Granularity Scale & Collectivity: When size does and doesn't matter

Alan Rector[1] Jeremy Roger[1] & Thomas Bittner[2]
[1]Department of Computer Science, University of Manchester, Manchester M13 9PL, UK
[2] IFOMIS Universityof Saarland, Germany
{rector|jrogers}@cs.man.ac.uk thomas.bittner@ifomis.uni-saarland.de

Abstract: Bridging levels of "granularity" and "scale" are frequently cited as key problems for biomedical informatics. However, detailed accounts of what is meant by these terms are sparse in the literature. We argue for distinguishing two notions: "size range", which deals with physical size, and "collectivity", which deals with aggregations of individuals into collections which have emergent properties and effects. We further distinguish these notions from "specialisation", "degree of detail", "density" and "connectivity." We argue that the notion of "collectivity" – molecules in water, cells in tissues, people in crowds, stars in galaxies – has been neglected but is a key to representing biological notions, that it is a pervasive notion across size ranges – micro, macro, cosmological, etc – and that it provides an account of a number of troublesome issues including the most important cases of when the part-whole relation is, or is not, transitive. Although examples are taken from biomedicine, we believe these notions to have wider application.

1. Introduction

It is a truism that a major challenge for bioinformatics is to bridge levels of granularity and scale, from molecular, to cellular, to organ, to organism, to ecology. However, it is rarely made clear exactly what is meant by "granularity" or "scale" or what the consequences are of differences in granularity and scale for which any explanation must account.

This paper argues that it would be clearer to distinguish unambiguously two dimensions. We term these two dimensions "collectivity" and "size range" despite the risk of adding yet further neologisms to the field[1]. The basic notion that we put forward is that entities considered individually at one level are considered as collectives with emergent properties at the next level – *e.g.* collectives grains of sand form a beach, collectives of stars form galaxies, collectives of cells form tissues. In general, for convenience, we shall refer to the "grains" of a "collective" and correspondingly to "granular parts". Alternatively we might refer to collectives as "emergent wholes". The notion of "collective" used here is similar to that of "groups" used by Artale [1, 2] and by Winston & Odell [12, 29], but neither they nor Padgham & Lambrix [13] investigate it extensively, although Winston and Odell put forward analogous reasoning for why the feet of geese do not form part of a flock. The notion of "granular parts" is also hinted at by the distinction between "constituent parts" and other forms of part-whole relation in the Foundational Model of Anatomy [19]. However, we suggest that this is a seriously under investigated aspect of representation and can be used to account for several important phenomena.

Our fundamental contention is that there are properties and effects of collectives that are emergent - *i.e.* that are not generally predictable from the properties of the individual grains and that therefore must be attributed to the collective as a whole. Some properties only make sense of a collective – *e.g.* the pattern of a tiling or the arrangement of cells in a tissue. It makes no sense to speak of the pattern of a single tile or the alignment of a single cell. In other cases the emergent properties are distinct from that of the grains even if related, *e.g.* the mood of a crowd is distinct from that of its constituent individuals, a beach has area; and galaxies have mass, tissues have strength, grow, etc. The fundamental point is that information is conveyed about and determined by the collective rather than its grains. Here we take as our prototype an classic hour glass. In some idealised world it might be possible to determine how long it took the sand to pass through an hour glass by examining the glass and the individual grains of sand and their initial configuration. In practice, no one would attempt such a feat . The time required for the sand to flow through the hourglass is a collective property of the sand in relation to the specific hour glass that contains it.

[1] Although we would prefer to reserve the term "granularity" for the notion here termed "collectivity", the term "granularity" has become so overloaded with different meanings in different fields that we reluctantly opt for a neologism rather than risk further confusion and controversy. "Scale" conforms more closely to "size". However, to avoid confusion we have likewise been explicit in this paper and used the term "size range".

Although the phenomenon of emergence is widely applicable, our fundamental motivations are biological. We seek:

1. To distinguish the way in which, for example, a cell is part of the body from the way a finger is part of the body – specifically that the loss of a cell not diminish the body whereas the loss of a finger does;
2. To use this to motivate an important criteria for when transitive relations should, or should not, be treated as transitive;
3. To represent loosely repetitive patterns in tissues – that the "cells in the mucosa are aligned" – and more generally patterns and other emergent properties of collectives;
4. To deal with the collective effects of cells, organelles, etc. – *e.g.* the process of secretion and regulation of hormones by the cells of endocrine organs or the collective strength of muscles made up of indeterminate numbers of muscle fibres.

More often than not, collectives are themselves portions of larger things – hence our reluctance to use the phrase "emergent whole". Galaxies are more than mere collectives of stars; tissues are more than collectives of cells; even a beach is more than a collective of sand. However in each case there is a sense in which we can treat the collective as a portion of the larger entity. If we can have independently measurable commensurable features for both the collective and the larger entity, we can speak of the proportion of the greater entity formed by the collective of grains just as we would speak of the proportion of water or salt in an amount of sea water, collagen in tissue, or the proportion of the mass of galaxy comprised of the visible stars.

Our goal is a set of broadly applicable principles. The paper follows broadly the intent and lessons, although not always the execution, of the *Open*GALEN Common Reference Model[15, 18]. As an illustration we present this paper and an implementation in the framework of OWL-DL[2]. However, the issues are general and independent of any particular implementation. http://www.cs.man.ac.uk/~rector/ontologies/collectivity-demo.owl and http://www.cs.man.ac.uk/~rector/ontologies/collectivity-demo-classified.owl

1.1 Outline of approach

We distinguish two notions often confused under the heading of "granularity":

Collectivity – *Grains vs Collectives* – the degree of collectivisation, *e.g.* with respect to water filling a lake, the relation 'filling' is to the water as, amongst other things, a collective of water molecules, not to the individual molecules themselves.

Size range– *Large vs Small* – the size of an object with respect to the phenomena that affect it, *e.g.* quantum scales of distance or relativistic scales of speed. However, less extreme differences in scale can have major effects. Surface tension is critical at the scale of a water flea's interaction with water but not at that for a human.

Furthermore we distinguish two types of parthood as subrelations of the basic mereological part-whole relation related to collectivity.

Granular parthood - *e.g.* the relation of the cells in the finger of the skin to the finger, in which the grains are parts of the whole by virtue of being grains in a collective that is part of the whole, and in which removing one granular part does not necessarily damage or diminish the whole..

Determinate parthood – *e.g.* the relation of the finger to the hand, in which the parts are directly part of the whole, and in which removing one determinate part necessarily damages or diminishes the whole.

Our major contentions are that:

1. Collectives

- 1a) "Collectives" are made up of "grains"
- 1b) "Collectives" are not mathematical sets – their identity is not determined by their membership.
- 1c) Being a "collective" ("collectivity") is independent of size
- 1d) There are emergent effects and characteristics of collectives as a whole not determinable from the characteristics of their grains.

2. Granular and determinate parts

- 2a) "Determinate parthood" is transitive; granular parthood is not.

[2] An OWL-DL ontologies illustrating the principles can be found at http://www.cs.man.ac.uk/~rector/ontologies/collectivity

2b) Damage to "determinate parts" is damage to the whole; damage to granular parts is not. More generally, many effects on determinate parts have corresponding or related effect on the whole; this is rarely true for granular parts.

