# Development of a methodology and an ontological schema for medical terminology

A thesis submitted to the University of Manchester for the degree of
Doctor of Medicine
in the Faculty of Medicine, Dentistry, Nursing and Pharmacy

**2004**

by
Dr Jeremy Edward Rogers MRCGP

*Faculty of Science and Engineering*
*Department of Computer Science*

# Table of Contents

## List of Figures

# Abstract


**THE UNIVERSITY OF MANCHESTER**

**ABSTRACT OF THESIS** submitted by **Dr Jeremy Edward Rogers** for the Degree of **Doctor of Medicine** and entitled:

Development of a methodology and an ontological schema for medical terminology.

Submitted September 2004

Medicine has a long tradition of attempting to codify its language and terminology. Traditional and familiar clinical terminologies, such as the International Classification of Diseases, are not sophisticated enough to support modern aspirations for healthcare information technology applications. Formal ontologies are proposed as a solution, but formal ontologies that are simple to use are not useful, while useful ontologies are often too complex to be directly useable.

Part One of this thesis describes some of the significant sources of complexity in formal ontology design, and some of the ways this complexity affects users of the ontology.

Part Two describes a methodology for reducing the cognitive load of interacting with a complex ontology: the semantic choices that are to be applied consistently throughout the ontology are made explicit as a metamodel. The metamodel is subsequently harnessed both *pre hoc* to guide user choices and *post hoc* to normalise the semantics of their expressions to a preferred form. ‘Semantic normalisation’ allows what the user writes to be decoupled from what is understood by their writings, thereby enabling authoring to take place using an intermediate representation.

Part Three of this thesis presents a series of experiments of opportunity to evaluate the methodology.

# Declaration and Copyright

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

## Acknowledgements

I very gratefully acknowledge the help and support of Professor Alan Rector without whose faith and encouragement neither the work, nor this thesis, would have happened.

Thanks are also due to the other computer scientists I had the privilege to work with, whose patient response to my ever-increasing list of user requirements resulted in sophisticated software to make possible the methodology described in this thesis. My special thanks go to Danny Solomon and Egbert van der Haring who did the bulk of the software development directly relevant to this thesis, also to Conrad Toft for an excellent reworking of the user interface. Thanks to Sean Bechhofer and (again) to Egbert for getting the original classifiers working, and to Angus Roberts for getting thrown in at the deep end.

Also thanks to the other clinicians who have worked with me over the years, particularly Pam Pole, Jonathan Fistein and Christopher Wroe, and to all the other GALEN-IN-USE consortium members across Europe, especially Pieter Zanstra and Jean Marie Rodrigues.

Finally, thanks to Henry for giving me the energy,
and Yvonne for giving me the motivation and support to finish the job.

*This work was supported by grants from the European Commission:*

*GALEN-IN-USE (HC1018)*
*SemanticMining (Network of Excellence 507505)*

## The Author

Jeremy Rogers obtained his MB ChB from Manchester University in 1989 and completed his postgraduate training as a General Practitioner, comprising a 3 year hospital rotation in Stepping Hill Hospital, Stockport and community training in New Mills, Derbyshire. He obtained his Membership of the Royal College of General Practitioners in 1994.

In 1994 he joined the Medical Informatics Group in the Computer Science Department, University of Manchester as a Clinical Research Fellow, where he has worked on the GALEN, GALEN-IN-USE, PRESTIGE, SynEX, PRODIGY, CLEF and SemanticMining projects.

His research interests are knowledge representation, clinical data entry, clinical data quality, clinical decision support and clinical information governance.

# Part One: Introduction and description of problem

This thesis is organised into three major parts:

Part One – introduction and description of the problem

Part Two – description of a methodology to address the problem

Part Three – experiments to evaluate the approach

# 1 Statement of the Problem and Thesis

## *1.1 Terminologies in Medicine*

Medicine has a long tradition of attempting to codify its terminology. The London Bills of Mortality, compiled weekly from the 1680s through to the 1830s, categorised the cause of all reported deaths in London, with the goal of monitoring disease prevalence trends generally and outbreaks of plague in particular. They are recognised as the forerunner of more modern controlled terminologies and coding systems for epidemiological monitoring such as the International Classification of Disease (ICD) [WHO 1975, WHO 1992]. A fragment of ICD-9 is shown in Figure 1:

- D24 Disorders Of Thyroid Gland
  - D240 Simple And Unspecified Goiter
    - D240.0 Goiter, Specified As Simple
    - D240.9 Goiter, Unspecified
  - D241 Nontoxic Nodular Goiter
    - D241.0 Nontoxic Uninodular Goiter
    - D241.1 Nontoxic Multinodular Goiter
    - D241.9 Unspecified Nontoxic Nodular Goiter
  - D242 Thyrotoxicosis With Or Without Goiter
    - D242.0 Toxic Diffuse Goiter
    - D242.1 Toxic Uninodular Goiter
    - D242.2 Toxic Multinodular Goiter
    - D242.3 Toxic Nodular Goiter, Unspecified Type
    - D242.4 Thyrotoxicosis From Ectopic Thyroid Nodule
    - D242.8 Thyrotoxicosis Of Other Specified Origin
    - D242.9 Thyrotoxicosis Without Mention Of Cause
  - D243 Congenital Hypothyroidism
- D244 Acquired Hypothyroidism
  - D244.0 Postsurgical Hypothyroidism
  - D244.1 Other Postablative Hypothyroidism
  - D244.2 Iodine Hypothyroidism
  - D244.3 Other Iatrogenic Hypothyroidism
  - D244.8 Other Specified Acquired Hypothyroidism
  - D244.9 Unspecified Acquired Hypothyroidism
- D245 Thyroiditis
  - D245.0 Acute Thyroiditis
  - D245.1 Subacute Thyroiditis
  - D245.2 Chronic Lymphocytic Thyroiditis
  - D245.3 Chronic Fibrous Thyroiditis
  - D245.4 Iatrogenic Thyroiditis
  - D245.8 Other And Unspecified Chronic Thyroiditis
  - D245.9 Thyroiditis, Unspecified
- D246 Other Disorders Of Thyroid
  - D246.0 Disorders Of Thyrocalcitonin Secretion
  - D246.1 Dyshormonogenic Goiter
  - D246.2 Cyst Of Thyroid
  - D246.3 Hemorrhage And Infarction Of Thyroid
  - D246.8 Other Specified Disorders Of Thyroid
  - D246.9 Unspecified Disorder Of Thyroid

**Figure 1: Fragment of the International Classification Disease version 9, showing part of the hierarchy of disorders of the Thyroid Gland**

Since the 1960s, however, codified medical data has increasingly been put to uses beyond epidemiology. Numerous schemes have been constructed internationally for new applications, for example to support financial billing (OPCS [OPCS 1990], CCAM [CCAM 1998]), resource management (HRGs) and direct clinical care (ICPC [Okkes 2000], READ [O'Neil 1995], MEDCIN). The range of clinical information that can be collected using a controlled terminology has correspondingly broadened from cause of death (ICD) to all aspects of clinical care (SNOMED CT).

The drive for comprehensive UK clinical data collection passed a landmark in 2002 with the first Wanless Report [Wanless 2002], which stated:

> "6.22 If these issues [*agreeing the benefits Information & Communication Technology (ICT) would deliver; staff training; stringent interoperability standards; and ring-fencing of significantly increased ICT budgets*] can be addressed …national, integrated ICT systems across the health service can lay the basis for the delivery of significant quality improvements and costs savings over the next 20 years. Without a major advance in the effective use of ICT … the health service will find it increasingly difficult to deliver the efficient,

> high quality service which the public will demand. This is a major priority which will have a crucial impact on the health service over future years."

By way of response, in 2003/4, the English National Health Service (NHS) committed GBP 5.5 billion to realising a 10-year vision of the future for its ICT [Booth 2003]. A cornerstone of this vision is the NHS National Clinical Record Service (NCRS): by 2006, all citizens of England[1] are to have a unified electronic patient record accessible from any point of healthcare delivery in the NHS. The longer term vision sees a progressive maturation of the patient record from a passive, purely medicolegal record into an active software agent, capable of supporting clinicians and patients in collaboratively individualising, planning, understanding and auditing their healthcare.

## *1.2 What is an ontology ?*

Ontology as a branch of philosophy has a long history stretching back to the ancient Greeks. In this context it means the metaphysical study of the nature and essential properties and relations of all beings, or of the principles and causes of being.

Ontology has a much shorter history as a branch of artificial intelligence and knowledge representation, where it carries a different meaning: the study of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest, and the relationships that hold among them.

This thesis is concerned with this second meaning: ontologies as a means to represent knowledge. In the context of medical terminologies and classifications, therefore, an ontology aims to analyse and represent both the concepts used within a particular medical discipline, and the relationships between those concepts.

Such an analysis will usually go further than, for example, the simple observation that an existing medical terminology has listed two concepts: ‘endocrine disease’ and ‘thyroid cancer’ and a relationship between them, such that one is a kind of the other. Rather, the thyroid gland itself will be identified as a commonly recurring concept across many phrases in the terminology, and the endocrine system as another, even though no specific code existed in the original scheme for either concept. Similarly, the separate concepts of disease and cancer will be identified. The ontological analysis would further represent that the thyroid was, structurally, *part of* the endocrine system, that cancer *is-A* disease, and that endocrine disease and thyroid cancer may be defined respectively as ‘disease *with locus* endocrine system’ and ‘cancer *with locus* thyroid gland’.

By these steps, the detailed reasons *why* ‘thyroid cancer’ was originally classified as a kind of ‘endocrine disease’ are represented explicitly. A common motivation for such a deconstructive ontological process is that the newly explicit information might be reused either to infer new relationships, or to validate existing ones.

Figure 2 gives a flavour of how an ontology might pursue its analysis to represent concepts relating to thyroid anatomy and pathology.

---

[1] The National Programme for IT (NP*f*IT) covers only the NHS for England. The Welsh and Scottish Assemblies are each independently persuing and finding their own health care IT programmes.

Gland isParentOf ThyroidGland.

ThyroidGland
always isPartOf EndocrineSystem;
sensibly hasPart Lobe.

(Lobe which isPartOf ThyroidGland)
name LobeOfThyroidGland

(NonNormalSwelling which isLocatedIn ThyroidGland)
name Goitre;
sensibly hasNodularity Nodularity.

(Goitre which hasNodularity (Nodularity which hasAbsoluteState nodular))
name NodularGoitre.

(Pathology which isLocatedIn EndocrineSystem) name EndocrineDisease.
(InflammationLesion which isLocatedIn ThyroidGland) name Thyroiditis.

*Thyroiditis appears as a single coded concept in external classifications or terminologies such as ICD.*

*Here it is explicitly defined within an ontology, in terms of more basic concepts also in the same ontology.*

**Figure 2: Illustrative extract of an ontological modelling of the thyroid gland and thyroid diseases**

## 1.3 The case for formal ontologies in medicine

A central element of the NHS Integrated Clinical Record Service, intended to achieve greater interoperability of that service with other components of the overall ICT strategy, is the mandated use across the entire NHS of a new kind of medical coding system: SNOMED CT [College of American Pathologists 2004]. Although this terminology may appear superficially similar to traditional schemes such as ICD, the technical characteristics of its construction are fundamentally different: it belongs to a class of knowledge representation paradigms now commonly known as ‘formal’ ontologies based on ‘description logics’ [Borgida 1994, Borgida 1996, Baader 2003a].

The detailed reasons for this move to a new technical foundation for clinical terminology are outside the scope of this thesis. A summary is presented here:

Traditional clinical terminologies such as ICD suffer from notoriously impoverished expressivity – clinicians can rarely say precisely what they needed, or wanted to say. For example, as illustrated in the extract of ICD relating to thyroid disease (Figure 1), there is no mechanism by which a clinician may further differentiate the types of acute thyroiditis by their recognised aetiology (e.g. autoimmune, infective, drug- or radiation-induced) even though these would typically have different treatments and outcomes. Similarly, there is no specific code for the clinical entity known as post-partum autoimmune thyroiditis, whose aetiology is unknown.

Traditional clinical terminologies like ICD are insufficient to support many desired applications [Lewis 2002]. As a result, and on a global scale, there is continuing activity to adapt, modify or create new classifications that will better support real, local applications [Chute 2000]. These endeavours traditionally employ entirely manual construction and maintenance techniques, requiring skilled curators. They are consequently very costly and labour intensive undertakings, prone to the errors typical of manual cataloguing.

Empirically, this effort has not resulted in a satisfactory solution to the problem: useful ‘intelligent’ systems driven by machine interpretation of coded clinical data – such as clinical decision support – have remained elusive [Broverman 2000]. Worse, the ensuing growth in the number of different schemes, or variants of the same scheme, has resulted in a Tower of Babel: data coded using different schemes cannot be merged, and algorithms written to analyse data from one scheme do not work on data expressed in another [Sittig 1994, Chute 2000].

Formal ontologies and description logics propose a unified solution to address many of these problems: an explicit model of the conceptual content of a domain of discourse (the ontology), expressed in a syntax with specified semantics (a *formal* ontology), that can be reasoned over by a computer algorithm (a description logic engine). They promise both hugely increased expressivity - users can create an almost infinite number of compositions by combining concepts already in the ontology - and a simultaneous reduction in the manual curation effort. This reduction in effort arises because the description logic engine (the classifier) automates the processes of examining the declared semantics of a given expression and comparing it with the explicit semantics of all the other expressions currently in the system, and thereby derives an automatic classification of new expressions with respect to the space of concepts previously encountered [Chandrasekaran 1999].

### 1.4 A problem with formal ontologies

During their construction, authors of both formal and informal ontologies frequently encounter situations where the same conceptual content can be represented in more than one way. For example, an ontology of medicine must decide whether, when representing ‘bilateral nephrectomy’, the ‘bilateral’ element attaches to the anatomy or to the surgical method. Is it a bilateral kidney, or a bilateral excision, or (redundantly) both? Such situations require arbitrary - and sometimes anti-intuitive - decisions to be made regarding how to model the domain.

Rigorously consistent application of such rules across an entire ontology – e.g. that ‘bilateral’ can modify a method and never anatomy - is required if that ontology is to remain predictably computable according to the specified formalism.

As an ontology grows ever larger, either to meet the demands of more use cases or to increase the detail with which it models the domain, so the number of arbitrary modelling choices and distinctions that have to be made increases. If, in a formal ontology, the underlying formalism is altered to increase its logical expressivity, the problem may be further magnified. The number of rules soon reaches a point where unaided humans can no longer consistently memorise and apply them, or be able to predict the effect of their modelling. This phenomenon has a greater consequence in formal ontologies compared with informal ontologies: the computer becomes unable to reason reliably across the ontology, because the various rules and distinctions are no longer consistently applied. Consistent application of formal rules of inference to an inconsistent knowledge base produces unpredictable results.

Users of ontologies may blame their difficulty in following the rules on a lack of adequate documentation. However, as an ontology grows larger and more complex, so an exhaustive documentation of all its rules or conventions becomes of decreasing value: few users will be unable to memorise all of that documentation. Moreover, precisely because the rules appear anti-intuitive or arbitrary, users do not always recognise *ad hoc* when a choice exists to be made or that the manual should be consulted for the appropriate rule to apply. In the context of biomedical ontologies such as SNOMED CT this is a particular risk, because many users (e.g. clinicians and nurses) will have almost no training or knowledge concerning ontology engineering.

## 1.5 Thesis

A synthesis of the problem, therefore, is:

> *Medicine needs useful formal ontologies, but formal ontologies that are simple to use are not useful, while useful ontologies appear to be too complex to be directly useable.*

This formulation of the problem hints at a possible solution: preserving the complexity and rigidity of a large ontology is desirable if our ability to compute over the result is not to be diminished or lost entirely. However, the user needs to be protected as much as possible from that complexity.

A potential methodology for preserving, whilst simultaneously hiding, ontological complexity arises from the following observation: many of the apparently arbitrary rules and conventions that are to be followed in a particular ontology can be explicitly described. Similarly, it is also possible to describe many of the possible alternative patterns that were considered but, ultimately, not chosen as the preferred form.

Explicit statement of these rules and conventions might allow for mechanisms to aid authors. These could include *pre hoc* constraints compelling users to follow the rules, as well as *post hoc* detection of 'alternative' expressions, triggering an automatic coercion to the preferred form. By these mechanisms, both authors and users could be freed from any requirement to know the more complex but regular rules, while these rules are still applied and checked for.

The thesis presented here is that:

> *Much of the complexity of formal ontologies arises from the consistent application of semantic patterns and choices. The cognitive load of using a complex formal ontology can be reduced if these patterns and choices are made explicit as a metamodel of the ontology, and where the metamodel is subsequently harnessed to guide user choices pre hoc and transform expressions post hoc to a preferred semantic form.*

In particular I shall demonstrate how one such formal ontology – the *Open*GALEN Common Reference Model of medicine (*Open*GALEN CRM) [OpenGALEN 2003] – can be presented in simpler form to facilitate authoring, using *post hoc* transformation algorithms.

## 1.6 Scope of Thesis

This thesis specifically does *not* seek to address the question of whether the *Open*GALEN CRM itself needs to be as complex as it is. Evidence for similar levels of complexity in other medical terminologies will be presented, but the goal of the thesis is not to justify or critique any particular level of complexity.

The thesis seeks only to propose a methodology whereby, given the complexity of ontologies such as the *Open*GALEN CRM, that complexity may be harnessed in order to construct a simpler interface with the user.

## *1.7 Setting*

This thesis presents a high level synthesis of the author's fulltime collaborative research work over a period of eight years (1994 to 2002) in projects[2] funded by the European Commission Research Framework III and IV programmes, and the UK Department of Health.

The core of the methodology, however, was developed by the author during the GALEN-IN-USE project (1996-1999), a pan-European collaboration funded for 42 months by the European Commission Research Framework IV, researching a new approach for constructing a multilingual European classification of surgical procedures [Rogers 2001, Solomon 2000]. At its centre was a large formal ontology of medicine, now known as the *Open*GALEN Common Reference Model. This ontology is expressed in GRAIL, a forerunner of more modern description logics [Rector 1997].

The chosen methodology for the project called for the meaning of code rubrics, taken from existing classifications of surgical procedures, to be expressed as structured representations ("dissections") with precise semantics using the *Open*GALEN CRM [Rogers 1997a].

This process is shown diagrammatically in Figure 3. The code 'O470', taken from the International Classification of Diseases version 9, corresponds to the text string rubric 'Appendectomy'. The apparent meaning of this rubric is expressed via an authoring process as a 'dissection' using a formal notation and an ontology of medical concepts:

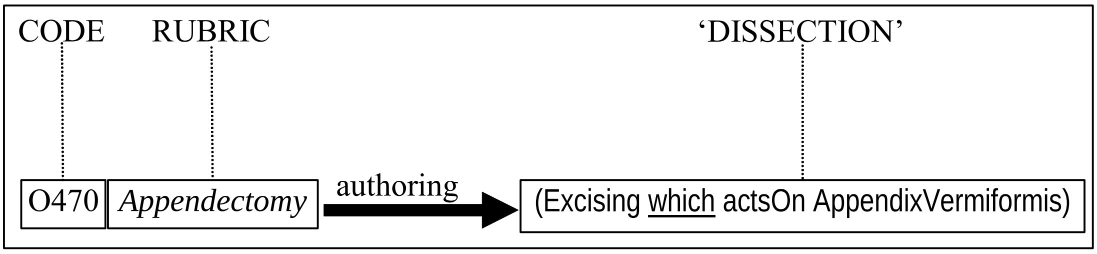



**Figure 3: Codes, rubrics and dissections**

More than 20 clinical authors distributed between the UK, Netherlands, Sweden, Finland, Greece, Germany, France and Italy were recruited to write the dissections.

It was originally conceived that dissections would be expressed directly in the GRAIL syntax and using the Common Reference Model ontology. However the clinical authors had little or no prior experience of formalisms such as GRAIL or of the particular ontology and modelling style of the Common Reference Model. They found the task too complicated. An alternative authoring paradigm was required instead, which is the subject of this thesis.

A new methodology was developed by the author of this thesis to present the dissection authors with a simpler syntax and an *ad hoc* user extendable vocabulary. A corpus of 20,782 rubrics, covering substantially the entire surgical subdomain, was eventually successfully 'dissected' by them using the new approach.

Two subsequent projects extending the methodology were funded by the UK Department of Health: the PRODIGY Drug Ontology [Wroe 2000] and the Prescribing Indicators Project [Rogers 2003].

[2] GALEN, GALEN-IN-USE, SynEx, PRESTIGE, PRODIGY, Drug Ontology, Prescribing Indicators

## 1.8 A note about 'users'

The terms 'user' and 'end-user' appear frequently throughout this thesis. In software engineering contexts, 'user' or 'end-user' refers to a person who uses a computer application, as opposed to those who purchased, developed or support it. The user may or may not know anything about how the application was engineered as a piece of software, and so may not realise when it isn't working properly or know how to fix it.

Although biomedical terminologies are not traditionally thought of as software applications, as they migrate from being human-only readable lists in books to becoming components of the software that interprets them, so the term 'user' is now also commonly used to denote a person who interacts with a computable ontology. However, many different types of user interaction with an ontology can be imagined, requiring varying levels of understanding of the underlying ontology engineering issues.

Within the body of this thesis, the term 'user' usually relates to those using the ontology specifically to assist them in the task of building and maintaining more traditional terminologies and classifications. However, particularly in Chapters 3 and 4 (sources and effects of ontological complexity) and in the final discussion, 'user' may also relate to clinicians using the ontology to represent part of a medical record. Also within this thesis the term 'user' is often qualified to indicate one of three broad subclasses according to their typical levels of understanding or experience of ontology engineering. These are listed below together with a description of a stereotype for each class:

**inexpert user** – *a user who has little or only limited understanding of ontology engineering either in general or with respect to the particular ontology they are using. One* who *records information using one or more terms from an ontology but who does not know how that data is subsequently to be processed. Exemplar: most busy clinicians when recording the characteristics of the patient before them, or of a procedure just performed, for inclusion in a medical record. Also within the context of the body of this thesis, the dissection authors were considered inexpert users.*

**expert user** – *a user who has a good understanding of ontology engineering both in the general case and specifically with respect to all aspects of the particular implementation of the clinical ontology before them. They are aware of the choices they have when forming an expression using the ontology, and of how those choices affect its subsequent analysis. Exemplar: an ontology engineer tasked with maintaining or extending an ontology, and with devising acceptable user interfaces by which inexpert users might use the ontology*

**intermediate level user** – *any user whose understanding is between expert and inexpert*

## 1.9 Assumed knowledge

Prior familiarity with the basic components from which ontologies are constructed, and with description logic technologies, would assist the reader of this thesis. A detailed account of either is beyond the scope of this thesis, and the reader is referred to [Gómez-Pérez 2004] and [Baader 2003a]. However, some notions particularly relevant to this thesis are presented in outline here:

### Primitives vs Compositions

An ontological analysis of the set of concepts within a domain of discourse will necessarily divide that set of concepts into two categories:

- Those concepts whose semantics cannot be expressed in terms of other concepts in the ontology. These are usually called primitive, or elementary concepts. Note that the notion of a primitive is different from that of a 'natural kind': a concept represented as a primitive in one ontology might representable as a composition in another. A natural kind can not be defined in *any* ontology. Hence, all natural kinds must be represented as a primitive in every ontology, but not all primitives in a given ontology are necessarily natural kinds.

- Those concepts that are expressed in terms of other concepts in the ontology. These are termed compositions, or composed concepts.

**Description logics, engines and classifiers**

A description logic is a formal specification of the rules for examining and comparing semantic descriptions relating to primitives and compositions within an ontology. When encoded as a software device, the result is a description logic engine. These are normally capable of, or aspire to, automating a number of types of inference, including:

- Detecting inconsistencies in the definition of a concept
  *e.g. a 'dead, alive person' would be a contradiction*
- Determining equivalence
  *e.g. 'left thumb' and 'thumb of left hand' refer to the same concept*
- Determining the classification of any newly presented concept with respect to all other concepts currently known to the engine
  *e.g. 'fracture of the femoral head' is a kind of 'fracture of a long bone'*

Because of this last function, description logic engines are often also often referred to as classifiers.

## *1.10 Organisation of Thesis in Three Parts*

The twelve chapters of this thesis are organised into three major parts and the appendices:

**Part One** – *introduction and description of the problem*

Chapter 1 - introduction (this chapter)
Chapter 2 - a brief review of other work in the field

Chapter 3 - the principal factors contributing to complexity in the design of formal clinical ontologies
Chapter 4 - how that complexity impacts on users.

**Part Two** – *description of a methodology to address the problem*

Chapter 5 - a summary of the methodology
Chapters 6 through 9 - documentation of the methodology in more detail.
Chapter 10 – implementation of the methodology

**Part Three** – *experiments to evaluate the methodology*

Chapter 11 - presentation of several experiments of opportunity to evaluate parts of the approach.
Chapter 12 - discussion and topics for further research.

**Glossary and Appendices**

A glossary of unfamiliar terms is provided

Three appendices present large datasets, or fragments of very large datasets, that are referenced in the text. A fourth appendix summarises a critical appraisal undertaken to consider the validity of results from the experiments of opportunity presented in Chapter 11.

## *1.11 Typographic conventions*

The following typographic and other annotational conventions are employed throughout this thesis:

Where an example or figure illustrates a flawed or otherwise not recommended way of representing knowledge, it is marked with this symbol: ⊗

**GRAIL expressions**

GRAIL expressions or fragments of expressions appear in this font [Arial Narrow].

GRAIL keywords appear underlined:

which, whichG, sensibly etc.

Knowledge names for concepts in GRAIL expressions appear in Arial Narrow, whilst names for semantic links appear in Arial Narrow italics:

Concept which *isLinkedTo* Concept

Where references to GRAIL concepts appear within the body text, they appear in Arial Narrow and surrounded by square brackets: [Leg], [Surgery], [*isLinkedTo*]

**Intermediate Representations**

Intermediate representation dissections appear in this font [Arial Narrow].

Knowledge names for descriptors in the intermediate representation appear in lower case only. Where they appear within the body text e.g. 'descriptor', they appear in this font but in single quotes.

Knowledge names for semantic links in the intermediate representation appear in upper case Arial Narrow with spaces replaced by an underscore: HAS_LOCATION.

**Other example representations**

In some cases, examples of possible representations are presented in an abbreviated pseudocode that is neither strictly GRAIL nor intermediate representation.

*These appear in this font and fully italicised.*

# 2 Introduction: Other Ontology Work

Recognition of the role of shared, robust ontologies in enabling machine communication has arrived gradually in many domains over the last decade [Chandresankar 1999], but ontologies are currently an area of very active research particularly in the context of the semantic web initiative, involving technologies such as OWL, TopicMaps and RDF [Horrocks 2003, Baader 2003b].

The broader biomedical terminology community has been at the forefront of this research, and the past decade has witnessed a detailed re-examination of its long-standing tradition for structuring biomedical domain knowledge as classifications. This is leading to a progressive migration to the new technologies, particularly in medicine [Noy1997, Chute 2000, Bodenreider 2001].

## 2.1 From Frames to Formalisms

Work within the artificial intelligence community to build ontologies for many domains, including medicine, began with early frame-based approaches [Minsky 1974]. This approach continues today through projects such as The MED [Cimino 1989, Cimino 2000], PROTÉGÉ [Gennari 2002] and the Digital Anatomist [Noy 2004].

Modern Frame-based systems based on the OKBC standard grew out of earlier systems such as KIF. They are constructed around the notion of a *frame*, a primitive object representing an entity in the domain of discourse. *Frames* are instances of one or more *classes* or *metaclasses* in the system. Each *frame* is associated with a number of *slots*, each of which may be filled with a number of *slot values* that are also *frames*. Thus each *slot* asserts a binary semantic relationship between two *frames*.

Among clinical terminologies, ICPC and SNOMED International [Côté 1993] were two of the first to present themselves as multi-faceted or frame-like schemes: new concepts outside either terminology could be expressed or defined by populating a simple frame with more basic concepts already within the terminology.

In 1998 Rossi Mori described the required characteristics of a new generation of semantically-based clinical terminologies [Rossi Mori 1998]. In the same year, Cimino described twelve desiderata of modern medical vocabularies [Cimino 1998a], including that concepts within a terminology should have a semantic definition. However, both stopped short of recommending that these should be based on a computable formalism [Rogers 1997b].

By the end of the 1990s two of the largest medical terminologies - SNOMED RT in the US [Spackman 1997], and Clinical Terms version 3 (CTV3) in the UK [O'Neil 1995, Brown 1998, Price 1998, Brown 1999] – were providing semantic definitions for concepts within the terminology. CTV3 had no underpinning formalism or logic engine; SNOMED RT was expressed in KRSS [Spackman 1997], a variant of KLONE.

A limitation of the frames paradigm, particularly in the context of authoring a large ontology, is that it has only one native capability for inference: the inheritance of slots down the asserted hierarchy. There is no in-built mechanism for a reasoner either to infer new relationships between concepts (including additional hierarchical relationships), or to check whether two declared relationships contradict each other. This limitation becomes significant if the domain being modelled is large or complex: humans make errors of omission or commission even in relatively small ontologies of only a few hundred concepts [see 11.3.2.3].

Formal ontology engineering primarily seeks to provide such inference and checking, thus offering increases in authoring efficiency - why bother to author by hand what can be inferred by the machine? - and accuracy. Research development of the *Open*GALEN CRM as a formal ontology began in 1992. In 1999 CTV3 and SNOMED RT merged to form SNOMED CT [College of American Pathologists 2004] and at the same time agreed to underpin subsequent development with a description logic formalism that assists in computing hierarchies and detecting conflicts. Other biomedical ontologies, such as the Gene Ontology, are also exploring this technology [Bada 2003, Wroe 2003].

In order to remain computationally tractable, however, current description logics offer a carefully restricted subset of logical constructors such that many desirable or 'true' statements can not be expressed [Baader 2003, Horrocks 2003]. For example, no description logic dialect currently supports 'shared variables' and, hence, can not express assertions such as that a procedure to incise and drain an abscess should properly be represented as an 'incision of an abscess:X followed by drainage of *the same* abscess:X'. Workarounds to selectively extend the expressivity of description logic systems have been described, particularly with regard to modelling of partonomy [Hahn 1999a, Hahn 1999b, Schulz 2000, Rogers 2000]. Ceusters and Smith have highlighted the expressive limitations of all such formalisms [Ceusters 2003].

## 2.2 Managing scale

The traditional manual techniques for ontology construction and maintenance are labour intensive and error-prone, requiring skilled curators. The scale of medicine makes this endeavour still more expensive. Semi-automatic approaches to ontology induction, often using natural language processing techniques, are an area of active research [Hahn 2003].

## 2.3 Managing mulitple ontologies

As medical ontologies have grown more numerous, so other researchers have explored techniques to manage the relationships between different ontologies: Oliver described requirements and techniques for documenting and managing change between diverging versions of the same ontology [Oliver 1998, Oliver 1999a, Oliver 1999b]. Noy designed an interactive environment for merging different ontologies, as well as a beginner's guide to ontology construction [Noy 1997, Noy 2001, Noy 2003,]. Gangemi has compared the implicit and explicit upper ontologies of several medical terminologies and integrated them within a single more abstract model [Gangemi 1999]. The National Library of Medicine's Unified Medical Language System [Lindberg 1993] uses lexical approaches to infer which medical phrases from traditional medical vocabularies mean the same thing, although others have demonstrated that this approach falls short of true concept equivalence [Pisanelli 1998, Cimino 1998b].

## 2.4 Managing complexity

The increasing number of available medical ontologies has been accompanied by a steady growth in their individual size and complexity. This poses new challenges for measuring and maintaining quality during the construction phase.

Measures of the quality of medical ontologies, or their fitness for purpose, have mostly been limited to comparisons of domain coverage under experimental conditions [Chute 1996, Brown 2001, Brown 2003], together with expositions of the philosophical or logical principles underpinning the ontological design and demonstrations that systems are in place to ensure these principles are adhered to. For example, the SNOMED Authority claims robust quality assurance measures are in place [College of American Pathologists 2004], but Ceusters has described errors of content within SNOMED-CT arising from a lack of tools to support consistent and systematic modelling [Ceusters 2004].

Measures of domain coverage or of philosophical correctness remain an unsatisfactory means to predict either fitness for purpose or end-user acceptance. Performance in the field remains the ultimate acid test of whether an ontology is 'good' or not. Studies of inter-rater variability in clinicians selecting codes from comparatively simple medical terminologies suggest that, even if the authors could deliver a more coherent ontology, users are prone to using it in incoherent ways [Bernstein 1997, Rogers 2003].

# 3 Sources of Complexity in Ontology Design

This thesis is concerned primarily with techniques to hide the complexity within an ontology from the users, in order to make it usable. This chapter describes some of the significant sources of that complexity identified by the author in the course of the research activity. Although the specific examples presented draw on the author's experience with a large medical ontology, the problems highlighted are believed to be generalisable to the design of large ontologies in other domains. In summary, they are:

- *Domain Complexity*
  - The medical domain is very large and inherently complex (see 3.1)
  - The medical domain is commonly viewed from many different ontological perspectives (see 3.2)
  - External or legacy knowledge bases that are to be expressed in terms of an ontology may require general categories or abstract axes of classification that are alien to most users (see 3.3)
  - Medical subspecialty ontologies do not have clear boundaries (see 3.4)
  - Large ontologies with many axes of classification are hard to navigate (see 3.5)
- *Artefactual Complexity*
  - Technologies for ontology engineering require formal syntactic notations that are often hard for humans to read (see 3.6)
  - Technologies to reason over ontologies have limitations; workarounds may introduce artefacts within the ontology (see 3.7)
- *Cognitive Complexity*
  - Natural Language may suggest ontological distinctions that are not relevant to the formal ontology, or that include some semantic redundancy (see 3.8)
  - Formal ontologies depend on greater precision and consistency of expression than humans normally recognise (see 3.9 and 3.10)
  - Constraint checking to enforce precision and consistency often only informs the user that there is a problem, not how to correct it (see 3.11)

These sources are discussed further in the following sections.

## *3.1 Domain Complexity: Scale of the medical domain*

The medical domain is very large. The November 2003AC release of the National Library of Medicine's Unified Medical Language System (UMLS) Metathesaurus, a project run since 1986, identifies 975,354 distinct concepts employed within the medical domain [NLM 2001] and currently listed in at least one controlled medical terminology. These concepts are further categorised under 134 high level semantic types (substance, disease, organism, anatomy, physiological process etc) which are themselves organised into a shallow monoaxial hierarchy. SNOMED CT, the mandated reference concept system for medicine in the United Kingdom National Health Service, contains fewer concepts (around 350,000) but they are organised into a much richer and more detailed class polyhierarchy.

Section 1.3 described how, within terminologies such as ICD, expressivity is often limited so that it may not be possible to record clinically significant details, such as the aetiology of thyroiditis. Paradoxically it is also the case that some of the distinctions between the listed concepts in such schemes can be subtle: for example, the September 2003 version of

READ codes lists 93,000 concepts, including a unique code and term for pulmonary fibrosis caused by inhaling aluminium particles: 'H430: aluminosis', and a different code and term preferred for pulmonary fibrosis caused specifically by inhalation of aluminium ore (bauxite) particles: 'H431:Bauxite fibrosis'.

Some terms are so highly specialised that many users of the scheme do not know what they mean, and only a tiny handful of patients could ever properly be assigned them. The same chapter of the READ codes (Respiratory System Diseases) includes a unique code (H35y5) for 'Pituitary snuff-takers' disease', a kind of allergic fibrosing alveolitis caused by inhaling the dried and powdered extract of human cadaver posterior pituitary glands. This preparation was used in the past to treat Diabetes Insipidus before the widespread availability of purified vasopressin nasal sprays. The 2004 release of SNOMED CT carries its own code for this concept. Only two recorded cases of Piuitary snuff-takers disease exist in the entire English speaking medical literature, the most recent in 1967 [Mahon 1967]. A more celebrated example of niche coding comes also from the READ codes, which were designed for, and are mainly used by, UK primary care physicians. Despite this context of use, the scheme includes (inherited from ICD9) 10 codes relating to specific types of accident sustained while travelling in a spacecraft.

The knowledge burden of coping with the size of the medical domain is compounded by the fact that it is also not static: new discoveries continually refine our understanding. New diseases and treatments are discovered or devised, whilst old ones are consigned to history. Existing classifications struggle to keep pace: OPCS version 4, the scheme used to record surgical activity in the NHS, has no codes for endoscopic procedures. Fully 25% of all the terms in ICD9 for psychiatric pathology are no longer recognised as real clinical entities by US psychiatrists.

The pace and scope of this refinement seems set to accelerate in the coming years as genomic and proteomic research begins to impinge on clinical medicine: In 1975 the International Classification of Disease version 9 (ICD9) contained only one code for Charcot-Marie-Tooth disease, the commonest genetically determined disorder of the peripheral nervous system with an incidence of 1 in 2500. By 1996, Steadmans Medical Dictionary recognised three clinical variants: one caused by sensorineural demyelination, one involving axonal loss, and a third characterised by anterior horn damage. By 2003 OMIM recognised forty-six distinct subphenotypes of the condition [OMIM 2004]: for twenty-six both the genetic and molecular basis has been established; in a further thirteen a distinct genotype is identified but the molecular basis remains unknown, while a further seven phenotypes are suspected to be genetically determined but no genetic locus (or molecular basis) is yet identified. This progressive genetic subdivision of clinical phenotypes is not restricted to the classical inherited diseases: OMIM also lists eight genetically determined phenotypes for migraine, six for asthma, three for a propensity to tuberculsosis, two for nocturnal enuresis and one for a familiar propensity to appendicitis.

These results suggest that the 20,000 or so discrete disease entities currently recognised in clinical medicine (13,390 in ICD-9-CM) may soon be multiplied by at least an order of magnitude as each is subdivided into genetic variants with different patterns of inheritance, prognosis or preferred treatment regime.

Finally, in addition to coping with the absolute and growing number of distinct medical categories, clinical descriptions of those categories may themselves be very detailed. Many diseases can be further qualified by a severity value, a stage, a prognosis, an aetiology, the presence or absence of key symptomatology and so on.

A successful ontology of medicine, therefore, strives to cope with all three aspects of the domain: size, change and detail of description.

### 3.2 Domain Complexity: Multiple ontological views

There is no single true ontology - for any domain - which is there to be discovered. Rather, many different ontological approaches can be used to describe any given domain and still produce a 'useful' result. In some cases, two different ontologies for the same domain may be equally suitable for one task, but for another task only one of the two is suitable. For example, both ICD version 9 and Clinical Terms Version 3 in medicine support aggregated statistical analysis of causes of death, but of the two only Clinical Terms Version 3 also has terms to record the symptoms associated with the same diseases.

However, more significant than any differences in the level of detail and scope offered by two ontologies are the differences in *how* they represent even the detail that they share. Consider the following example of semantically distinct ways of saying the same thing: In normal usage, humans would understand the phrases 'a half full glass of water' and 'a half empty glass of water' to be equivalent. We understand that, if somebody gave us the choice of receiving a real physical instance of either notion, we would receive the same object regardless – the same 'extension'. The two phrases, however, reveal two different intensions, or ontological views, of the world – one measures glasses of water on an 'emptiness' scale while the other measures them on a 'fullness' scale.

In medicine, as in other domains, it is common for different professionals to want to focus on different characteristics of the same event: physicians classify disease by organic pathology, nurses by the disability or suffering it causes; pathologists classify neoplasms by cell line and morphology, geneticists by cytogenetic markers, clinical oncologists by stage and progression, and patients by life expectancy. The general physician may only be interested in the location of a pain, whilst a pharmacologist may be more interested in comparing the mechanisms of action of different analgesic drugs in terms of the physiology of pain sensors and nerve conduction, with correspondingly less interest or precision in where the pain is located. A neurologist, by contrast again, may have no interest when examining a patient with referred or phantom limb pain in how pain is conducted from the periphery to the brain, but instead is especially interested in the process of cognitive pain localisation and perception. They may seek to draw a distinction, normally elided over in general medical practice (except in the case of referred pain), between the anatomical site from which the pain signal actually arises and the anatomical site where it is perceived to have arisen.

As an ontology seeks to support all these different viewpoints, so the number of options for describing concepts increases. Before long, the descriptive options required for one user with one viewpoint is overwhelmed by the union of all the descriptive options required by all other viewpoints.

### 3.3 Domain Complexity: Need for rule parsimony

The primary purpose of an ontology is normally to link concepts within to objects *outside* that ontology. This section demonstrates that, in order to author such links, the ontology may include concepts that are abstract or otherwise peculiar to other groups of users.

In a medical context, such a link might be between the CRM concepts [cough] and [wheeze] and the external identifier for a clinical proforma concerning asthmatic patients, as a result of which link the form might include options to specify whether a patient coughs or wheezes:

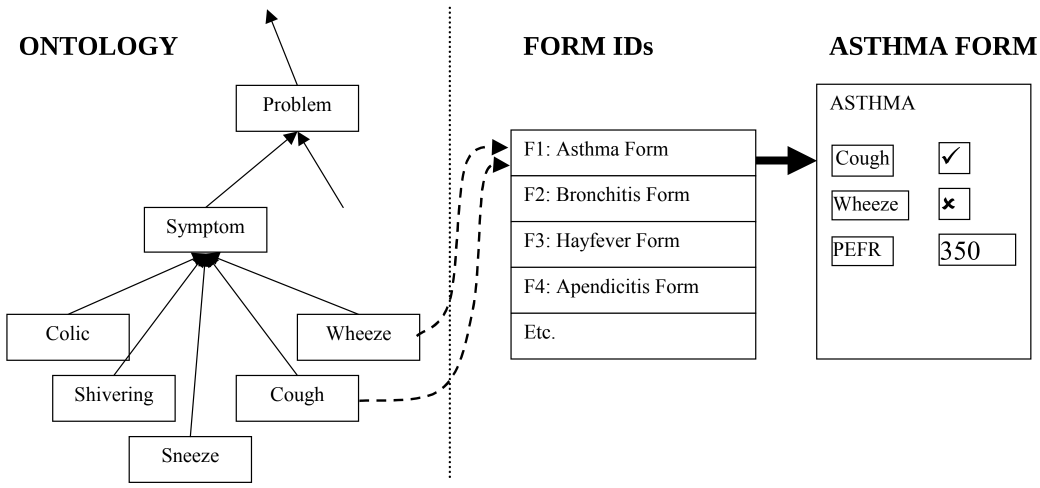



**Figure 4: Linking an ontology to data entry forms**

Similarly, the same concepts [cough] and [wheeze] may be linked to the external identifier for a prescribable drug, and as a result clinicians in this application are alerted not to use the drug in patients where those symptoms have previously been recorded.

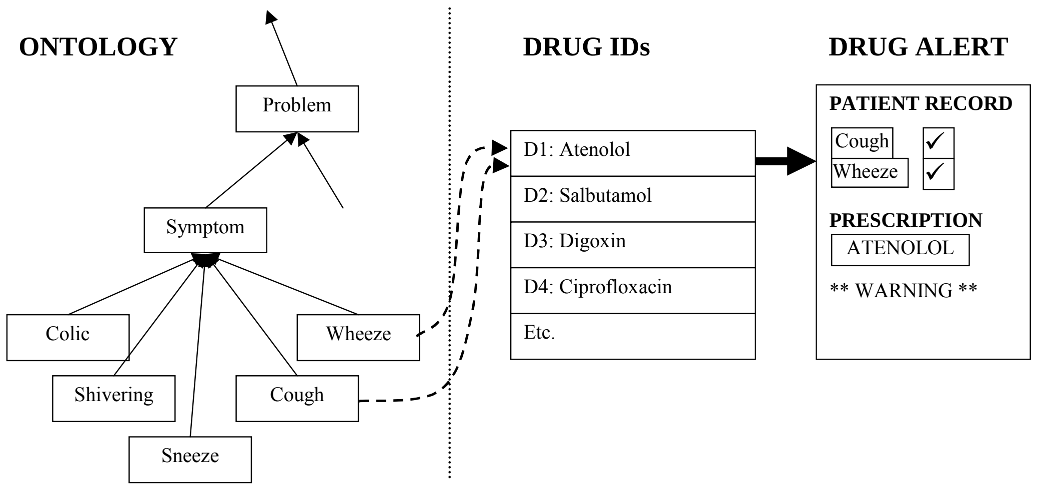



**Figure 5: Linking an ontology to drug alert application**

In both applications outlined (data entry or drug contraindication warning), the external links form a set of rules declaring the desired behaviour of the application in response to particular types of user input. Authoring such rule sets, however, poses a challenge: how to author the rules most parsimoniously.

In the case of a data entry form application, consider how to author a rule or rules such that all the anatomical structures that may be qualified as either left or right can be so qualified. Clearly, it is not desirable to have to explicitly and individually enumerate all the possible combinations:

> Form on [Hand] **should have** *options* [left, right] ⊗
> Form on [Palm of hand] **should have** *options* [left, right]
> Form on [Finger] **should have** *options* [left, right]
> Form on [Thumb] **should have** *options* [left, right]
> Form on [First metacarpal] **should have** *options* [left, right]
> Form on [Second metacarpal] **should have** *options* [left, right]
> Form on [Third metacarpal] **should have** *options* [left, right]
> Form on [Fourth metacarpal] **should have** *options* [left, right]
> Etc. etc.

Similarly, we would prefer not to have to separately list all the possible data entry form topics where a clinician might wish to comment on whether the patient has a cough or a wheeze:

> Form on [Asthma] **should have** *options* [cough, wheeze] ⊗
> Form on [Bronchitis] **should have** *options* [cough, wheeze]
> Form on [Lung cancer] **should have** *options* [cough, wheeze]
> Form on [Pulmonary tuberculosis] **should have** *options* [cough, wheeze]
> Form on [Allergic alveolitis] **should have** *options* [cough, wheeze]
> Etc etc.

The solution to this problem is to recognise that there are more parsimonious ways to author the required rules, for example:

> ALL Forms on [mirror-imaged body structures] OR [their subparts] **should have** *option* ANY [laterality]
> ALL Forms on [respiratory disease] **should have** *options* [cough, wheeze]
> ALL Forms on [symptoms] **should have** *option* ANY [severity]

However, in order to author such parsimonious rules, it is necessary that the concepts with which they are authored already exist or can be expressed in the underlying ontology. Equally importantly, these concepts must subsume both all that they should, and nothing that they should not: below the concept [Symptom] should be all possible symptoms, and no anatomical variant findings; [Respiratory Disease] should subsume pulmonary tuberculosis and lung cancer.

However, some concepts required for rule authoring, and the axes of classification they embody, may be very abstract: consider again the rule saying that, where an anatomical structure has a left and a right sided form, any proforma on which it appears should offer the option to supply the laterality. To write this rule requires an abstract superclass concept in the ontology equivalent to the notion of a 'structure that has potential-left-or-right-ness' and also all affected anatomical concepts must be qualified with this abstract property.

A further example is given by the representation of rules such as that proformas concerning *any* hollow object may offer the option that they contain other structures, whilst *only those* concerning certain subtypes of hollow object (e.g. tubes) should additionally offer options that they may be further qualified by a diameter. These and similar rules required development, within the CRM ontology, of a detailed model of variant topology: physical structures are grouped by whether they exist in one-, two- or three dimensions, and completely enclosed three dimensional spaces are distinguished from open-ended tubes, blind-ending tubes and anatomical potential spaces.

A consequence of parsimonious rule authoring for applications is that many concepts within an ontology must carry detailed descriptions of abstract properties (such as their topological characteristics). This can only increase any obfuscation of the model already caused by the integration of multiple ontological views (section 3.2).An important difference is that, typically, the obscure axes of description introduced into the model are not normally expected, or necessarily understood, by *any* of the regular end-users.

### 3.4 Domain Complexity: Lack of clear boundaries

The problem of any domain having different viewpoints is compounded in medicine by the fact that medicine is rarely treated as a single domain but instead as a multitude of smaller specialty subdomains. Many medical specialties choose to construct dedicated terminologies or ontologies specific to their subdomain.

However, the notion of an ontology being limited to a domain is somewhat artificial, implying a cleaner boundary between what is covered and what is not than exists in practice. In reality, ontologies that claim to cover only a particular domain inevitably touch on many related domains, each of which could be (and often is) independently represented using its own large and dedicated ontology.

For example, an ontology intended primarily to support the care of diabetic patients will inevitably require mention of many anatomical structures, such as the parts of the retina or the blood supply of the lower limb.

The natural tendency, when building an ontology for a specific purpose, is to focus on achieving consistent and coherent ontological style only for what is considered the core ontology – that part comprising the subtle differences and distinctions peculiar to the primary domain of discourse. The concepts more at the periphery of the domain, for example anatomy in the context of an ontology for diabetes, are often treated more casually.

Typically, this takes the form of a degree of ‘ontological simplification’, which becomes possible because the number of concepts involved from any single peripheral domain is small. So, for example, in modelling anatomy for the purposes of describing where diabetic lesions are, it is possible that the relationship between the anatomical structures involved and their more specific subparts would be represented using a single flavour of *isPartOf* link. This contrasts with what occurs in more complete models of all anatomy, where multiple flavours of *isPartOf* are often required [Artale 1996, Rogers 2000, Mejino 2003]. The modelling of anatomy in a model mainly about diabetes is likely to be less semantically precise than would be the case if the model was specifically about anatomy alone.

Decreasing semantic precision as you move further from the core of an ontology is responsible for at least some of the challenges that arise when two separate ontologies are required to interoperate for a new, common purpose. For example, an ontology used for describing diabetic patients and their clinical state might be required to interoperate with another ontology describing the indications and side effects of diabetic medication. Whilst both ontologies - one focussing on the patient’s clinical state and the other their drugs - might be expected to include some notion of anatomy, an ontology of anatomy will be peripheral to both. If what little anatomical content they share is represented using different assumptions and semantic simplifications, then there may be no simple or direct mapping between them.

Consider, for example, an imaginary diabetes disease ontology that allows abscesses to be localised to:

> *finger, hand, forearm, arm, shoulder, neck, face, scalp, chest, abdomen, back, thigh, calf, shin, forefoot or toe.*

A therapeutic ontology may, by contrast, suggest antibiotics for abscesses located in the skin. This exposes the contextual semantic simplification made in the disease ontology –

that 'finger' really means 'skin of finger'. Such an ontological gap between what is said and what was meant must, however, be bridged if the two ontologies are to work together. But gaps like this are often wider than they initially appear: the same picking list of anatomical parts might also be employed in the same disease ontology as the location of paraesthesia or ischaemic pain, and in these contexts it is doubtful that 'finger' still means 'skin of finger'.

In practice a major barrier, when trying to link or otherwise merge related ontologies, is the presence of multiple such ontological gaps scattered across the many peripheral ontologies that are central to neither concern but shared by both.

