

Addressing Biological Complexity to Enable Knowledge Sharing

Dean M. Jones, Pepijn R.S. Visser and Ray C. Paton

CORAL - Conceptualisation and Ontology Research at Liverpool

Dept. of Computer Science

University of Liverpool

Liverpool, U.K., L69 7ZF

dean@csc.liv.ac.uk

Abstract

Domain ontologies are now commonly used to enable heterogeneous information resources, such as knowledge-based systems, to communicate with each other. In this article, we present a classification of ontological mismatches, which represent various ways in which knowledge sharing can be impeded by different decisions made by the developers of different resources. We address some of the ways in which the complexity of biological knowledge will inevitably lead to such mismatches arising across different resources. As ontological mismatches can limit the potential for knowledge sharing, we assess the potential for the resolution of these problems.

1. Introduction

It is widely recognised that the principled sharing of knowledge across heterogeneous information systems requires the use of *domain ontologies* as the basis for achieving a common understanding of the domain. An ontology is the specification of a conceptualisation (Gruber, 1993), usually in the form of a logical theory that formally defines the meaning of the terms used in some domain. Where two resources, such as databases or knowledge-based systems, are based on different conceptualisations of the same domain, they are said to be *semantically heterogeneous*. Semantically heterogeneous resources make different ontological assumptions about a domain, i.e. there are *ontological mismatches* between the resources (Visser *et al.*, 1997; 1998). In order that heterogeneous resources can share knowledge, the ontological mismatches between resources need to be resolved. In the KRAFT¹ project (Gray *et*

al., 1997) we overcome ontological mismatches by defining mapping functions between the ontologies of each resource and a *shared ontology*. The mapping functions convert between the terminology of the individual resources and that of the shared ontology. Thus all messages that originate at resources are first converted into this common terminology and then into the terminology of their destination.

To enable the sharing of knowledge between biological resources, it is therefore necessary to define a shared ontology of the relevant biological domain. The ability to agree on a single ontology is notoriously difficult in biological domains. For example, one of the most common means of organising the terms defined by an ontology is a class hierarchy. Hafner and Fridman (1996) suggest that many biological substances cannot easily be represented in a class hierarchy as such substances may be subject to processes which change their identity. In this paper, we present further problem types that may be encountered when attempting to represent hierarchical relationships in biological knowledge. We assess the problems that this can cause for the sharing of biological knowledge between resources and discuss the potential for resolving these problems.

In the next section we describe several types of ontological mismatch. In section 3, we present five types of problem that can be encountered in defining class hierarchies in biological domains which can lead to such mismatches arising. We conclude with a discussion of the implications of this work for the sharing of knowledge between biological information systems.

¹ KRAFT - Knowledge Re-use and Fusion/Transformation
<http://www.csc.liv.ac.uk/~pepijn/kraft.html>

2. Ontology Mismatches

Following the definition of an ontology as an explicit conceptualisation of a domain, we assume the creation of an ontology to involve two sub-processes: the *conceptualisation* of a domain, followed by the *explication* of this conceptualisation. During conceptualisation, decisions are made about the classes, instances, relations, functions and axioms that are distinguished in the domain. It is common during this stage to assign attributes to the classes and to organise them hierarchically. *Conceptualisation mismatches* are mismatches between two (or more) conceptualisations of a domain. The conceptualisations differ in the ontological concepts distinguished or in the way these concepts are related, i.e. a conceptualisation mismatch can be either a *class mismatch* or a *relation mismatch*. First, we distinguish the following two types of class mismatch:

categorisation mismatch: two conceptualisations distinguish the same class but divide this class into different subclasses.

aggregation-level mismatch: two conceptualisations both recognise the existence of a concept but define the corresponding classes at different levels of abstraction.

We have also identified the following three kinds of relation mismatch:

structure mismatch: two conceptualisations distinguish the same set of classes but differ in the way these classes are structured by means of relations.

attribute-assignment mismatch: two conceptualisations differ in the way they assign attributes to a class, e.g. at different hierarchical levels.

attribute-type mismatch: two conceptualisations distinguish the same attribute but differ in the way this can be instantiated.