2c) A collective that is a "determinate part" of a whole remains a part of that whole regardless of the loss or gain of gains. (The issue of "empty collectives" is dealt with in Section 4.3.2)

There are two criteria of distinguishing granular and determinate parthood. The first is ontological; the second informational or cognitive:

1. Ontological – Whether there is a fixed, or nearly fixed number of parts – e.g. fingers of the hand, chambers of the heart, or wheels of a car such that there can be a notion of a single one being missing – or whether, by contrast, the number of parts is indeterminate – e.g. cells in the skin of the hand, red cells in blood, or rubber molecules in the tread of the tyre of the wheel of the car.

2. Informational – Whether the information to be conveyed pertains to the individual parts – e.g. the laceration to the fourth finger – or to the collective of parts – e.g. the arrangement of the cells in the skin of the finger.

These two criteria do not always correspond. In particular, we sometimes wish to refer to the collective properties of a fixed number of entities – *i.e.* to treat what are ontologically determinate parts informationally as being granular parts. We will return to this issue towards the end of this paper after the basic notions are established. (See Section 4.3.)

1.2 Other notions sometimes labelled "granularity"

We further distinguish "collectivity" and "size range" from four other notions with which they may be confused, and which other researchers have referred to as 'granularity' in addressing mereological issues [4, 23].

Specialisation – *Category vs kind*– the usual notion of "is-kind-of", *e.g.* that "mammal" is a generalisation including, amongst other things, dogs and elephants. Sometimes also labelled 'abstraction'.

Degree of detail – The amount of information represented about each entity, regardless of its level of specialisation. Crudely in an ontology represented in OWL, the number of axioms and restrictions concerning each entity.

Density – The number of semantically 'similar' concepts in a particular conceptual region. How "bushy" the subsumption graph is. High local density in an ontology usually co-occurs with high levels of specialisation and degree of detail, but in two different ontologies of the same overall depth, in a particular section one may find the same two categories separated by different numbers of intervening categories or possessing very different numbers of sibling categories.

Connectivity– The number of entities connected directly and indirectly to a given entity either through generalisation/specialisation or by other properties.

These notions will not be further discussed in this paper.

1.3 Criteria for success of the proposed approach

Our purpose in developing "ontologies" is to support information systems. The test of their adequacy is whether they can effectively represent the entities about which information must be communicated so that that communication is "faithful". This focuses our interest as much on the relations ("properties" in OWL; "roles" in most DLs; "attributes" in GRAIL) as on the entities related.

Our specific application is biomedicine, so that we will test our solution primarily with respect to well known biomedical knowledge resources including the Digital Anatomist Foundational Model of Anatomy [11, 19], the Open Biology Ontology (OBO) and more particularly the Gene Ontology [3, 27, 30] and *Open*GALEN [16-18]. Secondarily Johansson's paper provides a series of examples against which to test the notions put forward here {{ref}}.

More specifically, we seek a set of patterns and relations in OWL that are adequate to capture four notions and exclude their counterexamples:

1. Relation of faults and procedures to parts and wholes – *e.g.* that the disease of the part is a disease of the whole and certain procedure – *e.g.* repair – on a part is a procedures on the whole.
2. Patterns and characteristics of collectives *e.g.* that the cells of the intestine are typically aligned (with each other) or that the cells in bone are sparsely separated.

3. Collective or emergent effects of collectives, *e.g.* the total secretion of enzymes by the liver cells or the total force exerted by the cells in a muscle.
4. Persistent vs non-persistent parthood – *e.g.* that "Jack's finger" will still be referred to as "Jack's finger" even when it is severed from his hand. However, insulin secreted by a cell is not considered to be a part of that cell.

1.4 Independence of Collectivity and Size

1.4.1 "Collectivity" does not depend on physical size

Necessarily, grains are not physically larger than the collective of which they are members (except perhaps for some odd quantum cases). There is a tendency to talk of things as being at, for example, the "cellular level" or the "organ level" or the "subatomic" level, etc. However, such talk indicates a general tendency and conflates size and collectivity. Hairs are macroscopic entities of the same general size as small organs, yet most of the information we have to convey about hairs concerns the collective "hair" rather than individual "hairs". Sperm and eggs are both cells, but much of what we have to say about eggs pertains to individual eggs, whereas much more that we have to convey about sperm concern the collective, although we need a mechanism to cross levels of collectivity to speak of a single sperm fertilizing a single egg. Indeed, one of the issues in fertility research is to determine which factors depend on the collective of sperm and the fluids in which they are swimming, and which depend on the individual sperm cells themselves.

To extend the biological examples, within the cells there are both individual entities, such as the nucleus, and collectives such as mitochondria and chloroplasts. Within the nucleus there are a precisely countable number of chromosomes, which are usually treated individually, but uncountable collectives of macromolecules. In some circumstances, the same entities at the same size may be sometimes treated collectively and sometimes individually. The rigidity and shape of a chromosome are a collective property of the DNA molecules (and other supporting structures) that make it up; the "genes"[3] inheritance of characteristics is usually a feature of discrete sequences of base pairs (with complex dependence on context and regulation).

1.4.2 "Size range" does not depend on collectivity

There are many effects that are specific to physical size, distance, speed, density, etc. Most obviously, quantum and relativistic effects are generally relevant only for the very small, very large or the very rapidly moving[4]. Closer to everyday life, the surface tension and vortex effects that govern insects ability to fly, walk on walls, skim over water, etc. are highly relevant at their size range but almost irrelevant at the size of most mammals. Within biology, chemical bonding, van der Waals forces, other electrostatic forces, and many other effects are important at one physical size range but not at another. When they are relevant, they are relevant for both for individuals and collectives that conform to that size range.

2. Semi Formal Presentation

2.1 Notation

Neither of the XML concrete syntaxes for OWL is compact or readable enough for easy use in a paper, and even the official abstract syntax becomes bulky and difficult to read when there is any significant embedding. This paper therefore adopts the following conventions for a simplified syntax. In addition, this allows us to introduce syntax for two constructs not currently standard in OWL although likely in subsequent versions and supported by known description logics, qualified cardinality restrictions ("exactly 1") and general inclusion axioms ("propagates via").[5]

1. Subset and subproperties are indicated by indentation made explicit by '–'s. Where only two are involved a simple arrow is used, *e.g.* "Heart → Organ" for "Heart is a kind of Organ".
2. Properties are presented with their inverse separated by a slash, property modifiers – transitive, symmetric, functional, etc. are listed to the right, as in Table 1 above.
3. The OWL key words are adapted to a concise infix notation as shown in Table 2.
4. In complex expressions, indentation will be used rather than bracketing wherever the meaning is clear
5. Schema variables will be given in italics san serif in place of parts of names, *e.g.* *X,Y,Z* as in part_of *X*. Schema variables range over OWL class names.

---

[3] The definition of what constitutes a gene is problematic, at least in eukaryotic cells, but that need not concern us here.
[4] relative to the observer of course.
[5] "exactly n" and "propagates via" are special cases of the more general constructs known as "qualified cardinality restrictiosn" and "role inclusion axioms" respectively.