### 3.4.1 Ontological gaps: The mirage of single ontologies

Medicine has a long tradition of constructing terminologies or ontologies covering specific medical subdomains. The Unified Medical Language System (UMLS) is a testament both to the number of discrete medical concepts that have collectively been enumerated (more than 975,000) and to the difficulties of achieving interoperation between the different subdomain ontologies from which these concepts have been collated.

One solution posed to the problem of semantic gaps in medicine is to construct a single, all encompassing and coherent ontology to cover all of medicine in one structure. There would be no ontological gaps because interoperation with other ontologies would no longer be necessary.

However, this solution overlooks the risk that ontological gaps may exist within the single ontology itself as a result of its construction. A large ontology does not come into being instantaneously and at full scale, rather it is necessarily built iteratively in stages and grows larger as more is added to it. As with most large scale and complex engineering tasks, the prevailing approach is to modularise the construction. In essence, this amounts to building multiple discrete but overlapping subontologies, as a result of which ontological gaps are a risk. The potential for such gaps grows as the number of authors working on the project – either concurrently or sequentially – increases: with multiple authors come multiple, potentially conflicting assumptions and viewpoints.

### 3.4.2 Ontological Gaps: the mirage of ontology reuse

An alternative and superficially more evolutionary approach is to advocate the gradual replacement of semi- or un-structured, rough-modelling of peripheral concepts with complete and principled third party ontologies centred on those concepts. Thus, for example, instead of a small list of anatomical parts as possible locations for abscesses, the builder of a diabetic pathology ontology would eventually substitute an entire model of anatomy.

Such reuse and creeping integration has obvious attractions – notably the potential for reduced development cost and the possibility of achieving interoperability *post hoc*. However, there is also an obvious flaw: if all domain ontologies followed the guiding principle of inclusive reuse of peripheral ontologies then it would soon become impossible to build an ontology that did not model the entire world by extension. In reality the evolutionary approach may not be so different from the single universal ontology approach.

A further proposed compromise, therefore, is to include by reference only salient sections or elements of such well-formed peripheral ontologies. But semantic interoperability depends only partly on whether concept identifiers are shared: also important is a shared overall ontological schema. Sharing a reference to an identifier for the concept 'drinking glass' will not achieve interoperability between two ontologies if one describes them along a 'fullness' scale and the other along an 'emptiness' scale.

### 3.5 Domain Complexity: Multiaxial Navigation

Although almost all traditional medical terminologies include some classification of the concepts they contain, these classifications support only those certain analyses for which they were originally designed. For example, ICD9 includes a top-level abstract category called ‘Respiratory Diseases’ but has no similar category for ‘Occupational Diseases’. Even within the classification it *does* support, ICD sometimes offers an idiosyncratic view: the ‘Respiratory Disease’ category does not include pulmonary tuberculosis, or indeed any other infectious lung disease. These are instead aggregated under a separate top level ‘Infectious Disease’ category. This categorisation of medicine reflects the fact that infectious diseases and tuberculosis remain significant present-day clinical concerns in the 3rd world constituency for which WHO maintains ICD, but are correspondingly less significant in the rich industrialised nations where ICD underpins many medical billing systems.

Increasingly the requirement is for coded content in clinical records to be much more flexibly analysable. This requirement often translates into demand for multiaxial classifications of the concepts with which the record is represented: ‘pulmonary tuberculosis’ should appear in *both* the ‘Respiratory Disease’ and ‘Infectious Disease’ chapters. However, opening the door to multiple classification can lead to surprisingly complex structures, particularly as an ontology expands to accommodate more ontological viewpoints. For example, ‘pulmonary tuberculosis’ could also simultaneously be classified in future chapters for ‘Risk factors for cancer’, ‘Comorbidities of HIV’ and ‘Diseases with a genetic predisposition’ (see OMIM 607948, 607949 and 300259).

Users accustomed to simple monohierarchical schemes such as ICD9 may experience difficulty navigating such multiaxial structures to find terms [Bentley 1999]. Partly this relates to the significant expansion in navigational choices that ensue. However, a more important factor is that the simpler structure of ICD9 has its origin in the widely shared conventions and experience of how terms are organised in the medical domain. Although intuitive, the organisation of monohierarchical schemes is often heuristic and alogical. Whilst alogical multiaxial navigational schemes could also be constructed, when such schemes are devised to support improved analysis they typically make a point of striving for a more logical consistency and completeness of classification.

### 3.6 Artefactual Complexity: Formal Syntax

Formal expressions are necessarily precise and have to be syntactically unambiguous. The following example – a representation from the *Open*GALEN CRM of the familiar clinical notion of *tenesmus* – provides one extreme example of the surface complexity of a raw serialised notation (in this case, GRAIL):

```
(ClinicalSituation which <isCharacterisedBy (presence which isExistenceOf (ContractionProcess which <
   isSpecificFunctionOf SphincterAniMuscle hasImmediateConsequence Pain hasIntentionality
(Intentionality which hasAbsoluteState involuntary) hasDuration (Duration which hasAbsoluteState longTerm)
hasTemporalPattern (TemporalPattern which hasAbsoluteState ongoing) >)) isCharacterisedBy (presence
which isExistenceOf (UrgeToVoidUrineOrFaeces which hasProcessActivity (ProcessActivity which
hasQuantity (Level which hasMagnitude highLevel)))) isCharacterisedBy (presence which isExistenceOf
AbdominalStraining) >)
```

**Figure 6: Example of serialised GRAIL notation**

The native syntax in other representations (for example, OWL) can be still more opaque partially because they are designed by logicians, not ontologists, and therefore are phrased to expose the logic rather than the ontology. They are also normally encoded in surface syntaxes and representational languages (e.g. RDF/XML) that are explicitly intended to be optimised for machine, and not human, readability.

## *3.7 Artefactual Complexity: Formalism Limitations*

Any concept system intended to be reasoned over automatically will be subject to some expressive limitations: full first order logical inference is not computationally tractable in the general case. All practical computational formalisms must, therefore, to some extent be limited to less than full expressivity, though where they choose to impose the limitations may be different.

The GRAIL formalism underpins the *Open*GALEN Common Reference Model and the experiments detailed in this thesis. Among its more significant limitations are the complete absence of either negation or disjunction constructors, and a limited implementation of cardinality: only two values of cardinality are available (one or many), and the cardinality of a given semantic link is embedded within its definition and can not be varied according to the specific context in which that link appears. Qualified cardinality constraints that *are* sensitive to context (e.g. permitting the assertion that the semantic link *hasPart* takes at most 5 values when connecting Hand to one or more Digits, but up to 13 values when connecting Thorax to Rib) are already provided in some other logics, but are an area of continuing controversy in others such as OWL.

GRAIL's limitations can to a certain extent be worked around within the ontology, but such workarounds give rise to further complexity. Whilst choosing a different formalism might remove the need for the particular workarounds introduced in ontologies based specifically on GRAIL, it is likely that the different limitations of the alternative formalism will require other workarounds. Loom, for example, provides no native support for managing partonomy in anatomical relationships equivalent to that provided in GRAIL through its role inheritance constructors [Rector 1995a]. As a result Schulz and Hahn simulate it within the ontology using a technique ("SEP Triples") that introduces considerable complexity [Schulz 2000]. Similarly, although more modern description logic classifiers such as FaCT or RACER now implement true logical negation, they cope very badly with the kinds of qualified negation ('probably not', 'presumably not', 'definitely maybe') found in human discourse.

Three examples of the techniques employed in the *Open*GALEN CRM to workaround GRAIL's limited implementations of cardinality and of negation are presented below:

### 3.7.1 GRAIL cardinality artefact workaround: Feature-State

In the *Open*GALEN Common Reference Model (CRM), many axes of description that can describe phenomena are found beneath [Feature] (e.g. [Temperature], [Severity]), and the corresponding value choices along those axes are found beneath [State] (e.g. [Hot/Cold], [Mild/Moderate/Severe]). In general, [State] valuesets (e.g. [cold, warm, hot]) are mutually exclusive, and subclasses of [Feature] (e.g. [Temperature]) can have one value from each set (e.g. only one from [cold, warm, hot]). The *Open*GALEN CRM achieves this constraint through single valued cardinality of the *hasState* link, giving rise to a general schema for **Feature-States** as shown in Figure 7:

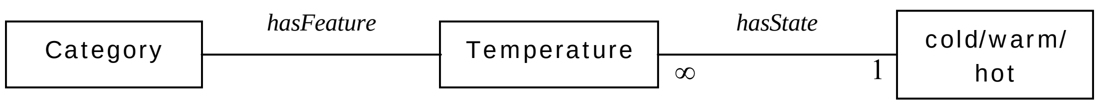



**Figure 7: General Feature-State schema**

However, some subclasses of [Feature] can take several values provided they are from sets measured along different dimensions. For example, the body [Temperature] of a patient could be [raised, increasing, lower] than the last reading and [higher] than expected (for the condition). In this scenario, [Temperature] is simultaneously described along four dimensions: by an absolute value, a trend, a comparison with an earlier reading and a comparison with the expected value.

To allow for such scenarios, the *Open*GALEN Common Reference Model has been adapted to allow a **[Feature]** such as **[Temperature]** to take orthogonal values simultaneously from more than one value set (e.g. **[rising]** from **[rising / constant / falling]** as well as **[low]** from **[high/low]**). This is achieved by enumerating the different dimensions as a family of single valued cardinality links, each a subtype of the original [*hasState*] link. Figure 8 shows the overall schema, together with a worked example in GRAIL syntax:

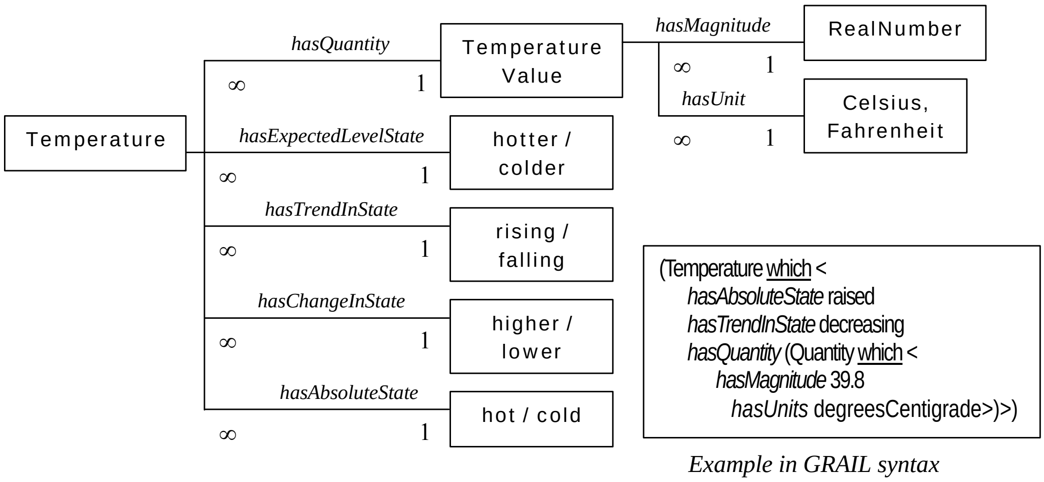



**Figure 8: Schema for different dimensions of features**

This schema prohibits semantically nonsense expressions such as: **[Temperature –high & low]**. However, other nonsense compositions are still possible. For example, one object could still be simultaneously linked to two different temperature entities: [**Temperature – high]** and **[Temperature – low]**. Attribute cardinality again provides a solution - the *hasFeature* link between **[Patient]** and **[Temperature]** can also be single valued (Figure 9):

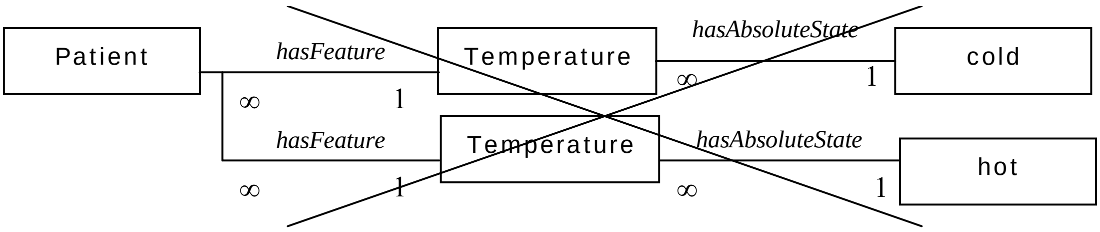



**Figure 9: Cardinality of hasFeature preventing nonsense composition ⊗**

However, whilst this schema prevents linking to more than one kind of the same **[Feature]**, it would also prevent linking to more than one *different* **[Feature].** If [*hasFeature*] is single valued, then it is not possible to describe a patient with a normal temperature but increasing weight (Figure 10):

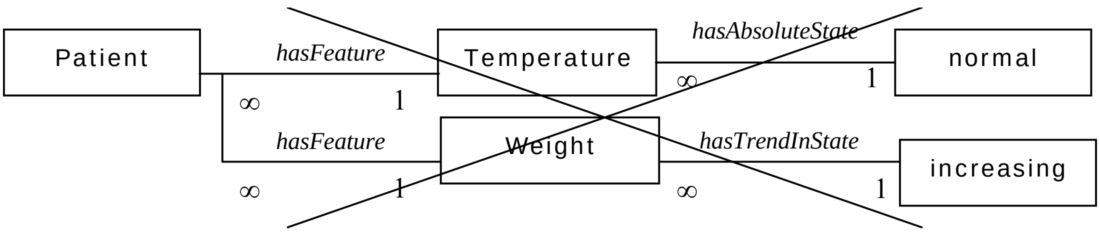



**Figure 10: Cardinality of hasFeature preventing sensible composition ⊗**

A further work-around is therefore required: the complete *Open*GALEN CRM schema for features and states employs a different, single-valued linking attribute for every Feature.

**[Frequency]** may only be linked to via [*hasFrequency*], **[Shape]** via [*hasShape*] and so on. Each such dedicated attribute is a child of [*hasFeature*], giving rise to expressions such as in Figure 11:

```
(Tumour which <
    hasTemperature (Temperature which <
        hasAbsoluteState  hot
        hasTrendInState  increasing>)
    hasTexture (Texture which
        hasAbsoluteState  smooth)
    hasMass (Mass  which
        hasQuantity (Quantity which <
            hasMagnitude 100
            hasUnit  gram>))>).
```

**Figure 11: Example of representation using full Feature-State schema**

This schema owes much to the work of Yuval Shahar and is sufficient to capture the key ideas from his work on temporal abstraction [Shahar, Tu et al. 1992; Shahar, Das et al. 1994].[3]

### 3.7.2 GRAIL cardinality artefact workaround: 'specific' semantic link types

The GRAIL formalism supports role inheritance of semantic links. However, like cardinality, this is a global property of a link and insensitive to the semantic contexts in which a link may appear.

A consequence of these limitations is that if the same flavour of semantic link can appear in a GRAIL ontology with different cardinality or role inheritance, then this can only be achieved by instantiating in the ontology a family of links possessing the different permutations of cardinality and transitivity needed. Thus, the single link flavour '*actsOn'*, used to link a **Process** with the **Structure** acted on by the process, is reified in the *Open*GALEN CRM as a tree of four link objects (Figure 12):

| LINK | CARDINALITY | TRANSITS |
|---|---|---|
| actsOn | Many - many | Includes is-part-of |
| actsMultiplyOn | Many - many | Includes is-part-of |
| actsSpecificallyAtLevel | One - many | *excludes* is-part-of |
| actsSpecificallyOn | One – many | Includes is-part-of |

**Figure 12: Cardinality and transitivity permutations reified as family of links**

…where, for example, the link *actsSpecificallyAtLevel* is reserved specifically for use with the process of 'amputating'. The normal role inheritance rule shared by the other flavours of *actsOn* holds that:

Process *actsOn* Part *isPartOf* Whole    subsumes    Process *actsOn* Whole

However, this rule does **not** apply to amputations. An amputation of the foot is *not* a kind of amputation of the leg, even though the foot clearly *isPartOf* the leg.

When modelling in GRAIL, therefore, the author must first determine what cardinality or transitivity constraints are appropriate in context and then know how to express that representational choice as a selection between different semantic link subtypes.

---

[3] The need for the workaround described was a known limitation of GRAIL, and was addressed in the specification for DAML+OIL by means of qualified cardinality restrictions. However, this construct at the time of writing was explicitly omitted from the specification for OWL.

### 3.7.3 GRAIL negation artefact workaround: Wrappers

In contrast with more modern logics, the GRAIL formalism itself provides no mechanism to represent negated phrases such as are commonly found in lists of medical diseases and procedures:

Ulcer **without** haemorrhage
Reduction of fracture **without** fixation

To address this difficulty, some aspects of the concepts of negation and inclusion are explicitly modelled as concepts within the *Open*GALEN Common Reference Model itself. The general pattern is to take an expression of the form

'**A** with **B** and without **C** and with **D**'

…and re-write it so that the inclusion of **A** is made explicit. This requires an additional concept **X**, not in the original statement:

'**X** with **A** with **B** without **C** and with **D**'

Within the *Open*GALEN CRM ontology, two concepts (**presence** and **absence**) model inclusion and negation respectively, giving rise to the final schema (Figure 13):

```
X with
    presence of A
    presence of B
    absence of C
    presence of D
```

**Figure 13: GRAIL 'wrapper' workaround for negation**

The additional **X** concept is required in order to provide a node in the overall graph structure from which the various 'with's and 'without's may be hung. The arms of the resulting graph structure represent the negated and included elements of the original string expression, while the repeating (X with presence/absence of…) pattern wraps the set of elements into a single concept. Schema workarounds such as this are therefore known as 'wrappers', with this example being designated the negation wrapper schema.

The wrapper workaround for negation comes at a considerable price: it is not sufficient, at least for a formal ontology that is to be reasoned over, to wrap only those concepts that include negation. All concepts, including the trivial cases where only one, non-negated concept is included, must be wrapped if automatic classification is to proceed as expected. Figure 14 demonstrates why this is necessary. The concept 'haemorrhage' should subsume 'ulcer with haemorrhage', but this can only occur in GRAIL if both concepts are expressed using the wrapper structure (first row of figure):

| | | |
|---|---|---|
| X with<br>presence of ulcer<br>presence of haemorrhage | will be subsumed by | X with<br>presence of haemorrhage |
| Ulcer with<br>presence of haemorrhage | will not be subsumed by | haemorrhage |

**Figure 14: Need for univeral use of wrappers**

## 3.8 Cognitive Complexity: Natural Language

A well recognised phenomenon of human discourse, in natural language, is that much relevant information is only implied by the context. Less well recognised is that some of the explicit semantic information may be redundantly repeated within an expression, because of normal patterns of syntactic agreement.

In normal speech we speak of inserting a device *into* a site, but removing it *from* the same location. This linguistic convention directly influenced the results obtained when participants of the GALEN-IN-USE project attempted to describe the surgical insertion or subsequent removal of a device, such as a heart valve or other prosthesis.

All dissections authored were collected centrally by the author of this thesis into a single Dissection Library. The TIGGER tool (see Chapter 10) allowed this library to be searched for all instances of a dissection, from any dissection author, where the descriptors inserting or removing were used. The set of dissections so identified were inspected by the author of this thesis, and the range of patterns used by dissection authors abstracted. These are summarised in Figure 15:

```
MAIN inserting                    MAIN removing
   ACTS_ON device                    ACTS_ON device
      INTO anatomy                      HAS_LOCATION anatomy

MAIN inserting                    MAIN removing
   ACTS_ON device                    ACTS_ON device
      HAS_DESTINATION anatomy           HAS_SOURCE anatomy
```

**Figure 15: Alternate schemas for inserting and removal**

Whilst the authors did not agree over the exact name of the semantic link between the device and the anatomy, there was universal agreement that the name of the link should be different depending on whether the procedure was an insertion or a removal. However, use of different links in this manner causes a potential problem: in a final, classified set of descriptions of such procedures we would expect that both 'insertion of a device into anatomy' and 'removal of a device from anatomy' could be classified to be kinds of 'some kind of procedure to a device in anatomy', as illustrated in Figure 16:

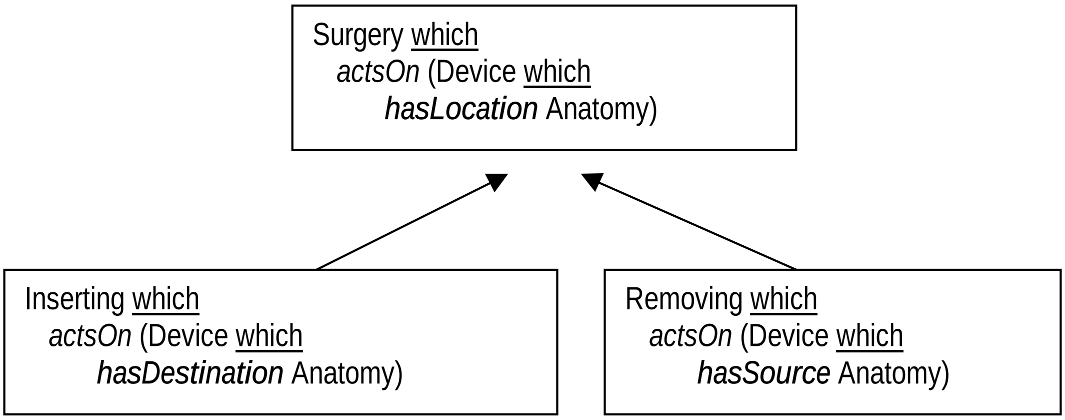



**Figure 16: Intended classification of insertions & removals**

Two solutions could be employed to achieve this intended result:

a) [*hasDestination*] and [*hasSource*] are retained as subtypes of a more general semantic link: [*hasLocation*], and [Removing] and [Inserting] are also subsumed by [Surgery]

b) Only a single semantic link [*hasLocation*] is used whilst [Removing] and [Inserting] remain subsumed by [Surgery]

On reflection it is obvious that, were option (a) pursued, then the concept immediately preceding the link (insertion or removal) predicts which specialised subtype of [*hasLocation*] should be used (either [*hasDestination*], or [*hasSource*]). This suggests that having two different links is simply duplicating a semantic distinction already represented explicitly by the choice of the deed itself.

Therefore, in the *Open*GALEN CRM, option (b) was selected and *all three* procedure descriptions are represented using a single semantic link, [*hasLocation*], as shown in Figure 17:

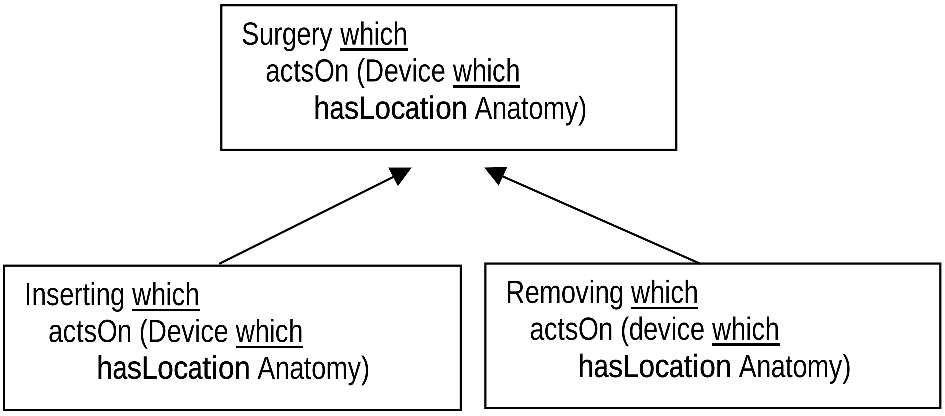



**Figure 17: OpenGALEN schema for insertions & removals**

In this scenario, therefore, the linguistically inspired intuition of the inexpert user is to insert more detail than is required in the *Open*GALEN ontology. Resisting this intuition requires an awareness of the complex relationship between what is intended by language, and the expressions by which it is communicated.

## *3.9 Cognitive complexity: Need for ontological precision*

Ontologies aim to provide a precise and unambiguous means to represent what was meant. Unfortunately, humans are often imprecise or ambiguous about what they want to say. A tension exists in computed ontologies: human users expect the machine to understand roughly what they meant, whilst machines require humans to mean precisely what they said. The larger the ontology, the greater the number of semantic contexts where this tension arises, and the lower the capacity of the user to remember all contexts or the preferred formulation in each.

In making their representational choices, therefore, users of an ontology typically make three kinds of error. They select, or construct as compositions, concepts whose meaning is:

- entirely different from that intended
- more specific than intended (excessive precision)
- less specific than intended (insufficient precision)

The remainder of this section considers each of these scenarios in turn. Several examples of complex *Open*GALEN CRM schemas are presented in detail. They will be referred to in subsequent chapters, where different mechanisms are presented for managing each in order to reduce the ontology user's cognitive burden.

### 3.9.1 Cognitive Complexity: choosing or constructing the wrong concept

The commonest reason a user selects entirely the wrong concept is when the knowledge name of a concept in a pre-enumerated list is misinterpreted.

For example, the 5 byte READ codes include the code and rubric pair '19C.. Constipation', and this is widely used in UK Primary Care to record patients where constipation is the presenting problem. However, 19C.. is in fact the common parent of '19C1. Constipated' and '19C2. Not constipated'. A correct interpretation of such patient records, semantically valid with respect to the true meaning of the recorded code, is therefore 'This patient may or may not be constipated'.

Another cause arises when a user seeks to construct a new composition, and where the ontology treats as different concepts the user perceives to be identical:

Returning to the half empty or full glass of water of section 3.2, a formal ontology takes no account of whether both concepts might reasonably be applied to the same physical instance of a particular glass of water[4]. Instead, it focuses exclusively on the explicit statement of two distinct measurement scales (emptiness, fullness) in the definitions of each concept, and notes that the two original notions are not equivalent because they have different intended meanings[5]. Computed ontologies restricted to reasoning about intended meaning, without reference to usage, will always hold that the two concepts are different.

Humans are less fussy and more capable than computers, and often fluidly shift the focus of their interpretation from intended meaning to actual usage. Many users would be surprised by an ontology that could not recognise a half full and half empty glass of water as ‘the same thing’. However, the widespread use of the rhetoric by which an optimist is spoken of as one who always sees the glass half full, and a pessimist as one who always sees it half empty, is perhaps recognition that the same distinction is semantically valid and understood at some level of human discourse.

#### 3.9.2 Cognitive Complexity: excessive precision

When trying to describe a real physical instance before them (e.g. a patient) by selecting concepts from even a very large pre-enumerated list, users will often find that there is no single term that exactly captures all the known properties of the subject to be described. Rather, they may be presented with a choice between either

> (a) a relatively abstract term that represents little of the interesting specific detail that they wish to represent about the subject

or

> (b) a very detailed term whose meaning captures much or all of the interesting detail known about the object, but also includes further detail that is not known to be either definitely true or false of the object.

Faced with this choice, many users will choose option (b).

For example, as a teaching exercise within a module on terminology written by the author of this thesis for the Royal College of Surgeons of Edinburgh distance learning course in medical informatics, the painting shown in Figure 18 was presented to thirty-nine informatics students.

| | |
|---|---|
| Woman……… | 23 students |
| Adult………….. | 4 students |
| Model…………. | 4 students |
| Person………… | 3 students |
| Man…………… | 2 students |
| Trumpeter…….. | 2 students |
| Sitter………….. | 1 student |
| TOTAL | 39 students |

**Figure 18: Coding 'Heart of the matter' by Rene Magritte**

[4] ie that they share a common extension of how they may be used

[5] ie they have different intensions

They were then asked to describe its subject matter by selecting terms from the Art and Architecture Thesaurus, a controlled vocabulary of terms for cataloguing works of art or architecture, maintained by the Getty Museum. The terms selected by the students for the human form in the painting were as shown.

Even though the figure's face is entirely obscured, the majority of students believed they could reliably discern its sex, but there was not universal agreement on what the sex was. Two students chose a term indicating that the figure could play the instrument in the painting. Less than a quarter of the students felt that the available information could not support being so precise on either the sex of the figure or their musical skills, and so opted for more abstract terms such as adult, person or model.

#### 3.9.3 Cognitive Complexity: insufficient precision

The potential for a user to select terms from an ontology that are less precise than needed arises particularly when the ontology requires distinctions to be made that are unusual to conscious thinking by all or most user groups. Two examples of such distinctions are presented below: the models of part-whole relationships and of the process-structure duality.

A further potential cause arises when a schema for a single topic integrates the fine grained semantic distinctions used only by a few specialists with the broader distinctions required by generalists. Schemas for cancer and pain are presented as examples of this phenomenon.

##### 3.9.3.1 Insufficient precision due to unfamiliar distinctions: Partonomy

It is inconceivable to have a model of the medical domain that did not include a significant number of concepts for anatomical parts. More than this, it is likely that such a model will require considerably more than mere mention that the anatomical parts exist; detailed information about how they relate to each other in an anatomical sense will be required.

Among the many possible anatomical relations found in standard medical anatomy texts - such as spatial (superior_to, lateral_to etc) or information about innervation and vascular supply – one in particular stands out as an absolute requirement for inclusion: information about how sub-parts of anatomy relate to larger parts. The need for part-whole, or partonomic, relationships to be explicitly declared within the model arises because, in medicine and many other domains, humans apply the rules such as that:

> 'A part of a part is a part of the whole'
> 'A disease of a part of an organ is a disease of the whole'
> 'An operation on a part of an organ is a surgery to the whole organ'

Thus, the *mitral valve* is part of the *left heart*, which is turn is part of the *heart*, and because of this the *mitral valve* is also part of the *heart*. Similarly, and because of that part-whole relationship, *mitral valve prolapse* is conventionally considered a kind of *heart disease*, and repair of such a prolapse is a kind of *cardiac surgery*.

However, in order to be able to apply these rules in all situations, it is necessary that the computer has access to a model that states exhaustively all the direct part-whole relationships that exist between all parts of anatomy mentioned in the model. By 'direct' part-whole relationships is meant that, for every piece of anatomy, only its largest subparts must be declared.

Thus, when considering the heart itself it is only necessary to say that it has two direct part-whole relationships: one each to the right and left hearts. Each of these will separately have further relationships to smaller pieces of anatomy. An abridged model of the heart to illustrate the principle is shown in Figure 19:

| | | |
|---|---|---|
| heart | HAS_PART | left heart, right heart |
| left heart | HAS_PART | left atrium, left ventricle, aortic valve, mitral valve |
| right heart | HAS_PART | right atrium, right ventricle, pulmonary valve, tricuspid valve |
| left atrium | HAS_PART | left auricle, mitral valve vestibule |
| right atrium | HAS_PART | right auricle, tricuspid valve vestibule |
| left ventricle | HAS_PART | aortic valve vestibule, left bundle branch, mitral chordae |
| right ventricle | HAS_PART | pulmonary valve vestibule, right bundle branch, tricuspid chordae |
| aortic valve | HAS_PART | aortic orifice, aortic ring, aortic sinus, aortic valve lunule |
| mitral valve | HAS_PART | mitral orifice, mitral papillary muscle, mitral ring |
| pulmonary valve | HAS_PART | pulmonary orifice, pulmonary ring, pulmonary valve lunule |
| tricuspid valve | HAS_PART | tricuspid orifice, tricuspid papillary muscle, tricuspid ring |

**Figure 19: Abridged model of cardiac anatomy**

This information alone is sufficient to infer, by applying the 'part of a part is part of the whole' rule, that e.g. both the tricuspid ring and all of the aortic valve lunules are parts of the heart.

However, the model fragment as declared above is simplistic, and (deliberately) omits the exceptions and special cases that apply when the heart, and other anatomy, is modelled in more detail.

For example, how should the ventricular septum be modelled? Is it part of both the left and right ventricles simultaneously? If so, then all the substructures within it – such as both the left and right bundle branches – would surely also be part of both ventricles. However, this would mean that a disorder of the right bundle branch would, through the application of the 'disease of a part is a kind of disease of the whole' rule, become a disorder of both the left and right ventricles – not a result that many clinicians would expect or support.

Similarly, is the pericardium actually part of the heart? Clinicians normally consider *pericardial disease* to be a kind of *cardiac disease*, from which observation we might infer that the pericardium is a kind of cardiac substructure. However, consideration of organogenesis and embryology suggests otherwise: the heart collides with and is enveloped by the pericardium, but developmentally the pericardium is no more part of the heart than the vertebral column is part of the spinal cord, or the peritoneum part of the liver. The NLM-funded UWDA Digital Anatomist project [Rosse 1998], released through UMLS, is one authoritative source that has taken an especially strong position on this matter, declaring that the pericardium is not anatomically part of the heart.

Consider also the concept of a leaking gastrojejunostomy, normally thought of as both a kind of gastric problem and a kind of jejunal problem. This traditional classification suggests that an anastomosis itself, as a structure, is part of both objects connected by it. However, this rule does not seem to apply so well to the notion of a leaking superficial colostomy which, though it would be accepted as a colonic problem, is less naturally conceived as a skin (dermatological) problem.

In modelling real world anatomical part-whole relationships in a formal ontology, therefore, the model author must cope with these exceptions. Typically, this requires that exceptions be distinguishable from the usual case. One method to achieve this is through the use of more than one flavour of the PART_OF link.

A detailed description of the different flavours of PART_OF used in the OpenGALEN CRM is outside the scope of this thesis but is given in [Rogers 2000]. For the purposes of this thesis it is sufficient to say that whereas the inexpert user might attempt to construct a complete model of anatomical part-whole relationships using but one link (PART_OF), the Digital Anatomist model employs 6 flavours of partonomic link whilst the OpenGALEN CRM is constructed using 16 different variants arranged into a hierarchy (Figure 20):

| | |
|---|---|
| part-of | Parent of all partitive relations |
| structure-part-of | Parent of all partitive relations between structures |
| arbitrary-part-of | Artificial clinical constructs (e.g. hand, foot and mouth) |
| segment-of | Division of a linear or tubular structure (e.g. artery, nerve) |
| solid-piece-of | Divisions of a homogenous 3 dimensional structure |
| layer-of | … plane of section parallel to flat surface (e.g. mucosa and stomach) |
| irregular-piece-of | … plane of section not parallel to a flat surface (e.g. lobe of liver) |
| pouch-of | … a pouch-like extension of a hollow structure (e.g. appendix) |
| component-of | Subcomponents of an assembly |
| func-component-of | …where the functions of the component are inherited to the whole |
| partitive-connection-of | …where the component is also connected to the whole |
| partitively-contained-in | …where the component is also contained in the whole |
| surface-of | Division within a planar structure (e.g. hypochondrium and abdomen) |
| makes-up | Division of a substance |
| subprocess | Division of a complex process |

**Figure 20: Abridged list of partonomic links in OpenGALEN CRM**

The choice of specific partonomic links, and the rules that govern which should be used in a given context, offers a greater level of ontological precision than an inexpert user would expect or easily comprehend.

#### 3.9.3.2 Insufficient precision due to unfamiliar distinctions: The Process-Structure dual

An important and very high level ontological decision within the *Open*GALEN Common Reference Model is that the world in general shall be described in terms of:

- *processes* (that have a duration in time) acting on
- *structures* (that have a size in space) made of
- *substances* (that have neither duration nor size),

These three fundamental and disjoint types of *phenomenon* all may additionally be described by various:

- *features* (temperature, colour etc.) that can take specified
- *state* values (hot, cold, red, severe) for a duration of time.

This particular 'upper ontology' [Rector 1996] is not unique to *Open*GALEN and other works do also successfully employ different upper ontologies. A detailed comparison or discussion of these various upper ontologies is beyond the scope of this thesis, and the interested reader is referred to other works such as by Asuncion Gómez-Pérez [Gómez-Pérez 2004]. This thesis proceeds from the assumption that all ontologies must make some basic high level decision of how to conceptualise the world. Having made such choices, the ontology should thence be populated in a manner that rigidly enforces consistency with those basic decisions, if computation over the population of terms is to behave consistently.

Rigid enforcement of the *Open*GALEN upper ontology can have anti-intuitive consequences. For example:

In medical speech we freely move between saying

'the patient has had *ulceration* of the lower legs for two months' (1)
'the *ulceration* is 10cm in diameter' (2)
'the *ulceration* is not responding to treatment' (3)

…without thinking about the fact that, in sentence (1), *ulceration* refers to the process of ulceration (which can be described by a duration), in sentence (2) it refers to the resulting lesion (described by a dimension), and in sentence (3) it is ambiguous between the two.

In a medical model, lesions as a category could be represented either as something (e.g. 'an ulcer') located in a part of the body, or as something (e.g. 'ulceration') acting on the same part of the body. Whilst a human might view either solution as semantically equivalent, an ontology based on the *Open*GALEN upper ontology could not: one is (fundamentally) a process while the other is the structure that results from that process. Although there is clearly a strongly implied causal relationship between the process of ulcerating and the resulting ulcer, they are most definitely not the same thing. The structure, for example, might reasonably be directly further described by its physical dimensions, whereas this would not be an appropriate thing to say of the process.

The importance of this fundamental ontological divide between structure and process becomes evident when you consider how the concept of "a large leg ulceration of 2 weeks duration" should be represented. A inexpert user might seek to write (Figure 21):

```
Ulceration
    hasLocation Leg
    hasDuration 2 weeks
```

**Figure 21: Naive representation of 'large leg ulcer of 2 weeks duration'**

However, if the upper ontology is to be followed, the duration and the size **can not** be attached to the same concept, because *size* is a property of structures and *duration* only of processes, and it is not possible to be simultaneously a structure and a process. The ulcerating process and the ulcer structure must be separated out, but the original phrase "a large leg ulceration of 2 weeks duration" is ambiguous on this specific distinction. It is a matter of opinion whether the concept as a whole should be represented as, primarily, a kind of structure or as a kind of process (Figure 22):

| 'ULCERATION' EXAMPLE AS PROCESS | 'ULCERATION' EXAMPLE AS STRUCTURE |
|---|---|
| *UlceratingProcess*<br>    *hasDuration 2 weeks*<br>    *hasConsequence (Ulcer Structure*<br>        *hasSize large*<br>        *hasLocation Leg>)* | *UlcerStructure*<br>    *hasSize large*<br>    *hasLocation Leg*<br>    *isConsequenceOf (UlcerationProcess*<br>        *hasDuration 2 weeks)* |

**Figure 22: Ulceration example as structure and process (pseudocode)**

Figure 23 demonstrates visually how each of these two alternative representations is, in fact, a graph transform of the other: the ulceration as process representation (top of figure) is transformed into the ulceration as structure representation (bottom of figure) by rotating a fragment through an arc indicated by the arrows. However, it remains true that they represent concepts with different intensions, within the semantics of the high level ontology adopted.

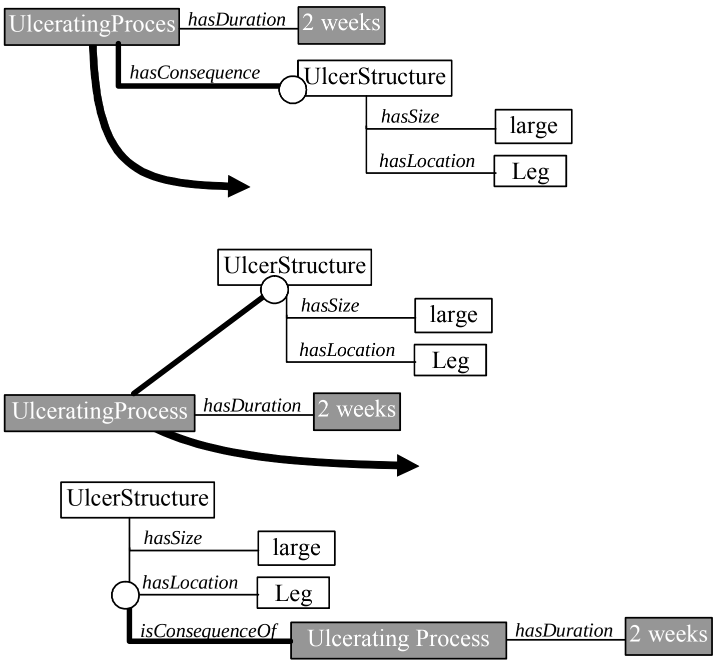



**Figure 23: Graph transformation of ulceration example**

Still further complexity arises from the fact that, were things to remain as outlined so far, the ontology now contains scope for a number of possible redundancies. If [UlcerStructure] existed in the model as a primitive descendent of [Lesion], and simultaneously [UlceratingProcess] existed under [PathophysiologicalProcess], what would be the difference between the composition [Lesion *isConsequenceOf* UlcerationProcess] and the primitive entity [UlcerStucture] ? Similarly, what would be the difference between [PathophysiologicalProcess *hasConsequence* UlcerStructure] and [UlceratingProcess] ? To remove this redundancy it is necessary that, for each such process-structure dual, only one is a primitive entity in the ontology while the other is expressed compositionally and in terms of the first.

In summary, high level decisions in an ontology seek to divide the world rigidly between categories. The schema for further axes of description of categories, such as size and disease course, are dependent on precise use of such high level distinctions.

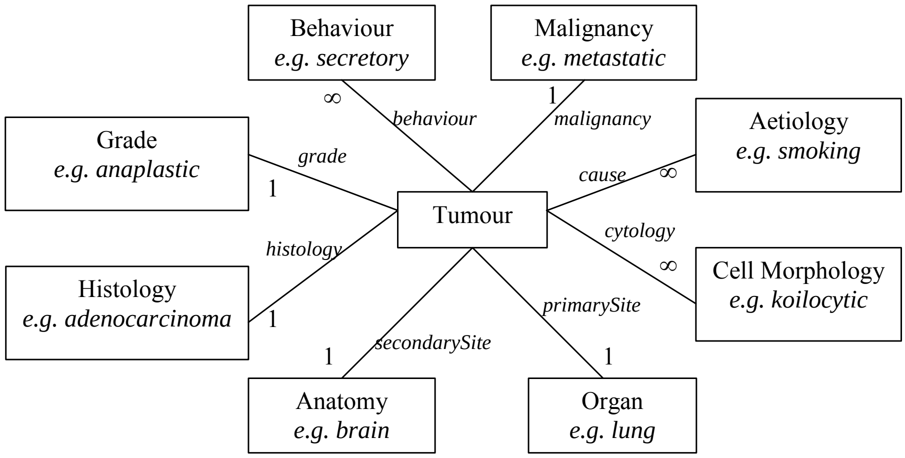
#### 3.9.3.3 Insufficient precision due to integrated schemas: Schema for Cancer

Most clinicians trying to describe a tumour will already be aware that such descriptions can be complex. For example many practitioners will recognise possible characterisation of a single tumour along a number of axes including:

*Anatomical structure from which tumour originally arose*
*Cell type from which tumour originally arose*
*Cytological features of the neoplastic cell line (e.g. spindle cell)*
*Functional features of the neoplastic cell line (e.g. gastrin secreting)*
*Histological features of the gross tumour (e.g. cystic)*
*Whether metastatic, or locally invasive, or not*
*Aetiology*

**Figure 24: Commonly recognised axes for describing a cancer**

However, whilst they may recognise the need for multiple axes of description, they may overlook the detail of how the axes relate to each other. A inexpert user’s schema for describing cancer, therefore, might appear as shown in Figure 25.




EXAMPLE IN SIMPLIFIED SYNTAX

*Tumour*

*hasPrimarySite Lung*
*hasSecondarySite Lymphnode*
*hasHistology Adenocyte*
*hasCytology Koilocytic*
*hasGrade Well Differentiated*
*hasBehaviour Mucin Secreting*
*hasMalignancy Metastatic*
*hasCause Smoking*

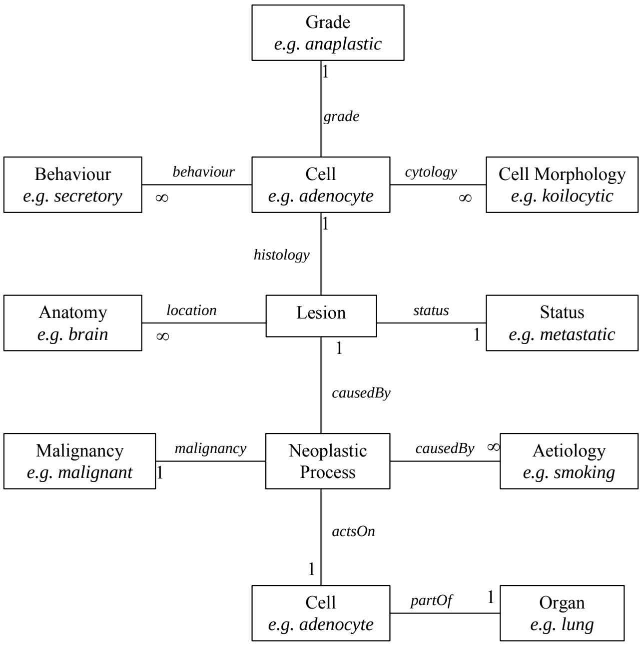
**Figure 25: Naive schema for cancer** ⊗

A richer model, however, might represent the fact that the original adenocyte from which the tumour arose was *partOf* the primary site. Similarly, that secreting mucin, being koilocytic and being well differentiated are properties of a concept not explicitly identified in the simple representation: the new neoplastic tumour cell line.

Being metastatic could then be represented as a *feature-of* of the neoplastic process that *causes* the tumour, rather than of the tumour directly: both the original tumour and any discrete metastases from it are metastatic (meaning they have the property of being able to metastasise). This distinction implies the existence of a ‘metastasising process’, which is the causative pathological process that give rise to an individual metastasis, and which could also be teased out.

Finally, the concept [Tumour] itself could be expressed in terms of the process-structure dual model already described, where a distinction is made between the physical object of a tumour, and the neoplastic process of cell division that causes it.

This more detailed analysis of the explicit and implicit concepts in the original representation, and of how they may relate to one another, gives rise to a more complex schema in the *Open*GALEN CRM, as illustrated below:




EXAMPLE IN SIMPLIFIED SYNTAX

*Lesion*
*causedBy Neoplasia*
*actsOn Adenocyte*
*partOf Lung*
*hasMalignancy malignant*
*causedBy Smoking*
*hasLocation Brain*
*hasHistology Adenocyte*
*hasBehaviour Mucin Secreting*
*hasCytology Koilocytic*
*hasGrade Well Differentiated*
*hasStatus metastatic*

**Figure 26: *Open*GALEN schema for cancer**

However, an ontology sufficiently expressive to represent tumours according to this more complex schema will often also be able to represent them using more naïve and less expressive schemata. Figure 27 illustrates how two authors, working within the same ontology, might approach the representation of a concept such as 'metastatic pancreatic adenocarcinoma'. Author A's representation is semantically less precise than required (in order that it may be recognised as representing the same concept as Author B) because, although both use almost exactly the same components to create the composition, they have employed schemas of differing sophistication.

**Author A: SIMPLE SCHEMA** ⊗

*Lesion*
*causedBy Neoplasia*
*actsOn Adenocyte*
*hasLocation Pancreas*
*hasMalignancy malignant*
*hasLocation Lymphnode*
*hasStatus metastatic*

**Author B: COMPLEX SCHEMA**

*Lesion*
*causedBy Neoplasia*
*actsOn Adenocyte*
*partOf Pancreas*
*hasMalignancy malignant*
*hasLocation Lymphnode*
*hasStatus metastatic*

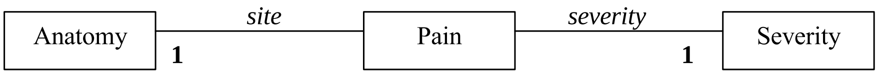
**Figure 27: Complex and simple representations of 'metastatic pancreatic adenocarcinoma'**

#### 3.9.3.4 Insufficient precision due to integrated schemas: Schema for Pain

Section 3.2 has discussed how the notion of pain can be approached by different user communities in different ways.

By contrast perhaps with cancer, probably the majority of clinicians seeking to represent the common notion of pain at a body site would opt for a very simple schema. [Pain] would be described only by its location and severity, resulting in a schema such as illustrated in Figure 28:

**SCHEMA**




**EXAMPLE IN GRAIL SYNTAX**

```
Pain which <
   hasLocation Knee Joint
   hasFeature Severe >
```

**Figure 28: Naive schema for pain** ⊗

However, examining the clinical domain of pain and sensation more generally, concepts that do not fit this simple schema are encountered:

| | |
|---|---|
| phantom limb pain - | *perception of pain occurs with no prior stimulus or sensation* |
| hyperaesthesia - | *stimulus and sensory processes are believed to be intact but perception is up-regulated* |
| synaesthesia - | *stimulus of one modality is perceived as another* |
| analgesia – | *acting by nerve blockade or (typically opiates) by modifying perception.* |

**Figure 29: Unusual pain concepts not possible to represent using naïve schema**

Physiological understanding of the mechanism of pain, therefore, makes a clear distinction between applying a painful stimulus, sensing the stimulus, transmitting the signal that a stimulus has been sensed, and finally interpreting or perceiving the signal received. Problems at each stage give rise to distinct subtypes of pain disorder, whilst pharmacological intervention at each stage gives rise to different kinds of analgesia or anaesthesia.

This more detailed examination of the phenomenon of pain in the clinical context reveals that it can be variously further described by the cause, any sites pain may radiate to, whether or not there is a real physical stimulus, and the extent to which that stimulus (if any) is perceived. The mechanisms of action of many analgesic drugs, for example, make reference to these concepts.

A coherent representation of pain covering all viewpoints, therefore, could combine all these disparate views of pain into a single integrated model. Pain itself would be decomposed into its constituent concepts.

Explicit mention is made within this analysis of the nerve impulse or ‘pain signal’ itself as a further significant concept. However, this ‘pain signal’ might itself be modelled in terms of other concepts already in the ontology for other reasons, such as nerve axons, nerve conduction or cell membrane depolarisation.

