BIT definitions were possible to be implemented in this tool, as its base language (OWL) is not expressive enough to to encode the full first-order logic language with which BIT is written, but the BIT ontology was used to guide the *Protégé* modelling.

It is now possible to execute queries about, for instance, the composition of the ventricular brain system, or about specialist knowledge about the domain. An example of an ontological query is: “Which structures compose the ventricular brain system?” .

Epistemic reasoning is possible in a similar way: the query “is the volume 6.52 of the right or left lateral ventricles classified as patient or control?”. In Protégé this query becomes “Lateral Ventricles and Vol value 6.52”, and produces the result: “Vol Right LV Control Barr”, which means that the classification of an individual whose lateral ventricle (LV) has a volume of 6.52 is “control” according to Barr (Steen et al., 2006).

The next section presents ongoing work on the application of the framework described above on the assessment of findings on neuroimages.

Epistemic classes as cited in Steen et al. (Steen et al., 2006)				
Epistemic subclass	**Epist. class**	**Onto. class**	**Θ (mm)**	**σ (mm)**
Barr et al. (Barr et al., 1997)	Control	RLV	6.52	2.69
Chua et al. (Chua et al., 2003)	Control	RLV	4.40	1.70
Degreef et al. (Degreef et al., 1992)	Control	RLV	6.75	6.75
Fannon et al. (Fannon et al., 2000b)	Control	RLV	4.70	1.90
Fannon et al. (Fannon et al., 2000a)	Control	RLV	4.70	1.90
James et al. (James et al., 1999)	Control	RLV	5.89	2.28
Lawrie et al. (Lawrie et al., 1999)	Control	RLV	3.50	2.00
Barr et al. (Barr et al., 1997)	Patient	RLV	8.22	4.22
Chua et al. (Chua et al., 2003)	Patient	RLV	5.10	1.70
Degreef et al. (Degreef et al., 1992)	Patient	RLV	8.15	3.05
Fannon et al. (Fannon et al., 2000b)	Patient	RLV	4.60	2.10
Fannon et al. (Fannon et al., 2000a)	Patient	RLV	5.80	2.90
James et al. (James et al., 1999)	Patient	RLV	8.12	3.32
Lawrie et al. (Lawrie et al., 1999)	Patient	RLV	3.70	2.60
Barr et al. (Barr et al., 1997)	Control	LLV	6.98	2.90
Chua et al. (Chua et al., 2003)	Control	LLV	5.00	2.00
Degreef et al. (Degreef et al., 1992)	Control	LLV	6.70	2.14
Fannon et al. (Fannon et al., 2000b)	Control	LLV	4.50	1.90
Fannon et al. (Fannon et al., 2000a)	Control	LLV	4.50	1.90
James et al. (James et al., 1999)	Control	LLV	6.16	2.30
Lawrie et al. (Lawrie et al., 1999)	Control	LLV	3.80	2.80
Barr et al. (Barr et al., 1997)	Patient	LLV	9.16	4.22
Chua et al. (Chua et al., 2003)	Patient	LLV	6.00	1.50
Degreef et al. (Degreef et al., 1992)	Patient	LLV	8.91	3.96
Fannon et al. (Fannon et al., 2000b)	Patient	LLV	5.10	3.00
Fannon et al. (Fannon et al., 2000a	Patient	LLV	6.20	3.20
James et al. (James et al., 1999)	Patient	LLV	9.66	4.00
Lawrie et al. (Lawrie et al., 1999)	Patient	LLV	3.90	2.50
Degreef et al. (Degreef et al., 1992)	Control	TV	1.62	0.32
Fannon et al. (Fannon et al., 2000b)	Control	TV	0.68	0.21
Fannon et al. (Fannon et al., 2000a)	Control	TV	0.70	0.35
James et al. (James et al., 1999)	Control	TV	1.60	0.35
Lawrie et al. (Lawrie et al., 1999)	Control	TV	0.40	0.20
Degreef et al. (Degreef et al., 1992)	Patient	TV	1.33	0.38
Fannon et al. (Fannon et al., 2000b)	Patient	TV	0.77	0.20
Fannon et al. (Fannon et al., 2000a)	Patient	TV	0.90	0.40
James et al. (James et al., 1999)	Patient	TV	2.08	0.66
Lawrie et al. (Lawrie et al., 1999)	Patient	TV	0.60	0.40