Abbreviated Informal	OWL Abstract Syntax	DL German Syntax
A AND B	intersectionOf(A B)	A ⊓ B
A OR B	unionOf(A B)	A ⊔ B
NOT A	complementOf(A)	¬A
has_property SOME C	restriction(has_property someValuesFrom(C))	∃ has_property . C
has_property ONLY C	restriction(has_property allValuesFrom(C))	∀ has_property . C
has_property EXACTLY-n C	restriction(has_property cardinality(1, C)[6]	∃!1 has_property C
B → A	subclassOf(B A)	B ⊑ A
A — B — — C	subclassOf(B A) subclassOf(C B)	B ⊑ A C ⊑ B
A ≅ B	equivalentClass(A B)	A ≅ B
P_1 propagates_via P_2	not applicable	$R_1 \circ R_2 \rightarrow R_2$

Table 2: Concise infix notation used in this paper with equivalents in OWL and standard DL notation

2.2 Basic properties and entities

We shall assume an upper ontology similar to DOLCE that includes the notions of "*Physical entity*" that includes both material entities, *i.e. "Physical objects"* and non-material entities such as holes and lines. We shall assume a distinction between "*Physical objects*" such as fingers and statues and "*Amounts of matter*" such as skin and clay as in DOLCE. However, we leave open until later the discussion of the controversy between cognitivist and realist over the nature of the link between physical objects and amounts of matter. However, we will take it that it is useful to distinguish to subproperties of gross-parthood, one between instances of "*Physical objects*" which we shall term "*determinate parthood*" and the other between instances of "*Amounts of matter*" which we shall call "*ingredienthood*". The common parent of *"determinate parthood"* and *"ingredienthood"* we shall term *"gross parthood"*.

The basic notions to be captured are that:

1. The parent part-whole relation, *"is part of"/"has part"* corresponds to the basic mereological relation and both it and the two subrelations *"is determinate part of"/"has determinate part"* and *"is ingredient of"/ "has ingredient"* satisfy the usual mereological axioms, *i.e.* that they are reflexive, transitive, and antisymmetric, i.e. everything is a part of itself[7]; parts of parts are parts of wholes; and nothing is a part of a part of itself.
2. The "*is grain of/has grain"* relation is irreflexive, antisymmetric, and non-transitive, *i.e.* that nothing can be a grain of itself; that grains of grains of a collective are not grains of the collective; and that nothing can be a grain of a grain of itself.
3. That the *"is grain of"* relation propagates via the "*is part of*" relation, *i.e.* that if an entity is a grain of collective that is part of a whole then that entity is also part of the whole. More formally: "*is grain of* ∘ *is part of* → *is part of*".

2.3 Approximation in OWL

Owl supports transitive properties (relations) and the notion of subproperties.It lacks the notion of propagates_via but this can be approximated by use of the role hierarchy by making is_grain_of a subproperty of is_part_of, which is a slightly stronger condition. This has the undesirable consequence that grains, which are analogous to members of a set, count as parts of the collective, which runs counter to the usual usage in for example Winston and Odell [12, 29]. However, in practice this causes little difficulty (See ???) It also lacks the the notions of reflexive, irreflexive and antisymmetric the consequences of which are discussed in ??? and ???. However, despite

[6] Not supported in the current OWL standard although proposed for extensions

[7] The usual formulation of the axiom the part-whole axioms in mereology is in terms of what is here called "reflexive parthood". "Proper parthood" is then defined as a part of the whole that is not equal to the whole.

these limitations, a sufficient representation of part-whole relations to cover the positive inferences from the more general axioms is possible. A demonstration following the development in this paper is available.[8]

Using the conventions described in The basic property hierarchy for the OWL approximation is presented in table 3a using the conventions described in .2.1 above. The additional properties of is_gross_part and is_ingredient_of are explained in 2.4.3 below.

Property	Transitive	Domain/ Range	Comments
is_part_of / has_part	Y	Physical_entity / Physical_entity	The generic part-whole relation Reflexive & antisymmetric properties not captured directly in OWL.
— is_gross_part_of/ has_gross_part	Y	Physical_entity / Physical_entity	The common parent of measurable portions and determinate parts.
— — is_determinate_part_of / has_determinate_part	Y	Physical_entity / Physical_entity	The relation between determinate parts and wholes, *e.g.* fingers and hands.
— — is_ingredient_of / has_ingredient	Y	Amount_of_matter/ Amount_of_matter	Further details deferred – see 2.4.3
— — …			See Section 4.3
— is_grain_of / has_grain	N	Physical_object / Collective	The relation between a grain and the collective. (NB: The status of collectives of non-material entities is left open at this point)

Table 3a

The corresponding entity hierarchy is described in table 3b.

Class	Use in this paper	Comments
Physical_entity	Domain/range of is_part_of and is_determinate_part_of	Common ancestor of all physical entities
— Physical_object	Domain for is_grain_of	Material physical entities
— Non_material_object	Excluded from domain for is_grain_of	Non-material physical entities, *e.g.* holes, lines, etc.
— Amount_of_matter	range for is_ingredient_of	Amounts of "stuff", roughly corresponding to mass nouns. (NB the Relation between Physical_object and Amount_of_matter depends on the debate between the cognitivist & realist stance and is not directly relevant to this paper. See 4.3)
— — Mixture	domain for is_ingredient_of	Abstract including solutions, suspensions etc.
Collective	Range of is_grain_of	Whether or not Collectives are considered physical and whether or not they are to be disjoint from Physical_object, is deferred. See 4.4.2

Table 3b

2.4 Basic schemas

2.4.1 Defining collectives

Normally collectives are defined using universal restrictions, *i.e.* collectives are defined following a schema where the upper case Italics indicates schema variables that range of class names.

Collective_of_X ≅ Collective AND has_grain ONLY *X*

[8] http://www.cs.man.ac.uk/~rector/ontologies/collectivity-demo.owl
http://www.cs.man.ac.uk/~rector/ontologies/collectivity-demo-classified.owl

There are two consequences of this schema:

1. Empty collectives are allowed. This is convenient when we want to talk about concentrations of zero or things that are empty or missing. We can define Non_empty_collective in the obvious way as:
   Collective AND has_grain SOME Anything[9]
2. A collective must be defined in terms of the most general type of which its grains are individuals, which may be a disjunction. (However, any collective defined in terms of a disjunction should be viewed with suspicion, as it is more likely to be more appropriately represented as a mixture.)

2.4.2 Reflexive parts

Because reflexive properties cannot be expressed directly in OWL it is useful to define two schemas for reflexive parthood:

Reflexive_part_of_*X* ≅ *X* OR is_part_of SOME *X*
Reflexive_gross_part_of_*X* ≅ *X* OR is_gross_part_of SOME *X*
Reflexive_determinate_part_of_*X* ≅ *X* OR is_determinate_part_of SOME *X*

Which schema is appropriate depends on the requirement. In simple "part explosions" only determinate parts are required. If both portions and determinate parts are required (see "Mixtures" below), then Reflexive_gross_part_of_*X* is required. If all parts, as in the digital anatomist or classic mereology, the most general notion of Reflexive_part_of_*X* is required.

In effect, these schemas allow much of the effect of reflexivity to be transferred to the entities from properties. It also makes it easy to express the notions related to SEP Triples [8, 9, 20]. Which is used in which circumstances depends on whether the class to be defined is to include only parts within the given level of collectivity or whether it is to include parts across levels of collectivity.