Explication mismatches are only evident from definitions based on the conceptualisation. We consider definitions to be 3-tuples, i.e. $Def = \langle T, D, C \rangle$ in which T is the definiendum, D is the definiens (to avoid confusion with the definiens we use the letter T - for term - to denote the definiendum), and C is the ontology-concept description to be defined. For practical reasons, we here assume C to be expressed in natural language. T is an atomic formula in a formal language and D is a (compound) formula in a formal language. The three components of a definition allow in principle for eight different relations between two definitions, and thus eight different types of mismatch.

However, if the term, definiens, and concept all differ we assume there to be no mismatch as there is no correspondence. Neither is there a mismatch if terms, definiens, and concepts are all the same as there is a complete match. This leaves the following six types of explication mismatch:

CT mismatch: two ontologies use the same definiens D but differ in both the concept C they define and the term T linked to the definiens.

CD mismatch: the same term is defined in two ontologies with different definiens and is used to denote distinct concepts. Note, that a CD mismatch implies that T is a homonym.

C mismatch: both ontologies have the same term T and definiens D but differ in the concept they define. Note that, like the CD mismatch, a C mismatch implies that T is a homonym.

TD mismatch: two ontologies define the same concept C but differ in the way they define it, both with respect to the term T used and the definiens D . Note that a TD mismatch implies that the two terms are synonyms (it is also possible that the two definiens contain synonyms).

T mismatch: two ontologies define the same concept C using the same definiens D but denote it using different terms.

D mismatch: two ontologies define the same concept C and use the same term T to denote the concept but use a different definiens.

It should be noted that each conceptualisation mismatch is also present in the explication of that conceptualisation. Consider, for instance, the attribute-type mismatch. This conceptualisation mismatch type occurs in the explication as a D or CD mismatch, depending on whether or not the type is specified in the description of the ontological concept C . We adhere to both sets of ontology mismatches as (a) this allows us to tell whether certain types of mismatches arise from conceptualisation or explication (which forms a basis for their resolution, see section 4), and (b) some mismatches are better understood at a conceptual level (i.e. in terms of classes and their hierarchical relations), whereas some mismatches are better understood in terms of ontology components (i.e. terms and definiens). For example, categorisation mismatches are more clearly understood as classes with different subclasses, but less obviously understood when expressed as a CD or D mismatch.

	Procaryotes	Eucaryotes
Organisms	bacteria and cyanobacteria	protists, fungi, plants and animals
Cell size	generally 1 to 10 μm in linear dimension	generally 10 to 100 μm in linear dimension
Metabolism	anaerobic or aerobic	aerobic
Organelles	few or none	nucleus, mitochondria, chloroplasts, endoplasmic reticulum, etc.
DNA	circular DNA in cytoplasm	very long linear DNA molecules containing many noncoding regions; bounded by nuclear envelope
RNA and protein	RNA and protein synthesized in same compartment	RNA synthesized and processed in nucleus; proteins synthesized in cytoplasm
Cytoplasm	no cytoskeleton: cytoplasmic streaming, endocytosis, and exocytosis all absent	cytoskeleton composed of protein filaments; cytoplasmic streaming; endocytosis and exocytosis
Cell division	chromosomes pulled apart by attachments to plasma membrane	chromosomes pulled apart by cytoskeleton spindle apparatus
Cellular organization	mainly unicellular	mainly multicellular, with differentiation of many cell types

Table 1: Comparison of Procaryotic and Eucaryotic Cells (from Alberts *et al.*, 1989, Table 1-1; p.19)

type. Indeed, we have identified exceptions for most of the defining characteristics listed in Table 1.

Concepts such as procaryotic and eucaryotic cells will commonly be formally defined in terms of a few necessary conditions and several sets of sufficient conditions. Such definitions will often cover only the most common instances of the class. Atypical instances may not therefore be covered by any of the definitions of in terms of sufficient conditions. Categorisation mismatches arise when different decisions are made about the form and number of definitions, giving rise to (at least) D mismatches between different ontologies. When defining a shared ontology, it is sensible to maximise the potential for sharing knowledge between heterogeneous resources. This will be facilitated if the number of sufficient definitions is maximised and the number of necessary conditions is minimised, as this allows communication about the greatest number of instances. However, care must be taken to ensure that the ontology defines only valid models of the domain. Definitions of concepts should not be relaxed to the extent that erroneous examples can be included as instances of a class.