Table 1: Epistemic classes as stated in Formula 41

5. An example of application: ontology and neuroimage analysis

In a previous work (Santos et al., 2009) we proposed an integrated framework for extracting and describing patterns of disorders from medical images using a combination of linear discriminant analysis and active contour models. We first used a multivariate statistical methodology to identify the most discriminating hyperplane separating two groups of images (healthy controls and patients with schizophrenia) contained in the input data. Second, we made explicit the differences found by the multivariate statistical method by subtracting the discriminant models of controls and patients, weighted by the pooled variance between the two groups. The proposed method was evaluated contrasting its results with well-known anatomical changes related to schizophrenia as described in the literature.

To illustrate the performance of this approach, we present below (in Figures 6, 7 and 8) some results on a magnetic resonance imaging (MRI) dataset that contains 44 patients with schizophrenia and 26 healthy controls. All these images were acquired using a 1.5T Philips Gyroscan S15-ACS MRI scanner (Philips Medical Systems, Eindhoven, The Netherlands), including a series of contiguous 1.2mm thick coronal images across the entire brain, using a T1-weighted fast field echo sequence (TE = 9ms, TR = 30ms, flip angle 30^{o}, field of view = 240mm, 256 x 256 matrix).

All images were reviewed by a MR neuro-radiologist. Ethical permission for this study was granted by the Ethics Committee of the Hospital das Clínicas, University of São Paulo Medical School.

<<Fig6.tiff>>

Figure 6: The most discriminant hyperplane found by the multivariate statistical approach. In the scatter plot, schizophrenia patients are coded with a cross whereas healthy controls are coded with a circle.

The statistical differences between the control and schizophrenia MRI samples captured by the discriminant analysis are illustrated in Figure 7. It shows the differences between the patient (on the top left) and control (on the right bottom) images captured by the multivariate statistical classifier using MR intensity features as inputs and all the spatially normalised samples for training.

These 14 images (from top left to bottom right) correspond to the 14 asterisks (on the horizontal axis, from left to right) shown on Figure 6 projected back into the image domain. We can interpret this mapping procedure as a way of defining intensity changes that come from “definitely schizophrenia” and “definitely control” samples captured by the knowledge extraction approach (Thomaz et al., 2007a).

The discriminant analysis is able to distinguish regions that have changed between images, however it is unable to measure the variation with respect to each voxel. In order to analyse the variation in a single point we use a voxel- to-voxel statistical test to obtain the effect size of the changes. Formula 43 represents the statistical test used, where x_c is the extreme control image model (bottom right image in Figure 7), x_p is the extreme patient image model (top left image in Figure 7), σ is the pooled variance between both groups, N_c is the number of control samples, and N_p is the number of patient samples.

<<Fig7.tiff>>

Figure 7: Statistical differences between the schizophrenia (on top left) and healthy controls (on bottom right) images captured by the discriminant analysis approach.

$$\text{Changes} = (x_c - x_p) \,/\, (\sigma^2/N_c + \sigma^2/N_p)^{1/2} \qquad (43)$$

A sample result of this procedure is shown in Figure 8 (and Figure 1), where one slice of the 3D MRI multivariate discriminant differences extracted is presented, superimposed on a control brain image randomly selected. In the picture, the gray-scale shows relative intensity change as a range of this thresholding. The dots in the images (Figures 1 and 8) show voxels whose difference of intensity between control and patient exceeded three pooled variances (i.e., differences that we consider as statistically significant). By normalising these resulting images with the Talairach atlas (Lancaster et al., 2000), it is possible to point out specific neuroanatomic structures where the changes occur. However, the findings on Figure 8 only show that there are differences between the two groups in the dataset, leaving aside any information about how the structures are affected (whether an enlargement or a shrinkage occurred).