2.4.3 Mixtures

Collectives and reflexive parts provide the basic mechanisms required, but almost all interesting cases involving collectives involve not just one collective but mixtures of collectives with other collectives and/or amounts of matter.

The relation between different collectives and amounts of matter in a mixture we term is_ingredient_of / has_ingredient, which is transitive. We place is_ingredient_of as a sibling of is_determinate_part_of and under is_gross_part_of because some classes and queries to be formulated include both, *e.g.* the gross parts of a car include both wheels and rubber; the gross parts of the arm include both the biceps and fascia.

The basic schema for mixtures is:

Mixture_of_X_1_and_X_2_and_…_and_X_n ≅
  Mixture AND has_ingredient SOME X_1 AND has_ingredient SOME X_2 AND…AND has_ingredient SOME X_n

Formally, the domain constraint on is_ingredient_of guarantees in this simple version that anything that has portions is a mixture. However, for clarity it is better to include Mixture as a conjunct explicitly. A Mixture can be defined by being an amount of matter that has ingredients[10].

Mixture ≅ Amount_of_matter AND has_ingredient SOME Amount_of_matter

For example, one might represent that blood is a mixture of – amongst other things – plasma, red cells and white cells:

```
Amount_of_blood →
  Mixture AND
  has_ingredient SOME Amount_of_plasma AND
  has_ingredient SOME (Collective AND has_grain ONLY White_blood_cell) AND
  has_ingredient SOME (Collective AND has_grain ONLY Red_blood_cell)
```

Note that, in common with most biomedical definitions, we have not closed the list of ingredients in the mixture. There is nothing in the above axiom to imply that blood does not contain other things, only that it does contain the ingredients mentioned. Nor have we made this a definition; it does not imply that any mixture of plasma, red cells and white cells is blood.

Strictly speaking we normally want the mixture to consist of exactly one amount or collective of each kind, so the above should strictly be:

```
Amount_of_blood →
  Mixture AND
```

[9] owl:Thing
[10] A given ontology might, for consistency, wish to insist that all amounts of matter were mixtures. That issue is deferred here

has_ingredient exactly-1 Amount_of_plasma AND
has_ingredient exactly-1 (Collective AND has_grain ONLY White_blood_cell) AND
has_ingredient exactly-1 (Collective AND has_grain ONLY Red_blood_cell)

When reducing such expressions when they appear in the rest of this paper to standard OWL, the "exactly-1" must be replaced by "SOME".

2.4.4 Proportions

Because the relative amounts in a mixture are so often important, and because the means of determining relative amounts vary – *e.g.* by weight, volume, activity, etc.– in a binary relational formalisms such as RDF or OWL, it is often appropriate to reify the relation has_ingredient, *i.e.* to re-represent it as a class, which we shall term Proportion.

The basic schema is:

Mixture_of_X_1_and_X_2_and_…_and_X_n ≘
Mixture AND
has_proportion EXACTLY-1 (Proportion AND is_of_ingredient SOME X_1 AND has_percentage VALUE p_1) AND
has_proportion EXACTLY-1 (Proportion AND is_of_ingredient SOME X_2 AND has_percentage VALUE p_2) AND
…AND
has_proportion EXACTLY-1(Proportion AND is_of_ingredient SOME X_n AND has_percentage VALUE p_n)

The example of blood extended to this schema becomes::

Amount_of_blood →:
Mixture AND
has_proportion EXACTLY-1 (Proportion AND is_of_ingredient SOME Plasma
AND has_percentage VALUE p_1) AND
has_proportion EXACTLY-1 (Proportion AND is_of_ingredient (Collective AND has_grain ONLY White_blood_cell)
AND has_percentage VALUE p_2)) AND
has_proportion EXACTLY-1 (Proportion AND is_of_ingredient (Collective AND has_grain ONLY Red_blood_cell)
AND has_percentage VALUE p_3))

In principle the p_i can be either numeric values or ranges, precise or fuzzy. In practice as OWL exists today, pending technical issues around user defined XML datatypes and issues with classifiers, they are confined to integers. Additional qualifiers on Proportion include the mechanism – *e.g.* by mass, volume or some other measure.

2.4.5 Allowing proportions and simple ingredients to coexist

It is possible to allow the two patterns – for simple ingredients and for proportions of ingredients to coexist if we arrange the property hierarchy as shown in Table 4. Care must be taken with the domain and range constraints.

Property	Transitive	Domain/ Range	Comments
is_ingredient_of / **has_ingredien**t	Y	Amount_of_matter / Amount_of_matter OR Proportion_of_matter	Ingredients of ingredients are ingredients of the whole
— of_mixture / **has_proportion**	N	Proportion / Amount_of_matter	Proportions of proportions are not proportions of the whole.
— is_proportion / **is_of_ingredient**	N	Amount_of_matter/ Proportion	

Table 4: Property hierarchy reconciling ingredients and proportions. Note that the relevant properties are the inverses (given in bold) to remain consistent with Table 3a.

The fact that proportions of proportions are not themselves the same proportions of the whole is reflected in the facts that has_proportion and is_of_ingredient are not transitive. Since the percentages attached to each proportion will have to be recalculated at each step down the chain, the relationship is not simply transitive but follows a more complex rule. That rule must be handled by reasoning mechanisms outside the scope of OWL or most other ontology languages. What can be captured in OWL is that ingredients of ingredients, by either mechanism, are ingredients of the whole, which is represented by the fact that the parent property, has_ingredient, is transitive.

2.4.6 Characteristics of collectives and patterns of collectives in mixtures

Characteristics of the collective itself. Members of a collective often have collective characteristics, *e.g.* that the cells of a tissue are aligned or that the atoms of a crystal form a particular lattice structure, that neurons fire synchronously or asynchronously, etc. Such characteristics pertain to the collective; they make no sense if applied to its individual grains. Nor do these characteristics depend on the collective's relation to any other entity of which it may be a part. Furthermore, just as collective's identity is not extensional, their characteristics are not universal over their extensions, *i.e.* they can be considered true even if they do not apply to every member of the collective, *e.g.* a crystal will still be said to have a particular alignment even if it has flaws.[11] Hence it is appropriate to represent such characteristics as properties of the collective, *e.g.*

```
Collective AND
  has_grain ONLY Cell AND
  has_pattern SOME Alignment[12]
```

Characteristics of the collective in relation to other entities. On the other hand, there are characteristics that pertain to the relation between a collective and other items in a mixture – *e.g.* that cells are suspended in plasma or that the water and alcohol molecules are intermingled in a miscible liquid. In this case the properties are best represented as additional characteristics of the Proportion, *e.g.*

```
Amount_of_blood →:
  Mixture AND
  has_proportion EXACTLY-1 (Proportion AND is_of_ingredient SOME Plasma
                                    AND has_percentage VALUE p1
                                    AND has_role SOME Suspensor_role) AND
  has_proportion EXACTLY-1 (Proportion AND is_of_ingredient (Collective AND has_grain ONLY White_blood_cell)
                                    AND has_percentage VALUE p2
                                    AND has_role SOME Suspensee_role)) AND
  has_proportion EXACTLY-1 (Proportion AND is_of_ingredient (Collective AND has_grain ONLY Red_blood_cell)
                                    AND has_percentage VALUE p3
                                    AND has_role SOME Suspensee_role))
```

The form above is chosen over a representation in the spirit of "Blood is plasma in which are suspended red and white cells" since this variant has the undesired implication that "Blood is a kind of Plasma" – a statement that is clearly false.

