

Centro Universitário da FEI
Rodolpho Freire

**Uma ontologia espacial sobre a neuroanatomia do sistema
ventricular cerebral e sua integração ao conhecimento sobre
neuroimagens em esquizofrenia**

São Bernardo do Campo, SP
2009

Rodolpho Freire

**Uma ontologia espacial sobre a neuroanatomia do sistema
ventricular cerebral e sua integração ao conhecimento sobre
neuroimagens**

Dissertação de Mestrado apresentada ao Departamento de Engenharia Elétrica do Centro Universitário da Fundação Educacional Inaciana como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Inteligência Artificial

Orientador: Prof.Dr.Paulo Eduardo Santos

São Bernando do Campo, SP

2009

Freire, Rodolpho

Uma ontologia espacial sobre a neuroanatomia do sistema ventricular cerebral e sua integração ao conhecimento sobre neuroimagens de esquizofrenia / Rodolpho Freire. São Bernardo do Campo, 2009.

118f. : il.

Dissertação - Centro Universitário da FEI

Orientador: Prof. Paulo Eduardo Santos

1 - Ontologias 2 - Segmentação de Imagens 3 - Análise Estatística

I. Santos, Paulo Eduardo, orient. II. Título.

CDU 681.32

Resumo

Este trabalho apresenta o desenvolvimento de uma ontologia espacial para descrever o sistema ventricular cerebral e o conhecimento presente em estudos sobre neuroimagens. Iniciamos propondo uma base para esta ontologia formulada em lógica de primeira ordem provendo formas de representar o conhecimento extraído de neuroimagens e presente nas meta-análises. Após a concepção da base em lógica de primeira ordem, implementamos a teoria no Protégé gerando consultas e formas de representação adequadas aos conceitos. Então, com a teoria implementada, discutimos como representar o conhecimento extraído de neuroimagens por meio de segmentação de imagens e técnicas de análise estatística. Como resultado final obtemos uma ontologia-epistemologia que representa a mereologia do sistema ventricular cerebral e pode ter novas classes inseridas por meio dos resultados da análise estatística, bem como representação de valores encontrados na literatura.

Palavras-chave: Ontologias, Segmentação de Imagens, Análise estatística.

Abstract

This work proposes the development of a spacial ontology in order to define the ventricular brain system and the knowledge present in research about neuroimages. Initially we have proposed a formal base for this ontology which is formulated in first order logic and, we have created ways of representing the knowledge extracted from neuroimages and that is contained in meta-analysis. After conceiving the first order logic basis, the theory was implemented in Protégé, creating queries and appropriate representation ways of the concepts. Then, with the theory implemented, we discussed how to represent the knowledge extracted from neuroimages through image segmentation and statistical analysis. As a final result, we have obtained an ontology-epistemology that represents the mereology of the ventricular brain system, which may insert new classes discovered by statistical analysis as well as representation of values found in literature.

Keywords: Ontologies, Image Segmentation, Statistical Analysis.

Agradecimentos

Agradecer uma jornada de pouco mais de dois anos é sempre uma tarefa complexa, pois sempre que citamos nomes podemos esquecer alguém que foi de grande importância. Então desde já digo que os agradecimentos aqui são extensivos a todos que tiveram direta ou indiretamente participação nessa jornada.

Agradeço primeiramente a Deus, sem o qual não existiriam motivos para agradecer.

Ao meu orientador Prof. Dr. Paulo Eduardo Santos, que além da orientação e das valiosas contribuições, se mostrou um amigo compreensivo e um grande motivador para esse trabalho e para a carreira acadêmica.

Agradeço aos mestres que nos abrilhantaram com seu conhecimento: Prof. Dr. Paulo Sérgio Rodrigues e Prof. Dr. Reinaldo Augusto da Costa Bianchi e em especial ao Prof. Dr. Flavio Tonidandel, membros da banca de qualificação

Aos meus companheiros de turma: Horst, Caruso, Haddad, Valquíria, Vanessa, Danilo Silva, Thiago, Rafael, Giulianno e especialmente o companheiro Danilo Nunes, que contribuiu de forma importante com os resultados, fornecendo e explicando os resultados das análises estatísticas entre outras incursões da vida de um mestrando.

Agradeço também aos amigos da Redecard: Edgard Dias e Edmilson Martinez, que contribuiriam com sua valiosa visão nas fases finais do trabalho, e a minha amiga Mariana que muitas vezes me ajudou nas revisões.

À minha família, minha mãe, que mesmo nas dificuldades e nem sempre podendo me dar o que gostaria sempre me deu apoio, bem como meus avós maternos, que como minha mãe, ajudaram na construção do meu caráter. Ao meu Pai, que mesmo tendo partido há tanto tempo, está ainda em minhas lembranças e no meu coração e ao meu grande amigo Marcos que faz parte dessa história tanto quanto os outros aqui citados.

A minha namorada Renata, que suportou e tolerou muitas vezes minha ausência e nem por isso deixou de estar ao meu lado, além de ser uma importante contribuidora ajudando na revisão do trabalho.

Ao Prof. Dr. Geraldo F. Busatto pelos direcionamentos e por conceder a base de imagens utilizada neste trabalho e em especial ao Prof. Dr. Paulo C. Sallet que direcionou as dúvidas de sua especialidade e esteve na banca de defesa, e ao Prof. Dr. Carlos Eduardo Thomaz que contribuiu com este trabalho tanto na banca de qualificação quanto na defesa.

À todos, um fraternal abraço !

“Há homens que lutam um dia, e são bons;
Há homens que lutam por um ano, e são melhores;
Há homens que lutam por vários anos, e são muito bons;
Há outros que lutam durante toda a vida, esses são imprescindíveis.”

Bertolt Brecht.

10.02.1898 - 04.08.1956

“Se hoje fosse o último dia de minha vida, queria fazer o que vou
fazer hoje? E se a resposta fosse Não muitos dias seguidos, sabia que
precisava mudar algo.”

Steve Jobs.

“A vida de um indivíduo só faz sentido se ajuda a tornar a vida das
demais criaturas mais nobres e mais belas.”

Albert Einstein

14.03.1879 18.04.1955

Lista de Figuras**Lista de Tabelas**

1	Introdução	p. 15
1.1	Objetivo	p. 16
1.2	Contribuições	p. 17
1.3	Organização do trabalho.	p. 18
2	Revisão Bibliográfica	p. 19
2.1	Ontologias	p. 19
2.1.1	A Construção da Ontologia - Conceitos	p. 21
2.1.2	Ontologias Biomédicas	p. 23
2.2	Epistemologia	p. 25
2.3	Alterações morfológicas associadas à casos de esquizofrenia	p. 27
2.3.1	Estruturas morfológicas e características cerebrais afetadas pela esquizofrenia	p. 28
2.3.2	Alterações no Sistema Ventricular Cerebral	p. 28

Sumário

2.3.3	Volume cerebral	p. 30
2.3.4	Lobo Temporal	p. 30
2.3.5	Lobo Temporal Medial	p. 31
2.3.6	Giro Temporal Superior - GTS	p. 31
2.3.7	Tálamo e Gânglios da Base	p. 32
2.3.8	Assimetria Cerebral, Lobos Frontais, Occipitais e Parietais.	p. 32
2.4	O sistema ventricular cerebral	p. 33
2.5	Limites fiat	p. 34
3	Métodos	p. 36
3.1	Técnicas e ferramentas de análise de imagens médicas.	p. 36
3.1.1	Segmentação de imagens	p. 37
3.1.2	Segmentação por crescimento de região	p. 39
3.1.3	SPM - <i>Statistical Parametric Mapping</i>	p. 41
3.2	Estatística multivariada (PCA+MLDA) e testes de tamanho efetivo	p. 43
3.3	Teoria de inclusão básica - BIT	p. 45
3.3.1	Relações Inversas	p. 50
3.3.2	Axiomas Adicionais	p. 51
3.3.3	Princípio da Suplementação Fraca - <i>Weak Supplementation Principle</i>	p. 51
3.3.4	Teoria de Classes	p. 52
3.3.5	Classes Discretas	p. 53
3.3.6	Relações mereológicas na teoria de classes	p. 54
3.3.7	Implicações utilizando a teoria de classes	p. 56
3.3.8	Considerações sobre a escolha da BIT	p. 57
3.4	O Protégé	p. 59

Sumário

4 Resultados	p. 62
4.1 Classificação dos elementos do SVC utilizando a teoria de classes	p. 63
4.2 Representando a propriedade de continuidade	p. 66
4.3 Representando limites fiat	p. 67
4.4 Representando o sistema ventricular cerebral	p. 71
4.4.1 Formalizando detalhes anatômicos e forames	p. 74
4.5 Incluindo conhecimento na representação formal - Classes epistêmicas.	p. 75
4.6 Implementação da formalização ontológica e epistêmica no Protégé	p. 79
4.6.1 Classes epistêmicas no Protégé	p. 81
4.6.2 Raciocinando sobre a formalização proposta	p. 83
4.7 Individualizando o conhecimento em neuroimagens	p. 89
4.8 Integrando a ontologia aos dados extraídos das análises estatísticas	p. 92
4.8.1 Identificando regiões cerebrais	p. 95
4.9 Realimentando a ontologia com a expectativa gerada pelas análises estatísticas. . .	p. 96
4.10 Incluindo probabilidades encontradas na literatura	p. 101
4.10.1 Considerações Finais	p. 104
5 Conclusões e trabalhos futuros	p. 105
Referências bibliográficas	p. 111
Referências	p. 111
A Apêndice	p. 117

LISTA DE FIGURAS

1	Tipos de Ontologia.	p. 23
2	O <i>Foundational Model Explorer</i> (FME)	p. 24
3	O Sistema Ventricular Cerebral	p. 34
4	Exemplo de segmentação de imagens.	p. 38
5	Expansão da matriz de vizinhança do algoritmo de crescimento de região.	p. 39
6	Exemplo de imagem modelo segmentada pelo algoritmo de crescimento de região	p. 41
7	Segmentação de imagens no SPM	p. 43
8	Pontos que representam alterações estatísticas relevantes.	p. 46
9	Relações mereológicas da BIT.	p. 46
10	Interface gráfica do Protégé	p. 60
11	Estruturas que compõem o sistema ventricular cerebral.	p. 64
12	Árvore taxonômica baseada na partonomia das regiões ventriculares.	p. 65
13	Local onde são traçados os limites fiats.	p. 69
14	Classes epistêmicas e ontológicas.	p. 76
15	As abas de propriedades do domínio no Protégé.	p. 79
16	Classes do sistema ventricular cerebral e epistêmicas implementadas no Protégé.	p. 81

Sumário

17	Telas do Protégé para inserção de propriedades de objetos, dados e indivíduos	p. 82
18	Exemplo de instanciação de um indivíduo.	p. 82
19	Exemplo de classe epistêmica	p. 83
20	Inferências realizadas pelo Protégé.	p. 84
21	Resultado da consulta que indica a composição anatômica do ventrículo lateral direito.	p. 85
22	Resultado da consulta que indica a massa intermediária intertalâmica.	p. 86
23	Resultado da consulta que apresenta a anatomia do aqueduto cerebral.	p. 87
24	Resultado da consulta que apresenta a anatomia do aqueduto cerebral.	p. 88
25	MRicro - Campos que indicam o tamanho do voxel	p. 90
26	Seleção do Ponto Semente e total de pixels	p. 90
27	Indivíduo relacionado a uma classe ontológica e epistêmica.	p. 92
28	Projeção no espaço do MLDA	p. 93
29	Pontos que representam alterações estatísticas relevantes.	p. 94
30	Segmentação por meio de modelos deformáveis.	p. 94
31	Distância ao ponto de origem dado pelo MRicro.	p. 95
32	Interface para conversão de pontos do GingerAle.	p. 96
33	Níveis do Atlas de Talairach.	p. 97
34	Regiões segmentadas com maiores valores de sensibilidade	p. 99
35	Estruturas cerebrais inseridas no Protégé.	p. 100
36	Classe epistêmica que representa um percentual de ocorrência e redução dado pela literatura.	p. 102
37	Consulta sobre classes epistêmicas que representam valores da literatura.	p. 103

LISTA DE TABELAS

1	Consolidação das alterações morfológicas causadas pela esquizofrenia.	p. 29
2	Propriedades lógicas das relações de classe.	p. 56
3	Classes e subclasses epistêmicas e ontológicas.	p. 80
4	Volumes ventriculares, média e razão do grupo de controle e pacientes.	p. 91
5	Regiões apontadas pelo atlas de Talairach	p. 96
6	Regiões que apresentaram maior índice de sensibilidade local.	p. 98
7	Regiões cerebrais encontradas nos cortes que apresentam mais alta sensibilidade . . .	p. 98
8	Estruturas alteradas pela esquizofrenia e valores de alteração encontrados nas meta-análises	p. 101

CAPÍTULO 1

INTRODUÇÃO

A esquizofrenia é uma doença que afeta em média 1% da população mundial, independente de classe social, sexo, religião ou condição socioeconômica. Em sua sintomatologia são observadas: alucinações, delírios e perda da noção da realidade, causando um deslocamento social do indivíduo. Esse efeito pode ser observado no filme “Uma mente Brilhante”, que retrata a vida de John Forbes Nash, matemático criador da teoria dos jogos e ganhador do prêmio Nobel em 1994, portador de esquizofrenia.

A esquizofrenia é considerada uma doença de difícil diagnóstico pois seus sintomas isoladamente são presentes em muitas outras patologias neurológicas, as alterações morfológicas em pacientes *in vivo* começaram a ser observadas na década de 1920 por estudos de pneumoencefalografia. Assim anteriormente a esse período as evidências eram baseadas em estudos *post mortem* [1]. Essa dificuldade de fundamentar, diagnosticar e tratar a esquizofrenia levou Plum¹ a umas das mais conhecidas declarações no meio da neurociência: “A esquizofrenia é o cemitério dos neuropatologistas”[1].

Com os avanços tecnológicos das últimas décadas, surgiram novas técnicas altamente eficientes de diagnóstico por imagem, que fornecem informações sobre o corpo humano com um alto grau de confiabilidade e que possibilitam realizar um diagnóstico *in vivo* rápido e preciso [2, 3].

Além das técnicas de diagnóstico por imagem, alguns sistemas computacionais foram construídos para dar apoio a essa tarefa, geralmente sistemas especialistas baseados em regras de complexidade

¹Plum,F, em sua publicação *Prospects for research on schizophrenia. V. Pharmacological observations, Drug-induced psychoses*, 1972.

alta e interface pouco amistosa com o usuário [4, 5].

Pesquisa e desenvolvimento em inteligência artificial permitiu a criação de sistemas inteligentes utilizando técnicas como aprendizado de máquina, probabilidades bayesianas, redes neurais, árvores de decisão, entre outras. Recentemente, com computadores mais poderosos, a visão computacional ganhou importância significativa na construção desses sistemas de apoio[6], permitindo a segmentação de regiões de interesse e aplicação de técnicas estatísticas que permitem classificar imagens evidenciando características discriminantes [7].

Com esses importantes avanços, aumentou-se a expectativa de criação de sistemas inteligentes que possam representar todo, ou parte, do conhecimento do especialista sobre o domínio. Nesse contexto, as ontologias ganharam espaço na representação do conhecimento em domínios complexos, possibilitando desta forma a realização de inferências sobre os objetos ou regiões que o compõem.

Basicamente, a ontologia em sistemas de informação é uma forma de descrever a partir de uma linguagem formal, os objetos de um domínio e seus relacionamentos. São amplamente utilizadas para representação de domínios espaciais em sistemas de informação geográfica (GIS) [8] e sistemas biomédicos [9].

Em nosso trabalho, o domínio de interesse é o sistema ventricular cerebral, região afetada pela esquizofrenia e evidenciada pelos estudos ao longo da história como veremos posteriormente. Ainda propomos a integração da ontologia com neuroimagens que são submetidas à técnicas de análise estatística e com o conhecimento oriundo da literatura, de forma a realimentá-lá com novas informações, o que nos disponibilizaria uma base de conhecimento consistente e atualizada sobre achados relevantes.

1.1 Objetivo

O objetivo deste trabalho é demonstrar a formalização de uma região cerebral afetada pela esquizofrenia (o sistema ventricular cerebral), por meio de uma ontologia espacial, e integrar essa teoria formal ao conhecimento extraído de neuroimagens e da literatura médica. A integração da ontologia com neuroimagens é feita em dois momentos distintos. O primeiro demonstra como extrair conhecimento e informações do domínio por meio de segmentação de imagens. O segundo, apresenta como podemos converter pontos (que representam diferenças entre grupos de paciente e controle saudáveis) em informação de alto nível, e como essa informação pode realimentar a ontologia.

A segmentação de imagens neste trabalho é feita em duas etapas, na primeira segmenta-se a neuroimagem em massa cinza, massa branca e líquido cefalorraquidiano por meio do *statistical parametric mapping* (SPM), um software que tem como plataforma o Matlab e é amplamente utilizado para análise de neuroimagens. A segunda etapa da segmentação é feita utilizando o algoritmo de crescimento de região com interação do usuário, que permite escolher o ponto de interesse da região qual deseja-se segmentar.

A descoberta de achados relevantes por meio de técnicas de análise estatística é feita através do *maximum uncertainty linear discriminant analysis* (MLDA). O MLDA é um classificador linear utilizado para separar em grupos diferentes conjuntos de amostras. Após a classificação, os voxels que caracterizam os fatores discriminantes são identificados pelo teste de tamanho efetivo de voxels. Os pontos encontrados são segmentados por meio de modelos deformáveis [7], o que nos permite extraí-los e mapeá-los em um atlas cerebral (neste caso o atlas de Talairach) para conhecer as regiões anatômicas representadas por esses pontos.

Dada essa etapa, demonstraremos como incluir as regiões descobertas na ontologia, de forma a realimentá-la, além de incluir o conhecimento encontrado na literatura. Para essa tarefa, utilizou-se uma ontologia espacial (a Teoria Básica de Inclusão [10]), que permite a formalização de conceitos anatômicos em lógica de primeira ordem. Durante o desenvolvimento da ontologia, propõe-se a extensão dessa teoria criando relações que definem limites entre estruturas e continuidade e formas de representação de conhecimento não inerentes ao domínio. Após a descrição formal da entidade, implementamos essa formalização em um navegador especializado para codificação de ontologias (Protégé), que nos permite a realização de inferências e testes de validade da ontologia.

1.2 Contribuições

As principais contribuições desta dissertação são:

1. Construção de uma ontologia espacial sobre o sistema ventricular cerebral.
2. Investigação sobre como assimilar o conhecimento de esquizofrenia presente na literatura à teoria formal.
3. Implementação da formalização no Protégé.

4. Descrição de estruturas cerebrais identificadas pela análise estatística por meio de um atlas cerebral e sua integração à ontologia.

1.3 Organização do trabalho.

Este trabalho está organizado da seguinte maneira:

- **Capítulo 2** - Revisão Bibliográfica.

Apresenta uma revisão da literatura para os temas que compõem esta dissertação: ontologias, epistemologia esquizofrenia, o sistema ventricular cerebral e limites fiat;

- **Capítulo 3** - Métodos.

Apresenta os métodos e ferramentas utilizadas neste trabalho para a construção dos resultados.

- **Capítulo 4** - Resultados.

Apresenta os resultados do trabalho. Em síntese, nesse capítulo são descritos os resultados da formalização em lógica de primeira ordem, a implementação da ontologia no Protégé e como podemos integrar a ontologia com técnicas de análise de imagem.

- **Capítulo 5** - Conclusões e trabalhos futuros.

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA

Nesta seção faremos uma revisão bibliográfica dos principais assuntos desta dissertação, abordando os temas:

1. Ontologias;
Conceitos sobre ontologias e trabalhos relacionados a ontologias biomédicas.
2. Epistemologia;
O que é a epistemologia e para que ela é usada junto à ontologia.
3. Esquizofrenia;
Uma visão geral da patologia e sua ação no cérebro.
4. Sistema Ventricular Cerebral;
A anatomia do sistema ventricular cerebral.
5. Limites fiat.
Conceito e utilização de limites fiat.

2.1 Ontologias

Smith [11] define a ontologia como um dos principais ramos da metafísica, uma área da filosofia que estuda fatos que estão além dos eventos que são observáveis no mundo natural. A palavra “On-

tologia” é derivada do grego (*ontos+logoi*) que significa “conhecimento do ser”.

A ontologia estuda os objetos de um domínio, sejam eles reais ou abstratos, vivos ou não, sendo responsável por prover formas de descrever e categorizar esses objetos e relações entre eles e com o mundo ao qual pertencem.

Segundo Uschold e Grüninger [12], uma ontologia pode ser representada por diferentes graus de formalismo do vocabulário utilizado para sua criação, classificando-as em:

- **Altamente informal;**

Baseada em linguagem natural.

- **Semi informal;**

Criada em linguagem natural restritiva e estruturada, o que permite redução de ambiguidades e aumento da clareza dos conceitos expressados.

- **Semi formal;**

Construída em uma linguagem artificial formalmente constituída.

- **Formal.**

Construção ontológica feita com uma linguagem artificial formalmente constituída com semânticas formais, teoremas e provas (completude e corretude).

As ontologias formais foram introduzidas por Edmund Husserl (1859-1938) como uma maneira de formalizar o conhecimento adquirido pela ontologia descritiva, que se utiliza da linguagem natural. As principais características das ontologias formais são serem construídas com elementos da lógica em conjunto com outras teorias como álgebra, mereologia, teoria dos conjuntos, topologia, etc. Isso torna as ontologias formais mais expressivas para representação de domínios complexos e programáveis computacionalmente [13, 14].

Com a construção de sistemas computacionais para automação de tarefas, iniciou-se a elaboração de sistemas que atuam em domínios diferentes com terminologias similares, porém com significados diferentes. Esse problema é conhecido por “Torre de Babel” [11].

Problemas de terminologia e incompatibilidades inicialmente eram resolvidos caso a caso. Segundo Smith [11], a primeira grande tarefa das ontologias foi melhorar esse processo com o uso de axiomas e teorias formais, que permitem compartilhar e reusar o conhecimento adquirido. Desde

então, a inteligência artificial vem explorando as ontologias formais com a mesma premissa básica da filosofia, que é criar uma definição formal de um domínio.

Basicamente a ontologia utilizada em sistemas de informação é similar a um modelo de dados. Entretanto, apesar da similaridade, a ontologia é uma forma de descrever (a partir de uma linguagem formal) os objetos de um domínio e seus relacionamentos. Compostas por quantificadores lógicos universais, as ontologias formais permitem inferências e extração de conhecimento a partir da formalização [11], enquanto o modelo de dados se limita a descrever relacionamentos cardinais entre as entidades [15, 16].

Podemos destacar como representantes do uso de ontologias, sistemas de web-semântica¹ [11], sistemas biomédicos [9] e sistemas de informações geográficas (GIS) [8].

2.1.1 A Construção da Ontologia - Conceitos

Uschold e Grüninger [12] apresentam em seu trabalho uma metodologia para construção de ontologias que consiste nos seguintes passos:

1. Identificação do objetivo e escopo da ontologia;

Nesta etapa questionamos o propósito da criação da ontologia e quais questões ela deve responder.

2. Construção da ontologia;

A construção da ontologia pode ser dividida em 3 fases:

- Captura da ontologia;

Nesta etapa descobrem-se os conceitos-chave e relacionamentos que compõem a ontologia, a descrição e identificação dos termos que representarão esses conceitos.

- Codificação da ontologia;

Trata da representação da ontologia através de uma linguagem formal para representação dos conceitos como classes, entidades e relações.

- Integração com ontologias existentes.

Permite e facilita a integração ou reusabilidade de conceitos com outras ontologias já construídas sobre o domínio de interesse.

¹A Web Semântica é uma extensão da Web atual que permite dar sentido semântico a informação para integrar informações entre páginas da Web [16].

3. Avaliação;

Consiste em verificar se a modelagem do domínio e a ontologia construída são suficientemente expressivas para responder a questionamentos feitos através de linguagem formal.

4. Diretrizes para a modelagem ontológica;

Define-se uma série de princípios para construção da ontologia. Segundo Gruber [14] esses princípios são: clareza, objetividade, coerência, completeza e extensibilidade além de evitar a utilização de simbologias próprias do domínio.

5. Documentação.

Elabora-se a documentação da ontologia, descrição dos conceitos e componentes da ontologia.

As ontologias formais são construídas, conforme já citado anteriormente, em conjunto com uma série de teorias, as quais podemos destacar:

1. **Taxonomia;**

Permite a classificação do conhecimento de um domínio de forma hierárquica (em classes e sub-classes) [17]. Na taxonomia temos relações como $Is_a(x,y)$ que indica que um entidade x é classificada como sub-classe de uma entidade y .

2. **Mereologia;**

É a teoria formal da parte e suas relações com o todo. A mereologia descreve as regiões por meio de relações independentes de formalismos matemáticos, como pontos e linhas da geometria tradicional.

Na mereologia encontramos as relações do tipo $P(x,y)$ (parte) indica que um objeto ou região x possui uma parte y , que pode ser inclusive a totalidade do observado. Relações do tipo $PP(x,y)$ (parte própria) representam que as partes constituintes do objeto são uma fatia de sua totalidade, ou seja, que x nunca será igual a y .

3. **Topologia;**

Permite a formalização de conceitos como conexidade e continuidade. Na topologia temos relações de localização como: $x\ Connect_a\ y$, que indica que um objeto x está conectado a um objeto y .

Ainda, as ontologias podem ser classificadas de acordo com a tarefa a ser realizada. Guarino [18] classifica as ontologias como:

- Ontologias de domínio e tarefa;
São ontologias criadas para formalizar um domínio ou uma tarefa de modo específico.
- Ontologias de aplicação;
São especializações das ontologias de domínio e tarefa, ontologias mais específicas para uma determinada aplicação.
- Meta-ontologias ou ontologias de nível-topo.
Provém uma conceitualização genérica do domínio.

A figura 1, apresenta um esquema gráfico dos tipos de ontologia:

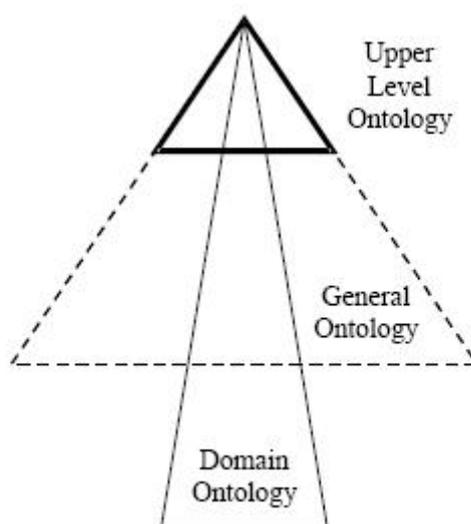


Fig. 1: Esquema gráfico que representa os tipos de ontologia: domínio, aplicação e nível topo.
Fonte: adaptado de Bodenreider *et al.* [19]

Dentre as ontologias criadas para representar diversos domínios, o interesse deste trabalho reside em ontologias biomédicas conforme descreveremos a seguir.

2.1.2 Ontologias Biomédicas

Podemos destacar como principais ontologias biomédicas o modelo de anatomia fundacional (FMA) ² [20], o GALEN [21, 22], a teoria básica de inclusão (BIT) ³ [10], a biblioteca aberta de

²Do inglês *Foundational Model Anatomy*

³Do inglês *Basic Inclusion Theory*

ontologias biomédicas (OBO) ⁴ [23] e a ontologia formal básica (BFO) ⁵ [24].

FMA é uma ontologia usada para a terminologia anatômica e possibilita a construção de um dicionário de dados para classificação de indivíduos desse domínio, como podemos observar na figura 2.

Desenvolvido inicialmente por anatomistas e posteriormente com a colaboração da comunidade de pesquisadores, foi redefinida recentemente visando fortalecer sua base lógica para aplicação em sistemas biomédicos. Assim, FMA se consolida entre as ontologias biomédicas como sendo uma importante ferramenta para formalização do conhecimento sobre anatomia [20, 10].

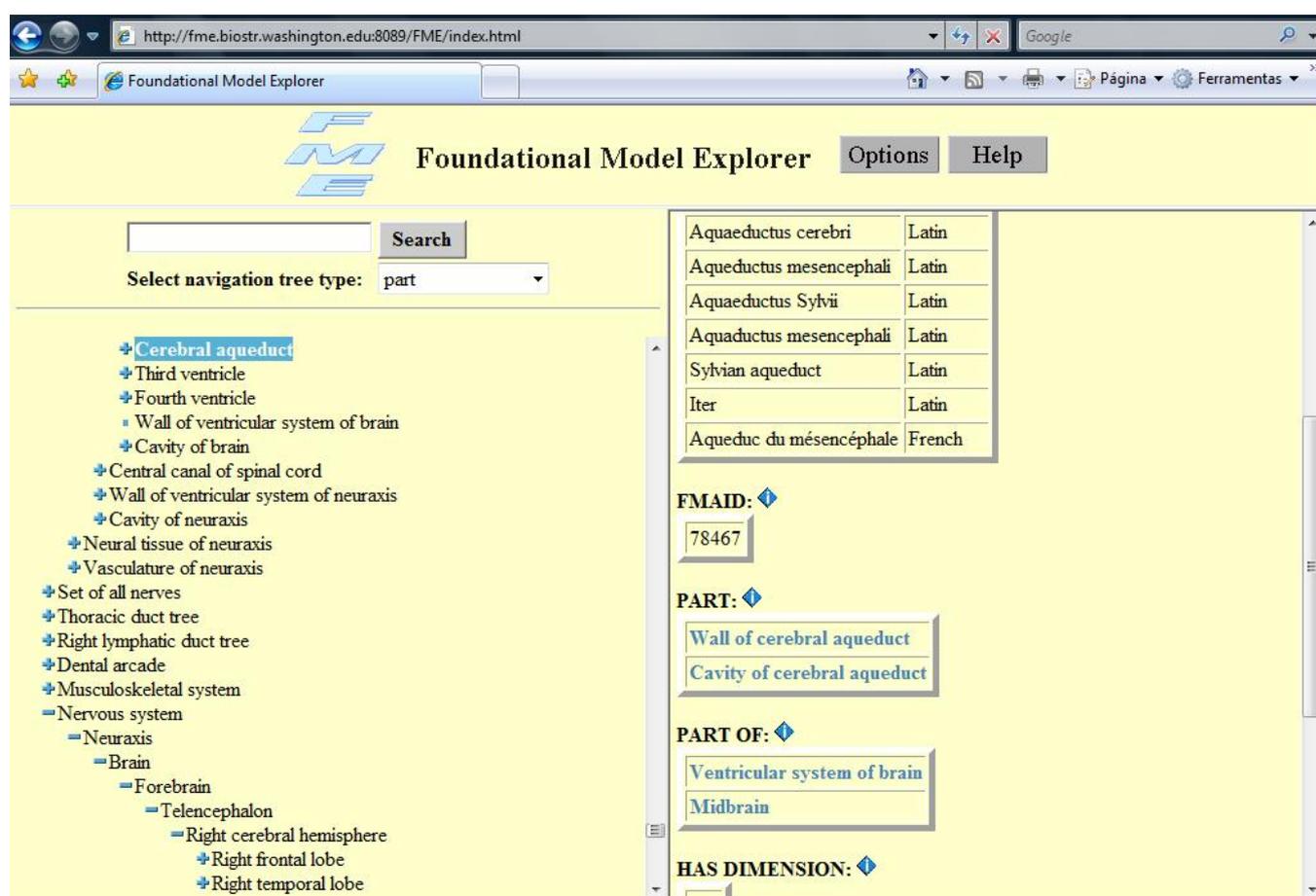


Fig. 2: O *Foundational Model Explorer* (FME), navegador que permite interação com a formalização do FMA.