The ontology delineates in Section 4 above is a step towards finding a solution to this issue, whereby the structures found by the image analysis could be compared to the body of knowledge relating schizophrenia (for instance) to neuroanatomical changes. This method provides, in an autonomous way, interpretations to the resulting images that are grounded on the literature available.

We currently can accomplish this for the ventricular brain system. More specifically, given a set of voxels representing statistically significant changes localised in the lateral ventricles

(for instance), the ontology can provide a description of the parts of the ventricles affected, how they are affected and a list of literature findings related to this portion of the brain.

Further research is needed to develop a complete ontology about the neuroanatomical structure changes related to schizophrenia. It is also important to note that some of these neuroanatomical changes may be present in other mental disorders, such as bipolar disorders.

<<Fig8.tiff>>

Figure 8: Effect size of the multivariate statistical differences comparing the intensity values described by the control and patient image models.

6. Discussion and open issues

The present work delineated some aspects in the development of a bio-ontology about a part of human neuroanatomy using the Basic Inclusion Theory (BIT) (Donnelly et al., 2006). We focused on the formalisation of the ventricular brain system (VBS), since it is regarded as the main anatomical structure related to schizophrenia. In order to accomplish this, two new relations had to be defined to represent, respectively, fiat boundaries and maximally continuous parts of the VBS. We left for future work the definition of fiat objects from fiat boundaries (Bittner and Winter, 1999).

Anatomical changes related to schizophrenia, as described in the literature, were also taken into account in the present paper. These literature findings were linked to the ontology by the assumption of epistemic classes, so that we could make a clear distinction between domain objects and knowledge about domain objects. Epistemic classes were characterised by the use of a primitive relation representing, in its arguments, that an epistemic class was knowledge about an ontological class. However, a rigorous treatment of this relation in terms of epistemic logics (Rescher, 2005, Baader et al., 2003) is still an open issue of this work.

Several pieces of evidence suggest that what we know as schizophrenia is in fact an ensemble of clinically discrete syndromes in which biological and environmental factors exert partial pathophysiological effecs (Joyce and Roiser, 2007; Fatemi and Folson, 2009). One of the strategies currently employed in disentangling such mechanisms is the search for endophenotipe, subtypes of psychoses found based on association with quantitative and stable clinical (e.g. neurocognitive performance) and neurobiologic markers (e.g genetic polymorphisms, neurochemical biomarkers and discrete patterns of neuroanatomic abnormalities obtained by neuroimaging) (Tan et al., 2008). The aim is to identify more homogeneous subgroups of psychoses in order to find their underlying etiopathogenic mechanisms. Searching for different structural brain abnormalities (as investigated in the present paper) is of utmost importance in advancing such models of research (Lawrie et al., 2008).

Subject to our current investigations are the complete description of neuroanatomy using BIT and the further interpretation of findings about schizophrenia from neuroimages, such as the change in size of the hippocampus, amongst others (Harrison, 1999). The interpretation of results obtained from neuroimaging statistical analysis, using the ideas discussed in this paper, was shown as an example in this paper (Section 5).

Key to the future development of this investigation is the development of theories capable of representing the shapes of brain regions and what changes they may suffer under psychiatric diseases.

Acknowledgements

This work was partially supported by Fapesp 2008/03995-5 (LogProb)

References

(Baader et al., 2003) Baader, F., Küsters, R., and Wolter, F. (2003). Extensions to description logics. In *The description logic handbook: theory, implementation, and applications*, pages 219–261. Cambridge University Press, New York, NY, USA.

(Barr et al., 1997) Barr, W., Ashtari, M., Bilder, R., Degreef, G., and Lieberman, J. (1997). Brain morphometric comparison of first-episode schizophrenia and temporal lobe epilepsy. *British Journal of Psychiatry*, 170:515–519.