However, the form is limited in complex cases, *e.g.* where one might want to say that the water plays the role of solute for sodium but suspensor for cells, because there are no "role value maps" in OWL, *i.e.* there is no way to say either that it is the same water that is the suspensor for the cells and the solute for sodium. Again, not all the semantics can be captured in OWL, but what can be is sufficient for most practical applications.

Note that for this purpose it would be necessary to reify Proportions even in a formalism supporting n-ary relations. Since there are an arbitrary number of ways by which a given proportion might be characterised, any fixed arity relation capturing only a fixed number of such characteristics would almost certainly become inadequate as the ontology evolved.

2.4.7 Emergent Effects of Collectives

Each cell in most glands secretes a portion of the hormone or other substance secreted; each granule in a synapse releases a portion of the neurotransmitter that fires the synapse; each muscle fibre exerts a measurable force when it contracts; each strand of a cable has its own tensile strength. However, in each of these cases, the information of interest is almost always about the collective effect. The collective effect is a function of the individual effects, but may be so highly non-linear that it would be difficult to predict, even if all the individual effects were known. The function is also highly variable for different collectives. Consider for example the different relationships between the collective strength of chains with respect to their links and of cables with respect to their strands (and how the strands are arranged within the cable). Furthermore, in many cases such as cables minor changes in the effects of individual grains are irrelevant provided the collective effect remains unchanged. Indeed the dynamics and relation of such individual effects to the collective effect is an important topic of systems biology.

Emergent effects are dealt with straightforwardly by schemas such as:

(*Collective_X* AND has_grain ONLY *Entity_Y*)→ has_effect *Effect_Z*

A simple example would be:

```
(Collective AND has_grain ONLY Pancreatic_eyelet_cell) →
                  has_effect SOME (Secretion AND has_target SOME Insulin
                                        AND has_rate VALUE r)
```

where r is a quantity with a numeric magnitude and units of type volume per unit time or weight per unit time.

The concern is not with the rate of secretion of individual eyelet cells, or indeed of individual eyelets, but with the rate of secretion of the entire collective of eyelet cells.

[11] How completely such characteristics are true belongs with a discussion of fuzziness or precision and is beyond the scope of this paper.

[12] For a discussion of the use of classes in value partitions, see Semantic Web Best Practice Committee's note http://www.w3.org/TR/swbp-specified-values/

3. Use and consequences

3.1 Propagation of faults

In general, faults propagate only across gross parthood, *e.g.* disorder to the liver is usually considered as a disorder of the digestive system, body, etc. whereas we would not normally consider a disorder of a single liver cell in this way. The liver cell is a grain of a collective that forms part of the liver (whether or not via a constitutes relation). Likewise, while we would consider a disorder of the metabolism of all, or a significant portion of, red cells – *e.g.* sickle cell anaemia – as a disorder of blood, we would not consider a disorder of the metabolism of a single red cell as a disorder of blood. Indeed, since both liver and red blood cells constantly die and are replenished, were we to consider the state of individual cells, all organisms would suffer from liver and blood disorders, which is clearly nonsense.

Hence the schema for disorders is normally

Disorder_of_*X* ≙ Disorder has_locus SOME Reflexive_gross_part_of_*X*.

Where has_locus is the property linking disorders to their anatomical or functional "site". This captures the above two examples and analogous cases while excluding the case of damage to individual cells, etc.

Note that the issue of propagation across boundaries of collectivity is orthogonal to the issue of whether the disorder applies to the entity as a whole or to its reflexive parts. There are disorders – gastritis, inflammatory bowel disease, septicaemia (infection of the blood), etc. that refer to the whole taken as a whole rather than its parts. For these cases, the appropriate schema excludes all parts, whether gross or granular:

Disorder_of_*X*_as_a_whole ≙ Disorder has_locus *X*.

Furthermore, the issue is not dependent on size. Analogies can be found at all physical size ranges.

3.2 Transitivity of part-whole relations

The issue of propagation of faults is closely related to the issue of transitivity of part-whole relations. Effectively, the argument in this paper is that most cases where the part-whole relation is not transitive involve transitions across levels of collectivity, *i.e.* involve chains of reasoning including the is_grain_of relation which is not transitive. Confusion arises because our usual language does not distinguish the broader is_part_of relation from its more specialised subrelations, here termed is_gross_part_of and is_grain_of. The is_grain_of relation marks boundaries between levels of collectivity, or what are often called levels of granularity. However, we argue that the critical issue of whether the part-whole relation is transitive is not one of physical size, *per se*, but of whether or not the relation deals with collectives or individuals.

As a partial validation of this view, consider the list of cases provided by Johansson of anomalies where the part-whole relation is not considered to be transitive. Table 5 lists these issues and whether or not they are accounted for by the distinction between gross parthood and granular parthood.

1.	A handle, x, can be part of a door, y, and a door can be part of a house, z, but yet the handle need not be (is not) a part of the house. That is, 'x < y' and 'y < z' but '¬(x < z)'. (Of course, 'part' cannot here and elsewhere in the list be synonymous with 'spatial part'.)	*Not accounted for*: confusion of direct and indirect partonomy.
2.	A platoon is part of a company, and a company is part of a battalion, but yet a platoon is not part of a battalion.	*Not accounted for*: confusion of direct and indirect partonomy
3.	A cell's nucleus is part of a cell, and a cell is part of an organ, but yet the nucleus is not part of an organ.	*Accounted for*. Cells are granular parts of the organ, not gross parts.
4.	Heart cells are parts of the heart, and the heart is part of the circulatory system, but yet the cells are not parts of the circulatory system.	*Accounted for*. Cells are granular parts of the Heart, not gross parts.
5.	Person P is part (member) of the football club FC, and FC is part (member) of the National Association of Football Clubs, NAFC, but yet P is not a part (member) of NAFC.	*Accounted for*. The person is a grain (member) of the football club, not a part of it and, similarly, the football club is a grain (member) of the association.
6.	Simpson's finger is part of Simpson, and Simpson is part of the Philosophy Department, but yet Simpson's finger is not part of the Philosophy Department.	*Accounted for*. Simpson is a member (grain) of the philosophy department (or possibly in some other relation to it), but not "part" of it in the sense used here.
7.	Hydrogen is part of water, and water is part of our	*Accounted for* and a false example. Hydrogen is not part of

	cooling system, but yet hydrogen is not part of our cooling system.	water. Hydrogen atoms are part of water molecules, collectives of which constitute water used in the cooling system
8.	Cellulose is part of trees, and trees are parts of forests, but yet cellulose is not part of forests.	*Accounted for*. Trees are grains for forests.
9.	A handle is part of a spoon, and a spoon is part of eating soup, but yet a handle is not part of eating soup.	*Not accounted for*; A different issue. Continuants and occurents cannot be parts of each other for reasons not discussed in this paper.
10.	This shard was part of a plate, and the plate was part of a dinner service, but yet the shard was not part of the dinner service.	*Probably accounted for*. Leaving aside issues of time and the status of the shard prior to the breakable, the plate might be considered part of a collective.
11.	This tree is part of the Black forest, and the Black forest is part of Germany, but yet this tree is not part of Germany.	*Accounted f*or. Trees are grains of forests. (Also the notion of geographical parthood might be treated differently by some authors)
12.	These grains of sand are part of the beach, and the beach is part of the island, but yet these grains of sand are not part of the island	*Accounted for*. The grains of sand are grains of the beach.