3.2 Multiple Sibling Instantiation

The type of classification mismatch arises when an instance of a class is also an instance of two (or more) of the immediate children of that class. In the classical view of

category membership, this situation should not be possible as sibling classes are distinguished from their immediate parent and each other by some *differentia* and therefore will not overlap in this way. This can perhaps be explained more clearly using an example. Alberts *et al.* (1989) describes the subclasses of the class 'remote signalling cell' as follows:

- (1) In **endocrine signaling**, specialized endocrine cells secrete **hormones**, which travel through the bloodstream to influence target cells that are distributed widely throughout the body.
- (2) In **paracrine signaling**, cells secrete **local chemical mediators**, which are so rapidly taken up, destroyed, or immobilized that the mediators act only on cells in the immediate environment, perhaps within a millimeter or so.
- (3) In **synaptic signaling**, which is confined to the nervous system, cells secrete **neurotransmitters** at specialized junctions called *chemical synapses*; the neurotransmitter diffuses across the synaptic cleft, typically a distance of about 50 nm, and acts only on the adjacent postsynaptic target cell. (Alberts *et al.*, 1989; p.682)

Remote signalling cells can be classified according to the mode of signalling in which they are involved, giving the three subclasses 'endocrine cell', 'paracrine cell' and 'nerve cell'. Subsequently, however, another type of remote signalling cell is identified:

3. Classification Mismatches

Although the representation of hierarchical knowledge is widely seen as an important aspect in the design of a formal ontology, very little advice is available on the problems that may be encountered during this process and how such problems can be addressed. Furthermore, the advice that is available is often based on an unrealistic view of the structure of categories. For example, some recommendations that arose from the development of the MENELAS ontology are given in Bouaud *et al.* (1995). These principles are largely based on the classical model of category membership, where some entity is deemed to be a member of a category if and only if it exhibits the relevant necessary and sufficient features. It has long been recognised that this is an inadequate model of the way in which category membership is decided. There has, however, been little work on how the complexity of category structure can be addressed in the development of shared ontologies. Here we describe some common problems that are encountered in the formal representation of hierarchical knowledge. Firstly, as there is often confusion regarding the terms that are used to discuss hierarchical representations of knowledge, definitions of some terms will be useful. Note that these definitions are based on those used in the ONTOLINGUA representation language (Gruber, 1993):

class: a partitioning to which some assertion or set of assertions might apply.

instance: a member of a class which is either a subclass or an individual.

individual: a member of a class that is not itself a class.

We now describe five types of *classification problem*. A classification problem can be informally defined as a situation in which there is confusion about the hierarchical relationships that should be included in an ontology (for a fuller discussion, see Jones, 1997; Jones and Paton, 1998). The resolution of these problems will be based on decisions made by the ontology designer. Different designers will make different decisions, leading to classification mismatches (and therefore also explication mismatches) between the alternative ontologies. In the following sections, we give examples of classification problems from biological domains and discuss the options available to the designer of a shared ontology in resolving these mismatches. Where it is feasible to do so, we describe the representation in a shared ontology

that will maximise the potential for knowledge sharing across heterogeneous resources.