(Bittner and Winter, 1999) Bittner, T. and Winter, S. (1999). On ontology in image analysis. In ISD '99: *Selected Papers from the International Workshop on Integrated Spatial Databases, Digital Inages and GIS*, pages 168–191, London, UK. Springer-Verlag.

(Bota and Swanson, 2008) Bota, M. and Swanson, L. (2008). Bams neuroanatomical ontology: design and implementation. *Frontiers in neuroinformatics*, 2(2).

(Bronzino et al., 1989) Bronzino, J., Morelli, R. A., and Goethe, J. (1989). Overseer: a prototype expert system for monitoring drug treatment in the psychiatric clinic. *IEEE Transactions on Biomedical Engineering*, 36(5):533–540.

(Casati and Varzi, 1999) Casati, R. and Varzi, A. C. (1999). *Parts and Places: the structures of spatial representation.* MIT Press, Cambridge, Mass.

(Chen et al., 2006) Chen, H., Fuller, S. S., Friedman, C., and Hersh, W., editors (2006). *Medical Informatics, volume 8 of Integrated Series in Information Systems*, chapter Biomedical Ontologies, pages 211–236. Springer US.

(Chua et al., 2003) Chua, S., Lam, I., Tai, K., Cheung, C., Tang, W. N., Chen, E., Lee, P., Chan, F. L., Lieh-Mak, F., and McKenna, P. (2003). Brain morphological abnormality in schizophrenia is independent of country of origin. *Acta Psychiatrica Scandinavica*, 108:269–275.

(Cohn and Hazarika, 2001) Cohn, A. G. and Hazarika, S. M. (2001). Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae*, 46(1-2):1–29.

(Cohn and Renz, 2008) Cohn, A. G. and Renz, J. (2008). Qualitative spatial representation and reasoning. In F. van Hermelen, V. Lifschitz, B. P., editor, *Handbook of Knowledge Representation*, pages 551–596. Elsevier, Amsterdam, NL.

(Costa and Cesar-Jr., 2006) Costa, L. and Cesar-Jr., R. M. (2006). *Shape analysis and classification: theory and practice*. CRC Press, Boca Raton, FL.

(Degreef et al., 1992) Degreef, G., Ashtari, M., Bogerts, B., Bilder, R. M., Jody, D. N., Alvir, J. M. J., and Lieberman, J. A. (1992). Volumes of ventricular system subdivisions measured from magnetic resonance images in first-episode schizophrenic patients. *Archives of general psychiatry*, 49(7):531–537.

(Donnelly et al., 2006) Donnelly, M., Bittner, T., and Rosse, C. (2006). A formal theory for spatial representation and reasoning in biomedical ontologies. *Artificial intelligence in medicine*, 36(1):1–27.

(Fannon et al., 2000a) Fannon, D., Chitnis, X., Doku, V., Tennakoon, L., O'Ceallaigh, S., Soni, W., Sumich, A., Lowe, J., Santamaria, M., and Sharma, T. (2000a). Features of structural brain abnormality detected in first-episode psychosis. *American Journal of Psychiatry*, 157:1829–1834.

(Fannon et al., 2000b) Fannon, D., Tennakoon, L., O'Ceallaigh, S., Doku, V., Soni, W., Chitnis, X., Lowe, J., Sumich, A., and Sharma, T. (2000b). Third ventricle enlargement and developmental delay in first-episode psychosis: preliminary findings. *British Journal of Psychiatry*, 177:354–359.

(Fatemi and Folson, 2009) Fatemi, S.H. and Folsom, T.D. The neurodevelopmental hypothesis of schizophrenia revisited. *Schizophrenia Bulletin*, 35(3): 528–548, 2009.