GALEN [21, 22], assim como o FMA, também foi inicialmente construído por especialistas e

⁴Do inglês *Open Biomedical Ontologies*

⁵Do inglês *Basic Formal Ontology*

passou por adaptações que o tornou mais robusto para o uso em domínios biomédicos, tendo amplas aplicações como farmacologia, patologias e anatomia [21, 10].

OBO é uma biblioteca aberta para várias ontologias no domínio biológico e médico. A biblioteca OBO Relacional [23] é uma ontologia baseada em relações mereológicas, taxonômicas e temporais.

BFO [24] é uma ontologia criada com bases mereológicas que conta com relações de parte-todo, continuidade, definições de limites entre objetos além de relações temporais e de diferenciação entre o objeto e região espacial que ele ocupa. Outra característica da BFO é que suas relações são formuladas em lógica de primeira ordem.

BIT [10] assim como BFO, é uma mereologia criada com base em relações espaciais formalizadas em lógica de primeira ordem o que permite que as relações sejam facilmente utilizadas em conjunto com outras teorias complementares. BIT também conta com uma teoria taxonômica e de instanciação de classes que permite uma melhor organização taxonômica do domínio.

Por exemplo, podemos expressar (usando BIT) o axioma $PP(\text{ventriculo_direito}, \text{coracao})$ que nos diz que o ventrículo direito é parte própria do coração, com sua extensão taxonômica podemos ter a relação $P_{12}(\text{ventriculo_direito}, \text{coracao})$ que expressa que todo ventrículo direito é parte própria de um coração e todo coração tem um ventrículo direito como parte própria.

Utilizando a teoria de classes, ainda podemos expressar que todo coração é um órgão por meio da relação Is_a . Mais detalhes sobre BIT, e uma discussão sobre a sua escolha em relação às demais ontologias, será apresentada no capítulo 3.

Entretanto, apenas a formalização do domínio muitas vezes é insuficiente para representar todos os conceitos. Na próxima seção abordaremos a epistemologia, uma teoria que se ocupa da representação do conhecimento, característica presente em nosso trabalho, mas não representada pela teoria ontológica.

2.2 Epistemologia

Como vimos na seção 2.1, as ontologias são usadas para descrever o domínio, para formalizar e representar suas características. Entretanto não é escopo da teoria ontológica pela sua própria definição, representar o conhecimento que possa ser gerado sobre domínio de interesse e que não seja explícito em sua estrutura (o que sabemos *versus* o que desejamos e podemos saber) [25].

A estruturação e organização do conhecimento sobre o domínio é a tarefa da epistemologia, ramo da filosofia que trata da investigação sobre a natureza, fontes e validade do conhecimento. Em computação, podemos citar a epistemologia presente em discussões sobre as novas gerações de GIS [26], representação de conhecimento humano em sistemas de informação geográfica [27], sistemas biomédicos abordando terminologia médica [28] e classificação de diferenças de terminologia em diagnósticos [29].

Nosso trabalho busca o desenvolvimento de uma ontologia que permita inferir fatos sobre o sistema ventricular cerebral, entretanto é desejável que também possamos incluir e inferir fatos sobre esquizofrenia, exaustivamente pesquisados e publicados em trabalhos de meta-análise ou explorados por meio de técnicas de análise de imagens.

O conhecimento presente nesses trabalhos, entretanto, não representa o domínio (sistema ventricular cerebral) propriamente dito e sim conceitos desse domínio, como por exemplo: grupos de pacientes e grupos de controle, onde o especialista baseado na observação e em características desses grupos pode concluir o diagnóstico e descobrir novos fatos referente à patologia.

Para exemplificar (de modo genérico) podemos conceber o sistema ventricular cerebral de um indivíduo x_i qualquer. Esse sistema ventricular cerebral é uma representação do domínio formalizado pela ontologia, ou seja, um indivíduo ontológico que representa características desse domínio.

Ao analisarmos características dessa entidade (sistema ventricular cerebral do paciente x_i) em relação a outras entidades, podemos por exemplo, classificá-lo como pertencente ao grupo de pacientes. Essa classificação é baseada em uma série de características presentes no domínio (por exemplo, o volume ventricular), mas não são mais representações desse domínio e sim de um conhecimento externo aos elementos *qua* elementos (como grupos de pacientes ou controle).

Em nosso trabalho, propomos a formalização do sistema ventricular cerebral e do conhecimento obtido através de meta-análises e análises estatísticas, o que caracteriza a necessidade da inclusão da epistemologia juntamente com a ontologia na construção de uma representação formal. Sem uma representação epistêmica seria incorreto designar como indivíduos da ontologia o conhecimento das análises médicas, já que não são características do domínio de interesse.

Também não poderíamos designar essa informação como classes ontológicas, o que nos levaria a uma “confusão Kantiana” [25]. A confusão Kantiana significa tratar conceitos como objetos ou indivíduos. Por exemplo, tratar a região que o objeto ocupa como o objeto, caracteriza uma confusão Kantiana, ou ainda formalizar na ontologia (como se fossem características do domínio) o conheci-

mento extraído da literatura.

Trabalhos de meta-análises, bem como alterações morfológicas causadas pela esquizofrenia, serão abordados na próxima seção.

2.3 Alterações morfológicas associadas à casos de esquizofrenia

A esquizofrenia, inicialmente denominada por Emil Kraepelin como “*dementia preacox*”, [1, 30] é uma desordem mental que atinge aproximadamente 1% da população mundial [1], porém suas causas e efeitos ainda são objetos de pesquisas [1].

A dificuldade de fundamentar, diagnosticar e tratar a esquizofrenia é caracterizada pelo conjunto de sintomas inerentes da esquizofrenia serem comuns a muitas outras patologias neurológicas como o mal de Alzheimer, distúrbios de humor e hidrocefalia, além da sintomática proveniente do uso de substâncias como agentes quimioterápicos e corticosteróides [31].

A esquizofrenia apresenta um conjunto de sinais e sintomas que incluem alterações morfológicas, alterações comportamentais e psicológicas.

Nas últimas décadas, diversos estudos realizados *post mortem* e posteriormente *in vivo* [1, 2, 31] com o auxílio de técnicas de diagnóstico por imagem, comprovaram a presença de alterações morfológicas e morfométricas em diversas estruturas do cérebro de pacientes esquizofrênicos [31, 30, 32].

A ressonância magnética (RM) teve um papel importante para realização e avanço dos estudos realizados *in vivo*, por ser uma técnica de diagnóstico mais precisa em relação a técnicas como a pneumoencefalografia, ventriculografia, tomografia computadorizada e o ultra-som [2, 33, 1]. A vantagem da ressonância magnética provém de sua alta resolução e a possibilidade de executar cortes finos em qualquer orientação [2], além da ausência de radiações ionizantes presentes na técnica da tomografia computadorizada [34].

A seguir apresentaremos as principais alterações associadas à esquizofrenia, compiladas em trabalhos de meta-análise.

2.3.1 Estruturas morfológicas e características cerebrais afetadas pela esquizofrenia

Nesta seção descreveremos algumas das estruturas e características cerebrais que são afetadas pela esquizofrenia. Tomaremos como base alguns estudos baseados em meta-análises, de modo a ter uma visão ampla da ação da patologia o que nos auxiliará posteriormente na formalização do domínio.

Entre as principais alterações nas estruturas cerebrais descritas em [30] estão:

- Aumento dos ventrículos laterais e terceiro ventrículo;
- Redução do volume cerebral;
- Redução do lobo temporal e estruturas límbicas.

Existem diversas publicações que abordam a dilatação ventricular e alterações morfométricas nas estruturas cerebrais em pacientes esquizofrênicos. Nesta seção utilizaremos como principais referências dois trabalhos de meta-análise, o de Shenton *et al.* [1] e Lawrie e Abukmeil [35] entre outras referências. A tabela 1⁶ resume as alterações morfológicas que serão abordadas no decorrer da seção.

2.3.2 Alterações no Sistema Ventricular Cerebral

O alargamento dos ventrículos laterais foi inicialmente reportado em estudos realizados *post mortem* por Hecker em (1871) e por Kraepelin (1913) [30]. Esses estudos foram posteriormente reafirmados por exames realizados *in vivo* a partir de pneumoencefalografia, e recentemente com tomografia computadorizada e ressonância magnética [2].

Shenton *et al.* [1] aborda inicialmente estudos realizados com imagens de tomografia computadorizada, em que 75% dos casos foram observados alargamento dos ventrículos laterais. Este artigo também relata 55 estudos com imagens de ressonância magnética (IRM), onde 80% dos casos indicam alargamento ventricular significativo. Lawrie e Abukmeil [35] descrevem um aumento médio

⁶É importante ressaltar que este trabalho não classifica a literatura médica em casuística (literatura com base em casos de esquizofrenia medidos pelo autor) e em literatura de revisões bibliográficas, entretanto para fins de trabalhos futuros é importante que esse fato seja considerado com apoio dos especialistas.

Estrutura Afetada	Referências
Alargamento dos ventrículos laterais	[2, 30, 1], [36, 35], [37, 38]
Alargamento do terceiro ventrículo	[2, 30, 1]
Alargamento do quarto ventrículo	[2, 30, 1],
Redução do volume cerebral	[2, 30, 35], [1]
Redução dos lobos temporais	[2, 35, 1],
Redução do complexo amígdala hipocampo	[30, 35, 1],
Reduções da amígdala	[35]
Redução do hipocampo	[2, 30, 35],
Redução do giro parahipocampal	[2, 30, 35],
Redução do giro temporal superior	[2, 30, 1], [35, 39],
Alterações no talâmo	[30, 35, 1],
Aumento de volume da gânglia basal	[30, 35, 1]
Assimetria cerebral	[31]
Alterações dos lobos frontais	[35, 1]
Alterações dos lobos occipitais	[30, 35, 1]
Alterações dos lobos parietais	[2, 36, 1]

Tab. 1: Consolidação das alterações morfológicas causadas pela esquizofrenia apresentadas neste capítulo.

de 40% nos ventrículos laterais, mais relevantes no corpo e nos cornos occipitais dos ventrículos laterais.

Esse aumento do volume dos ventrículos laterais, em algumas de suas divisões anatômicas, também é relatado pela meta-análise de Shenton *et al.* [1]. Em seu estudo é apontado o alargamento do corno temporal do sistema ventricular sendo mais observado no ventrículo lateral esquerdo.

Segundo Sallet [30], os alargamentos ventriculares são apresentados em todas as divisões morfológicas do ventrículo lateral e independem do sexo do paciente como sugerem alguns estudos. Sallet apresenta dados de estudos sobre o alargamento ventricular do corpo do ventrículo lateral, corno frontal, corno occipital e corno temporal. Entretanto, Sallet enfatiza o aumento significativo do corpo do ventrículo lateral e cornos occipitais, indo ao encontro das observações de Lawrie e Abukmeil [35].

Valores de alteração no volume ventricular também são abordados por Steen *et al.* [38], que apresenta em sua meta-análise alterações de volume nos ventrículos lateral esquerdo e direito apontando um aumento médio de 33,7% para os ventrículos laterais esquerdo e 24,7% no ventrículo lateral direito.

O terceiro ventrículo também aparece nos estudos de alargamento ventricular em pacientes esquizofrênicos. Shenton *et al* [1] relata um estudo com 33 IRM, onde em 73% dessas imagens foram observados alargamento do terceiro ventrículo e 27% não apresentaram alargamento. Steen *et al* [38] reporta em sua meta-análise 8 estudos que abordam volumes ventriculares do terceiro ventrículo, uma alteração média de 25,3% no terceiro ventrículo.

Segundo Shenton *et al.* [1], o quarto ventrículo apresentou 20% de efeitos de alargamento contra 80% de amostras que não apresentaram alterações de volume, valores de alterações de volume não foram encontrados.

O alargamento ventricular é evidenciado ainda em muitos estudos como o de Elkis [32], Vita [40] e Harrison [31]. A diferença de volume dos ventrículos laterais e o terceiro ventrículo em grupos de paciente e controle é abordada por Steen *et al.* [38], que apresenta em seu trabalho de meta-análise, valores de evidência de alargamento ventricular de vários autores, como podemos observar na tabela 1.

2.3.3 Volume cerebral

A perda de volume cerebral é verificada com relevância em muitos estudos de meta-análise. Shenton *et al.* [1] relata que em 50 estudos conduzidos com imagens de ressonância magnética, o volume total cerebral menor que o normal são observados em 22% dos casos e 78% deles não existem alterações significativas.

Lawrie e Abukmeil [35] em seu estudo, observam uma redução do volume total do cérebro de 3% entre pacientes e grupo de controle. Castro [2] aborda a diminuição de volume e o aumento de líquido cefalorraquidiano nas regiões cerebrais.

Essa redução do volume cerebral está relacionada aos sinais abordados anteriormente: o alargamento ventricular, aumento dos ventrículos (onde é produzido parte do líquido cefalorraquiano) e perda de volume de algumas estruturas cerebrais [2] como veremos nas próximas seções.

2.3.4 Lobo Temporal

Shenton *et al.* [1] reporta a redução do volume nos lobos temporais em um estudo com 51 IRM, onde 61% apresentaram redução no volume e 39% não apresentaram alterações. Os autores também destacam que os estudos no lobo temporal podem ser conflitantes devido às diferenças metodológicas

e apresenta um estudo realizado por ela mesma com 31 IMR seguindo a mesma metodologia de análise nas imagens. Nesse estudo 55% das imagens apresentaram redução de volume dos lobos temporais e 45% não apresentaram diferenças.

Alterações nos lobos temporais são verificadas por Lawrie e Abukmeil [35], que relatam uma redução de 6% para o lobo temporal esquerdo e 9.5% para o lobo temporal direito. Castro [2] em seu estudo também relaciona a perda de volume dos lobos temporais, estendendo ainda sua análise à substância branca e cinzenta dos lobos temporais esquerdo e direito.

2.3.5 Lobo Temporal Medial

Segundo Shenton *et al.* [1], as estruturas do lobo medial temporal incluem o complexo amígdala-hipocampo e o giro para-hipocampal. O complexo amígdala-hipocampo é composto pela amígdala e o hipocampo devido a dificuldade de separá-los nos cortes de IMR, assim muitos estudos consideram-nas como uma única estrutura.

Entretanto, existem estudos que isoladamente tratam de diferenças morfológicas na amígdala e no hipocampo. Em seu estudo, Shenton *et al.* [1] relata que em 49 estudos de IMR das estruturas que compõem o lobo medial temporal, 74% reportaram alterações e em 26% não apresentaram diferenças.

Nos estudos de Lawrie e Abukmeil [35] para o complexo amígdala-hipocampo, foram observadas reduções de até 6.5% no tamanho total da estrutura. As reduções do complexo amígdala-hipocampo também são observadas nos estudos de Castro [2], Sallet [30] e Shenton *et al.* [39].

As estruturas amígdala e hipocampo são observadas separadamente por Lawrie e Abukmeil [35] e apresentam alterações morfométricas em ambos os casos. Na amígdala foi observada uma redução de em média 10% e no hipocampo de 7 a 8.5% em homens. Sallet [30] cita estudos que apontam uma redução bilateral de hipocampos de aproximadamente 4% e com a inclusão da amígdala os valores sobem para a média de 8%, enquanto o giro para-hipocampal esquerdo apresentou uma redução de até 14% e o direito de 9% [35], alterações essas também observadas por Shenton *et al.* [39] e Castro [2].

2.3.6 Giro Temporal Superior - GTS

Shenton *et al.* [1] apresenta 15 estudos de IMR, onde 67% reportaram reduções de volume no GTS e 33% não apresentaram alterações. Lawrie e Abukmeil [35] apontam que a diferença de

metodologias utilizadas pelos grupos de pesquisa para essa estrutura dificulta as comparações, mas apresentam que existe forte evidência de redução de volume na parte anterior e posterior, indo ao encontro de Shenton *et al.* [39] que aponta 15% de redução no giro temporal superior esquerdo.

2.3.7 Tálamo e Gânglios da Base

Em 12 estudos do tálamo, 42% apontaram redução de volume e 58% não foram observadas diferenças [1]. Lawrie e Abukmeil [35] apontam uma redução de 6 a 7% no volume dessa estrutura, redução também observada por Sallet [30].

Em 25 estudos de IMR sobre os gânglios da base, 68% reportaram aumento e 8% não reportaram alterações [1]. Sallet [30] explica que o aumento dos Gânglios da base pode ser devido à utilização de medicação. Esse aumento é relatado por Lawrie e Abukmeil [35] porém não há evidência de valores de redução nesse estudo.

2.3.8 Assimetria Cerebral, Lobos Frontais, Occipitais e Parietais.

A assimetria cerebral é definida como o torque do lobo frontal direito em relação ao lobo occipital esquerdo [41].

Alterações na assimetria cerebral são relatadas por Lawrie e Abukmeil [35] e Harrison [31].

Alterações em estruturas consideradas na assimetria cerebral são relatadas por Shenton *et al* [1] que observa que, em 50 estudos de IMR, 60% apresentaram redução dos lobos frontais, enquanto Lawrie e Abukmeil [35] apontam uma redução de até 5.5% nessa estrutura.

Nos lobos occipitais em 9 estudos de IMR, 44% reportaram reduções de volume e 56% não reportam diferenças [1], esses achados são corroborados por Sallet [30] e Lawrie e Abukmeil [35], este último relata reduções nos lobos occipitais de 6.5% em homens e 8.5% em mulheres.

Em relação aos lobos parietais, segundo Shenton *et al.* [1], em 15 estudos de IMR 60% reportaram achados positivos de perda de volume, também observadas por Honnea [36] e Castro [2].

Entre outras estruturas cerebrais afetadas pela esquizofrenia podemos ressaltar alterações na assimetria do *planum temporale* [1, 30], redução do volume do corpo caloso [30, 1], alterações morfológicas no *cavum septum pellucidum* [1, 31], bulbos olfatórios [1] e cerebelo [1, 36].

Esta seção mostrou as alterações morfológicas causadas pela esquizofrenia de acordo com difer-

entes autores. Iniciamos a seção apresentando uma compilação das principais alterações (veja tabela 1), e discutimos suas características e valores estatísticos baseado no trabalho de Shenton *et al.* [1], e de Lawrie e Abukmeil [35]. Outras referências complementaram o trabalho dando uma visão mais ampla sobre as alterações morfológicas apresentadas.

Sallet [30] em seu estudo cita as alterações ventriculares como uma das mais importantes. Como podemos perceber na tabela 1, o alargamento ventricular (ventrículos laterais, terceiro e quarto ventrículo) é apresentado em diversas literaturas.

Assim entendemos que mesmo não sendo uma alteração morfológica presente apenas na esquizofrenia [32, 31] o alargamento ventricular é uma alteração de alta relevância em seu diagnóstico [31, 1], sendo o motivador da construção de uma ontologia-epistemologia que represente a entidade e o conhecimento gerado pelas análises sobre o domínio. Esta ontologia será discutida no capítulo 4.

Na próxima seção faremos uma breve descrição das partes anatômicas que constituem o sistema ventricular cerebral.

2.4 O sistema ventricular cerebral

O sistema ventricular cerebral é uma cavidade no cérebro composta por um conjunto de estruturas morfológicas denominadas ventrículos e pequenos canais que fazem a ligação entre eles, denominados forames. Nos ventrículos estão situados os plexos coróides que são responsáveis pela produção de líquido cefalorraquiano, chegando a 30% do total produzido no corpo humano.

O sistema ventricular é composto pelos ventrículos laterais esquerdo e direito, terceiro ventrículo, quarto ventrículo, forame de Monro, aqueduto cerebral, forame de Luskha e forame de Magendie. Podemos observar na figura 3 um molde do sistema ventricular cerebral e sua anatomia.

Abaixo descreveremos de forma geral o sistema ventricular cerebral e os detalhes anatômicos de cada ventrículo.

Os ventrículos laterais são subdivididos em corpo, corno anterior (ou frontal), corno posterior (ou occipital) e corno inferior (ou temporal). Os ventrículos laterais se comunicam com o terceiro ventrículo através dos forames de Monro [43, 33].

O terceiro ventrículo é subdividido em recesso óptico, recesso suprapineal e infundíbulo e se comunica com os ventrículos laterais pelos forames de Monro (ou intraventriculares) e com o quarto

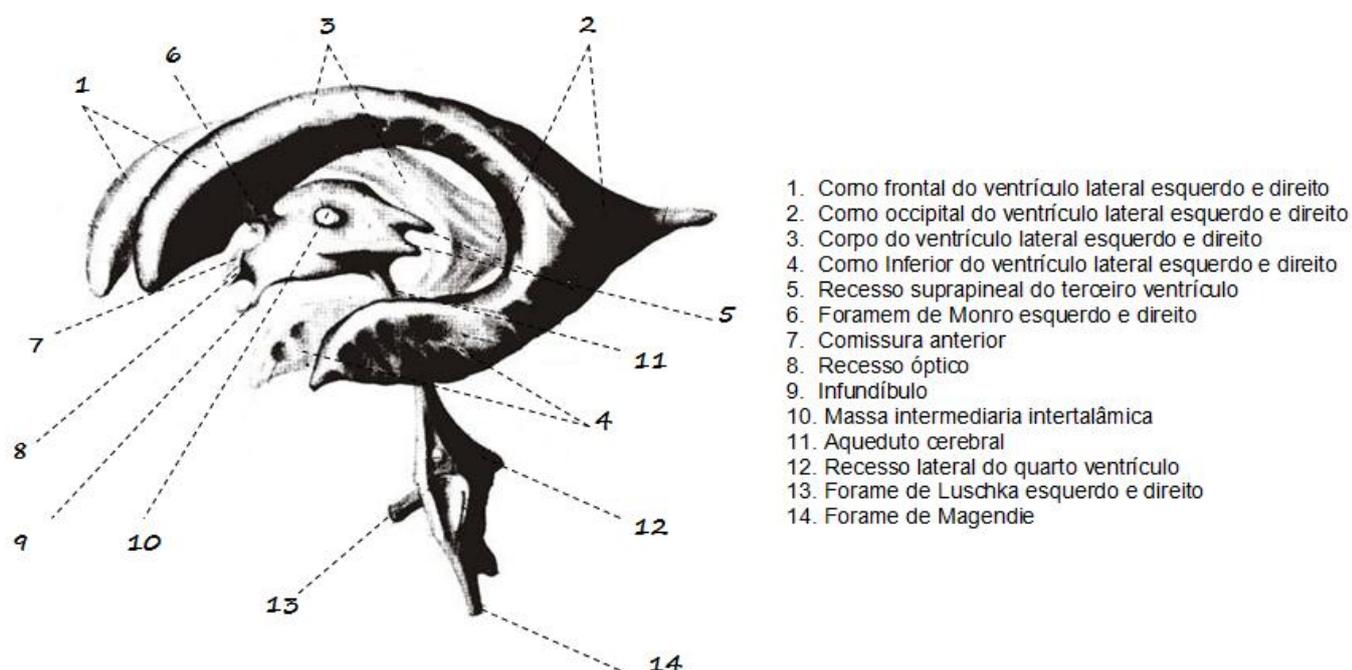


Fig. 3: O Sistema Ventricular Cerebral
 Fonte: adaptado de Maciel [42]

ventrículo através do aqueduto cerebral. Cortando o terceiro ventrículo temos a massa intermediária intertalâmica [43, 33].

O quarto ventrículo é composto pelo recesso lateral e os forames de Luschka e se comunica com o terceiro ventrículo através do aqueduto cerebral e com o espaço subaracnóide através do forame de Magendie [43, 33].

O objetivo deste trabalho é formalizar o sistema ventricular cerebral, devido a sua importância no diagnóstico da esquizofrenia. Como vimos nesta seção, o sistema ventricular cerebral não contém divisões concretas entre suas estruturas anatômicas, utilizamos então (para separar as divisões anatômicas do sistema ventricular cerebral) limites fiat, que apresentaremos na próxima seção.

2.5 Limites fiat

Os limites fiat são utilizados para delimitar regiões e objetos que tem seus limites aceitos de forma consensual, porém sem um limite concreto observável [44, 45].

Podemos exemplificar um limite fiat como a fronteira entre países, que muitas vezes não tem

limites concretos que definam onde termina o território de um e inicia o de outro. Esses limites são amplamente utilizados em sistemas de informação geográfica [8] e sistema biomédicos para delimitação de estruturas anatômicas [46].

Varzi [44] apresenta em seu trabalho definições para limites entre estruturas. Em síntese essas definições são baseadas em relações como parte tangencial ($TP(x,y)$, que significa que x é parte tangencial de y) e conexão ($C(x,y)$, que significa que x está conectado a y). O autor também aborda nesse trabalho a coincidência entre limites e outras propriedades inerentes dessas formalizações.

Em nosso trabalho usaremos limites fiat, dado que os limites entre as subdivisões anatômicas do sistema ventricular são limites aceitos consensualmente entre os especialistas e não possuem delimitações concretas quando analisamos isoladamente a entidade. A formalização dos limites fiat neste trabalho é baseada na BIT, que não conta com algumas relações utilizadas nos trabalhos de Varzi [44] e Smith e Varzi [45], por esse motivo, nos inspiramos nos trabalhos desses autores para prover uma conceitualização de limites fiat utilizando as relações da BIT.

No próximo capítulo apresentaremos as ferramentas utilizadas nesse trabalho para formalização da ontologia (descrevendo a BIT e apresentando o Protégé) e técnicas de análise e segmentação de imagem que usaremos para formalizar os conceitos apresentados nesse capítulo.

CAPÍTULO 3

MÉTODOS

Nesta seção apresentaremos os métodos e ferramentas que utilizamos direta ou indiretamente para geração dos resultados desse trabalho, passaremos pelos conceitos da visão computacional (segmentação de imagens médicas), análise discriminante estatística multivariada, testes de tamanho efetivo e também por ferramentas como o *statistical parametring mapping*(SPM) e o Protégé.

3.1 Técnicas e ferramentas de análise de imagens médicas.

A presença das imagens no diagnóstico médico se torna cada vez mais frequente. Com os avanços tecnológicos das últimas décadas, surgiram novas técnicas altamente eficientes de diagnóstico por imagens, que puderam prover informações sobre o corpo humano com um alto grau de confiabilidade e que possibilitam ao médico fazer um diagnóstico mais rápido e preciso [2, 3]. Entre essas técnicas podemos destacar a ressonância magnética, devido ao alto grau de detalhes que sua imagem possui.

Recentemente a segmentação de imagens, área da computação que estuda e desenvolve técnicas de extração de contornos em imagens, ganhou importância significativa no apoio às análises de imagem no diagnóstico médico, possibilitando destacar e observar mais criteriosamente objetos de interesse na imagem[3].

Nessa seção apresentaremos os seguintes conceitos:

1. Segmentação de imagens e o algoritmo de segmentação por crescimento de região, utilizado

neste trabalho para segmentação de regiões de interesse em neuroimagens;

2. O *Statistical Parametric Mapping* (SPM), utilizado para segmentação de neuroimagem em massa cinza, massa branca e líquido cefalorraquidiano;
3. Conceitos de técnicas de análise estatística multivariada e testes de tamanho efetivo (PCA+MLDA).
4. A teoria de inclusão básica (BIT), mereologia escolhida para criação da representação formal do sistema ventricular cerebral, apresentada no capítulo 2.4;
5. O Protégé, uma ferramenta de edição e manutenção de ontologias, escolhida para inclusão e testes da ontologia-epistemologia.

3.1.1 Segmentação de imagens

A análise computacional de imagens corresponde a programas capazes de extrair o objeto de interesse do fundo da imagem, o que possibilita mensurar e visualizar esse objeto [3]. Esse passo é chamado de segmentação de imagens. Na análise de imagens médicas a segmentação de imagens é amplamente usada para [3, 47]:

- Localização e diagnóstico de tumores;
- Cirurgia guiada por computador;
- Estudo da estrutura anatômica.

Existem diversas técnicas de segmentação de imagens, cada uma apresenta vantagens que dependem das características da imagem a ser segmentada e do resultado esperado da segmentação.

Segundo Pham *et al.* [48], dentre os métodos mais utilizados para segmentação de imagens médicas, podemos destacar a segmentação por limiar adaptativo, crescimento de região, *k-means*, segmentação por redes neurais *self-organized maps* (SOM) e os modelos deformáveis. Os resultados da segmentação de alguns desses métodos podem ser verificados na figura 4.

Na área médica, onde o conhecimento do especialista é de extrema importância para análise dos dados extraídos pela segmentação, é recomendado que sejam utilizadas técnicas que permitam maior

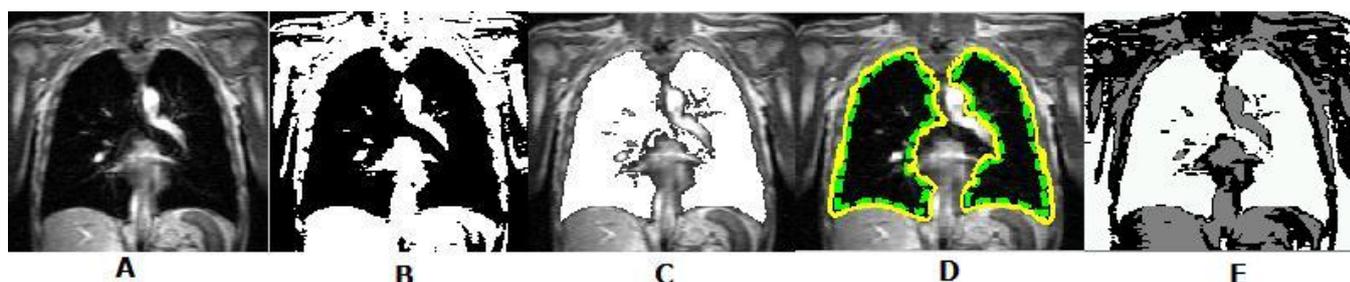


Fig. 4: Exemplo de segmentação de imagens criado pelo autor baseado na imagem modelo do Matlab. O original (A), segmentação por limiar adaptativo (B), segmentação por crescimento de região (C), modelos deformáveis (D) e segmentação k-means (E).

interação do usuário para obtenção de melhores resultados. Entretanto mesmo com a interação do especialista, essas técnicas devem ser precisas para reproduzir os objetos de interesse muitas vezes representados na imagem de forma diferente do esperado, devido à ruído, baixo contraste da imagem ou até mesmo características da patologia apresentada [3].