3.1 Atypical Instances

The existence of atypical instances has long been recognised (e.g. Rosch, 1975). Typicality ratings result from differing judgements of the degree to which an instant is a member of a class. As an example, consider one of the most basic distinctions made in cell biology, the classification of cells as either procaryotes or eucaryotes. The distinguishing features of these cells are given in Table 1 (reproduced from Alberts *et al.*, 1989). This table is introduced with the sentence: "the major existing eucaryotes have in common both mitochondria and a whole constellation of other features that distinguish them from procaryotes (Table 1-1)" (Alberts *et al.*, 1989; p.19). Table 1 is therefore taken to describe the distinguishing features of the two child classes of the class 'cell'. According to the classical Aristotelean theory of concepts, the properties described here should be interpreted as the defining features of two classes being described. However, many of the listed properties are generalisations which will not apply to every instance of each class. For example, it is clear that the property 'contains a nucleus' is one of the defining characteristics of eucaryotic cells: "Eucaryotic cells, by definition and in contrast to procaryotic cells, have a *nucleus*" (Alberts *et al.*, 1989; p.15). Although this is presented as a defining feature of eucaryotic cells, it is not always the case, as shown by the eucaryotic red blood cells: "erythrocytes (or red blood cells) are very small cells, usually with no nucleus" (Alberts *et al.*, 1989; p.25). The location of genetic material and associated processes - such as translation and transcription - is also given as a distinguishing feature of the two subclasses of cell, as the following description shows:

in procaryotic cells, there is no compartmentalization - the translation of RNA sequences into protein begins as soon as they are transcribed ... in eucaryotes, however (except in mitochondria and chloroplasts, which in this respect as in others are closer to bacteria), the two steps in the path from gene to protein are kept strictly separate (Alberts *et al.*, 1989; p.26).

It is clear that although the location of genetic processes is presented as a distinguishing feature of cell type, there are atypical instances - mitochondria and chloroplasts - which do not exhibit the features associated with their corresponding

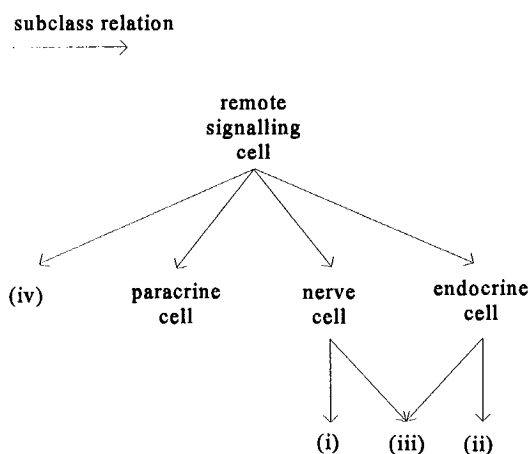


Figure 1: Possible classifications of ‘neuroendocrine cell’

The linking function of the hypothalamus is mediated by cells that have properties of both nerve cells and endocrine cells; for this reason they are called *neuroendocrine cells*. (Alberts *et al.*, 1989; p.683)

Neuroendocrine cells secrete hormones into the bloodstream like endocrine cells and respond to synaptic stimulation like nerve cells. Four possible representations of the class ‘neuroendocrine cell’ are identified in Fig. 1. The class ‘neuroendocrine cell’ could be categorised as either an instance of ‘nerve cell’ (i), as an instance of ‘endocrine cell’ (ii) or as an instance of both of these (iii). However, as neuroendocrine cells do not satisfy all of the properties of either of these classes, they are typical of neither and could be defined as a new subclass of ‘remote signalling cell’ (iv).

This type of classification problem results from the difficulty of determining those attributes of individuals that contribute to their identity as a kind. As discussed in Guarino *et al.* (1994), Strawson (1959) defines sortal predicates, i.e. those that allow us to identify a thing as a particular kind, as providing a principle for distinguishing and counting individuals, e.g. ‘apple(X)’. This contrasts with non-sortal predicates, which supply such a principle only for individuals already distinguished, e.g. ‘red(X)’. According to Guarino *et al.* (1994) a predicate is sortal if it is (a) countable, i.e. the predicate allows a given object to be identified amongst other kinds of objects, and (b) temporally stable, i.e. if the predicate holds for an object at a given time, it also holds for the same object at another time. Further, a substantial sortal is one that is ontologically rigid i.e. it cannot lose the

property without losing its identity. Substantial sortals are predicates that identify the type of an entity and should therefore be defined as classes. Non-substantial sortals should be defined as roles on those classes. In our example, the way in which a cell is behaving is a non-substantial sortal and therefore does not determine its identity. The location of a cell is the property that determines its type and therefore, as neuroendocrine cells are both part of the nervous system and of the endocrine system, they are both a kind of nerve cell and a kind of endocrine cell. The correct representation for this class would be (iii) in Fig. 1.