(First et al., 1993) First, M., Opler, L., Hamilton, R., Linder, J., Linfield, L., Silver, J., Toshav, N., Kahn, D., Williams, J., and Spitzer, R. (1993). Evaluation in a inpatient setting of DTREE, a computer-assisted diagnostic assessment procedure. *Comprehensive psychiatry*, 34(3):171–75.

(Gattaz et al., 1988) Gattaz, W. F., Rost, W., Kohlmeyer, K., Bauer, K., Hubner, C., and Gasser, T. (1988). CT scans and neuroleptic response in schizophrenia: a multidimensional approach. *Psychiatry Research*, 26(3):293–303.

(Gray, 1918) Gray, H. (1918). *Anatomy of the Human Body*. Philadelphia: Lea & Febiger. re-edited by Bartleby.com in 2000, available from www.bartleby.com/107.

(Harrison, 1999) Harrison, P. J. (1999). The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain*, 122(4):593–624.

(Harrison and Owen, 2003) Harrison, P. J. and Owen, M. J. (2003). Genes for schizophrenia: recent findings and their pathophysiological implications. *Lancet*, 361:417–19.

(James et al., 1999) James, A., Crow, T., Renowden, S., Wardell, A., Smith, D., and Anslow, P. (1999). Is the course of brain development in schizophrenia delayed? Evidence from onsets in adolescence. *Schizophrenia Research*, 40:1–10.

(Joyce and Roicer, 2007) Joyce EM, Roiser JP. Cognitive heterogeneity in schizophrenia. *Current Opinion in Psychiatry*, 20(3): 268–272, 2007.

(Kandel et al., 2000) Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of Neural Science.* McGraw-Hill, New York, USA.

(Knublauch et al.2004) Knublauch, H., Fergerson, R.W., Noy, N.F., and Musen, M.A. (2004). The protégé owl plugin: An open development environment for semantic web applications. 3rd International Semantic Web Conference (ISWC 2004), pages 229–243. Springer

(Lancaster et al., 2000) Lancaster, J., Woldorff, M., Parsons, L., Liotti, M., Freitas, C., Rainey, L., Kochunov, P., Nickerson, D., Mikiten, S., and Fox, P. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10(3):120–31.

(Lawrie et al., 1999) Lawrie, S. M., Whalley, H., Kestelman, J. N., Abukmeil, S. S., Byrne, M., Hodges, A., Rimmington, J. E., Best, J. J. K., Owens, D. G. C., and Johnstone, E. C. (1999). Magnetic resonance imaging of brain in people at high risk of developing schizophrenia. *Lancet*, 353:30–33.

(Lawrie et al., 2008) Lawrie, S.M., Hall, J., McIntosh, A.M., Cunningham-Owens, D.G., Johnstone, E.C.. Neuroimaging and molecular genetics of schizophrenia: pathophysiological advances and therapeutic potential. *British Journal of Pharmacology*, 153: S120–S124, 2008.

(Martin et al., 2001) Martin, R. F., Mejino, J. L. V., Bowden, D. M., Brinkley, J., and Rosse, C. (2001). Foundational model of neuroanatomy: Implications for the human brain project. In *Proc. of the American Medical Informatics Association Fall Symposium*, pages 438–442, Philadelphia, USA. Hanley & Belfus.

(Nuechterlein and Dawson, 1984) Nuechterlein, K. H. and Dawson, M. E. (1984). A heuristic vulnerability/stress model of schizophrenic episodes. *Schizophrenia Bulletin*, 10:300–12.

(Parent and Carpenter, 1996) Parent, A. and Carpenter, M. (1996). *Carpenter's Human Neuroanatomy*. Williams & Wilkins, Baltimore, USA.

(Razzouk, 2001) Razzouk, D. (2001). *Construção de uma base de conhecimento de um sistema de apoio a decisao*. PhD thesis, UNIFESP, São Paulo, Brazil.

(Rescher, 2005) Rescher, N. (2005). *Epistemic Logic: A Survey Of the Logic Of Knowledge.* University of Pittsburgh Press, Pittsburgh, USA.