Olabarriaga [49] e Xiaohan e Yla-Jaaski [50] discutem a importância da interação do usuário no processo de segmentação de imagens médicas. Estes autores afirmam que a intervenção humana é de grande importância para avaliação dos resultados da segmentação, e, se necessário, efetuar a correção ou ajuste do resultado da segmentação manualmente.

Assim, indo ao encontro dos trabalhos de Olabariaga [49] e Xiaohan e Yla-Jaaski [50], propomos em nosso trabalho a utilização do algoritmo de crescimento de região com interação do usuário para escolha do ponto inicial de segmentação.

Com a segmentação do algoritmo de crescimento de região é possível destacar a região ventricular nas neuroimagens e descobrir a quantidade de voxels¹ que compõem essa região, permitindo calcular seu volume. Conhecido o volume ventricular de cada neuroimagem, podemos integrar esse dado a ontologia-epistemologia por meio de indivíduos que representam ou o domínio, ou um conhecimento sobre o domínio. Detalhes sobre o algoritmo de crescimento de região serão discutidas na próxima seção.

¹O voxel é um pixel volumétrico, enquanto o pixel tem duas dimensões (x e y) o voxel conta com três dimensões (x, y, z).

3.1.2 Segmentação por crescimento de região

O algoritmo de segmentação por crescimento de região parte do princípio que a região de interesse em uma imagem é composta por uma conexão de pixels ou voxels de valor próximo [51]. Sendo assim, um pixel/voxel escolhido em uma região de interesse pode representar todos os pixels que compõem toda a região.

O algoritmo de crescimento de região é inicializado com a interação do usuário, que escolhe um voxel dentro da região de interesse. A partir da escolha desse ponto (chamado de semente), o algoritmo inicia as seguintes etapas:

- Inicializa um contador do número de voxels que compõem a região ;
- Inicia a verificação dos voxels vizinhos de acordo com uma matriz de vizinhança pré-definida (um exemplo de expansão da matriz de vizinhança pode ser visto na figura 5);
- Caso o voxel visitado atenda ao critério de agregação:
 1. Incorpora o voxel visitado à região;
 2. Atualiza o contador do número de voxels que compõem a região;
 3. Atualiza a média da região.

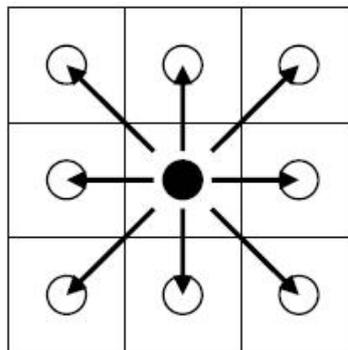


Fig. 5: Exemplo de expansão da matriz de vizinhança do algoritmo de crescimento de região.

Fonte: adaptado de Mat-Isa *et al.* [52]

Caso o voxel não atenda ao critério de agregação pré-estabelecido, o algoritmo para a verificação da vizinhança naquela direção e continua sua execução em outras direções, até que a região não possa mais crescer. O critério de agregação de voxels pode ser feito por:

- Limiar pré estabelecido pelo usuário;
- Gradiente da imagem;
- Estatística dos pixels/voxels agregados.

A escolha do melhor critério depende das características da imagem a ser segmentada. Por exemplo, se a região de interesse de uma imagem tem tons de cinza, cor ou textura com valores muito similares ao fundo da imagem, a escolha de um critério de agregação pelo limiar entre voxels pode não ser efetivo, nesse caso, o critério de agregação pelo gradiente da imagem poderia se mostrar mais eficiente.

O algoritmo usado para as segmentações de imagem neste trabalho utiliza como critério de agregação um limiar definido pelo usuário. Esse valor indica a similaridade que o valor de um voxel $v(x,y)$ tem em relação à média de voxels que compõem a região, como indicado na fórmula 3.1:

$$V \in \mathcal{R} \leftrightarrow [f(V) - Media] < T; \quad (3.1)$$

Sendo V representa o voxel visitado, $f(V)$ representa o valor do voxel visitado, $Media$ é a média da região e T um limiar definido pelo usuário.

Assim, o algoritmo executa o procedimento para todos os voxels da vizinhança, até encontrar voxels que subtraídos da média da região tenham valores diferentes do limiar definido pelo usuário, ou no momento em que todos os voxels da imagem foram analisados. A figura 6 mostra um exemplo de segmentação de uma imagem modelo utilizando o algoritmo de crescimento de região que utilizamos nesse trabalho aplicando diferentes valores de limiar.

Como podemos perceber no exemplo, quanto maior o limiar (diferença entre o pixel verificado e a média de pixels), mais pixels são agregados à região. Em contrapartida, com valores de limiar muito baixos, pixels que efetivamente sejam da região de interesse podem não ser agregados. O algoritmo 1 apresenta o pseudocódigo do algoritmo de crescimento de região *slice a slice*.

Neste trabalho, usaremos o algoritmo de crescimento de região para segmentar e conhecer a quantidade de voxels das regiões ventriculares observadas em cortes de IMR. Nas neuroimagens as regiões ventriculares são bem delimitadas e destacadas do fundo da imagem, o que torna o algoritmo de crescimento de região ideal para segmentação sem grandes perdas de informação.

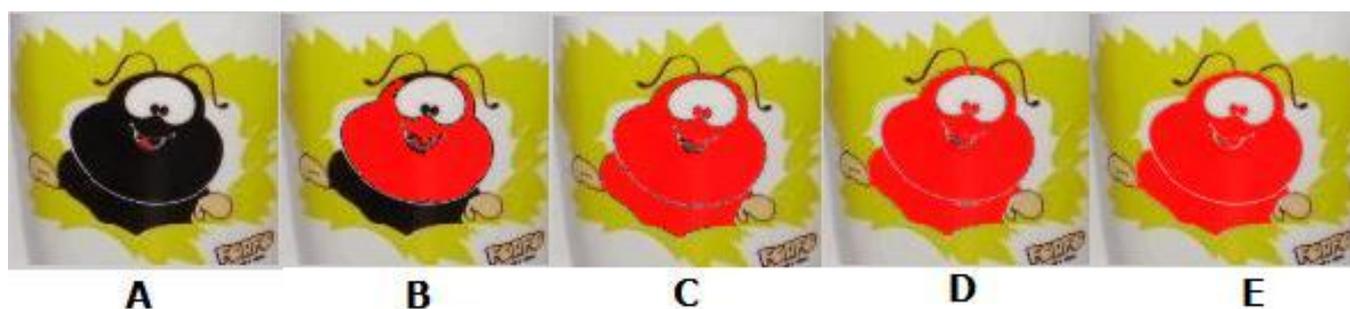


Fig. 6: Exemplo de imagem modelo segmentada pelo algoritmo de crescimento de região, a imagem (A) apresenta a imagem original, (B) segmentada com um limiar de 0.1, (C) com limiar de 0.2, (D) com limiar de 0.3 e (E) com limiar de 0.4.

Algorithm 1 Algoritmo de crescimento de região

```

região=semente
CresceReg(v(x,y,z))
  Início
  SE  $v(x,y,z) \notin \mathcal{R}$ 
    Início
     $CritAgreg = [f(V) - Media]$ 
    SE  $CritAgreg < Threshold$ 
      Início //adiciona pixel a região
      CresceReg(x,y-1,z)
      CresceReg(x,y+1,z)
      CresceReg(x-1,y,z)
      CresceReg(x+1,y,z)
    fim
  fim
fim
  fim
fim

```

Porém antes da segmentação por crescimento de região, a imagem passa por técnicas de pré-processamento utilizando o *statistical parametring mapping* (SPM), utilizado para normalização espacial e pré-segmentação das neuroimagens utilizadas neste trabalho, conforme abordaremos na próxima seção.

3.1.3 SPM - *Statistical Parametring Mapping*

O *Statistical Parametring Mapping*(SPM) é um pacote para Matlab² amplamente utilizado para pré-processamento de imagens médicas e análise de diferenças individuais entre um determinado voxel em duas ou mais imagens [53].

²O Matlab é um programa utilizado para cálculos matemáticos, que conta com diversos pacotes que permitem trabalhar com aplicações de visão computacional, engenharia, etc. Disponível em: www.mathworks.com

Em nosso trabalho utilizamos o SPM para segmentar neuroimagens separando a massa branca, massa cinzenta e líquido cefalorraquidiano, porém o SPM conta com outras técnicas de pré-processamento de imagens. Entre as funções disponíveis para tal estão a normalização, segmentação e suavização.

A normalização espacial consiste no alinhamento das imagens a serem estudadas tomando como referência uma imagem modelo composta por médias de imagens cerebrais de adultos jovens normais [54, 55] e aproximadas ao espaço descrito no atlas de Talairach e Tournoux [55]. Assim com base na imagem de referência, a normalização espacial corrige diferenças de escala, rotação, translação entre as amostras por meio de transformações apropriadas.

A etapa da normalização espacial garante que a estrutura anatômica a que pertence um voxel $v(x, y, z)$ na imagem modelo estará correlacionada com as amostras para identificação de padrões (que um voxel na imagem modelo corresponda ao mesmo voxel na imagem de interesse), assim pode-se evitar que diferenças de interesse no estudos realizados não sejam apresentadas por estarem sendo comparados voxels de regiões diferentes [56, 57].

Após as imagens normalizadas, o SPM também possibilita a segmentação da imagem em massa branca, massa cinzenta e líquido cefalorraquidiano. A figura 7 mostra um exemplo dessa segmentação. Nesta etapa também é efetuada uma equalização do *bias* (intensidade das imagens) do grupo de amostras e remoção de artefatos, como a caixa craniana [55, 57]³.

A etapa de suavização não é utilizada diretamente por esse trabalho, mas pode ser feita através da aplicação de um filtro gaussiano isotrópico na imagem para eliminação de ruídos. Sobre a variação do filtro aplicado, Mechelli [54] cita que diversos estudos de diferenças estatísticas entre grupos de neuroimagens tem utilizado filtros com 12-mm.

Nosso trabalho utiliza o SPM para pré-processamento de uma base de imagens em que aplicamos o algoritmo de crescimento de região, e também as imagens submetidas a análise estatística que utilizamos (provenientes do trabalho de Santos *et al.* [7]), passam pela normalização espacial do SPM.

Com o SPM também é possível realizar testes de tamanho efetivo, os trabalhos de Aguiar [57] e Oliveira [56] abordam com mais detalhes as metodologias de pré-processamento de neuroimagens (MBV-padrão e MBV-otimizado) e a análise estatística univariada utilizando o SPM, que não são

³A eliminação de artefatos também pode ser executada pelo software MRicro utilizando a função *skull strip* [58], disponível para download em: <http://www.sph.sc.edu/comd/rorden/mricro.html>

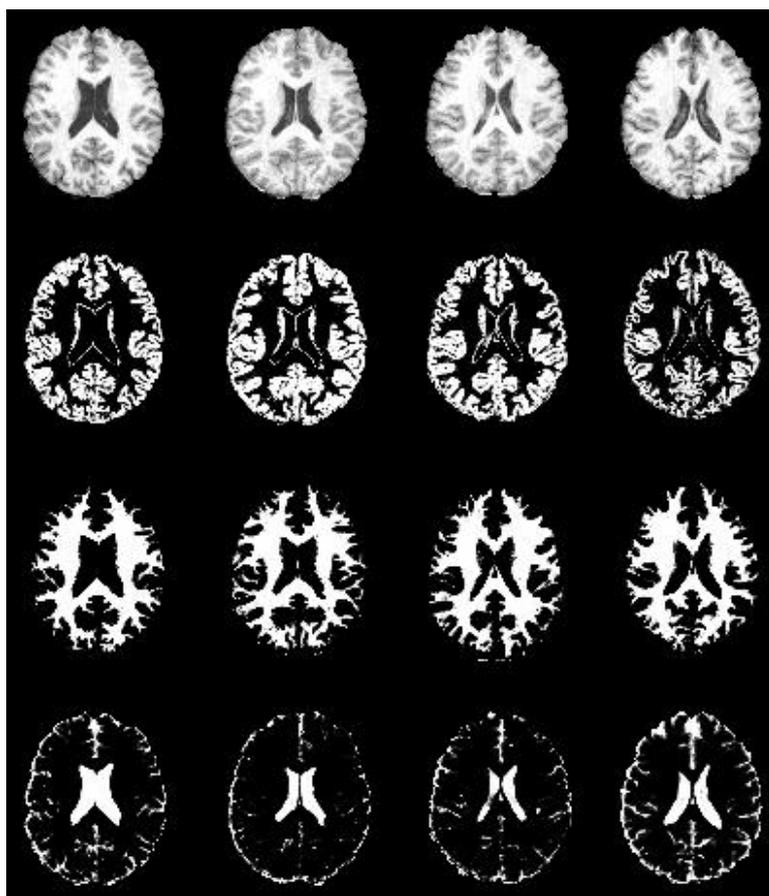


Fig. 7: Segmentação de imagens no SPM, na primeira fileira estão as imagens originais, na segunda fileira a segmentação por massa cinza, na terceira fileira a segmentação por massa branca, e na quarta fileira a segmentação pelo líquido cefalorraquidiano.

foco de discussão deste trabalho.

Na próxima seção apresentaremos o MLDA, técnica de análise estatística aplicada em algumas das neuroimagens utilizadas neste trabalho para descoberta de informação discriminante entre grupos de neuroimagens que podem ser inseridas na ontologia.

3.2 Estatística multivariada (PCA+MLDA) e testes de tamanho efetivo

Neste trabalho utilizamos o resultado de um grupo de neuroimagens submetidas às técnicas de análise estatística. O principal objetivo de submeter um grupo de imagens a essas análises é encontrar alterações relevantes entre esses grupos.

Por meio das análises estatísticas é possível encontrar alterações morfológicas relevantes que podem ser inseridas na ontologia juntamente com o conhecimento presente na literatura sobre essas alterações, ou até mesmo descobrir alterações não perceptíveis nesses grupos. Utilizamos para esta tarefa, o modelo estatístico discriminante (PCA+MLDA) e testes de tamanho efetivo.

A técnica *maximum uncertainty linear discriminant analysis* (MLDA)[59], é uma técnica de análise estatística multivariada amplamente utilizada para discriminação de diferenças entre grupos de imagens.

O MLDA se baseia na técnica de análise discriminante linear (LDA) [60], porém é mais efetivo para análises onde o número de variáveis é maior que o número de amostras devido ao problema conhecido por número pequeno de amostras ⁴. Nas imagens utilizadas neste trabalho temos um número total de 902629 variáveis (voxels) a serem comparadas em relação a 25 imagens do grupo de controle e 43 imagens do grupo de pacientes, totalizando 68 amostras, características que tornam o MLDA ideal para utilização na análise.

O MLDA pode ser aplicado diretamente no conjunto de amostras, porém ainda assim podemos ter uma análise custosa computacionalmente [61]. Para resolver esse possível problema de recursos computacionais para o processamento do MLDA, aplica-se a análise de componentes principais ou *principal component analysis* (PCA) em conjunto com o MLDA [61, 57, 56]. A junção dessas duas técnicas é conhecida por modelo estatístico discriminante (SDM)⁵.

O PCA permite a redução dos dados observados [62]. Basicamente podemos representar um conjunto de amostras usando apenas as características mais importantes desse grupo e descartando as demais, reduzindo assim a dimensionalidade dos dados observados. Essa tarefa é executada descartando as variâncias menos significativas de cada amostra [56, 61].

A análise de componentes principais é aplicada na matriz que contém nosso conjunto de amostras (neuroimagens) para redução de sua dimensão, preservando os valores de variância mais significativos de cada amostra e descartando as variâncias nulas, diminuindo a dimensionalidade dos dados à serem analisados pelo MLDA. Assim podemos aplicar o PCA em conjunto com o MLDA para análises consistentes sem perda significativa de características e custo computacional reduzido [61, 57, 56].

Os detalhes matemáticos dos cálculos do PCA são abordados com mais detalhes nos trabalhos

⁴O problema do número pequeno de amostras (SSS) ou *small sample size* é um problema matemático relacionado a inversão da matriz de covariância S_w no LDA, mais detalhes sobre esse problema são apresentados nos trabalhos de Thomaz *et al.* [59], Kitani [61] e Oliveira [56]

⁵Em inglês *statistical discriminant model*

de Fukunaga [63] e Kitani e Thomaz [64] e o MLDA é abordado no trabalho de Thomaz *et al.* [59], Kitani [61], Aguiar [57], Oliveira [56].

Como resultado da aplicação do modelo estatístico multivariado nas imagens que compõem os grupos de controle e paciente, obtemos a característica mais discriminantes de cada imagem projetada no espaço do MLDA. Pode-se então descobrir o hiperplano de separação entre os grupos, médias locais de cada grupo e o extremo de cada grupo dado por três vezes o desvio padrão ($\pm 3\sqrt{\sigma^2}$) [7].

Podemos também reprojetar esses pontos em seu espaço original reconstruindo a imagem, assim as características do extremo de cada grupo (que representa as imagens mais distantes entre si) podem ser comparadas. Para descobrirmos quais desses pontos apresentam alterações relevantes, essas imagens são submetidas a testes de tamanho efetivo. Utilizando esse tipo de análise, podemos efetuar a comparação entre voxels dos modelos de imagem que representam os extremos dos grupos de paciente e controle através da equação 3.2:

$$diff = \frac{C - P}{\sqrt{\frac{\sigma^2}{n_c} + \frac{\sigma^2}{n_p}}} \quad (3.2)$$

Sendo C o modelo imagem reconstruído a partir do ponto extremo do grupo de controle e P o modelo de imagem reconstruído a partir do grupo de pacientes, o termo σ representa a variância ponderada de cada grupo projetado no espaço do MLDA e os termos n_c o número de indivíduos que compõem os grupos de controle e n_p o número de indivíduos que compõem os grupos de pacientes [7]. Assim obtemos as diferenças significativas⁶, como podemos observar na figura 8.

Na próxima seção abordaremos a teoria de inclusão básica, e apresentaremos uma discussão sobre sua escolha para a formalização do sistema ventricular cerebral.

3.3 Teoria de inclusão básica - BIT

A teoria de inclusão básica (BIT)⁷ é uma teoria formal proposta por Donnelly *et al.* [10], concebida para prover bases para o raciocínio espacial qualitativo na área biomédica. Esta seção apresenta as relações que compõem a BIT como podemos observar na figura 9).

⁶Para as imagens utilizadas nesse trabalho foi levado em consideração diferenças positivas e negativas com $|diff| > 3$ [7]

⁷Sigla do inglês *basic inclusion theory*

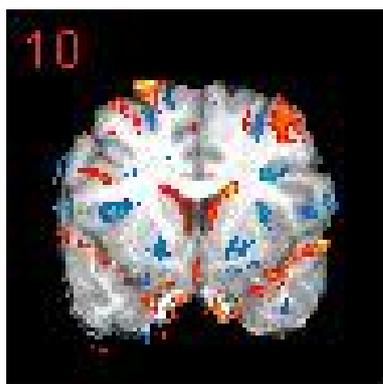


Fig. 8: Pontos que representam alterações estatísticas relevantes evidenciadas pelos testes de tamanho efetivo entre os extremos do grupo de paciente e controle encontrados pelo MLDA.

Fonte: adaptado de Santos *et al.* [7]

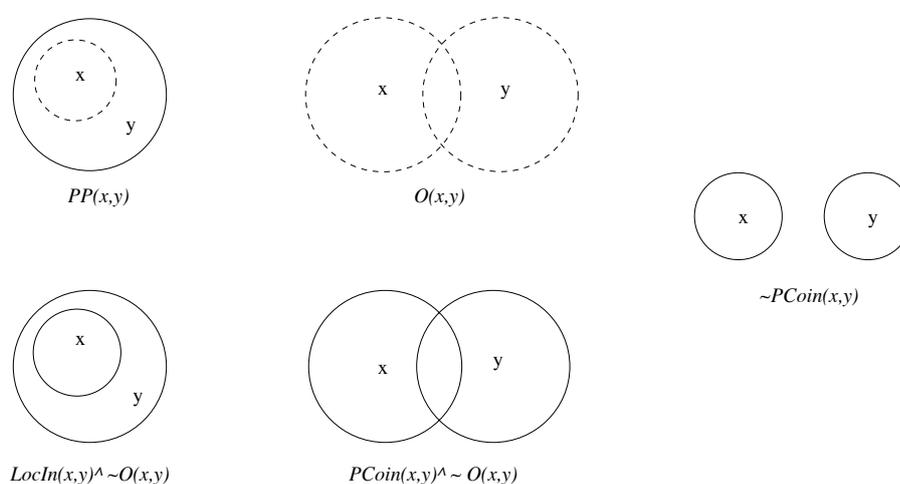


Fig. 9: Relações mereológicas da BIT.

Fonte: adaptado de Donnelly *et al.* [10]

Relação Parte

A relação parte (P)⁸ entre dois indivíduos x e y , indica que x é parte de y . A relação parte é uma relação primitiva em muitas mereologias e essencial para definir outras relações como “parte própria” (PP) e “sobreposição” (O) [10]. Representamos a relação parte conforme demonstrado na fórmula 3.3:

$$P(x,y) \quad (3.3)$$

A relação parte é reflexiva (fórmula 3.4), ou seja, todo objeto é parte dele mesmo, antiasimétrica

⁸do inglês *Part*

(fórmula 3.5) que nos diz que se x é parte de y e y é parte de x , então x e y são idênticos, e transitiva (fórmula 3.6) que representa que se x é parte y , e y é parte de z , então x é parte de z .

$$P(x,x) \tag{3.4}$$

$$P(x,y) \wedge P(y,x) \rightarrow (x = y) \tag{3.5}$$

$$P(x,y) \wedge P(y,z) \rightarrow P(x,z) \tag{3.6}$$

Relação r - Regiões espaciais

Como apresentamos no capítulo 2.2 existe uma importante diferenciação entre o objeto e a região espacial que ele ocupa. Utilizando BIT podemos representar a região espacial de um objeto por meio da relação r , que representa a região espacial de um objeto [10].

$$r(r(x)) \equiv r(x) \tag{3.7}$$

$$P(x,y) \rightarrow (P(r(x),r(y))) \tag{3.8}$$

A fórmula 3.7 representa que a região espacial de x é sua própria região espacial, e a fórmula 3.8 nos diz que se x é parte de y , então a região espacial de x é parte da região espacial de y .

Relação Parte Própria

A relação parte própria (PP)⁹ entre dois indivíduos x e y significa que x é parte própria de y , se x for qualquer parte de y , mas que não represente todo y . A relação de parte própria e representada pela fórmula 3.9

$$PP(x,y) \equiv P(x,y) \wedge x \neq y \tag{3.9}$$

A relação de parte própria é transitiva, então se x é parte própria de y e y é parte própria de z então x é parte própria de z (fórmula 3.10). A fórmula 3.11 significa que se x é parte própria de y então y não é parte própria de x o que caracteriza assimetria. A relação de Parte própria também é irreflexiva,

⁹Do inglês *Proper Part*

em outras palavras, nenhuma região é sua parte própria (fórmula 3.12)

$$PP(x,y) \wedge PP(y,z) \rightarrow PP(x,z) \quad (3.10)$$

$$PP(x,y) \rightarrow \neg PP(y,x) \quad (3.11)$$

$$\neg PP(x,x) \quad (3.12)$$

Relação Sobreposto

A relação de sobreposição (O)¹⁰ entre dois indivíduos x e y , significa que x e y sobrepõem-se se existe uma região z que faz parte de x e y , simbolicamente temos:

$$O(x,y) \equiv \exists z(P(z,x) \wedge P(z,y)) \quad (3.13)$$

A relação sobreposto é simétrica: se x é sobreposto a y então y é sobreposto a x (fórmula 3.14); e reflexiva (fórmula 3.15) que significa que toda região se sobrepõe a si mesma.

$$O(x,y) \rightarrow O(y,x) \quad (3.14)$$

$$O(x,x) \quad (3.15)$$

Ainda podem ser derivadas as seguintes asserções para as relações de parte e parte própria:

$$PP(x,y) \rightarrow O(x,y) \quad (3.16)$$

$$O(x,y) \wedge P(y,z) \rightarrow O(x,z) \quad (3.17)$$

A fórmula 3.16 significa que se x é parte própria de y , então x sobrepõe y e a fórmula 3.17 nos diz que se x sobrepõe y e y é parte de z então x sobrepõe z .

¹⁰Do inglês *Overlap*

Relação localizado em

A relação “Localizado Em” (*LocIn*)¹¹ entre dois indivíduos x e y , representa que: x está localizado em y se a região de ocupação de x é parte da região y (fórmula 3.18)

$$LocIn(x,y) \equiv P(r(x),r(y)) \quad (3.18)$$

A fórmula 3.19 significa que todo indivíduo está localizado nele mesmo, garantindo assim a reflexividade da relação *LocIn*, enquanto a fórmula 3.20, significa que se x está localizado em y e y está localizado em z então x está localizado em z (transitividade).

$$LocIn(x,x) \quad (3.19)$$

$$LocIn(x,y) \wedge LocIn(y,z) \rightarrow LocIn(x,z) \quad (3.20)$$

Também são derivadas as seguintes asserções para as relação de parte e parte própria:

$$P(x,y) \rightarrow LocIn(x,y) \quad (3.21)$$

$$PP(x,y) \rightarrow LocIn(x,y) \quad (3.22)$$

$$LocIn(x,y) \wedge PP(y,z) \rightarrow LocIn(x,z) \quad (3.23)$$

$$PP(x,y) \wedge LocIn(y,z) \rightarrow LocIn(x,z) \quad (3.24)$$

As fórmulas 3.21 e 3.22 significam que se x é parte ou parte própria (respectivamente) de y , então x está localizado em y . A fórmula 3.23 nos diz que, se x é parte de y e y está localizado em z , então x está localizado em z , enquanto de forma similar, a fórmula 3.24 significa que se x é parte própria de y e y está localizado em z , então x está localizado em z .

Relação coincidência parcial

A relação *PCoin*¹² entre dois indivíduos x e y , representa que x e y coincidem parcialmente se a região espacial de x e a região espacial de y se sobrepõem. No caso da coincidência parcial as regiões

¹¹do inglês *Located In*

¹²do inglês *partially coincide*

físicas não têm uma parte em comum sobrepostas como na relação de sobreposição (O), apenas suas regiões espaciais (r) se sobrepõem. Formalmente:

$$PCoin(x,y) \equiv O(r(x),r(y)) \quad (3.25)$$

A relação $PCoin$ é reflexiva, assim todo indivíduo coincide parcialmente com si mesmo (fórmula 3.26). A fórmula 3.27 indica que a relação $PCoin$ é simétrica, assim se x coincide parcialmente com y então y coincide parcialmente com x .

$$PCoin(x,x) \quad (3.26)$$

$$PCoin(x,y) \rightarrow PCoin(y,x) \quad (3.27)$$

E são derivadas as asserções abaixo para as relações sobreposição e localizado em :

$$O(x,y) \rightarrow PCoin(x,y) \quad (3.28)$$

$$LocIn(x,y) \rightarrow PCoin(x,y) \quad (3.29)$$

A relação 3.28 significa que se x e y sobrepõe-se, então x e y se coincidem parcialmente, já a fórmula 3.29 nos diz que, se x está localizado em y , então x coincide parcialmente com y .

3.3.1 Relações Inversas

Dada uma relação binária S , sua relação inversa S^{-1} é definida como: $S^{-1}(x,y) \leftrightarrow S(y,x)$. Por exemplo, a relação de parte própria entre dois elementos x e y (x é parte própria de y), sua relação inversa só é possível caso y seja parte própria de x , como apresentamos na fórmula 3.30.

$$PP^{-1}(x,y) \leftrightarrow PP(y,x) \quad (3.30)$$

Por exemplo, a asserção $PP^{-1}(coracao,ventriculo_esquerdo)$ indica que o coração tem o ventrículo esquerdo como uma de suas partes próprias [10].

No caso de relações simétricas, sua inversa corresponde a ela mesma. Por exemplo a relação de

sobreposição 3.13 onde temos que $O(x,y) \rightarrow O(y,x)$, então, $O = O^{-1}$.

3.3.2 Axiomas Adicionais

Os axiomas do BIT construídos em lógica de primeira ordem permitem que sejam derivados um número infinito de axiomas suplementares que podem enriquecer a linguagem e tornar a mereologia mais expressiva [65]. Apresentaremos abaixo alguns axiomas propostos por Donnelly *et al.* [10] para essa finalidade.

3.3.3 Princípio da Suplementação Fraca - *Weak Supplementation Principle*

O princípio da suplementação fraca é proposto em mereologias [13, 10] para desambiguar situações oriundas da relação parte própria. Basicamente o princípio da suplementação fraca garante que se um indivíduo x é parte própria de y , e não compromete todo y , assim sempre existirá uma região z que faz parte da composição de x além de y , formalmente:

$$PP(x,y) \rightarrow \exists z[PP(z,y) \wedge \neg O(z,x)] \quad (3.31)$$

Donnelly *et al.*[10], baseado no princípio da suplementação fraca, deriva o axioma que restringe um indivíduo a não ter apenas uma parte própria (fórmula 3.32) :

$$PP(x,y) \rightarrow \exists z[PP(z,y) \wedge (z \neq x)] \quad (3.32)$$

Logo, se um indivíduo x possui uma parte própria, existe pelo menos mais uma região z que é parte própria de x , e todas as regiões que são partes próprias de x , se somadas, equivalem ao tamanho total de x e não se sobrepõem [13, 10, 66]. Formalmente essa dedução mereológica pode ser expressa em termos: Se x é parte de y , então y tem partes próprias $x_1, x_2, x_3, \dots, x_n$ tal que nenhum x_n se sobreponham e a soma de $x_1, x_2, x_3, \dots, x_n$ é igual a y .