The cumulative effect of extending or merging naïve schemata such as Figure 28 illustrates, until they provide a single coherent representation of specialist perspectives of pain, is often a considerably more baroque schema as illustrated in Figure 30:

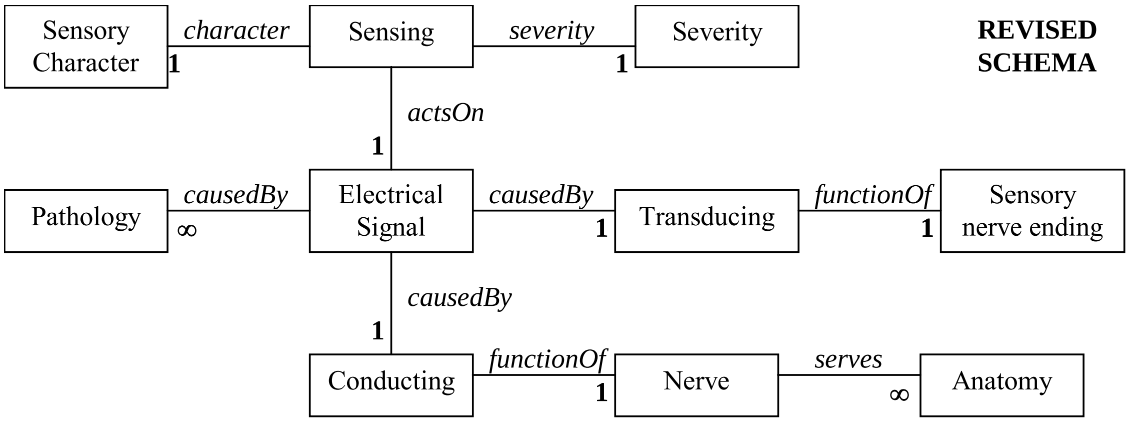



**Figure 30: OpenGALEN schema for pain**

As in the preceding cancer example, an expressive ontology will often concurrently support use of both simple and complex schemas. However, by contrast with the cancer example, specialists and generalists seeking to represent a common concept such as ‘mild ischaemic cardiac pain’ may arrive at solutions that differ not only in schema but also in the semantic components with which they are constructed (Figure 31).

**Author A: SIMPLE SCHEMA** ⊗

*Pain*
    *hasLocation Heart*
    *causedBy Ischaemia*
    *hasSeverity Mild*

**Author B: COMPLEX SCHEMA**

*Sensing*
    *hasCharacter Painful*
    *hasSeverity Mild*
    *actsOn ElectricalSignal*
        *causedBy Transducing*
            *functionOf PainReceptor*
        *causedBy Ischaemia*
        *causedBy Conducting*
            *isFunctionOf VagusNerve*
                *serves Heart*

**Figure 31: Complex and simple representations of 'mild ischaemic cardiac pain'**

### 3.9.4 Cognitive Complexity: combinations of schemas

Pain and cancer have previously been presented as examples of familiar clinical concepts requiring ontological schemas more complex than intuition might expect, in order to accommodate multiple views. However, it should be noted that at least part of their surface complexity derives from the requirement that they should be consistent with other higher level schemas (e.g. feature-state, process-structure previously described in section 3.7) and which pervade the whole model.

This point is demonstrated in the cancer example revisited below in Figure 32, which is presented here in raw GRAIL syntax. Elements required by the feature-state schema are in bold, those required by the process-structure schema are italicised and on a grey highlight.

```
BodyStructure which
    hasUniqueAssociatedProcess (NeoplasticProcess which <
        actsSpecificallyOn (Adenocyte which isSpecificStructuralComponentOf Lung)
        hasMalignancy (Malignancy which hasAbsoluteState malignant)
        hasMetastaticStatus metastatic>)
    hasUniqueAssociatedProcess MetastasesProcess
    hasLocation lymphnode
    hasComponent (Adenocyte which
        hasSpecificFunction (Secreting which actsSpecificallyOn Mucin)
        hasCellMorphology (CellMorphology which <
            isConsequenceOf (CellMorphologyChangeProcess which
                    hasCompleteness (Completeness which hasQuantity
                        (Level which hasMagnitude midLevel)))
            hasAbsoluteState koilocytic>)
    isConsequenceOf  (Inspiration which actsSpecificallyOn Smoke)
```

**Figure 32: Representation of cancer with highlighting of elements arising from different schema**

## 3.10 Cognitive Complexity: Requirement for ontological consistency

The preceding sections have documented specific situations where users of a formal ontology may mistakenly provide too much ontological detail (language artefacts in inserting *vs* removing; ), too little detail (precise flavours of part-whole relation), or the wrong detail (half full, rather than half empty).

This section discusses the particular importance of applying ontological choices and schemas consistently, particularly in the context of automating the detection of semantically equivalent expressions.

### 3.10.1 Consistency and canonisation

An important functionality of a computed compositional terminology is to detect when two different expressions mean the same thing [Brown 1999, Chute 2000, Spackman 2001]. One technique for achieving this functionality is to transform any expression to a canonical form: if two different expressions transform to the same canonical form, they are semantically equivalent.

Consider, for example, the concept of ‘colitis’. This can be constructed within the OpenGALEN Common Reference Model from the primitive concepts for the [Colon] and [InflammationLesion]:

(InflammationLesion which *hasLocation* Colon) name Colitis.

A characteristic of GRAIL, and of most description logics, is that both primitive concepts (undefined concepts such as [InflammationLesion] and [Colon]) and composed concepts (those defined in terms of other concepts in the system, such as [Colitis]) can take equal part in the formation of any further new compositions.

The existence of the newly named composed concept [Colitis] in the OpenGALEN ontology therefore now allows for two apparently different mechanisms by which the concept of ‘radiation colitis’ might be composed:

(Colitis which *isConsequenceOf* Radiation) name RadiationColitis. (1)

(InflammationLesion which <
*isConsequenceOf* Radiation
*hasLocation* Colon>) name RadiationColitis. (2)

However, because the original definition of [Colitis] remains available to the system, the canonisation process first expands (1) to a fully atomised form in which it is re-expressed entirely in terms of primitives. This re-write results in (1) being transformed into (2). In this case, therefore, simple atomisation of two apparently different expressions results in them being identified as semantically equivalent.

Complete atomisation of a composition is, however, not in itself always sufficient to recognise semantic equivalence between compositions. More complex expressions may also require the fully atomised re-write to be reduced further, for example by removing redundant information.

To illustrate this problem, consider first the CRM schema governing how laterality (left or right) is assigned. The English phrase ‘left knee’ is represented as:

Knee which hasLeftRightSelector leftSelection.

However, if the ontology allowed laterality to be assigned to processes as well as structures, then the phrase ‘left knee arthrotomy’ could be expressed in two different, fully atomised and fully canonised forms, one attaching laterality to the anatomy, the other to the act:

(Incising which *actsSpecificallyOn*
(Knee which *hasLeftRightSelector* leftSelection)).

(Incising which <
*actsSpecificallyOn* Knee
*hasLeftRightSelector* leftSelection>).

The GRAIL formalism could not identify these expressions as semantically equivalent. For this reason, the CRM schema prohibits laterality as a modifier of processes; only physical structures can have laterality.

Consider now the phrase ‘left sided arthrotomy’. Superficially this phrase contains no explicit mention of a piece of anatomy to which the laterality could be assigned. To comply with the laterality schema, however, the model author should infer the existence of an unspecified joint to which the laterality is assigned, such that the correct fully atomised representation of this phrase would be:

(Incising which *actsSpecificallyOn*
(Joint which *hasLeftRightSelector* leftSelection)). (3)

However, consider now what may happen if the ontology already also contains a composed concept [Arthrotomy], which has previously been defined as:

(Incising which *actsSpecificallyOn* Joint) name Arthrotomy. (4)

A different model author trying to represent ‘left sided arthrotomy’ may be tempted to use the defined concept [Athrotomy] within the expression and to write:

(Arthrotomy which *hasLeftRightSelector* leftSelection).

..but this breaks the schema for laterality, because the atomised rewrite that results from substituting the atomic definition of [Arthrotomy] in (4) would be:

(Incising which <
*actsSpecificallyOn* Joint
*hasLeftRightSelector* leftSelection>).

…within which expression the laterality is assigned to the [Incising] process, not the anatomy. If the concept [Arthrotomy] is to be used at all, and if the canonisation process is to work, the second model author must write:

(Arthrotomy which *actsSpecificallyOn*
(Joint which *hasLeftRightSelector* leftSelection)).

But this expression now contains a duplication of the information that a joint is involved, as evidenced by the fully atomised re-write that results from substituting [Arthrotomy] with its original definition from (4):

((Incising which < (5)
*actsSpecificallyOn* Joint)
*actsSpecificallyOn* (Joint which *hasLeftRightSelector* leftSelection)>).

However, the full canonisation process detects and removes such redundancy. A simplified illustration of the process, whereby (5) is transformed into (3) and thus their semantic equivalence is detected, is demonstrated in outline below:

(Incising which < (5a – re-write )
*actsSpecificallyOn* Joint
*actsSpecificallyOn* (Joint which *hasLeftRightSelector* leftSelection)>).

(Incising which < (5b – compare criteria)
*actsSpecificallyOn* Joint
*actsSpecificallyOn* (Joint which *hasLeftRightSelector* leftSelection)>).

(Incising which < (5c – spot redundancy)
*actsSpecificallyOn* Joint
*actsSpecificallyOn* (Joint which *hasLeftRightSelector* leftSelection)>).

(Incising which < (5d -remove redundancy)
*actsSpecificallyOn* (Joint which *hasLeftRightSelector* leftSelection)>).

(Incising which *actsSpecificallyOn* (5e – rewrite as (3) )
(Joint which *hasLeftRightSelector* leftSelection)).

The full details of the mechanisms to reduce expressions to a canonical form and detect semantic equivalence are outside the scope of this thesis. The example above is presented to demonstrate the important point that the consistent behaviour of such functionality comes with conditions: within the CRM ontology, pre-formed compositions (such as [Arthrotomy]) may only be used within the definition of larger compositions if proper account is consistently taken of the various ontological schemata in operation, such as those already described. The author of any such expression must therefore:

1. recognise those pre-formed concepts that are composed, not primitive
2. have access to their definition
3. understand which parts of their definition must be redundantly re-expressed

Failure to do this may result in a composition whose canonical form does not comply with the various ontological schema in force, and which can not therefore be recognised as semantically equivalent to compositions written by other model authors.

This requirement on model authors is a significant cognitive burden on all but the most expert users of an ontology, and this is therefore a significant barrier to the direct involvement of domain experts in ontology development and maintenance. Section 9.4 of this thesis describes an approach that significantly reduces this burden of understanding.

## *3.11 Cognitive Complexity: Schema Constraints & Metamodels*

As an ontology grows larger, and more detailed, so the number of occasions where ontological choices have to be made also increases. Keeping track of the consequent “ontological commitments” is taxing enough for the authors of the ontology. More problematic, however, is that inexpert users become increasingly unlikely to realise where important ontological choices exist. Frequently one of many alternative semantic patterns will occur naturally and immediately to them, and they may never be aware that other patterns could be considered, or that another pattern is in fact the preferred schema.

Some form of user guidance is therefore desirable as a means to prevent unwitting use of non-preferred schemas. One form of guidance could be implemented as *pre hoc* restriction of the semantic choices open to users: for example, users might only be offered the choice to describe drinking glasses along an **emptiness** scale. Another form of guidance is *post hoc* validation of choices made: users that try to describe drinking glasses along a **fullness** scale are advised that this is not acceptable.

The GRAIL formalism uses a single mechanism (sanctioning) to provide both forms of guidance. A description of GRAIL sanctioning is presented here partly to illustrate the difficulties of constructing such guidance systems. However, the discussion is also relevant to section 9.3.2 of this thesis, where a new use for the knowledge expressed in the CRM as GRAIL sanctions is proposed: as a fundamental resource within algorithms to hide some of an ontology’s complexity from its end-users.

### 3.11.1 Schema Constraints: GRAIL Sanctioning and Metamodels

In GRAIL, all compositions between any two entities using any semantic link are prohibited unless explicitly permitted. The grammatical and sensible operators act to provide that permission and for this reason are more usually known as GRAIL ‘sanctions’.

Briefly, if a GRAIL author attempts to construct a composition of the form:

A which *semanticLink* B

..but no sanction yet exists in the model such that

A’ sensibly *semanticLink’* B’ *(A’ subsumes A, semanticLink’ subsumes semanticLink, B’ subsumes B)*

..then the GRAIL engine will reject the candidate composition as being unsanctioned.

Declarations of sensible sanctions, however, are themselves contingent on a prior grammatical sanction supporting the same semantic association. Thus, attempts to assert:

A’ sensibly *semanticLink’* B’

..will also be rejected by a GRAIL engine without the prior declaration of a grammatical sanction of the form:

A” grammatically *semanticLink*” B”  (A” subsumes A’ etc)

In declaring sanctions, the usual practice is for grammatical sanctions to be asserted in terms of relatively abstract categories and links, and sensible sanctions between more concrete categories.

Thus:

Structure grammatically *isPartOf* Structure.

.. supports many subsequent sensible anatomical modelling sanctions such as:

Finger sensibly *isPartOf* Hand.
Lobe sensibly *isPartOf* Lung.
Tail sensibly *isPartOf* Pancreas

..whilst at the same time an attempt to add nonsense sensible sanctions such as:

Bone sensibly *isPartOf* Urine.

..would be rejected, because no grammatical sanction is in place to support a sensible sanction using *isPartOf* between a structure (bone) and a substance (urine)

Because grammatical sanctions are placed at a more abstract level in the ontology, and because they act to permit further sensible sanctions to be declared at a more specific level, the GRAIL constraint model is usually described as having two levels of sanction.

Newly constructed compositions, on presentation to the reasoner, are therefore first checked to determine whether they are sanctioned and, if they are not, they are rejected. The author of a new composition can chose between two levels of sanction checking: compositions expressed using the which operator effectively require that sanctioning exists at both the grammatical and more semantically precise sensible levels to permit it; those expressed using the whichG operator require only the (lesser) grammatical level of sanctioning.

From the above exposition it can be seen that the set of sensible sanctions in force in a GRAIL ontology acts to specify and limit how that ontology may be extended. In that regard they can be viewed as defining a partial metamodel of the ontology. Similarly, because the set of grammatical sanctions acts to specify how the set of sensible sanctions can itself be extended, they in turn define a metamodel of the metamodel, or a meta-metamodel of the ontology itself.

### 3.11.2 Cognitive complexity: interpreting rejection

Within such a constrained environment, where all candidate compositions are checked for whether or not they are permitted by a metamodel, authors face a challenge when a proposed composition is rejected. Two possible explanations for any such rejection present themselves:

If the author is certain that the candidate expression is well-formed, semantically complete and consistent with respect to the rest of the model, then the implication is that the existing metamodel constraints are wrong. In order to proceed, the author must determine how to edit the metamodel (ie alter the GRAIL sanctions).

The alternative explanation is that the expression is semantically wrong in some respect and itself needs re-formulating.

Deciding which explanation is the correct one requires expertise and familiarity with the underlying metamodel and schema. A characteristic of a constraint model that only checks compositions *post hoc* after they have been presented is that, while they may inform the author that there is a semantic problem, they are less able to advise how to correct it.

# 4 Analysis: Effects of complexity on the user

Chapter 3 described the different sources of complexity that are taken into account when an ontology is built. This chapter discusses how the combined effect of these complexities is an object that can be particularly hard to understand for users.

Following an analysis of themes identified by informal observational study of users of the *Open*GALEN Common Reference Model, and of their common mistakes or questions, a categorisation of their principle difficulties is proposed below under three headings, described further in this chapter:

- Unreadable syntax
- Navigation
- Authoring

## 4.1 Unreadable syntax

As described in section 3.6, expressions written in the complete GRAIL syntax are often difficult to read, even for experts or by the original author of an expression.

Unsurprisingly, users of the final ontology wishing to inspect formal concept definitions struggle to make sense of such expressions; the surface syntax alone can be sufficient to obscure the semantics. Poor readability is a significant barrier to cooperative working (where a model author wishes to read and comment on another's work) and to re-use, where a second author may wish to adapt another's work.

## 4.2 Navigation

### 4.2.1 Polyhierarchy disorientation

The *Open*GALEN Common Reference Model provides for a complex organisation of its concepts into a multiaxial hierarchy, as described in section 3.5. However, presenting terms to users as such a polyhierarchy appears to be an additional barrier to user acceptance of the ontology [Bentley 1999], especially if most intended users are already familiar with, and therefore expect to see, the traditional monohierarchical view of medicine that is embodied in mainstream medical classification systems.

In a formal ontology such as the CRM, the axes of classification presented as choices to the user range from those familiar to many, through those only familiar to highly specialised users, to those understood by and relevant only to the original ontology authors and of no interest to any user (such as topography, as previously discussed in 3.3). Navigating the concept space by moving up and down the hierarchy in search of a specific concept can feel like trying to find a way through a maze: at each turn there are many choices, but you can not be sure which are blind alleys and which will lead you to your goal. Additionally, because the knowledge names given to concepts in the hierarchy can be ambiguous (see 3.5), the user's interpretation of a navigational choice can be entirely different from its real meaning; not only is the concept space a maze, but the signs pointing the way are misleading.

The abridged hierarchy of cardiovascular anatomy from the *Open*GALEN CRM (Figure 34) demonstrates this point: a user seeking to traverse the classification from abstract chapter or section headings to more detailed concepts, in search of a specific term, is at all stages presented with multiple navigational choices, some of which can be difficult to comprehend.

For example, the highlighted concept [CVSComponentOfThorax] might have been expected to subsume the [Heart] but does not: this is because the concept [Thorax] here relates to the thoracic cage only, and not to the thorax as a major division of the body (ie not to the thoracic cage and its contents). Consequently, it subsumes only the intercostal vasculature, including the concept [IntercostalArtery] in bold.

Looking under the alternative broader heading of [ArterialStructure] for these same intercostal arteries, the hierarchy broadens out into variety of different abstract groupings of arteries; the concept [IntercostalArtery] can in fact be reached by three other routes (Figure 33) descending from the original [CardiovascularSystemComponent] concept, all passing through [Artery]:

- ArterialStructure
  - AfferentTubularPartOfCardiovascularSystem
    - DivisionOfArtery
      - Artery
        - ArteryWhichHasLaterality
          - MirrorImagedArtery
            - **IntercostalArtery**
        - DeepArtery
          - **IntercostalArtery**
        - PeripheralArtery
          - **IntercostalArtery**

**Figure 33: Other navigational routes to [IntercostalArtery]**

The fact that a code can be reached by more than one navigational path in the *Open*GALEN CRM might be imagined to be a good thing, making it more rather than less likely that the concept will be found. However, although four navigational routes lead down from [CardiovascularSystemComponent] to [IntercostalArtery], this must be set against the fact that the unabridged, and fully expanded, version of the hierarchy shown in Figure 34 occupies more than 2500 lines even though only 1112 unique concepts are present: most of them can be reached by more than one route. Successfully navigating through the choices depends on understanding and correctly interpreting the sometimes obscure concepts and axes of classification encountered along the way.

```
  + BodyStructure
- BodySystemAnatomy
  - CardiovascularSystemComponent
    + AbnormalVascularStructure
    - ArterialStructure
      + AbnormalArterialStructure
      - AfferentTubularPartOfCardiovascularSystem
        + DivisionOfArtery
        + Arteriole
        - Artery
          + AnatomicalEndArtery
          + ArterialAutograft
          + ArterialBranch
          + ArterialGraft
          + ArteryOfHeadOrNeck
          + ArteryOfNasalPassage
          - ArteryWhichHasLaterality
            - MirrorImagedArtery
              + ArteryOfExtremity
                AnteriorDivisionOfInternalIliacArtery
                InfraOrbitalArtery
              + IntercostalArtery
                InternalMammaryArtery
       ……..snip……
                ZygomaticotemporalArtery
          + ArteryWhichHasNoLaterality
          + CentralArtery
          - DeepArtery
            + BranchOfAorta
            + DeepArteryOfLowerExtremity
       ……..snip……
            + IntercostalArtery^
              InternalCarotidArtery
          + IntraAbdominalSystemicArtery
          + IntraThoracicSystemicArtery
          - PeripheralArtery
            + BranchOfAbdominalAorta
            + BranchOfThoracicAorta
            + IntercostalArtery^
       ……..snip……
              UmbilicalArtery^
      CVSComponentOfAbdomen
    + CVSComponentOfLowerLimb
    - CVSComponentOfThorax
      + IntercostalArtery^
      + IntercostalVein
    + CVSComponentOfUpperLimb
    + IntraAbdominalCardiovascularStructure
    + IntraThoracicCardiovascularStructure
    + TubularCardiovascularComponent
    + VascularStructure
    + VenousStructure
    + (TissueCell < isStructuralComponentOf CarotidBody >)
      AorticPulmonaryBody
      CardiacConduit
      CarotidBody
      GlomusJugulare
    + NAMEDCVSBodyPart
  + ChemoreceptorSystemComponent
……..snip……
```

**Figure 34: Hierarchy of cardiovascular anatomy (abridged)**

### 4.2.2 Concept clutter

The *Open*GALEN Common Reference Model aims to be able to express, in a single polyhierarchical data structure, the union of all those concepts that any user might want for an application and in any medical subdomain, together with the union of all possible ontological perspectives by which any individual concept might be viewed. The issues of *Domain Complexity* described in sections 3.1 to 3.5 make achieving this aim an elusive goal.

Specific individual users, however, will always have highly application and sub-domain specific requirements both in terms of which concepts they are interested in, and from which perspective.

An inevitable consequence of supporting multiple ontological views across a very large domain, therefore, is that a particular user approaching the CRM for a specific purpose will find that the majority of the terms encountered are irrelevant to that purpose, being only relevant to other possible users. As the level of irrelevant ‘conceptual clutter’ increases, it progresses from initially being a curious distraction, through being a source of irritating confusion, until the stage is reached where the terms relevant to one specific user are completely obscured by the terms required for all other uses. This occurs at all levels of any hierarchical classification of the concepts in the domain: additional leaf concepts are included for other specialists, and additional abstractions higher up the hierarchy are included to provide multiple ontological views both for those same specialists and to enable links to external applications, corpora or rule bases as described in section 3.3.

For example, the ability to classify anatomical structures according to their detailed topological characteristics is included in the *Open*GALEN CRM solely for the purpose of building and maintaining the ontology in the first place, but is unlikely to be of interest or value to an end-user clinician. The subsumption hierarchy (Figure 35) shows the concept [Heart] together with 8 of its direct parents, two of which (in bold) owe their presence to this topological modelling.

    InternalOrgan
    **TubularCardiovascularComponent**
    IntraThoracicCardiovascularStructure
    MuscleTissueStructure
    IntegumentaryStructure
    FibrousBodyStructure
    IntraMediastinalStructure
    **TubularBodyPart**
Heart

**Figure 35: Parents of [Heart]**

The combination of uncertainty in making choices navigating the polyhierarchy and the obscuring effect of conceptual clutter means that users who fail to find a concept are unable to conclude that the concept is genuinely not present in the concept system. They fear that they may have been looking in the wrong place, or under the wrong name, or simply did not notice the concept in the lists of concepts presented.

## 4.3 Authoring

### 4.3.1 Pressure for parsimony – new primitive or composition ?

On those occasions that users were certain a concept was *not* already present in the model, they found themselves unsure regarding how to add it safely without introducing semantic redundancy. Should they be trying to assemble a new composition from smaller more basic terms that *did* already exist in the CRM, or should a new primitive concept be added and, if so, where in the overall concept hierarchy ?

Because of these concerns, success in identifying a concept not already in the model may simply result in a new obligation to search for more concepts: if the required (missing) concept might properly be created as a composed concept in terms of other more basic concepts already present in the model, then the user must now establish whether all of those basic concepts exist or not, before they can assemble the originally desired concept from its parts.

### 4.3.2 Schema uncertainty – no conceptual cookbook

Even before consideration of whether the component concepts for a potential new composition can be found, users complain that it is difficult for them to know the correct ontological schema, or "recipe", by which they should be combined to construct the new composition. The problems of *Cognitive Complexity* detailed in sections 3.9, 3.10 and 3.11 and Artefactual Complexity detailed in sections 3.7 and 3.8 come into play.

On occasions when they were aware of the correct schema, however, users often perceived that it was arbitrarily complicated, over-engineered, and difficult to follow.

# Part Two: An Intermediate Representation

Part One of this thesis has described multiple sources of complexity that can make an ontology inaccessible for even expert users.

This thesis is **not** concerned with whether ontologies must be complex, or whether useful or usable ontologies might be constructed by simpler approaches.

Given that complex ontologies exist,
Part Two sets out a methodology for coping with them, and describes the workflows and tools developed to implement that methodology

# 5 Easing the user’s task - an Intermediate Representation

The *Open*GALEN Common Reference Model of medicine is a complex ontology that has become increasingly opaque and inaccessible to its intended users. Part Two of this thesis describes a methodology to solve this problem: instead of interacting directly with the ontology, users instead access it indirectly via an ‘intermediate representation’ that is simpler both syntactically and ontologically. Expressions in the intermediate representation may then be automatically re-written to the final representation.

Part Two of this thesis may be considered in three sections:

*Section One*: chapter 5 (this chapter) provides an overview of the approach

*Section Two*: chapters 6 through 8 describe how a complex ontology such as the *Open*GALEN Common Reference Model is simplified to derive an intermediate representation.

*Section Three*: chapter 9 describes how expressions in the intermediate representation are transformed back into the complex ontology

## 5.1 Illustrative example of intermediate and final representation

To demonstrate the overall goal, Figure 36 provides an illustration of the relative simplicity of the GALEN intermediate representation (IR) when compared with the native OpenGALEN ontology representation into which it is transformed. Both expressions shown seek to represent the same underlying concept: an open surgical procedure to remove atheroma from a coronary artery.

The essential characteristics of the intermediate representation as compared with the target representation are evident from this illustration:

- **simpler syntax**: the IR has no brackets and suppresses keywords (e.g. extrinisically, which, whichG)
- **simpler ontology**: the IR version employs just 3 concepts and 2 links, compared with 13 concepts and 13 links for the same concept in GRAIL
- **simpler semantics**: the GRAIL representation mentions the concept [heart] twice and also includes an additional conceptual graph fragment (SurgicalDeed which....isEnactmentOf) that precedes what appeared to be the start of the expression (Dividing) in the intermediate representation

| INTERMEDIATE REPRESENTATION | GRAIL REPRESENTATION |
|---|---|
| RUBRIC "Open coronary artery endarterectomy" | (SurgicalDeed which isMainlyCharacterisedBy |
| SCHEME “CTV3” CODE "X00tN" | (performance whichG isEnactmentOf |
| MAIN removing | ((Removing which playsClinicalRole SurgicalRole) whichG < |
| HAS_APPROACH open approach | hasSpecificSubprocess |
| ACTS_ON atheroma | (SurgicalApproaching whichG hasSurgicalOpenClosedness |
| HAS_ LOCATION coronary artery | (SurgicalOpenClosedness whichG hasAbsoluteState surgicallyOpen)) |
| | actsSpecificallyOn |
| | (Atheroma which hasSpecificLocation |
| | (TunicaIntima whichG isSpecificLayerOf CoronaryArtery))>))) |
| | hasProjection (('READ' schemeVersion 'default') code 'X00tN' 'code'); |
| | extrinsically hasDissectionRubric 'Open coronary artery endarterectomy'. |

**Figure 36: Comparison of intermediate and GRAIL representations**

Within the GALEN-IN-USE project, the goal was to author and quality assure some 20,000 representations (dissections) of similar, or greater, complexity. The right hand example in Figure 36 illustrates why performing this authoring directly in GRAIL, using unskilled authors, proved impossible.

## 5.2 Need for automation in transforming from Intermediate to Final

For a representation to be *intermediate*, information in that representation needs to be transformed into a final *target* representation. Such a transformation might – at least in principle - be performed entirely manually, semi-automatically or fully automatically.

The GALEN Intermediate Representation was intended for a very large knowledge acquisition effort. In this context an entirely manual transformation of each representation to the final GRAIL form could not have been resourced, nor might one reasonably expect that it would in fact be applied systematically or without error.

For these reasons the transformation process needed to be automated as much as possible.

## 5.3 Design Approach: Reversible Systematic Simplification

The intermediate representation was engineered as a *systematic* simplification of the intended target GRAIL representation and CRM ontology: all design decisions concerning simplification of the syntax of the intermediate representation, the presentation and content of its ontology, or the range of ontological schema permitted, were devised to facillitate or enable the specification and implementation of algorithms to reverse that simplification, resulting in the capability to perform a fully automatic transformation from IR back to the target ontology.

A significant advantage of the approach is that expressions phrased in the intermediate representation, though less rich than the final target representation, still carry significant semantic information. This can be used to make the transformation from intermediate to final representations sensitive to the semantic context. The particular ontological commitments of the target ontology can be enforced by trapping variant schemas and transparently re-writing them.

An important caveat of this approach is that it relies on the consistent use of any ontological commitments by the original authors of the target ontology. The more a target ontology is inconsistent in part or in whole with regard to the application of schema during its construction, the harder it becomes to systematically simplify it.

The detailed methodology by which the intermediate representation was engineered as a systematic simplification of the final target representation and ontology is described in subsequent chapters. Figure 37 provides an initial schematic overview. Smaller copies of Figure 37 introduce each of the following chapters 6 through 9, with highlighting (white text on black background) to emphasise the particular section under discussion.

The central column in Figure 37 identifies three broad techniques used to address the user difficulties described in chapter 4. The techniques are summarised here as:

- *Syntactic Compression (Chapter 6)* – representations are authored using a simplified syntax, omitting brackets and keywords and providing simpler access to any workarounds
- *Ontological Compression (Chapter 7)* – the full detail of the target ontology is hidden from the user; only a subset of concepts and semantic links identified as more relevant to the user task is presented as an ‘authoring ontology’, which is presented using a simpler navigational hierarchy

- *Semantic Compression (Chapter 8)* – the requirement to comply precisely with all the semantic schema of the target ontology is relaxed, as is the requirement for maximal parsimony and non-redundancy
- *Expansion (Chapter 9)* – the three forms of compression outlined are matched by a corresponding expansion and re-write process.

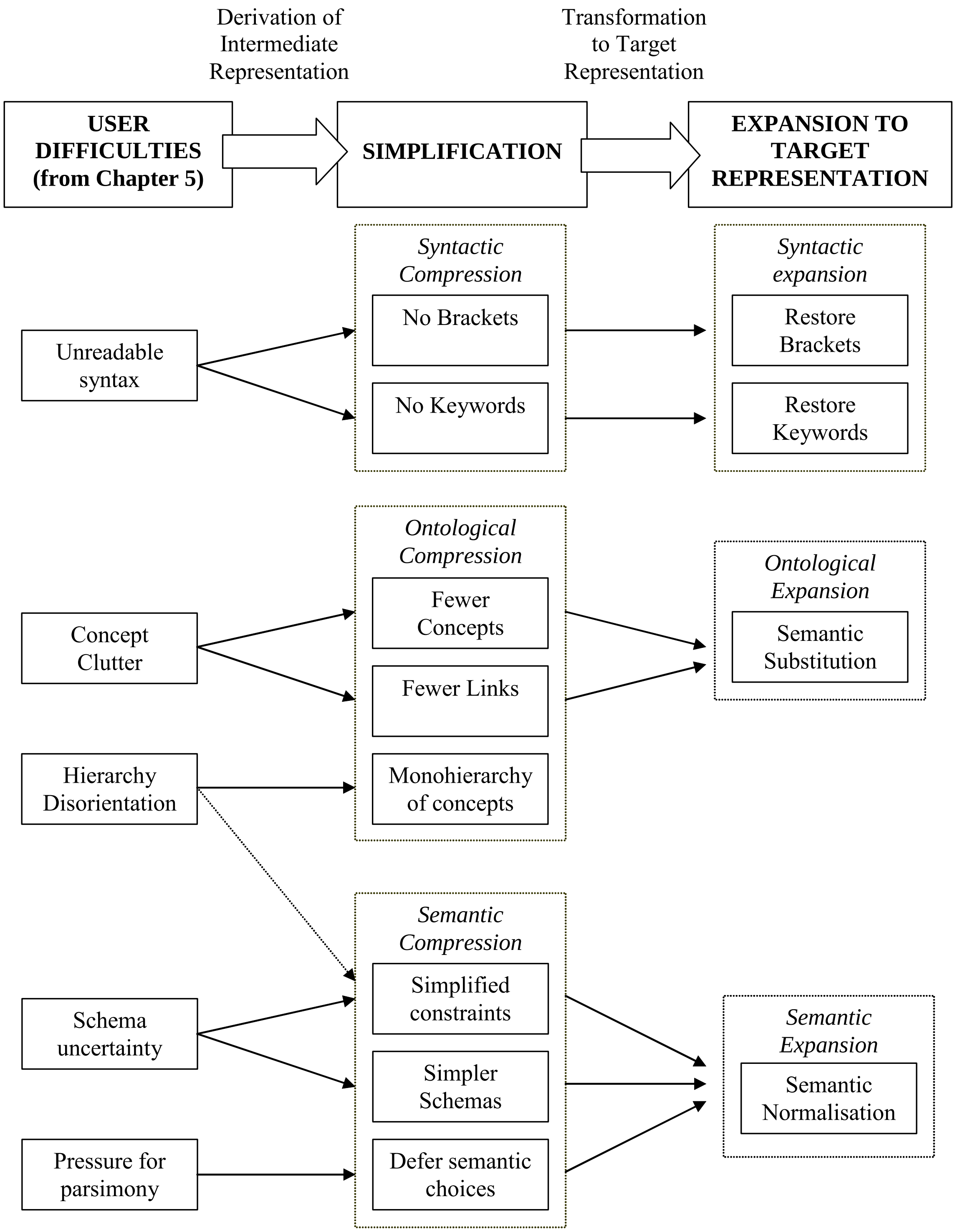
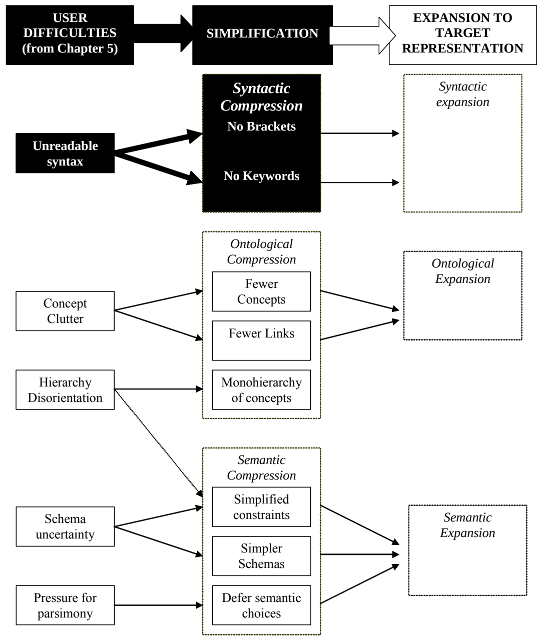



**Figure 37: Schematic diagram of methodology**

# 6 Simplification Techniques: Syntactic Compression

## *6.1 Introduction*




An important first goal of the intermediate representation was to address the problem of *unreadable syntax*: the initial surface readability of GRAIL expressions is poor. Expressions written by one author can not easily be comprehended by another, and this is a significant barrier to collaborative working and quality assurance.

Two techniques were used to achieve a simpler surface appearance:

- substituting visible tokens (brackets) with invisible ones (tabbed indents)
- restricting the expressivity of the IR to only two GRAIL keywords and then suppressing display of even these

These approaches are discussed in more detail in the following sections. The complex representation of *tenesmus* originally presented in section 3.6 (reproduced below as Figure 38) is used as a worked example throughout this chapter to demonstrate their effect. The starting point of syntactic compression is raw GRAIL syntax:

(ClinicalSituation which <isCharacterisedBy (presence which isExistenceOf (ContractionProcess which <isSpecificFunctionOf SphincterAniMuscle hasImmediateConsequence Pain hasIntentionality (Intentionality which hasAbsoluteState involuntary) hasDuration (Duration which hasAbsoluteState longTerm) hasTemporalPattern (TemporalPattern which hasAbsoluteState ongoing) >)) isCharacterisedBy (presence which isExistenceOf (UrgeToVoidUrineOrFaeces which hasProcessActivity (ProcessActivity which hasQuantity (Level which hasMagnitude highLevel)))) isCharacterisedBy (presence which isExistenceOf AbdominalStraining) >)

**Figure 38: Worked example in serialised GRAIL notation**

Chapters 6 through 8 follow the expression in Figure 38 as a worked example, showing how it is progressively simplified to that shown in Figure 39:

```
MAIN contraction process
    IS_SPECIFIC_FUNCTION_OF sphincter ani muscle
    HAS_IMMEDIATE_CONSEQUENCE pain
    HAS_FEATURE involuntary
    HAS_FEATURE long term
    HAS_FEATURE ongoing
WITH urge to void urine or faeces
    HAS_FEATURE high activity level
WITH abdominal straining
```

**Figure 39: Worked example in final intermediate representation**

## 6.2 Syntactic Compression : Elimination of bracketing

The standard GRAIL notation includes the use of both rounded and angle brackets to determine the precise semantics of the expression. This is demonstrated in Figure 40, where bracketing determines whether [red] describes the finger or the hand:

| GRAIL SYNTAX | MEANING |
|---|---|
| Finger which isPartOf (Hand which hasColour Red) | Finger of red hand |
| Finger which <isPartOf Hand hasColour Red> | Red finger of hand |

**Figure 40: Role of brackets in GRAIL**

The presence of brackets in the GRAIL notation becomes an increasing burden as the statements being represented become more complex. Not only must authors ensure that brackets are balanced (in the absence of tools supporting this task), but comprehending the resulting expressions becomes increasingly difficult.

The surface unreadability of complex expressions such as Figure 38 had, in fact, already been addressed in the native GRAIL authoring and browsing environment through the introduction of 'pretty printing'. The bracketing of the original serialised form is retained, but tabbed indentation serves as an additional but more visually convenient signal of the same information (Figure 41).

Although pretty printing results in a *more* readable form for human readers and authors, the notation is still not so easily readable that the nature of the concept represented in Figure 41 is always quickly grasped. Additionally, as already stated, the pretty printed version simultaneously and redundantly employs two different methods to display the graph structure: bracketing and tabbed indenting.

To simplify this state of affairs and remove the visually redundant information, the intermediate representation notation does away with the brackets completely at both authoring and review time, using only tabbed indenting. With this simplification, the expression from Figure 38 above can be re-written as Figure 42.

```
(ClinicalSituation which <
    isCharacterisedBy
        (presence which isExistenceOf
            (ContractionProcess which <
                    isSpecificFunctionOf SphincterAniMuscle
                    hasImmediateConsequence Pain
                    hasIntentionality
                            (Intentionality which hasAbsoluteState involuntary)
                    hasDuration
                            (Duration which hasAbsoluteState longTerm)
                    hasTemporalPattern
                            (TemporalPattern which hasAbsoluteState ongoing)>))
    isCharacterisedBy
        (presence which isExistenceOf
            (UrgeToVoidUrineOrFaeces which hasProcessActivity
                    (ProcessActivity which hasQuantity
                            (Level which hasMagnitude highLevel))))
    isCharacterisedBy
        (presence which isExistenceOf AbdominalStraining)>)
```

**Figure 41: Pretty Printed version of worked example**

ClinicalSituation which
  *isCharacterisedBy* presence which
    *isExistenceOf* ContractionProcess which
      *isSpecificFunctionOf* SphincterAniMuscle
      *hasImmediateConsequence* Pain
      *hasIntentionality* Intentionality which
        *hasAbsoluteState* involuntary
      *hasDuration* Duration which
        hasAbsoluteState longTerm
      *hasTemporalPattern* TemporalPattern which
        *hasAbsoluteState* ongoing
  *isCharacterisedBy* presence which
    *isExistenceOf* UrgeToVoidUrineOrFaeces which
      *hasProcessActivity* ProcessActivity which
        *hasQuantity* Level which
          *hasMagnitude* highLevel
  *isCharacterisedBy* presence which
    *isExistenceOf* AbdominalStraining

**Figure 42: Modified pretty printed version of worked example**

## *6.3 Syntactic Compression: Elimination of keywords*

The full GRAIL notational syntax recognises 109 keywords. However, of these, only the 12 operators shown in Figure 43 make up the basic core and are sufficient for almost all knowledge authoring purposes:

| CORE GRAIL OPERATORS | EFFECT |
|---|---|
| newSub | Assert existence of a new primitive category, and its parent |
| newAttribute | Assert a new semantic link type as a child of an existing link |
| specialisedBy | Assert for transitivity of semantic links |
| addSub | Assert child relationship with existing category |
| which | Constructor for compositions |
| whichG | Less constrained constructor for compositions |
| name | Assign name to a composition |
| grammatically | Assert a constraint at 'grammar' level (lowest) |
| sensibly | Assert a constraint at 'sensible' level (highest) |
| necessarily | Assert an axiom bidirectionally |
| topicNecessarily | Assert an axiom unidirectionally |
| extrinsically | Assert non-classificatory fact |

NB The remaining 97 operators include preformed combinations of the core 12 (e.g. sensiblyAndTopicNecessarily), operators for GQL (GRAIL Query Language) and various constructors introduced to assist linking applications concepts to external applications.

**Figure 43: Core GRAIL operators**

Using the 12 core operators, the GRAIL author is able to declare primitive category and semantic link spaces for the user, together with constraints governing how users may combine entities and links to create new compositions. A significant point, however, is that in making such new compositions, the user requires only the operator(s) for making compositions. Most users will make no use of (or should be actively discouraged from using) those other operators that exist either to add new primitive entities or links, or to alter the constraint system controlling new compositions.

Further simplifications of the intermediate representation notation therefore become possible: firstly, the user can be restricted to using only the composition-forming operators. Secondly, the 'pretty printing' can be extended to suppress entirely the display of these keywords; authors can signal the need for a compositional operator by graphical means only: continuing their composition on a new line but with an additional indent level.

In GRAIL, two distinct compositional operators are available: which or whichG. Therefore, a simpler notation that indicates only *where* one of these operators should appear requires a mechanism to determine *which* to implement. This mechanism is described in Section 9.2.2.

It is therefore possible to omit display of keywords altogether from the intermediate representation and, as a result, the worked example *tenesmus* expression used throughout this chapter becomes yet simpler to read (Figure 44):

```
ClinicalSituation
    isCharacterisedBy presence
        isExistenceOf ContractionProcess
            isSpecificFunctionOf SphincterAniMuscle
            hasImmediateConsequence Pain
            hasIntentionality Intentionality
                    hasAbsoluteState involuntary
            hasDuration Duration
                    hasAbsoluteState longTerm
            hasTemporalPattern TemporalPattern
                    hasAbsoluteState ongoing
    isCharacterisedBy presence
        isExistenceOf UrgeToVoidUrineOrFaeces
            hasProcessActivity ProcessActivity
                    hasQuantity Level
                            hasMagnitude highLevel
    isCharacterisedBy presence
        isExistenceOf AbdominalStraining
```

**Figure 44: Worked example with GRAIL operators suppressed**

## 6.4 Syntactic Sugar: more readable typography

A final alteration to the intermediate representation surface appearance is not syntactic but typographical. By convention, knowledge names for concepts and semantic links in the OpenGALEN CRM are written with 'space removed word capitalisation', such that the string 'first second third' becomes 'FirstSecondThird'. The above example, even after removing brackets and keywords, demonstrates that it can still be difficult to distinguish the concepts from the links.

The Intermediate Representation therefore adopts a different typography: concept names appear all in lower case and with spaces persisting, whilst semantic links appear all in upper case and with spaces substituted by underscores. The tenesmus example, after syntactic compression and typographic change, becomes:

```
clinical situation
    IS_CHARACTERISED_BY presence
        IS_EXISTENCE_OF contraction process
            IS_SPECIFIC_FUNCTION_OF sphincter ani muscle
            HAS_IMMEDIATE_CONSEQUENCE pain
            HAS_INTENTIONALITY intentionality
                    HAS_ABSOLUTE_STATE involuntary
            HAS_DURATION duration
                    HAS_ABSOLUTE_STATE long term
            HAS_TEMPORAL_PATTERN temporal pattern
                    HAS_ABSOLUTE_STATE ongoing
    IS_CHARACTERISED_BY presence
        IS_EXISTENCE_OF urge to void urine or faeces
            HAS_PROCESS_ACTIVITY process activity
                    HAS_QUANTITY level
                            HAS_MAGNITUDE high level
    IS_CHARACTERISED_BY presence
        IS_EXISTENCE_OF abdominal straining
```

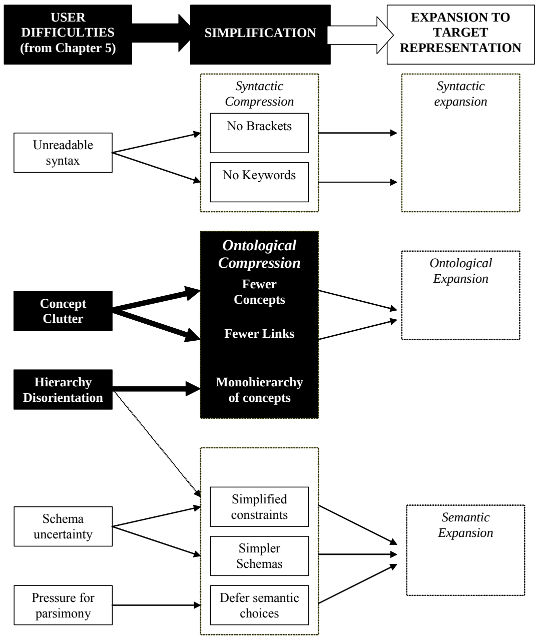
**Figure 45: Worked example following typographic conventions**

## 6.5 Summary

Expressions representing detailed concepts are necessarily complex. When written in raw GRAIL syntax such expressions are difficult to author and even more difficult to read. The intermediate representation increases initial readability of representations by *syntactic compression* that suppresses the display of both bracketing and keywords, and by a different typographic convention.

# 7 Simplification Techniques: Ontological Compression

## *7.1 Introduction*




This chapter describes techniques to make it easier to find concepts and links within the ontology of concepts.

In summary, they are:

- Fewer concepts
- A shallow monohierarchy
- Fewer Links

The problem of *Concept Clutter* (most concepts being irrelevant to specific users) is reduced by ensuring that only those terms likely to be needed by a particular group of users are presented to them, and correspondingly terms not relevant to the task in hand are hidden.

*Hierarchy Disorientation* (multiaxial classification with obscure abstractions) is reduced by presenting the restricted term list in a less complex navigational structure.

A further burden for the inexpert user, in addition to navigating a large multiaxially classified concept space, is navigating around and choosing between the relatively large number (by comparison with other compositional systems) of different semantic link types available in the complete *Open*GALEN model.

In total, the *Open*GALEN model provides more than 400 different flavours of semantic link. Although approximately half are required to work around various limitations of the underlying GRAIL formalism (as described in section 3.6), many of the semantically ‘genuine’ flavours are grouped in closely related families – such as the partonomic links described in 3.9.3.1. The differences between members within a semantic link family can be subtle.

Therefore, in addition to making it easier to find concepts, a further explicit goal was to reduce the need to be familiar with either the precise semantic differences between similar semantic links, or with those created as workarounds.

## *7.2 Ontological Compression: Fewer concepts*

The expectation was that surgical procedure descriptions could be expressed compositionally using a much smaller number of concepts than was present in the full *Open*GALEN ontology, perhaps only a few thousand. It was anticipated, for example, that many anatomical terms would be reused over and again across many descriptions of surgical procedures, although equally it was expected that some anatomical terms (such as those for the different nerve plexuses) might rarely occur. By contrast very few of the large number of terms for different species of microorganism would be likely to be needed.

Therefore, the ontology presented to the authors was tailored to be highly task specific – to contain only those terms actually used by authors given the task they were undertaking. For this reason, the resulting tailored ontology is known as an ‘authoring ontology’.

An attempt to anticipate the set of terms needed by authors, before they commenced any work, led to the compilation of a list of some 600 candidate terms comprising anatomy (leg, hand, intestine) and deeds (excising, opening, amputating) [Rossi Mori 1997]. This *pre hoc* list was, however, largely ignored by users partly due to difficulties navigating a completely flat list of even such a modest number of terms, and partly due to lack of ownership. Ontology tailoring therefore became an iterative and *ad hoc* activity: the authors effectively chose to start work with no predefined set of terms and instead created their own as they progressed. Terms ('descriptors') used were collected from individual authors, with a cumulative list redistributed periodically to all authors so that each could benefit from the work of the community of authors. The list of used descriptors stabilised after about 5,000 surgical procedure rubrics had been represented, with the eventual cumulative authoring ontology numbering 4168 terms required to represent 20,782 surgical procedure rubrics..

## 7.3 Ontological Compression: Reducing Hierarchy Disorientation

As described in section 3.5, navigating around the *Open*GALEN ontology is difficult not only because of the large number of concepts in the CRM (more than 20,000 in version 6) but also because they are organised into a polyhierarchy including many abstract categories.

The weaknesses of monohierarchies as a means of organising large numbers of concepts are well described, particularly where the hierarchy is intended to be used both for human navigation at data capture and also subsequently for aggregation or statistical analysis. However, whatever the logical or philosophical merits of polyhierarchies, many users perceive them as unnecessarily confusing and complex. Therefore, since the authoring ontology was expected to be relatively small and also since any categorial structuring was intended exclusively to assist human navigation (and not also statistical grouping), a monohierarchy of the descriptors seemed a viable proposition.

### 7.3.1 A categorial structure

An initial categorisation of the authoring ontology 'descriptors' was informed by reference to CEN ENV 1828, a pre-standard subsequently voted a full standard that proposes a categorial structure and semantic schema for representations of surgical procedures [CEN 1999]. Central to the CEN model is that all procedures must comprise a deed that must 'act on' a direct object but that may also sometimes 'act on' an indirect object. Both direct and indirect objects can be either anatomy, pathology or devices. This outline schema for describing surgical procedures implies a simple categorisation of the concepts employed.

The first draft of a categorial structure for the authoring ontology therefore comprised:

- 'deeds' (e.g. excising, opening, removing, injecting)
- 'anatomy' (e.g. leg, heart valve, tongue, liver)
- 'pathology' (e.g tumour, cyst)
- 'features' (e.g. partial, total, radical, elective)

However, as the authoring ontology (especially the list of anatomical parts) grew larger and as the ENV 1828 inspired schema for surgical procedures itself had to be extended, additional categories and subcategories were created. The final categorial structure comprised 115 distinct descriptor categories (see Appendix One) organised into a monohierarchy with a maximum of 6 levels, under eight top level categories: Characteristic, Quantity, RouteOrApproach, SignOrSymptom, Pathology, Process, Structure or Substance.

### 7.3.2 Applying the categorial structure

The final authoring ontology for surgical procedures contained approximately 4000 different descriptors, categorised into one of the 115 categories detailed in Appendix One. Whilst it might have been feasible to assign and maintain this categorisation entirely manually, such an endeavour is time consuming and error prone.

Section 9.3.1 will describe in detail how each descriptor in the authoring ontology has a declared mapping to a corresponding concept in the target ontology. The primary purpose of these mappings is to enable intermediate representation expressions to be re-expressed using semantically equivalent terms from the target ontology. However, they may also be used to automate descriptor categorisation.