When addressing domains such as biology that include highly complex entities, it is often difficult to identify those predicates which are sortal. Conflicting decisions made by the developers of different resources will result in mismatches between ontologies. Of course, the difficulty of resolving these mismatches will depend on the degree to which the representations conflict.

3.3 Context-sensitive Membership

The validity of many classifications is dependent upon the context in which they are used. As an example of this, consider the kinds of chemical bond that can exist between biological molecules. There are two classes of chemical bond, covalent and non-covalent. The subclasses of the non-covalent bond class are described by: “non-covalent bonds encountered in biological molecules are usually classified into three types: **ionic bonds**, **hydrogen bonds**, and **van der Waals attractions**” (Alberts *et al.*, 1989; p.88). Sometimes, however, hydrophobic forces are included as another kind of non-covalent bond: “Another important weak force is created by the three-dimensional structure of water, which tends to force hydrophobic groups together” (Alberts *et al.*, 1989; p.88). It appears to be difficult to say with any certainty whether or not ‘hydrophobic force’ is an instance of ‘non-covalent bond’: “This expulsion from the aqueous solution generates what is sometimes thought of as a fourth kind of weak non-covalent bond.” (Alberts *et al.*, 1989; p.88). This is a weaker assertion than stating that hydrophobic forces are atypical instances of non-covalent bonds. Only under certain conditions are hydrophobic forces considered to be non-covalent bonds.

Some definitions of a class may include context-sensitive instances as members while others may not. This may lead to (at least) D mismatches between different

ontologies. In order to maximise the potential for knowledge sharing, context-sensitive members are generally included as instances of a class in a shared ontology, as far as this is consistent with the tasks to be supported by the ontology.

3.4 Excluded Instances

As an example of an excluded instance, consider the class of small organic molecules, the subdivision of which is described in Alberts *et al.*, (1989) as follows: “cells contain just four major families of small organic molecules: the simple **sugars**, the **fatty acids**, the **amino acids**, and the **nucleotides**” (p.43). It appears from this that the class ‘small organic molecule’ has four child subclasses. However, this is not strictly the case as “some cellular compounds do not fit into these categories” (Alberts *et al.*, 1989; p.43). In an ontology that includes only the above four subclasses of ‘small organic molecule’, any of these cellular compounds will be an excluded instance.

Scientific theories are rarely concerned with individuals but largely refer to abstract classes. The name of a molecule refers to a class rather than some individual. Defining a type of molecule as an individual should be avoided as many of the properties associated with being an individual do not apply to classes, and vice versa. We cannot include the names of molecules as individuals that are instances of a class such as ‘small organic molecule’. This correlates with the suggestion in section 3.1 that having decided to include a class in an ontology, it should be defined as completely as possible. This entails that all known subclasses of ‘small organic molecule’ should be included. It should be noted however that the solution being proposed here is not that the subclasses of all classes should be exhaustively specified, since this will often be impractical, e.g. consider exhaustively specifying the subclasses of the class ‘thing’. The scope of an ontology is defined by the task for which the ontology is used. If a task demands that the class ‘small organic molecule’ is defined in an ontology, the task will probably also demand that this is specified as fully as possible in order that non-intended models are excluded. Few tasks will demand that the class ‘thing’ is exhaustively specified.

3.5 Non-instance Similarity

As suggested in section 2.2, it is often difficult to decide which predicates contribute to the identity of the individuals in a domain. Even when those predicates that are substantial

sortals have been identified, entities can be misidentified as belonging to the particular class. Two individuals may be judged to be more or less similar to each other, depending on which properties are taken to be relevant to the comparison, the relative weighting given to the selected properties, and so forth.