Assim, é derivada a fórmula 3.33 que garante que todo indivíduo tenha uma parte própria. Por exemplo, a formação celular dos órgãos, dado que todo órgão é formado por milhares de células. Assim, toda a célula é parte própria de um órgão, e a soma total de todas as células que compõem o órgão equivale ao órgão, e elas não se sobrepõem [10].

$$\forall y \exists x PP(x, y) \quad (3.33)$$

3.3.4 Teoria de Classes

Assim como as principais ontologias biomédicas previamente apresentadas nesse trabalho, a BIT implementa uma teoria taxonômica para prover uma classificação geral dos indivíduos de forma hierárquica, essa extensão da mereologia do BIT é chamada de *BIT+Cl*[10].

A importância de instanciar os indivíduos e criar árvores taxonômicas está na generalidade das estruturas. Quando dizemos: “*O hemisfério direito é parte própria do cérebro*”, estamos nos referindo à partonomia entre **um** hemisfério direito específico e **um** cérebro específico, mas também podemos dizer que, **todo** o cérebro possui um hemisfério direito.

Outro exemplo possível e mais genérico seria admitir uma classe “*Sistema Nervoso*”, onde suas especializações seriam subclasses com o nome das estruturas anatômicas que compõem o sistema nervoso, ou ainda uma classe “*Orgão*”, onde cada órgão individualmente seria uma subclasse. O BIT permite a criação de classes a partir da relação binária *Inst* (fórmula 3.34 que significa que *x* é um indivíduo da classe *A*):¹³.

$$Inst(x, A) \quad (3.34)$$

Temos como propriedades da relação de instanciação:

$$\exists x Inst(x, A) \quad (3.35)$$

$$\exists A Inst(x, A) \quad (3.36)$$

A fórmula 3.35 garante que toda classe tenha algum membro e a fórmula 3.36 nos diz que todo indivíduo é membro de uma classe.

Podemos citar como exemplo a fórmula 3.37, que instância um coração do indivíduo hipoteticamente chamado de Rodolpho (*coracao_{Rodolpho}*), à classe que representa o órgão coração (*CORACAO*), genericamente:

¹³Classes são representadas neste trabalho por letras maiúsculas

$$Inst(coracao_{Rodolpho}, CORACAO). \quad (3.37)$$

A partir da instanciação de indivíduos em classes pela relação *Inst*, Donnelly *et al.* [10] propõem a relação de englobamento *Is_a*, dada pela fórmula 3.38, que representa que a classe *A* é subclasse de *B* se para todo indivíduo *x* que for instância de *A*, *x* também for uma instância de *B*, ou seja:

$$Is_a(A, B) \equiv \forall x(Inst(x, A) \rightarrow Inst(x, B)) \quad (3.38)$$

A relação *Is_a* pode ser exemplificada com base no exemplo anterior (fórmula 3.37), onde instanciamos um indivíduo à classe *CORACAO*, e no exemplo , instanciamos a classe *CORACAO* como subclasse da classe *ORGAO*, assim temos que todo coração é um órgão.

$$Is_a(CORACAO, ORGAO) \quad (3.39)$$

3.3.5 Classes Discretas

Segundo Donnelly *et al.*[10] consideramos classes discretas quando diferentes indivíduos pertencentes a mesma classe não se sobrepõem. Formalmente representado pela fórmula 3.40

$$Discrete(A) \equiv \forall x \forall y (Inst(x, A) \wedge Inst(y, A) \wedge x \neq y) \rightarrow \neq O(x, y) \quad (3.40)$$

Como exemplo de classes discretas podemos citar o coração, pois dois corações distintos pertencem a mesma classe mas não se sobrepõem. Geralmente classes não-discretas são classes genéricas que possuem diversas subclasses, por exemplo, o sistema respiratório é sobreposto ao sistema de órgãos[10].

Caso uma classe *A* seja discreta, qualquer englobamento de *A* também será discreto [10], formalmente:

$$Discrete(B) \wedge Is_a(A, B) \rightarrow Discrete(A) \quad (3.41)$$

Assim, dadas as relações de instanciação (3.34) e englobamento (3.38) e suas definições é proposto a instanciação das relações espaciais que compõem a BIT para representação dos conceitos

taxonômicos, como veremos na próxima seção [10].

3.3.6 Relações mereológicas na teoria de classes

Dada a relação binária R (qualquer relação mereológica do BIT), utilizada em conjunto com as relações de instanciação, poderemos representar conceitos genéricos em relação a classes e não somente a indivíduos. Donnelly *et al.* [10] definem três formas de representação da relação R em relação a teoria de classes (R_1, R_2, R_{12}).

A fórmula 3.42 restringe a relação ao primeiro argumento (A), ou seja, todo indivíduo instanciado à classe A tem um indivíduo instanciado à classe B . Por exemplo, a fórmula 3.43 significa que todo aparelho reprodutor feminino é parte própria de um corpo humano, mas nem todo corpo humano possui um aparelho reprodutor feminino como parte própria, ou seja:

$$R_1(A, B) \equiv \forall x(Inst(x, A) \rightarrow \exists y[Inst(y, B) \wedge R(x, y)]) \quad (3.42)$$

$$PP_1(APARELHO_REPRODUTOR_FEMINIMO, CORPO_HUMANO) \quad (3.43)$$

A fórmula 3.44 restringe a relação ao segundo argumento (B), ou seja, todo indivíduo instanciado à classe B tem um indivíduo instanciado à classe A . Por exemplo, a fórmula 3.45 nos diz que todo coração tem células como parte própria, mas nem todas as células são parte própria do coração, isto é:

$$R_2(A, B) \equiv \forall y(Inst(y, B) \rightarrow \exists x(Inst(x, A) \wedge R(x, y))) \quad (3.44)$$

$$PP_2(CELULA, CORACAO) \quad (3.45)$$

A fórmula 3.46 restringe a relação aos dois argumentos, em outras palavras, toda instância de A tem uma instância de B como parte própria, e toda instância de B tem uma instância de A como parte própria. Por exemplo, a fórmula 3.47 indica que todo ventrículo lateral é parte própria do sistema ventricular e todo sistema ventricular tem um ventrículo lateral como parte própria.

$$R_{12}(A, B) \equiv R_1(A, B) \wedge R_2(A, B) \quad (3.46)$$

$$PP_{12}(\text{Ventriculo_Lateral}, \text{SistemaVentricular}) \quad (3.47)$$

Segundo Donnely *et al.*[10], as propriedades das relações mereológicas como a transitividade, assimetria e irreflexividade, não necessariamente se mantêm nas relações sobre classes. A seguir destacaremos as propriedades lógicas das relações instanciadas.

Transitividade

Dada uma relação transitiva R (qualquer relação transitiva da BIT), seus equivalentes R_1 , R_2 e R_{12} também são transitivos. Na BIT as relações transitivas são: parte (P) (fórmula 3.48), parte própria (PP) (fórmula 3.49) e localizado em ($LocIn$) (fórmula 3.50):

$$P_i(A, B) \wedge P_i(B, C) \rightarrow P_i(A, C), i = 1, 2, 12 \quad (3.48)$$

$$PP_i(A, B) \wedge PP_i(B, C) \rightarrow PP_i(A, C), i = 1, 2, 12 \quad (3.49)$$

$$LocIn_i(A, B) \wedge LocIn_i(B, C) \rightarrow LocIn_i(A, C), i = 1, 2, 12 \quad (3.50)$$

Reflexividade

Se R é uma relação reflexiva, seus equivalentes R_1 , R_2 e R_{12} também são reflexivos. Na BIT a relação parte (fórmula 3.51), sobreposição (fórmula 3.52), localizado em (fórmula 3.53) e coincidência parcial (fórmula 3.54), são relações reflexivas.

$$P_i(A, A), i = 1, 2, 12 \quad (3.51)$$

$$O_i(A, A), i = 1, 2, 12 \quad (3.52)$$

$$LocIn_i(A, A), i = 1, 2, 12 \quad (3.53)$$

$$PCoin_i(A, A), i = 1, 2, 12 \quad (3.54)$$

Simetria

Seja R uma relação simétrica, seus equivalentes R_1 , R_2 não são simétricos e seu equivalente R_{12}

é simétrico. Na BIT as relações simétricas são O e $PCoin$.

Assim, para a propriedade de simetria não podemos afirmar que $R_1(A, B) \rightarrow R_1(B, A)$. Por exemplo, a asserção $O_1(Mao, Nervo)$ (toda mão sobrepõe algum nervo), não tem um correspondente simétrico em $O_1(Nervo, Mao)$ (todo nervo sobrepõe uma mão), mas tem sua equivalência se restringirmos o argumento ao segundo elemento, como observamos na relação 3.55:

$$R_1(A, B) \leftrightarrow R_2(B, A) \quad (3.55)$$

Assim a relação $O_1(MAO, NERVO)$ (toda mão sobrepõem algum nervo) teria como equivalente simétrica a relação $O_2(NERVO, MAO)$ (toda mão sobrepõem algum nervo, mas nem todo nervo é sobreposto por uma mão).

Nem todas as propriedades lógicas das relações são transferidas para suas equivalentes na teoria de classes. A tabela 2 apresenta as propriedades de uma relação R e suas equivalentes (R_1, R_2 e R_{12}) e quais propriedades são transferidas da relação sobre indivíduos para a relação sobre classes.

Propriedades lógicas de R	Sobre Classes			
	Sobre indivíduos	R_1	R_2	R_{12}
Reflexivo	+	+	+	+
Irreflexivo	-	-	-	-
Simétrico	-	-	+	+
Asimétrico	-	-	-	-
Antiasimétrico	-	-	-	-
Transitivo	+	+	+	+

Tab. 2: Tabela que indica quais propriedades lógicas sobre relações de classe tem herança da relação equivalente sobre indivíduos, e quais não herdam propriedades. Herdar a característica é representado por (+) e não herdar é representado por (-).

Fonte: adaptado de Donelly *et al.* [10]

3.3.7 Implicações utilizando a teoria de classes

As implicações simples das relações sobre indivíduos da BIT, como por exemplo, a fórmula $PP(x, y) \rightarrow LocIn(x, y)$ (fórmula 3.22), tem equivalência em todas as relações derivadas pela teoria de classes ao utilizarmos BIT+Cl. Por exemplo, a fórmula 3.56 significa que a implicação: a classe A é parte própria da classe B , então A está localizado em B , tem validade para todas as variações da

teoria de classes (1,2,12).

$$PP_i(A,B) \rightarrow LocIn_i(A,B), i = 1, 2, 12 \quad (3.56)$$

Entretanto, segundo Donnelly *et al.* [10], algumas implicações que envolvem diferentes tipos de relações de classes podem não ser dedutíveis no geral.

Por exemplo, a fórmula 3.59 significa que toda próstata está localizada em alguma cavidade pélvica, e a fórmula 3.60 que toda pélvis feminina é parte própria de uma cavidade pélvica.

$$LocIn_1(PROSTATA,CAVIDADE_PELVICA) \quad (3.57)$$

$$PP_2(CAVIDADE_PELVICA,PELVIS_FEMININA) \quad (3.58)$$

Baseado nessas asserções, não podemos inferir que toda próstata esta localizada em uma pélvis feminina (fórmula 3.59), tampouco que toda pélvis feminina tem alguma próstata localizada nela (fórmula 3.60).

$$LocIn_1(PROSTATA,PELVIS_FEMININA) \quad (3.59)$$

$$LocIn_2(PROSTATA,PELVIS_FEMININA) \quad (3.60)$$

A BIT foi a ontologia escolhida nesse trabalho para a representação do sistema ventricular cerebral, a sua escolha se baseou em algumas características em relação a outras ontologias apresentadas brevemente no capítulo 2.1.2, que discutiremos na próxima seção.

3.3.8 Considerações sobre a escolha da BIT

Nesta seção apresentamos a *Basic Inclusion Theory* (BIT), uma teoria formal concebida para a utilização no domínio biomédico. Entretanto existem várias maneiras de construir ontologias por meio de outras teorias formais, como as já citadas no capítulo 2.1.2: o modelo de anatomia fundacional (FMA) [20], GALEN [21, 22], a biblioteca aberta de ontologias biomédicas relacional (OBO) [23] e a ontologia formal básica (BFO) [24].

Donnelly *et al.* [10] define a mereologia como sendo a base para o raciocínio espacial qualitativo na área biomédica por permitir inferências mais complexas sobre o posicionamento de entidades, bem como uma formalização completa e livre de ambiguidades.

O FMA [20] e o GALEN [21, 22] assim como a BIT possuem relações mereológicas, entretanto a sua formalização em lógica descritiva dá margens a ambiguidade no processo de raciocínio espacial [10, 67], onde podemos citar por exemplo a relação *part_of* aplicadas em diferentes contextos do conhecimento anatômico, como veremos futuramente nessa seção.

O OBO [23] é composto por diversas ontologias construídas para domínios específicos, e conta com uma teoria composta por relações mereológicas [23]. Entretanto, assim como o FMA e o GALEN, as relações mereológicas propostas pelo OBO podem levar, assim como o FMA a alguma ambiguidade. Um exemplo sobre possíveis ambiguidades utilizando relações de partonomia será descrito a seguir.

As relações de parte-todo no FMA são representadas pela relação *part_of*, assim para representar a sentença: “A pélvis feminino é parte de um corpo”, utilizaríamos: *Pelvis_Feminino part_of corpo*. Utilizando o GALEN a relação apropriada seria *isDivisionOf*, enquanto, o OBO utilizaria a relação *part_of*. Sabemos porém com base no conhecimento da anatomia que, nem todo corpo tem como parte própria um pélvis feminino. A BIT permite que representemos esse fato por meio da relação PP_1 .

Outra sentença sobre anatomia como: ventrículo esquerdo é parte do coração, também seria representado pelas relações *part_of* no FMA e no OBO e *isDivisionOf* [10, 23], enquanto na BIT usamos a relação PP_{12} .

A BFO [24] é criada em lógica de primeira ordem e composta por relações mereológicas e conta com relações parte-todo, temporais e topológicas. Podemos citar a relação de parte e parte própria, sobreposição, definição de limites de entidades, e localização temporal. A BFO é, em termos de relações para indivíduos, mais completa que a BIT, entretanto, a BFO não possui claramente formalizadas relações que permitam asserções sobre classes, como por exemplo as relações $R_{1,2,12}$ da BIT.

Neste trabalho optamos pela BIT para construir nossa representação formal pela sua construção em lógica de primeira ordem, suas relações mereológicas e taxonômicas claramente construídas ¹⁴, a

¹⁴A mereologia e taxonomia, juntamente com a topologia são consideradas extremamente relevantes para a construção de uma ontologia consistente como vimos no capítulo 2.1.1

teoria de classes e capacidade permitir inferências sobre classes com base em relações espaciais.

Uma desvantagem da BIT é o fato de não possuir relações que definam limites e continuidade entre estruturas anatômicas (como a relação *BoundaryOf(x,y)* da BFO), pois entendemos que o limite de cada estrutura anatômica é de importância fundamental para representação da entidade de interesse, já que o conhecimento anatômico é baseado em divisões, pontos e superfícies [21]. Para podermos formalizar os limites e continuidade entre estruturas, propomos uma extensão da BIT, que será apresentada no capítulo 4.

Na próxima seção apresentaremos o Protégé, ferramenta utilizada para apresentação computacional da teoria formal criada nesse trabalho.

3.4 O Protégé

O Protégé é uma ferramenta integrada para construção de ontologias e sistemas baseados em conhecimento desenvolvido pela *Stanford Medical Informatics - SMI* da Universidade de Stanford ¹⁵.

Usando o Protégé podemos editar e criar ontologias por meio da *Ontology Web Language (OWL)*. A OWL foi proposta pelo W3C e desenvolvida para ser utilizada para Web Semântica, é uma linguagem capaz de representar classes, instâncias, indivíduos e suas propriedades, além de derivar consequências lógicas não explícitas da interação entre os componentes.

A complexidade do desenvolvimento em OWL pode ser consideravelmente diminuída através do desenvolvimento por meio de sua interface gráfica (figura 10), que permite a criação de um conjunto de classes e suas propriedades, gerando o código OWL em *background*.

Escolhemos o Protégé para esse trabalho, pois ele tem sido o editor de ontologias amplamente utilizado em projetos acadêmicos, governamentais e em corporações devido às suas características e funcionalidades, que incluem a existência de diversos *plugins* que agregam funcionalidades e permitem o desenvolvimento de ontologias para domínios específicos. Além dessas características, o Protégé é uma ferramenta intuitiva de fácil uso que permite demonstrar os conceitos apresentados nesse trabalho.

Além de sua ampla utilização, o Protégé conta com documentação e suporte da comunidade, e mesmo sendo independente de domínio (permite a construção de diversos tipos de ontologia), o

¹⁵Disponível para download em: <http://protege.stanford.edu/download/registered.html>, atualmente na versão 4.0 beta

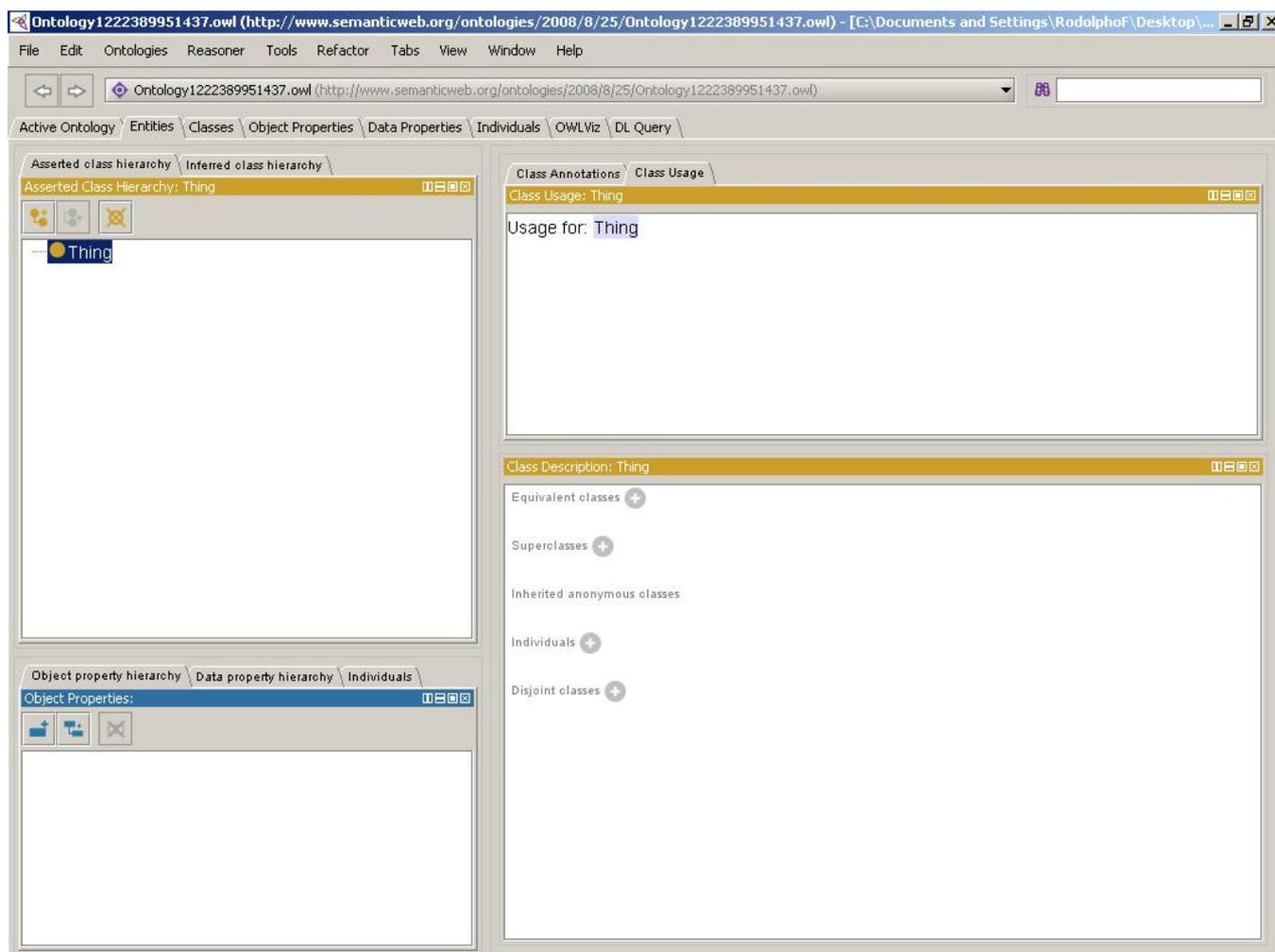


Fig. 10: Interface gráfica do Protégé

desenvolvimento do Protégé foi historicamente desenvolvido para aplicações biomédicas [16].

As ontologias biomédicas abordadas neste trabalho (exceto a BIT) tem uma implementação disponível em Protégé: o FMA ¹⁶, o GALEN ¹⁷, a BFO ¹⁸ e o OBO. ¹⁹.

O Protégé 4 possui dois sistemas de raciocínio automático nativos, que nos permitem fazer inferências sobre o domínio, são eles o Pellet [68], e o *Fast Classification of Terminologies*(FaCT) [69].

¹⁶Disponível em: <http://protege.cim3.net/cgi-bin/wiki.pl?FoundationalModelAnatomy>, último acesso 10/05/2009

¹⁷Disponível em: <http://krono.act.uji.es/Links/ontologies/full-galen-protege.zip/view>, último acesso 10/05/2009

¹⁸Disponível em: <http://www.ifomis.org/bfo/1.1>, último acesso 10/05/2009

¹⁹Disponível em: http://www.bioontology.org/wiki/index.php/OboInOwl:Main_Page, último acesso 10/05/2009

Ambos permitem a classificação da ontologia e verificação da sua consistência, entretanto o FaCT pode ser utilizado com várias lógicas descritivas. Em contrapartida o Pellet tem sua utilização restrita a linguagem OWL.

Sobre a diferença entre os resultados fornecidos por esses sistemas de inferência, Gardiner *et al.* [70] apresentam em seu trabalho o estudo de alguns sistemas de raciocínio automático usados para a linguagem OWL, entre eles o FaCT e o Pellet. Os autores concluem que todos os sistemas de raciocínio automático testados fornecem resultados corretos, sendo a principal diferença entre eles a performance (que varia de acordo com o número de indivíduos e classes a serem inferidas a partir do modelo definido).

No próximo capítulo apresentaremos os resultados, a formalização do sistema ventricular utilizando o BIT, sua inclusão no Protégé e sua integração com neuroimagens.

CAPÍTULO 4

RESULTADOS

Este capítulo utiliza os conceitos abordados nos capítulos 2 e 3 para construir uma representação formal do sistema ventricular cerebral e do conhecimento proveniente da literatura. Também nesta seção propomos a extensão da teoria de inclusão básica (BIT) para representar limites e continuidade entre as estruturas.

Durante esta formalização utilizaremos a seguinte terminologia para representar os elementos que compõem o sistema ventricular cerebral:

- Ventriculos laterais esquerdo e direito por “VLE” e “VLD” ;
- Terceiro Ventrículo por “TV”;
- Quarto Ventrículo por “QV”;
- Forame de Monro do ventrículo lateral esquerdo e direito por “FME” e “FMD”;
- Aqueduto Cerebral por “AC”;
- Forame de Luschka esquerdo por “FLE”;
- Forame de Luschka direito por “FLE”;
- Forame de Magendie por “FM”
- Corpos do ventrículo lateral esquerdo e direito por “CVLE e “CVLD”;

- Cornos frontais do ventrículo lateral esquerdo e direito por “CFVLE” e “CFVLD”;
- Cornos occipitais do ventrículo lateral esquerdo e direito por “COVLE” e “COVLD”;
- Cornos laterais do ventrículo lateral esquerdo e direito por “CLVLE” e “CLVLD”;
- Recesso Suprapineal do terceiro ventrículo por “RS”;
- Comissura Anterior do terceiro ventrículo por “CA”;
- Recesso óptico do terceiro ventrículo por “RO”;
- Infundíbulo do terceiro ventrículo por “I”;
- Recesso lateral do quarto ventrículo por “RL”.

Após a formalização, apresentaremos a criação de classes no Protégé que permitam a instanci-
ação de indivíduos de acordo com os conceitos que eles representam (ontológicos ou epistêmicos).
Por último, demonstraremos como a ontologia-epistemologia pode ser integrada a fatos extraídos da
análise estatística de neuroimagens, por meio de um processo de realimentação e sua integração com
o conhecimento encontrado na literatura.

4.1 Classificação dos elementos do SVC utilizando a teoria de classes

Nesta seção proveremos duas representações taxonômicas dos elementos que compõem o sis-
tema ventricular cerebral. A primeira representação é baseada no tipo das estruturas que compõem
o sistema ventricular cerebral. Como vimos até então, o sistema ventricular cerebral é composto ba-
sicamente por dois tipos de estruturas, os ventrículos (ventrículo lateral, terceiro ventrículo e quarto
ventrículo) e os forames (forame de Monro e aqueduto cerebral, forames de Luschka e forame de
Magendie). As fórmulas de 4.1 a 4.12 formalizam essa árvore taxonômica:

$$PP_{12}(VENTRICULOS, SVC) \quad (4.1)$$

$$Is_a(VLD, VENTRICULOS) \quad (4.2)$$

$$Is_a(VLE, VENTRICULOS) \quad (4.3)$$

$$Is_a(TV, VENTRICULOS) \quad (4.4)$$

$$Is_a(QV, VENTRICULOS) \quad (4.5)$$

$$PP_{12}(FORAMES, SVC) \quad (4.6)$$

$$Is_a(FME, FORAMES) \quad (4.7)$$

$$Is_a(FMD, FORAMES) \quad (4.8)$$

$$Is_a(AC, FORAMES) \quad (4.9)$$

$$Is_a(FLD, FORAMES) \quad (4.10)$$

$$Is_a(FLE, FORAMES) \quad (4.11)$$

$$Is_a(FM, FORAMES) \quad (4.12)$$

Podemos visualizar essa estrutura na figura 11.

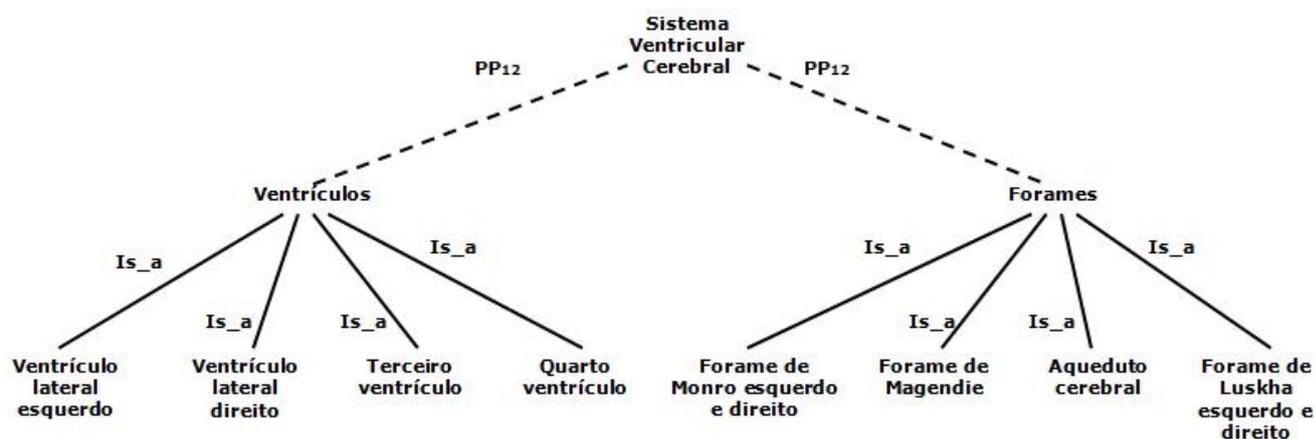


Fig. 11: Árvore taxonômica das estruturas que compõem o sistema ventricular cerebral.

Outra representação taxonômica proposta nesse trabalho é baseada na mereologia do sistema ventricular cerebral, cada estrutura e subestrutura anatômica que compõem o sistema ventricular cerebral é instanciada como uma classe como podemos observar na figura 12.

$$PP_{12}(VLD, SVC) \quad (4.13)$$

$$PP_{12}(VLE, SVC) \quad (4.14)$$

$$PP_{12}(FMD, SVC) \quad (4.15)$$

$$PP_{12}(FME, SVC) \quad (4.16)$$

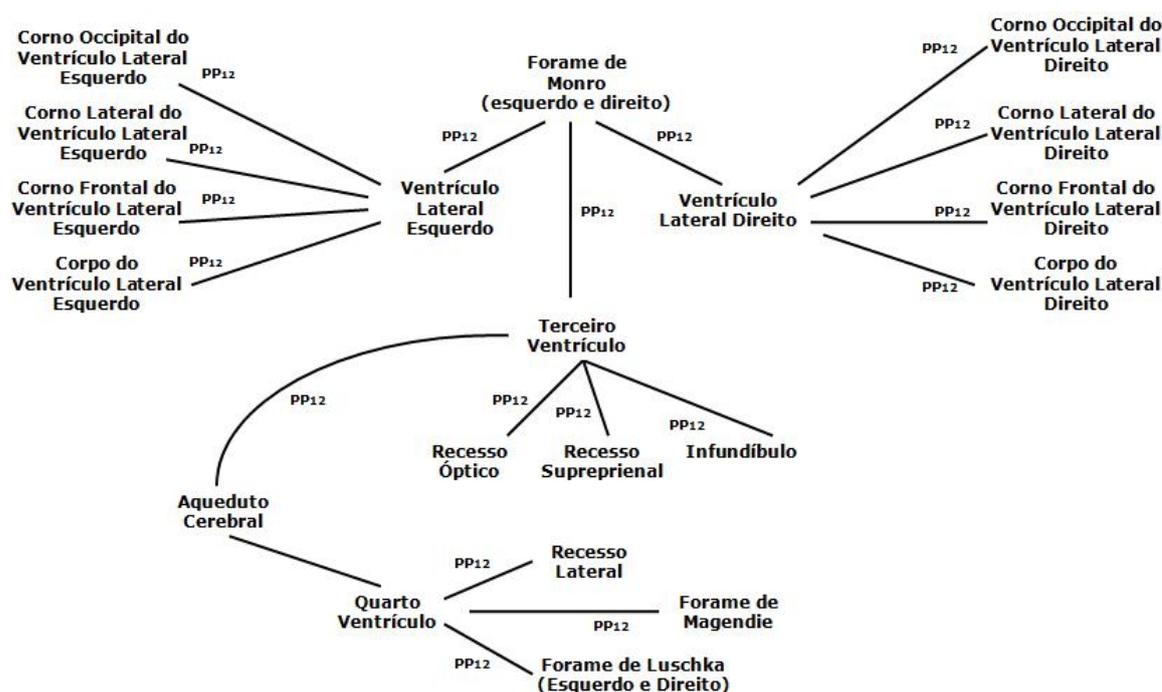


Fig. 12: Árvore taxonômica baseada na partonomia das regiões ventriculares.

$$PP_{12}(AC, SVC) \quad (4.17)$$

$$PP_{12}(COVLD, VLD) \quad (4.18)$$

$$PP_{12}(CVLD, VLD) \quad (4.19)$$

$$PP_{12}(CLVLD, VLD) \quad (4.20)$$

$$PP_{12}(CFVLD, VLD) \quad (4.21)$$

$$PP_{12}(COVLE, VLE) \quad (4.22)$$

$$PP_{12}(CVLE, VLE) \quad (4.23)$$

$$PP_{12}(CLVLE, VLE) \quad (4.24)$$

$$PP_{12}(CFVLE, VLE) \quad (4.25)$$

$$PP_{12}(FMD, CFVLD) \quad (4.26)$$

$$PP_{12}(FMD, TV) \quad (4.27)$$

$$PP_{12}(FME, CFVLE) \quad (4.28)$$

$$PP_{12}(FME, TV) \quad (4.29)$$

$$PP_{12}(RO, TV) \quad (4.30)$$

$$PP_{12}(RS, TV) \quad (4.31)$$

$$PP_{12}(I, TV) \quad (4.32)$$

$$PP_{12}(RL, TV) \quad (4.33)$$

$$PP_{12}(FLD, QV) \quad (4.34)$$

$$PP_{12}(FLE, QV)PP_{12}(FM, QV) \quad (4.35)$$

As fórmulas 4.13 a 4.16 representam que os ventrículos são parte própria do sistema ventricular cerebral, e todo sistema ventricular cerebral possui os ventrículos. De fato, as estruturas mesmo que alteradas morfológicamente, no caso do alargamento dos ventrículos laterais presentes na esquizofrenia, não passam a pertencer a uma nova região. O mesmo se dá com as divisões anatômicas dos ventrículos e os forames representados pelas fórmulas de 4.17 a 4.12.