The membership of each of the 115 descriptor categories can be defined in terms of concepts and hierarchies in the target ontology. All descriptors with a mapping can then be assigned to a descriptor category on the basis of that definition. Typical descriptor category definitions are of the form:

> *For all descriptors (D) possessing a mapping to a concept (C) in the target ontology and where (C) is subsumed by any member of a set of concepts [X,Y,Z] in the target ontology, then (D) is assigned to the descriptor category ($Cat^D$) where $Cat^D$ is mapped to [X,Y,Z].*

For example, because the descriptor category ‘ BloodVesselOfHeadOrNeck’ is defined thus:

> *For all descriptors (D) possessing a mapping to a concept (C) in the OpenGALEN ontology and where (C) is subsumed by any member of a set of concepts [(BloodVessel whichG IsDivisionOf HeadOrNeck)] in the OpenGALEN ontology, then (D) is assigned to the descriptor category (BloodVesselOfHeadOrNeck) where (BloodVesselOfHeadOrNeck) is mapped to [(BloodVessel whichG IsDivisionOf HeadOrNeck)] .*

..the descriptor ‘carotid artery’ is assigned to it. This descriptor is mapped to the primitive concept [CarotidArtery] in the *Open*GALEN ontology, which is in turn subsumed by [(BloodVessel whichG IsDivisionOf HeadOrNeck)].

Descriptors that do not yet have a mapping to a *Open*GALEN ontology concept can not be automatically categorised in this way, though they may have a provisional category assigned to them by hand, for example by the author who first created the descriptor.

Some descriptor categories represent the aggregation of several classes from the *Open*GALEN ontology: the descriptor category ‘TemporalMarker’ includes descriptors such as ‘acute’, ‘daily’ and ‘permanent’. Its definition accordingly is:

> *For all descriptors (D) possessing a mapping to a concept (C) in the target ontology and where (C) is subsumed by any member of a set of concepts [UrgencyState, TimePeriod, RevisionStatus] in the target ontology, then (D) is assigned to the descriptor category (TemporalMarker) where (TemporalMarker) is mapped to [UrgencyState, TimePeriod, RevisionStatus] .*

An additional complication arises because the *Open*GALEN ontology is multiaxial whereas the categorial structure is mono-axial. Many descriptor mappings therefore meet the definition of more than one descriptor category. For example, the descriptor ‘ectopic adrenal gland’ maps to a concept in the target ontology that is simultaneously a kind of [Gland], a kind of [EndocrineSystemComponent] and a kind of [PathologicalStructure]. As a result of these three subsumptions it fulfils the definitions of three different descriptor categories: ‘Gland, ‘EndocrineSystemComponent’ and ‘Lesion’. Because the authoring ontology navigational hierarchy is intended to be monoaxial it is not permitted to have the one descriptor appear in all three categories. The 115 categories are, therefore, additionally assigned a ranking so that in the (frequent) event of a descriptor being a candidate for more than one category, the highest ranking descriptor category wins.

In addition to suppressing selected axes of multiple classification, this descriptor categorisation mechanism can also transform them. For example, the *Open*GALEN ontology view of anatomy is as morphology, in which pathological forms of anatomical structures are subsumed by the unmodified structures (e.g. [ThyroidGland] subsumes [EctopicThyroidGland]). The authoring ontology view of anatomy is as normative anatomy, in which one descriptor category holds normal forms of anatomy and a separate category ('Lesion') contains the pathological variants. The automated descriptor categorisation performs this transform.

## 7.4 Ontological Compression: Fewer semantic links

In addition to reducing the number of concepts present in the authoring ontology, an additional goal was to reduce the number of different semantic link types that a user must be familiar with, and the need to know when and how to use each correctly. By contrast to the way authoring ontology descriptors were collated entirely cumulatively and *post hoc* from the ongoing collected work of the authors, the set of links available to authors was successfully primed at the start with an initial set anticipating likely expressive needs. Only a small number of further links were added to this initial set in the light of actual use.

The content of the priming set was informed partly by the requirements of CEN ENV 1828 and partly by detailed knowledge of the target (*Open*GALEN CRM) ontology.

In choosing the authoring ontology links, two different classes of target ontology links that might be systematically simplified were identified:

- those that were artefacts resulting from workarounds
- those from semantically related families

### 7.4.1 Suppression of Artefactual Links

As previously discussed, one significant limitation of the GRAIL formalism is its impoverished model of link cardinality, a consequence of which is the presence of a large number of artefact semantic links in the *Open*GALEN ontology, primarily numerous subclasses of [*hasFeature*] as described in section 3.6.

The intermediate representation, by contrast, provides no syntactic means at all to state cardinality. However, the appropriate cardinality can sometimes be inferred from the semantic context and need not be declared by the user. The mechanism by which this inference occurs is described in 9.3.2.5.

In particular, this mechanism makes it possible to hide the artefact variant cardinality links of the *Open*GALEN target ontology, such as the three subtypes of [*actsOn*] detailed in 3.7.2. The authoring ontology correspondingly contains only a single ACTS_ON link, and its inverse IS_ACTED_ON_BY. By the same mechanism, approximately 34 other variant cardinality links in the *Open*GALEN ontology, and their 34 inverses, are hidden from the user in the authoring ontology.

Section 8.3.2 describes a technique for semantic compression of the Feature-State workaround schema, which allows for a further significant reduction in the number of semantic link types required in the authoring ontology: more than 200 different subtypes of the 'hasFeature' link in the *Open*GALEN ontology are collapsed into one HAS_FEATURE link within the authoring ontology.

### 7.4.2 Ontological Compression of Link Families

A number of semantic link type 'families' can be identified within the complete set present in the OpenGALEN ontology. The set of partitive links, for example (already detailed in section 3.9.3.1 but reproduced below) are clearly recognisable as more specific flavours of the single '*hasPart'* link commonly present in simpler ontologies. The subtypes of a '*causes'* link, also illustrated below, form another family.

As stated in 3.9.3.1, explanation of why the *Open*GALEN ontology employs 16 subtypes of '*hasPart'* link is outside the scope of this thesis [Rogers 2000]. It is sufficient for this thesis to note that because semantic rules exist in the *Open*GALEN ontology determining which of the 16 flavours should be used, the authoring ontology can free inexpert users from that choice and complexity. Thus the authoring ontology does in fact contain only one partitive semantic link type: HAS_PART and its inverse IS_PART_OF. Similarly, the various subtypes of 'causes' are condensed into one CAUSES link in the authoring ontology and its inverse CAUSED_BY.

| FAMILY OF PARTITIVE LINKS | FAMILY OF CAUSAL LINKS |
|---|---|
| part-of | hasConsequence |
| structure-part-of | hasComplication |
| arbitrary-part-of | hasImmediateConsequence |
| segment-of | specificallyTriggers |
| solid-piece-of | hasLateConsequence |
| layer-of | hasRequiredConsequence |
| irregular-piece-of | |
| pouch-of | |
| component-of | |
| func-component-of | |
| partitive-connection-of | |
| partitively-contained-in | |
| mixed-throughout | |
| suspended-in | |
| surface-of | |
| makes-up | |

**Figure 46: Families of partitive and causal semantic links in CRM**

### 7.4.3 Simplified organisation of links

Following conflation of families of links, and removal of artefactual links, the number of links in the authoring ontology was 64, compared with the 400 or so in the target *Open*GALEN ontology. In addition to this reduction in the number of links, the authoring ontology also simplified their organisation. The 400 *Open*GALEN ontology links are organised into a multiaxial hierarchry, whereas the 64 authoring ontology links are a flat list. Each of the 400 *Open*GALEN ontology links is required to have an explicit inverse (e.g. *hasPart / isPartOf, hasLocation / isLocationOf* etc). Within the authoring ontology, even when a link and its inverse are present, that association is not in fact explicitly represented: no use case was encountered, within the limits of the goals for the intermediate representation, for links and there inverses to be associated. Additionally, a link can exist in the authoring ontology without any corresponding inverse (e.g. HAS_LATERALITY is included, but there is no link equivalent to IS_LATERALITY_OF).

## *7.5 Ontological Compression: Refining the ontology*

Tailoring of the authoring ontology went beyond more than a simple collation and sharing of the set of terms and links actually used by at least one author: software tools (TIGGER) were constructed to screen candidate new descriptors and allow manual exclusion of typographic or lexical variants of descriptors already encountered (artery,Artery, artary, arteries). Synonymous links such as HAS_SOURCE and FROM (used by two different authors between the notions of removing an implant and the site of the implant) were also sought out.

The usual method for managing synonymy in a controlled vocabulary is to maintain sets of alternative terms that point to a preferred concept ID within the same ontology. In the authoring ontology, however, this method was used almost exclusively to represent only typographical or lexical variants of a term as trivial synonyms. True synonymy – for example the equivalence between ‘knee cap’ and ‘patella’ – was managed by an alternative method (see section 7.5.2).

Authors therefore received a refined ‘canonical’ cumulative authoring ontology, comprising a list of descriptor and links encountered so far, via an interface that presented only preferred descriptors and links whilst still having a record of previously encountered typographical variants. Thus, in the event that an author failed to find a descriptor that was in fact present, and instead attempted to create a new one, the authoring tool would alert them if they tried to create any previously encountered typographical or lexical variant. A similar alert system for new links was not implemented as creation of new links occurred very infrequently.

This methodology allowed for the bottom-up discovery of the particular set of concepts needed by a community of authors, but at the expense of a central collation and dissemination effort. This is discussed further in chapter 12.

### 7.5.1 Removing ambiguity

The TIGGER tool also supported management of descriptor or link ambiguity: sets of dissections from different authors, using the same descriptor or link but in obviously semantically different ways, could be identified. For example, the descriptor ‘bladder’ was used by different authors in representations of both urinary and gastrointestinal surgical procedures, in the former to denote the urinary bladder and in the latter the gall bladder. Similarly, the authoring ontology link HAS_ORIGIN was used by some authors to denote the species from which a biograft originated, and by others to denote the pre-operation position of a body part in a transposition procedure. TIGGER allows the corpus of all intermediate representations to be searched for all occurrences of a link or descriptor, so that such ambiguity may be detected by manual inspection.

Descriptors identified as overtly ambiguous were flagged as ‘banned’ (henceforth neither presented to authors as part of the authoring ontology, nor permitted to be recreated), and split into unambiguous new descriptors (e.g. ‘urinary bladder’ and ‘gallbladder’). Ambiguous links were simply split; no mechanism to ban them was implemented.

In addition to maintaining a curated authoring ontology, the corpus of dissections authored so far by all authors was also periodically processed to realign it with the most recent version of the curated authoring ontology. Dissections previously authored using ambiguous descriptors or links were returned to their respective authors with appropriate global substitutions, whilst all dissections were fed back to authors with global substitution of all typographical or lexical variants by preferred forms.

### 7.5.2 Allowed redundancy

In theory, rigorous curation could have resulted in a highly compact authoring ontology. However, whilst removal of all typographical and lexical variants from the surface display would improve navigability, removal of all synonyms is not always desirable. A pragmatic decision to allow a certain amount of redundancy through synonymy was taken: dissection writing interfaces display to authors and accept different descriptors for certain concepts: for example ‘femur’ and ‘os femoris’ are both displayed.

Similarly, whilst the single HAS_FEATURE link *could* be used to assign both age and sex:

```
MAIN surgery
    HAS_PATIENT human
        HAS_FEATURE male
        HAS_FEATURE less than 12 years old
```

… many authors preferred to use separate HAS_SEX and HAS_AGE links:

```
MAIN surgery
    HAS_PATIENT human
        HAS_SEX male
        HAS_AGE less than 12 years old
```

..and for this reason these links were preserved.

### 7.6 Summary

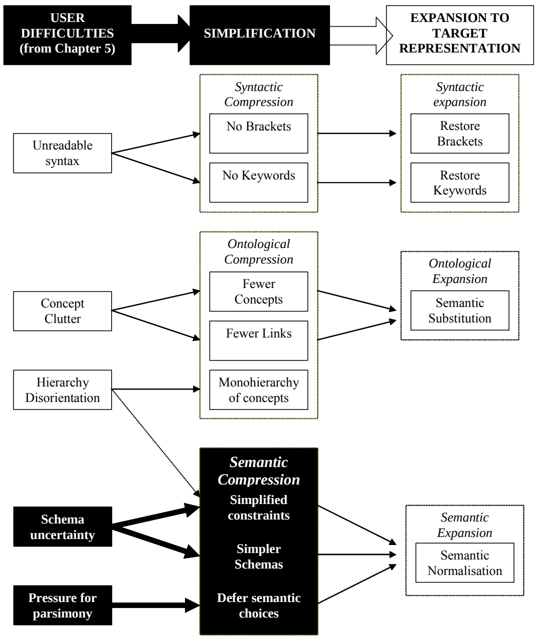
In even modest sized ontologies, new users experience difficulty finding the concepts they need for a particular purpose amongst the (often much more) numerous list of concepts they do not need.

Reducing the category and semantic link spaces to only those actually used by authors undertaking a specific task may assist them in finding concepts (or at least being certain whether those they need are truly absent). Authors can be further assisted in finding the terms they require by organising the authoring ontology into a simple monohierarchy.

# 8 Simplification Techniques: Semantic Compression

## 8.1 Introduction

This chapter describes techniques used to manage *Pressure for Parsimony* (should a missing concept be added as a new primitive or composition ?) and *Schema Uncertainty* (which schema do I use to add a composition ?). Their purpose is to reduce the need for authors to comply with the various ontological schema of the target representation or (where this is not possible) to make it easier for them to do so.




*Pressure for Parsimony* is addressed by relaxing the author-time requirement for precision and consistency and allowing the concept list to include some redundancy.

*Schema Confusion* is addressed through a simplified system of constraints and a simplification of compositional schema.

## 8.2 Semantic Compression: deferring compositional choice

Section 3.10.1 (Consistency & Canonisation) set out the problems that arise if new compositions can be expressed interchangeably in terms of either existing compositions or their more basic primitives, and if the capability to compute semantic equivalence through a process of canonisation is to be retained. Key issues were the need to know in advance which concepts are compositions, have access to their definitions, and know how those definitions relate to the various ontological schemata in force, in order to determine which parts of the schema need redundant re-expression.

An important goal of the authoring ontology and the intermediate representation, however, is freeing the inexpert user as much as possible from such requirements. The authoring ontology therefore includes single descriptors whose semantic meaning can also be represented in terms of other descriptors and semantic links in the authoring ontology. These are known as '*Decomposable Descriptors*'.

For example, Figure 47 shows three different ways by which different authors, or the same author on different days, might reasonably approach the representation of the English phrase 'left knee arthrotomy' using the authoring ontology:

```
MAIN arthrotomy                MAIN incising                  MAIN arthrotomy
   ACTS_ON left knee joint        ACTS_ON knee joint             ACTS_ON knee joint
                                     HAS_LATERALITY left            HAS_LATERALITY left
```

**Figure 47: Three ways to represent 'left knee arthrotomy'**

These three representations differ in the degree to which the descriptors with which they are expressed are 'primitive' – for example, 'left knee joint' in the far left expression is clearly equivalent to the fragment 'knee joint HAS_LATERALITY left' present in the other two. However, there is no information within the authoring ontology itself that represents this equivalence. In the context of the Intermediate Representation, preserving this expressive freedom reduces *Pressure for Parsimony*: users do not need to worry whether a term they have added could, in fact, be expressed in more primitive terms.

However, although decomposable descriptors are allowed and can be supported, a further goal of authoring ontology curation (section 7.5) is to pragmatically limit their growth: if the authoring ontology contained very large numbers of trivial decomposable descriptors – for example, a 'left' form of every piece of anatomy, such as 'left leg', 'left hand' etc. – then the descriptor list would quickly grow to an unmanageable size. There is a trade off between improved accessibility when words or phrases entered directly by an author are more likely to recognised, and reduced accessibility when an author can not think of the term and needs to search through all available descriptors to find the most appropriate one.

Where a decision is made to removing an undesirable decomposable descriptor from the authoring ontology, TIGGER provides a mechanism to globally replace all individual occurrences of such descriptors with their semantically equivalent graphs, expressed in terms of other descriptors and links in the authoring ontology.

In addition to including decomposable descriptors, the authoring ontology also includes single primitive descriptors whose semantic equivalent in the target ontology is a composed, rather than a primitive, concept in that ontology. For example, the semantically equivalent representations in the target ontology for the descriptors 'arthrotomy' and 'left knee joint' are shown in Figure 48:

| DESCRIPTOR | TARGET ONTOLOGY COMPOSITION |
|---|---|
| left knee joint | (KneeJoint which hasLeftRightSelector leftSelection) |
| arthrotomy | (Incising which actsSpecificallyOn Joint) |

**Figure 48: Compositions in target ontology equivalent to primitive descriptors in authoring ontology**

Section 9.3.1.2 details how information concerning equivalences between decomposable descriptors and primitives in the authoring ontology, and ontological schema in the OpenGALEN ontology, are maintained outside the authoring ontology and used during the rewrite operation.

## 8.3 Semantic Compression: Workaround artefacts

Section 3.6 described how one source of complexity in ontology construction stems from the need to work around various limitations of the GRAIL formalism underpinning the OpenGALEN ontology. The following sections describe the techniques by which such 'artefact' schema can be systematically abbreviated. Note that such semantic compression has a further effect: not only does it reduce *Schema Confusion* by freeing the user of the need to be familiar with arcane schema, it also improves surface readability and reduces *Unreadable Syntax*.

The worked *tenesmus* example previously used in Chapter 6 to illustrate syntactic simplification is revisited to illustrate semantic compression.

### 8.3.1 Compressing the negation wrapper

The CRM negation wrapper was detailed in 3.7.3. The representation of *tenesmus* provides a real example of this wrapper. Figure 49 reproduces this worked example, with elements relevant to the negation wrapper highlighted in bold:

```
clinical situation
    IS_CHARACTERISED_BY presence
        IS_EXISTENCE_OF contraction process
            IS_SPECIFIC_FUNCTION_OF sphincter ani muscle
            HAS_IMMEDIATE_CONSEQUENCE pain
            HAS_INTENTIONALITY intentionality
                HAS_ABSOLUTE_STATE involuntary
            HAS_DURATION duration
                HAS_ABSOLUTE_STATE long term
            HAS_TEMPORAL_PATTERN temporal pattern
                HAS_ABSOLUTE_STATE ongoing
    IS_CHARACTERISED_BY presence
        IS_EXISTENCE_OF urge to void urine or faeces
            HAS_PROCESS_ACTIVITY process activity
                HAS_QUANTITY level
                    HAS_MAGNITUDE high level
    IS_CHARACTERISED_BY presence
        IS_EXISTENCE_OF abdominal straining
```

**Figure 49: Worked example highlighting wrapper for negation**

As previously described, the negation wrapper schema has a very regular pattern. This lends itself to being captured using a much abbreviated syntax: the user need only signify firstly that a negation wrapper construct is to be used at all, following which they need specify only which branches of the wrapped concepts are negated, and which are not.

For the reasons described in section 3.7.3, **all** dissections expressed in the Intermediate Representation are expanded into negation wrapper constructs. Negated branches are identified using the WITHOUT link, whilst MAIN or WITH signals absence of negation. Rewritten in this abbreviated syntax, the tenesmus example shown originally in Figure 38 is further reduced to Figure 50:

```
MAIN ContractionProcess
        isSpecificFunctionOf SphincterAniMuscle
        hasImmediateConsequence Pain
        hasIntentionality Intentionality
            hasAbsoluteState involuntary
        hasDuration Duration
            hasAbsoluteState longTerm
        hasTemporalPattern TemporalPattern
            hasAbsoluteState ongoing
WITH UrgeToVoidUrineOrFaeces
            hasProcessActivity ProcessActivity
                hasQuantity Level
                    hasMagnitude highLevel
WITH AbdominalStraining
```

**Figure 50: Worked example with wrapper compression**

It can be seen that the effect of this semantic compression is to replace the 10 wrapper links or concepts in the original expression to 3 links (MAIN, WITH and WITH) with corresponding improvement in overall readability as a by-product.

### 8.3.2 Compressing the Feature-State cardinality workaround

Section 3.7.1 has described the Feature-State workaround for cardinality, whilst Section 3.9.4 demonstrated how consistent use of the Feature-State schema significantly contributes to the surface complexity and verbose nature of many desired expressions. Fortunately, the regular nature of the workaround lends also itself to a more compressed syntax.

Essentially, once it is known precisely which subtype of [State] is to be assigned – for example [red, pink, severe, mild, radiolucent, radioopaque] - it is possible to work out all the preceeding semantic elements required to fit the schema. The user no longer needs to specify exactly which flavour of the [*hasFeature*] semantic link type should be used, nor to identify which descendent of [Feature] should be used as the intermediary concept.

Representations can therefore be simplified still further, such that the worked 'tenesmus' example may now be rewritten in its final and most compact form, as shown in Figure 51:

```
MAIN contraction process
    IS_SPECIFIC_FUNCTION_OF sphincter ani muscle
    HAS_IMMEDIATE_CONSEQUENCE pain
    HAS_FEATURE involuntary
    HAS_FEATURE long term
    HAS_FEATURE ongoing
WITH urge to void urine or faeces
    HAS_FEATURE high level
WITH abdominal straining
```

**Figure 51: Final IR format of worked example**

## *8.4 Semantic Compression: Simplified constraints*

The GRAIL metamodelling functionality, implemented through the constraint system declared using the keywords grammatical and sensible - has been described in section 3.11. Its stated goal is that *all and only all* meaningful compositions in a GRAIL ontology should be permitted [Goble 1996, Rector 1997]. The corollary is that all nonsense or meaningless compositions should be rejected because the constraint system does not permit them. In practice, the OpenGALEN ontology constraints fall short of excluding all nonsense compositions but, in pursuit of approximating to its stated ideal, more than 24,000 separate constraint rules are in operation.

The original purpose and ultimate goal of the GRAIL constraint system is to drive a 'predictive data entry system' [Goble 1994, Rector 1995b], i.e. that for any given concept that a user wishes to talk about, the constraint system dictates *a priori* both which semantic link types can be selected from an interface picking list and also the range of values each link can take, in order to further describe that concept. Because the options offered are contextually 'sensible', users tend to perceive the system less as limiting or dictating their options, and more as anticipating or predicting what they might want to say.

Predictive data entry, however, equates to a sophisticated system for limiting *Schema Confusion*: the author is told in advance what more they are permitted to say, and how.

An obvious limitation of the predictive approach is that, until the constraint system is substantially complete, it is difficult or impossible to create such an interface for use by naïve authors. This effectively prevents any early end-user testing and feedback such as could inform construction of either a complex ontology (e.g. regarding scope of terms required) or its constraints (e.g. scope of compositional patterns either required or expected).

The intermediate representation constraints were designed as a compromise constraint system. As in the *Open*GALEN ontology sanctions, their goal was simultaneously to reduce the ability of authors to create nonsense expressions and to drive a predictive data entry interface. However, the authoring ontology provides significantly less expressivity for authoring constraints: in the *Open*GALEN ontology, a rule can be authored in terms of any of its 20,000 concepts and 400 semantic links. In the intermediate representation, although any authoring ontology link may be used, constraints are not authored between descriptors but between the descriptor categories. Finally, whereas the *Open*GALEN ontology offers two levels of constraint (grammatical and sensible) the intermediate representation has only one level. The reduced expressivity, together with the less ambitious goals, result in an constraint system for the authoring ontology that runs to just over 200 individual rules, expressed in terms of 115 descriptor categories and 80 semantic links.

The following extract (Figure 52) of the authoring ontology constraints illustrates the general form:

| | | |
|---|---|---|
| BodySubstance | IS_PART_OF | Anatomy |
| BodySubstance | IS_PART_OF | Lesion |
| Deed | BY_MEANS_OF | BodySubstance |
| Device | HAS_LOCATION | BodySubstance |
| Pathology | HAS_LOCATION | BodySubstance |

**Figure 52: Extract of authoring ontology constraints**

The first two constraints dictate that, for any descriptor of category ‘BodySubstance’ (e.g. blood, pus, articular cartilage, stroma, mucous), it is permissible to further describe it using only one of the semantic links also in the authoring ontology: IS_PART_OF. Further, this single link can only be used to describe a ‘BodySubstance’ as being part of either some piece of ‘Anatomy’ or a ‘Lesion’. These constraints, therefore, effectively exclude compositions such as:

| | | |
|---|---|---|
| blood | IS_PART_OF | mucous |

The final three constraints dictate that a ‘BodySubstance’ can be the value following one of two semantic links types (BY_MEANS_OF or HAS_LOCATION) but only if these link from one of three other categories of descriptor, as appropriate.

In common with the GRAIL constraint model, the intermediate representation constraint model also supports inheritance of constraints over a hierarchy. For intermediate representation constraints, inheritance is over the descriptor category hierarchy (see section 7.3.1 and Appendix One). The semantics of, for example:

| | | |
|---|---|---|
| BodySubstance | IS_PART_OF | Anatomy |

Is in fact to say that:

> *Any* descriptor of category ‘BodySubstance’, *or* any descriptor of a category subsumed by ‘BodySubstance’, may be combined using the link IS_PART_OF with *any* descriptor of category ‘Anatomy’ *or* any descriptor of a category subsumed by ‘Anatomy’.

Thus, in order to establish the full set of constraints that apply to descriptors of category ‘BodySubstance’ it is necessary to consider not only those already listed and applying directly on that category, but also other constraints authored using more general ancestors of ‘BodySubstance’, for example ‘Substance’ – a category that subsumes both ‘BodySubstance’ and ‘Chemical’ in the hierarchy of categories. These are listed in Figure 53:

| | | | | | |
|---|---|---|---|---|---|
| Anatomy | CAUSED_BY | Substance | Substance | CAUSES | Process |
| Anatomy | CAUSES | Substance | Substance | HAS_DESTINATION | Lesion |
| Pathology | CAUSED_BY | Substance | Substance | HAS_DESTINATION | Structure |
| Lesion | HAS_DESTINATION | Substance | Substance | HAS_DESTINATION | Substance |
| Structure | HAS_DESTINATION | Substance | Substance | HAS_FUNCTION | Process |
| Anatomy | HAS_PART | Substance | Substance | HAS_FEATURE | Feature |
| Feature | IS_FEATURE_OF | Substance | Substance | HAS_LOCATION | Structure |
| Process | ACTS_ON | Substance | Substance | IS_ACTED_ON_BY | Deed |
| Anatomy | IS_MADE_OF | Substance | Substance | IS_CONTAINED_IN | Structure |
| Device | IS_MADE_OF | Substance | Substance | IS_MADE_OF | Material |
| Structure | HAS_PROXIMITY | Substance | Substance | HAS_PROXIMITY | Structure |

**Figure 53: Authoring ontology constraints on category 'Substance'**

A comparison between these authoring ontology constraints, and the GRAIL sanctions in the target *Open*GALEN ontology, reveals that the former duplicate most of the sense and scope of the latter's grammatical sanctions. In that regard an important function of the authoring ontology constraints is to enforce the same high level schema and meta-metamodel of the target ontology.

### 8.4.1 Authoring ontology Constraints and *Concept Clutter*

Although mainly intended as a means to guide and constrain authors to a particular semantic style, a side effect of using the intermediate constraints to drive a predictive user interface is that the list of descriptors presented to authors at any one time is trimmed still further, so that they are never presented with the full authoring ontology as a picking list. Instead, they are presented only with a subset determined by which categories may be used after a specified semantic link. This further reduces *Concept Clutter*.

### 8.4.2 *Schema Confusion*: The Style Guide

A final resource made available to authors to assist them in determining the most appropriate semantic representation in a given situation was a corpus of previously authored idealised representations, indexed by both semantic links used or by individual descriptor. An author struggling with how to represent the semantic relationship between, for example, an anastomosis and the structures it joins, can search the style guide corpus for any representation that employs the descriptor ‘anastomosis’. Such representations in the style guide contain implicit semantic rules that can be inferred by authors, whether or not they are (or can be) made explicit and enforced via any intermediate representation authoring interface.

## 8.5 Summary

Basic constraints guide intermediate representation authors in their choices of semantic links and values to further describe concepts, thereby assisting them in following the high level schema of the target ontology.

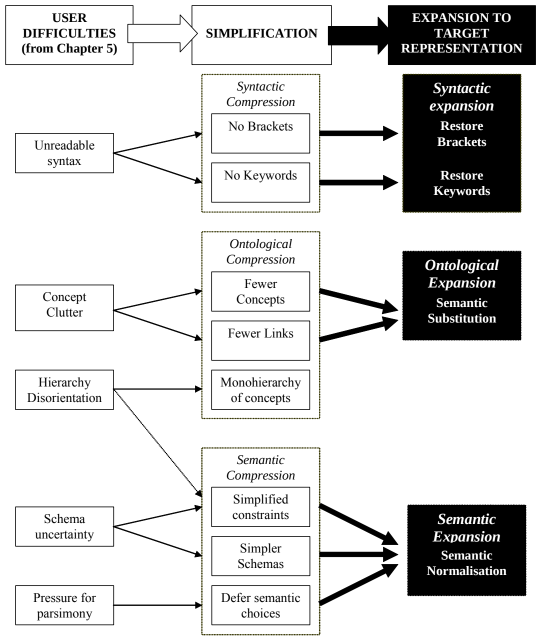
Deferring the choice of whether something is a primitive or a composition in the OpenGALEN ontology relieves some *Pressure for Parsimony*.

# 9 Reversing the simplification: From IR to GRAIL

The previous three chapters have described measures taken to make it easier for relatively inexpert users to author representations, by inviting them to use a simpler intermediate representation. For an intermediate representation to be 'intermediate', however, a mechanism is required by which expressions in the intermediate representation are re-written using the target representation syntax and ontology.

This chapter describes that mechanism.

## 9.1 Introduction




The previous three chapters have detailed systematic transformations to simplify certain aspects of the target *Open*GALEN ontology and GRAIL representation (its syntax, large vocabulary etc). Correspondingly, part of the re-write methodology amounts to formalising the reverse of these transformations.

- The syntactic compression of Chapter 6 is reversed by a *syntactic expansion* in which brackets and keywords are restored
- The ontological compression by means of the authoring ontology of Chapter 7 is reversed by an *ontological expansion:* an exhaustive declaration of descriptor mappings and link mappings allows expressions using the authoring ontology to be *semantically substituted* with equivalent concepts from the OpenGALEN ontology.
- The semantic compression of Chapter 8 is reversed by a *semantic expansion*. Compressed syntaxes for common patterns (wrappers, feature-state) are expanded to their full form. The Pressure for Parsimony of Chapter 4, in addition to being managed *pre hoc* by restrictive constraints that prohibit some alternate schemas is additionally managed post-hoc by *semantic normalisation*.

These three processes are described further in the following sections.

## 9.2 Syntactic expansion

Chapter 6 described how the surface syntax of the OpenGALEN ontology was systematically simplified to derive the intermediate representation syntax by hiding both bracketing and keywords. This section describes how these elements are restored.

### 9.2.1 Syntactic Expansion: reinstating brackets

The first step in rewriting intermediate representation expressions using the more complex GRAIL syntax requires a straightforward replacement of the tabbed indentation employed in the intermediate representation with explicit use of bracketing, such that:

| ORIGINAL INTERMEDIATE REPRESENTATION | AFTER REINSTATING BRACKETS |
|---|---|
| MAIN excising<br>ACTS_ON leg<br>IS_ACTED_ON_BY ischaemia<br>HAS_LATERALITY left<br>WITH applying<br>ACTS_ON bandage | MAIN (excising<br>ACTS_ON (Leg <<br>IS_ACTED_ON_BY ischaemia<br>HAS_LATERALITY left>))<br>WITH (applying<br>ACTS_ON bandage) |

**Figure 54: Restoration of bracketing**

### 9.2.2 Syntactic Expansion: reinstating keywords

The second step of the re-write is to re-instate the GRAIL keywords. As explained in Section 6.3, only two GRAIL keywords need to be considered: which and whichG. *Where* one or other keyword should appear in an expanded target ontology expression is predicted by each new level of tabbed indentation in the source intermediate representation expression, as shown in Figure 55:

| AFTER REINSTATING BRACKETS | LOCATION OF GRAIL KEYWORDS |
|---|---|
| MAIN excising<br>ACTS_ON leg<br>IS_ACTED_ON_BY ischaemia<br>HAS_LATERALITY left<br>WITH applying<br>ACTS_ON bandage | MAIN (excising which./whichG<br>ACTS_ON (Leg which./whichG <<br>IS_ACTED_ON_BY ischaemia<br>HAS_LATERALITY left>))<br>WITH (applying which./whichG<br>ACTS_ON bandage) |

**Figure 55: Reinstating GRAIL operator keywords**

The remaining issue is to determine whether, in each instance, the which or the whichG operator should be used. Section 3.11 outlined the difference between the which and whichG operators: they have a similar function (specifying compositions comprising two concepts and a semantic link) but they invoke different levels of check against the metamodel and meta-metamodel of the ontology.

The crudest mechanism by which the rewrite mechanism could interact with the target ontology sanctioning system would be to largely ignore it: the whichG constructor could be used throughout, thereby requiring only that all compositions should be consistent with the meta-metamodel (grammatical level of sanctioning). The effect of permanently relaxing the constraint model in this way would be that only the most grossly semantically incorrect expressions – for example 'Colour whichG *isPartOf* Structure', or 'Time whichG *isDurationOf* Leg' - would still be rejected because they contravened the highest level ontological commitments of the CRM, encoded primarily by means of the meta-metamodelling grammatical level sanctions. This relaxation would therefore still allow many nonsense expressions: 'Ear whichG *isPartOf* Leg' is grammatically sanctioned in the CRM meta-metamodel.

Whilst this simplification might serve the purposes of the user, and fulfil the goal of freeing them from having to determine why compositions were rejected (since almost none would ever be rejected), it is less useful for those maintaining the target ontology itself. The corpus of expressions written in intermediate representation can be mined to identify where the *Open*GALEN ontology constraints are incomplete: if an intermediate representation user attempts to create an obviously plausible composition (e.g. they try to combine [Asthma] with [*hasSeverity*] and [severe]), but the *Open*GALEN ontology constraints reject it, then this suggests that those constraints need changing.

The intermediate representation, therefore, employs a more sophisticated and less permanently relaxed strategy (detailed further in 9.3.2): the which operator is used if a metamodel (sensible) sanction exists that would allow it, and the whichG if only a meta-metamodel (grammatical) level sanction does. The composition is rejected if no sanction is found in either the metamodel or the meta-metamodel. However, whenever the whichG operator must be used, it is used with the *least* specific semantic link type so sanctioned. Conversely, the which operator is used with the *most* specific link type so sanctioned. (The mechanism by which this is achieved is presented in 9.3.2.3).

Underlying this approach - selective relaxation of the constraint model - is the fundamental assumption that the inexpert user, despite their naivety, is usually right: unless a candidate composition written in the intermediate representation breaks the top level ontological meta-metamodel of the target ontology, it is assumed to be semantically valid and the error to lie in a missing metamodel (sensible) sanction in the target ontology. The list of occasions where it has been necessary to relax the constraint model to the grammatical level of sanctioning can be studied elsewhere and at a later date by experts. If many independent users have constructed similar composition fragments that appear semantically valid, but these were all rejected, it suggests new sanctions that should be added to the existing set of permissions in either the metamodel or the meta-metamodel.

## 9.3 Ontological Expansion

Chapter 7 described the creation of an authoring ontology, comprising a set of concept labels and semantic links that are entirely distinct from those in the target ontology. This section describes the steps necessary to enable expressions written using the authoring ontology to be re-written using the target *Open*GALEN ontology.

### 9.3.1 Semantic Substitution: Descriptor mappings

Given that the authoring and target ontologies are independent, but ultimately any representations using the authoring ontology are to be expressed in terms of the target ontology, it follows that an exhaustive set of mappings must be authored: for each authoring ontology descriptor a semantically equivalent concept in the target ontology must be declared.

As will be demonstrated, a byproduct of these mapping are that they allow the *Pressure for Parsimony* to be managed, by allowing:

- authoring ontology *redundancy* (e.g. descriptors such as ‘femur’ and ‘os femoris’ co-existing within the authoring ontology);
- *decomposable descriptors* (whose meaning can also be represented using a combination of other descriptors and links in the authoring ontology).

In many cases the mapping is trivial: the concepts may have the same, or similar, name in both ontologies, for example the authoring ontology descriptor ‘carotid artery’ is mapped to the OpenGALEN ontology concept [CarotidArtery]. Tools exist that propose candidate semantically equivalent mappings based on lexical matching of knowledge names. Such lexically proposed mappings must be manually validated. Where no lexical match can be found, a mapping must be determined and asserted entirely by hand.

One important advantage of the bottom-up development of an authoring ontology, independent of any final target ontology, is that dissection authors can write expressions using concepts not yet modelled in the target ontology. A further advantage of this situation is that target ontology development can therefore be scoped to meet user requirements, as they become apparent. However, a consequence is that some descriptor mappings can not be declared until the target ontology has been appropriately modified or extended.

In other cases, however, a descriptor may correspond to a concept that, though not currently installed in the target ontology, can be created as a new composition within it in terms of other concepts that *are* already present. For example, 'left eyebrow' might be a new descriptor and, although no directly equivalent concept already existed in the target ontology, the concepts [left] and [eyebrow] were already present.

This situation is managed smoothly: descriptor mappings are not required to correspond to fixed or persistent identifiers for concepts already installed in the target ontology. They are, in fact, all treated as representations of concepts that *might be* evaluated in the target ontology. Following canonisation of such a representation (as described in 3.10.1), the classification engine will add it **if and only if** it is found that such a concept is not already installed in the target ontology. Similarly, if after canonisation and evaluation it is found that, in fact, a mapping can not be evaluated by the target ontology (perhaps it is not sanctioned, or is expressed in terms of a target ontology primitive that no longer exists) then an attempt to use that mapping will fail and an alert will be raised.

An example of typical descriptor mappings is shown in Figure 56:

| AUTHORING ONTOLOGY DESCRIPTOR | SEMANTICALLY EQUIVALENT CONCEPT IN OPENGALEN ONTOLOGY |
|---|---|
| left | LeftSelection |
| left knee joint | (KneeJoint which *hasLeftRightSelector* leftSelection) |
| knee joint | KneeJoint |
| femur | Femur |
| os femoris | Femur |
| excising | Excising |
| angioscope | Device which<br>*isSpecificPhysicalMeansOf* (NonDirectInspecting which<br>*actsSpecificallyOn* BloodVessel) |
| subluxation | BodyStructure which <<br>*hasUniqueAssociatedProcess* (DislocationProcess which<br>*hasCompleteness* (Completeness which<br>*hasAbsoluteState* partial))<br>*hasPathologicalStatus* pathological > |

**Figure 56: Examples of typical descriptor-to-CRM mappings**

#### 9.3.1.1 Descriptor redundancy: Mappings of semantically equivalent descriptors

From the examples shown in Figure 56, it can be seen in the first instance how simple descriptor redundancy within the authoring ontology is resolved: semantically equivalent descriptors in the authoring ontology (e.g. femur, os femoris) are mapped to the same concept (e.g. femur) in the *Open*GALEN ontology. Simply by substituting with declared mappings, two apparently different intermediate representation expressions formulated using redundantly equivalent descriptors will be re-written as the same expression, as illustrated in Figure 57:

| ORIGINAL EXPRESSIONS USING AUTHORING ONTOLOGY | AFTER RE-WRITE BY SUBSTITUTION WITH MAPPED TARGET ONTOLOGY CONCEPT |
|---|---|
| excising ACTS_ON femur | Excising ACTS_ON Femur |
| excising ACTS_ON os femoris | Excising ACTS_ON Femur |

**Figure 57: Descriptor redundancy managed through common mappings**

Semantic equivalence in the authoring ontology is thus represented not by the usual method of declaring which terms are synonymous references to a preferred concept within that ontology, but by asserting that they are individually semantically equivalent to a concept in an *external* ontology – the target ontology. This indirection may potentially go further: Figure 58 shows two different authoring ontology descriptors that have two different target ontology mappings but, when those mappings are evaluated, are found to reduce to the same canonical form (see 3.10.1):

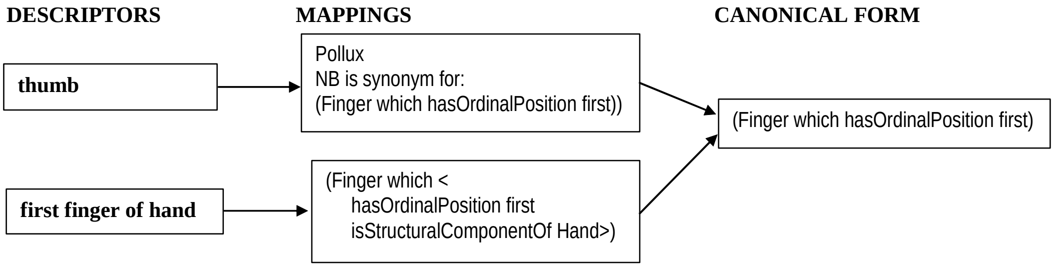



**Figure 58: Role of canonisation in determining equivalent descriptor mappings**

Thus semantic equivalence of descriptors may, ultimately, be inferred only through semantic computation within the target ontology, rather than by simple lookup in tables such as Figure 56.

#### 9.3.1.2 Mappings of decomposable descriptors

The inclusion of decomposable descriptors within the authoring ontology (section 8.2) introduces a more complex challenge to resolve. Two authors may write expressions in the authoring ontology that are semantically equivalent to the human reader but which, even after simple mapping substitution as above, remain superficially different. Figure 59 provides an example of this phenomenon.

Figure 59 also hints at what is still required such that the final rewrites of the two original expressions are identifiable as semantically equivalent: the authoring ontology link HAS_LATERALITY needs to be rewritten using the target ontology semantic link [*hasLeftRightSelector*]. The next section describes how intermediate representation semantic links are mapped to the target ontology, including how the requirement for rewriting HAS_LATERALITY is fulfilled.

| ORIGINAL EXPRESSIONS USING AUTHORING ONTOLOGY | AFTER RE-WRITE BY SUBSTITUTION WITH MAPPED TARGET ONTOLOGY CONCEPT |
|---|---|
| excising<br>ACTS_ON left knee joint | Excising<br>ACTS_ON (KneeJoint which<br>*hasLeftRightSelector* leftSelection) |
| excising<br>ACTS_ON knee joint<br>HAS_LATERALITY left | Excising<br>ACTS_ON (KneeJoint<br>HAS_LATERALITY leftSelection) |

**Figure 59: Expansion of decomposable descriptors**

### 9.3.2 Semantic Substitution: Link Mappings

Section 7.4 described how several 'families' of semantic link types can be identified within the OpenGALEN ontology (e.g. the '*hasPart*' and '*causes*' families). To simplify the authoring ontology, each of these families were conflated into a single semantic link. (HAS_PART, CAUSES) and their corresponding inverses (IS_PART_OF, CAUSED_BY). Separately identified were those semantic links whose existence within the *Open*GALEN ontology was largely an artefact of the GRAIL formalism's limited implementation of cardinality (e.g. the many subtypes of [*hasFeature*]). These artefactual links were also conflated (e.g. into the single authoring ontology link, HAS_FEATURE)

A many-to-one conflation of multiple target ontology links into single authoring ontology links correspondingly means (potentially) a one-to-many mapping from authoring ontology to target ontology links. Each authoring ontology link therefore requires both a list of candidate mappings to one or more target ontology links, and an algorithm for choosing between them. The following sections provide further detail of how candidate mapping sets are represented, and of the algorithm (Figure 64) for selecting from a set of candidate mappings.

#### 9.3.2.1 Link mappings: one-to-one mappings

For some authoring ontology links the candidate mapping list is trivially short: a one-to-one mapping exists between them and their direct semantic equivalent in the *Open*GALEN ontology. Figure 60 shows examples of two authoring ontology links with a one-to-one mapping to the target ontology:

| AUTHORING ONTOLOGY LINK | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY |
|---|---|
| OCCURS_DURING | *OccursDuring* |
| IS_BRANCH_OF | *IsBranchOf* |

**Figure 60: Two examples of one-to-one link mappings**

In either of the above cases, whenever one of the authoring ontology links is encountered in an expression, the rewrite will in all cases attempt to substitute only the single mapping choice presented. However, as described in section 9.2.2, an important issue is whether the GRAIL <u>which</u> or <u>whichG</u> operator should be used with the candidate link mapping.

To make the choice, the target ontology is queried to establish whether a metamodel (sensible) sanction exists to permit use of the specific link mapping candidate between the two concepts that map to the descriptors on either side of the original link. If such a sanction exists, the <u>which</u> operator is used in the final rewrite. If not, the target *Open*GALEN ontology is requeried to determine whether a meta-metamodel (grammatical) sanction exists. If one does, then the <u>whichG</u> operator will be used. If neither type of sanction exists, the rewrite algorithm for the entire expression fails and an error is raised.

Thus, given the particular set of descriptor and link mappings in Figure 61, the rewrite of the intermediate representation fragment:

artery IS_BRANCH_OF aorta

…would query the *Open*GALEN ontology to establish whether a sensible or grammatical sanction existed to permit:

ArterialStructure <u>which</u> *isBranchOf* Aorta

| AUTHORING ONTOLOGY LINK OR DESCRIPTOR | DESCRIPTOR OR CANDIDATE LINK MAPPING IN TARGET ONTOLOGY |
|---|---|
| IS_BRANCH_OF | *IsBranchOf* |
| bacteria | Bacterium |
| artery | ArterialStructure |
| aorta | Aorta |

**Figure 61: A set of descriptor and link mappings**

…and finding that only a grammatical sanction exists, the final rewrite would be:

ArterialStructure whichG isBranchOf Aorta

By contrast, an attempt to rewrite the fragment:

bacteria IS_BRANCH_OF aorta

..would fail and report an error, since there is not even a meta-metamodel grammatical sanction that permits an [Organism] to be a *branchOf* a [Structure].

This basic algorithm is extended when an authoring ontology link corresponds not to a single link in the OpenGALEN ontology, but instead can best be represented as a 'chain' of links and entities. To cope with this scenario, chains of arbitrary length can be represented explicitly as illustrated by the following table:

| AUTHORING ONTOLOGY LINK | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY (mappings are 'chains') |
|---|---|
| BYPASSES | *isSpecificPhysicalMeansOf* Bypassing *actsSpecificallyOn* |
| BY_APPROACH_TECHNIQUE | *hasSpecificSubprocess* SurgicalApproaching *hasSpecificSubprocess* |

**Figure 62: Authoring ontology links mapped to target ontology chains**

The algorithm to select which or whichG then runs in a similar way. To rewrite the fragment:

stent BYPASSES aorta

…the OpenGALEN metamodel is first queried whether the following is sanctioned:

Stent which *isSpecificPhysicalMeansOf* (Bypassing whichG *actsSpecificallyOn* Aorta)

(NB note that whichG is always inserted within chain definition)

…and if this fails, the following is tried instead against the meta-metamodel:

Stent whichG *isSpecificPhysicalMeansOf* (Bypassing whichG *actsSpecificallyOn* Aorta)

#### 9.3.2.2 Link Mappings: context sensitive mappings

Simple cases where an authoring ontology link has only one candidate mapping are the minority: usually there is more than one candidate link mapping. A relatively simple example of this is given by the authoring ontology link HAS_LATERALITY, used in practice by authors to assign variously the values left, right, bilateral, medial and lateral to anatomical structures.

In the target *Open*GALEN ontology, the mutual exclusivity of these laterality value sets is enforced using a mechanism similar to that used for assigning [Features] to [States]: each value set (left vs right, medial vs lateral etc) has its own semantic link with single cardinality. Thus, determining the correct rewrite of the authoring ontology link HAS_LATERALITY depends directly on the context in which it is found: from which value set is the descriptor following it drawn? If the descriptor is either 'left' or 'right', then the rewrite should be *hasLeftRightSelector* whereas if the descriptor following is either 'medial' or 'lateral' then the link rewrite should be *hasMedialLateralSelector*, and so on.

In the general case, lists of candidate mappings for each authoring ontology link were drawn up iteratively by empirical analysis of how that link was actually used by the users. The set of semantic contexts in which each link appeared was enumerated and, for each such context, a set of target ontology links for consideration (known as 'candidate link mappings') was specified together with an order in which to consider them. The combined set of semantic contexts and candidate link mappings for each authoring ontology link are known as 'the link mappings' for that link. The complete list of authoring ontology links, together with their respective link mappings, are presented in Appendix Two.

Figure 63 sets out the link mappings for the link HAS_LATERALITY:

| AUTHORING ONTOLOGY LINK | CONTEXT | | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY |
|---|---|---|---|
| | LEFT | RIGHT | |
| HAS_LATERALITY (Context #1) | Any | BilateralUnilateralSelector | *hasBilateralUnilateralSelector* |
| HAS_LATERALITY (Context #2) | Any | Any | *hasLeftRightSelector*<br>*hasMedialLateralSelector*<br>*hasPositionalSelector* |

**Figure 63: Candidate link mappings for HAS_LATERALITY**

Two separate contexts (the rows in the table) are shown in Figure 63 for the HAS_LATERALITY link, and each context is associated with a single candidate link mapping set, in the right hand column of the row. Each link context is defined as applying for every occurrence of that link between a preceeding ('left' context) concept of a certain type, and a following ('right' context) concept of another type.

Left and Right contexts are expressed in terms of concepts from the target ontology, rather than descriptors or descriptor categories. Therefore, a given left context applies when the target ontology concept mapped to the descriptor preceeding the link is subsumed by the target ontology concept defining the left context. Similarly, a right context applies if the concept mapped to the descriptor following the link is subsumed by the concept defining the right context.

This choice – to represent contexts in terms of target ontology concepts, not descriptors – was made because descriptors, and particularly their very simple monoaxial categorial structure, are optimised only for navigation. The target ontology, which can be accessed via the descriptor mappings, provides a much richer means to specify the semantic context in which an authoring ontology link appears.

For any given authoring ontology link with more than one possible context, the set of contexts is considered in turn in the declared order until one is found that applies.

If a context is found to apply (left and right contexts are satisfied) but none of the candidate mappings considered is sanctioned either sensibly or grammatically, then the rewrite for the entire expression fails and an error is raised. If all contexts are considered, but none applies, the rewrite for the entire expression also fails and an error is raised.

In the table above, Context #1 is recognised whenever the descriptor to the right of HAS_LATERALITY (its 'right' context) is mapped to a concept in the *Open*GALEN ontology that is either subsumed by, or equal to, the concept [BilateralUnilateralSelector]. The left context is set to 'any', indicating that Context #1 applies provided only the right context is satisfied, and regardless of the descriptor to the left of the link.

In the event Context #1 is recognised, the candidate link mapping set has only one candidate [*hasBilateralUnilateralSelector*] and the rewrite algorithm proceeds as for the previous section. However, if the first context is not recognised (ie if the descriptor following HAS_LATERALITY is neither 'bilateral' or 'unilateral'), then the rewrite algorithm moves on to consider whether the Context #2 applies.