(Rosse and Mejino, 2003) Rosse, C. and Mejino, J. (2003). A reference ontology for biomedical informatics: the foundational model of anatomy. *Journal of Biomedical Informatics*, 36(6):478–500.

(**Santos and Cabalar, 2008)** Santos, P. and Cabalar, P. (2008). The space within fisherman's folly: Playing with a puzzle in mereotopology. *Spatial Cognition and Computation*, 8(1 & 2):47–64.

(Santos et al., 2009) Santos, P., Thomaz, C., Santos, D., Freire, R., Sato, J., Louza, M., Sallet, P., Busatto, G., and Gattaz, W. (2009). Exploring the knowledge contained in neuroimages: statistical discriminant analysis and automatic segmentation of the most significant changes. *Unpublished draft*.

(Smith and Ceusters, 2007) Smith, B. and Ceusters, W. (2007). Ontology as the core discipline of biomedical informatics: Legacies of the past and recommendations for the future direction of research. In Crnkovic, G. and Stuart, S., editors, *Computing, Philosophy, And Cognitive Science*, pages 104–122.Cambridge Scholars Press, Newcastle, UK.

(Smith et al., 2005) Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A. L., and Rosse, C. (2005). Relations in biomedical ontologies. *Genome Biology*, 5(6).

(Smith and Varzi, 2000) Smith, B. and Varzi, A. (2000). Fiat and bona fide boundaries. *Philosophy and Phenomenological Research*, 60(2):401–420.

(Spitzer and Endicott, 1968) Spitzer, R. L. and Endicott, J. (1968). Diagno: computerized program for psychiatric diagnosis utilizing the differential diagnostic procedure. *Archives of General Psychiatry*, 18:747–756.

(Steen et al., 2006) Steen, R. G., Mull, C., McClure, R., Hamer, R. M., and Lieberman, J. A. (2006). Brain volume in first-episode schizophrenia. *British journal of psychiatry*, 188:510–518.

(Tan et al., 2008) Tan, H.-Y., Callicott, J.H., Weinberger, D.R. Intermediate phenotypes in schizophrenia genetics redux: is it a no brainer? *Molecular Psychiatry*, 13: 233-238, 2008.

(Thomaz et al., 2007a) Thomaz, C. E., Boardman, J. P., Counsell, S., Hill, D. L. G., Hajnal, J. V., Edwards, A. D., Rutherford, M. A., Gillies, D. F., and Rueckert, D. (2007a). A multivariate statistical analysis of the developing human brain in preterm infants*. Image Vision Computing*, 25(6):981–994.

(Thomaz et al., 2007b) Thomaz, C. E., Duran, F. L. S., Busatto, G. F., Gillies, D. F., and Rueckert, D. (2007b). Multivariate statistical differences of MRI samples of the human brain. *Journal of Mathematical Imaging and Vision*, 29(2-3):95–106.

(Varzi, 1997) Varzi, A. (1997). Boundaries, continuity, and contact. NOUS, 31 (1): 26-58

(Varzi, 1996) Varzi, A. C. (1996). *Reasoning about space: The hole story*. Logic and Logical Philosophy, 4:3–39. 23A

A Key terms and definitions

1. spatial ontology: logical formalisms that make clear the spatial foundations of domains;
2. Basic Inclusion Theory: a region-based formalism for building biomedical ontologies, which makes a clear distinction between mereological and location relations;
3. Fiat boundaries: limits that are commonly accepted but which do not have a concrete existence;
4. Meta analysis: a statistical procedure for combining the results of independent studies (usually from the published research material);
5. Kantian confusion: usually referred to the confusion between objects and concepts;
6. Epistemic logics: logical formalisms that represent (and reason about) knowledge and how knowledge is acquired;
7. Linear discriminant analyses: set of statistical methods used to find the linear combination of features which best separate two groups of data points;
8. Active contour models: also known as snakes are methods for segmenting an object outline from a possibly noisy image by means of minimizing an energy function associated to the current contour.