Assim, com a formalização baseada na teoria de classes construída, é necessário criar formas de representar os limites de cada estrutura e a continuidade entre as estruturas do sistema ventricular cerebral. Nas seções 4.2 e 4.3 apresentaremos a proposta de expansão da BIT para relações de limites e continuidade entre estruturas anatômicas.

4.2 Representando a propriedade de continuidade

Uma descontinuidade pode ser definida como uma fronteira material ou não, entre duas regiões ou objetos de forma que essas regiões fiquem desconexas. O sistema ventricular cerebral, como discutido na seção 2.4 é contínuo nos ventrículos laterais, quarto ventrículo e nas regiões dos forames, porém é descontínuo no terceiro ventrículo pela massa intermediária intertalâmica. Essa propriedade de existência ou não de continuidade, é uma característica importante a ser expressa nas formalizações anatômicas do sistema ventricular cerebral.

Santos e Cabalar propõem em [71] uma maneira de representar a descontinuidade baseada no trabalho de Varzi [44]. Basicamente, a proposta é construir por meio de relações mereológicas uma ontologia capaz de representar as regiões contínuas. Neste trabalho usaremos esse mesmo princípio, entretanto, sendo o sistema ventricular cerebral uma região imaterial, qualquer descontinuidade seria caracterizada por uma estrutura essencialmente material, como no caso do terceiro ventrículo.

Para representar a existência de uma descontinuidade x em um objeto y definimos a relação 4.36.

$$Desc(x, y) \quad (4.36)$$

Dada a definição da relação que representa a descontinuidade, podemos então definir as propriedades:

$$Desc(x,y) \rightarrow LocIn(x,y) \quad (4.37)$$

A fórmula 4.37 nos diz que se x é uma descontinuidade na região y então x está localizado em y .

$$Desc(x,y) \rightarrow \neg Desc(y,z) \quad (4.38)$$

A fórmula 4.38 expressa que o local de existência de uma descontinuidade não é uma descontinuidade.

Assim, dada a definição de uma descontinuidade, podemos definir a continuidade de uma região através da fórmula 4.39. Esta fórmula define que uma região x é contínua à região y se x for parte própria de y ou x for parte de y e para toda a região z que é uma descontinuidade, esta não se localiza na região x .

$$PCont(x,y) \equiv [PP(x,y) \vee P(x,y)] \wedge \forall z \neg Desc(z,x) \quad (4.39)$$

Tendo definido a continuidade entre regiões, é necessário garantir a maximidade da região x , em outras palavras, que não exista nenhum outro segmento da região x que possa representá-la por inteiro. Essa representação é feita pela fórmula 4.40, que define que uma região é contínua e máxima se a região x for contínua a região y e não existir nenhuma região z que seja parte própria de y e contínua em y , ou seja:

$$Segment(x,y) \equiv PCont(x,y) \wedge \neg \exists z [PP(y,z) \wedge PCont(z,y)] \quad (4.40)$$

4.3 Representando limites fiat

Como abordamos no seção 2.4, o sistema ventricular cerebral é uma cavidade disposta no cérebro, responsável pela produção de líquido cefalorraquiano composta pelos ventrículos laterais esquerdo

e direito, terceiro ventrículo, quarto ventrículo e os forames, além de características anatômicas de cada estrutura.

Os ventrículos, suas particularidades anatômicas e os forames não têm limites físicos concretos para indicar sua divisão morfológica, muitas vezes são classificadas por características visuais ou funcionais, logo, usaremos limites fiat para representar as regiões de cada subestrutura.

Para definirmos uma relação que represente a existência de um limite fiat, tomaremos como base o trabalho de Santos e Cabalar [71]. Assim, podemos definir a relação $LFiat$ (fórmula 4.41), que assume a existência de um limite fiat x em uma região y . É importante ressaltar que o limite fiat não é uma descontinuidade na região e nem se situa de forma concreta nela.

$$LFiat(x,y) \quad (4.41)$$

Dada a definição de existência de um limite, podemos então definir as propriedades abaixo:

$$LFiat(z,x) \rightarrow PCoin(z,x) \quad (4.42)$$

A fórmula 4.42 representa que todo limite fiat z do objeto x coincide parcialmente com o objeto x :

$$LFiat(x,y) \rightarrow \neg \exists z LFiat(z,x) \quad (4.43)$$

A fórmula 4.43 garante a não reflexividade da relação $LFiat$, em outras palavras, o hospedeiro de um limite fiat não é um limite fiat.

$$LFiat(x,y) \wedge LFiat(x,z) \rightarrow \neg Desc(y,z) \quad (4.44)$$

A fórmula 4.44 nos diz que se x é um limite fiat em y e x é um limite fiat em z , então as regiões y e z são contínuas.

$$LFiat(x,y) \wedge LFiat(y,z) \rightarrow PCoin(x,y) \quad (4.45)$$

Finalmente, a fórmula 4.45 representa que se x é um limite fiat em y e y é um limite fiat em z então os limites fiat x e y coincidem parcialmente.

Os limites fiat utilizados neste trabalho são representados por Z_i (o índice i é um número que identifica o limite fiat) e estão compreendidos de Z_1 a Z_{21} . A imagem 13 apresenta o sistema ventricular cerebral, com a indicação do local onde são traçados cada limite fiat utilizado em nosso trabalho.

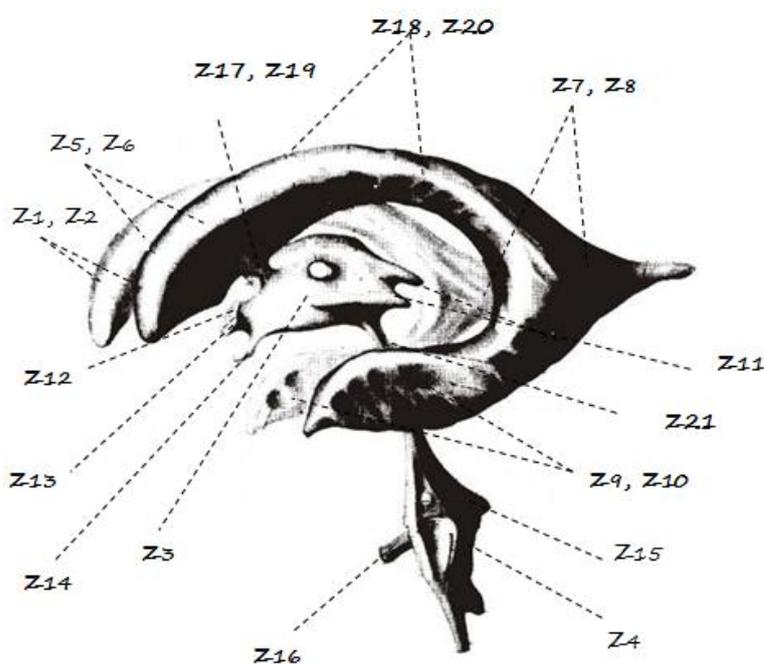


Fig. 13: Local onde são traçados os limites fiats para regiões do sistema ventricular cerebral.

A seguir apresentamos qual a região hospedeira de cada limite fiat que usamos na delimitação das estruturas do sistema ventricular cerebral, apresentado na figura 13.

- Z_1 é um limite fiat e sua região hospedeira é o ventrículo lateral esquerdo (VLE);
- Z_2 é um limite fiat e sua região hospedeira é ventrículo lateral direito (VLD);
- Z_3 é um limite fiat e sua região hospedeira é terceiro ventrículo (TV);
- Z_4 é um limite fiat e sua região hospedeira é quarto ventrículo (QV);

- Z_5 é um limite fiat e sua região hospedeira é corno frontal do ventrículo lateral esquerdo (CFVLE);
- Z_6 é um limite fiat e sua região hospedeira é corno frontal do ventrículo lateral direito (CFVLD);
- Z_7 é um limite fiat e sua região hospedeira é corno occipital do ventrículo lateral esquerdo (COVLE);
- Z_8 é um limite fiat e sua região hospedeira é corno occipital ventrículo lateral direito (COVLD);
- Z_9 é um limite fiat e sua região hospedeira é corno lateral do ventrículo lateral direito (CLVLD);
- Z_{10} é um limite fiat e sua região hospedeira é corno lateral do ventrículo lateral esquerdo (CLVLE);
- Z_{11} é um limite fiat e sua região hospedeira é recesso suprapineal do terceiro ventrículo (RS);
- Z_{12} é um limite fiat e sua região hospedeira é forame de Magendie do quarto ventrículo (FM);
- Z_{13} é um limite fiat e sua região hospedeira é recesso óptico do terceiro ventrículo (RO);
- Z_{14} é um limite fiat e sua região hospedeira é infundíbulo do terceiro ventrículo (I);
- Z_{15} é um limite fiat e sua região hospedeira é recesso lateral do quarto ventrículo (RL);
- Z_{16} é um limite fiat e sua região hospedeira é forame de Luschka direito com o quarto ventrículo (FLD);
- Z_{17} é um limite fiat e sua região hospedeira é forame de Luschka esquerdo com o quarto ventrículo (FLE);
- Z_{18} é um limite fiat e sua região hospedeira é forame de Monro esquerdo (FME);
- Z_{19} é um limite fiat e sua região hospedeira é corpo do ventrículo lateral esquerdo (CVLE);
- Z_{20} é um limite Fiat e sua região hospedeira é forame de Monro direito (FMD);
- Z_{21} é um limite fiat e sua região hospedeira é corpo do ventrículo lateral direito (CVLD);
- Z_{22} é um limite Fiat e sua região hospedeira é aqueduto cerebral(AC).

As definições apresentadas neste trabalho vão ao encontro da proposta do trabalho de Varzi [44] e são similares as definições propostas por Smith e Varzi [45], porém, esses trabalhos apresentam as definições de limites fiat baseadas na relação de sobreposição (O) e em algumas relações não presentes na BIT. Entretanto conforme já abordamos, a BIT permite que trabalhem com regiões espaciais por meio da relação r (como visto na seção 3.3) que é a base de relações que usamos para derivar a representação de limites fiats usada nesse trabalho.

4.4 Representando o sistema ventricular cerebral

Para representar o sistema ventricular cerebral, é necessário formalizar os limites fiat que delimitam cada estrutura que o compõe, conforme apresentado na figura 13. Esses limites são traçados com o conhecimento a priori do domínio e representados abaixo pelas fórmulas de 4.46 a 4.67:

$$LFiat(Z_1, VLE) \quad (4.46)$$

$$LFiat(Z_2, VLD) \quad (4.47)$$

$$LFiat(Z_3, TV) \quad (4.48)$$

$$LFiat(Z_4, QV) \quad (4.49)$$

$$LFiat(Z_5, CFVLE) \quad (4.50)$$

$$LFiat(Z_6, CFVLD) \quad (4.51)$$

$$LFiat(Z_7, COVLE) \quad (4.52)$$

$$LFiat(Z_8, COVLD) \quad (4.53)$$

$$LFiat(Z_9, CLVLD) \quad (4.54)$$

$$LFiat(Z_{10}, CLVLE) \quad (4.55)$$

$$LFiat(Z_{11}, RS) \quad (4.56)$$

$$LFiat(Z_{12}, FM) \quad (4.57)$$

$$LFiat(Z_{13}, RO) \quad (4.58)$$

$$LFiat(Z_{14}, I) \quad (4.59)$$

$$LFiat(Z_{15}, RL) \quad (4.60)$$

$$LFiat(Z_{16}, FLD) \quad (4.61)$$

$$LFiat(Z_{17}, FLE) \quad (4.62)$$

$$LFiat(Z_{18}, FME) \quad (4.63)$$

$$LFiat(Z_{19}, CVLE) \quad (4.64)$$

$$LFiat(Z_{20}, FMD) \quad (4.65)$$

$$LFiat(Z_{21}, CVLD) \quad (4.66)$$

$$LFiat(Z_{22}, AC) \quad (4.67)$$

Assim, dada as definições dos limites fiat usados para delimitar o sistema ventricular cerebral, temos condições de representar cada ventrículo individualmente. Entretanto, ao formalizarmos as regiões do sistema ventricular cerebral, percebemos que definir a entidade por meio de relações mereológicas, não garante que todos os aspectos e conhecimento adquirido sobre o domínio estarão presentes. Por exemplo, a esquizofrenia altera o tamanho dos ventrículos, fato esse que somente por relações mereológicas não conseguimos expressar consistentemente.

As relações de tamanho nos auxiliarão nessa tarefa, além de serem úteis para a desambiguação de regiões. Por exemplo, podemos definir além do posicionamento dos forames qual dos forames possíveis ele é dado o tamanho dos ventrículos aos quais ele se conecta, ou ao escolher uma região ventricular sabemos exatamente qual ela é pelo seu tamanho e não somente pelas definições mereológicas.

Para essa tarefa usaremos as relações de tamanho introduzidas em [72, 73]. Assim, o volume de cada região ventricular (que representaremos por φ) pode ser formalizado nas relações que definem os ventrículos, através da relação 4.68.

$$Vol(x) = \varphi \quad (4.68)$$

Sabemos a partir do conhecimento anatômico, que os ventrículos laterais esquerdo e direito são maiores em relação ao terceiro e quarto ventrículos, e o terceiro ventrículo é maior que o quarto: “Ventrículo Lateral esquerdo e direito (VLE e VLD) > Terceiro Ventrículo (TV) > Quarto Ventrículo (QV)”.

Assim, podemos representar os volumes de cada ventrículo por:

- Volume do ventrículo lateral direito representado por Vol_VLD;

- Volume do ventrículo lateral esquerdo representado por Vol_VLE ;
- Volume do terceiro ventrículo representado por Vol_TV ;
- Volume do quarto ventrículo representado por Vol_QV .

Assim, através dos limites fiat e das relações de tamanho, podemos criar uma representação livre de ambiguidades e expressiva para representar cada ventrículo levando em consideração sua posição e seu tamanho.

$$\begin{aligned} Inst(x, VLD) \leftarrow (Vol(x) = \varphi) \wedge [\varphi > Vol_TV] \wedge [\varphi > Vol_QV] \wedge \\ LFiats(Z_2, x) \wedge Segment(x, SVC) \end{aligned} \quad (4.69)$$

A fórmula 4.69 nos diz que x é o ventrículo lateral direito, se seu volume φ for maior que o volume do terceiro ventrículo e maior que o volume do quarto ventrículo e z_2 for limite fiat em x e x for contínua e máxima com relação ao sistema ventricular cerebral.

$$\begin{aligned} Inst(x, VLE) \leftarrow (Vol(x) = \varphi) \wedge [\varphi > Vol_TV] \wedge [\varphi > Vol_QV] \wedge \\ LFiats(Z_1, x) \wedge Segment(x, SVC) \end{aligned} \quad (4.70)$$

A fórmula 4.70 representa que x é o ventrículo lateral esquerdo, se seu volume φ for maior que o volume do terceiro ventrículo e maior que o volume do quarto ventrículo e Z_1 for limite fiat em x e x for contínua e máxima ao sistema ventricular cerebral.

$$\begin{aligned} Inst(x, TV) \leftarrow (Vol(x) = \varphi) \wedge [(\varphi < Vol_VLE) \wedge (\varphi < Vol_VLD)] \wedge \\ (\varphi > Vol_QV) \wedge LFiats(Z_3, x) \wedge Segment(x, SVC) \end{aligned} \quad (4.71)$$

A fórmula 4.71 define que x é o terceiro ventrículo, se seu volume φ for menor que o volume do ventrículo lateral esquerdo e direito e maior que o volume do quarto ventrículo e Z_3 for limite fiat em x e x for contínua e máxima ao sistema ventricular cerebral.

$$\begin{aligned} Inst(x, QV) \leftarrow (Vol(x) = \varphi) \wedge [(\varphi < Vol_VLE) \wedge (\varphi < Vol_VLD)] \wedge \\ (\varphi < Vol_TV) \wedge LFIat(Z_4, x) \wedge Segment(x, SVC) \end{aligned} \quad (4.72)$$

A fórmula 4.71 nos diz que x é o quarto ventrículo, se seu volume φ for menor que o volume do ventrículo lateral esquerdo e direito e menor que o volume do terceiro ventrículo e Z_4 for limite fiat em x e x for contínua e máxima ao sistema ventricular cerebral.

Após a formalização dos ventrículos é necessário formalizar cada detalhe anatômico dos ventrículos, como veremos na próxima seção.

4.4.1 Formalizando detalhes anatômicos e forames

Como visto no capítulo 2.4, a estrutura do sistema ventricular cerebral é composta pelos ventrículos laterais esquerdo e direito, que são subdivididos pelo corpo, corno lateral, corno occipital e corno frontal. O terceiro ventrículo é subdividido em comissura anterior, recesso óptico, recesso suprapineal e infundíbulo. O quarto ventrículo em recesso lateral e o forame de Luschka. Ainda temos os forames de Monro esquerdo e direito que se conectam com o corno frontal do ventrículo lateral esquerdo e direito.

Essas características tornam necessária a formalização dos detalhes anatômicos que compõem o sistema ventricular cerebral, representado pelas fórmulas de 4.73 a 4.89.

$$Inst(x, COVLD) \leftarrow Segment(x, SVC) \wedge LFIat(Z_8, x) \quad (4.73)$$

$$Inst(x, CLVLD) \leftarrow Segment(x, SVC) \wedge LFIat(Z_9, x) \quad (4.74)$$

$$Inst(x, COVLE) \leftarrow Segment(x, SVC) \wedge LFIat(Z_7, x) \quad (4.75)$$

$$Inst(x, CLVLE) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{10}, x) \quad (4.76)$$

$$Inst(x, FM) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{12}, x) \quad (4.77)$$

$$Inst(x, RO) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{13}, x) \quad (4.78)$$

$$Inst(x, RS) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{11}, x) \quad (4.79)$$

$$Inst(x, I) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{14}, x) \quad (4.80)$$

$$Inst(x, RL) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{15}, x) \quad (4.81)$$

$$Inst(x, CFVLD) \leftarrow Segment(x, SVC) \wedge LFIat(Z_6, x) \wedge LFIat(Z_{19}, x) \quad (4.82)$$

$$Inst(x, CVLD) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{21}, x) \quad (4.83)$$

$$Inst(x, CVLE) \leftarrow Segment(x, SVC) \wedge LFIat(Z_{19}, x) \quad (4.84)$$

$$Inst(x, CFVLE) \leftarrow Segment(x, SVC) \wedge LFIat(Z_5, x) \wedge LFIat(Z_{17}, x) \quad (4.85)$$

$$Inst(x, FME) \leftarrow LFIat(Z_{18}, x) \wedge LFIat(Z_5, x) \wedge LFIat(Z_3, x) \wedge Segment(x, SVC) \quad (4.86)$$

$$Inst(x, FMD) \leftarrow LFIat(Z_{20}, x) \wedge LFIat(Z_6, x) \wedge LFIat(Z_3, x) \wedge Segment(x, SVC) \quad (4.87)$$

$$Inst(x, AC) \leftarrow LFIat(Z_{22}, x) \wedge LFIat(Z_3, x) \wedge LFIat(Z_4, x) \wedge Segment(x, SVC) \quad (4.88)$$

$$Inst(x, FLD) \leftarrow LFIat(Z_{16}, x) \wedge LFIat(Z_5, x) \wedge Segment(x, SVC) \quad (4.89)$$

$$Inst(x, FLE) \leftarrow LFIat(Z_{17}, x) \wedge LFIat(Z_5, x) \wedge Segment(x, SVC) \quad (4.90)$$

Nos focamos, até então, em formalizar a estrutura anatômica do sistema ventricular cerebral, bem como suas propriedades de continuidade e limites fiat e usamos a teoria de classes da BIT para construir uma representação da entidade de forma hierárquica. Na próxima seção abordaremos como podemos inserir o conhecimento encontrado na literatura na representação formal do sistema ventricular cerebral.

4.5 Incluindo conhecimento na representação formal - Classes epistêmicas.

Com o crescimento do uso de técnicas de diagnóstico por imagem, um extenso número de trabalhos de meta-análise foi publicado. Esses trabalhos são baseados em estudos realizados com neuroimagens e fornecem importantes informações, percentuais de alterações de volume, número de indivíduos afetados em um grupo com características em comum, entre outros.

No entanto representar conhecimento (em nosso trabalho, informações presentes na literatura) não é tarefa da ontologia. As alterações apontadas nas meta-análises são submetidas a análise do especialista, em função de um grupo e geralmente essas representam a média das alterações percebidas nesse grupo.

A representação desse tipo de informação é tarefa para epistemologia (como abordamos na seção

2.2), que se ocupa da representação do conhecimento. Também como abordamos na seção 2.2 representar conceitos como indivíduos ou objetos nos leva ao problema conceitual da confusão Kantiana (em nosso trabalho, representar conhecimento como uma entidade anatômica ou um indivíduo).

A BIT não possui relações ou axiomas específicos que nos permitam separar o domínio ontológico do domínio epistêmico, assim, para podermos expressar os conceitos epistêmicos, propomos nesse trabalho uma representação baseada em classes. As classes ontológicas representarão o domínio e sob elas, estarão as classes epistêmicas que representam o conhecimento sobre o domínio. Um esquema gráfico pode ser observado na figura 14.

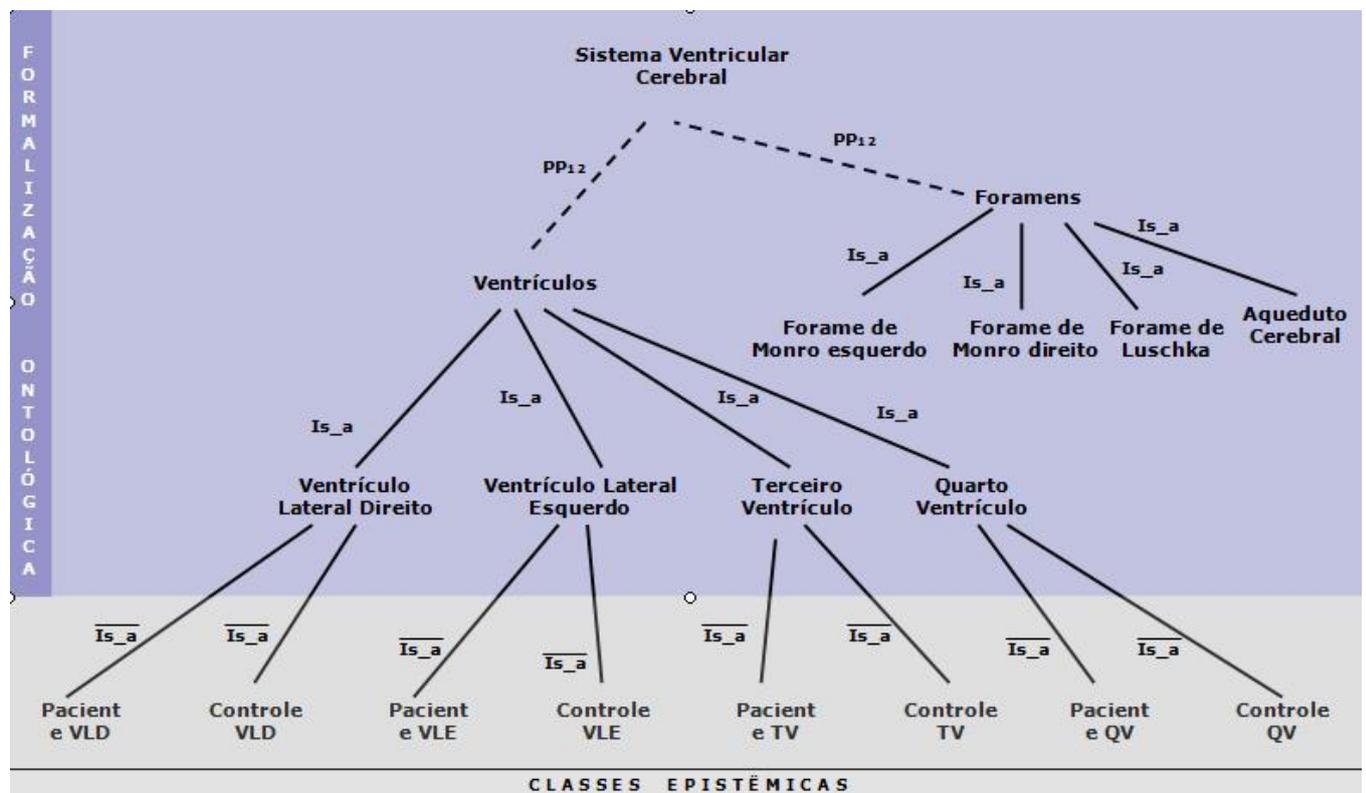


Fig. 14: Classes epistêmicas e ontológicas.

Assim garantimos que indivíduos que representam conhecimento sobre o domínio, ficarão instanciados sob as classes epistêmicas e indivíduos que representam características do domínio sob as classes ontológicas. Porém a instanciação tradicional da BIT por meio do axioma *Is_a* (como visto na seção 3.3, fórmula 3.38) só permite que uma classe *A* seja instanciada à classe *B* se elas tiverem o indivíduo *x* em comum.

Essa representação não nos atende diretamente, pois um indivíduo epistêmico *x* deve ser instan-

ciado somente a uma classe epistêmica *EPIST*, e não deve ser instanciado a classe ontológica *ONT*. Além dessa característica, é necessário também indicar que o indivíduo x traga um conhecimento sobre a classe ontológica *ONT*.

Essa definição pode ser representada pela lógica epistêmica. Spies [74] argumenta que a maioria das fontes bibliográficas sobre gerenciamento do conhecimento não contemplam conceitos epistêmicos. A integração entre o domínio epistêmico e ontológico é proposta por Tabet [75], que aborda a utilização da lógica epistêmica em conjunto com a ontologia, basicamente por meio das relações que representam a crença em determinado fato.

Assim, nos inspiramos na lógica epistêmica e na integração proposta por Tabet [75] como alternativa para representação do conhecimento junto a ontologia proposta nesse trabalho, representando o conhecimento referente domínio por meio da relação κ . A Relação κ neste trabalho, representa que o indivíduo x traz um conhecimento sobre uma classe A . Em síntese, ele permite que o indivíduo que representa conhecimento seja instanciado a uma subclasse epistêmica e relacionado a uma classe ontológica, formalmente representamos:

$$\kappa(x, A) \tag{4.91}$$

Para a representação dos indivíduos epistêmicos, propomos uma nova relação para BIT que chamaremos de $\overline{Is_a}$. Basicamente essa relação indica (além da condição de instanciação dos indivíduos), que o indivíduo x traz um conhecimento sobre uma classe ontológica (dado através da relação 4.91), como podemos verificar na fórmula 4.92.

$$\overline{Is_a}(EPIST, ONT) \equiv \forall x(Inst(x, EPIST) \rightarrow \neg Inst(x, ONT) \wedge \kappa(x, ONT)) \tag{4.92}$$

Podemos então criar as classes que representarão o conhecimento presente na literatura, que chamaremos de classes epistêmicas. Para criarmos as classes epistêmicas usadas nesse trabalho levamos em consideração o conhecimento sobre alterações ventriculares contidos na literatura onde esse conhecimento é geralmente separado em grupos de paciente e controle.

As fórmulas de 4.93 a 4.100 apresentam a instanciação das classes epistêmicas que representam o conhecimento encontrado na literatura, no caso, as classes de controle e paciente dos ventrículos

laterais direito, esquerdo, terceiro e quarto ventrículo as suas respectivas classes ontológicas.

$$\overline{Is_a}(CONTROLE_VLD, VLD) \quad (4.93)$$

$$\overline{Is_a}(PACIENTE_VLD, VLD) \quad (4.94)$$

$$\overline{Is_a}(CONTROLE_VLE, VLE) \quad (4.95)$$

$$\overline{Is_a}(PACIENTE_VLE, VLE) \quad (4.96)$$

$$\overline{Is_a}(CONTROLE_TV, TV) \quad (4.97)$$

$$\overline{Is_a}(PACIENTE_TV, TV) \quad (4.98)$$

$$\overline{Is_a}(CONTROLE_QV, QV) \quad (4.99)$$

$$\overline{Is_a}(PACIENTE_QV, QV) \quad (4.100)$$

Criadas as classes epistêmicas que representam os principais grupos estudados na literatura (paciente e controle), podemos criar sub-classes epistêmicas que representaram cada informação contida nas meta-análises. Nesse caso utilizamos a relação Is_a , já que a herança de indivíduos entre classes epistêmicas é permitida. A fórmula 4.101 é um modelo para representar os valores encontrados na meta-análise de Steen *et al.* [38], que contém um compêndio de diversos autores que relatam alterações de volume no ventrículo lateral direito, esquerdo e terceiro ventrículo.