In the case of the HAS_LATERALITY link, the Context #2 applies if *any* descriptor appears either to the left of the link (its 'left' context) or to its right (its 'right' context). Thus, for any given instance of the HAS_LATERALITY link in an expression, if the first context does not apply then the second context, which amounts to a 'no specific context', always will.

#### 9.3.2.3 Link Mappings: multiple candidate mappings

Context #2 – the 'no specific context' context - for the authoring ontology link HAS_LATERALITY presents a candidate link mapping set comprising an ordered choice of three different *Open*GALEN ontology links. Multiple candidate mappings are processed according to the rewrite flowchart algorithm shown in Figure 64.

Beginning at the top of a candidate link mapping list, each candidate link mapping is considered in turn. The target *Open*GALEN ontology is queried to determine whether a sensible sanction exists to permit its use. If such a sanction exists, that candidate is selected and the remaining candidates are not considered. If no sensible sanction exists, the next candidate in the list is considered. This proceeds until either a sensibly sanctioned candidate is found, or the end of the list is reached.

If no sensibly sanctioned candidate is found, and the end of the list is reached, then the algorithm restarts but *working backwards from the last candidate mapping*. Additionally, this time the target ontology is queried for the existence of a grammatical sanction. As before, the algorithm selects the first candidate encountered where such a sanction exists. If no grammatically sanctioned candidate is encountered either, and the top of the list is reached, then the rewrite for the entire expression fails and an error is raised.

The algorithm described in section 9.3.2.1 is, therefore, the result of running the algorithm described in this section when the link context and candidate mapping lists both have a list size of one. The complete algorithm is presented in Figure 64 as a flowchart.

It may also be seen how, for the decomposable descriptor example in 9.3.1.2, the algorithm results in HAS_LATERALITY being rewritten as hasLeftRightSelector, as was required: Context #1 does not apply (the descriptor following the link is neither 'unilateral' nor 'bilateral'), whilst Context #2 (the 'no specific context' context) does. Because the *Open*GALEN ontology contains a sensible sanction that allows a knee joint to be linked to 'leftSelection' using hasLeftRightSelector, it is selected for the final rewrite.

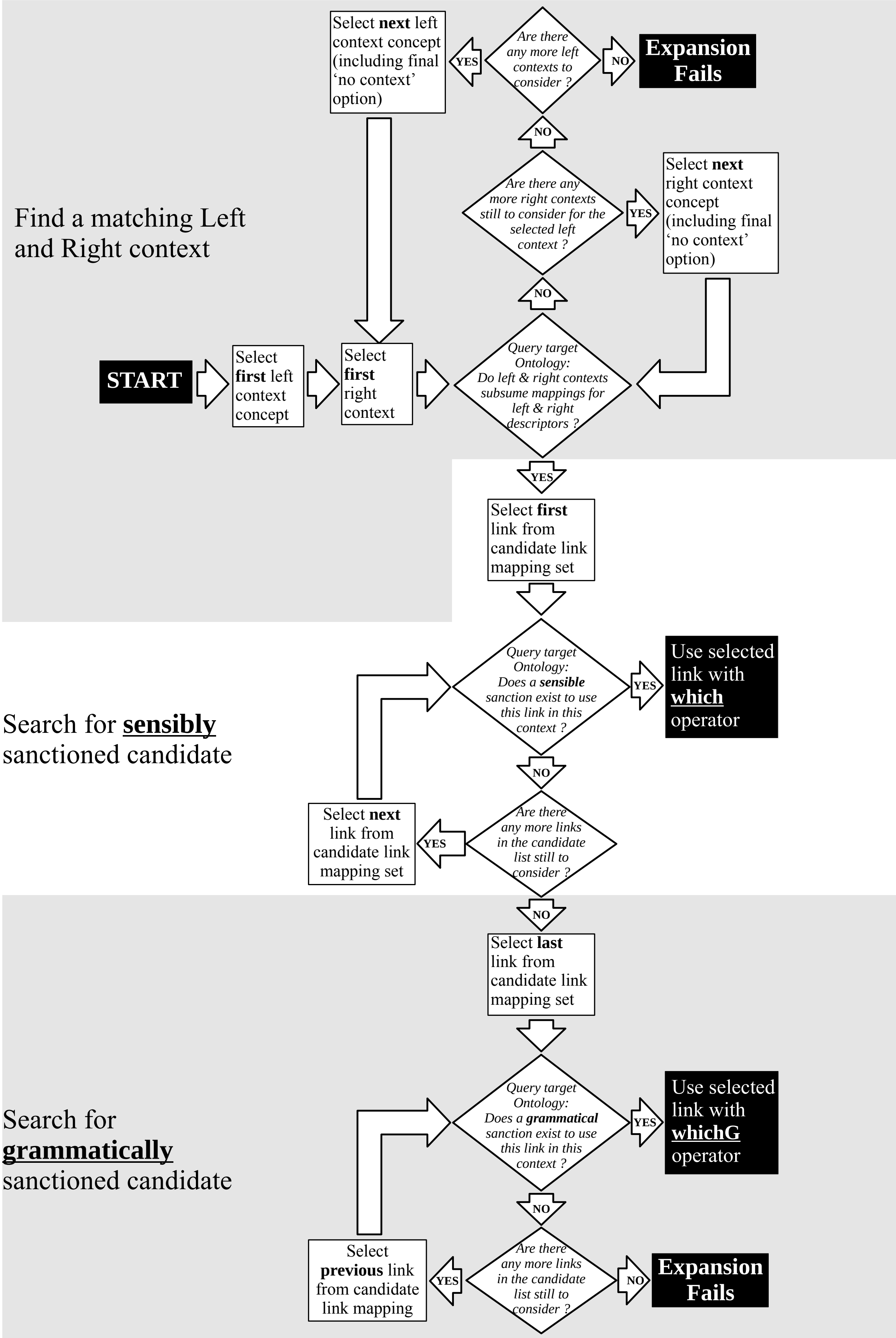



**Figure 64: Flow chart of link mapping selection algorithm**

#### 9.3.2.4 Link Mappings: clarification of algorithm

From the preceding section and from Figure 64 it can be seen that, once a matching left-right context pair has been identified, processing the candidate link mapping set for that context pair occurs in two phases – first seeking a candidate permitted by a sensible sanction (the metamodel), then (if the first phase fails) one permitted by a grammatical sanction (the meta-metamodel). Although both phases superficially appear to use one common candidate link mapping list, running first from top to bottom and subsequently from bottom to top, this view is perhaps misleading. Although the candidate mappings are presented and maintained as a single ordered list, they could be more properly treated as two separate lists, one for each phase of processing.

A candidate mapping list is typically constructed such that the bottom candidate in the list (to be considered last in phase one) could never be selected in phase one, while *only* the bottom candidate (and therefore first to be considered) could ever be selected in the second phase. This behaviour occurs because the bottom candidate is usually the common hierarchical ancestor of all those candidates above it. For example, in context 2 of HAS_LATERALITY (above) the final OpenGALEN ontology link [*hasPositionalSelector*] is the common ancestor of both preceding links in the list.

Given this arrangement, during the first phase of the algorithm the final ancestor candidates could never be chosen: if the bottom link were sensibly sanctioned, then one or all of the preceding candidates descended from it and already considered in phase one would also be sensibly sanctioned by inheritance, and therefore one of them would have been selected before the bottom (ancestor) link candidate was ever considered. Similarly, for phase two when the list is read in reverse in search of a grammatically sanctioned candidate, a similar situation applies: if a grammatical sanction exists, it will be found to hold for one of the ancestor links at the bottom of the list, and considered first, before any of the more specific descendent links above it are ever considered.

The reason for this arrangement lies in the fact that, in the OpenGALEN ontology, expressions that can only be composed under grammatical sanctioning are most commonly abstract expressions, usually formed using relatively abstract concept categories and similarly abstract semantic links between them. To illustrate this, Figure 65 sets out the different ways that the authoring ontology link IS_PART_OF should be expanded for five different expressions.

| ORIGINAL EXPRESSION | REWRITTEN EXPRESSION IN OPENGALEN ONTOLOGY |
|---|---|
| myocardium IS_PART_OF ventricle | Myocardium which *isSpecificLayerOf* HeartVentricle |
| lobe IS_PART_OF lung | Lobe which *isSpecificSolidDivisionOf* Lung |
| patella IS_PART_OF knee joint | Patella which *isStructuralComponentOf* KneeJoint |
| skin IS_PART_OF arm | SkinCovering which *isSurfaceDivsionOf* Arm |
| anatomy IS_PART_OF anatomy | BodyStructure whichG *IsDivisionOf* BodyStructure |

**Figure 65: Different expansions of IS_PART_OF link**

The first four rewrites listed are each sanctioned at the sensible level, but using different and specific flavours of partitive link, in accordance with the *Open*GALEN ontology partonomy model outlined in 3.9.3.1. The rewrite of the final expression, however, is only sanctioned at the grammatical level and is written using [*IsDivisionOf*], the common ancestor of all the other partitive semantic links. A significant effect of, and also the main driver for, using [*IsDivisionOf*] in this way is to ensure that, when automatically classified within the *Open*GALEN ontology, the rewritten final expression subsumes the other four.

#### 9.3.2.5 Link Mappings: inference of cardinality

In the same way that the ordering of the candidate link mapping set determines how semantically specific the selected link mapping will be, so the ordering also determines the cardinality. Figure 66 shows an extract of the candidate link mappings for the authoring ontology link ACTS_ON. The first three candidate mappings in the list are different cardinality variants of the same flavour of semantic relationship. Which candidate mapping is selected depends, as before, on the nature of the CRM metamodel.

In addition to allowing that processes may act on structures, the CRM Metamodel further expresses the constraint that most processes may act on *one and only one* structure at a time. Within GRAIL the only means to express these twin constraints in the CRM metamodel is by means of a single sensible sanction using the semantic link [*actsSpecificallyOn*]:

Process sensiblly actsSpecificallyOn Structure.

…where [*actsSpecificallyOn*] has many-One cardinality. However, the CRM Metamodel allows a small number of processes to act on more than one thing at a time – for example [Ventilation] can act on several discrete components of the respiratory tract at the same time. In these cases the cardinality constraint previously described is overridden by the addition of a more specific sensible sanction, such as:

Ventilation sensibly actsMultiplyOn RespiratoryTractComponent

…where [*actsMultiplyOn*] has many-many cardinality. It may be seen, therefore, how these two CRM Metamodel statements have the effect that ACTS_ON will be expanded to [*actsSpecificallyOn*] and its associated many-One cardinality for all contexts between [Process] and a [Structure], except [Ventilation] of some [RespiratoryTractComponent].

| AUTHORING ONTOLOGY LINK | CONTEXT | | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY |
|---|---|---|---|
| | LEFT | RIGHT | |
| ACTS_ON | any | any | *actsMultiplyOn*<br>*actsSpecificallyOn*<br>*actsOn*<br>*LocativeAttribute* |

**Figure 66: Extract of candidate link mappings for ACTS_ON**

## *9.4 Semantic Expansion*

Syntactic expansion and semantic substitution are not in themselves sufficient to achieve a satisfactory rewrite. The issue of *Schema Confusion* still remains: if the goal is to rewrite intermediate representation expressions such that the derived expressions can be integrated and classified coherently with other concepts already in the target ontology, then all rewritten expressions must comply with the same ontological style and schemas in that ontology. This includes complying with the various ontological schemas detailed in Chapter 3, such as for feature-states, negation, partonomy or processes as well as the more complex schema such as for pain or cancer.

Although the intermediate representation expressions are guided at author time towards a particular ontological style by means of the authoring ontology constraints detailed in section 8.4, these constraints are very much simpler than that in the target ontology. Even with these constraints, the intermediate representation still permits variant semantic patterns and schemas that would not be permitted in, and are not compliant with, the target *Open*GALEN ontology.

There is, therefore, a requirement for some *post hoc* mechanism to coerce expressions to the required semantic style and schema. Known collectively as 'semantic normalisation', two separate techniques can be identified:

- Semantic expansion for canonisation
- Semantic suppression of linguistic distinctions

### 9.4.1 Semantic Normalisation for Canonisation

The intermediate representation allows for expressions with more relaxed ontological style and less precise semantics. A consequence of this is that the authoring environment does not allow the model author to meet the stated requirements (as described in section 3.10.1) for canonisation to work correctly: the authoring ontology contains primitive descriptors mapped to composed concepts in the *Open*GALEN ontology, as well as decomposable descriptors that can be composed in terms of other descriptors and links *within* the authoring ontology. However, such composed concepts are not readily identifiable to the author, and their definitions are not accessible. Finally, since an explicit goal of the intermediate representation is to free casual authors from the need to know the schema that should be complied with, they are not in a position where they might understand either when or how parts of expressions should be redundantly re-expressed.

Thus, if different intermediate representation expressions with equivalent semantics are to be successfully expanded to the target ontology and then reduced to a common canonical form, an important goal of semantic normalisation must be to manage these issues on behalf of the author. The solution adopted to achieve this goal is to extend the link mapping mechanism to insert the elements that need redundant re-expression, as illustrated in the following example:

Intermediate representation authors might choose to say that a condition (such as inflammation) is caused directly by a microorganism, or alternatively by an infection at a particular place and by a microorganism. Within the *Open*GALEN ontology, however, the convention is that only infection lesions can be caused by microorganisms. Infection lesions may of course have secondary effects, but the association between a secondary effect and the original microorganism should always be through the intermediate concept of an infection lesion.

Figure 67 demonstrates this point: the rewrite for the authoring ontology link CAUSED_BY is variable, depending on its right context.

| ORIGINAL EXPRESSIONS | SCHEMA-COMPLIANT RE-WRITTEN EXPRESSION |
|---|---|
| inflammation<br>CAUSED_BY virus | Inflammation which *isConsequenceOf*<br>(InfectionLesion whichG *isConsequenceOf* Virus) |
| inflammation<br>CAUSED_BY bacterial vaginosis | Inflammation which *isConsequenceOf*<br>(InfectionLesion which <<br>*hasLocation* Vagina<br>*isConsequenceOf* GardnerellaVaginalis>) |

**Figure 67: Two different expansions of CAUSED_BY**

In the first example, where the right context is a kind of microorganism, CAUSED_BY is mapped into a target ontology chain:

*isConsequenceOf* InfectionLesion *isConsequenceOf*

In the second example, the right context (bacterial vaginosis) maps to a pathological lesion whose full compositional definition includes both a causative organism and an anatomical location. In the second example, therefore, CAUSED_BY has a different mapping to a single target ontology link:

*[isConsequenceOf].*

Figure 68 is an extract of the link mappings for the authoring ontology link CAUSED_BY, showing those elements that ensure that the intended rewrites of Figure 67 are achieved in the general case: Context #1 traps those situations where a composition involves an infection and an organism, but where these are being composed in accordance with the target ontology schema. Context #2 traps and normalises those situations where the target ontology schema is not being followed (something, other than an infection lesion, is described as caused by a microorganism). Context #3 applies to the remainder of uses of the link.

| AUTHORING ONTOLOGY LINK | 'LEFT' CONTEXT | 'RIGHT' CONTEXT | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY |
|---|---|---|---|
| CAUSED_BY (Context #1) | InfectionLesion | MicroOrganism | *IsConsequenceOf* |
| CAUSED_BY (Context #2) | Any | MicroOrganism | *isConsequenceOf* InfectionLesion *isConsequenceOf* |
| CAUSED_BY (Context #3) | any | Any | *IsConsequenceOf* |

**Figure 68: Extract of link mappings for CAUSED_BY**

### 9.4.2 Semantic Normalisation: Cancer Schema

A further example of *post hoc* semantic normalisation can be seen for the complex schema for cancer described in section 3.9.3.3. This schema (Figure 69) holds that the primary site can not be directly associated with the notion of the tumour: the schema requires that a tumour *arises* in a cell which is *partOf* the tissue *of* an organ. Further, the process-structure dual between tumour(lesion) and neoplastic(process) is also enforced:

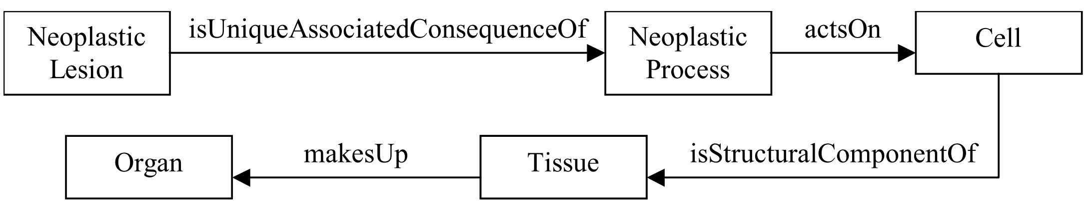



**Figure 69: Diagram of abridged target ontology schema for cancer**

The users, however, more naturally tend to model that a tumour can be *locatedIn* any one of these concepts as shown in Figure 70:

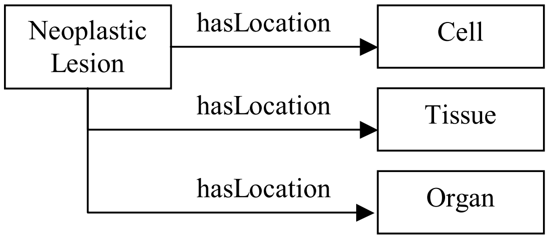



**Figure 70: Diagram of typical naive schema for cancer**

…and hence will write, in intermediate representation, a variety of representations such as those shown in Figure 71:

| PHRASE | CANDIDATE COMPOSITION |
|---|---|
| 'Adenocarcinoma' | malignancy HAS_LOCATION adenocyte |
| 'Bone cancer' | neoplasm HAS_LOCATION bone tissue |
| 'Lung cancer' | neoplasm HAS_LOCATION lung |

**Figure 71: Possible naive expressions for cancer using HAS_LOCATION link**

Normalising such expressions to the preferred target ontology schema for cancer can be achieved through the link mappings for the link HAS_LOCATION, an extract of which is shown in Figure 72. The first three contexts test for each of the three potential user patterns, and (if selected) their candidate link mapping acts to insert the required portion to comply with the preferred schema:

| 'LEFT' CONTEXT | 'RIGHT' CONTEXT | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY |
|---|---|---|
| NeoplasticLesion | Cell | *hasUniqueAssociatedProcess* NeoplasticProcess *actsSpecificallyOn* |
| NeoplasticLesion | Tissue | *hasUniqueAssociatedProcess* NeoplasticProcess *actsSpecificallyOn* Cell *IsDivisionOf* |
| NeoplasticLesion | BodyStructure | *hasUniqueAssociatedProcess* NeoplasticProcess *actsSpecificallyOn* Cell *isStructuralComponentOf* Tissue *makesUp* |
| NeoplasticLesion | any | *hasUniqueAssociatedProcess* NeoplasticProcess *actsSpecificallyOn*<br>*hasUniqueAssociatedProcess* NeoplasticProcess *LocativeAttribute* |

**Figure 72: Extract of link mappings for HAS_LOCATION**

### 9.4.3 Semantic Normalisation: HAS_LOCATION

Although a link named [*hasLocation*] exists in the target ontology, it should more appropriately have been renamed [*has-locus*] since it is reserved exclusively for linking pathological lesions to body sites and is not a generic locator in the traditional spatio-temporal reasoning sense. The similarly named authoring ontology link HAS_LOCATION is by contrast used by intermediate representation authors with a much broader semantic interpretation, covering 46 specific left-right context pairs including between:

- pathological features or states (e.g. large size, or a pathological increase) and physical structures
- pain and body sites

- prosthetic devices and the organs they replace
- processes (such as spastic contraction) and organs (such as muscles)

A selection of link mappings for HAS_LOCATION are shown in Figure 73 to demonstrate how these different uses are trapped and normalised.

| 'LEFT' CONTEXT | 'RIGHT' CONTEXT | CANDIDATE LINK MAPPING SET IN OPENGALEN ONTOLOGY |
|---|---|---|
| Pain | BodyStructure | *actsOn* ExternalStimulusSensation<br>*actsOn* PainSignal<br>*isConsequenceOf* NerveConduction<br>*hasUniqueAssociatedDisplacement*<br>Displacement *isDisplacementFrom* |
| Pain | BodySystem | *actsOn* ExternalStimulusSensation<br>*actsOn* PainSignal<br>*isConsequenceOf* NerveConduction<br>*hasUniqueAssociatedDisplacement*<br>Displacement *isDisplacementFrom*<br>BodyStructure *IsDivisionOf* |
| Pain | Tissue | *actsOn* ExternalStimulusSensation<br>*actsOn* PainSignal<br>*isConsequenceOf* NerveConduction<br>*hasUniqueAssociatedDisplacement*<br>Displacement *isDisplacementFrom*<br>BodyStructure *isMadeOf* |
| Pain | any | *no mapping* |
| Prosthesis | any | *hasFunction* GeneralisedProcess<br>*isFunctionOf* |
| Spasm | TubularBodyStructure | *actsSpecificallyOn* Diameter *isDiameterOf* |
| Spasm | any | *Involves* |
| DimensionChanging | TubularBodyStructure | *actsSpecificallyOn* Diameter *isDiameterOf* |
| DimensionChanging | LinearBodyStructure | *actsSpecificallyOn* Length *isLengthOf* |
| DimensionChanging | MuscleTissueStructure | *IsFunctionOf* Muscle<br>*isStructuralComponentOf* |
| DimensionChanging | Abdomen | *actsSpecificallyOn* Diameter *isDiameterOf* |
| DimensionChanging | LaminarPhysicalStructure | *actsSpecificallyOn* VerticalDepth<br>*isVerticalDepthOf* |
| DimensionChanging | SkinCovering | *actsSpecificallyOn* VerticalDepth<br>*isVerticalDepthOf* |
| DimensionChanging | SolidBodyStructure | *actsSpecificallyOn* Volume *isVolumeOf* |
| DimensionChanging | any | *no mapping* |
| Process | any | *actsMultiplyOn*<br>*actsSpecificallyOn*<br>*actsOn*<br>*isSpecificFunctionOf*<br>*isFunctionOf*<br>*LocativeAttribute* |
| Any | Any | *hasMultipleLocation*<br>*hasSpecificLocation*<br>*hasLocation*<br>*LocativeAttribute* |

**Figure 73: Larger extract of contexts and link mappings for HAS_LOCATION**

#### 9.4.4 Semantic Normalisation: Link Mappings as Schema Constraints

Rows four through seven of Figure 73 show that the contexts for normalising HAS_LOCATION to the Pain schema of section 3.9.3.4 define candidate link mappings only when the right context is either a [BodyStructure], a [BodySystem], or a [Tissue]. In the event that any other type of concept is to the right of HAS_LOCATION, a ‘no mapping’ state is explicitly declared (row seven).

This construct has the effect of preventing the final default ‘any-any’ context for this link being used in the specific case where the left context is a kind of pain, but where no right context match were to be found. This behaviour has much in common with that produced by the target ontology schema constraints (GRAIL sanctions) of section 3.11: intermediate representation compositions that breach permitted semantic patterns are rejected *post hoc*.

A further advantage of explicitly defining those contexts that have no mapping (and where, therefore, expansion will fail by design rather than by default) is that it aids the detection of existing links being put to new semantic uses: if new uses were always routinely expanded by the default ‘any-any’ context, they could go unnoticed and incorrectly expanded for an indefinite period. Overriding this behaviour, such that the link has no default behaviour and can only be expanded when encountered in previously defined context, means new uses are detected as soon as they are devised.

# 10 Implementation

This chapter describes the workflows followed by the author both in designing, refining and implementing the methodology described in Chapters 5-9 and in the subsequent experiments described in Chapter 11. The author of this thesis worked closely with software engineers to iteratively specify, test and debug what became a highly sophisticated software environment[6] to support the implementation of the methodology. Where relevant, the workflow descriptions below reference the major components of that environment, with screeshots. The reader is also referred to [Rogers 2001].

## *10.1 Overview of workflow*

Four major tasks, and supporting tools, can be identified within the methodology for hiding ontological complexity described in this thesis:

1. Authoring dissections using the SPET software
2. Extending the common reference model using OpenKnoME
3. Linking dissections to the common reference model using TIGGER
4. Quality assurance of the whole using TIGGER and OpenKnoME

The first of these tasks was performed by the geographically distributed dissection authors. The remaining three tasks, constituting a central support activity, were performed by the author of this thesis. Although all software tools were constructed by other researchers, the author was closely involved in the specification of both the overall architecture and the individual tools and their data exchange formats, and in their subsequent testing.

The four tasks, and their relationship to each other, can be further subdivided into discrete activities and supporting tools, summarised diagrammatically in Figure 74. The following sections outline further the typical experimental workflow followed in implementing the methodology, accompanied by screenshots showing the progress of authoring and processing a single dissection: “Operation on papillary muscle”.

## *10.2 Central collection and dissemination of dissection library*

20,782 surgical procedure rubrics were authored during the GALEN-IN-USE project by the distributed dissection authors using the SPET software tool (Figure 75). These dissections were normally emailed to the author as they were written, so that a regular cumulative release of the Dissection Library (Figure 76) could be distributed to all authors.

[6] TIGGER and OpenKnoME are available as open source software from www.topthing.com. The SPET tool is a component of a software suite now called the Classification Workbench, from www.kermanog.com.

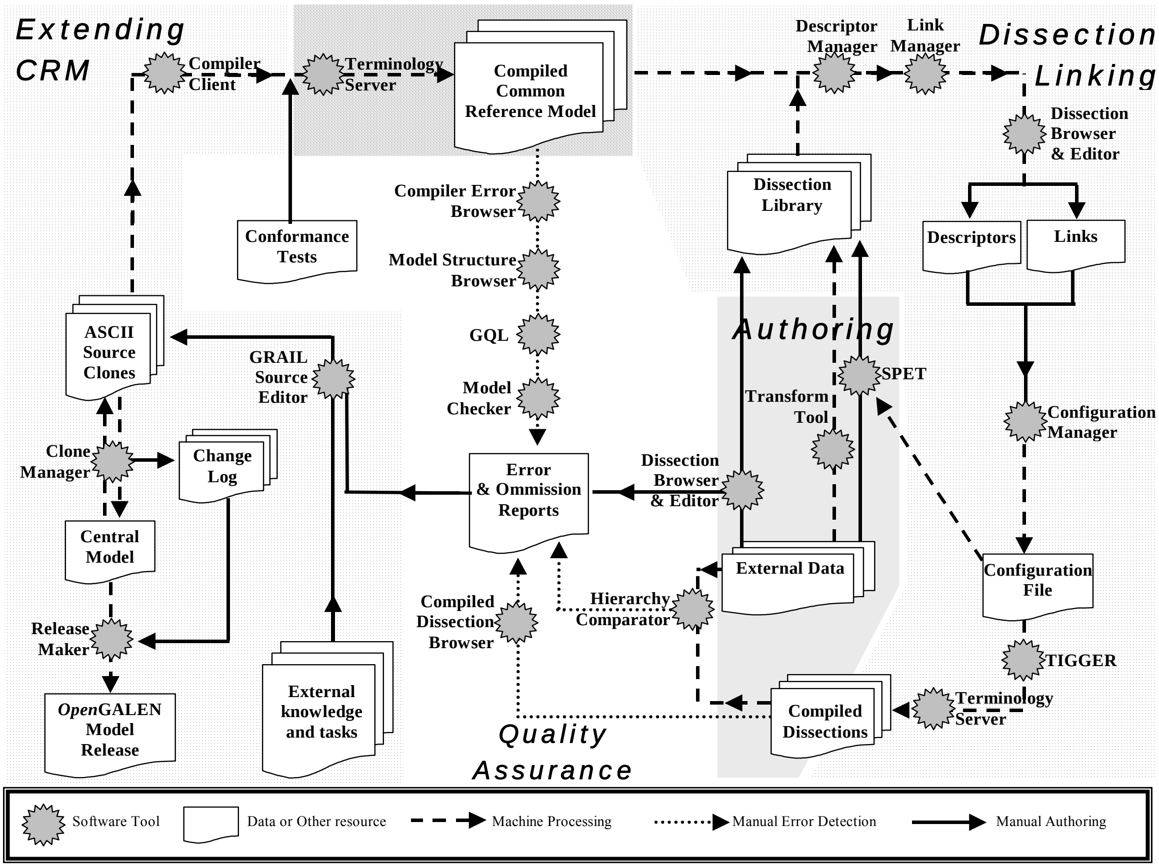



**Figure 74: Diagrammatic representation of tasks and workflows in this thesis Shaded areas indicate four major tasks**

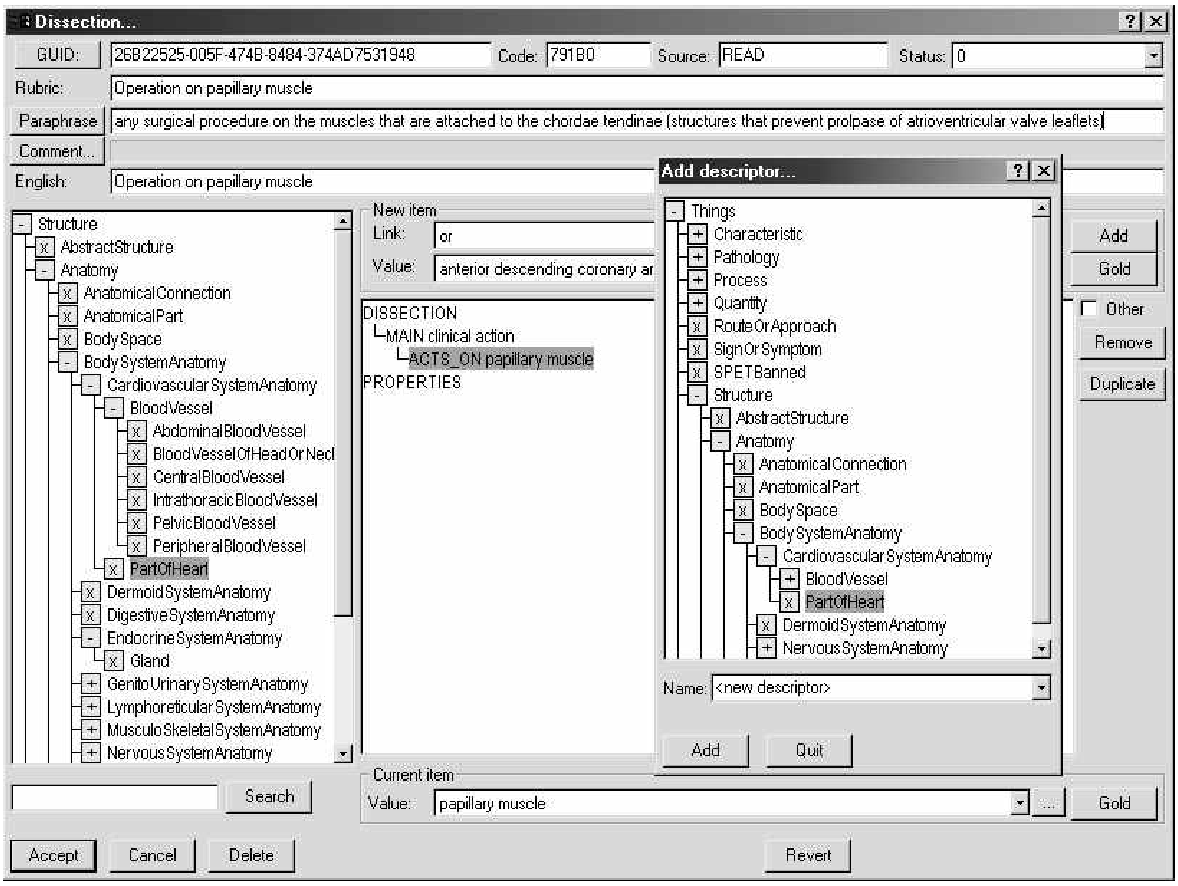



**Figure 75: SPET dissection authoring interface**

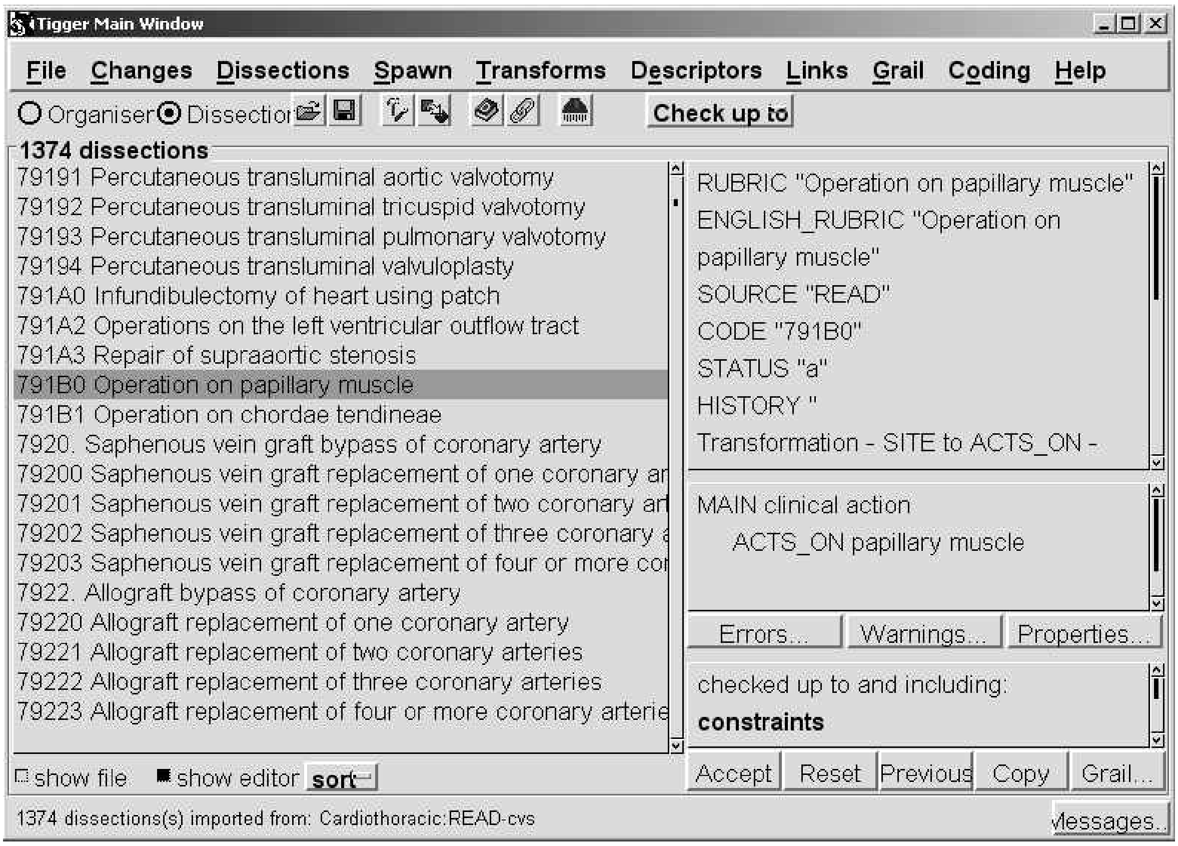



**Figure 76: TIGGER Dissection Browser (Dissection library at left, highlighted dissection shown right)**

## *10.3 Extraction and dissemination of term (descriptor) and link lists*

As described in this thesis, the final methodology allowed dissection authors to write dissections using terms drawn from an *ad hoc* authoring ontology, rather than using the *pre hoc* term list of the target ontology. Figure 75 shows the dialogue within the SPET dissection authoring interface whereby dissection authors may enter a new descriptor to their local authoring ontology, and suggest an appropriate descriptor category for it.

The **Descriptor and Link Manager** (Figure 77 and Figure 78) components of **TIGGER** supported the iterative extraction, collation, curation and dissemination of the cumulative list of all authoring ontology descriptors and links used by dissection authors as they worked their way through authoring representations of surgical procedure rubrics. Curation tasks included identifying lexical variant and synonym descriptors used by different authors, declaration of preferred descriptors, and assigning new descriptors to high level categories within a simple navigational classification (see 7.3 and Appendix One).

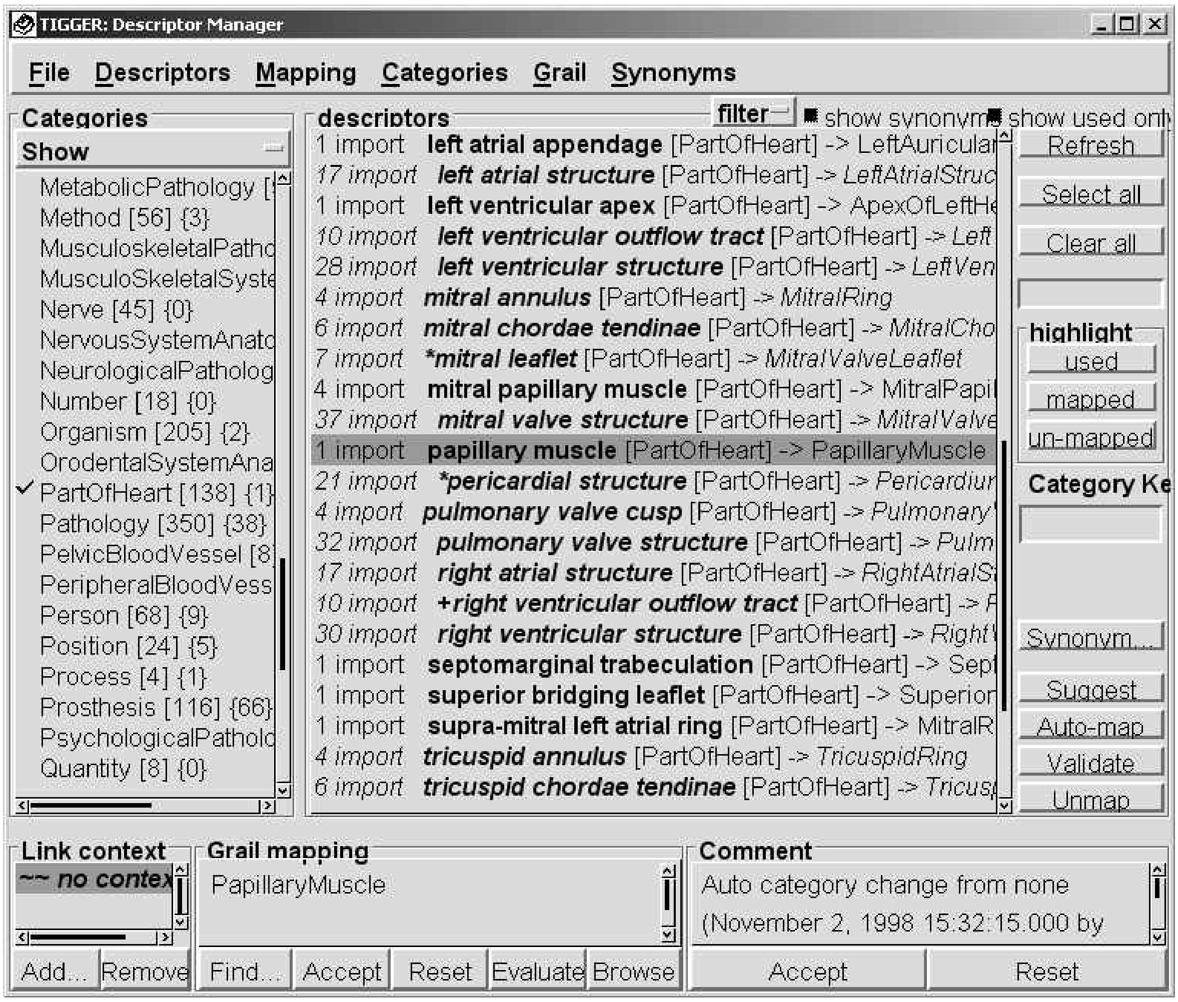



**Figure 77: TIGGER Descriptor Manager**

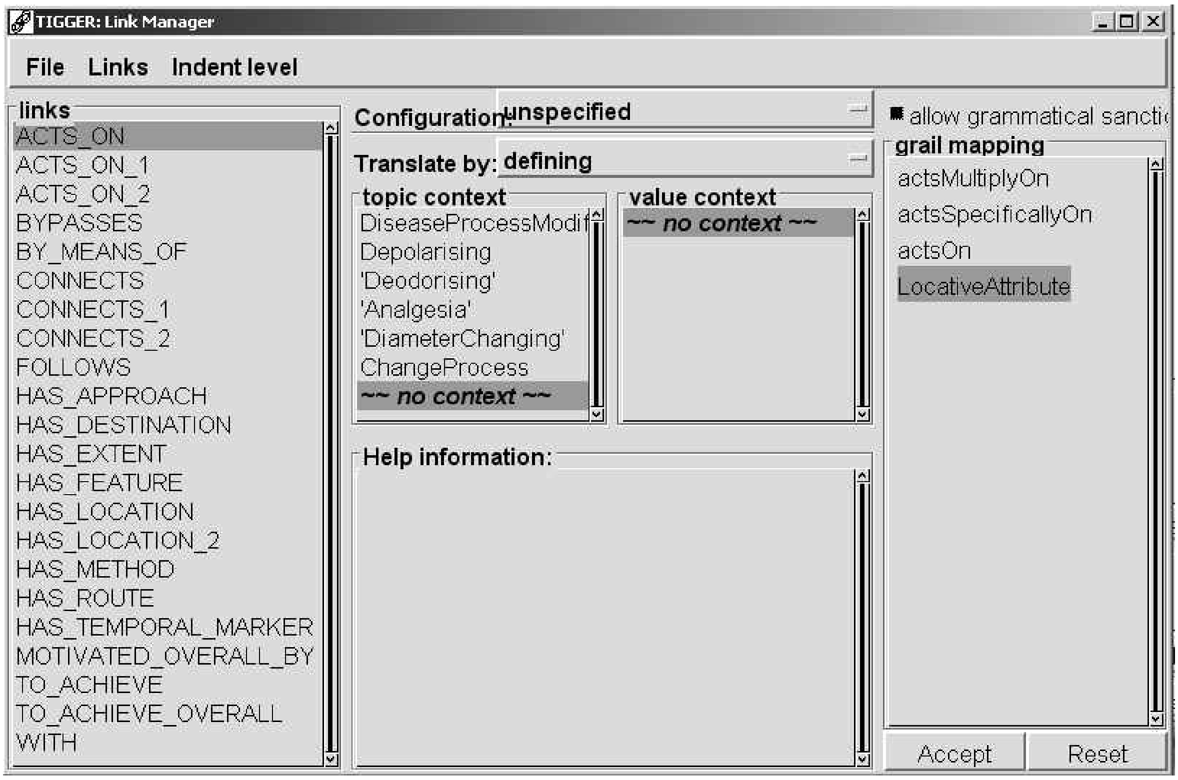



**Figure 78: TIGGER Link Manager**

### *10.4 Mapping descriptors and links*

In order to re-express dissections within a **Compiled Common Reference Model**, the author of this thesis examined each new descriptor and link as it was encountered and declared mappings between them and the common reference model (Figure 77 and Figure 78). To assist this mapping process, the **Dissection Browser and Editor** component of **TIGGER** (Figure 76) allowed the author to search the growing dissection library for all dissections, written by any author, using a particular descriptors or link. This allowed the intended meaning of a descriptor or link to be checked for any one author, and the scope of intended meanings to be determined across different authors (see 7.5.1).

However, because the authoring ontology was decoupled from the development of the common reference model, dissection authors could write dissections covering types of procedures or anatomical structures not yet modelled in the common reference model. This decoupling had the effect that, in the early phases of the project, new descriptors and links were encountered that could not yet be mapped to the common reference model.

The cumulative set of encountered descriptors and links, together with their categorisation, mapping and synonym information, was periodically released to all authors as a **Configuration File**, which could be imported into the **SPET** tool to provide a growing curated authoring vocabulary.

### *10.5 Analysis of dissections for common or variant patterns*

The **Dissection Browser and Editor** component of **TIGGER** (Figure 76) allowed detailed inspection and manipulation of the dissection library corpus. In addition to allowing sets of dissection using a particular link or descriptor to be identified and browser, the **Transformation Tool** also allowed the author to search for all instances of a dissection that used a particular semantic pattern.

Where such inspections revealed that different authors had used the same descriptor or link for very different intended meanings, or were using different semantic patterns for the same meaning, global descriptors replace or graph transformations were performed on the **Dissection Library** to harmonise those differences.

### *10.6 Studying classifications of compiled dissections*

Both the **TIGGER** and **SPET** tools included the functionality to import all, or part, of the **Dissection Library** and, using the mapping information contained in the **Configuration File**, re-express each dissection as a candidate GRAIL expansion (Figure 79) for integration within a **Compiled Common Reference Model**.

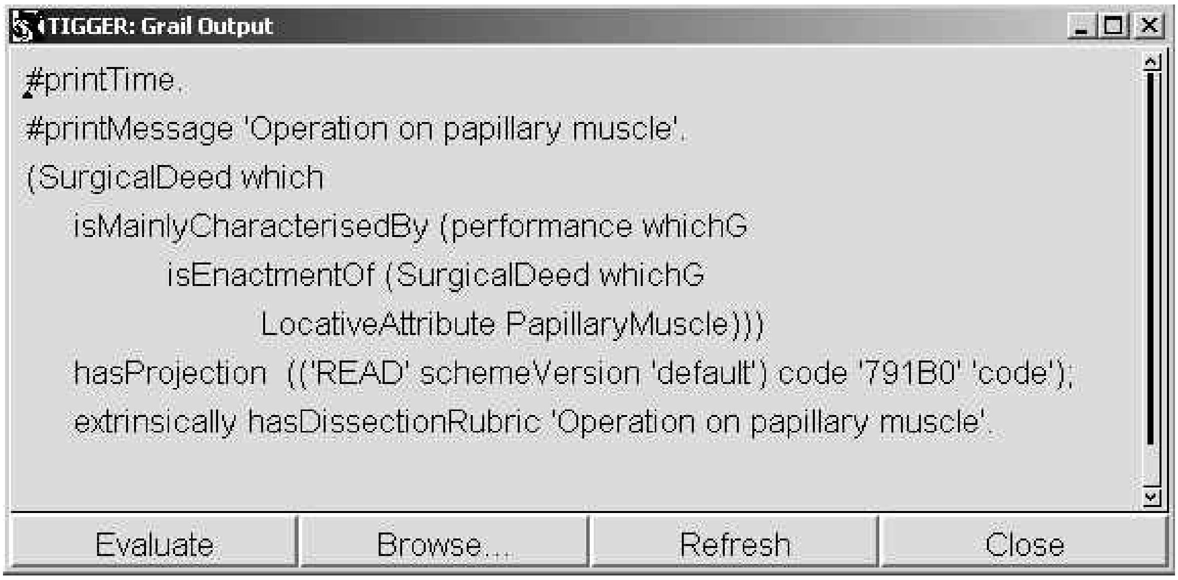



**Figure 79: Candidate GRAIL expansion (generated by clicking on GRAIL.. button in Figure 76)**

A **Terminology Server** integrated all candidate GRAIL expressions into **Compiled Dissections,** classified as a polyhierarchy. This may be inspected visually for obvious errors, using the **Compiled Dissection Browser (**Figure 80**).**

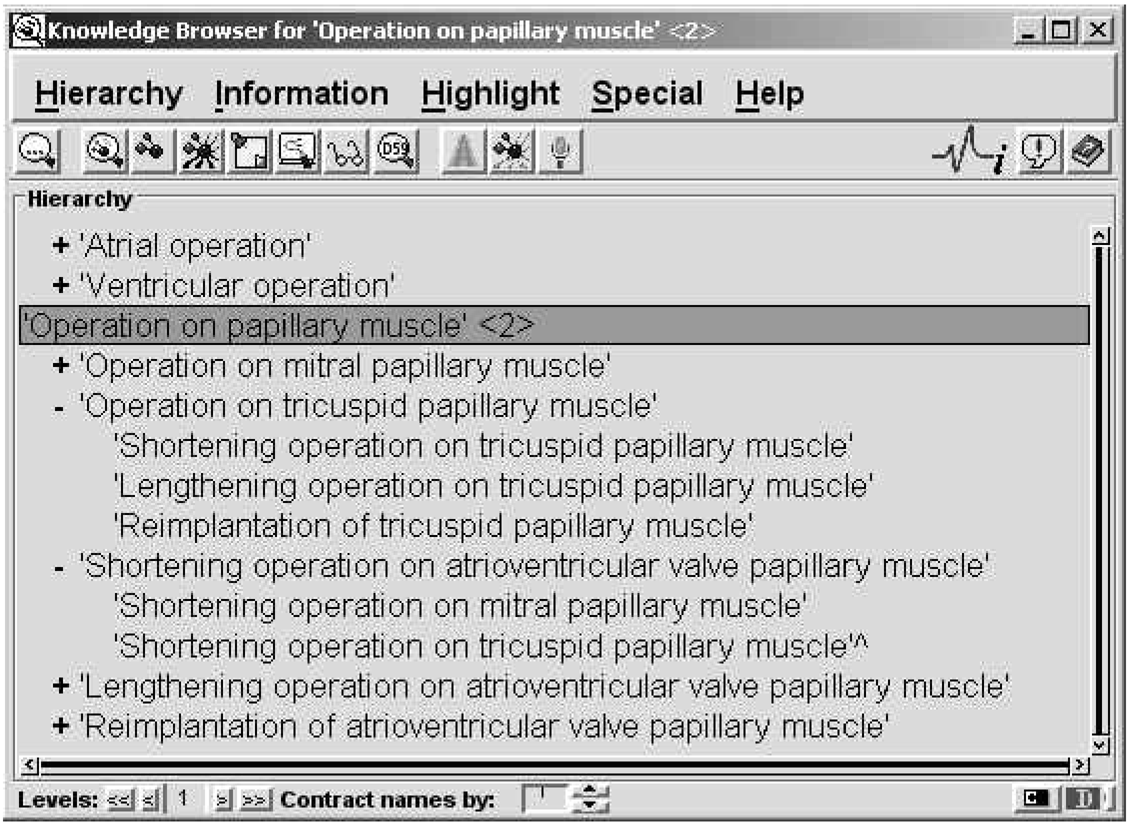



**Figure 80: TIGGER Compiled Dissection Browser (result of compiling all dissections in Figure 76)**

Alternatively the **Hierarchy Comparator (**Figure 81**)** can automatically compare the hierarchy inferred by the **Terminology Server** with other suggested hierarchies for the same classification, such as their classification in an original medical terminology scheme. The Hierarchy Comparator examines all parent-child relationships in each hierarchy and then assigns them to one of five sets

- unique to the inferred (GRAIL) hierarchy
- unique to the external (other) hierarchy

- a parent-child relationship from one hierarchy is precisely inverted in the other (contradictory)
- both concepts in a parent-child relationship from the external hierarchy map to a single concept in the inferred (GRAIL) hierarchy
- represented in both hierarchies (common)

Each member of each set may then be examined by hand and further categorised according to user defined result categories (e.g. valid or invalid parent-child relationship)

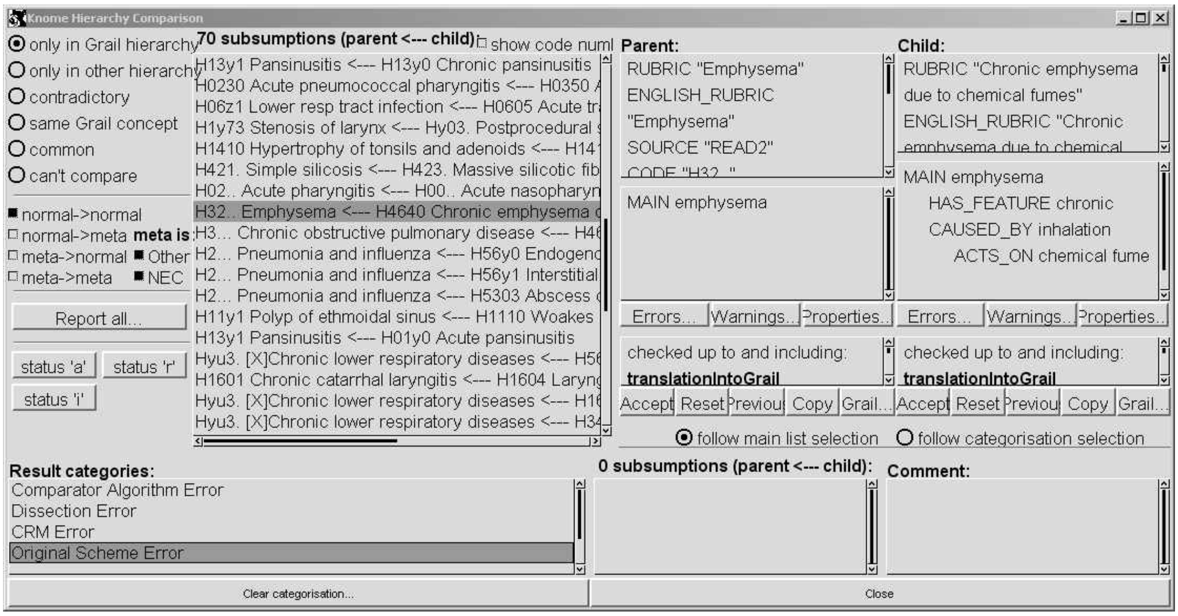



**Figure 81: Hierarchy Comparator Tool**

## *10.7 Authoring new common reference model content*

One possible output of the descriptor and link mapping step, or of studying compiled dissections, is identification of errors in the Common Reference Model. Commonly these will be errors of omission, typically where a new descriptor can not yet be mapped because no equivalent entity exists in the Common Reference Model. However, errors of commission may also be identified, for example where a compiled dissection misclassifies because the Common Reference Model contains a false assertion. Other mechanisms for identifying errors in the Common Reference Model include independently motivated direct inspection and intermittently applied quality assurance steps such as **Terminology Server** conformance testing, schema consistency checks using GQL, and post-compilation model integrity checks.