Table 5: Johansson's list of cases for non-transitivity of part-whole relations

We would argue that cases 4)-8) and 11)-12) are clearly accounted for by the distinction between gross and granular parthood.

Of the remainder, for cases 1 and 2, Johansson puts forward the argument that there is a narrow, non-transitive subproperty of parthood, which we usually term "direct parthood", that is not transitive and that the problem arises out of a confusion of the direct subproperty and the parent transitive property. He draws support for this distinction from Simons [21]and Casati & Varzi [5]. This seems to us entirely correct. However, Johansson also includes case 3 in this category. We would argue that it was better accounted for by the distinction between gross and granular parthood.

Case 9) Johannson explains by noting that two notions of parthood being used are fundamentally different. Again we would agree, a point we would signify by the incompatibility of parthood for occurents and continuants, *i.e.* "eating" and "spoon".

Case 10 is dealt with cursorily but seems clearly to raise a host of questions, not least whether the shard *per se* existed prior to the shattering of the plate.

Thus of Johansson's twelve cases, at least eight can be accounted for by the distinction between determinate and granular parthood; two are accounted for by confusion of direct and transitive parthood. Of the remainder, one is accounted for by the distinction between parthood for occurents and continuants (which is outside the scope of this paper), and one (case 10) is decidedly peculiar on several fronts. Johansson's thesis is that intransitive parthood predicates are not binary predicates. Our argument is that for the cases where it applies, the distinction between gross and granular parthood – *i.e.* between parthood without levels of collectivity rather than across them – is simpler, easier to apply, and arguably more fundamental.

3.3 Persistent and non-persistent part-hood

It is a general pattern that things continue to be spoken of as 'parts' even after they have been separated from the whole. Thus we speak of "John's finger" even after it has been amputated. Even if it has failed to develop we may speak of it as being absent. By contrast, we do not speak of the secretions from an individual cell as remaining part of that cell, although we might speak of them as being from one or more organs or even parts of organs. Hence we might legitimately seek to distinguish, for example, testosterone produced by the adrenal gland from testosterone produced by the testes, or oestrogen from the ovary from oestrogen from adipose tissue. However, we would be unlikely to distinguish testosterone originating from individual cells. Likewise, although we might talk of "cells from John's liver" following a biopsy, we would be unlikely to consider these as parts of John or his liver, present or missing, in the same sense as we would his amputated finger.

As in the above cases, we would argue that "persistent parthood" is something that pertains to things arising from gross parts but not from granular parts. This point, we accept, remains somewhat speculative and requires further investigation. (Note, we find "persistent parthood" as used here closer to common clinical usage than "permanent parthood" as advocated in Smith *et al.*) [25].

4. Discussion

4.1 Biomedical cases

4.1.1 Tissues and substances

A major motivation for the current work is to deal with specific problems in the adequate representation of the biological notions of tissue and substance. In this formulation both are "mixtures" some of whose "ingredients" are "collectives".[13] The schemas offered here provide both for properties that are intrinsic to the collective – *e.g.* arrangements and patterns – and for properties of the relation of the collective to the rest of the mixture, *e.g.* the proportion, distribution, etc. The claim is not that tissues *are* collectives, but that they are best viewed as amounts of matter some of whose ingredients are collectives. Much work remains to be done to describe patterns within tissues, but the schemas given provide a starting point. The "Mixture" and the "proportion" are suitable reified entities to be described – although one might want to change the labelling of the entities we here call "proportions" to indicate the wider range of information potentially expressed about them. Whether the limitation of OWL and related languages to binary predicates will prove a major barrier to extending the representation of tissues remains to be seen. At this stage it is postulated that they are "good enough" for classification, but that further representation and inference mechanisms are likely to be required for applications such as detailed representation of developmental morphology.

4.1.2 Why do current bio ontologies not make the distinction between granular and determinate parts?

An obvious question is: "If the distinction between determinate and granular parthood is so important, why is it not already standard?" The simplest answer is that few of the large bio-ontologies built to date have been required or used to support inferences that require this distinction.

In the Foundational Model of Anatomy [19, 26], the distinction is prefigured by the notion of "constituent parts". However, the FMA is based exclusively on structure rather than function, so that the issue of emergent effects does not arise. Even when dealing with structure, the FMA does not represent attributes that apply to collectives such as the alignment of cells in the mucosa of the intestine (although the example is due to Cornelius Rosse[14]). Likewise, the FMA does not support detailed cardinality with respect to parts, so the distinction between fixed numbers of parts – *e.g.* fingers – and indeterminate numbers of parts – *e.g.* cells – does not arise. However, these limitations do present difficulties. The issue of the status of tissues and their structure is a significant problem and has, for example, plagued discussions in the SAEL consortium[15] in its efforts to reconcile various anatomic representations in mouse and man. The notions in this paper provide a framework for representing a number of the important notions raised in those discussions and a route towards reconciliation of some of the controversies.

In principle, the *Open*GALEN ontology supports the distinction between collectives (termed "multiples") and determinate parts (termed "components"). However, in practice it has often been elided. The prime use for *Open*GALEN has been for defining surgical procedures and the drug actions and usages. In the first case attention is confined to determinate parts; in the second almost exclusively to granular parts (*e.g.* receptors). In very few cases is their room for confusion; hence the lack of distinction has not proved troublesome. Were the *Open*GALEN model to be extended to include stronger modelling of physiology and function, then it is almost certain that the distinctions presented in this paper would become critical.

In SNOMED-CT, the primary use for anatomy is for the site, or locus, of diseases and the target of surgical and other interventions. Both uses are predominantly on the level of gross anatomy where collective effects are uncommon. Although notions such as "hair loss" must be defined as being literally "loss of at least one hair", in practice no inferences turn on the detailed representation.

Does this neglect of the distinction between determinate and granular parts mean that the distinction is purely "academic"? We believe not. It merely reflects the current state of the art whereby representations are typically restricted to a single level of "collectivity", or if you prefer, "granularity".

As the demand for stronger functional representation across "levels of granularity" grows, including through the interoperation of extant 'single level' ontologies, so too will the need for a precise language to describe individual and collective effects and to distinguish them clearly from effects of physical size.

4.2 Collectives and Normalisation of Ontologies

To support modularisation and maintenance, a major goal of the *Open*GALEN ontologies is to maintain a "normalised" structure in their implementation in which all primitives form disjoint trees and all multiple

---

[13] The label "ingredient" is perhaps not ideal here. No better has yet been suggested, but the authors are open to suggestions.
[14] Private communication, 2004.
[15] http://www.sofg.org/sael/

classification is the result of inference rather than assertion [14]. The schema put forward here all lend themselves to normalisation in this sense. At least in its cognitivist/multiplicative versions, the different aspects of each entity are clearly factored so that they can be described independently.