A fórmula 4.101 representa a instanciação de uma subclasse epistêmica A à uma classe epistêmica B se e somente se, existir um indivíduo x que tenha o volume de σ , este volume é obtido de acordo com dados na literatura médica. As abreviações ¹ no primeiro e segundo argumentos da relação Is_a , bem como os valores de σ estão na tabela 3

$$Is_a(A, B) \leftrightarrow \exists x Vol(x) = \sigma \quad (4.101)$$

Por exemplo, a fórmula 4.102 representa a subclasse epistêmica $Barr_C_VLD$ instanciada a classe epistêmica $Controle_VLD$ se e somente se o volume do indivíduo for igual a 6.52, essa representação corresponde a primeira linha da tabela 3.

¹Devido ao grande número de fórmulas optamos por utilizar os termos genéricos para facilitar a compreensão do conceito e a leitura.

$$Is_a(Barr_C_VLD, Controle_VLD) \leftrightarrow \exists x Vol(x) = 6.52 \quad (4.102)$$

Assim concluímos a representação ontológica do sistema ventricular cerebral, onde formalizamos domínio e suas relações e a epistêmica, onde buscamos maneiras de representar o conhecimento presente na literatura. No próximo capítulo apresentaremos a implementação dessa teoria formal utilizando o Protégé.

4.6 Implementação da formalização ontológica e epistêmica no Protégé

Como já abordamos na seção 3.4, o Protégé é uma ferramenta amplamente utilizada para construção de ontologias para diversos domínios, porém historicamente construída para utilização no domínio biomédico. O Protégé trabalha com quatro principais características do domínio, apresentadas na figura 15.



Fig. 15: As abas de propriedades do domínio no Protégé.

- **Class - Classes;**
Representam o domínio de estudo, nesse caso o sistema ventricular cerebral e suas divisões anatômicas.
- **Object Properties - Propriedades do Objeto;**
Neste trabalho, representam as relações e axiomas da BIT e derivados para representar o domínio.
- **Data Properties - Propriedade de Dados;**
Representa-se o tipo de dados (Inteiro, Caracter, etc) utilizado para representar alguma informação do domínio.
- **Individuals - Indivíduos.**
Representa os indivíduos. Os indivíduos são representações únicas e singulares do domínio, ou representações associadas ao conhecimento especialista sobre o domínio

Classes epistêmicas criadas com base na meta-análise de Steen <i>et al.</i> [38]				
Sub-classe epistêmica (A)	Conceito	Classe epistêmica (B)	Classe ontológica	σ
Barr_C_VLD	Controle VLD, segundo Barr	CONTROLE_VLD	VLD	6.52
Chua_C_VLD	Controle VLD, segundo Chua	CONTROLE_VLD	VLD	4.40
Degreef_C_VLD	Controle VLD, segundo Degreef	CONTROLE_VLD	VLD	6.75
Fannon_C_VLD	Controle VLD, segundo Fannon	CONTROLE_VLD	VLD	4.70
Fannonb_C_VLD	Controle VLD, segundo Fannon(b)	CONTROLE_VLD	VLD	4.70
James_C_VLD	Controle VLD, segundo James	CONTROLE_VLD	VLD	5.89
Lawrie_C_VLD	Controle VLD, segundo Lawrie	CONTROLE_VLD	VLD	3.50
Whith_C_VLD	Controle VLD, segundo Whith	CONTROLE_VLD	VLD	3.11
DeLisi_C_VLD	Controle VLD, segundo DeLisi	CONTROLE_VLD	VLD	5.60
Barr_P_VLD	Paciente VLD, segundo Barr	PACIENTE_VLD	VLD	8.22
Chua_P_VLD	Paciente VLD, segundo Chua	PACIENTE_VLD	VLD	5.10
Degreef_P_VLD	Paciente VLD, segundo Degreef	PACIENTE_VLD	VLD	8.15
Fannon_P_VLD	Paciente VLD, segundo Fannon	PACIENTE_VLD	VLD	4.60
Fannonb_P_VLD	Paciente VLD, segundo Fannon(b)	PACIENTE_VLD	VLD	5.80
James_P_VLD	Paciente VLD, segundo James	PACIENTE_VLD	VLD	8.12
Lawrie_P_VLD	Paciente VLD, segundo Lawrie	PACIENTE_VLD	VLD	3.70
Whith_P_VLD	Paciente VLD, segundo Whith	PACIENTE_VLD	VLD	4.99
DeLisi_P_VLD	Paciente VLD, segundo DeLisi	PACIENTE_VLD	VLD	6.75
Barr_C_VLE	Controle VLE, segundo Barr	CONTROLE_VLE	VLE	6.98
Chua_C_VLE	Controle VLE, segundo Chua	CONTROLE_VLE	VLE	5.00
Degreef_C_VLE	Controle VLE, segundo Degreef	CONTROLE_VLE	VLE	6.70
Fannon_C_VLE	Controle VLE, segundo Fannon	CONTROLE_VLE	VLE	4.50
Fannonb_C_VLE	Controle VLE, segundo Fannon(b)	CONTROLE_VLE	VLE	4.50
James_C_VLE	Controle VLE, segundo James	CONTROLE_VLE	VLE	6.16
Lawrie_C_VLE	Controle VLE, segundo Lawrie	CONTROLE_VLE	VLE	3.80
Whith_C_VLE	Controle VLE, segundo Whith	CONTROLE_VLE	VLE	2.81
DeLisi_C_VLE	Controle VLE, segundo DeLisi	CONTROLE_VLE	VLE	5.54
Barr_P_VLE	Paciente VLE, segundo Barr	CONTROLE_VLE	VLE	9.16
Chua_P_VLE	Paciente VLE, segundo Chua	CONTROLE_VLE	VLE	6.00
Degreef_P_VLE	Paciente VLE, segundo Degreef	CONTROLE_VLE	VLE	8.91
Fannon_P_VLE	Paciente VLE, segundo Fannon	CONTROLE_VLE	VLE	5.10
Fannonb_P_VLE	Paciente VLE, segundo Fannon(b)	CONTROLE_VLE	VLE	6.20
James_P_VLE	Paciente VLE, segundo James	CONTROLE_VLE	VLE	9.66
Lawrie_P_VLE	Paciente VLE, segundo Lawrie	CONTROLE_VLE	VLE	3.90
Whith_P_VLE	Paciente VLE, segundo Whith	CONTROLE_VLE	VLE	4.96
DeLisi_P_VLE	Paciente VLE, segundo DeLisi	CONTROLE_VLE	VLE	6.84
Cahn_C_TV	Controle TV, segundo Cahn	CONTROLE_TV	TV	0.62
Degreef_C_TV	Controle TV, segundo Degreef	CONTROLE_TV	TV	1.62
DeLisi_C_TV	Controle TV, segundo DeLisi	CONTROLE_TV	TV	0.91
Fannon_C_TV	Controle TV, segundo Fannon	CONTROLE_TV	TV	0.68
Fannonb_C_TV	Controle TV, segundo Fannon(b)	CONTROLE_TV	TV	0.70
James_C_TV	Controle TV, segundo James	CONTROLE_TV	TV	1.60
Lawrie_C_TV	Controle TV, segundo Lawrie	CONTROLE_TV	TV	0.40
Lim_C_TV	Controle TV, segundo Lim	CONTROLE_TV	TV	0.27
Cahn_P_TV	Paciente TV, segundo Cahn	PACIENTE_TV	TV	0.85
Degreef_P_TV	Paciente TV, segundo Degreef	PACIENTE_TV	TV	1.33
DeLisi_P_TV	Paciente TV, segundo De Lisi	PACIENTE_TV	TV	0.91
Fannon_P_TV	Paciente TV, segundo Fannon	PACIENTE_TV	TV	0.77
Fannonb_P_TV	Paciente TV, segundo Fannon(b)	PACIENTE_TV	TV	0.90
James_P_TV	Paciente TV, segundo James	PACIENTE_TV	TV	2.08
Lawrie_P_TV	Paciente TV, segundo Lawrie	PACIENTE_TV	TV	0.60
Lim_P_TV	Paciente TV, segundo Lim	PACIENTE_TV	TV	0.34

Tab. 3: Classes e subclasses epistêmicas e ontológicas usadas na abreviação da fórmula 4.101. Na primeira coluna temos as subclasses epistêmicas representadas por (A), na segunda coluna o conceito representado pela subclasse, na terceira coluna a classe epistêmica (B) à que pertence a subclasse (A), na quarta coluna a classe ontológica a que pertence o conhecimento, e finalmente, na quinta coluna (σ) os valores de volume ventricular.

Dado esses conceitos, demonstraremos nesta seção como foi feita a implementação da ontologia-epistemologia utilizando o Protégé. Primeiramente, como podemos observar na figura 16, foram criadas as classes que representam o sistema ventricular cerebral, suas divisões anatômicas e limites fiat.



Fig. 16: Classes do sistema ventricular cerebral e epistêmicas implementadas no Protégé, na imagem pode ser observada a classe Z_1 , que representa o limite fiat do ventrículo lateral direito

Inserimos então as fórmulas e relações que sustentam a teoria apresentada neste capítulo. A figura 17(A) indica a guia onde inserimos as relações da BIT e outras relações criadas que suportam o domínio como os limites fiats e continuidade, e a composição dessas propriedades. Os tipos de dados são criados conforme a figura 17(B), o tipo de dados neste trabalho representa os volumes ventriculares e percentuais encontrados na literatura.

A inserção de indivíduos é feita na aba *Individuals*, em nosso trabalho, indivíduos podem ser ontológicos ou epistêmicos, dependendo do tipo de conceito que representam. A figura 17(C) apresenta a tela do Protégé onde fazemos a inserção de indivíduos.

4.6.1 Classes epistêmicas no Protégé

Sendo o Protégé destinado a implementação de ontologias, a criação de classes epistêmicas não é suportada pela ferramenta de forma desejada (instanciada a uma classe ontológica sem herdar indivíduos ou características). No Protégé uma subclasse sempre herda características de sua superclasse, sendo assim, os indivíduos instanciados a uma subclasse epistêmica seriam herdados pelas classes ontológicas, fato que queremos evitar separando a ontologia da epistemologia.

Ao tentarmos codificar essa restrição, o teste da ontologia pelo Protégé aponta uma falha de

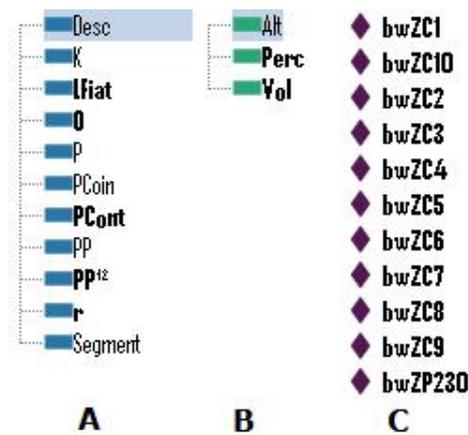


Fig. 17: Telas do Protégé para inserção de propriedades de objetos (A), propriedades de dados (B) e criação de indivíduos (C)(os indivíduos que aparecem nesta imagem, são neuroimagens que tem seu volume ventricular conhecido através do algoritmo de crescimento de região como veremos futuramente).

consistência, a alternativa encontrada para utilização do Protégé nessa formalização foi a criação de classes epistêmicas separadas das entidades ontológicas. Entretanto, separar as classes não permitiria que os indivíduos da epistemologia fossem relacionados à ontologia, ou ainda fazer consultas epistêmicas baseadas nas características do domínio como por exemplo: “*quais indivíduos do ventrículo lateral esquerdo são do grupo de pacientes*”.

Podemos minimizar essa limitação por meio da relação κ abordada anteriormente neste capítulo. Criada no Protégé, ela indica que o indivíduo instanciado às classes epistêmicas contém um conhecimento referente ao domínio ontológico. A imagem 18 apresenta um exemplo de indivíduo instanciado a classe epistêmica e a definição da relação κ que indica sobre qual classe do domínio ele representa conhecimento.



Fig. 18: Indivíduo de exemplo (A) instanciado a classe epistêmica *Barr_C_VLE* e sua relação com a ontologia por meio da relação κ (B).

A figura 18 apresenta um indivíduo de exemplo, instanciado a classe epistêmica *Barr_C_VLE* e a relação κ , que relaciona o indivíduo com a classe *Ventriculo_Lateral_Esquerdo*, indicando que

esse indivíduo traz um conhecimento sobre essa classe.

As classes que formalizam o conhecimento encontrado na literatura, representadas pela fórmula 4.101, são criadas no domínio das classes epistêmicas e descritas em suas propriedades os valores encontrados na literatura, conforme podemos observar na figura 19.

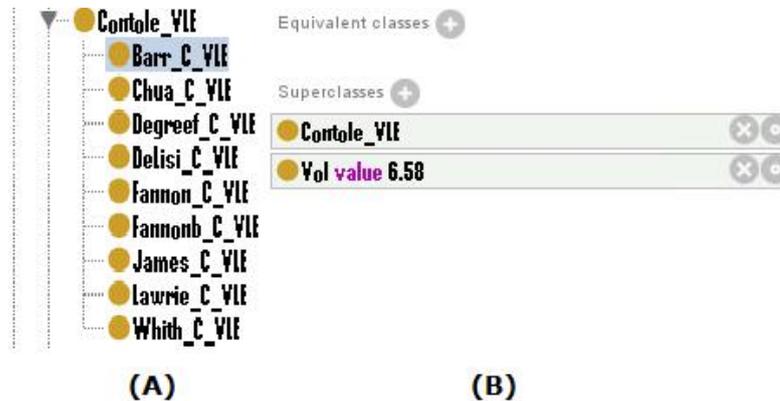


Fig. 19: Classe epistêmica (A) e descrição do conhecimento encontrado na meta-análise de Steen *et al.* [38] segundo Barr. O valor de volume ventricular encontrado é representado nas propriedades de classe (B).

4.6.2 Raciocinando sobre a formalização proposta

Com as classes, restrições e tipo de dados implementados no Protégé, podemos executar o sistema de raciocínio automático (FaCT ou Pellet). A execução do sistema de raciocínio automático processa a subsunção de classes (classificação da ontologia), atribui as propriedades de objetos, instanciação de indivíduos e permite que sejam executadas consultas.

A figura 20 apresenta uma parte das relações inferidas pelo sistema de raciocínio automático para o corno frontal do ventrículo lateral direito, no caso as classes disjuntas a essa estrutura como a classe *C_Direito* (corpo do ventrículo lateral direito), *CL_Direito* (corno lateral do ventrículo lateral direito), *CO_Direito* (corno occipital do ventrículo lateral direito) e o limite fiat que delimita o corno frontal do ventrículo Z_6 .

Ainda na imagem 20, abaixo a esquerda estão as propriedades de classe, o quadro delineado com linha pontilhada foi inferido, e apresenta a classe que representa o forame de Monro direito que se conecta com o corno frontal do ventrículo lateral direito.

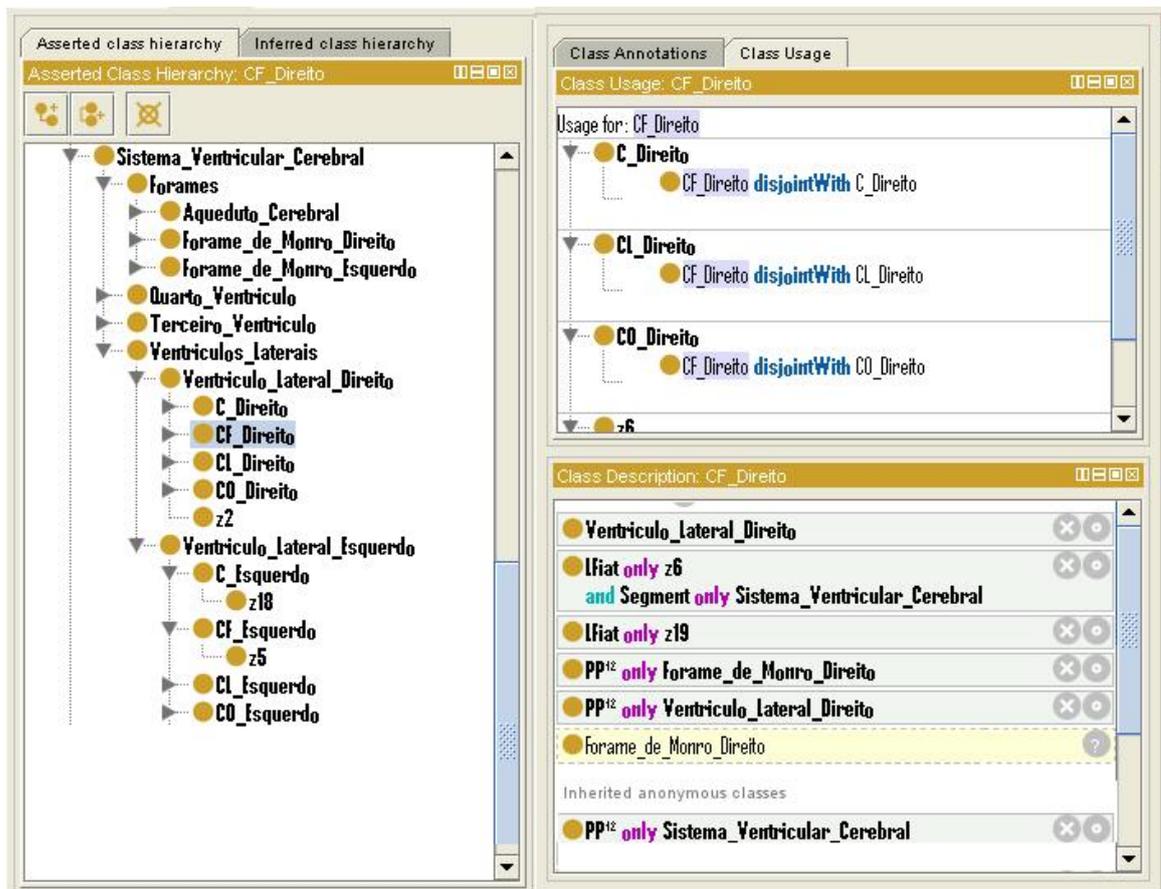


Fig. 20: Parte das inferências realizadas sobre a classe que representa o corno frontal do ventrículo lateral direito. Acima a esquerda é apresentada a subsunção de classes, contendo classes disjuntas e limites fiat, abaixo delineado pela linha pontilhada a classe inferida que representa o forame de Monro direito, que se conecta com o corno frontal do ventrículo lateral

Podemos então gerar algumas consultas para verificar a consistência da ontologia. As figuras apresentam as consultas realizadas no Protégé seguidas pelo resultados da inferência.

A primeira consulta é sobre a composição anatômica do ventrículo lateral esquerdo, apresentada na figura 21. A consulta realizada foi: *PP₁₂ only Ventrículo_Lateral_Direito*, que significa: quais estruturas são partes próprias do sistema ventricular cerebral e que estruturas o sistema ventricular tem como parte própria ?. Obtemos como resultado as estruturas anatômicas que compõem o ventrículo lateral direito (*C_Direito*, *CF_Direito*, *CL_Direito* e *CO_Direito*).

The screenshot shows the Protégé query results window. The query is *PP₁₂ only Ventrículo_Lateral_Direito*. The results are displayed in a tree view under the heading "Query results".

- Ventriculos_laterais** (Super classes)
 - Ventrículo_Lateral_Direito** (Sub classes)
 - C_Direito**
 - CF_Direito**
 - CL_Direito**
 - CO_Direito**
- Descendant classes**
 - C_Direito**
 - CF_Direito**
 - CL_Direito**
 - CO_Direito**
 - Nothing**
 - z20**
 - s**

Fig. 21: Resultado da consulta que indica a composição anatômica do ventrículo lateral direito.

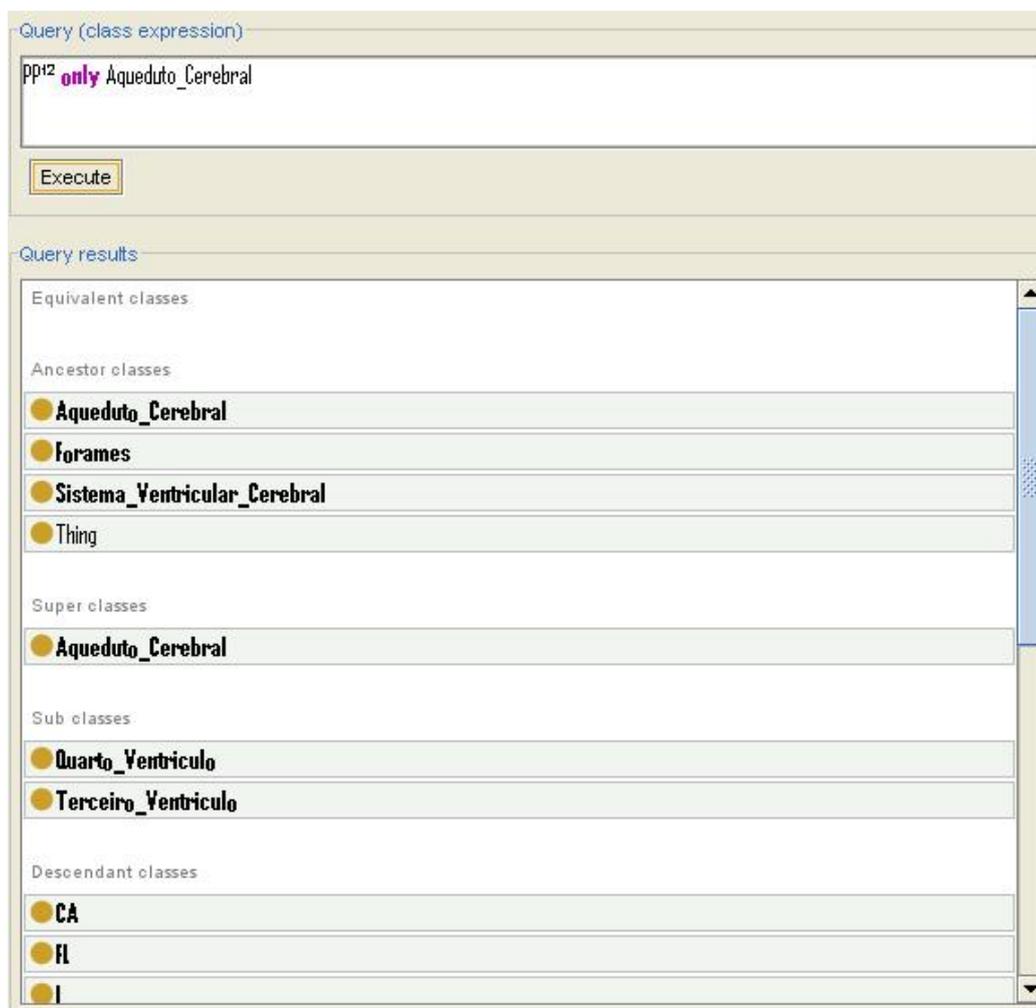
A segunda consulta é sobre qualquer região do sistema ventricular cerebral que não seja contínua e máxima, apresentada na figura 22. A consulta realizada foi: *not(Segment only Sistema_Ventricular_Cerebral)*, que significa: qual estrutura não é contínua ao sistema ventricular cerebral ? O resultado obtido é a massa intermediária intertalâmica, representada no Protégé por MI.

The image shows a screenshot of the Protégé software interface. At the top, there is a text box labeled "Query (class expression)" containing the query: `not (Segment only Sistema_Ventricular_Cerebral)`. Below the text box is an "Execute" button. Underneath is a section titled "Query results" which contains several categories of class information:

- Equivalent classes:** (empty)
- Ancestor classes:**
 - Thing
- Super classes:**
 - Thing
- Sub classes:**
 - MI
- Descendant classes:**
 - MI
 - Nothing
- Instances:** (empty)

Fig. 22: Resultado da consulta que indica a massa intermediária intertalâmica, que é uma descon- tinuidade no sistema ventricular cerebral.

A terceira consulta, apresentada na figura 23, é sobre a composição anatômica do aqueduto cerebral. A consulta realizada foi *PP12 only Aqueduto_Cerebral*, que significa: quais estruturas são partes próprias do aqueduto cerebral e quais estruturas o aqueduto cerebral tem como parte própria ?. O resultado obtido é o terceiro e o quarto ventrículo.



The screenshot displays a query execution window. At the top, the query is defined as "Query (class expression)" with the expression "PP12 only Aqueduto_Cerebral". Below the query is an "Execute" button. The results section, titled "Query results", is divided into several categories:

- Equivalent classes:** (empty)
- Ancestor classes:**
 - Aqueduto_Cerebral
 - Forames
 - Sistema_Ventricular_Cerebral
 - Thing
- Super classes:**
 - Aqueduto_Cerebral
- Sub classes:**
 - Quarto_Ventriculo
 - Terceiro_Ventriculo
- Descendant classes:**
 - CA
 - FL
 - I

Fig. 23: Resultado da consulta que apresenta a anatomia do aqueduto cerebral.

A quarta consulta, como podemos observar na figura 24, é sobre a composição anatômica do forame de Monro direito. A consulta realizada foi: *PP12 only Forame_de_Monro_Direito*, que significa: quais estruturas são partes próprias do forame de Monro direito e quais estruturas o forame de Monro direito tem como parte própria ?. O resultado obtido é o corno frontal do ventrículo lateral direito e o terceiro ventrículo.

The screenshot shows a query execution interface with the following sections:

- Query (class expression):** Contains the query `PP12 only Forame_de_Monro_Direito`.
- Execute:** A button to run the query.
- Query results:** A scrollable list of class relationships:
 - Ancestor classes:**
 - Forame_de_Monro_Direito
 - Forames
 - Sistema_Ventricular_Cerebral
 - Thing
 - Super classes:**
 - Forame_de_Monro_Direito
 - Sub classes:**
 - CF_Direito
 - Terceiro_Ventriculo
 - Descendant classes:**
 - CA
 - CF_Direito
 - I
 - Nothing

Fig. 24: Resultado da consulta que apresenta a anatomia do forame de Monro direito.

Com a implementação da formalização concluída, podemos fazer a integração da ontologia-epistemologia com neuroimagens. Na próxima seção, apresentaremos uma prova de conceito que aborda como extrair informações de neuroimagens e inseri-las como indivíduos na teoria formal proposta neste trabalho.

4.7 Individualizando o conhecimento em neuroimagens

Como vimos no capítulo 1, o advento das técnicas de diagnóstico por imagem facilitou o processo de diagnóstico do especialista, principalmente na esquizofrenia, já que as alterações causadas por essa patologia só podiam ser observadas *post mortem*.

Dada a importância das imagens no diagnóstico da esquizofrenia, entendemos que a proposta desse trabalho, de construir uma teoria formal sobre uma das estruturas cerebrais consideradas como de grande importância para o diagnóstico da esquizofrenia, deva discutir formas de integração com informações e conhecimento que possam ser extraídos de neuroimagens.

Nesta seção demonstraremos (por meio de uma prova de conceito), como técnicas de segmentação de imagens podem ser usadas para extrair características, conhecimento de neuroimagens e como esses indivíduos podem ser inseridos na ontologia-epistemologia.

A proposta desta prova de conceito é extrair o volume dos ventrículos laterais em cada imagem, por meio do algoritmo de crescimento de região² e inseri-los na ontologia-epistemologia de acordo com o conceito que representam (conhecimento sobre o domínio ou informação sobre o domínio). Para essa tarefa utilizamos 20 imagens de ressonância magnética cerebral, sendo 10 do grupo de pacientes e 10 do grupo de controle.

O volume extraído de cada neuroimagem representará um indivíduo ontológico, por sua vez esse mesmo indivíduo representa o conceito do grupo em que ele está classificado (pacientes ou controle), logo podemos inserir essas neuroimagens como indivíduos sob duas diferentes abordagens (ontologia e epistemologia). Ainda podemos conhecer a média de volume de cada grupo e se o indivíduo se encontra acima ou abaixo dessa média, de forma a dar maior granularidade ao conhecimento.

Para cada neuroimagem será executada a segmentação em massa branca, cinzenta e líquido cefalorraquidiano através do SPM (um exemplo de segmentação foi apresentado na seção 3.1.3, imagem 7). Essa etapa é realizada para que as imagens que apresentam a segmentação por líquido cefalorraquidiano sejam submetidas a segmentação pelo algoritmo de crescimento de região, já que os ventrículos, como vimos no capítulo 2.4 são cavidades repletas desse líquido.

Após a segmentação pelo líquido cefalorraquidiano, aplicamos o algoritmo de crescimento de região nos cortes onde o sistema ventricular é evidente e como resultado obtemos a quantidade de

²A extração do volume dos ventrículos não é feita utilizando marcações de regiões de interesse (ROI's), por ser uma técnica específica da área médica, utilizamos de forma mais simples o volume total dos ventrículos laterais.

voxels que compõem a região do ventrículo em cada corte (figura 26). O cálculo do volume dos ventrículos é feito somando os voxels que compõem a região em cada corte e multiplicando o número total de voxels pelo valor do volume individual de um voxel (o volume do voxel é obtido no programa MRicro no campo *size(mm)*, destacado na imagem 25).

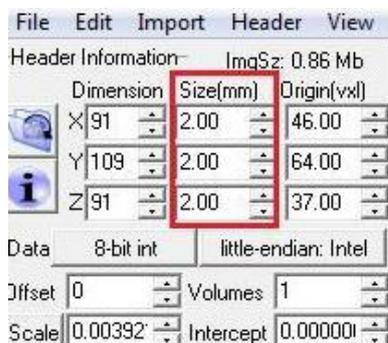


Fig. 25: MRicro - Campos que indicam o tamanho do voxel.

Podemos observar na figura 26(A), a seta que indica o ponto selecionado pelo usuário ao executar o algoritmo para um corte específico e na imagem 26(B), após a execução do algoritmo, a imagem de saída segmentada e um contador que acumula a quantidade de voxels que compõem a região.