The **KnoME** component of ***Open*KnoME** presents a rich environment for managing numerous (more than 1400) **ASCII Source Files** for the **Common Reference Model** (Figure 82 and Figure 83) including support through the **Clone Manager** for distributed and collaborative modelling; up to 4 expert ontologists have been working under the author's supervision at one time on the same sources. ***Open*KnoME** includes integrated support for error and change logging and management (Figure 84).

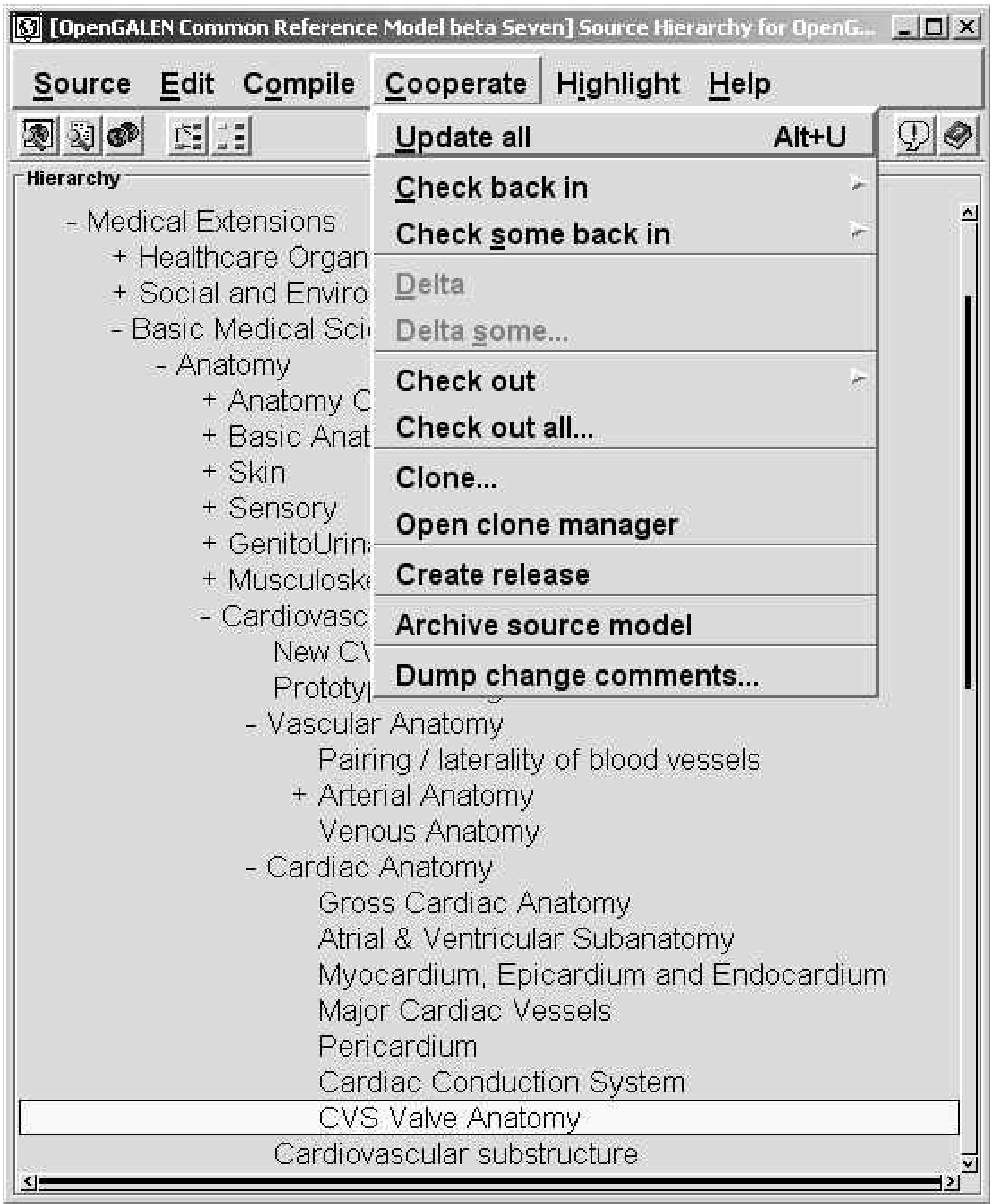



**Figure 82: GRAIL Source Manager, with menu of collaborative modelling support functions**

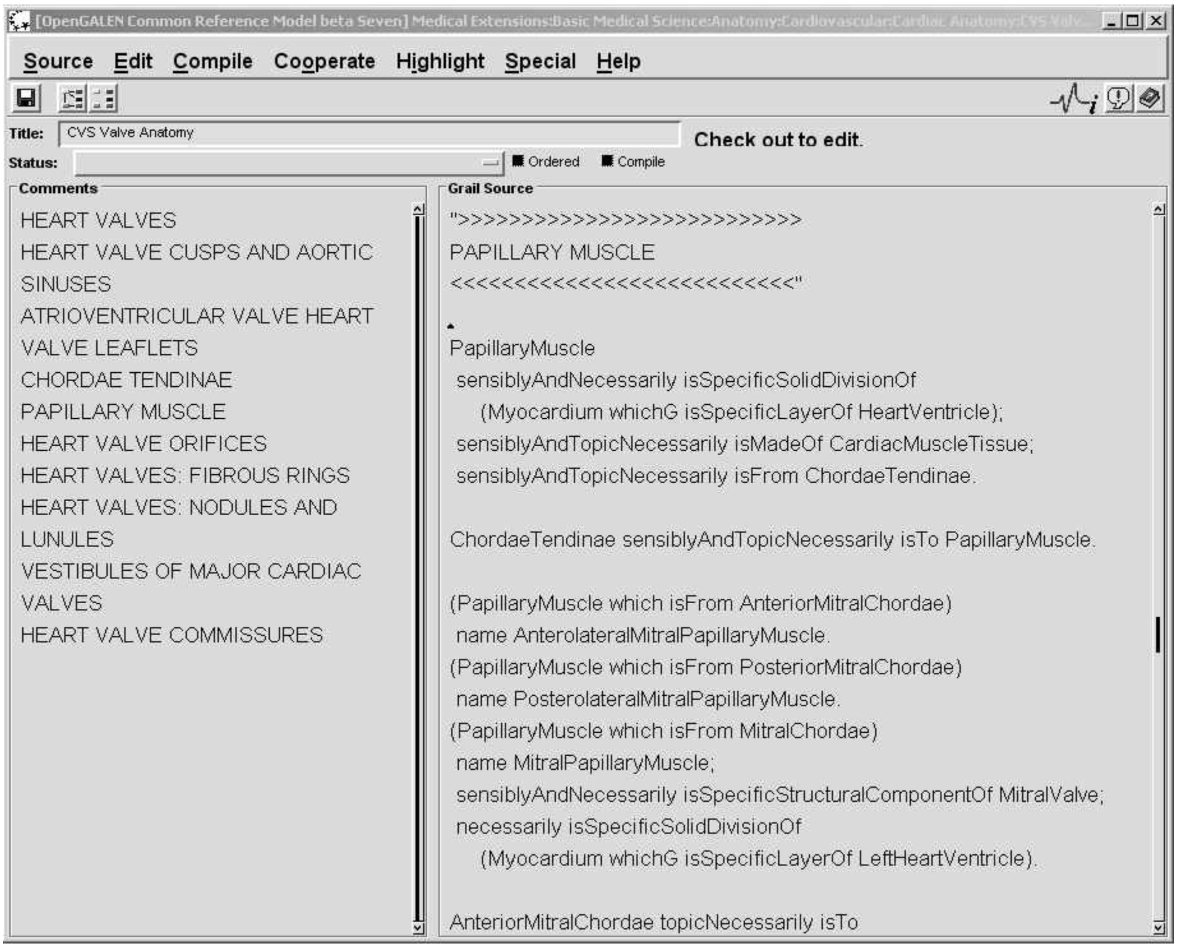



**Figure 83: GRAIL Source editor showing GRAIL code describing anatomy of papillary muscles**

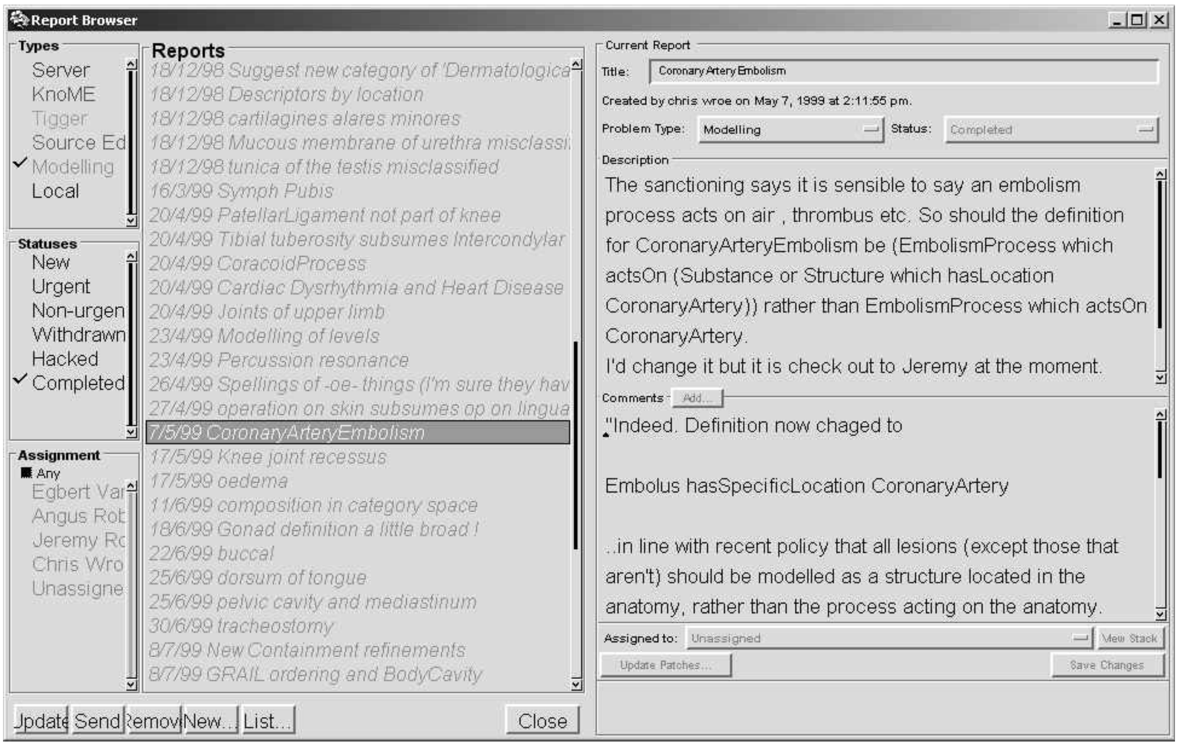



**Figure 84: OpenKnoME Error Manager showing reported CRM errors and responses**

# Part Three: Evaluation of Methodology

Part One of this thesis described why ontologies become complex to use

Part Two set out a methodology for hiding that complexity

Part Three describes a series of experiments to evaluate the methodology

# 11 Evaluation and Results

## *11.1 Introduction*

Part One of this thesis described the complexities that face the user of an ontology, categorised into domain, artefactual and cognitive complexity. The potential consequences when such complexity overwhelms the users were described in Chapter 4.

Part Two described how those same complexities were used to achieve a systematic and reversible simplification of the whole.

This chapter describes several experiments that were undertaken to evaluate whether the proposed simplification was successful. Also discussed are the likely resources required to implement the methodology, and some limitations of the methodology that were encountered.

**The experiments described here were necessarily experiments of opportunity conducted within the framework of the GALEN-IN-USE project, whose limited resources did not allow for more comprehensive evaluation. Appendix Four summarises a critical appraisal undertaken to consider the internal or external validity of results from these experiments, whilst each report below includes an exposition of the princple threats identified to the validity of the experiments.**

**Chapter 12 includes a discussion of further potential experiments, including those that might yield results with greater validity.**

The experiments provide information on four phenomena:

- Time to train new users and subsequent productivity
- Reproducibility of representations
- Semantic utility of modelling work
- Generalisability of methodology
- Resource requirements of methodology
- Known limitations of current methodology

In summary, the results of these experiments show that:

- The simplified representation was more accessible to users (than the native GRAIL representation and target CRM ontology): there was a significant reduction in the time to train and overall productivity was higher.
- Reproducibility of semantic expressions by different users is difficult to assess
- The syntactic, ontological and semantic expansion processes described in Part Two of this thesis produce the desired result: polyhierarchies were computed in the target ontology from original representations expressed with a third party authoring ontology. These were substantially similar to those created by hand for the same concepts. The computed polyhierarchies may be easier to debug and maintain, particularly for larger groupings of concepts.
- Representations authored using the <u>complete</u> methodology – simplified syntax, authoring ontology and the expansion processes – result in a hierarchy of considerable complexity and which superficially appears similar to that which would be expected. However, a more rigorous assessment of the quality of such classifications remains to be performed.

- The approach can be generalised, and extended, to cover other tasks and subdomains
- The central support cost is front-loaded. A single central expert is capable of supporting roughly 20 distributed model authors
- Some limitations in the current methodology would be improced by more sophisticated strategies for mapping descriptors and links

## *11.2 Time to train and productivity*

Construction of the *Open*GALEN Common Reference Model – the target ontology - was performed directly by experts in the domain (clinicians) who had subsequently also been trained in concept modelling techniques. This contrasts with other approaches where specialists in concept modelling (who know nothing of the specific domain) interview the domain experts (who correspondingly know nothing of modelling) and transcribe the knowledge elicited.

An empirical finding has been that only clinicians with a particular way of thinking about the world can be trained to also be successful model authors directly in the target ontology. Primarily, they must already possess a natural aptitude for classifying and segmenting the world, as this is not a skill or perception that can easily be taught to those who do not already have it. However, even with such an aptitude it still requires an apprenticeship of many months, working in the same room as an experienced model author, before new model authors appreciate and can manage all the complexities detailed in Chapter 3.

An important measure of the success of the approach presented in this thesis, therefore, was the extent to which the time to train or apprentice could be reduced. Could domain experts contribute usefully to the expanding knowledge base more rapidly? Could the requirement for extended co-working with an experienced model author be reduced?

**Method**

20 domain expert authors, geographically distributed between 8 different countries (France, UK, Netherlands, Greece, Italy, Spain, Germany, Sweden), were recruited during the GALEN-IN-USE project. Although some had prior experience developing or maintaining medical terminologies, most did not. None had prior concept modelling experience, and none were specifically selected for natural modelling aptitude.

The task of these recruits was to represent as dissections the semantic meaning embodied by the phrases (‘rubrics’) taken from several national classifications of surgical procedures (see section 1.7 and Figure 3). Only one such classification scheme was processed in its entirety during the project (CCAM, from France). However, substantially the whole of the surgical subdomain was covered, including significant extracts from six different surgical procedures classifications (READ, CCAM, ICPM, NCSP, ICD-9-CM, WCC) originally in seven original natural languages (English, French, Swedish, Dutch, German, Italian, Greek).

**Phase I: Prototyping** In the first year, authoring was performed by a subgroup of the 20 domain expert authors. Simple text editors only were used, coupled with a web-based syntax checker prior to central submission of work.

**Phase II: Deployment** After the first year a custom editing tool (SPET – the surgical procedure entry tool) was constructed. This tool was created to more effectively deliver the authoring ontology that was concurrently developed and refined using the TIGGER tool (see Chapter 10). A major initial goal of SPET was to reduce the high rate of syntax and typographical errors seen using text editors alone. The SPET additionally provided a mechanism to feed the cumulative authoring ontology descriptor list back to the authors as it developed.

**Results**

Phase I prototyping participants received no training, but were issued with a brief style guide. During this period, approximately 3000 representations of surgical procedure rubrics (‘dissections’) were authored.

For deployment, Phase II participants (all 20 participating domain experts) received an augmented copy of the style guide, the SPET software, and a single three-day training session. Over the following three years, 20,782 surgical procedure rubrics were represented as dissections in the intermediate representation using a shared, common authoring ontology that comprised only 4168 descriptors. The Phase II experiment provides clear confirmation that domain experts can become productive with very little training. Additionally, the need for a lengthy and closely supervised apprenticeship is removed: after the initial three-day training, authors were supervised entirely remotely.

### 11.2.1 Threats to Internal and External Validity

**Selection bias** –two different groups were compared for time to train: GRAIL apprentices were selected by extended personal interview for natural ontology engineering aptitude whilst intermediate representation authors were self-selected researchers who already had prior understanding and interest in classification issues but not necessarily in ontology engineering. Neither group was representative of other possible populations of clinical ontology users, wheither randomly selected or comprised of other kinds of intact users groups (e.g. practising clinicians)

The GRAIL apprentice group (selected for aptitude) was set the harder of the two tasks, such that any selection bias introduced is more likely to have caused an under- rather than overestimation of any difference in time to train.

From the size of the effect observed in the reported study, and from intuition, it seems likely that most other potential subjects populations would also find the Intermediate Representation easier to learn that GRAIL. However the size of the difference and the absolute length of time taken to train in either may be very different.

Similarly, although the productivity might be expected to be higher for any population using the Intermediate Representation, the size of the difference observed here may not necessarily be indicative of what would be achieved in other populations: it seems very likely that most populations would achieve zero or close to zero productivity using GRAIL, but general clinicians may find the Intermediate Representation more challenging than the participants in these experiments and, as a consequence, be less productive.

**Practice effects** - training for all dissection authors in the use of the intermediate representation followed an earlier attempt to train some of those authors to write directly in GRAIL. However, that earlier attempt was very short lived such that any practice effects are considered to have been minimal.

**Experimental morbidity** – all subjects completed the training phases. However, some withdrew from the productivity phase as a result of external factors (typically lack of time or completion of contract), such that individual productivity in the Intermediate Representation group varied from as few as 50 dissections to as many as several thousand. For this reason, productivity can only be measured for the group as a whole.

**Maturation** – those authors who participated for the longest period of time are likely to have become more proficient, both as a result of practise and as a result of improvements in the authoring software (SPET) and in the content of the authoring ontology.

**Selection-Maturation Interaction** - it is possible that the observed productivity of trained dissection authors as a group conceals a highly skewed distribution whereby a small number of adept and matured authors produced the majority of dissections. It is not known, for example, whether doubling the size of the authoring group would have doubled their overall productivity.

**Multiple Intervention Interaction** – the assessment of productivity was conducted following exposure of the study groups to a training programme. This observed productivity may not be generalisable to untrained user populations such as, for example, UK general clinicians who traditionally receive little or no detailed training in the proper use of clinical terminologies for clinical recording.

## 11.3 Reproducibility and semantic utility

The described reduction in training time, and overall productivity, would be of little value if the efforts of the authors were not usable or useful.

An initial small-scale experiment attempted to measure the extent of reproducibility when using the intermediate representation. Two later experiments investigated whether the approach produced semantically useful expressions.

### 11.3.1 Reproducibility

#### Method

During the Phase I prototyping (and using word processors to write their dissections), authors from four of the piloting centres were invited to author, in parallel, dissections for 8 centrally selected phrases listed in Figure 85:

*Other osteotomy for correcting position*

*Total excision of tibia and fibula*

*Right ventricular infundibulectomy*

*Laparoscopic transcystic biliary tract exploration with retrograde endoscopic sphincterotomy*

*Echography of the scrotum*

*Circumcision for phimosis, with or without frenuloplasty*

*Bypass from aorta to iliac artery*

*Other operations on infrarenal abdominal aorta and iliaca arteries and distal connections*

**Figure 85: Dissections authored in parallel to study reproducibility**

The 32 separate dissections resulting were collated, printed out and inspected by the author of this thesis.

#### Results

The four independent authors produced essentially the same result for only one of the eight rubrics, “Echography of the scrotum”, as shown in Figure 86:

| | |
|---|---|
| MAIN imaging<br>ACTS_ON scrotum<br>BY_MEANS_OF ultrasound equipment | MAIN imaging<br>ACTS_ON scrotum<br>BY_MEANS_OF echograph |
| MAIN imaging<br>ACTS_ON scrotum<br>BY_MEANS_OF ultrasound machine | MAIN investigating<br>ACTS_ON Anatomy: scrotum<br>BY_METHOD ultrasonography |

**Figure 86: Four dissections received for 'Echography of the scrotum'**

More diversity of representation was encountered than expected for the remaining seven rubrics. For example, the rubric "Circumcision for phimosis, with or without frenuloplasty" produced the four results shown in Figure 87.

For the most part, this representational variability was thought to be related to fundamental differences in what the eight original phrases inspired the authors to try and represent. For example, in the circumcision example, only two of the authors thought it was important to represent *why* the procedure was performed (those segments of graphs beginning with the link MOTIVATED_BY).

| | |
|---|---|
| MAIN excising<br>    ACTS_ON preputium<br>        HAS_PATHOLOGY phimosis | MAIN excising<br>    ACTS_ON prepuce_of penis<br>    MOTIVATED_BY curing<br>        ACTS_ON phimosis<br>WITH fashioning<br>    ACTS_ON penis_fraenum |
| MAIN circumcising<br>    ACTS_ON penis<br>    BY_TECHNIQUE removing<br>        ACTS_ON prepuce<br>    MOTIVATED_BY caring<br>        ACTS_ON phimosis | MAIN removing<br>    ACTS_ON Anatomy: prepuce |

**Figure 87: Four dissections received for 'Circumcision for phimosis, with or without frenuloplasty'**

#### 11.3.1.1 Threats to Internal and External Validity

**Confounding** - a fundamental weakness exists in the experimental design. It was wrongly assumed that identical natural language phrases equated to an identical input for each author in the study.

This assumption overlooks the fact that each author can only represent what they individually *think* the phrase means. Since each will apply individually different semantic interpretations to each phrase, what each thinks a phrase means will be different. The more ambiguous the input phrase, or the more complex the concept concerned, the greater the variability of interpretation encountered.

In order to bypass this individual semantic processing of the original phrases, it would be necessary to provide as the common input a complete and conceptually unambiguous representation of the meaning to be represented. This would essentially mean providing the desired output of the experiment as its input.

**Confounding/Maturation** - The reproducibility experiments were conducted early in the GALEN-IN-USE project, before much of the methodology and supporting authoring tools had been developed and before the participants in the experiment had had an opportunity to gain experience. Although the results of the experiment are thought to be internally valid with respect to the performance of untrained authors operating in an unconstrained and unguided authoring environment, it would be interesting to know whether reproducibility improved as a result of improved tooling for authors, and as a result of their own expertise acquired over time.

**Selection bias** – as previously discussed, the study group of Intermediate Representation authors was unrepresentative of other possible populations of clinical ontology users, wheither randomly selected or comprised of other kinds of intact users groups (e.g. practising clinicians). However, although an even more specialist subject population might have scored higher, it would be difficult for another population to achieve lower reproducibility than was observed here.

### 11.3.2 Semantic utility

#### 11.3.2.1 Large scale interdigitation

A total of 20,782 disssections of rubrics from surgical procedures were authored during the GALEN-IN-USE project by authors working in different sites, using the SPET tool and the shared cumulative authoring ontology. These dissections covered substantially the entire surgical subdomain, however some areas were covered more than once: different authors working on different coding schemes chose to process equivalent chapters. Thus, the cardiovascular surgery chapters from each of NCSP, CCAM, CTV3 and ICD were separately processed.

##### 11.3.2.1.1 Method

A total of 3617 dissections covering cardiovascular surgical procedure rubrics were collected centrally by the author of this thesis. 2999 of the 3617 dissections were successfully expanded to the target ontology using the methodology and tools described in Chapter 10. The resulting polyhierarchy was examined informally by a clinician using the Compiled Dissection Browser (see 10.6).

##### 11.3.2.1.2 Result

The result may best be described as an 'interdigitation' of the concepts represented in the rubrics from the original four, separate schemes. A fragment is presented in Appendix 3.

The computed hierarchy had a maximum of 12 levels between the top level concept 'cardiovascular procedure' and the most distant leaf concepts. On many occasions two different rubric dissections expand to the same target ontology concept.

A thorough validation of the inferred hierarchy was not possible within the resources available, although the cross validation of the inferred polyhierarchy for the 1374 CTV3 dissections (documented in 11.3.2.3) allows these to be used as a reference 'backbone' classification within the overall inferred oragnisation of all 2999 cardiovascular procedures processed. Gross inspection of the interdigitated polyhierarchy suggests that, although there are some apparent errors (usually missed classifications rather than misclassifications), it is broadly correct. The depth of the inferred hierarchy – 12 levels – compares favourably with the original schemes which, individually, are typically shallower.

##### 11.3.2.1.3 Threats to Internal and External Validity

**Experimenter bias** – the assessment of whether the interdigitated classification of cardiovascular procedures was 'correct' was made informally and subjectively by the author. As a consequence it has very limited internal validity.

##### 11.3.2.1.4 Conclusion

In the absence of a thorough examination of the inferred hierarchy, the semantic utility of the efforts of the different authors remains in question.

The preliminary examination is, however, encouraging and suggests that at least a reasonable first approximation of a merged polyhierarchical classification could be computed, faster than could probably be achieved by a human working alone, particularly given the fact that the original 3617 rubrics appear in five different languages (French, English, Swedish, Greek, German).

##### 11.3.2.1.5 Additional Benefits

Some unexpected benefits were derived directly from having a corpus of dissections in the intermediate representation and authoring ontology. Firstly, the international community of authors found it interesting to have a formalised and explicit ‘interlingua’ and interchange format that enabled them more easily to exchange and compare their work. Secondly, the systematic use of the same controlled vocabulary allowed a novel form of indexing of the corpus – it is possible, for example, to retrieve all the dissections that use the descriptor ‘heart’ without any reference to the target ontology.

A detailed exploration of these issues is, however, beyond the scope of this thesis.

#### 11.3.2.2 CCAM

Within the corpus of 20,782 dissections authored, one national classification of surgical procedures was processed in its entirety: the French Classification Commune des Acts Medicaux [CCAM 1998]. An initial pilot reported positively on the potential value of intermediate representation dissections during the construction of this new classification. A dissection step was therefore included in the final methodology to service a subsequent French government contract for the construction of the entire classification.

##### 11.3.2.2.1 Method

Clinical domain experts funded by the French Ministry of Health proposed 7478 rubrics across 16 chapters. These were then dissected (using the English language authoring ontology described in this thesis) by a team of three trained dissectors in the University of St Ettiene, France [Trombert-Paviot 2000] using the SPET tool (section 10.2). These dissections were expressed using 2336 different descriptors, 65 synonym descriptors, and 59 different intermediate representation links.

Their work was checked for correctness and completeness by a senior clinician, before transmission to a team of computational linguists at the University of Geneva. The dissections were expanded to the target ontology, but not classified**.** The explicit semantics of the resulting candidate GRAIL expressions were used as input to natural language generation software, which generated one or more phrases in French for each expression [Rassinoux 2000].

Generated phrases were compared by the original clinical panels with their original rubrics [Rodrigues 2000].

##### 11.3.2.2.2 Results

20% of the original rubrics required revision to correct ambiguity, incompleteness, errors of omission or commission, and inconsistencies that were discovered as a result of this comparison.

##### 11.3.2.2.3 Additional Benefit

Although this methodology was conceived primarily as a means to check proposed rubrics for ambiguity, a byproduct was a detailed conceptual index of the rubrics. The CCAM approach is currently under close scrutiny by the German centre of classification as a technology to develop a new German national classification of surgical procedures [Personal Communication 2004], and also by WHO for further development of the International Classification of Healthcare Interventions (ICHI) [Personal Communication 2005].

##### 11.3.2.2.4 Threats to Internal and External Validity

**Experimenter bias** – determination that certain CCAM rubrics were ambiguous was the subjective consensus decision of a panel of experts, and carries correspondingly greater internal validity than if it had been by a single rater.

However, measuring semantic utility of dissections by inspection of something generated from them overlooks those situations where the ‘right’ answer (a generated phrase concordant with the intention of the original rubric) may be obtained for the ‘wrong’ reason: natural language is inherently ambiguous, and it is possible to generate the same ‘concordant’ phrase from both a semantically correct dissection and from some semantically nonsensical graph transformations of the first.

**Selection bias** – the production of CCAM rubrics within University of Ste Ettienne followed a rigorous two step quality control methodology unique to that site: dissections were initially written by one of three junior staff, and each was then inspected and signed off by a single senior clinician before despatch for language generation and subsequent inspection by the CCAM rubric review panels. The results of semantic utility tests for any particular set of authored dissections can not be generalised to all dissections without reference to the internal quality control procedures already applied prior to testing.

#### 11.3.2.3 Cross Validation

Contemporaneously with the work described in this thesis, an exhaustive decomposition of the rubrics of surgical procedures was being authored independently within the NHS’s Clinical Terms Version 3 (CTV3) [Price 1998]. This work had many similarities with the work in this thesis, for example it also followed a schema for representing surgical procedures adapted from CEN ENV 1828. A significant difference was that the semantic declarations in CTV3 were not intended specifically to support any computerised classification of the concepts represented.

The existence of the CTV3 resource offered the possibility to conduct a cross-validation experiment [Rogers 1998].

##### 11.3.2.3.1 Method

A flowchart of the method is present in Figure 88. Compositional information relating to 162 rubrics from the *Endocrine Procedure* subchapter of CTV3 was obtained from the NHS as an extract of their ‘atomic qualifier’ data file. A series of scripts and transforms was written using the TIGGER Transformation Tool to convert between the CTV3 and GALEN schemata (Figure 89), and the data from 149 CTV3 rubrics were automatically re-written in the Intermediate Representation syntax, and using the GALEN interpretation of ENV 1828 (Figure 89).

Data for 13 CTV3 rubrics could not be converted due to ontological style differences. For example, the authors in CTV3 had not always made the same choices as in GALEN regarding whether, or how, to decompose a given concept into constituent atoms. The treatments of the modifiers unilateral/bilateral illustrate this: in CTV3, they modify the structure operated upon, whilst in GALEN they modify the deed itself. Further, the CTV3 authors did not express unilateral/bilateral as separate atoms: they instead remain embedded within primitive entities. For example, CTV3 uses the primitive descriptor ‘bilateral adrenal glands’ in dissecting ‘Bilateral adrenalectomy’:

```
READ_MAIN excision action
    SITE bilateral adrenal glands
```

By contrast, the equivalent GALEN dissection would be:

GALEN_MAIN excision action
    HAS_LATERALITY bilateral
    ACTS_ON adrenal gland

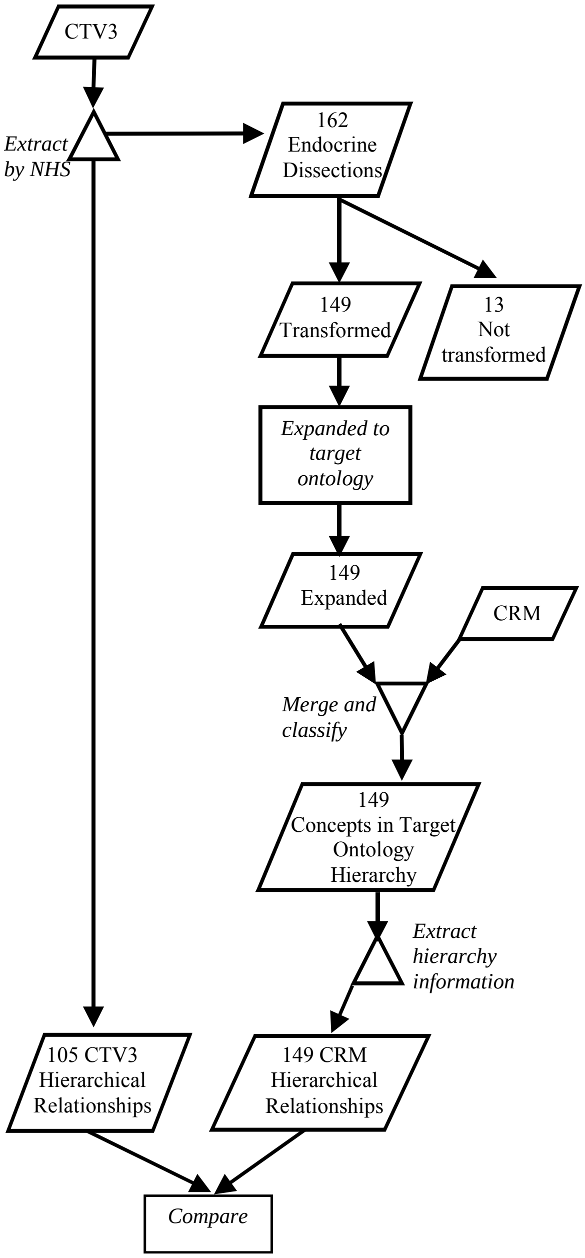



**Figure 88: Flowchart of methodology for cross-validation**

| | |
|---|---|
| Extract of 'atomic qualifier' data file from CTV3 | 71000\|Transethmoidal hypophysectomy\|X9002\|Approach\|X812M\|Transethmoidal\|9\|A\|F<br>71000\|Transethmoidal hypophysectomy\|X900S\|Method\|X793K\|Excision - action\|9\|A\|F<br>71000\|Transethmoidal hypophysectomy\|X9019\|Site\|Xa06A\|Pituitary structure\|9\|A\|F |
| GALEN 'dissection' resulting from scripted transformation of same data | RUBRIC 'Transethmoidal hypophysectomy'<br>SOURCE 'READ' CODE '71000'<br>READ_MAIN excision action<br>    HAS_APPROACH transethmoidal<br>    SITE pituitary structure |

**Figure 89: Transformation of READ templates into dissections in the GALEN intermediate representation**

Automatic transformation between such dissections - to recognise them as semantically equivalent - would have required linguistic and graph manipulation tools to manage their semantic and structural differences. These tools were outside the scope of this limited experiment.

Note that the CTV3 representations did not use the native GALEN authoring ontology for surgical procedures. They were instead expressed using 71 similar, but different, unique concept labels and 7 unique semantic link labels drawn from the body of CTV3. The CTV3 descriptors and links were therefore treated as an alternative authoring ontology.

The CTV3 authoring ontology was mapped to the target ontology, using the same tools and techniques devised for mapping authoring ontologies described in Chapter 10. With the link and descriptor mappings in place, 149 (91%) of the original 162 CTV3 dissections were expanded into GRAIL and presented to a GRAIL reasoner. A new classification of the original 149 CTV3 rubrics was thus derived entirely by automatic analysis, by integrating CTV3's semantic dissections of each rubric into the existing semantic content of the GALEN CRM.

The machine-derived multiaxial classification of the 149 processed dissections was then compared with the multiaxial classification of the same 149 rubrics that had previously been authored manually by the CTV3 authors, and whose specification had accompanied the 'atomic qualifier' data as a 'hierarchical relationship' file (Figure 88). This comparison was performed manually, using the Hierarchy Comparison tool (Figure 90). There were insufficient funds to allow any experimental iteration, whereby errors or differences might have been corrected and the comparison run again.

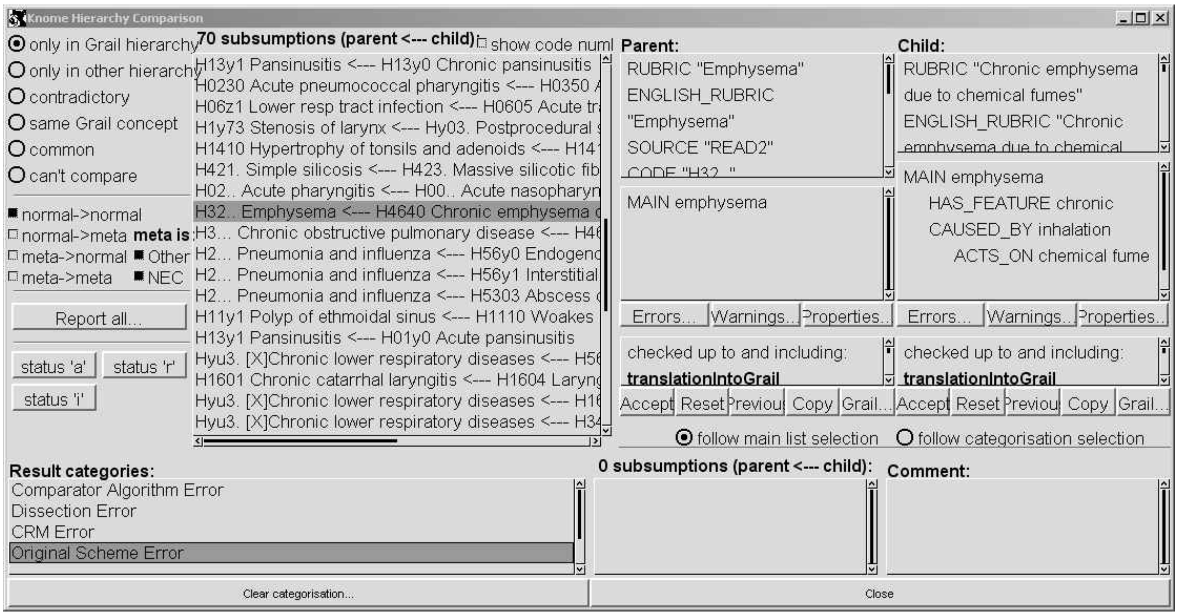



**Figure 90: Hierarchy Comparator Tool**

### 11.3.2.3.2 Results

The manual classification of the CTV3 rubrics placed the 149 rubrics in a subsumption lattice comprising 111 parent child relationships. The computed CRM hierarchy placed the same concepts in a lattice containing 149 parent child relationships.

95 out of 111 parent child relationships in the manually created lattice (86%) were also found in the computed lattice. The differences between the two lattices (asserted and computed) were:

- 52 parent-child relationships were present in the computed lattice, but not in the manual lattice

- 14 parent-child relationships were present in the manual CTV3 lattice but not in the computed CRM lattice
- 2 relationships were present in both lattices, but inverted (the parent concept in the manually created lattice had become the child concept in the computed lattice)

These 68 differences between the two schemes were manually reviewed, and categorised as arising from one of five main causes, as set out in Figure 91:

| CAUSE OF MISSED CLASSIFICATION | Relation only in computed lattice | Relation only in manual lattice | Relation inverted between lattices |
|---|---|---|---|
| CTV3 Dissection errors | 24 | 8 | 2 |
| Incorrect mappings from CTV3 authoring ontology to the target ontology | 2 | 1 | 0 |
| Differences of opinion between anatomical models | 8 | 2 | 0 |
| Errors or inconsistencies in underlying knowledge models | 9 | 2 | 0 |
| GALEN transitivity and partonomy | 9 | 1 | 0 |
| **TOTAL** | **52** | **14** | **2** |

**Figure 91: Breakdown of missed and mis- classifications and their cause**

*CTV3 dissection errors*: Correct classification will not occur if the semantic information in the CTV3 dissections is incorrect.

Two 'inverted' classifications were encountered, in which the parent-child relationship asserted manually between two rubrics in the CTV3 hierarchy was precisely inverted in the derived hierarchy within the target ontology. In one case this occurred because a 'total excision of …' rubric had not been supplied with any atomic qualifier in the CTV3 dissection to indicate the extent of the procedure. In the other a rubric reading 'removal of thyroid nodule' had been given the wrong atomic qualifier (it was stated to be a kind of excising instead of a removing) in the CTV3 dissection.

Six relationships present in the CTV hierarchy could not be inferred by the classifier because, in each parent-child concept pair, both parent and child had semantically identical CTV3 dissections. The classifier therefore determined that the parent and child were, in fact, the same concept and therefore collapsed the pair into a single node in the computed hierarchy.

Two 'CTV3-only' relationships were missed in the computed CRM lattice because, in both cases, a property ascribed in the parent term dissection had not been carried through to the dissection for its child.

Twenty-four 'GALEN only' relationships were inferred because five CTV3 dissections each contained such semantic omissions. The target ontology concept generated from only partially complete semantic information will necessarily be more general than the true meaning of the rubric. The derived classification of the concept is correspondingly higher in the automatically derived hierarchy: spuriously with respect to the rubric's meaning as understood by humans, but correct with respect to the meaning as it was presented to the computer. In this higher position, these concepts then acquire children they did not have in the manually asserted hierarchy - often the terms that were siblings in the manual hierarchy.

For example, both dissections for partial- and total substernal thyroidectomy omitted to mention their approach element (‘substernal’). The generated target ontology concepts then corresponded to the more general notions of partial- and total thyroidectomy, regardless of approach. The two terms between them subsequently acquired eleven false children, including ‘Hemithyroidectomy’ and ‘Lobectomy of thyroid gland’.

The remaining three semantic omissions in the CTV3 data were:

> ‘excision of thyroglossal fistula’ did not include fistula as the pathology atom (three false children);
>
> ‘removal of thyroid nodule’ did not include the nodule (six false children);
>
> ‘Surgical biopsy of endocrine system NOS’ did not include ‘NOS - Operation’ as a classification atom (four false children).

*Incorrect descriptor-to-CRM mappings*: a single ‘CTV3-only’ relation and two ‘GALEN only’ relations were attributed to incorrect mappings of two CTV3 authoring ontology descriptors to the CRM. For example, the CTV3 descriptor ‘modification’ had been manually mapped to the very general CRM notion of ‘any form of surgery’, when a more correct and specific interpretation would have been ‘any form of morphological change’.

*Differences of opinion between anatomical models*: Two ‘CTV3 only’ relations arose because of differences of opinion regarding anatomy. One concerned CTV3’s classification of the thymus as an endocrine gland; although historically a clinically customary classification of the thymus, recent anatomical and endocrine thinking no longer supports this. The target ontology GALEN model, therefore, does not consider the Thymus to be an endocrine organ.

Eight ‘GALEN only’ parent-child relations also reflected anatomical dispute: six of these concerned the thyroglossal tract and associated structures and lesions, which GALEN modelled as part of the thyroid gland whilst CTV3 did not.

*Errors or inconsistencies in underlying knowledge models*: The classifier missed two ‘CTV3 only’ relations because of errors in either the CTV3 or GALEN model. In one, ‘persistent patent thyroglossal duct’ was not flagged as pathological in the CRM and (therefore) was not classified as a subtype of ‘thyroglossal duct pathology’ in that target ontology. In the other, modelling in the CTV3 thesaurus was inconsistent regarding the notions ‘excising’ and ‘removing’; within the CTV3 authoring ontology, excising was given as a more specific form of removing. However, one rubric reading ‘removing of…’ had been classified manually as more specific than one reading ‘excising of…’.

Nine ‘GALEN only’ relations appeared justifiable on formal and semantic grounds. Study of the rubrics themselves confirmed they might be genuine omissions from the CTV3 hierarchy. For example, ‘Endocrine surgical biopsy’ might reasonably subsume biopsies of the adrenal, parathyroid and thyroid glands. These subsumptive relations were not present in the manual classification, but were suggested by the automatic classification.

*GALEN role inheritance and partonomy*: a consequence of the OpenGALEN CRM use of the GRAIL role inheritance mechanism is that ‘excision of gland’ will subsume ‘excision of part of gland’ and ‘excision of lesion in gland’. This is an unnatural classification for most clinicians, for whom the phrase ‘excision of gland’ normally means excision of the whole gland and explicitly not only some part of it, or only of a lesion within it. Construction of a formalism and ontology that do not behave this way is a fundamental partonomic modelling problem and the subject of continuing debate.

One ‘CTV3 only’ classification was attributed to this difference between the two systems, whilst nine ‘GALEN only’ parent-child relations arose: ‘parathyroidectomy’, ‘thyroidectomy’, ‘hypophysectomy’ and ‘adrenalectomy’ each subsumed excision of lesions located in the corresponding gland, whilst 'thyroid incision' also in error subsumed 'incision of thyroid lesion'.

**Summary of Results**

Independently authored third party data was successfully imported into the intermediate representation environment. A total of 126 possible valid parent-child relationships were identified by manual review of both the original asserted lattice and the differences detected during the comparison with the computed lattice. The asserted lattice contained 111 of these 126 valid relationships (88%) and 2 incorrect relationships. The computed lattice contained 112 of the 126 valid relationships (88%) but also a further 28 incorrect relationships arising from errors in the semantics declared in the imported data, and 9 relationships whose validity depends on your partonomic world view. 95 of the 126 true relationships (75%) were found in both the asserted and computed lattices.

#### 11.3.2.4 Large scale cross validation

As a result of the success of the first cross-validation experiment detailed above, another researcher used the same tools and techniques to cross-validate a larger section of CTV3. This experiment is reported briefly here.

##### 11.3.2.4.1 Method

Compositional information relating to 2606 rubrics from five further surgical subdomains within CTV3 (*Cardiovascular, Male Genital, Gynaecological, Ear and Lymphatic procedures)* was obtained. The CTV3 authoring ontology covering these rubrics was mapped to the CRM target ontology, as before in 11.3.2.

2377 (91%) of the rubric dissections were successfully expanded into the target ontology. A derived multiaxial classification was computed, and this was compared with the multiaxial classification of the same rubrics that had been manually asserted by the CTV3 authors.

##### 11.3.2.4.2 Summary of Results

Of the parent-child relationships compared:

- 1549 were found to appear in both the manual and the computed hierarchies
- 2626 were found only in the derived (computed) hierarchy
- 363 were found only in the manually asserted CTV3 hierarchy
- 6 were inverted

Due to the scale of the differences, and the limited resource, no detailed analysis was conducted to determine which of the differences represented missing valid relationships, and which included invalid relationships.

##### 11.3.2.4.3 Threats to Internal and External Validity

**Experimenter bias** – for both the small and large scale cross validation experiments, the manual review and categorisation of differences between the inferred and asserted hierarchies was performed by a single researcher, and therefore potentially subject to experimenter bias.

**Fatigue** – the process of categorising differences between the inferred and asserted hierarchy is lengthy and repetitive, and therefore subject to a fatigue effect.

#### 11.3.2.4.4 Conclusion – Cross Validation

The results of both cross-validation experiments demonstrated that the manually crafted classifications were incomplete. The observed error rate - approximately 15% of all valid parent-child relationships were missing in the manually crafted classification – is similar to error rates detected in other domains [Wroe 2003, Zanstra 1999].

Both experiments confirm that some form of authoring ontology can exist independently of a target ontology: the imported data had been authored using CTV3, an authoring ontology entirely independent of the CRM. However, these experiments can neither support nor refute the use of the particular authoring ontology developed during this thesis or of the particular methodology proposed for constructing one.

The hierarchy comparison shows the computed result to be better than had been done by hand, and for this reason.the experiments are taken to directly support the validity of two further elements of the overall methodology presented in this thesis - specifically an intermediate representation as a syntactic form, and a semantically normalised expansion to a target ontology.

However, whilst the computed hierarchy did find many parent-child links missing from the manual classification, errors in the declared semantics led to a different collection of relationships being missed in the computed hierarchy, as well as some additional invalid relationships being inluded. This result might be misinterpreted as indicating that the methodology or a semantically based approach more generally, was no better than an entirely manual technique. This conclusion overlooks several significant factors:

- Manual quality assurance of classification hierarchies is better at identifying concepts appearing where they should not than those not appearing where they should; incorrect information is more obvious than missing information. Most errors in the manual classification were of missing information (missed classifications). By contrast, each original semantic error declared in either the CRM or the dissections tends to give rise <u>simultaneously</u> to both missed classifications and misclassifications (errors of inclusion) in the computed classification. Although these experiments could not explore this issue further, it is for this reason possible that a semantic based approach might be more accurately quality assured, as computed misclassifications exist to more prominently point the way to the underlying errors.
- The true comparative cost of manual versus computed crafting and curating of classifications is not known. One unpublished report suggests that the semantic approach is less costly than continued manual maintenance, even for very small domains of only a few hundred concepts [Zantra 2004]: the manual addition of 20 new rubrics to the 2000-term musculoskeletal chapter of ICPM-DE required 6 full day meetings involving 10 specialists, and a total of more than 5 man months of effort spread over 8 months. The semantic-based computed approach required an initial one-time investment of 2.5 man months to represent the entire chapter of rubrics, after which the 20 new additions could be integrated very quickly.
- The ability to compute a (mostly) correct polyhierarchy is not the only value added by a semantically-based approach. Concept-based indexing, multilingual natural language generation [Wagner 1999] and the ability to dynamically extend or filter the set of concepts and associated classification are three additional benefits that accrue, and which are notoriously difficult to deliver using manual techniques.