4.3 Cognitivist vs Realist / Multiplicative vs unitary representation

4.3.1 "Amounts of matter" and "Physical objects": The "constitutes" relationship

The discussion so far has made no link between entities of type Amount_of_matter and entities of type Physical_object. This relation is a matter of controversy between the cognitivist / multiplicative view represented by Guarino and Welty in OntoClean and DOLCE [6, 10, 28] and Smith and his colleagues' in the Basic Formal Ontology (BFO) [22, 24]. Fundamentally, given a statue made of clay, Guarino and Welty's 'Cognitivist/Multiplicative' view is that there are two entities – a "Statue" and an "Amount of clay" and that the " "'Amount of clay' *constitutes* the 'Statue'". Smith's 'Realist/Unitary' view is that there is a single entity and that the "'Amount_of_clay' *is* the 'Statue'", or more precisely that the "'Amount_of_clay' *is* (during some time span) the 'Statue'".

4.3.2 Size of collectives: Empty, small, and determinate collectives.

From a cognitivist, or perhaps better stated "informationalist" viewpoint, there is no problem with empty collectives. There is information to be conveyed about them – that they are empty – therefore it is appropriate to represent them. Likewise, the number of grains in a non-empty collective is irrelevant to whether or not it can be considered a collective. If there is information to be conveyed about the collective properties of some entities, it is irrelevant that in a particular case, the grains of the collective happens to contain to be only a few, or even a single entity.

This view also means that there is no problem with the notion of "determinate collective" "Collectives" as discussed in this paper have so far an indeterminate number of grains. There are, however, collective effects of determinate collections of entities – the collective grip of the fingers, acuity of the eyes, the total capacity of the plates in a dinner service, etc. Note that in each of these cases, the collective effect is not determined by the precise number in the collective even though there is a 'normative' number. For example, a grip has strength whether one or more fingers is missing (or indeed a supernumerary finger were present), a person's visual acuity is typically recorded whether a person has one or two functioning eyes, as being the best visual acuity with all the available eyes.

From the cognitivist or "informationalist" perspective there is no problem – there is distinct information to be conveyed both about the collective and the individual entities that comprise it, hence it is appropriate to represent them separately. However, the realist must find a way of reconciling the collective and deterministic parthood without introducing additional entities. From the point of view of the formal theory, there need be nothing to prevent the same entity being a determinate and granular part of the same whole, indeed to do so would significantly increase the complexity of the axiomitization.

Most other issues discussed in this paper are largely independent of this controversy. For purposes of this paper and presentation in OWL, the factorisation provided by the Cognitivist/Multiplicative view is clearer and briefer, so we shall adopt it here. To do so requires adding the relation constitutes/is_constituted_by to table 3a at the point marked by the elision as one of the additional kinds of "gross parthood". The domain of constitutes is Physical_object, and the range is Amount_of_matter. Since the domain and range are different, and in most formulations disjoint, constitutes/is_constituted_by is non-transitive.

4.4 Other unresolved issues

4.4.1 Operations on Collectives

The most common requirement for operations on collectives is for variants of union and flattening. The collective of members of several collectives – *e.g.* the cells in the skin of the thumb and forefinger – can be easily expressed. Likewise, where collectives are nested, the flattened version can be easily captured – *e.g.* the collective of all cells in the collective of pancreatic eyelets. Although logically possible, the authors have encountered no practical applications requiring such intersections of collectives.

4.4.2 Are collectives of physical entities physical? material?

Whether collectives of physical entities should or should not count as physical has been deliberately left open in this paper. Likewise, whether empty collectives should be material. Because the schema for collectives uses "only" (allValuesFrom) rather than "some" (someValuesFrom) it is perfectly reasonable to assert axioms of the form, for example, that "all collectives of only physical entities are physical", "all non-empty collectives of only physical entities are material." These axioms seem both natural and helpful in biological applications. Similarly, it seems natural to treat non-empty collectives as material, and empty collectives as non-material, analogous to holes.

4.5 Representation in OWL: loss over a full first order theory

The primary goal of this paper is to provide a basis for a representation in description logics and OWL in particular. These languages are deliberately limited with respect to first order logic in order to make them computationally tractable. What is lost in the reduction?

1. The inability to represent irreflexive and antisymmetric properties means that certain incorrect representations cannot be excluded (inferred to be unsatisfiable). If one is willing to accept that no collective can be a grain of another collective without being an ingredient of something else – a desirable restriction in our formulation, then the effect of the irreflexivity of is_grain_of can be made by making its domain NOT Collective and its range Collective. No such solution is possible for antisymmetry, so ontologies represented in GRAIL cannot exclude cycles in the part-whole relationship, although this can be checked for separately by tools.
2. The inability to represent reflexive properties requires making "proper parthood" primitive defining the usual "reflexive parthood" via schemas as described in 2.4.2.
3. The lack of "qualified cardinality constraints" including "EXACTLY-n" means that it is usually most expedient to approximate the relation between ingredients and wholes by simple existential restrictions. In theory this means that the formal model cannot exclude having two identical ingredients.

5. Conclusion: A basis for describing tissues and biological phenomena at multiple "granularities"

The word "granularity" has been used in so many different ways by so many different authors in so many different contexts that to try to enforce a single meaning on the term seems unlikely to succeed. We have therefore used the words "collectivity" and "size range" to distinguish two notions that are often lumped together under the general heading of "granularity". We have labelled the relation between grain and collective is_grain_of rather than the more familiar is_member_of to avoid confusion with mathematical sets defined extensionally. Correspondingly we propose two subrelations:

1. *"Determinate parthood"* – the relation between fingers and hands;
2. *"Granular parthood"* – the relation between cells of the skin of the hand and the hand. For convenience we also define an intermediate;

We argue that the distinction between determinate and granular parthood are useful approaches to two further troublesome problems:

1. When to treat parthood as transitive. .
2. When to treat parthood as persistent.

We argue that determinate parthood can be treated as transitive and persistent, whereas granular parthood cannot, although both imply the parent mereological parthood relation which is, of course, transitive. An implementation using the OWL property hierarchy is presented within a cognitivist framework analogous to DOLCE [7, 28]. The elaboration of the techniques within a realist framework remains to be demonstrated.

In an area where the language is fraught, we invite alternative suggestions for the labelling of any of the notions in this paper. However, whatever the labelling, we suggest that the central notion of collectives and grains is ubiquitous and accounts for important phenomena both in biomedical and broader ontologies and accounts for the criteria set out in the introduction in Section 1.3.

We argue that the two notions of collectivity and size are effectively independent and that boundaries between levels of collectivity occur at all size ranges. In general, notions such as "cellular scale", "atomic scale", "cosmic scale" are nominally focused on size but often conflate the two notions. For example, on the cellular scale one may want to refer to the collectives of organelles such as mitochondria or macromolecules. Furthermore, at least in biomedical applications, it is frequently necessary to refer both to individual grains and to the collectives that they form – *e.g.* both to "the sperm in the seminal fluid" and to "the individual sperm that fertilises the egg".

Our primary motivation has been to provide a basis for representation of the structure of biological materials and substances – *e.g.* the pattern of arrangement of cells in a tissue or the concentration of red cells in blood. To represent information in standard formalisms, there must be entities in the representation to which the information applies. In the representation presented this role is played by the classes Mixture, Proportion and Collective – respectively for the material as a whole, the relation of each ingredient to the mixture, and the ingredients themselves respectively. These notions have been used in limited scale representations. The next stage is to use them to try to provide a comprehensive account of some small set of tissues for a practical application. Likewise, the applicability of these representations to broader areas outside biomedicine remains to be demonstrated.