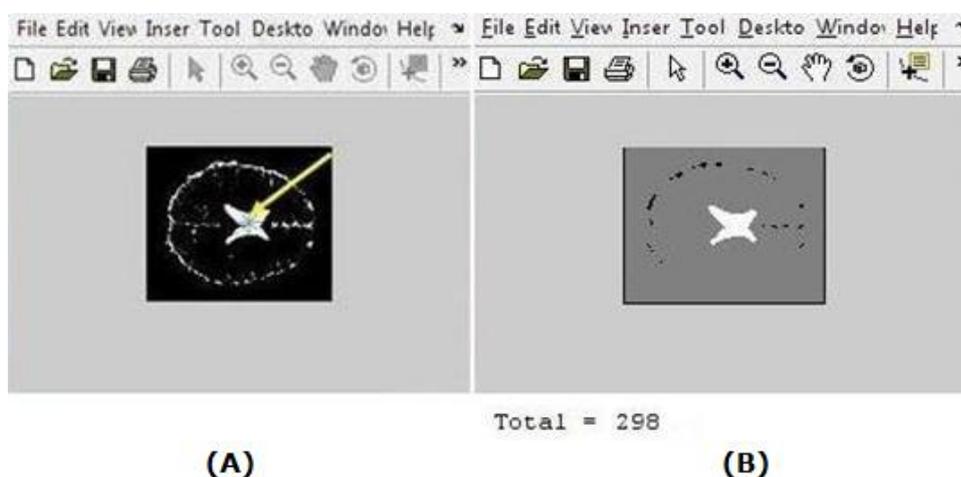


Fig. 26: Seleção do ponto semente durante a execução do algoritmo de crescimento de região (A) e total de pixels após a execução do algoritmo (B).

A tabela 4 apresenta os volumes ventriculares, médias e uma razão da média pelo volume da imagem, que indica os indivíduos que estão acima ou abaixo das médias do seu grupo.

Cada volume ventricular de uma neuroimagem, independente do grupo a que ela pertence (paciente ou controle), é um indivíduo ontológico, pois apresenta uma característica do domínio. Entre-

Volumes do grupo de controle				
Imagem	Total de Voxels	Volume (cm ³)	Média	Razão
bwZC1.img	4681	37448	22565,6	1,659517141
bwZC2.img	4159	33272	22565,6	1,474456695
bwZC3.img	640	5120	22565,6	0,226894033
bwZC4.img	2668	21344	22565,6	0,945864502
bwZC5.img	3097	24776	22565,6	1,097954408
bwZC6.img	1303	10424	22565,6	0,461942071
bwZC7.img	2747	21976	22565,6	0,973871734
bwZC8.img	2531	20248	22565,6	0,897294998
bwZC9.img	4363	34904	22565,6	1,546779168
bwZC10.img	2018	16144	22565,6	0,715425249
Volumes do grupo de controle				
bwZP27.img	2145	17160	23160	0,740932642
bwZP28.img	3375	27000	23160	1,165803109
bwZP29.img	4197	33576	23160	1,449740933
bwZP30.img	3336	26688	23160	1,152331606
bwZP31.img	2878	23024	23160	0,994127807
bwZP32.img	2852	22816	23160	0,985146805
bwZP33.img	3412	27296	23160	1,178583765
bwZP34.img	1900	15200	23160	0,656303972
bwZP35.img	2690	21520	23160	0,929188256
bwZP36.img	2165	17320	23160	0,747841105

Tab. 4: Volumes ventriculares, média e razão do grupo de controle e pacientes.

tanto o mesmo indivíduo quando observado sob a ótica epistêmica não representa mais o domínio, mas sim o conceito “paciente” ou o conceito “controle”. Assim os indivíduos que compõem o grupo de controle e pacientes, podem ser instanciados também em classes epistêmicas adequadas aos conceitos que representam.

A figura 27 apresenta a inclusão de um indivíduo no domínio da ontologia e da epistemologia, mais especificamente, o indivíduo que representa o volume ventricular da neuroimagem de nome bwCZ1 (do lado esquerdo da imagem), instanciado sob a classe ontológica que representa os ventrículos laterais (do lado esquerdo da imagem, representado por *Ventriculos_Laterais*) e sob a subclasse epistêmica que representa que o volume ventricular dessa neuroimagem está abaixo da média do seu grupo (no caso, paciente, dado pela classe *Abaixo_da_Media_Paciente*, no lado direito da imagem).

A relação κ descrita no Protégé pela sentença “*K only Ventriculos_Laterais*”, indica que o indivíduo bwCZ1 traz um conhecimento sobre a classe *Ventriculos_Laterais*, expresso nesse caso pelo grupo a que ele pertence (pacientes).

Assim todas as imagens segmentadas pelo algoritmo de crescimento de região que tiveram seus volumes calculados, apresentadas na tabela 4 são inseridas no Protégé sob a classe que re-



Fig. 27: Indivíduo relacionado a uma classe ontológica e epistêmica.

presenta o ventrículo lateral, e sob as classes epistêmicas que representam seus conceitos (são essas classes: *Abaixo_da_Media_Paciente*, *Abaixo_da_Media_Control*, *Acima_da_Media_Paciente* e *Acima_da_Media_Paciente*).

Essa diferenciação de como classificamos o indivíduo pelo conceito que ele representa é essencial quando pensamos em análises mais realistas aos especialistas sobre uma base de imagens. Por exemplo, um especialista poderia desejar aplicar técnicas de análises estatísticas sobre grupos de pacientes e controles e extrair alterações relevantes a partir de características dos indivíduos que compõem os grupos (como idade, sexo, escolaridade, hábitos, etc). Essas características de indivíduos estariam representadas por uma ontologia modelada para essa finalidade, entretanto a classificação de uma determinada alteração observada em função dessa características, levaria a criação de indivíduos essencialmente epistêmicos.

Este processo entretanto estaria limitado a associação de indivíduos as classes existentes, como abordamos no capítulo 2.3.1 a esquizofrenia afeta diversas estruturas do cérebro, não apenas o sistema ventricular cerebral. Na próxima sessão abordaremos como podemos inserir novas classes na ontologia, descobertas a partir da análise estatística.

4.8 Integrando a ontologia aos dados extraídos das análises estatísticas

Na seção 3.2 apresentamos o PCA+MLDA, que em conjunto com o teste de tamanho efetivo são técnicas que ao serem aplicadas aos grupos de neuroimagens (pacientes e controles), é possível obter como resultado as principais diferenças entre esses grupos.

Conhecer essas diferenças por meio dessas técnicas, significa poder corroborar achados listados

na literatura médica e, até mesmo descobrir novas alterações que eventualmente não sejam perceptíveis e integrar essas descobertas a probabilidades de ocorrência disponíveis na literatura médica.

Essas probabilidades de ocorrência de determinada alteração morfológica são encontradas nas meta-análises, esses trabalhos como já abordamos anteriormente, contém informações de estudos realizados em diferentes grupos de indivíduos e apontam o percentual de ocorrência de alterações nesses grupos.

Utilizaremos como base dos resultados desta seção o trabalho de Santos *et al.* fei09. Em síntese o trabalho propõe a aplicação do PCA+MLDA e testes de tamanho efetivo em um conjunto de 68 imagens, compostas por grupos de pacientes e controles saudáveis. Após a descoberta desses pontos, eles são segmentados por meio do algoritmo de modelos deformáveis, que agrupam pontos próximos compondo uma região na imagem e calcula-se sua pertinência estatística em relação a tabela 1, apresentada na seção 2.3.1.

Ao aplicar o PCA+MLDA no conjunto de amostras se encontra o hiperplano de separação entre as amostras e os extremos dos grupos, como podemos observar na imagem 28.

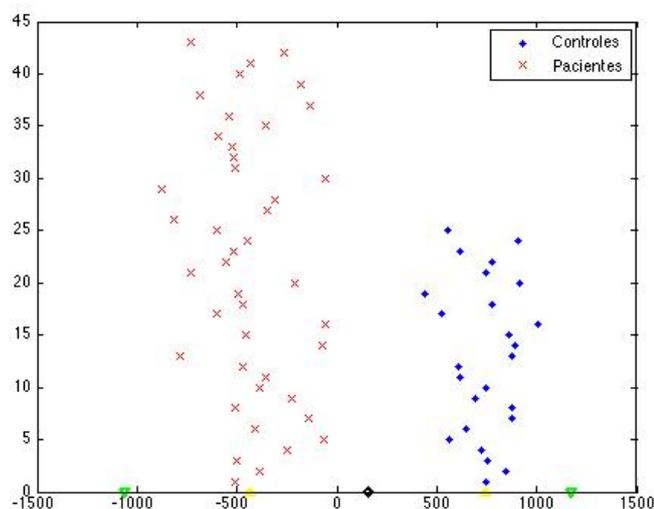


Fig. 28: Projeção no espaço do MLDA das imagens originais. Os pontos verdes correspondem ao extremo de cada grupo dado por três vezes o desvio padrão de cada grupo ($+3\sqrt{\sigma_p}$ e $-3\sqrt{\sigma_c}$), os pontos amarelos correspondem a média de cada grupo, e o ponto preto ao hiperplano de separação.

Fonte: Adaptado de Santos [76]

Após a aplicação do modelo discriminante estatístico (PCA+MLDA), as imagens são reconstruídas levando-se em consideração os extremos, pois essas imagens possuirão a maior diferenças entre

os grupos. Assim, com as imagens reconstruídas é aplicado testes de tamanho efetivo que identifica as diferenças de voxels mais significativas nas imagens. A imagem 29 apresenta os cortes axial, coronal e sagital com as alterações relevantes encontradas.

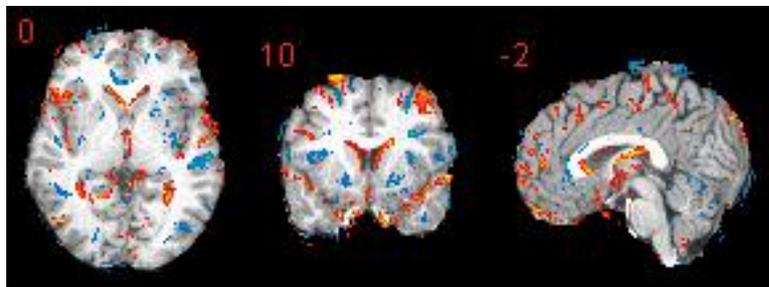


Fig. 29: Pontos identificados pelo testes de tamanho efetivo que representam alterações estatísticas relevantes entre os extremos do grupo de paciente e controle encontrados pelo MLDA para os cortes axial, coronal e sagital.

Fonte: Adaptado de Santos *et al.* [7]

Após a identificação dos pontos relevantes pelos testes de tamanho efetivo, foi aplicada a segmentação por modelos deformáveis, cujo objetivo é obter as relações espaciais entre os pontos descobertos. Em síntese, a segmentação utilizando modelos deformáveis pode agrupar os pontos que pertencem a uma mesma região espacial na imagem. A imagem 30 apresenta um exemplo de um corte de neuroimagem com os pontos mais significativos descobertos pela análise do PCA+MLDA e testes de tamanho efetivo, e a segmentação por meio de modelos deformáveis.

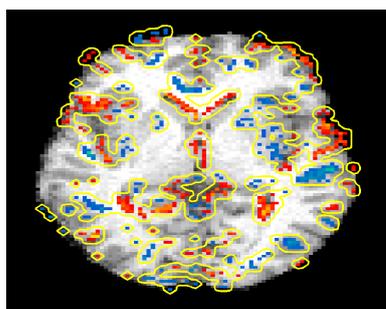


Fig. 30: Segmentação por meio de modelos deformáveis das regiões encontradas pela análise do PCA+MLDA e testes de tamanho efetivo.

Fonte: adaptado de Santos *et al.* [7]

Após as análises estatísticas e segmentação utilizando modelos deformáveis, é necessário representar as estruturas anatômicas correspondentes. Na próxima seção exploraremos como converter os pontos encontrados pelas análises estatísticas em informação sobre regiões cerebrais.

4.8.1 Identificando regiões cerebrais

A transformação de pontos de interesse no espaço das neuroimagens nas estruturas anatômicas correspondentes pode ser feita através de uma transformação do espaço MNI (orientação dos pontos nativa das neuroimagens utilizadas nesse trabalho) para o espaço de Talairach. Para tal transformação utilizamos a aplicação *GingerALE*³ do projeto *Brain Map*, que faz a tradução dos pontos MNI para Talairach automaticamente e gera um arquivo que pode ser carregado no cliente do atlas, que fornece as regiões.

Por exemplo, queremos saber qual região cerebral corresponde a um ponto de interesse em uma imagem MNI, dado que o ponto de origem nas imagens utilizadas nesse trabalho é (46;64;37), e sendo o ponto de origem, sua distância ($i; j; k$) é dada por (0;0;0) e tamanho de cada voxel é de (2;2;2).

Então se quisermos descobrir a distância do ponto (48;72;40) dado em coordenadas (x, y, z), a distância ($i; j; k$) ao ponto de origem (46;64;37) será (4;16;6)*mm*. Essa distância ao ponto de origem é dada também pela aplicação MRicro como podemos observar na figura 31.

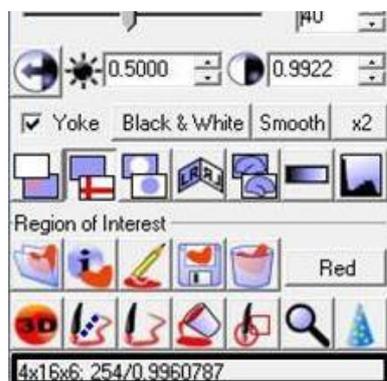


Fig. 31: Destacado na imagem, distância ao ponto de origem dado pelo MRicro.

Esse ponto pode ser carregado na aplicação *GingerALE*, através da opção *Tools - Convert Foci*, escolhendo a opção *MNI(SPM) to Talairach*, apresentada na figura 32.

³Aplicação disponível para download em: <http://www.brainmap.org/icbm2tal/> último acesso em 21/04/2009

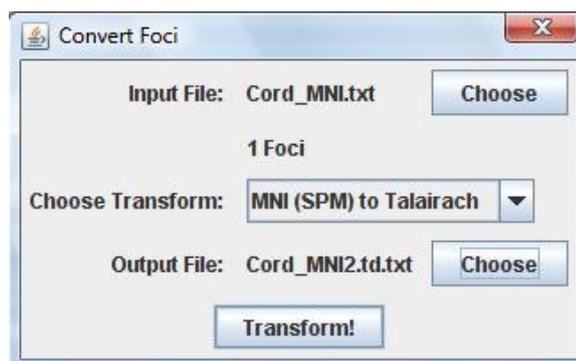


Fig. 32: Interface para conversão de pontos do GingerAle.

Após a transformação, as distâncias no espaço de Talairach obtidas que correspondem à distância MNI (4;16;6) são: (2.65;12.6;10.92), que consultadas no atlas de Talairach apontam as regiões cerebrais as quais esses pontos pertencem, como podemos observar na tabela 5.

Granularidade da informação	Região
Hemisphere level	Right Cerebrum
Lobe level	Sub-lobar
Gyrus level	Lateral Ventricle
Tissue level	Cerebro-Spinal Fluid
Celular level	*

Tab. 5: Regiões apontadas pelo atlas de Talairach para o ponto de exemplo.

O atlas de Talairach apresenta a informação classificada em 5 níveis [77] representados na figura 33. Devido a granularidade, a informação pode não existir em alguns níveis e nesses casos é apresentado um “*”, como podemos perceber nos resultados de nosso exemplo na tabela 5

Então, assim como em nosso exemplo, podemos descobrir a qual região cerebral pertence os pontos encontrados pelas análises estatísticas aplicadas ao grupo de imagens. A próxima seção aborda o resultado dessa conversão de pontos relevantes em regiões cerebrais e a realimentação da ontologia a partir desses pontos.

4.9 Realimentando a ontologia com a expectativa gerada pelas análises estatísticas.

Santos *et al.* [7] apresentam os resultados da aplicação do PCA+MLDA e testes de tamanho efetivo em um grupo de 68 neuroimagens compostas por dois grupos (paciente e controle). Após a

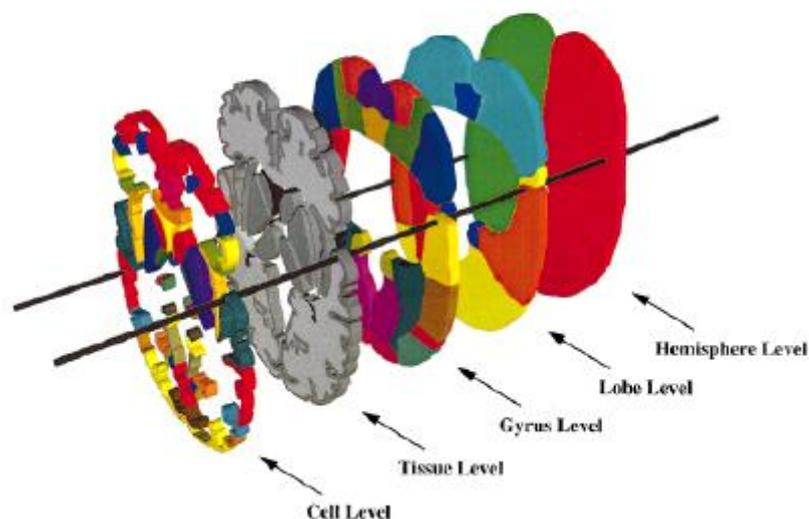


Fig. 33: Níveis do Atlas de Talairach.
 Fonte: Adaptado de Lancaster *et al.* [77]

aplicação das técnicas de análises estatísticas os pontos descobertos como alterações relevantes são segmentados por meio de modelos deformáveis. Para cada região descoberta pela segmentação, é calculada a sensibilidade localmente para determinar quantos pontos nessas regiões fornecem respostas que vão ao encontro de nosso *gold standart* representado pela tabela 1.

Obtemos então, um ranking dos cortes que contém as regiões com maiores índices de sensibilidade, apresentados na tabela 6, e os cortes das neuroimagens equivalentes são apresentadas na figura 34.

As regiões cerebrais correspondentes aos diversos pontos que compõem as regiões segmentadas podem ser descobertas através da conversão dos pontos, conforme apresentamos em nosso exemplo anterior. A tabela 7 apresenta as regiões correspondentes a esses pontos e em quais cortes elas estão evidentes.

Posição	Corte	Sensitividade
1	27	0.36
2	37	0.36
3	33	0.29
4	38	0.29
5	50	0.29
6	38	0.29
7	28	0.29
8	29	0.29
9	44	0.22
10	27	0.22

Tab. 6: Regiões que apresentaram maior índice de sensibilidade local.
Fonte: adaptado de Santos *et al.* [7]

Região cerebral	Corte
Cuneus	38
Lobo frontal	27;28;29;27
Ventrículo Lateral	27;37;33;38;38
Lobo límbico	27;37;33;38;38;28;29
Mesoencéfalo	27;37;28;29
Lobo occipital	37;33;38;50;38;44
Lobo Parietal	50
Precuneus	50
Região sub-giral	37;33;38;50;38;28;29
Região sub lobar	27;37;33;38;38
Lobo temporal	27;37;33;38;50;38;28;29;44;27
Tálamo	37
Giro frontal inferior	27;28;29
Giro lingual	37;33;38
Giro frontal medial	28;29
Giro frontal médio	27;28
Giro occipital médio	38
Giro temporal médio	37;50;38;44
Giro parahipocampal	27;37;33;38;28;29
Giro subcaloso	29
Giro temporal superior	27;50;28;29;44;27

Tab. 7: Regiões cerebrais encontradas nos cortes que apresentam mais alta sensibilidade.

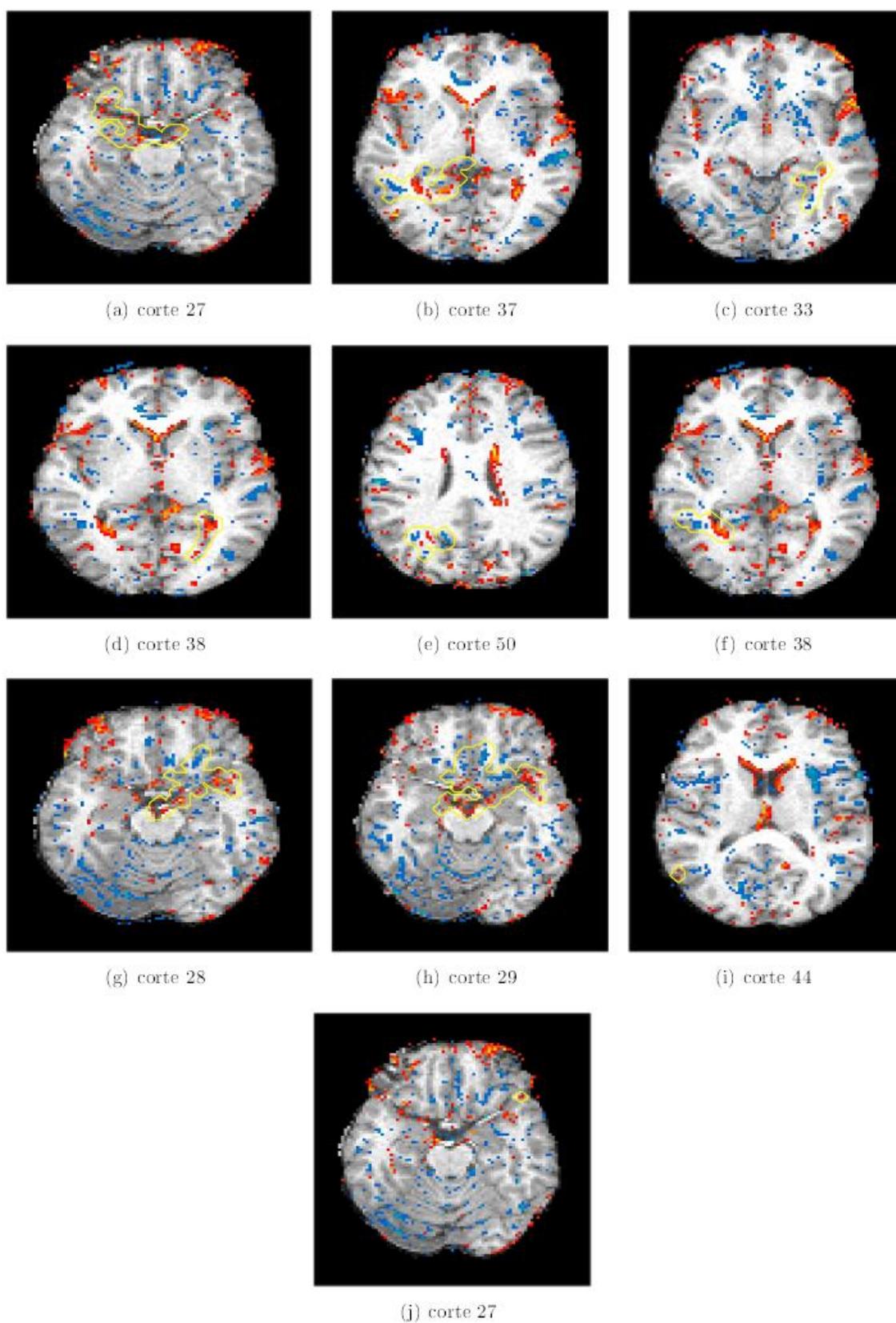


Fig. 34: Cortes que contêm as regiões com maiores valores de sensibilidade de acordo com a tabela 6.

Fonte: adaptado de Santos *et al.*[7]

Assim, as regiões encontradas e apresentadas na tabela 7 podem ser utilizadas para realimentar a ontologia. Cada região descoberta é inserida como uma nova classe ontológica. A figura 35 apresenta as regiões identificadas, inseridas como classes na ontologia. A inserção de novas classes na ontologia permite que novos conceitos epistêmicos sejam inseridos, novas formalizações mereológicas sejam inseridas na ontologia, e conseqüentemente permite que uma gama maior de consultas à base de conhecimento possam ser realizadas.



Fig. 35: Estruturas cerebrais descobertas pela análise estatística inseridas no Protégé.

Na próxima seção veremos como incluir o conhecimento encontrado nas meta-análises na ontologia epistemologia, baseado nas regiões descobertas pelas análises estatísticas.

4.10 Incluindo probabilidades encontradas na literatura

A proposta dessa seção é integrar os resultados das análises estatísticas com resultados das meta-análises, o que nos permite criar uma base de conhecimento que pode ser realimentada com novas alterações oriundas das análises estatísticas e integrada com o conhecimento presente na literatura médica.

Para representar os percentuais de ocorrência das alterações morfológicas presentes na literatura, são criadas classes epistêmicas que representam esses valores e são relacionadas (como vimos neste capítulo), com a ontologia por meio da relação κ . A tabela 8 representa os valores apresentados na seção 2.3.1 de porcentagem de alterações e valores de redução.

Estrutura	% de ocorrência	Média	Fonte
Ventrículos laterais	80%	(+)40%	[1], [35]
Terceiro ventrículo	73%	(+)25.3%	[1], [38]
Quarto ventrículo	20%	N/C	[1]
Lobo temporal	61%	(-)6%(LTE)*	[1],
	55%	(-)9.5%(LTD)**	[35]
Complexo amígdala hipocampo	74%	(-)6,50%	[1],
		(-)8%	[35],
			[30]
Amígdala	N/C	(-)10%	[35]
Hipocampo	N/C	(-)7%	[1],
		(-)8,5%	[35]
		(-)4%	
Giro parahipocampal	N/C	(-)14%(GPD)***	[35]
		(-)9% (GPE)****	
Giro temporal superior	67%	(-)15%	[1], [35]
Talámo	42%	(-)6%	[1],
		(-)7%	[35]
Gânglios da base	68%	N/C	[1]
Lobos frontais	60%	(-)5,50%	[1],
		N/C	[35]
Lobos occipitais	44%	(-)6,5%	[1],
		(-)8,5%	[35]
Lobos Parietais	60%	N/C	[1]

Tab. 8: Estruturas alteradas pela esquizofrenia e valores de alteração encontrados nas meta-análises. A segunda coluna apresenta as porcentagens de ocorrência e a terceira coluna a média das alterações, o sinal (+) representa aumento volumétrico, o sinal (-) diminuição. (Os asteriscos representam: (*) - Lobo temporal esquerdo, (**) LTD - lobo temporal esquerdo, (***) GPD - Giro parahipocampal direito, (****) GPE - Giro parahipocampal esquerdo)

A figura 36 mostra a classe epistêmica que representa as alterações percentuais do ventrículo lateral apresentadas na tabela 8. Foi atribuído a essa classe um indivíduo de exemplo, que se relaciona com sua correspondente ontológica pela relação κ .

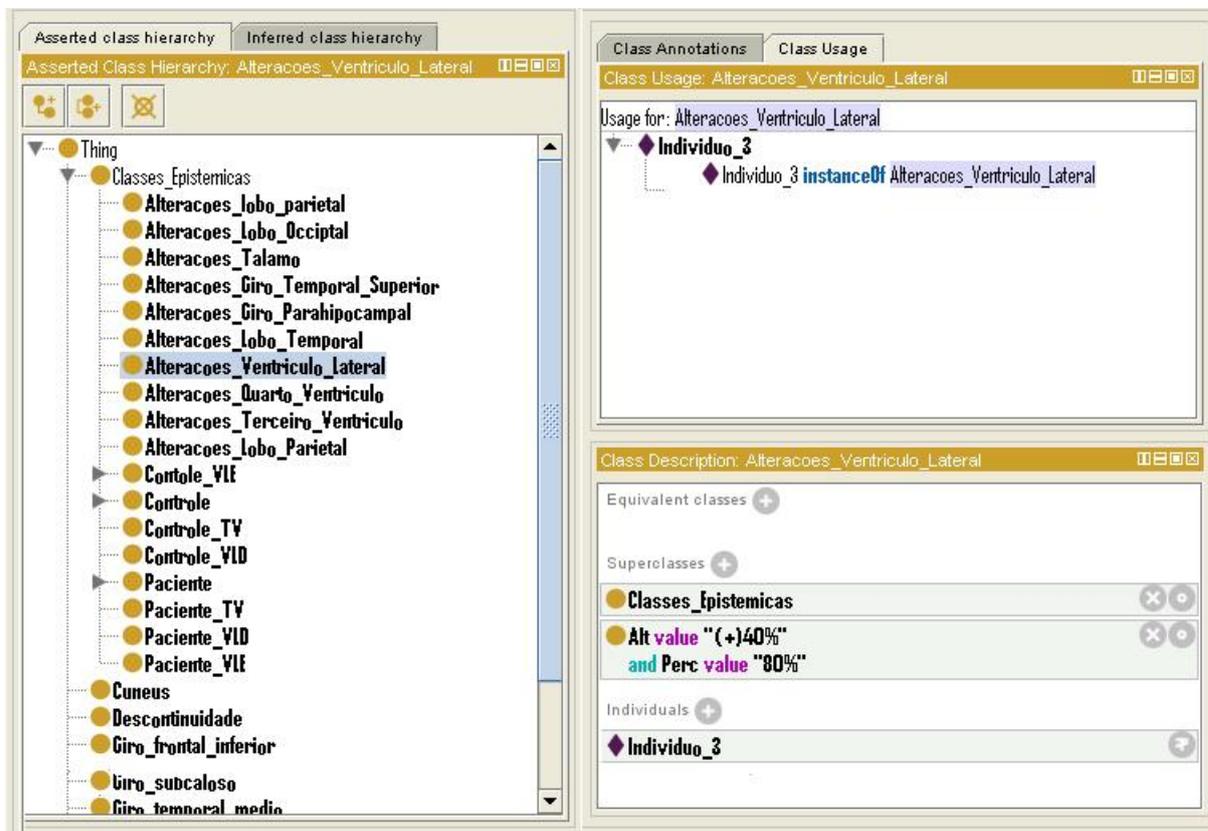


Fig. 36: Classe epistêmica que representa um percentual de ocorrência e redução dado pela literatura.

Assim como para o exemplo, as probabilidades listadas na tabela 8 são inseridas no Protégé como classes epistêmicas, e nas suas propriedades descrito os valores encontrados nas meta-análises. Ao inserir um indivíduo de exemplo, ele é relacionado com a ontologia pela relação κ , abordada anteriormente na seção 4.7

A figura 37 apresenta o resultado de uma consulta sobre os valores percentuais inseridos na imagem 36. Foi relacionado o indivíduo de exemplo (*individuo_3*) com a classe que representa os valores de alteração encontrados na literatura. A consulta é sobre os valores de alteração de 40% de aumento de volume e que seja um conhecimento relacionado aos ventriculos laterais, e o resultado obtido é o indivíduo de exemplo que foi relacionado a classe.

Query (class expression)

Alt value "(+)40%" and K only Ventriculos_Laterais

Execute

Query results

Equivalent classes

Ancestor classes

- Thing

Super classes

- Thing

Sub classes

- Nothing

Instances

- ◆ Individuo_3

Fig. 37: Consulta que apresenta o indivíduo relacionado com a classe epistêmica que representa valores percentuais de alteração encontrados na literatura e sua correspondente na ontologia.