As an example, consider the information given in section 2.1 regarding the way in which mitochondria and chloroplasts are similar to bacteria. It is clear that the former pair exhibit many of the distinguishing features of procaryotic cells. Table 2 details the similarities between procaryotic cells and the organelles mitochondria and chloroplasts, based on some of the distinguishing features of procaryotic cells outlined in Table 1. According to the definitions given in Table 1, these are the features that would be the most important in deciding on membership of the category ‘procaryotic cells’. Although “mitochondria and chloroplasts show important differences from ... present day aerobic bacteria and cyanobacteria” (Alberts, *et al.*, 1989; p.18) and would probably not be judged to be typical instances of the class ‘procaryotic cell’ it is very possible they would be included as atypical instances, based for example on the similarities described in Table 2. However, mitochondria and chloroplasts are organelles and as such do not belong to the class ‘cell’ and should not be considered even as atypical members of the class of procaryotic cells.

The similarities outlined in Table 2 can be explained in terms of biological theory, as “chloroplasts share a common ancestry with cyanobacteria and evolved from procaryotes that made their home inside eucaryotic cells” (Alberts, *et al.*, 1989; p.18). Mitochondria and chloroplasts are descendents of procaryotes but are not classified as such: “Although they seem to have originated as symbiotic bacteria, they have undergone large evolutionary changes” (Alberts, *et al.*, 1989; p.18-19). This explanation of the classification of chloroplasts and mitochondria as organelles also explains the degree of similarity between these organelles and procaryotic cells. The identification of such cases will help prevent the inclusion of erroneous instances of classes in ontological definitions.

4. Conclusions

This paper has shown that the formal representation of complex biological concepts is far from straightforward, mainly because it is often difficult or impossible to identify

	Procaryotic Cells	Mitochondria and Chloroplasts
Size	generally 1 to 10 μm in linear dimension	"Mitochondria ... often resemble bacteria in size and shape"
Metabolism	anaerobic or aerobic	"Many present-day bacteria respire like mitochondria" "Chloroplasts carry out photosynthesis in much the same way as procaryotic cyanobacteria"
DNA	circular DNA in cytoplasm	"[chloroplasts] contain DNA that is nearly indistinguishable in nucleotide sequence from portions of a bacterial chromosome"
RNA and protein	RNA and protein synthesized in same compartment	"in eucaryotes, however (except in mitochondria and chloroplasts, which in this respect as in others are closer to bacteria), the two steps in the path from gene to protein are kept strictly separate"

Table 2: Comparison of Procaryotic Cells with Mitochondria and Chloroplasts
(after Alberts *et al.*, 1989, Table 1-1; p.19)

a set (or sets) of defining features for a concept. We have identified several types of problem that can arise in the formal representation of hierarchical representations and have analysed some of the issues involved in selecting one representation above another. As such choices exist in the representation of biological knowledge, there will inevitably be ontological mismatches between different repositories of biological knowledge. Such mismatches can often be difficult to resolve.

We have suggested that some of the problems involved in the representation of hierarchical knowledge in a biological ontology will need to be resolved on the basis of the tasks that are required to be supported by the ontology. This is in accordance with the latest work on ontological engineering (e.g. Grüninger and Fox 1995; Uschold 1996; Fernández, Gómez-Pérez and Juristo 1997) which suggests that ontologies should be developed and assessed in relation to their intended purpose. In domains such as molecular biology, formal definitions of terms are not common and therefore are not available to the ontological engineer. In such situations, we cannot expect to be able to objectively develop ontologies that can be used for all knowledge sharing tasks. Instead, ontologies should be defined for use on a task-specific basis, e.g. for the exchange of certain kinds of knowledge and data.

Biological knowledge is continually evolving, seemingly at an increasing pace. It is unlikely that it will be

possible to agree on standard descriptions of biological domains. O'Leary (1997) argues that the lack of ontology stationarity is an impediment in the development of standard ontologies for knowledge sharing. This, along with the highly subjective nature of ontology development, means that different representations of the same domain will be available at any one time. For the use of ontologies to become widespread practice, it will be necessary to develop methods which with different ontologies can be evaluated (for examples of such assessments, see Fridman Noy and Hafner, 1997; Visser and Bench-Capon, 1998) and to agree on ways in which ontologies should be described. This will allow developers of knowledge sharing systems to select the most suitable ontologies for their particular needs from those that are available.