Acknowledgements

This work supported in part by the Semantic Mining Network of Excellence ((NoE 507505) sponsored by the European Commission, by the CLEF project (G011852) sponsored by the UK Medical Research Council, and by the CO-ODE/HyOntUse (GR/S44686/1) projects sponsored jointly by the UK Joint Infrastructure Services Committee (JISC) and UK Engineering and Physical Sciences Research Council (EPSRC).

References

[1] A. Artale, E. Franconi, and N. Guarino, "Open problems for part-whole relations," *International Workshop on Descripition Logics*, Boston, MA (1996) http://www.dl.kr.org/dl96/

[2] A. Artale, E. Franconi, and L. Pazzi, "Part-whole relations in object-centered systems: An overview," *Data and Knowledge Engineering*, vol. 20 (1996) 347-383.

[3] M. Bada, D. Turi, R. McEntire, and R. Stevens, "Using reasoning to guide annotation with gene ontology terms in GOAT," *ACM SIGMOD Record*, vol. 33 (2004) 27-32.

[4] T. Bittern and B. smith, "A theory of granular partitions," in *Foundations of Geographic Information Science*, M. Duckham, M. Goodchild, and M. Worboys, Eds. (Taylor & Francis Books, London, 2003) 117-151.

[5] R. Casati and A. C. Varzi, *Parts and Places* (Clarendon Press:Oxford,1999).

[6] N. Guarino and C. Welty, "An overview of OntoClean," in *Handbook of Ontologies*, S. Staab and R. Studer, Eds. (Springer Verlag, 2004) 151-159.

[7] N. Guarino and C. Welty, "Towards a methodology for ontology-based model engineering.," *ECOOP-2000 Workshop on Model Engineering*, Cannes, France (2000)

[8] U. Hahn, S. Schulz, and M. Romacker, "Partonomic reasoning as taxonomic reasoning in medicine," *Proc. of the 16th National Conf. on Artificial Intelligence & 11th Innovative Applications of Artificial Intelligence (AAAI-99/IAAI-99)*, Orlando FL (AAAI Press/MIT Press, 1999) 271-276

[9] U. Hahn, S. Schulz, and M. Romacker, "Part-whole reasoning: a case study in medical ontology engineering," *IEEE Intelligent Systems and their Applications*, vol. 14 (1999) 59-67.

[10] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltraman, "WonderWeb Deliveable 18," vol. 2004: WonderWeb consortium, 2003.

[11] J. L. V. Mejino and C. Rosse, "Conceptualization of anatomical spatial entities in the Digital Anatomist Foundation Model," *Journal of the American Medical Informatics Association*(1999) 112-116.

[12] J. J. Odell, "Six different kinds of composition," *Journal of Object Oriented Programming*, vol. 5 (1994) 10-15.

[13] L. Padgham and P. Lambrix, "A Framework for Part-of Hierarchies in Terminological Logics," E. Sandewall and P. Torasso, *KR-94*, 1994) 485-96

[14] A. Rector, "Modularisation of Domain Ontologies Implemented in Description Logics and related formalisms including OWL," J. Genari, *Knowledge Capture 2003*, Sanibel Island, FL (ACM, 2003) 121-128

[15] A. Rector, A. Gangemi, E. Galeazzi, A. Glowinski, and A. Rossi-Mori, "The GALEN CORE Model Schemata for Anatomy: Towards a re-usable application-independent model of medical concepts," P. Barahona, M. Veloso, and J. Bryant, *Twelfth International Congress of the European Federation for Medical Informatics, MIE-94*, Lisbon, Portugal (1994) 229-233

[16] A. L. Rector, P. E. Zanstra, W. D. Solomon, J. E. Rogers, R. Baud, W. Ceusters, W Claassen, J. Kirby, J.-M. Rodrigues, A. R. Mori, E. v. d. Haring, and J. Wagner, "Reconciling Users' Needs and Formal Requirements: Issues in developing a Re-Usable Ontology for Medicine," *IEEE Transactions on Information Technology in BioMedicine*, vol. 2 (1999) 229-242.

[17] J. Rogers and A. Rector, "The GALEN ontology," J. Brender, J. Christensen, J.-R. Scherrer, and P. McNair, *Medical Informatics Europe (MIE 96)*, Copenhagen (IOS Press, 1996) 174-178

[18] J. Rogers and A. Rector, "GALEN's model of parts and wholes: Experience and comparisons," *Journal of the American Medical Informatics Association*(2000) 819-823.

[19] C. Rosse, I. G. Shapiro, and J. F. Brinkley, "The Digital Anatomist foundational model: Principles for defining and structuring its concept domain," *Journal of the American Medical Informatics Association*(1998) 820-824.

[20] S. Schulz and U. Hahn, "Parts, locations, and holes - Formal reasoning about anatomical structures," *Artificial Intelligence in Medicine Europe (AIME-2001)*, Cascais, Portubal (Springer, 2001) 293-303

[21] P. Simons, *Parts. A study in Ontology* (Clarendon:Oxford,1987).

[22] B. Smith, "The basic tools of formal ontology," N. Guarino, *Formal Ontology in Information Systems (FOIS)*, Amsterdam (IOS Press (Frontiers in Artificial Intelligence and Applications), 1998) 19-28

[23] B. Smith, "Granular partitions and vagueness," in *Formal Ontology and Information Systems*, C. Welty and B. Smith, Eds. (Academic Press, New York, 2001) 225-237.

[24] B. Smith, "The logic of biological classification and the foundations of biomedical ontology," D. Westerstahl, *10th International Conference in Logic Methodology and Philosophy of Science*, Oviedo Spain (Elsevier-North-Holland, 2004)

[25] B. Smith, W. Ceusters, B. Klagges, J. Kohler, A. Kumar, J. Lomax, C. Mungall, F. Neuhause, A. Rector, and C. Rosse, "Relations in Biomedical Ontologies," *Genome Biology*(2005) (in press).

[26] B. Smith and C. Rosse, "The role of foundational relations in the alignment of biomedical ontologies," M. Fieschi, E. Coiera, and Y.-C. J. Li, *Medinfo-2004*, San Francisco, CA (IMIA, IOS Press, 2004) 444-448

[27] The Gene Ontology Consortium, "Gene Ontology: tool for the unification of biology," *Nature Genetics*, vol. 25 (2000) 25-29.

[28] C. Welty and N. Guarino, "Supporting ontological analysis of taxonomic relationships," *Data and Knowledge Engineering*, vol. 39 (2001) 51-74.

[29] M. Winston, R. Chaffin, and D. Hermann, "A taxonomy of part-whole relations," *Cognitive Science*, vol. 11 (1987) 417-444.

[30] C. Wroe, R. Stevens, C. A. Goble, and M. Ashburner, "An Evolutionary Methodology To Migrate The Gene Ontology To A Description Logic Environment Using DAML+OIL," *Proceedings of the 8th Pacific Symposium on Biocomputing (PSB)*, Hawaii (2003) 624-635