4.10.1 Considerações Finais

Neste capítulo apresentamos os resultados desta dissertação, iniciamos com uma formalização do sistema ventricular cerebral em lógica de primeira ordem. Os resultados propostos nesta seção permitem que a formalização seja facilmente ampliada ou utilizada em conjunto com outras teorias, bem como estendida para representação de outros conceitos.

Apresentamos também uma solução para a representação de conhecimento (epistemologia) junto a ontologia. Neste trabalho nos inspiramos na lógica epistêmica [75] através da relação κ , que nos permitiu instanciar e trabalhar com conceitos epistêmicos de forma independente da ontologia.

Após as representações da ontologia-epistemologia, implementamos a teoria formal no Protégé, as classes (ventrículos laterais, terceiro ventrículo, etc), sub-classes (corno frontal do ventrículo lateral esquerdo e direito, massa intertalâmica, recesso lateral, etc) e os limites fiat que delimitam cada estrutura ($Z_1, Z_2 \dots Z_{21}$). Implementamos também as classes epistêmicas e subclasses epistêmicas, que permitiram a inserção de conhecimento extraído de técnicas de análise de imagem.

A implementação no Protégé permitiu a execução de consultas que demonstrem a validade da representação proposta, nessa etapa obtivemos resultados corretos, tanto nas consultas realizadas sobre conceitos ontológicos, como as consultas realizadas sobre conceitos epistêmicos.

A integração com neuroimagens foi feita demonstrando como podemos extrair conhecimento por meio de técnicas de segmentação de imagem (em nosso trabalho, o algoritmo de crescimento de região com interação do usuário). Também apresentamos como técnicas de análise estatística (PCA+MLDA e testes de tamanho efetivo), que combinadas com a segmentação por modelos deformáveis e a identificação dos pontos em regiões cerebrais que permitiram a inserção de novas classes na ontologia.

A partir da expansão da ontologia, foi consolidada a tabela 8 com informações encontradas na literatura médica, e inseridas na teoria formal, por meio de classes epistêmicas, o que permite a criação de uma base de conhecimentos que integra análises de um grupo de imagens com a literatura disponível sobre esquizofrenia. No próximo capítulo apresentaremos as conclusões e trabalhos futuros propostos para o tema desta dissertação.

CAPÍTULO 5

CONCLUSÕES E TRABALHOS FUTUROS

Em um primeiro momento, esse trabalho se focou em investigar como formalizar conceitos de neuroanatomia por meio de uma ontologia espacial e compreender como inserir o conhecimento presente na literatura médica nessa formalização. Para tal, estudamos os aspectos de construção de ontologias computacionais e características de alterações morfológicas causadas pela esquizofrenia.

Durante nossos estudos, o sistema ventricular cerebral se mostrou uma das mais importantes alterações morfológicas causada pela esquizofrenia, listada em diversas literaturas médicas, o que nos fez escolher essa estrutura anatômica como objeto da ontologia proposta nesse trabalho. Dado esse fato, compreendemos a anatomia do sistema ventricular cerebral e como poderíamos representar sua estrutura contínua e uniforme, que conta com diversas divisões anatômicas, tarefa que pode ser realizada por meio de limites fiat. Esses conceitos foram apresentados no capítulo 2.

Durante o estudo das diversas ontologias biomédicas que permitem a criação de uma teoria com base mereológica consistente, duas alternativas se mostram mais adequadas, a teoria básica de inclusão (BIT) e a ontologia formal básica (BFO). A escolha da BIT para essa representação se deu devido a sua extensão chamada de BIT+CI, que consiste em uma teoria de classes que permite uma representação consistente da entidade de interesse, além de outras características como relações específicas para regiões espaciais, como a relação de coincidência parcial (*PCoin*) e localizado em (*LocIn*). A BIT e considerações sobre sua escolha foram apresentadas no capítulo 3.

No desenvolvimento dos resultados deste trabalho encontramos a necessidade de expandir a BIT,

de forma a compreender conceitos de continuidade e limites fiat. Propomos a extensão por meio das relações *LFiat*, *Desc*, *PCont* e *Segment* (conceituadas na seção 4.2). Essas propostas também são listadas como trabalhos futuros do artigo principal que introduz a BIT[10].

Entretanto, a formalização de limites entre estruturas pode ser estendida. Em nosso trabalho tivemos a necessidade de representar limites fiat, que como conceituamos na seção 2.5 são utilizados para delimitar regiões ou objetos que tem seus limites aceitos de forma consensual. Porém existem objetos (em nosso caso, estruturas anatômicas) que tem limites concretos (por exemplo, a margem de um rio), esse tipo de representação é feito pelos limites *bona fide*, então uma formalização mereológica das demais estruturas anatômicas identificadas em neuroimagens levaria a expansão da teoria criando-se relações que representem *limites bona fide*.

A BIT se mostrou expressiva para a formalização dos conceitos apresentados nesse trabalho, entretanto a ausência de algumas relações como parte tangencial (*TP*), externamente conectado (*EC*) não permitiram uma formalização mais abrangente dos limites fiat. As alternativas para essa formalização são a expansão das relações mereológicas da BIT ou a utilização de outra ontologia como a BFO [24], porém levaria a necessidade da criação de uma teoria de classes expressiva como a da BIT.

Porém a ontologia não é, conceitualmente, a teoria formal usada para representação de conhecimento de fatos encontrados na literatura. Buscamos então, alternativas de representação desses fatos da literatura na epistemologia. Então, nossa proposta inicial era construir uma formalização em lógica de primeira ordem que permitisse a implementação dessa teoria independente de ferramentas, e escolhemos o Protégé para os testes de consistência e consultas sobre a formalização proposta. Como vimos na seção 3.4, o Protégé tem o uso difundido para ontologias biomédicas, tem uma documentação de fácil acesso e conta com grande aceitação da comunidade.

Encontramos obstáculos na formalização utilizando o Protégé no que tange as representações epistêmicas, sendo o Protégé uma ferramenta para implementação de ontologias, a herança de indivíduos por classes é obrigatória. A alternativa foi a criação de classes epistêmicas separadas das suas respectivas classes ontológicas. Nessa solução, a relação κ se mostrou indispensável para relacionar o conceito e o indivíduo a sua classe na ontologia.

Outra alternativa de implementação seria a utilização do Power Loom [15] e da linguagem Prolog, que poderia se mostrar uma alternativa viável para implementação de regras mais claras. Porém a cogitação de outra ferramenta para a formalização dos conceitos apresentados foi descartada devido a complexidade em realizar uma comparação entre diversas plataformas, bem como o tempo de

implementação, o que entendemos que caracteriza um outro trabalho.

Ainda como proposta inicial, demonstramos a integração com o conhecimento extraído de neuroimagens por meio de segmentação de imagens. A seção 4.7 apresenta uma prova de conceitos sobre essa integração, demonstrando como um dado extraído de uma neuroimagem pode representar um indivíduo na ontologia e na epistemologia, e a seção 3.1 abordamos o algoritmo de crescimento de região, usado neste trabalho para extração de dados relevantes em neuroimagens.

A prova de conceitos tinha por objetivo demonstrar como podemos inserir conhecimento extraído de neuroimagens. O objetivo foi alcançado, porém inicialmente a proposta consistia em trabalhar os volumes como os cálculos executados em meta-análises, entretanto, não encontramos uma literatura específica que elucidava as técnicas de delineamento das regiões de interesse e entendendo que essa é uma tarefa do especialista, simplificamos a prova de conceitos de forma a trabalhar com todo o volume do ventrículo lateral.

Durante o desenvolvimento desse trabalho foi identificado a possibilidade de inserir não só informações pela prova de conceitos proposta, que consiste em uma segmentação simples, mas também através de técnicas de análise estatística. Esses resultados são apresentados na seção 4.9 e se baseiam na pesquisa de mestrado do aluno Danilo Nunes Silva, que gerou o artigo *Exploring the knowledge contained in neuroimages: statistical discriminant analysis and automatic segmentation of the most significant changes* [9]. Assim, abordamos o PCA+MLDA e testes de tamanho efetivo na seção 3.2.

A importância da integração da ontologia-epistemologia com técnicas como o MLDA e testes de tamanho efetivo, está na descoberta de novas regiões que podem realimentar a ontologia, o que prevê formas de comparação das análises entre as amostras com a literatura médica disponível.

Ainda durante o desenvolvimento da teoria formal, percebemos que a representação do conhecimento por meio da epistemologia era uma tarefa pouco abordada nos trabalhos da área. A alternativa de representação proposta nesse trabalho na seção 4.5, foi inspirado na lógica epistêmica, proposta no trabalho de Tabet [75].

A alternativa foi através da utilização da relação κ , que em nosso trabalho significa que um indivíduo contém conhecimento sobre uma classe ontológica. A utilização da lógica epistêmica se mostrou uma alternativa viável e foi de grande importância para representações epistêmicas, entretanto a utilização da lógica epistêmica pode ser expandida, levando em consideração as implicações sobre suas relações. Entretanto devido a complexidade dessa tarefa ela não foi realizada neste trabalho, e enten-

demos que deve ser explorada adequadamente em uma pesquisa a parte.

A implementação da epistemologia encontrou também a limitação do Protégé para implementação da herança diferenciada entre classes proposta nesse trabalho. Sendo baseado em uma linguagem criada para representação ontológica, a característica de subsunção da classe epistêmica à ontológica de não permitir a herança do indivíduo classificado no domínio da ontologia (fórmula 4.92), não pode ser implementado como desejado. A alternativa foi a criação das classes epistêmicas separadas das ontológicas integradas por meio da relação κ , como abordamos no capítulo 4.6.1. Nessa etapa mais uma vez a formalização inspirada na lógica epistêmica foi de extrema importância para representação dos conceitos no Protégé.

A integração com as regiões encontradas pelas análises estatísticas e segmentação das regiões por meio de modelos deformáveis, proposta na seção 4.7, consiste em descobrir a região cerebral por meio da conversão dos pontos encontrados como alterações significativas, do espaço das imagens para coordenadas do atlas Talairach. Esses pontos convertidos são os que apresentaram mais alto valor de sensibilidade (0.71), o que significa que 71% dos destes pontos estão cobertos na literatura, que tem como referência a tabela 1 [7].

Com os pontos identificados, as regiões descobertas foram usada para realimentar a ontologia, o que garante uma importante característica citada na seção 2.1, a extensibilidade. A próxima etapa consistiu em representar no domínio epistêmico os valores percentuais de alteração da estrutura, encontrados nas meta-análises.

A representação desses valores é uma contribuição importante desse trabalho, pois ao descobrirmos as regiões que representam as alterações descobertas pelas análises estatísticas, podemos integrar esse conhecimento a valores de probabilidade de ocorrência exhaustivamente abordados na literatura médica por meio das meta-análises. Existia então, a necessidade de prover uma forma de representação para esse conhecimento e demonstrar como ele poderia ser integrado a ontologia-epistemologia.

A forma de como representar o conhecimento encontrado em meta-análises foi por meio de classes epistêmicas, que tem valores descritos em suas propriedades, essa representação, assim como a própria abordagem da epistemologia na representação trabalho evita a confusão Kantiana, como abordamos na seção 2.2

Concluimos que esse trabalho proveu formas de representação do domínio e do conhecimento por meio da integração da ontologia com a epistemologia. Na ontologia, propomos a extensão da BIT de forma a prover uma representação adequada de limites fiat e continuidade e utilizamos relações de

tamanho e a teoria de classes para uma representação consistente do sistema ventricular cerebral.

Na integração com a epistemologia abordamos a alternativa da lógica epistêmica para representação de conhecimento e demonstramos como podemos inserir conhecimento presente nas meta-análises e extraídos de técnicas de análise e segmentação de imagens.

As consultas realizadas sugerem a validade dessa proposta. Todas as consultas apresentaram resultados corretos, o que permite que a formalização proposta seja utilizada como base de conhecimento para análise da relevância dos pontos encontrados e conhecer qual os valores de ocorrência dessas alterações.

Sendo essa a primeira modelagem com o objetivo de formalizar o sistema ventricular como entidade e dar base a formalização epistêmica do conhecimento sobre esquizofrenia, vemos uma gama de possibilidades para trabalhos futuros:

1. Estudo mais detalhado da utilização de predicados da lógica epistêmica na ontologia-epistemologia como indicador da inserção de conhecimento.

A separação entre ontologia epistemologia nesse trabalho, foi desenvolvida inspirado na proposta encontrada em [75], entretanto, não foi amplamente discutida o que traz a possibilidades de um estudo mais aprofundado sobre sua implicações e possibilidades de utilização em conjunto com a ontologia.

Entendemos que essa é uma parte importante, principalmente na para ontologias biomédicas, pois nem sempre se deseja ter apenas a formalização do domínio, mas também do conhecimento que pode ser extraído das análises e da literatura. A adoção da lógica epistêmica como forma de representação de conhecimento integrada a ontologia deverá levar em consideração suas implicações lógicas e verificação da consistência da teoria.

2. Extensão da modelagem ontológica

A expansão da modelagem ontológica visa formalizar todo o cérebro fazendo uso da mereologia. Para tal é necessário estudar como podemos representar limites *bona fide* pois nem toda estrutura cerebral é composta por limites consensuais, mas também por limites físicos concretos.

Também faz-se necessário a formalização das características e relacionamentos entre as estruturas da anatomia cerebral, o que permite que sejam executadas consultas sobre a localização e relacionamento entre as diversas estruturas anatômicas do cérebro.

3. Modelagem de características de grupos na ontologia-epistemologia

Modelar na ontologia características sobre os grupos de indivíduos (idade, sexo, hábitos, etc), permite inferir quais alterações encontradas pelas análises estatísticas ou de segmentação de imagem são comuns a determinadas características e fazer comparações com novas descobertas ou literatura. Através do processo de inferência deverá ser possível descobrir inclusive novas características que sejam comuns a determinados grupos ou alterações.

4. Ontologia da Imagem

A ontologia da imagem [78] consiste em criar uma ontologia para a representação do domínio que vemos na imagem e conseguir diferenciar regiões, e prover formas de interligá-la com a ontologia do objeto possibilitando inferências mais complexas.

Por exemplo, uma região de pontos agrupadas pelo MLDA nem sempre representa alterações em apenas uma região cerebral, esses pontos podem representar alterações entre duas ou mais regiões. A ontologia da imagem poderia prover formas de descrever essa região de pontos como uma região formada por partes de duas ou mais regiões, permitindo inferências específicas para determinados cortes da imagem de ressonância magnética.

REFERÊNCIAS

- [1] SHENTON, M. E. et al. A review of mri findings in schizophrenia. *Schizophrenia Research*, v. 49, n. 1-52, p. 1–52, 2001.
- [2] CASTRO, C. C. Ressonância magnética na esquizofrenia, um estudo morfométrico. *Radiol Bras*, v. 34, n. 3, p. 141–144, 2001.
- [3] OLABARRIAGA, S. D.; SMEULDERS, A. M. W. Interaction in the segmentation of medical images: A survey. *Medical Image Analysis*, v. 5, p. 127–142, 2001.
- [4] RAZZOUK, D. et al. How do experts recognize schizophrenia: the role of the disorganization symptom. *Revista Brasileira de Psiquiatria*, v. 28, n. 1, p. 119–128, 2006.
- [5] RAZZOUK, D.; SHIRAKAWA, I.; MARIA, J. de J. Sistemas inteligentes no diagnóstico da esquizofrenia. *Revista Brasileira de Psiquiatria*, v. 22, p. 0, 2000.
- [6] KONONENKO, I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in Medicine*, v. 23, n. 1, p. 89–109, 2001.
- [7] SANTOS, P. E. et al. Exploring the knowledge contained in neuroimages: statistical discriminant analysis and automatic segmentation of the most significant changes. 2009.
- [8] FONSECA, F.; EGENHOFER, M.; BORGES, K. A. Ontologias e interoperabilidade semântica entre sigs. In: *Workshop Brasileiro de Geoinformática*. [S.l.: s.n.], 2001.
- [9] BITTNER, T.; GOLDBERG, L. J.; DONNELLY, M. Ontology and qualitative medical images analysis. In: *Interdisciplinary Ontology Conference*. [S.l.: s.n.], 2008. v. 1.
- [10] DONNELLY, M.; BITTNER, T.; ROSSE, C. A formal theory for spatial representation and reasoning in biomedical ontologies. *Artificial Intelligence in Medicine, Elsevier Intell Med*, v. 36, n. 1, p. 1–27, January 2006. ISSN 0933-3657.

- [11] SMITH, B. Blackwell guide to the philosophy of computing and information. In: _____. [S.l.]: Oxford: Blackwell, 2003. cap. Ontology, p. 155–166.
- [12] USCHOLD, M.; GRÜNINGER, M. Ontologies: principles, methods, and applications. *Knowledge Engineering Review*, v. 11, n. 2, p. 93–155, 1996.
- [13] BORST, P.; AKKERMANS, H.; TOP, J. Engineering ontologies. *International Journal of Human-Computer Studies*, v. 46, p. 365–406, 1997.
- [14] GRUBER, T. Encyclopedia of database systems. In: _____. [S.l.]: Springer, 2008. cap. Ontology, p. –.
- [15] SANCHES, A. C. D. *Construção e estudos sobre ontologias de bioquímica*. Dissertação (Mestrado) — Universidade de São Paulo, 2005.
- [16] CASARE, S. J. *Uma ontologia funcional de reputação de agentes*. Dissertação (Mestrado) — Universidade de São Paulo, 2005.
- [17] PÉREZ, A. G.; BENJAMINS, V. R. Overview of knowledge sharing and reuse components: Ontologies and problem-solving methods. In: *In*. [S.l.: s.n.], 1999.
- [18] GUARINO, N. Formal ontology and information systems. In: *Formal Ontology in Information Systems*. [S.l.: s.n.], 1998.
- [19] BODENREIDER, O.; MITCHELL, J. A.; MCCRAY, A. T. Biomedical ontologies. *Pacific Symposium on Biocomputing*, v. 0, p. 89–92, 2005.
- [20] ROSSE, C.; MEJINO, J. L. V. A reference ontology for biomedical informatics: the foundational model of anatomy. *J. of Biomedical Informatics*, Elsevier Science, San Diego, USA, v. 36, n. 6, p. 478–500, December 2003. ISSN 1532-0464.
- [21] RECTOR, A. L.; ROGERS, J.; POLE, P. A. The galen high level ontology. In: *Proc of MIE 96*. [S.l.]: IOS Press, 1996. p. 174–178.
- [22] RECTOR, A. L.; ROGERS, J. E. Ontological issues in using a description logic to represent medical concepts: Experience from galen. In: *IMIA Working Group 6 Workshop*. [S.l.: s.n.], 1999.
- [23] SMITH, B. et al. Relations in biomedical ontologies. *Genome Biology*, -, p. –, 2005.
- [24] GRENON, P. *BFO in a Nutshell : A Bi-categorical Axiomatization of BFO and Comparison with DOLCE*. [S.l.], 2003.
- [25] SMITH, B.; CEUSTERSB, W. Ontology as the core discipline of biomedical informatics; legacies of the past and recommendations for the future direction of research. In: *Computing, Philosophy, and Cognitive Science - Cambridge Scholars Press*. [S.l.: s.n.], 2006.
- [26] CAMARA, G.; M.V.MONTEIRO, A.; MEDEIROS, J. S. Representações computacionais do espaço: Um diálogo entre a geografia e a ciência da geoinformação. In: *Workshop sobre novas tecnologias em ciências geograficas*. [S.l.: s.n.], 2000.

- [27] BITTNER, T. On ontology and epistemology of rough location. *Lecture Notes In Computer Science*, v. 1661, p. 433 – 448, 1999.
- [28] BODENREIDER, O.; SMITH, B.; BURGUN, A. The ontology-epistemology divide: A case study in medical terminology. *Proceedings of FOIS*, v. 1, p. 185–195, 2004.
- [29] CEUSTERS, W.; SMITH, B. Referent tracking for treatment optimisation in schizophrenic patients: A case study in applying philosophical ontology to diagnostic algorithms. *Web Semantics: Science, Services and Agents on the World Wide Web*, v. 4, p. 229–236, 2006.
- [30] SALLET, P. C. *As Esquizofrenias segundo a Classificacao das Psicoses Endogenas de Karl Leonhard and sua Correlacao com Imagens Cerebrais por meio de Ressonancia Magnetica (MRI)*. Tese (Doutorado) — Universidade de São Paulo, 2002.
- [31] HARRISON, P. J. The neuropathology of schizophrenia. a critical review of the data and their interpretation. *Brain Oxford Journals*, v. 122, n. 4, p. 593–624, 2005.
- [32] ELKIS, H. et al. Neuroimagem estrutural and psicopatologia: sintomas positivos and negativos and dilatacao ventricular na esquizofrenia. *Revista Brasileira de Psiquiatria*, v. 23, p. 19–23, 2001.
- [33] SILVA, A.; CORRÊA, M. J. U.; BASTOS, A. M. Sistema ventricular: anatomia e patologia aplicadas ao diagnóstico por imagem. *Jornal Brasileiro de neurocirurgia*, v. 2, p. 60–65, 2003.
- [34] JÚNIOR, E. A.; YAMASHITA, H. Aspectos basicos de tomografia computadorizada e ressonancia magnetica. *Revista Brasileira de Psiquiatria*, v. 23, p. 0, 2001.
- [35] LAWRIE, S. M.; ABUKMEIL, S. S. Brain abnormality on schizofrenia, a systematic and quantitative review of volumetric magnetic ressonance imaging studies. *The British Journal of Psychiatry*, v. 172(2), n. 9, p. 110–120, 1998.
- [36] HONEA, R. et al. Regional deficits in brain volume in schizophrenia:a meta-analysis of voxel-based morphometry studies. *The American Journal of Psichiatria*, v. 162, n. 3, p. 2233–2245, 2005.
- [37] OKUBO, Y.; SAIJO, T.; ODA, K. A review of mri studies of progressive brain changes in schizophrenia. *Journal of Medical and Dental Sciences*, v. 48, n. 48, p. 61–67, 2001.
- [38] STEEN, R. G. et al. Brain volume in first-episode schizophrenia. *Br J Psychiatry*, v. 188, p. 510–518, 2006.
- [39] SHENTON, M. E. et al. Abnormalities of the left temporal lobe and thought disorder in schizophrenia, a quantitative magnetic resonance review. *The New England Journal of Medicine*, v. 327, n. 9, p. 604–612, 1992.
- [40] VITA, A. et al. Cerebral ventricular enlargement as a generalized feature of schizophrenia: a distribution analysis on 502 subjects. *Schizophrenia Research*, v. 44, n. 44, p. 25–34, 2000.
- [41] CROW, T. J. Assimetria cerebral e lateralização da linguagem: deficits nucleares na esquizofrenia como indicadores de predisposicao genetica. *Revista de Psiquiatria do Rio Grande do Sul*, v. 26, p. 122–134, 2001.

- [42] MACIEL, E. P. *Esclerose múltipla: correlação clínica, líquido cefalorraquidiano e neuroimagem*. Tese (Doutorado) — Faculdade de Ciências Médicas, da Universidade Estadual de Campinas, 2002.
- [43] MACHADO, A. *Neuroanatomia Funcional*. [S.l.]: ATHENEU EDITORA, 2003.
- [44] VARZI, A. C. Boundaries, continuity, and contact. In: *Time, Space and Movement: Meaning and Knowledge in the Sensible World (Proceedings of the 5th International Workshop*. [S.l.: s.n.], 1995. p. 79–100.
- [45] SMITH, B.; VARZI, A. C. Fiat and bona fide boundaries. In: *Philosophy and Phenomenological Research*. [S.l.: s.n.], 2000. p. 401–420.
- [46] FIELDING, J. M.; MARWEDE, D. The image as spatial region: Location and adjacency within the radiological image. *Formal ontology in information systems*, IOS Press, v. 150, p. 89–100, 2006.
- [47] HINSHAW, K. P.; ALTMAN, R. B.; BRINKLEY, J. F. Shape-based models for interactive segmentation of medical images. In: *Image processing*. [S.l.]: SPIE, 1995. p. 771–780.
- [48] PHAM, D. L.; XU, C.; PRINCE, J. L. A survey of current methods in medical image segmentation. In: *Annual Review of Biomedical Engineering*. [S.l.: s.n.], 2000.
- [49] OLABARRIAGA, S. D. et al. An intelligent interactive segmentation method for the joint space in osteoarthritic ankles. *Lecture Notes in Computer Science*, Volume 1613, p. 394–399, 1999.
- [50] XIAOHAN, Y.; YLA-JAASKI, J. Interactive surface segmentation for medical images. In: *Proceedings of 3rd International Conference on Signal Processing*. [S.l.: s.n.], 1996.
- [51] ADAMS, R.; BISCHOF, L. Seeded region growing. *IEEE Transactions Pattern Analysis and Machine Intelligence*, v. 16(6), p. 641–647, 1994.
- [52] MAT-ISA, N. A. et al. Seeded region growing features extraction algorithm; its potential use in improving screening for cervical cancer. *International Journal of the Computer the Internet and Management (IJCIM)*, v. 13(1), p. 61–70, 2005.
- [53] BASSITT, D. P. et al. Insight and regional brain volumes in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, v. 1, p. 5862, 2007.
- [54] MECHELLI, A. et al. Voxel-based morphometry of the human brain: Methods and applications. *Current Medical Imaging Reviews*, v. 1, p. 1–9, 2005.
- [55] ASHBURNER, J. et al. *SPM5 Manual, The FIL Methods Group*. [S.l.], 2006.
- [56] OLIVEIRA, S. H. A. *Análise estatística multivariada de imagens cerebais de controle e pacientes com transtorno obsessivo compulsivo*. Dissertação (Mestrado) — Centro Universitário da FEI, 2007.

- [57] AGUIAR, N. A. O. *Comparação entre as análises estatísticas univariada e multivariada no diagnóstico da doença cerebral de Alzheimer*. Dissertação (Mestrado) — Centro universitário da FEI, 2007.
- [58] JENKINSON, M.; PECHAUD, M.; SMITH., S. Mr-based estimation of brain, skull and scalp surfaces. In: *Eleventh Annual Meeting of the Organization for Human Brain Mapping*. [S.l.: s.n.], 2005.
- [59] THOMAZ, C.; KITANI, E.; GILLIES, D. A maximum uncertainty lda-based approach for limited sample size problems - with application to face recognition. *Proceedings of SIBGRAPI 05, Journal of Brazilian Computer Society (JBACS) IEEE CS Press*, v. 12, p. 89–96, 2005.
- [60] R.A.FISHER. The use of multiple measures in taxonomic problems. *Annals of Eugenics - Cambridge University Press*, v. 7, p. 179–188, 1936.
- [61] KITANI, E. K. *ANÁLISE DE DISCRIMINANTES LINEARES PARA MODELAGEM E RECONSTRUÇÃO DE IMAGENS DE FACES*. Dissertação (Mestrado) — Centro Universitário da FEI, 2007.
- [62] HARDLE, W.; SIMAR, L. *Applied multivariate statistical analysis*. [S.l.]: Springer Verlag, 2007.
- [63] FUKUNAGA, K. *Introduction to statistical pattern recognition*. 2. ed. [S.l.]: Boston: Academic Press, 1990.
- [64] KITANI, E. C.; THOMAZ, C. E. *Um Tutorial sobre Análise de Componentes Principais para o Reconhecimento Automático de Faces*. [S.l.], 2006. Disponível em http://www.fei.edu.br/cet/Tutorial_ReconhecimentoFaces.pdf , último acesso 30/03/2009.
- [65] RUSSEL, S.; NORVING, P. *Inteligencia Artificial*. [S.l.]: Elsevier, 2004.
- [66] MARQUES, J. O. A. *Forma e Representação no Tractatus de Wittgenstein*. Tese (Doutorado) — Unicamp, 1998.
- [67] BITTNER, T.; DONNELLY, M. Logical properties of foundational relations in bio-ontologies. *Artificial Intelligence in Medicine*, v. 39, p. 197–216, 2007.
- [68] PARSIA, B.; SIRIN, E. Pellet: An owl dl reasoner. In: *Proceedings of the 3th International Semantic Web Conference*. [S.l.: s.n.], 2004.
- [69] HORROCKS, I. Fact and ifact. In: *In Proceedings of the International Workshop on Description Logics (DL99)*. [S.l.: s.n.], 1999. p. 133–135.
- [70] GARDINER, T.; TSARKOV, D.; HORROCKS, I. Framework for an automated comparison of description logic reasoners. In: *5th International Semantic Web Conference*. [S.l.: s.n.], 2006.
- [71] SANTOS, P.; CABALAR, P. The space within fisherman’s folly: Playing with a puzzle in mereotopology. *Spatial Cognition and Computation*, v. 8, p. 47 – 64, 2008.

- [72] SCHMIDTKE, H.; WOO, W. A size-based qualitative approach to the representation of spatial granularity. In: *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 2007.
- [73] BITTNER, T.; DONNELLY, M. Formal theory of qualitative size and distance relations. *Proceedings of the 21 International Workshop on Qualitative Reasoning*, v. 1, p. 31–49, 2007.
- [74] SPIES, M. Epistemic logic and knowledge management. *WM 2003: Professionelles Wissensmanagement - Erfahrungen und Visionen, Beiträge der 2. Konferenz Professionelles Wissensmanagement*, v. 28, p. 421–421, 2003.
- [75] TABET, C. Semantics and ontology: On the modal structure of an epistemic theory of meaning. *Pre-Proceedings of the 26th International Wittgenstein Symposium*, v. 1, p. 337–341, 2003.
- [76] SANTOS, D. N. *Descrição automática das diferenças estatisticamente mais relevantes entre pacientes com esquizofrenia e controles saudáveis*. Dissertação (Mestrado) — FEI, 2009.
- [77] LANCASTER, J. L. et al. Automated talairach atlas labels for functional brain mapping. *Human Brain Mapping*, v. 10, p. 120–131, 2000.
- [78] MARWEDE, D.; FIELDING, J. M. Entities and relations in medical imaging: An analysis of computed tomography reporting. *Applied Ontology*, v. 2, p. 67–79, 2007.

APÊNDICE A

APÊNDICE

Trabalhos Publicados Pelo Autor

1. Freire,R., Santos,D.N., Santos,M.V, Santos,P.E. - An exercise on developing an ontology-epistemology about schizophrenia and neuroanatomy. - Nature Precedings (2009)
Disponível em: <http://dx.doi.org/10.1038/npre.2009.3483.1>