#### 11.3.3 Conclusion – Semantic Utility

Two significant weaknesses are inherent in any effort to assess the semantic utility of authored representations: firstly, significant human resource is required to undertake a manual review of even modest numbers of results generated by processing the explicit semantics of dissections, such as the large scale integration, the CCAM experience and the CTV3 cross validations reported here.

Secondly, there is no agreement in the ontology community regarding the most valid or complete indicator of semantic utility. A more comprehensive and less resource intensive methodology than comparing classifications is yet to be devised.

### 11.4 Generalisability of methodology

The approach described in this thesis was conceived, and developed, to support the specific task of authoring dissections of rubrics taken from surgical procedure classifications. The potential for wider applicability beyond this subdomain and task was investigated in a further experiment, described briefly below.

#### 11.4.1 Applicability to Disease

Rubrics for the cardiovascular and respiratory disease chapters of version 2 of the READ codes (a monoaxial predecessor of CTV3, still used extensively in UK primary care) were obtained. A manually asserted but monoaxial classification of these rubrics was also obtained from the same source.

Unlike the cross-validation experiments, no 3rd party semantic decomposition was available. Therefore, the author of this thesis used the intermediate representation syntax without alteration, and the TIGGER toolset described in Chapter 10, to write new semantic dissections *de novo,* representing the perceived meaning of each rubric.

A new authoring ontology, for diseases, was extracted from the corpus of disease dissections. There was considerable overlap between this new disease authoring ontology and the original surgical procedure authoring ontology. For example, both ontologies had numerous descriptors for anatomical structures and in many cases their names were identical. Where both ontologies contained common descriptors, and where the surgical procedure authoring ontology already contained a mapping from the descriptor to the target ontology, the mapping was imported into the disease authoring ontology. New mappings were authored for descriptors that were unique to the disease authoring ontology.

The concept represented by each disease dissection was expanded into the target ontology, and a classification of all the resulting concepts was computed by a GRAIL classification engine. This was then compared with the manually asserted hierarchy, using the same tools and techniques described in the cross-validation experiment.

The results were broadly similar to those encountered in the cross-validation experiments, although the experiment is complicated by some of the idiosyncratic properties of the READ coding scheme, including the use of subclass relationships as the means to store synonyms. The results presented below have also not been checked for rater bias or variability, and should therefore be treated with caution.

**Cardiovascular**

The cardiovascular disease chapter of the READ codes places 918 concepts in a monoaxial hierarchy with 465 manually asserted parent-child relationships between them. The computed lattice placed the same 918 concepts in a lattice of 490 parent-child relationships, of which 356 relationships were also represented in the READ lattice (agreement with 78% of the READ lattice). 189 differences were found between the two lattices, and these were further analysed:

- 143 differences were attributed to errors in the READ lattice
    - 117 relationships were missing
    - 19 were included in error
    - 1 relation was included, but inverted
    - 6 relations stated as parent-child in READ were in fact to synonyms
- 24 differences were attributed to errors in the target CRM ontology, resulting in 23 missed parent-child classifications and one misclassification within the inferred hierarchy..
- 22 differences were attributed to an error in the authoring of one or more original dissections, resulting in 7 misclassifications, 12 missed classifications, 2 inverted relations and 1 instance where parent and child concept pair were mistakenly represented as the same concept in the inferred hierarchy.

**Respiratory**

The respiratory disease chapter of the READ codes placed 656 concepts in a manually asserted monoaxial hierarchy with 404 parent-child relationships between them. The computed lattice placed the same 656 concepts in a lattice of 391 parent-child relationships, of which 318 relationships were also represented in the READ lattice (84% agreement with the READ lattice). 115 differences were found between the two lattices, and these were further analysed:

- 80 differences were attributed to errors in the READ lattice
    - 70 relationships were missing
    - 5 were included in error
    - 1 relation was included, but inverted
    - 4 relations stated as parent-child in READ were in fact to synonyms
- 16 differences were attributed to errors in the target CRM ontology, resulting in 12 missed parent-child classifications, one inverted relation and 3 instances where parent and child concept pairs were mistakenly represented as the same concept in the inferred hierarchy
- 19 differences were attributed to an error in the authoring of one or more original dissections, resulting in 13 missed parent-child classifications in the inferred hierarchy, one inverted relation and 5 pairs of parent and child concepts being collapsed into one concept.

**Summary**

Following the analysis of differences between the inferred and asserted classifications, a gold standard set of polyhierarchical true relations was inferred: any relationship that was common to both classifications was assumed, without further inspection, to be true. To this initial set was then added all relationships present in one hierarchy but not in the other but which, on inspection, was confirmed to be a valid parent-child relationship. The performance of the manual and inferred classification strategies with respect to the gold standard is presented in Figure 92:

| | Number of Nodes examined | Gold Standard true relations | Common Relations | Relations in READ Lattice | | | Relations in Inferred Lattice | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Total | True | False | Total | True | False |
| Cardio | 918 | 511 | 356 | 465 | 439 | 26 | 490 | 474 | 16 |
| Resp | 656 | 423 | 318 | 404 | 394 | 10 | 391 | 389 | 2 |
| **TOTAL** | **1574** | **934** | **674** | **869** | **833** | **36** | **881** | **863** | **18** |

**Figure 92: Analysis of disease modelling experiment results**

These combined data show that the inferred hierarchy has slightly but significantly greater recall than the manual hierarchy (92% of all true relations present compared to 89%, $\chi^2$=5.76 p<=0.025) and slightly but significantly higher precision (98% of stated relationships in the inferred hierarchy are correct, compared to 96% in the manual hierarchy $\chi^2$=6.45 p<=0.025). However, an interesting observation is that, although both hierarchy construction strategies (manual vs inferred) are reasonably precise, only 72% of all correct relations are present in both, suggesting that they may be precise in different ways.

Note that these results suggest a relatively small difference exists between the number of relations present in the original explicitly monohierarchical READ classification (869 relations), and the derived 'Gold Standard' polyhierarchical classification (934 relations). This may indicate the influence of a selection bias, in that the particular set of terms chosen for the experiment more naturally form a relatively shallow and not very dense polyhierarchy. A similar experiment conducted on a set of more closely related terms may produce significantly different scores for precision and recall.

#### 11.4.2 Applicability to drug information

The methodology and tools developed during the work described in this thesis were used in a subsequent project to represent, and reason over, information about prescribable drugs [Solomon 1999, Wroe 2000, Rogers 2003]. A detailed description of this project is outside the scope of this thesis. Only the key results are presented here.

For the work on drug information, the unaltered intermediate representation was sufficient to allow definitional descriptions of pharmacological products to be represented. For example, the notion of an antipseudomonal drug in powder form for injection could be represented as:

```
MAIN drug
HAS_FORMULATION powder
HAS_ROUTE injection
HAS_DRUG_FEATURE physiological action
    WHICH_IS life damaging process
        ACTS_ON pseudomonas
```

However, additional support was required for the appropriate GRAIL constructs needed to assert non-definitional links between a product so described, and those properties that are true of the product, but not part of its definition. For example links to the set of conditions it may be used to treat, or those that are side effects.

The intermediate representation syntax was therefore extended to provide this support. The following example shows how the original definition of an injectable antipseudomonal agent (above) is further linked to the notion of being indicated for pseudomonal infections, via the only Intermediate Representation keyword: PROPERTIES, which signals that everything preceding it is part of the definition of a concept in IR, whilst everything following it should be treated as necessary properties of the concept so defined:

```
MAIN drug
HAS_FORMULATION powder
HAS_ROUTE injection
HAS_DRUG_FEATURE physiological action
   WHICH_IS life damaging process
      ACTS_ON pseudomonas
PROPERTIES
   HAS_DRUG_FEATURE indication
      FOR managing
         ACTS_ON infection
                CAUSED_BY pseudomonas
```

With this modest extension of functionality, the intermediate representation then proved sufficiently expressive to support authoring of detailed descriptions of more than 1100 pharmacological products listed in the British National Formulary, together with their mechanisms of action, formulations, mode of delivery, indications, side effects, contraindications and interactions.

This work was performed by three model authors with considerable experience of working directly with the target ontology. However, because of the complexity and length of the statements to be expressed, all three model authors doubted that it would have been cognitively possible to have performed this task working directly in the target ontology.

The authoring ontology for this work included many unique new descriptors (principally, long lists of pharmacological compounds and formulations) but also some that were common to the prior surgical procedure and disease work. Several new authoring ontology links were created (e.g. HAS_DRUG_FEATURE) but some of those originally devised for the surgical procedure authoring ontology were re-used (e.g. ACTS_ON, CAUSED_BY).

The resulting corpus of drug information, expressed entirely in intermediate representation and an authoring ontology, was expanded to the target ontology. A multiaxial classification of prescribable products was constructed.

#### 11.4.3 Conclusion

These experiments provide further evidence to support the validity of the overall paradigm presented in this thesis: representations of the semantics of two large corpora of phrases or domain knowledge were represented entirely within the intermediate representation paradigm. Subsequently these representations were expanded to the target ontology and, in the case of the work on diseases, the resulting classification properly compared favourably with a previously authored classification and with an inferred 'Gold Standard'. The team who worked on the drug information do not believe that the task could have been completed using the target ontology directly.

Additionally, both experiments provide evidence for the generalisability of the approach: two new subdomains were processed with very little modification of the methodology.

### *11.5 Resource requirements*

A full assessment of the overall approach should include the cost of the central support services. This is the effort required both to coordinate the construction and maintenance of both an authoring ontology, and the normalisation algorithms that determine the final semantic utility of any authored dissections.

Two issues are addressed in this section: the size of the effort required to create one authoring ontology, and the extent to which that effort could be re-used in constructing other authoring ontologies.

### 11.5.1 Initial effort

The central effort supporting the distributed authors comprised three primary tasks:

- iterative population of the authoring ontology including manually declaring, or validating lexically suggested, mappings to the target ontology
- parallel development directly in GRAIL of the CRM and its metamodels, including extensions such that the CRM contained concepts equivalent to new descriptors, and the metamodel permitted appropriate compositions
- authoring of the semantic normalisation algorithms

During the work on surgical procedures described in this thesis, a central effort of approximately 2-3 person years was required. However, the target ontology was relatively underdeveloped at the start and more than half the central effort was expended developing it, particularly the modelling of human anatomy.

One premise of the overall methodology was that, as more medical subdomains were tackled, so the target ontology should grow more comprehensive and mature, and proportionately less new modelling would be required. To date, subsequent experiments appear to bear out this expectation. For example, the anatomy modelling originally done to support surgical procedures proved to be widely reusable without significant alteration or enhancement in both the work on diseases and drug information documented above: 33% of all 772 descriptors used in the 1865 cardiovascular and respiratory disease dissections described in section 11.4.1 had previously been encountered, and mapped, during the authoring of the 20,782 surgical procedure dissections. Similarly, 5.6% of all 8164 descriptors used within the 26,970 drug ontology dissections had been previously encountered. The additional target ontology modelling effort required to support the (limited) disease work amounted to approximately 1 man month.

The effort required to populate an authoring ontology – to collect new descriptors as authors begin working in a new subdomain, and feed the canonical list back to them periodically – appears to be front-loaded. The typical distribution of the frequency with which descriptors are used divides them into a core of very frequently used descriptors that will be encountered *en masse* early on in the collation exercise, and a long tail of much less frequently used descriptors that will be encountered sporadically over a much longer period of time. This phenomenon has the effect that the rate of growth of new descriptors is very fast in the early stages of working in a new domain, but then rapidly reduces.

By contrast, the work required to author semantic normalisation rules is generally back-loaded: it is necessary to have a substantial corpus of different patterns of expression before such rules can be identified and tested.

### 11.5.2 Reuseable effort

The generalisability experiments revealed a high degree of overlap between subdomains with regard to both the authoring ontology descriptors and links employed, and their mappings. In both experiments reported – for dieases, and for drug information – the new authoring ontologies were extensively pre-populated from the cumulative list by including common descriptor categories (e.g. anatomy, disease or drugs).

The great majority of descriptors encountered in any subdomain were found to be semantically unambiguous: their unique mapping to the target ontology remained valid even when used in different authoring ontologies. A brain is a brain, whether it is the target of a surgical procedure, the location of a disease, or the site of action of a drug. Exceptions were only very rarely encountered. For example, the descriptor ‘suspension’ denotes a deed in the surgical procedure authoring ontology, but is a kind of pharmacological formulation in the drug authoring ontology.

Similarly, the generalisability experiments demonstrated both the existence of a core of links (e.g. ACTS_ON, HAS_PART, CAUSES, CAUSED_BY) common to all authoring ontologies, and further that the semantic normalisation rules required to expand these common links are also re-usable across subdomains and tasks.

#### 11.5.3 Conclusion

The central effort is modest when compared to the peripheral effort it is able to support: a ratio of one central model author to between 10 and 20 distributed authors was achieved.

This compares favourably with the alternative, where the centre would necessarily become involved in a lengthy apprenticeship of all participants learning to use the target ontology directly, during which time both centre and trainee would achieve low productivity.

### *11.6 Known limitations of current methodology*

Three limitations of the described methodology were encountered during the course of the experiments of opportunity:

#### 11.6.1 Descriptor mappings

It is possible that an Intermediate Representation author might represent as a composition of authoring ontology descriptors and links a concept that is represented in the target ontology as a primitive. An example might be an author who represents the concept 'hepatic artery' as 'artery SERVES liver' whereas, in the target ontology [HepaticArtery] is a primitive entity. The descriptor mapping methodology described above does not allow for this possibility: whilst primitive descriptors can be mapped to composed entities in the target ontology, no mechanism exists to map a composed entity in the authoring ontology to a primitive target ontology entity.

In practice the situation where such a mapping was required rarely arose, and it usually indicated a semantic error on the part of the intermediate representation author. However a mechanism to support such mappings would be a useful extension of the work described in this thesis.

#### 11.6.2 Link mappings – right contexts

The described algorithm for selecting a link mapping from a candidate link mapping set, detailed in 9.3.2, depends mainly on a consideration of the nature of the concepts on either side of the link: its left and right contexts respectively.

The described algorithm, however, is naïve for the right context: only the concept to which the right-hand descriptor is directly mapped is considered. This algorithm is blind to the situation where the right-hand descriptor does not stand alone, but appears as the head of a subgraph within the representation. In such situations the 'true' right context of the link should be equal to the target ontology composition that results from expanding the entire subgraph to the right of the link.

For example, in Figure 93, the current naïve algorithm would expand the link HAS_LOCATION identically in both representations, because the left and right contexts for the link are identical. The target ontology metamodel does not allow that a vegetation lesion can sensibly be located on *any* kind of valve (e.g. not including valve components of machines), and therefore HAS_LOCATION would in both cases be expanded to the most abstract link available.

| CURRENT (NAÏVE) ALGORITHM | POTENTIAL IMPROVED ALGORITHM |
|---|---|
| MAIN removing<br>ACTS_ON vegetation<br>HAS_LOCATION valve | MAIN removing<br>ACTS_ON vegetation<br>HAS_LOCATION **valve**<br>**IS_PART_OF vein**<br>**HAS_FEATURE xenograft** |

**Figure 93: Determination of right context by subgraph**

By contrast, an improved algorithm might treat the right hand representation differently: the entire subgraph in bold type, corresponding to the much more specific notion of a xenograft venular valve, would be considered as the right context. In this scenario, HAS_LOCATION might be expected to expand to a more specific semantic link.

### 11.6.3 Link Mappings – users can not be semantically general

Section 9.3.2.4, clarifying the link mapping algorithm, exposed one limitation of the overall intermediate representation approach.

One of the primary goals of the semantic expansion process is to compensate for the inability of inexpert users to be semantically specific, particularly with regard to semantic links. The approach described is based on the assumption that a model author can use a relatively non-specific semantic link type, from which the system will infer an appropriate, maximally specific mapping.

This obviously works **less** well when the users' intent was in fact to be semantically general rather than specific. Further, in an ontology where there is a rich hierarchy of semantic link types, the choice is not binary between either an entirely abstract mapping or an entirely specific one: there are degrees of abstraction that some authors may wish to exploit.

The current semantic expansion algorithm is therefore limited in two respects: firstly, it provides no means to access the range of intermediate interpretations between maximally specific and maximally abstract. Secondly, the algorithm for selecting between maximally abstract or specific is fixed and can not be overridden. Ideally, a mechanism is needed for expert users to override the default selection behaviour and indicate when they intend to be general rather than specific, and to specify exactly how general they want to be.

### 11.6.4 Link Mappings – Cardinality

The intermediate representation itself includes no facility for authors to specify the cardinality of specific authoring ontology links. A number of surgical procedure rubrics were encountered where it was necessary to provide this capability, but the only mechanism available to do this was to copy the workaround used in GRAIL models: create variant flavours of authoring ontology link, with specific cardinality properties. ACTS_ON_1 and ACTS_ON_2 (see Appendix 2) are examples of authoring ontology inks that exist for this reason. Their link mappings are constructed to force the choice of a manyMany variant of [*actsOn*].

A further limitation already mentioned is that GRAIL itself does not support qualified cardinality constraints. It is likely that the issue of how to manage cardinality will require re-examination if the methodology is applied to more sophisticated description logic.

# 12 Discussion

The thesis presented in the introduction was that:

> *Much of the complexity of formal ontologies arises from the consistent application of semantic patterns and choices. The cognitive load of using a complex formal ontology can be reduced if these patterns and choices are made explicit as a metamodel of the ontology, and where the metamodel is subsequently harnessed to guide user choices pre hoc and transform expressions post hoc to a preferred semantic form.*

Part One of the thesis has described some of the significant sources of complexity in formal ontologies, including the cognitively difficult task of applying ontological choices consistently over a very large domain. Part Two described a methodology to engineer a reversible systematic simplification of a complex target ontology, including *pre hoc* constraints and *post hoc* semantic normalisations driven by an underlying metamodel. This methodology allowed the population of the OpenGALEN CRM to proceed as two parallel and loosely coupled activities: a small central team of expert users extending the CRM and its metamodel directly in GRAIL, and a much larger and distributed team of inexpert users performing the bulk representation of the semantics of 20,782 surgical procedure rubrics.

The experiments of opportunity described in Part Three provide considerable evidence both that this methodology increased the accessibility of the target ontology to a group of users, and that useful work was produced.

This chapter discusses further work that would be necessary to more strongly prove the thesis, and describes current and future prospects that such further work may be carried out. Finally, the implications of this thesis for programmes to introduce ontology-driven point of care applications to support clinicians, such as the National Programme for Information Technology, are discussed.

## *12.1 Further Work*

Chapter 11 described several experiments of opportunity to evaluate elements of the overall methodology. This section considers options for further evaluations.

### 12.1.1 Time to train

Greater internal and external validity for the time to train evaluation would be gained by a more representative sampling of subjects to reflect other ontology user groups, such as practising clinicians. Further, rather than a cross-over study and associated risk of a practice effect, the sample should be randomised to either GRAIL training and authoring, or Intermediate Representation training and authoring. However, any measure of time to train must first provide a clearer definition of when either group of authors may be considered to be ‘trained’, a status especially hard to define in the context of a GRAIL apprenticeship.

One approach might be to measure the amount of central training investment that must occur before subjects reach a predetermined productivity level, for example the ability to author a test set of dissections to a given quality within a set period of time. Such an experiment, however, would require an agreement on how to determine and measure the quality element of the exit test, and repeated testing of individual subjects until they met the requirement to be considered ‘trained’ would obviously introduce a new testing effect.

A central difficulty with such a study remains the cost of recruiting sufficient clinical time both for the initial training and subsequent bulk dissection writing (particularly for the GRAIL arm of such a study) through which the effectiveness of training is to be assessed.

### 12.1.2 Productivity

As above, greater internal and external validity would be gained by a more representative sampling of subjects from other ontology user groups, such as practising clinicians.

The crude measure of cumulative productivity from the entire group of dissection authors could be significantly improved by experiments in which productivity was recorded on a per subject basis (rather than for the whole group) and was measured not as total cumulative output but time taken to perform a standard modelling task. This could be further refined to a time series study, allowing greater exploration of issues such as rate of increase in individual productivity following training, and a measure of the distribution of productivity achieved within and across different subject populations.

As for time to train, a central anticipated difficulty would be the cost of recruiting and retaining sufficient clinician time. Additionally, defining a standard task necessarily requires definition of the quality with which that task should be completed and, consequently, a methodology for measuring dissection quality (see 12.1.4).

### 12.1.3 Reproducibility

As described in Chapter 11, achieving a common input presents a central difficulty for any experiment to assess how reproducible the ouput of dissection authors is. An alternate experimental design might provide as the input a template comprising both a model phrase and a matching pre-authored representation. A more representative sample of clinical authors might then be asked to represent phrases similar to or derived from the template phrases. A refinement of this methodology could re-test individual subjects intermittently as a time series to see whether reproducibility improves as authors become more experienced. Some experimental designs could blind the subjects to the fact that they were being tested.

However, even these designs carry a potential flaw: the arbiter of whether two different authors have produced the same output should not be that their dissections have the same surface appearance, but that they are understood to be semantically identical according to some formal algorithm for comparing semantics.

Further, whilst GRAIL and similar formalisms may allow us to determine whether different authors produce semantically identical output, a more useful measure might be of the extent to which representations from different authors are similar, if not actually identical. Algorithms for measuring semantic similarity are much less well developed than algorithms for proving semantic identity

### 12.1.4 Semantic Utility and Dissection Quality

A fundamental problem for knowledge representation projects is that no gold standard representation of ‘the truth’ exists against which to test authored work for recall and precision. Further, *post hoc* manual verification of complex knowledge bases is both extremely costly and highly prone to observer bias, such as the significantly greater reliability of human verifiers when detecting errors of inclusion compared to errors of omission. Rater fatigue compounds this problem, especially for very large knowledge bases.

As a proxy for validation against a gold standard, cross validation against other works may be attempted, as described in 11.3.2.3. However, opportunities for such experiments remain typically obstructed by intellectual property and licensing constraints. Further, as the reported experiments demonstrate, these works are theselves only ‘bronze standards’: they have themselves been only manually checked, and still contain their own relatively easily detectable errors of inclusion and ommision. The reported experiments have, typically, involved less than a thousand concepts per experiment. A larger scale cross-validation would be valuable, for example against a much larger fragment of SNOMED CT.

Notwithstanding legal concerns, large scale cross validation would require very significant resource for the pain-staking manual review of the differences detected by automated hierarchy comparison. This is particularly true for any experiment involving inspection and categorisation of differences by multiple raters in parallel, such as to minimise observer bias and fatigue effects.

#### 12.1.5 Assessing the utility of collaborative work

The CTV3 cross validation for surgical procedures described in 11.3.2.3, and the generalisation to disease experiments in 11.4.1, both demonstrate that the work of *an individual* author, using their own authoring ontology, are sufficient to compute a result similar to that which had already been created entirely by hand. No experiment has yet been performed to demonstrate that the work of *a group of different authors*, using this methodology, can be treated as a semantically coherent set of representations. Although it may be possible to expand the output of different authors to one target ontology using the same normalisation and expansion rules, this does not guarantee that any such merger is without significant semantic error or inconsistency.

In this regard, although the CCAM work employed multiple authors, their output was never validated specifically for whether it could be reasoned over coherently. In particular no single classification of the concepts they represented was ever computed, or checked. An experiment similar to that described in section 11.3.2 (interdigitation of cardiovascular procedures) but accompanied by a thorough, multiply rated manual check of the combined hierarchy as above, remains to be performed.

A further problem in the experiments to date is that the work of different authors relates to different classifications, whilst no authoritatively quality assured hierarchy exists covering the union of their dissection output. An experiment therefore to be considered would take all dissections from a single subtree of an already hand-crafted ‘bronze standard’ source (such as one chapter of SNOMED or CTV3) and randomise the nodes of that tree to different authors for dissection. The merged corpus of dissections could then be processed to derive an inferred hierarchy, and this compared with the ‘bronze standard’.

## *12.2 Intermediate level users*

A design goal for the methodology was to hide as much of the complexity of the target ontology as possible, and allow a group of users to interact with the target ontology more easily. Whilst the experiments of opportunity suggest this methodology was successful with regard to these specific aims, they also exposed a weakness with respect to users of intermediate expertise: some authors, presented with an incorrect computed classification of their dissections were suspicious that the error lay in the semantic expansion process, but found themselves unable to explore their theory.

In part this was because the tooling was not designed to allow them to do so, but a further factor worthy of exploration is this: a perhaps inevitable result of making explicit the metamodel for a complex ontology, and devising new algorithms to more tightly bind it to the model, is that the resulting combined construct is considerably more complicated than the complex object that was originally to be hidden. In order to present the wolf in sheep’s clothing, the wolf has had to evolve. Thus, whilst the methodology may have increased accessibility for the inexpert users, at the same time it has increased the distance anybody must travel in order to become an expert.

Further research is therefore needed to determine how the methodology might be adapted to allow for improved debugging facilities, and provide interested users with a migration pathway from the inexpert to the expert user experience.

## 12.3 Implications for ontology driven point-of-care applications

The introduction to this thesis described how traditional medical terminologies, normally deployed as auditing and epidemiological tools, are being re-engineered as clinical ontologies in order to meet the challenges of constructing point-of-care clinical information systems. This section explores the implications of this thesis for that re-engineering, with specific focus on SNOMED CT© and its role in the English National Programme for Information Technology (NP*f*IT).

### 12.3.1 SNOMED CT© and sources of complexity

Any single clinical ontology that aims to cover substantially all of medicine and all user groups is at risk of many of the kinds of complexity described in Chapter 3: the scale and inherent complexity of the domain are inescapable, as are the lack of clear boundaries within it and the different perspectives of users. All description logics, seen from the point of view of ordinary clnicians, have opaque syntax. Strategies to work around limitations in the formalism, such as SNOMED CT's role grouping, serve to further confuse.

### 12.3.2 SNOMED CT© Metamodel

A central tenet of this thesis is that, once an ontology becomes complex, mechanisms are required to guide data entry *pre hoc* and normalise entered data *post-hoc*. In this thesis, such mechanisms are driven by, and therefore dependent on the existence of, an explicit machine readable metamodel. In this regard SNOMED CT's continuing lack of any such explicit metamodel, and the corresponding absence of tools to guide, constrain or normalise user input according to that metamodel, are causes for concern.

The July 2004 SNOMED CT© release was the first to explicitly identify a small subset of its attributes – initially just 54 – as 'approved' and, for each, to specify some domain and range constraints. Whist these constraints may be interpreted in spirit as roughly similar to the GRAIL grammatical sanctions (meta-metamodel) within the *Open*GALEN Common Reference Model, it remains a significant weakness that they are presented only as human readable text and not also as a standardised computer readable format. Further, at the time of writing, the SNOMED CT© metamodel is still evolving with some of the initial set of 54 approved links already retired from that list six months later, and new ones from the previously unapproved list marked as approved.

### 12.3.3 Quality of SNOMED CT© content

SNOMED CT©, like all ontologies, is subject to the insoluble problem of there being no "Gold Standard" reference to compare its content with in order to automatically detect any errors of inclusion or omission. Validating SNOMED CT© content is therefore necessarily a largely manual process of systematic human inspection and review. Other researchers are actively exploring automatic or semiautomatic methods of SNOMED validation and cross-validation, and are reporting significant content errors [Ceusters 2004, Bodenreider 2005]. These errors of inclusion and omission are considered likely to adversely affect the ability of any reasoner to reliably derive canonical forms for post coordinated SNOMED CT© expressions and, as a result, to detect semantic equivalence detection or compute post-hoc dynamic classifications.

#### 12.3.3.1 Quality Assurance of generative formal ontologies

Systematic manual review of the static content within an ontology like SNOMED CT at a specific point in time (e.g. of centrally released content) is not necessarily an appropriate test of an object that is, fundamentally, dynamic. For a generative ontology, where new compositions may be created by end-users and classified *post hoc* (‘post-coordinated’), ideally a guarantee could be given that all such new compositions would be correctly classified. However, it is physically impossible to exhaustively review and affirm all possible computed results, because their number in even a modest compositional scheme grows as a combinatorial explosion. Such an assurance, thereore, can not be directly verified.

As a currently incomplete, partly incorrect, and still evolving product, users of SNOMED CT© are therefore potentially faced with additional sources of complexity not previously described: the challenge of working around limitations and errors in the knowledge content of any one release of an ontology, and of adapting those workarounds in response to changes in those errors as they are (hopefully) corrected across successive versions of the product.

#### 12.3.4 Impact on SNOMED CT© users

As a result of the issues raised above, current versions of SNOMED CT© are likely to exhibit many of the kinds of complexity identified in this thesis, together with additional issues due to the numerous factual errors of omission and inclusion in current versions of the product. There is some evidence that clinical end-users are already suffering some of the consequences of ontological complexity outlined in chapter 4: specialty subsets have been requested (to reduce concept clutter), whilst Bentley’s 1999 paper documenting user difficulty navigating polyhierarchies related specifically to users of CTV3, the scheme from which much of SNOMED CTs current content was directly copied.

The phenomenon of intercoder variability – one external manifestation of concept clutter and polyhierarchy disorientation - is already familiar to users of traditional terminologies generally, and has been specifically reportd in the context of direct coding by clinicians using SNOMED CTs ancestor schemes, such as version 2 of the READ codes [Rogers 2002, Rogers 2003]. Inconsistent coding between clinicians representing clinical data using SNOMED CT is a serious risk, especially whilst SNOMED CT content continues to include many semantically duplicate concepts and links arising from the merger of SNOMED RT and CTV3.

The potential for schema uncertainty in SNOMED CT© users was formally recognised in January 2005 in an NHS Information Authority consultation document [Cheetham 2005] enumerating the different potential schemas already existing within SNOMED CT© for recording the laterality of procedures. The same consultation document expresses the case for a ‘near to user’ form of SNOMED CT© similar to an intermediate representation. Clinical end-users would still be obliged to use SNOMED CT© terms, but might be permitted to compose laterality using more than one schema that would be coerced *post hoc* to a canonical form.

#### 12.3.5 Conclusion

Overall, given the work of this thesis, the current quality of SNOMED CT© content, and the limited understanding of clinical ontologies in the clinical community at large, it seems optimistic to hope that clinical information coded at the point of care by clinicians using SNOMED CT© could be of high semantic quality in the forseable future. In particular, it is inlikely to support the correct operation of sophisticated clinical applications such as are envisaged in the NPfIT procurement documents.

Three factors are likely to significantly degrade the semantic quality of clinician entered SNOMED CT© coded entries, and these will manifest themselves unpredictably as mis- and missed post-coordinated classification:

- inter-rater inconsistency in term selection
- schema confusion and inconsistency when framing compositions
- errors in the current SNOMED CT© core content

This thesis proposes a methodology that addresses the first two of these factors. Debugging the SNOMED CT© core content itself is necessarily a lengthy and mainly manual process. Although the SNOMED editorial strategy explicitly calls for the willing participation of clinical users in identifying errors in the course of operational use, this strategy presupposes that an incentive can be found to recruit and engage clinicians to enter coded information whilst, for the reasons given above, that data will be of limited value.

## 12.4 Future Prospects

During the period of work described in this thesis, the bioinformatics research community has embraced informal ontologies as a tool to organise its very large information space (e.g. The Gene Ontology, MGED and SAEL). However, a migration pathway to formal approaches is now being considered (e.g. Open Biological Ontologies, GONG). Within the medical world, SNOMED CT© remains a likely *de facto* international standard, but the issues described in the preceding section 12.3 are causing growing concern.

In parallel with these interests, the field of ontology engineering research has expanded significantly and broadened its appeal outside biomedicine. An approach founded on formal ontologies is being adopted within the Semantic Web initiative, embodied in the agreed standard representation, the description logic OWL. Much of the current activity has focused on optimising algorithms for reasoning over ontologies so that they are computationally tractable. The problem of how authoring, and using, a complex ontology can be made a cognitively tractable task remains.

GRAIL as a formalism is now more than ten years old, and much of the tooling used to implement the methodology described in this thesis (particularly the TIGGER) is bound to that formalism, esoteric and no longer maintained. However, the issues described in Part One of this thesis and the associated challenge of offering users a less demanding interaction with a complex ontology are likely to become increasingly relevant.

A substantial reimplementation of both the tools outlined in Chapter 10 and the methodologies of Part Two of this thesis are planned within a mainstream OWL and JAVA environment (the PROTÉGÉ OWL Tab), and parallel work has begun to express large parts of the *Open*GALEN CRM - the target ontology in this thesis – in the OWL formalism for further investigation of the behaviour of large and complex ontologies. However, a number of significant issues have been identified with the OWL formalism itself and with the current OWL tooling, and these remain to be addressed before the work of this thesis could continue within an OWL environment. Chiefly these issues concern:

- Exploring the scaling behaviour of OWL classifiers when reasoning over ontologies as large and rich as the *Open*GALEN CRM.
- Achieving a computationally efficient OWL implementation of functionality equivalent to GRAIL's role inheritance.
- Developing tools or methodologies by which structures similar to the *Open*GALEN CRM metamodel might be represented and used in any reimplementation of semantic normalisation algorithms.
-

## References

Artale A, Franconi E, Guarino N, Pazzi L (1996). **Part-Whole Relations in Object-Centered Systems: An Overview**. *Data and Knowledge Engineering (DKE) Journal* 20: 347-383, North-Holland, Elsevier.

Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider P. (2003a) **The Description Logic Handbook**. Cambridge University Press. ISBN: 0521781760

Baader F, Horrocks I, Sattler U. (2003b) **Description logics as ontology languages for the semantic web**. In Dieter Hutter and Werner Stephan, editors, *Festschrift in honor of Jörg Siekmann, Lecture Notes in Artificial Intelligence*. Springer, 2003. To appear

Bada M, McEntire R, Wroe C, Stevens R. (2003). **GOAT: The Gene Ontology Annotation Tool**. *Proceedings of the 2003 UK e-Science All Hands Meeting*, 514-519. Nottingham, UK

Bernstein RM, Hollingworth GR, Viner G, Shearman J, Labelle C, Thomas R. (1997) **Reliability issues in coding encounters in primary care using an ICPC/ICD-10-based controlled clinical terminology**. *Journal of the American Medical Informatics Association, Symposium Supplement* 1997, Vol 21: p 843 and D004493, 1997.

Bentley TE, Price C, Brown PJB. (1999) **Managing Hierarchies in Modern Healthcare Terminologies** In Lorenzi NM. (Ed) *Proceedings of the 1999 AMIA Annual Fall Symposium Philadelphia*, Hanley & Belfus, 1999: 1025.

Bodenreider O. (2001) **Medical Ontology Research: A Report to the Board of Scientific Counselors of the Lister Hill National Center for Biomedical Communications**. http://etbsun2.nlm.nih.gov:8000/publis-ob-offi/pdf/2001-MOR-BoSC.pdf

Bodenreider O, Smith B, Kumar A, Burgun A. (2005) **Investigating subsumption in SNOMED CT: an exploration into large description logic-based biomedical terminologies**. *Submitted to Artifical Intelligence in Medicine* *http://etbsun2.nlm.nih.gov:8000/pubs/pdf/2005-aiim-ob.pdf (visited 8/2/05)*

Borgida A, Patel-Schneider PF, (1994) **A Semantics and Complete Algorithm for Subsumption in the CLASSIC Description Logic**. *Journal of Artificial Intelligence Research* 1 (1994) 277-308

Borgida A. (1996) **On the relative expressive power of Description Logics and Predicate Calculus**. *Artificial Intelligence* 82 (1996) 353-367

Booth N. (2003) **Sharing patient information electronically throughout the NHS** *BMJ*, Jul 2003; 327: 114 - 115.

Broverman CA, Kahn MG. (2000) **An Invitational Workshop: Towards Representations for Sharable Guidelines**. *Proc GLIF Workshop*; Boston MA 2000. http://www.glif.org/workshop/

Brown PJB, O'Neil MJ, Price C. (1998) **Semantic representation of disorders in version 3 of the Read Codes**. *Meth Inform Med* 1998; **37**: 415-9.

Brown PJB, Price C. (1999) **Semantic Based Concept Differential Retrieval and Equivalence Detection in Clinical Terms Version 3 (Read Codes)**. In Lorenzi NM. (Ed)

*Proceedings of the 1999 AMIA Annual Fall Symposium* Philadelphia, Hanley & Belfus, 1999: 27-31

Brown PJB, Odusanya L. (2001) **Does Size Matter? – Evaluation of Value Added Content of Two Decades of Successive Coding Schemes in Secondary Care.** In: Bakken S, (Ed). *Proceedings of the 2001 AMIA Fall Symposium.* Philadelphia: Hanley and Belfus, 2001: 71-5

Brown PJB, Warmington V, Laurence M, Prevost AT. (2003) **Randomised crossover trial comparing the performance of Clinical Terms Version 3 and Read Codes 5 byte set coding schemes in general practice.** *BMJ*, May 2003; 326: 1127 – 0

CCAM Departement De Sante Publique Et D'information Medicale (1998) **Classification Commune des Actes Medicaux** http://www.ccam.sante.fr/

CEN European Committee for Standardisation , Technical Committee 251 (1999) **EN 1828 Health informatics - Categorial structure for classifications and coding systems of surgical procedure**. CEN TC251 WGII 1999 & http://www.centc251.org/WGII/WGIIdocs03-98.htm (visited Jul 29 2004)

Ceusters W, Smith B, Flanagan J. (2003) **Ontology and Medical Terminology: Why Description Logics are not enough**. In *Proceedings of Towards an Electronic Patient Record (TEPR 2003)*, San Antonio, 10-14 May 2003

Ceusters W, Smith B, Kumar A, Dhaen C. (2004) **Ontology-Based Error Detection in SNOMED-CT** In *Proceedings of Medinfo 2004*, San Francisco (in press)

Chandrasekaran B, Josephson JR, Benjamins VR (1999) **What are ontologies, and why do we need them ?** *IEEE Intelligent Systems*, 14, 20-26

Cheetham E, Arrowsmith I. (2005) **National Programme for Information Technology NP*f*IT SNOMED CT Post-coordination rules**. Draft guidance document v 0.98

Chute CG, Cohn SP, Campbell KE, Oliver DE, Campbell JR. (1996) **The content coverage of clinical classifications**. *J Am Med Inform Assoc*. 1996;3: 224-33.

Chute CG. (2000) **Clinical Classification and Terminology: Some History and Current Observations**. *J Am Med Inform Assoc*. 2000 May; 7 (3): 298–303

Cimino JJ, Hripcsak G, Johnson SB, Clayton PD. (1989) **Designing an Introspective, Multipurpose, Controlled Medical Vocabulary**. In: Kingsland LC (Ed). *Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care*. New York: IEEE Computer Society Press, 1989: 513-8.

Cimino JJ. (1998a) **Desiderata for Controlled Medical Vocabularies in the Twenty-First Century**. *Methods Inform Med* 1998; **37**: 394-403.

Cimino JJ (1998b) **Auditing the Unified Medical Language System with Semantic Methods.** *J Am Med Inform Assoc.* 1998:**5**(1): 41-51

Cimino JJ. (2000) **From data to knowledge through concept-oriented terminologies: experience with the Medical Entities Dictionary**. *Journal of the American Medical Informatics Association*; 2000;**7**:288-297

College of American Pathologists (2004). **Snomed Clinical Terms® Technical Reference Guide, January 2004 release**. *College of American Pathologists, Northfield, Ill*. Pp 12-15

Côté RA, Rothwell DJ, Palotay JL, Becket RS, Brochu L. (1993) **The Systematized Nomenclature of Human and Veterinary Medicine: SNOMED International**. 4 vols. Northfield, Ill: *College of American Pathologists*, 1993.

Gangemi A, Pisanelli DM, Steve G, (1999) **An Overview of the ONIONS project: Applying Ontologies to the Integration of Medical Terminologies**, *Data and Knowledge Engineering* , vol.31, pp. 183-220, 1999.

Gennari J, Musen MA, Fergerson RW, Grosso WE, Crubézy M, Eriksson H, Noy NF, Tu SW. (2002) **The Evolution of Protégé: An Environment for Knowledge-Based Systems Development.** Project Technical Report SMI-2002-0943. http://smi.stanford.edu/pubs/SMI_Abstracts/SMI-2002-0943.html (site visited 29 Jul 2004)

Goble C, Crowther P, Solomon D. (1994) **A medical terminology server**. In D. Karagiannis, (Ed) *Proc. 5th DEXA, volume LNCS Vol 856*, pages 661--670. Springer-Verlag, 1994.

Goble C, Haul C, Bechhofer S. (1996) **Describing and Classifying Multimedia using the Description Logic GRAIL**. *SPIE Conference on Storage and Retrieval of Still Image and Video IV* San Jose, 1996

Gómez-Pérez A, Fernández-López M, Corcho O. (2004) **Ontological Engineering**. Springer 2004 ISBN: 1-85233-551-3 pp 71-104 & http://www2.springeronline.com/sgw/cda/pageitems/document/cda_downloaddocument/0,11996,0-0-45-108719-0,00.pdf (visited 29 Jul 2004)

Hahn U, Schulz S, Romacker M. (1999a). **Partonomic reasoning as taxonomic reasoning in medicine**. *Proc.16th National Conf. on Artificial Intelligence & 11th Innovative Applications of Artificial Intelligence* (AAAI-99/IAAI-99), Orlando FL, AAAI Press/MIT Press.

Hahn U, Schulz S, Romacker M. (1999b). **Part-whole reasoning: a case study in medical ontology engineering**. *EEE Intelligent Systems & their applications* 14(5);: 59-67

Hahn U, Schulz S. (2003). **Towards A Broad-Coverage Biomedical Ontology Based On Description Logics**. *Pacific Symposium on Biocomputing* 2003: 577-588

Horrocks I, Patel-Schneider PF, van Harmelen F.(2003) **From SHIQ and RDF to OWL: The making of a web ontology language**. *Journal of Web Semantics*, 1(1):7-26, 2003

Lewis A. (2002) **Health informatics: information and communication**. *Advances in Psychiatric Treatment* (2002) 8: 165-171

Lindberg DA, Humphreys BL, McCray AT. (1993) **The Unified Medical Language System.** *Methods Inf Med* 1993;32(4):281-91.

Mahon WE, Scott DJ, Ansell G, Manson GL, Fraser R. (1967) **Hypersensitivity to pituitary snuff with miliary shadowing in the lungs** .*Thorax* 1967; 22: 13-20.

Mejino JLV, Agoncillo AV, Rickard KL, Rosse C. (2003) **Representing Complexity in Part-Whole Relationships within the Foundational Model of Anatomy**. *Annual Fall Symposium of American Medical Informatics Association*, Washington DC. Hanley & Belfus Inc. Philadelphia PA 450-454.

Minsky M. (1974) **A Framework for Representing Knowledge**. In P. Winston (Ed.) *The Psychology of Computer Vision.* McGraw-Hill, 1975

NLM National Library of Medicine (2001) **UMLS Knowledge Sources. 12th ed**. Bethesda (MD): *National Library of Medicine*; 2001

Noy NF. Hafner C. (1997) **The State of the Art in Ontology Design: A Survey and Comparative Review**. *AI Magazine*, 18 (3), Fall 1997, 53-74.

Noy NF, McGuiness DL. (2001) **Ontology Development 101: A Guide to Creating Your First Ontology**. Available as SMI technical report SMI-2001-0880 (2001) from Stanford Medical Informatics Group website (visited July 29 2004)

Noy NF, Musen MA. (2003) **The PROMPT Suite: Interactive Tools For Ontology Merging And Mapping.** *International Journal of Human-Computer Studies*, 59/6 pp. 983-1024. Also Available as SMI technical report SMI-2003-0973 from Stanford Medical Informatics Group website (visited July 29 2004)

Noy NF, Musen MA, Mejino JLV, Rosse C. (2004) **Pushing the Envelope: Challenges in a Frame-Based Representation of Human Anatomy**. *Data and Knowledge Engineering* 48:335-359

OPCS Office of Population Censuses and Surveys. (1990) **Classification of Surgical Operations and Procedures, 4th Revision**. London, England: *Her Majesty's Stationery Office*, 1990.

Okkes I, Jamoulle M, Lamberts H, Bentzen N.(2000) **ICPC-2-E: the electronic version of ICPC-2. Differences from the printed version and the consequences**. *Fam Pract* 2000; 17: 101–107.

Oliver DE. (1998). **Synchronization of diverging versions of a controlled medical terminology**. *Stanford, CA, Stanford University School of Medicine*, Technical Report SMI 98-0741

Oliver DE, Shahar Y (1999a) **Change Management of Shared and Local Health-Care Terminologies**. *IMIA Working Group 6 Meeting on Health Concept Representation and Natural Language Processing*, Phoenix, AZ, 1999.

Oliver DE, Shahar Y, Musen MA, Shortliffe EH. (1999b) **Representation of Change in Controlled Medical Terminologies**. *Artificial Intelligence in Medicine* 15(1):53-76, 1999.

OMIM (2004) *Online Mendelian Inheritance in Man* Website (visited 16 Apr 2004) http://www3.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM&cmd=Limits

O'Neil MJ, Payne C, Read JD. (1995) **Read Codes Version 3: A User Led Terminology**. *Methods Inf Med*. 1995;34: 187-92.

OpenGALEN [2002] **The GALEN Common Reference Model Version 6**. www.opengalen.org/open/crm visited March 3rd 2004

Pisanelli DM , Gangemi A , Steve G. (1998). **An Ontological Analysis of the UMLS Metathesaurus**, *JAMIA* vol. 5 S4, pp. 810-814, 1998

Price C, Brown PJB, Bentley TE, O'Neil MJ. (1998) **Exploring the ontology of surgical procedures in the Read Thesaurus**. *Methods Inf Med* 1998; 37: 420-5.

Rassinoux AM, Ruch P, Baud RH, Lovis C. (2000). **Semantic handling of medical compound words through sound analysis and generation processes**. *Proc AMIA Symp*. 2000;:675-9

Rector AL. (1995a) **Coordinating Taxonomies: Key To Re-usable Concept Representations**. *Fifth conference on Artificial Intelligence in Medicine Europe (AIME '95).* Pavia, Italy, Springer. 17-28.

Rector AL, Glowinski AJ, Nowlan A, Rossi Mori A. (1995b) **Medical-concept Models and Medical Records: An Approach Based on GALEN and PEN&PAD**. *J Am Med Inform Assoc*. 1995 2:1;19-35

Rector A, Rogers JE, Pole P. (1996) **The GALEN High Level Ontology**. *Fourteenth International Congress of the European Federation for Medical Informatics*, MIE-96, Copenhagen, Denmark

Rector A, Bechhofer S, Goble C, Horrocks I, Nowlan W, Solomon W. (1997) **The GRAIL concept modelling language for medical terminology**. *Artif Intell Med.* 1997 Feb;9(2):139-71.

Rector AL, Zanstra PE, Solomon WD, Rogers JE et el. (1998). **Reconciling Users' Needs and Formal Requirements: Issues in developing a Re-Usable Ontology for Medicine.** *IEEE Transactions on Information Technology in BioMedicine, Special issue on the EU Healthcare telematics programme*; Vol. 2 No.4 pp.229-242

Rodrigues JM, Trombert-Paviot B, Baud RH, Wagner J, Meusnier-Carriot F. (2000) Galen-In-Use : **Using artificial intelligence terminology tools to improve the linguistic coherence of a national coding system for surgical procedures**. *Medinfo*. 1998;9 Pt 1:623-7

Rogers JE, Solomon WD, Rector AL, Pole PM, Zanstra P, van der Haring E. (1997a) **Rubrics to Dissections to GRAIL to Classifications** C. Pappas et al. (Eds.) *Medical Informatics Europe '97*, IOS Press;:241-245.

Rogers J, Rector A. (1997b) **Terminological Systems: Bridging the Generation Gap**. *Proc AMIA Symp. 1997*;: 610-614

Rogers JE, Price C, Rector, AL, Solomon WD, Smejko N (1998) **Validating Clinical Terminology Structures: Integration and Cross-Validation of Read Thesaurus and GALEN.** *Annual Fall Symposium of American Medical Informatics Association*, Orlando FLA. Hanley & Belfus Inc. Philadelphia PA;:845-849

Rogers J.E., Rector, AL. (2000) **GALEN's Model of Parts and Wholes: Experience and Comparisons**. *Annual Fall Symposium of American Medical Informatics Association*, Los Angeles CA. Hanley & Belfus Inc. Philadelphia PA;:714-8

Rogers JE, Roberts A, Solomon WD, van der Haring E, Wroe CJ, Zanstra PE, Rector AL. (2001) **GALEN Ten Years On: Tasks and Supporting tools**. *Proceedings of MEDINFO2001*, V. Patel et al. (Eds) IOS Press;: 256-260

Rogers JE, Wroe C, Roberts A, Swallow A, Cantrill J, Stables D, Rector A. (2002) **Feasibility and Reliability of Computerised Review of Long Term Prescribing: Final Report (Supplementary)**. http://www.cs.man.ac.uk/mig/projects/old/prescribing-indicators/FinalReportRTF.zip (visited 02-02-05)

Rogers JE, Wroe CJ, Roberts A, Swallow A, Stables D, Cantrill JA, Rector AL (2003) **Automated quality checks on repeat prescribing**. *Br J Gen Pract* 2003; 53: 838-844

Rosse C, Shapiro L, Brinkley J. (1998) **The digital anatomist foundational model: Principles for de-fining and structuring its concept domain**. *Annual Fall Symposium of American Medical Informatics Association*, Orlando FLA. Hanley & Belfus Inc. Philadelphia PA;:820-824

Rossi Mori A, Gangemi A, Steve G, Consorti F, Galeazzi E. (1997) **An Ontological Analysis of Surgical Deeds**. *Proceedings of Artificial Intelligence in Medicine Europe (AIME-97)*, Grenoble, France, 1997

Rossi Mori A, Consorti F, Galeazzi E, (1998) **Standards to support development of terminological systems for healthcare telematics**. *Meth Inform Med* 1998; 37: 551-563

Schulz S, Hahn U, Romacker M. (2000) **Modelling Anatomical Spatial Relations with Description Logics**. *Annual Fall Symposium of American Medical Informatics Association*, Los Angeles CA. Hanley & Belfus Inc. Philadelphia PA;: 779 - 783

Sittig DF. (1994) **Grand challenges in medical informatics.** *Journal of the American Medical Informatics Association*, vol. 1, pp. 412-413, 1994.