5. Acknowledgements

The authors would like to thank members of the KRAFT project, especially those at Liverpool, who have contributed to the work on mismatches.

6. References

- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. (1989) *Molecular Biology of the Cell*, Garland, New York.
- Bouaud, J., Bachimont, B., Charlet, J. and Zweigenbaum, P. (1995) "Methodological Principles for Structuring an

- 'Ontology'", *IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, August 19-20th.
- Fernández, M., Gómez-Pérez, A. and Juristo, N. (1997) "METHONTOLOGY: From Ontological Art Towards Ontological Engineering", presented at *AAAI-97 Spring Symposium on Ontological Engineering*, Stanford University, March 24-26th.
- Fridman Noy, N. and Hafner, C.D. (1997) "The State of the Art in Ontology Design: A Survey and Comparative Review", *AI Magazine*, **18**(3), 53-74.
- Gruber, T.R. (1993) "A Translation Approach to Portable Ontologies", *Knowledge Acquisition*, **5**(2), 199-220.
- Grüninger, M. and Fox, M.S. (1995) "Methodology for the Design and Evaluation of Ontologies", *IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing*, August 19-20th, Montreal.
- Gray, P.M.D., Preece, A.D., Fiddian, N.J., Gray, W.A., Bench-Capon, T.J.M., Shave, M.J.R., Azarmi, N., Wiegand, M.E., Ashwell, M., Beer, M.D, Cui, Z., Diaz, B.M., Embury, S.M., Hui, K.Y., Jones, A.C., Jones, D.M., Kemp, G.J.L., Lawson, E.W., Lunn, K., Marti, P., Shao, J. and Visser, P.R.S. (1997) "KRAFT: Knowledge Fusion from Distributed Databases and Knowledge Bases", in *Proc. Database and Expert System Applications (DEXA' 97)*, Toulouse, France.
- Guarino, N., Carrara, M. and Giaretta, P. (1994) "An Ontology of Meta-Level Categories", in J. Doyle, E. Sandewall and P. Torasso (eds.) *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference*, Morgan Kaufmann, San Mateo, CA.
- Hafner, C.D. and Fridman, N. (1996) "Ontological Foundations for Biology Knowledge Models", *Proc. 4th International Conf. on Intelligent Systems for Molecular Biology*, St Louis, AAAI Press, 78-87.
- Jones, D.M. (1997) "Representation and Visualisation of Scientific Theories for Knowledge-based Systems", PhD Thesis, University of Liverpool.
- Jones, D.M. and Paton, R.C. (1998) "Some Problems in the Formal Representation of Hierarchical Relationships", in *Proc. Conference on Formal Ontology in Information Systems - FOIS'98*, June 6-8th, Trento, Italy, IOS Press.
- O'Leary, D.E. (1997) "Impediments in the use of Explicit Ontologies for KBS Development", *International Journal of Human-Computer Studies*, **46**, 327-337.
- Rosch, E. (1975) "Cognitive Representation of Semantic Categories", *Journal of Experimental Psychology*, **104**, 192-233.
- Strawson, P.F. (1959) *Individuals: An Essay in Descriptive Metaphysics*, Routledge, London.
- Uschold, M. (1996) "Building Ontologies: Towards A Unified Methodology", in *Proc. Expert Systems 96*, Cambridge, December 16-18th.
- Visser, P.R.S., Jones, D.M., Bench-Capon, T.J.M., and Shave, M.J.R. (1997) "An Analysis of Ontology Mismatches: Heterogeneity versus Interoperability", presented at *AAAI 1997 Spring Symposium on Ontological Engineering*, Stanford University, CA.
- Visser, P.R.S and Bench-Capon, T.J.M. (1998) "A Comparison of Four Ontologies for the Design of Legal Knowledge Systems", *Artificial Intelligence and Law*, accepted for publication.
- Visser, P.R.S., Jones, D.M., Bench-Capon, T.J.M. and Shave, M.J.R. (1998) "Assessing Heterogeneity by Classifying Ontology Mismatches", in *Proc. Conference on Formal Ontology in Information Systems - FOIS'98*, June 6-8th, Trento, Italy, IOS Press.