Solomon WD, Wroe CJ, Rector AL, Rogers JE, Fistein JL, Johnson P. (1999) **A Reference Terminology for Drugs**. In *Annual Fall Symposium of American Medical Informatics Association*, Washington DC. Hanley & Belfus Inc. Philadelphia PA;:152-155

Solomon W.D., Roberts A., Rogers J.E., Wroe C.J. Rector, A.L. (2000) **Having our cake and eating it too: How the GALEN Intermediate Representation reconciles internal complexity with users' requirements for appropriateness and simplicity**. *Annual Fall Symposium of American Medical Informatics Association*, Los Angeles CA. Hanley & Belfus Inc. Philadelphia PA;:819-23

Spackman KA, Campbell KE, Côté RA. (1997) **SNOMED RT: A Reference terminology for healthcare**. In *Annual Fall Symposium of American Medical Informatics Association 2001*, Hanley & Belfus Inc. Philadelphia PA;:640-4

Spackman KA. (2001). **Normal forms for description logic expression of clinical concepts in SNOMED RT**. In *Annual Fall Symposium of American Medical Informatics Association 2001*, Hanley & Belfus Inc. Philadelphia PA;:627-631

Trombert-Paviot B, Rodrigues JM, Rogers JE, Baud R, van der Haring E, Rassinoux AM, Abrial V, Clavel L, Idir H. (2000) **GALEN: a third generation terminology tool to support a multipurpose national coding system for surgical procedures**. *Int J Med Inf*. 2000 Sep;58-59:71-85.

Wagner JC, Rogers JE, Baud RH, Scherrer J-R (1999) **Natural language generation of surgical procedures** *International Journal of Medical Informatics*; Vol. 53, NOS. 2,3;:175-192

Wanless D. (2002) **Securing our future health: taking a long-term view. Final report**. London: *HM Treasury*, 2002. http://www.hm-treasury.gov.uk/consultations_and_legislation/wanless/consult_wanless_index.cfm. Site visited August 2nd 2004

Wroe C, Stevens R, Goble CA, Ashburner M. (2003) **A Methodology to Migrate the Gene Ontology to a Description Logic Environment Using DAML+OIL** *Pacific Symposium on Biocomputing* 8:624-635(2003).

Wroe CJ, Solomon WD, Rector AL, Rogers JE (2000) **Inheritance of Drug Information** In *Annual Fall Symposium of American Medical Informatics Association*, Los Angeles CA. Hanley & Belfus Inc. Philadelphia PA;:1158

World Health Organization. (1975) **International Classification of Diseases, 9th Revision, Clinical Modifications**. Geneva, Switzerland: WHO, 1975.

World Health Organization.(1992) **International Statistical Classification of Diseases and Related Health Problems, 10th Revision**. Geneva, Switzerland: WHO, 1992.

Zanstra PE (1999) **Dutch Classification of Procedures for Oral Hygeinists (chapter 2)**. Personal communication

Zanstra PE (2004) **Experiences with GALEN-IN-USE in the Netherlands – maintenance of ICPM-DE.** Personal communication.

## Glossary

**Authoring Ontology** – *a controlled vocabulary of concepts and links whose content is optimised for a specific authoring task, and is independent from but linked to a* ***target ontology***

**Candidate Mapping Set –** *a set of alternative expressions in a target ontology, one or zero of may be considered the correct re-expression of an authoring ontology expression, but which is to be determined*

**Canonisation –** *the process of reducing a composed concept to its minimally non-redundant canonical form, and an important stage in the process of detecting compositions that are semantically equivalent: two compositions with the same canonical form are semantically equivalent*

**CCAM** - *Classification Commune des Acts Medicaux. A new classification of surgical procedures, mandate for use across the French healthcare sector in 2003/4*

**Composed Concept**- *a concept that not only can be expressed in terms of other concepts and links already in the overall concept system, but that is already so expressed. The subsequent behaviour of composed concepts in a system (e.g. for post-coordination) may be determined partly or wholly with reference to that atomised expression. Also known variously as decomposed, atomised, defined or dissected concept*

**Composition –** *see* ***composed concept***

**Concept** - *a basic unit of thought*

**Controlled Vocabulary** - *a terminology where user interaction with it is constrained in some way. The commonest constraint is not being able to add new terms to the vocabulary without the permission of the designated keeper of the term list. Other controls include identifying which term in a set of synonyms is the preferred one*

**Decomposable Descriptors** – *a primitive descriptor in an authoring ontology whose meaning can also be expressed as a composition in terms of other descriptors and links in the authoring ontology*

**Description Logics** – *are knowledge representation languages employing a subset of first order logic for tailored for expressing knowledge about concepts and concept hierarchies*

**Descriptor** – *a term for a concept in an authoring ontology*

**Descriptor Category** – *a supercategory assigned to a descriptor*

**End-user** - *the ultimate user (qv) for which something is intended*

**Hierarchy** - *any way of ordering terms and/or concepts, through a system of links, into a tree-like structure, where objects at the top of the tree are in some way more general or abstract than those lower down. The nature of each link between each level in the tree may be explicit or only implied, and more than one flavour of semantic link can be used to build the tree (in which case it may be called a mixed hierarchy).*

**ICT** – *Information Communication & Technology*

**Intermediate Representation** – *is any data structure constructed from input data to a program or algorithm and from which part or all of the output of the program or algorithm is constructed. Within this thesis, intermediate representation means a syntactically and ontologically simple representation within which users represent expressions that are subsequently transformed into a more complex target ontology.*

**International Classification Of Diseases (ICD)** – *a well-established terminology of medical disease, presented as a monoaxial classification, and maintained by the World Health Organisation. Widely used internationally for comparison of healthcare system performance and billing, it is now in its 10th revision*.

**Metamodel** – *an explicit specification of the preferred or allowed schema by which a model, or an ontology, may be extended*

**Monoaxial Hierarchy** - *a hierarchy constructed so that each node in the tree only ever has one parent (even if more than one parent would conceptually be more correct), though each node can still have more than one child*

**Multiaxial Hierachy** - *a hierarchy constructed such that each node in the tree may have more than one parent as well as more than one child. Being multiaxial does not automatically imply completeness of classification*

**Natural Kind –** *an entity, or concept, that can not be provided with a complete definition in terms of more basic concepts, in any ontology. The classic exemplar is an elephant: it is not possible to construct a list of characteristics that 'define' an elephant such that any object that possesses all those characteristics must, by definition be an elephant. You either are an elephant or you aren't one; there is no test for elephant-ness.*

**Ontology** – *a system of concepts and links, connected to a terminology*

***Open*GALEN** – *a registered Dutch not-for-profit organisation that provides the legal point of origin for opensource licensing of the OpenGALEN Common Reference Model*

**Primitive Concept**- *a concept that can not be completely expressed or defined in terms of other concepts and links already in the overall system of concepts. Therefore, in order to represent the entire meaning of that concept in the system, it is necessary that a new single entity is added to the system as a free-standing concept. Also known as an atom or atomic concept*

**Ontological Schema** – *a specification of a preferred general pattern of concepts and links to be used when forming a new composition representing a particular class of concepts*

**Rubric –** *a text phrase indicating the meaning of a code in a medical terminology. Thus, in the expression 'D245.0 Acute Thyroiditis' from the International Classification of Diseases version 9 (ICD9), the code is 'D245.0' and the text string ' Acute Thyroiditis' is the rubric denoting the intended meaning of that code, wherever it appears.*

**Semantic Link –** *a binary relationship between two concepts e.g. hasPart, is-A, hasLocation*

**Semantic Normalisation** – *a process by which expressions that follow a variety of ontological schema are re-expressed in terms of a preferred schema*

**SNOMED CT** – *a large proprietary ontology of medicine, owned by the College of American Pathologists and mandated for use in the United Kingdom NHS*

**Subsumption Hierarchy** - *a special case of a hierarchy in which the relationship between every parent-child pair in the tree is always and only that the child is a true 'kind-of' its parent. This is the definition of a true classification*

**Target Ontology** – *an ontology to which expressions written using authoring ontology are to be transformed e.g. by* ***semantic normalisation***

**Term -** *a text string; one of many possible words or phrases with which you usually label a concept*

**Terminology** - *a fixed list of many terms. Strictly speaking should exclude any link to a separate list of concepts although, commonly, objects that are called 'terminologies' usually do include such links either explicitly (e.g. as a link to an alphanumeric concept identifier) or implicitly (e.g. by superimposing some sort of hierarchy, which places the terms in an organisation that is clearly inspired by an analysis of the underlying concepts). Terminologies may include information about parts of speech for language analysis*

**Uniaxial Hierarchy** – *see monoaxial hierarchy*

**User** - *the person who uses a computer application, as opposed to those who developed or support it. The end-user may or may not know anything about computers, how they work, or what to do if something goes wrong. End-users do not usually have administrative responsibilities or privileges*

> **inexpert user** – *a user who has little or only limited understanding of ontology engineering either in general or with respect to the particular ontology they are using. One records information using one or more terms from an ontology but who does not know how or why that data is subsequently to be processed. Exemplar: most busy clinicians when recording the characteristics of the patient before them, or of a procedure just performed, for inclusion in a medical record.*

> **expert user** – *a user who has a good understanding of ontology engineering both in the general case and specifically with respect to all aspects of the particular implementation of the clinical ontology before them. They are aware of the choices they have when forming an expression using the ontology, and of how those choices affect its subsequent analysis. Exemplar: an ontology engineer tasked with maintaining or extending an ontology, and with devising acceptable user interfaces by which inexpert users might use the ontology*

> **intermediate level user** – *any regular user whose understanding is between expert and inexpert*

> **specialist user** – *a healthcare worker operating in a highly specialised area of medicine and correspondingly requiring the use of terms within the ontology that are unlikely to be used, or understood, by the generalist healthcare worker*

# Appendix One: Hierarchy of descriptor categories

Hierarchy of 115 descriptor categories; first level categories are displayed in bold.

```
Thing
 - Characteristic
     Age
   - Feature
       AbstractFeature
       BloodFeature
       Extent
       InvestigationResult
       Method
     Laterality
     Occupation
     Position
     Sex
     TemporalMarker
   RouteOrApproach
   SignOrSymptom
 - Process
     AbstractProcess
     BodyProcess
     Deed
   - HealthCareAct
       InvestigationAct
       TreatmentProcess
     NonHealthCareAct
 - Quantity
     Number
     OrdinalPosition
     Range
 - Substance
     BodySubstance
   - Chemical
     - Drug
 - Pathology
     BiochemicalPathology
     BreastPathology
     CardiovascularPathology
     DigestiveSystemPathology
     EndocrinePathology
   - GenitourinaryPathology
       FemaleGenitourinaryPathology
       MaleGenitourinaryPathology
     HaematologicalPathology
     HepatoBilaryPathology
     InfectiousPathology
     Lesion
     MetabolicPathology
     MusculoskeletalPathology
     NeurologicalPathology
     PsychologicalPathology
     RenalPathology
     ReproductiveSystemPathology
     RespiratorySystemPathology
     SensorySystemPathology
     SkinPathology
 - Structure
     AbstractStructure
   - Device
       Material
       Prosthesis
   - Organism
       OrganDonor
       Person
   - Anatomy
       AnatomicalConnection
       AnatomicalPart
       BodySpace
       BodySystem
     - BodySystemAnatomy
       - CardiovascularSystemAnatomy
         - BloodVessel
             AbdominalBloodVessel
             BloodVesselOfHeadOrNeck
             CentralBloodVessel
             IntrathoracicBloodVessel
             PelvicBloodVessel
             PeripheralBloodVessel
           PartOfHeart
         DermoidSystemAnatomy
         DigestiveSystemAnatomy
       - EndocrineSystemAnatomy
           Gland
       - GenitoUrinarySysAnatomy
           FemaleGenitalSystem
           LowerUrinaryTract
           MaleGenitalSystem
           UpperUrinaryTract
       - LymphoreticularSysAnatomy
           LymphNode
       - MusculoSkeletalSysAnatomy
           Bone
           Joint
       - NervousSystemAnatomy
           Nerve
         OrodentalSystemAnatomy
         RespiratorySystemAnatomy
       - SensorySystemAnatomy
           EarAnatomy
           EyeAnatomy
       Graft
 - Banned
     Ambiguous
     BadMapping
     BannedFromSchema
     PleaseDecompose
   Untyped
```

## Appendix Two: Task ontology Links and Candidate Link Mapping Sets

| TASK ONTOLOGY LINK | LEFT CONTEXT | RIGHT CONTEXT | CANDIDATE LINK MAPPING SET |
|---|---|---|---|
| ACTS_ON | DiseaseProcess | ClinicalSituation | *actsOn* |
| | DiseaseProcess | any | *actsOn ClinicalSituation isMainlyCharacterisedBy presence isExistenceOf* |
| | Depolarising | any | *isSpecificFunctionOf*<br>*isFunctionOf*<br>*LocativeAttribute* |
| | any | any | *actsMultiplyOn*<br>*actsSpecificallyOn*<br>*actsOn*<br>*LocativeAttribute* |
| ACTS_ON_1 | any | any | *actsMultiplyOn*<br>*actsOn*<br>*LocativeAttribute* |
| ACTS_ON_2 | any | any | *actsMultiplyOn*<br>*actsOn*<br>*LocativeAttribute* |
| ACTS_ON_3 | any | any | *actsMultiplyOn*<br>*actsOn*<br>*LocativeAttribute* |
| BYPASSES | any | any | *isSpecificPhysicalMeansOf Bypassing actsSpecificallyOn* |
| BY_APPROACH_TECHNIQUE | any | any | *hasSpecificSubprocess SurgicalApproaching hasSpecificSubprocess* |
| BY_MEANS_OF | any | any | *hasSpecificPhysicalMeans*<br>*hasPhysicalMeans* |
| BY_TECHNIQUE | any | any | *hasSpecificSubprocess* |

| | | | |
|---|---|---|---|
| | | | *hasSubprocess*<br>*hasSpecificPhysicalMeans*<br>*hasPhysicalMeans* |
| CAUSED_BY | InfectionLesion | MicroOrganism | *isSpecificConsequenceOf*<br>*isSyndromeElementOf*<br>*isConsequenceOf* |
| | InfectionLesion | Helminth | *isSpecificConsequenceOf*<br>*isSyndromeElementOf*<br>*isConsequenceOf* |
| | any | MicroOrganism | *isSpecificImmediateConsequenceOf InfectionLesion*<br>*isSpecificConsequenceOf*<br>*isConsequenceOf InfectionLesion isSpecificConsequenceOf*<br>*isSpecificConsequenceOf InfectionLesion isSpecificConsequenceOf* |
| | any | Helminth | *isSpecificImmediateConsequenceOf InfectionLesion*<br>*isSpecificConsequenceOf*<br>*isConsequenceOf InfectionLesion isSpecificConsequenceOf*<br>*isSpecificConsequenceOf InfectionLesion isSpecificConsequenceOf* |
| | any | any | *isSpecificImmediateConsequenceOf*<br>*isSpecificConsequenceOf*<br>*isSyndromeElementOf*<br>*isConsequenceOf* |
| CAUSES | any | any | *hasImmediateConsequence*<br>*hasSpecificConsequence*<br>*hasSpecificImmediateConsequence*<br>*hasSyndromeElement*<br>*hasConsequence* |
| CONNECTS | Joint | any | *hasSpecificStructuralComponent*<br>*hasStructuralComponent* |
| | Valve | any | *isPartitiveConnectionOf* |
| | BodyConnection | any | *isPartitiveConnectionOf* |
| | Junction | any | *isConnectionOf* |

| | | | |
|---|---|---|---|
| | Meatus | any | *isPartitiveConnectionOf* |
| | any | any | *isConnectionOf* |
| CONNECTS_1 | Joint | any | *hasSpecificStructuralComponent*<br>*hasStructuralComponent* |
| | Valve | any | *isPartitivelyFrom* |
| | BodyConnection | any | *isPartitivelyFrom* |
| | Junction | any | *isFrom* |
| | Meatus | any | *isPartitivelyFrom* |
| | any | any | *isFrom* |
| CONNECTS_2 | Joint | any | *hasSpecificStructuralComponent*<br>*hasStructuralComponent* |
| | Valve | any | *isPartitivelyTo* |
| | BodyConnection | any | *isPartitivelyTo* |
| | Junction | any | *isTo* |
| | Meatus | any | *isPartitivelyTo* |
| | any | any | *isTo* |
| CONNECTS_3 | Joint | any | *hasSpecificStructuralComponent*<br>*hasStructuralComponent* |
| | Valve | any | *isGammaPartitiveConnectionOf* |
| | BodyConnection | any | *isGammaPartitiveConnectionOf* |
| | Junction | any | *isGammaConnectionOf* |
| | Meatus | any | *isGammaPartitiveConnectionOf* |
| | any | any | *isGammaConnectionOf* |
| CONTAINS | any | any | *specificallyNonPartitivelyContains*<br>*specificallyPartitivelyContains*<br>*contains* |
| FOLLOWED_BY | any | any | *isSpecificallyFollowedBy*<br>*isFollowedBy*<br>*hasSpecificComplication*<br>*hasComplication* |

| | | | |
|---|---|---|---|
| FOLLOWS | any | any | *specificallyFollows*<br>*isSpecificComplicationOf*<br>*isComplicationOf*<br>*follows* |
| HAS_APPROACH | any | ArbitraryBodyConstruct | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans Route passesThrough* |
| | any | Route | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans* |
| | any | SurgicalOpening | *hasSpecificSubprocess SurgicalApproaching hasSurgicalOpenClosedness SurgicalOpenClosedness hasAbsoluteState* |
| | any | Device | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans* |
| | any | AnteroRetrogradeSelector | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans TranstubalRoute hasAnteroRetrogradeSelector* |
| | any | Process | *hasSpecificSubprocess SurgicalApproaching hasSubprocess* |
| | any | SurgicalIncision | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans Route passesThrough* |
| | any | TubularBodyStructure | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans Route passesThrough* |
| | any | LaminarPhysicalStructure | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans Route passesThrough* |
| | any | any | *no mapping* |
| HAS_DESTINATION | any | any | *isActedOnSpecificallyBy Transport hasSpecificConsequence Displacement isDisplacementTo* |
| HAS_DONOR | any | DonorStatus | *hasDonorOrigin* |
| | any | any | *no mapping* |
| HAS_EXTENT | any | Distribution | *hasDistribution* |
| | any | Completeness | *hasCompleteness* |
| | any | StructuralExtent | *hasStructuralExtent* |
| | any | SurgicalExtent | *hasSurgicalExtent* |
| | any | any | *no mapping* |
| HAS_FEATURE | | | |
| HAS_FUNCTION | PhysicalPresentation | any | *hasSpecificFunction* |

| | | | |
|---|---|---|---|
| | | | *hasFunction* |
| | any | any | *hasSpecificFunction*<br>*isSystemDefinedBy*<br>*hasFunction* |
| HAS_LATERALITY | any | BilateralUnilateralSelector | *hasBilateralUnilateralSelector* |
| | any | any | *hasLeftRightSelector*<br>*hasProximalDistalSelector*<br>*hasMedialLateralSelector*<br>*hasSuperiorInferiorSelector*<br>*hasAnteriorPosteriorSelector*<br>*hasPositionalSelector* |
| HAS_LOCATION | CellulitisLesion | Tissue | *hasSpecificLocation*<br>*involves* |
| | CellulitisLesion | BodyStructure | *hasSpecificLocation ConnectiveTissue makesUp*<br>*involves ConnectiveTissue makesUp* |
| | CellulitisLesion | any | *no mapping* |
| | ArteriosclerosisLesion | Arterial Structure | *hasSpecificLocation TunicaIntima isSpecificLayerOf* |
| | ArteriosclerosisLesion | any | *involves* |
| | InflammationLesion | any | *hasSpecificLocation*<br>*involves* |
| | NonnormalBodyConnection | any | *hasMultipleLocation*<br>*hasSpecificLocation*<br>*hasLocation*<br>*LocativeAttribute* |
| | BodyConnection | any | *hasPartitiveConnection* |
| | Prosthesis | any | *hasFunction GeneralisedProcess isFunctionOf* |
| | Congestion | VenousStructure | *isVolumeOf LiquidBlood isSpecificallyNonPartitivelyContainedIn* |
| | Congestion | ArterialStructure | *isVolumeOf LiquidBlood isSpecificallyNonPartitivelyContainedIn* |
| | Congestion | BloodVessel | *isVolumeOf LiquidBlood isSpecificallyNonPartitivelyContainedIn* |
| | Congestion | any | *isVolumeOf LiquidBlood isSpecificallyNonPartitivelyContainedIn*<br>*BloodVessel serves* |

| | | | |
|---|---|---|---|
| | Permeability | any | *isPermeabilityOf* |
| | Organomegaly | any | *isSizeOf* |
| | Spasm | TubularBodyStructure | *actsSpecificallyOn Diameter isDiameterOf* |
| | Spasm | any | *involves* |
| | DimensionChanging | TubularBodyStructure | *actsSpecificallyOn Diameter isDiameterOf* |
| | DimensionChanging | LinearBodyStructure | *actsSpecificallyOn Length isLengthOf* |
| | DimensionChanging | MuscleTissueStructure | *isFunctionOf Muscle isStructuralComponentOf* |
| | DimensionChanging | Abdomen | *actsSpecificallyOn Diameter isDiameterOf* |
| | DimensionChanging | LaminarPhysicalStructure | *actsSpecificallyOn VerticalDepth isVerticalDepthOf* |
| | DimensionChanging | SkinCovering | *actsSpecificallyOn VerticalDepth isVerticalDepthOf* |
| | DimensionChanging | SolidBosyStructure | *actsSpecificallyOn Volume isVolumeOf* |
| | DimensionChanging | any | *no mapping* |
| | Gas | any | *isNonPartitivelyContainedIn*<br>*isContainedIn* |
| | SkinTag | any | *hasSpecificProximity* |
| | Teratoma | BodyStructure | *hasUniqueAssociatedProcess NeoplasticProcess actsSpecificallyOn GermCell isStructuralComponentOf Tissue makesUp* |
| | Teratoma | any | *no mapping* |
| | Papilloma | BodyStructure | *involves NeoplasticProcess actsSpecificallyOn Cell isStructuralComponentOf Epithelium makesUp* |
| | Papilloma | any | *no mapping* |
| | NeoplasticLesion | Cell | *hasUniqueAssociatedProcess NeoplasticProcess actsSpecificallyOn* |
| | NeoplasticLesion | Tissue | *hasUniqueAssociatedProcess NeoplasticProcess actsSpecificallyOn Cell IsDivisionOf* |
| | NeoplasticLesion | BodyStructure | *hasUniqueAssociatedProcess NeoplasticProcess actsSpecificallyOn Cell isStructuralComponentOf Tissue makesUp* |
| | NeoplasticLesion | any | *hasUniqueAssociatedProcess NeoplasticProcess actsSpecificallyOn hasUniqueAssociatedProcess NeoplasticProcess LocativeAttribute* |
| | Pain | BodyStructure | *actsOn ExternalStimulusSensation actsOn PainSignal isConsequenceOf NerveConduction hasUniqueAssociatedDisplacement Displacement isDisplacementFrom* |

| | | | |
|---|---|---|---|
| | Pain | BodySystem | *actsOn ExternalStimulusSensation actsOn PainSignal isConsequenceOf NerveConduction hasUniqueAssociatedDisplacement Displacement isDisplacementFrom BodyStructure IsDivisionOf* |
| | Pain | Tissue | *actsOn ExternalStimulusSensation actsOn PainSignal isConsequenceOf NerveConduction hasUniqueAssociatedDisplacement Displacement isDisplacementFrom BodyStructure isMadeOf* |
| | Pain | any | *no mapping* |
| | TissueNecrosis | BodyStructure | *actsOn Tissue makesUp* |
| | TissueNecrosis | any | *no mapping* |
| | Feature | any | *isColourOf*<br>*isFeatureOf* |
| | Process | any | *actsMultiplyOn*<br>*actsSpecificallyOn*<br>*actsOn*<br>*isSpecificFunctionOf*<br>*isFunctionOf*<br>*LocativeAttribute* |
| | any | SurfaceVisibilityStatus | *hasSurfaceVisibility* |
| | any | BodySpace | *isPartitivelyContainedIn*<br>*isNonPartitivelyContainedIn*<br>*isContainedIn*<br>*hasMultipleLocation*<br>*hasSpecificLocation*<br>*hasLocation*<br>*LocativeAttribute* |
| | any | any | *hasMultipleLocation*<br>*hasSpecificLocation*<br>*hasLocation*<br>*LocativeAttribute* |
| HAS_LOCATION_1 | Joint | any | *isAlphaConnectionOf*<br>*isConnectionOf* |

| | | | |
|---|---|---|---|
| | Valve | any | *isAlphaPartitiveConnectionOf* |
| | BodyConnection | any | *isAlphaPartitiveConnectionOf* |
| | Junction | any | *isAlphaPartitiveConnectionOf* |
| | Meatus | any | *isAlphaPartitiveConnectionOf* |
| | Process | any | *actsMultiplyOn*<br>*actsOn*<br>*isSpecificFunctionOf*<br>*isFunctionOf*<br>*LocativeAttribute* |
| | InflammationLesion | any | *hasMultipleLocation* |
| | any | any | *hasAlphaConnection*<br>*hasMultipleLocation* |
| HAS_LOCATION_2 | Joint | any | *isConnectionOf*<br>*isBetaConnectionOf* |
| | Valve | any | *isBetaPartitiveConnectionOf* |
| | BodyConnection | any | *isBetaPartitiveConnectionOf* |
| | Junction | any | *isBetaPartitiveConnectionOf* |
| | Meatus | any | *isBetaPartitiveConnectionOf* |
| | Process | any | *actsMultiplyOn*<br>*actsOn*<br>*isSpecificFunctionOf*<br>*isFunctionOf*<br>*LocativeAttribute* |
| | InflammationLesion | any | *hasMultipleLocation* |
| | any | any | *isBetaConnectionOf*<br>*hasMultipleLocation* |
| HAS_LOCATION_3 | Joint | any | *isConnectionOf*<br>*isGammaConnectionOf* |
| | Valve | any | *isGammaPartitiveConnectionOf* |
| | BodyConnection | any | *isGammaPartitiveConnectionOf* |
| | Junction | any | *isGammaPartitiveConnectionOf* |

| | | | |
|---|---|---|---|
| | Meatus | any | *isGammaPartitiveConnectionOf* |
| | Process | any | *actsMultiplyOn*<br>*actsOn*<br>*isSpecificFunctionOf*<br>*isFunctionOf*<br>*LocativeAttribute* |
| | InflammationLesion | any | *hasMultipleLocation* |
| | any | any | *hasSpecificLocation*<br>*hasLocation*<br>*hasGammaConnection* |
| HAS_METHOD | Incising | SurgicalIncision | *hasSpecificImmediateConsequence* |
| | Incising | NAMEDEponymousMethod | *hasSubprocess* |
| | Incising | any | *no mapping* |
| | any | Energy | *hasPhysicalMeans*<br>*hasSpecificPhysicalMeans* |
| | any | ClinicalRole | *playsClinicalRole* |
| | any | any | *hasSubprocess*<br>*hasSpecificSubprocess* |
| HAS_NUMBER | NumericQuantity | any | *hasMagnitude* |
| | TemporalQuantity | any | *hasMagnitude* |
| | Duration | MagnitudeValueType | *hasQuantity TemporalIntervalValue hasMagnitude* |
| | Duration | any | *no mapping* |
| | any | UsefulnessState | *hasUsefulness Usefulness hasAbsoluteState* |
| | any | any | *hasNumber* |
| HAS_PATIENT | | | *isCharacterisedBy* |
| HAS_POSITION | any | Selector | *hasLeftRightSelector*<br>*hasBilateralUnilateralSelector*<br>*hasProximalDistalSelector*<br>*hasMedialLateralSelector*<br>*hasSuperiorInferiorSelector*<br>*hasAnteriorPosteriorSelector* |

| | | | |
|---|---|---|---|
| | | | *hasCentralPeripheralSelector hasSuperficialDeepSelector hasPositionalSelector* |
| | any | topInteger | *hasNumber* |
| | any | any | *hasPosition* |
| HAS_PROXIMITY | | | *hasSpecificProximity hasProximity* |
| HAS_QUANTITY | any | ImpreciseNumber | *hasNumber* |
| | any | Range | *hasRange* |
| | any | any | *hasQuantity* |
| HAS_ROUTE | any | ArbitraryBodyConstruct | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans Route passesThrough* |
| | any | Route | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans* |
| | any | SurgicalOpening | *hasSpecificSubprocess SurgicalApproaching hasSurgicalOpenClosedness SurgicalOpenClosedness hasAbsoluteState* |
| | any | Device | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans* |
| | any | AnteroRetrogradeSelector | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans TranstubalRoute hasAnteroRetrogradeSelector* |
| | any | TubularBodyStructure | *hasSpecificSubprocess SurgicalApproaching hasPhysicalMeans Route hasColinearityWith* |
| | any | any | *no mapping* |
| HAS_SIZE | any | any | *hasSize Size hasAbsoluteState* |
| HAS_TEMPORAL_MARKER | any | OrdinalPosition | *hasOrdinalPosition* |
| | any | SurgicalPrimarySecondaryStatus | *hasSurgicalPrimarySecondaryStatus* |
| | any | TimePeriod | *occursDuring* |
| | any | RevisionStatus | *hasRevisionStatus* |
| | any | ImmediacyStatus | *hasImmediacy Immediacy hasAbsoluteState* |
| | any | TimingAppropriateness | *hasTimingAppropriateness* |
| | any | LifecycleDiseaseStageState | *hasLifecycleDiseaseStage LifecycleDiseaseStage hasAbsoluteState* |

| | | | |
|---|---|---|---|
| | any | Permanence | *hasPermanence* |
| | any | TimeOfOccurence | *occursDuring* |
| | any | Duration | *hasDuration* |
| | any | Reversibility | *hasReversibility* |
| | any | StartingProcess | *isActedOnBy* |
| | any | TemporalPattern | *hasTemporalPattern* |
| | any | any | *no mapping* |
| IS_ACTED_ON_BY | any | Process | *isActedOnSpecificallyBy*<br>*isActedOnBy* |
| | any | any | *no mapping* |
| IS_BRANCH_OF | any | any | *isBranchOf* |
| IS_CONTAINED_IN | Tissue | HollowStructure | *isMixedThroughout*<br>*isContainedIn* |
| | Tissue | any | *specificallyMakesUp*<br>*makesUp* |
| | Muscle | any | *isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isSolidRegionOf* |
| | Tendon | any | *isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isSolidRegionOf* |
| | Nerve | any | *isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isSolidRegionOf* |
| | BloodVessel | any | *isSpecificStructuralComponentOf* |

| | | | |
|---|---|---|---|
| | | | *isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isSolidRegionOf* |
| | GeneralisedSubstance | GeneralisedSubstance | *isDissolvedWithin*<br>*isMixedThroughout* |
| | GeneralisedSubstance | any | *isNonPartitivelyContainedIn* |
| | any | HollowStructure | *isSpaceDefinedBy*<br>*isMixedThroughout*<br>*isContainedIn* |
| | any | Liquid | *isDissolvedWithin*<br>*isInSuspensionWithin* |
| | any | SolidStructure | *isMixedThroughout*<br>*isSolidRegionOf* |
| | any | any | *no mapping* |
| IS_FEATURE_OF | Concentration | any | *isConcentrationOf* |
| | Frequency | any | *isFrequencyOf* |
| | TemporalPattern | any | *isTemporalPatternOf* |
| | Size | any | *isSizeOf* |
| | PoisoningProcess | any | *actsSpecificallyOn* |
| | any | any | *isFeatureOf*<br>*involves* |
| IS_FUNCTION_OF | any | any | *isSpecificFunctionOf*<br>*isFunctionOf* |
| IS_LOCATION_OF | any | any | *isSpecificLocationOf*<br>*isMultipleLocationOf*<br>*isLocationOf*<br>*InverseLocativeAttribute* |
| IS_MADE_OF | any | any | *isSpecificallyMadeOf*<br>*isMadeOf* |
| IS_PART_OF | Integument | any | *isSpecificSurfaceDivisionOf* |

| | | | |
|---|---|---|---|
| | | | *isSurfaceDivisionOf*<br>*isSpecificLayerOf*<br>*isLayerOf*<br>*isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isArbitraryComponentOf*<br>*isSolidDivisionOf* |
| | SoftTissue | any | *specificallyMakesUp*<br>*makesUp*<br>*isArbitraryComponentOf* |
| | Tissue | any | *specificallyMakesUp*<br>*makesUp*<br>*isArbitraryComponentOf* |
| | BodyLayer | any | *isSpecificLayerOf*<br>*isLayerOf*<br>*isSolidRegionOf* |
| | SkinCovering | any | *isSpecificSurfaceDivisionOf*<br>*isSurfaceDivisionOf*<br>*IsDivisionOf* |
| | any | LinearPhysicalStructure | *isSpecificLinearDivisionOf*<br>*isLinearDivisionOf*<br>*isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isArbitraryComponentOf*<br>*IsDivisionOf* |
| | any | LaminarPhysicalStructure | *isSpecificSurfaceDivisionOf*<br>*isSurfaceDivisionOf*<br>*isSpecificLayerOf* |

| | | | |
|---|---|---|---|
| | | | *isLayerOf*<br>*isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isArbitraryComponentOf*<br>*IsDivisionOf* |
| | any | any | *isSpecificStructuralComponentOf*<br>*isStructuralComponentOf*<br>*isSpecificSolidDivisionOf*<br>*isSolidDivisionOf*<br>*isArbitraryComponentOf*<br>*IsDivisionOf* |
| IS_SERVED_BY | any | any | *isSpecificallyServedBy*<br>*isServedBy* |
| MOTIVATED_BY | any | any | *hasSpecificGoal*<br>*hasGoal* |
| MOTIVATED_OVERALL_BY | any | any | *hasSpecificGoal*<br>*hasGoal* |
| OCCURS_DURING | any | any | *occursDuring* |
| OR_MAIN | any | any | *isMainlyCharacterisedBy performance isEnactmentOf* |
| REPLACES | ImplantableDevice | any | *hasFunction GeneralisedProcess isFunctionOf* |
| | any | any | *no mapping* |
| SERVES | any | any | *specificallyServes*<br>*serves* |
| TO_ACHIEVE | any | any | *hasSpecificGoal*<br>*hasGoal*<br>*isSpecificallyToDetermine*<br>*isToDetermine* |
| TO_ACHIEVE_OVERALL | any | any | *hasSpecificGoal*<br>*hasGoal* |

| | | | |
|---|---|---|---|
| | | | *isSpecificallyToDetermine*<br>*isToDetermine* |
| WITH | any | any | *isCharacterisedBy performance isEnactmentOf* |
| WITHOUT | any | any | *isCharacterisedBy nonPerformance isEnactmentOf* |
| WITH_GUIDANCE_BY | any | any | *hasSubprocess Guiding hasPhysicalMeans* |
| WITH_MAIN | any | any | *isMainlyCharacterisedBy performance isEnactmentOf* |
| WITH_OPTIONALLY | any | any | *isCharacterisedBy performance isEnactmentOf* |

# Appendix Three: Merged Classification of Surgical Procedures (Abridged Fragment)

2999 of 3617 dissections obtained, representing the meaning of rubrics relating to cardiovascular surgery from four different classifications of surgical procedures, were expressed in intermediate representation and using the surgical procedure task ontology (see section 11.3.2).

They were expanded to the target ontology, and a subsumption hierarchy of all the dissections computed, based only on an examination of the explicit semantics. The hierarchy was multiaxial: a concept in the target ontology may have more than one parent.

The full hierarchy occupies 9251 lines and 165 pages in the format presented here; only the first 200 lines is presented. For clarity of reading, this has been further abridged by deleting 2 lines where the rubric length forced a line wrap in the text. This, only 198 lines of text are presented.

The complete computed hierarchy has a maximum of 12 levels between the top level concept ‘cardiovascular procedure’ and the most distant leaf concepts. The abridged fragment shown here has 7.

Many concepts in the hierarchy have more than one parent, and appear more than once in the displayed hierarchy; this why a classification of only 2999 concepts requires 8452 lines to display.

On many occasions two different rubric dissections expand to the same target ontology concept. Where this occurs, the target ontology concept appears with the number of different dissections that relate to it shown in brackets after.

- 'Cardiovascular procedure'
  - 'Patch enlargement of conduit'
  - 'Excision of carotid body tumour'
  - 'Vascular cannulation procedure'
    - 'Vascular cannula adjustment'
    - 'Flushing cannula'
    - 'Vascular cannula removal'
      - 'Removal of intravenous cannula'
        - 'Removal of central venous line'
    - 'Vascular cannula unblockage'
  - 'operation on vessels'
    - 'Operation för truncus arteriosus communis'
    - 'percutaneous transluminal angioplasty'
      - 'transluminal coronary angioplasty'
      - 'single vessel percutaneous transluminal coronary angioplasty [PTCA] without mention of thrombolytic agent'
        - 'single vessel percutaneous transluminal coronary angioplasty [PTCA] with thrombolytic agent'
      - 'multiple vessel percutaneous transluminal coronary angioplasty [PTCA] performed during single operative episode'
    - 'correctie van truncus arteriosus'
      - 'Repair of common arterial trunk'
    - 'incision, excision, and occlusion of vessels'
      - 'Excision of artery'
        - 'Attention to arteriovenous shunt'
        - 'Excision of arteriovenous fistula'
        - 'Prosthetic graft thrombectomy'
        - 'Excision of intracranial aneurysm'
          - 'Excision of aneurysm of cerebral artery'
        - 'Resection of pulmonary artery' <3>
          - 'Exérèse d''un obstacle endo-artériel pulmonaire, sans CEC, par thoracotomie'
          - 'Exérèse d''un obstacle endo-artériel pulmonaire, avec CEC, par thoracotomie'
          - 'Keuhkovaltimon sisäkalvon ja kroonisen tukoksen poisto'

'Resection of systemic to pulmonary shunt aneurysm'
'Exérèse d''un obstacle endo-artériel pulmonaire, par voie endovasculaire'
'Resection and end-to-end anastomosis of pulmonary trunk'
'Excision of vegetations from pulmonary trunk'
'Resection of thoracic aorta'
'Resection of coarctation'
'Extended resection of thoracic aorta'
'Radically extended resection of thoracic aorta'
'Excision of ductus arteriosus'
'Excision of vegetations from ductus arteriosus'
'Resection of aneurysm of ductus arteriosus'
'Excision of arteriovenous graft'
'Prosthetic graft thrombectomy'
'Excision of vein'
'Harvesting of vein'
'Attention to arteriovenous shunt'
'Excision of arteriovenous fistula'
'Onttolaskimon aukaisu rintaontelossa'
'Excision of varicose vein'
'Excision of portal vein'
'Partial excision of portal vein'
'Multiple phlebectomy'
'Local excision of varicose vein of leg'
'Excision of perforating vein'
'Exérèse d''une tumeur des vaisseaux pulmonaires, par thoracotomie'
'Annan reparation av truncus arteriosus communis'
'Bypass of renal artery'
'Bypass graft of the descending thoracic aorta'
'other operations on vessels'
'operaties van bloedvoorziening van hart, niet gespecificeerd' <2>
'herstel van aneurysma van bloedvaten van hart'
'repair of aneurysm of coronary vessel' <2>
'Áëëåò ÅðåìâÜóåéò Êáñäéáêþí Áããåßùí'
'Áëëåò ÅðåìâÜóåéò Êáñäéáêþí Áããåßùí'
'Overige operaties van bloedvoorziening van hart'
'Coronary artery operation' <2>
'transluminal coronary angioplasty'
'transluminal coronary angioplasty' <2>
'removal of coronary artery obstruction' <2>
'open chest coronary artery angioplasty'
'single vessel percutaneous transluminal coronary angioplasty [PTCA] without mention of thrombolytic agent'
'single vessel percutaneous transluminal coronary angioplasty [PTCA] with thrombolytic agent'
'other specified removal of coronary artery obstruction'
'percutaneous transluminal angioplasty'
'single vessel percutaneous transluminal coronary angioplasty [PTCA] without mention of thrombolytic agent'
'single vessel percutaneous transluminal coronary angioplasty [PTCA] with thrombolytic agent'
'Correction d''une anomalie de naissance d''une artère coronaire, par thoracotomie, avec CEC'
'Plastie d''élargissement du tronc commun de l''artère coronaire gauche, par thoracotomie, avec CEC'
'Rekanalisering av koronarartärer'
'Annan rekanalisering av koronarartär(er)'
'Angioplasty of coronary artery'
'Percutaneous transluminal balloon angioplasty of one coronary artery' <3>
'Percutaneous transluminal balloon angioplasty of bypass graft of coronary artery'
'Open angioplasty of coronary artery'
'Exploration of coronary artery with inspection only' <3>
'Rekonstruktion av koronarartär vid skada'
'Rekonstruktion vid anomal koronarartäravgång'
'Annan operation vid anomal koronarartär'
'Repair of arteriovenous fistula of coronary artery'
'Open intracoronary repair of coronary artery fistula'
'repair of aneurysm of coronary vessel' <2>
'Annan rekonstruktion av koronarartär vid skada'
'Sutur av koronarartär med bypass'

'Sutur av koronarartär med patch'
'Sutur av koronarartär'
'Ligatur av anomal koronarartär'
'Slutning av koronarfistel'
'Obliteration av koronarfistel'
'Fermeture d''une fistule coronaro-cardiaque, par thoracotomie, sans CEC'
'Fermeture d''une fistule coronaro-cardiaque, par thoracotomie, avec CEC'
'Fermeture d''une fistule coronaro-cardiaque, par thoracotomie, sans CEC'
'Fermeture d''une fistule coronaro-cardiaque, par thoracotomie, avec CEC'
'Áïñôïóôåöáíéáßá ÐáñÜêáìøç ÔåóóÜñùí Þ ÐåñéóóïôÝñùí Óôåöáíéáßùí Áñôçñéþí??????' <4>
'Coronary artery bypass graft x 1' <7>
'Saphenous vein graft bypass of coronary artery' <5>
'Prosthetic replacement of four or more coronary arteries' <5>
'Allograft replacement of two coronary arteries' <5>
'Coronary artery graft placement'
'reconstructie van coronairostium'
'Rekonstruktion des Koronarostiums'
'a. mammaria - a. coronaria anastomose, enkelvoudig' <2>
'a. mammaria - a. coronaria anastomose, meervoudig'
'a. mammaria - a. coronaria anastomose, meervoudig'
'Transponering av koronarartär'
'Reimplantation of coronary artery'
'Coronary interposition technique' <2>
'Vidgning av koronartär med patch'
'Anastomos mellan arteria gastroepiploica och koronarartärer'
'Annan anastomosoperation mellan arteria gastroepiploica och koronarartärer'
'Anastomos mellan arteria gastroepiploica och koronarartärer'
'Anastomos mellan arteria mammaria interna och koronarartärer '
'Annan anastomosoperation mellar arteria mammaria interna och koronarartärer'
'Anastomos mellan arteria mammaria och koronarartärer'
'Anastomoser mellan arteria mammaria interna och koronarartärer bilateralt'
'Double anastomosis of mammary arteries to coronary arteries'
'Single anastomosis of mammary artery to left anterior descending coronary artery'
'Koronar bypass med fritt artärtransplantat'
'Andra koronara bypassoperationer med fritt artärtransplantat'
'Koronar bypass med fritt artärtransplantat med arteria gastroepiploica'
'Koronar bypass med fritt artärtransplantat med arteria mammaria'
'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat)' <2>
'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) dreifach und mehr'
'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) einfach'
'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) zweifach'
'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) dreifach und mehr'
'Anlegen eines aortokoronaren Bypass' <2>
'Anlegen eines aortokoronaren Bypass mit Endarterektomie (TEA)'
'Anlegen eines aortokoronaren Bypass mit Aortenklappenersatz'
'Anlegen eines aortokoronaren Bypass mit Aortenklappenersatz durch Allotransplantat'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Allotransplantat'
'Anlegen eines aortokoronaren Bypass mit Aortenklappenersatz durch Kunstprothese'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Kunstprothese'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Allotransplantat'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Kunstprothese'
'Anlegen eines aortokoronaren Bypass mit Mitralklappenersatz'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Allotransplantat'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Kunstprothese'
'Anlegen eines aortokoronaren Bypass mit Mitralklappenersatz durch Allotransplantat'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Allotransplantat'
'Anlegen eines aortokoronaren Bypass mit Mitralklappenersatz durch Kunstprothese'
'Anlegen eines aortokoronaren Bypass mit Aortenklappen- und Mitralklappenersatz durch Kunstprothese'
'Aortocoronaire bypass, enkelvoudig' <2>
'Ligatur och bypass-rekonstruktion vid anomal koronarartär'
'Overige gespecificeerde aortocoronaire bypass'

    'Aortocoronaire bypass, meervoudig'
    'Aortocoronaire bypass, meervoudig'
'Anlegen eines aortokoronaren Bypass onA mit sonstiger autogener Arterie'
    'Anlegen eines aortokoronaren Bypass onA mit autogener Vene und Arterie'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit autogener Vene und Arterie'
        'Anlegen eines zweifachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines dreifachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines vierfachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines fünffachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit autogener Vene und Arterie'
    'Anlegen eines aortokoronaren Bypass onA mit A. mammaria interna'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit A. mammaria interna'
        'Anlegen eines einfachen aortokoronaren Bypass mit A. mammaria interna'
        'Anlegen eines zweifachen aortokoronaren Bypass mit A. mammaria interna'
        'Anlegen eines dreifachen aortokoronaren Bypass mit A. mammaria interna'
        'Anlegen eines vierfachen aortokoronaren Bypass mit A. mammaria interna'
        'Anlegen eines fünffachen aortokoronaren Bypass mit A. mammaria interna'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit A. mammaria interna'
    'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit sonstiger autogener Arterie'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit autogener Vene und Arterie'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit A. mammaria interna'
    'Anlegen eines einfachen aortokoronaren Bypass mit sonstiger autogener Arterie'
        'Anlegen eines einfachen aortokoronaren Bypass mit A. mammaria interna'
    'Anlegen eines zweifachen aortokoronaren Bypass mit sonstiger autogener Arterie'
        'Anlegen eines zweifachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines zweifachen aortokoronaren Bypass mit A. mammaria interna'
    'Anlegen eines dreifachen aortokoronaren Bypass mit sonstiger autogener Arterie'
        'Anlegen eines dreifachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines dreifachen aortokoronaren Bypass mit A. mammaria interna'
    'Anlegen eines vierfachen aortokoronaren Bypass mit sonstiger autogener Arterie'
        'Anlegen eines vierfachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines vierfachen aortokoronaren Bypass mit A. mammaria interna'
    'Anlegen eines fünffachen aortokoronaren Bypass mit sonstiger autogener Arterie'
        'Anlegen eines fünffachen aortokoronaren Bypass mit autogener Vene und Arterie'
        'Anlegen eines fünffachen aortokoronaren Bypass mit A. mammaria interna'
    'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit sonstiger autogener Arterie'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit autogener Vene und Arterie'
        'Anlegen eines sechsfachen aortokoronaren Bypass und mehr mit A. mammaria interna'
'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat)' <2>
    'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) dreifach und mehr'
    'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) einfach'
    'Revaskularisation mit freiem A. mammaria interna-Transplantat (IMA-Transplantat) zweifach'

ETC.

# Appendix Four: Validity analysis for experiments of opportunity

The table below summarises the results of a critical appraisal of the internal validity of the experiments of opportunity (Chapter 11)

| THREAT | DEFINITION | Time To Train | Productivity | Reproducibility | SemanticUtility | Cross Validation |
|---|---|---|---|---|---|---|
| **Confounding** | *Extraneous variable changes systematically with the independent variable, and therefore prevents us from inferring a causal effect between the independent and dependent variables* | None identified | None identified | Low scores may be a function of lack of experience, and have improved with greater experience | None identified | None identified |
| **History** | *Unanticipated events occurring while the experiment is in progress, but affect the experimental groups differently* | short period; GRAIL training occurs over longer period, during which model may have improved | None identified | Subjects involved in experiment for too short a time period for this effect to be relevant | Subjects involved in experiment for too short a time period for this effect to be relevant | Not Relevant - Only one subject |
| **Maturation** | *Processes within the subjects operating as a function of time.* | Possibility of fatigue in GRAIL group making that training process even more prolonged | Most productive IR authors benefitted from improvements and practise | Subjects involved in experiment for too short a time period for this effect to be relevant | Subjects involved in experiment for too short a time period for this effect to be relevant | Not Relevant - Only one subject |
| **Non-random sampling** | *Biases resulting from selection or creation of groups that are not equivalent.* | Although different groups, nature of bias more likely to result in underestimation of difference in time to train | Although different groups, nature of bias more likely to result in underestimation of difference in time to train | Although different groups, nature of bias more likely to result in underestimation of difference in time to train | CCAM dissections subject to higher QA process | Not Relevant - Only one subject |
| **Selection-maturation interaction** | *Nonequivalent groups of different ages creating a bias such that selection and maturation interact.* | More likely to result in underestimation of differences between two interventions | Perhaps small group of adept authors produced all the dissections | No maturation effect | No maturation effect | No maturation effect or sampling bias |
| **Testing (practice effect)** | *The effect of taking one test upon the scores of a subsequent test.* | Minimal prior exposure to GRAIL for some of the IR training groups. | No pretest | No pretest | No pretest | No pretest |
| **Instrumentation (instrument drift, fatigue effect)** | *An effect due to changes in a measuring instrument or changes in observers or scorers.* | Single measurement applied to 2 groups not to individuals. One subjective judgement: productive = trained | Constant measuring instrument | Constant measuring instrument | Possible fatigue of CCAM raters | Categorisation of differences in hierarchies is time consuming and subject to fatigue |
| **Statistical regression** | *An effect operating where subjects selected on the basis of extreme scores regress toward the mean of that variable.* | Not relevant | Not relevant | Not relevant | Not relevant | Not Relevant - Only one subject |
| **Experimental mortality** | *The differential loss of subjects from one or more groups on a nonrandom basis.* | No loss of subjects occurred | Not all authors were equally productive. | No loss of subjects occurred | No loss of subjects occurred | Not Relevant - Only one subject |
| **Experimenter bias** | *A bias caused by the expectations of either the experimenter or the subjects or both.* | Subjective score of when trained | Automated scoring | Automated scoring | Single rater, unclear criteria for measuring 'utility'. Right answer for wrong reasons | Single rater bias |