# Continua in Biological Systems

**Stefan Schulz[1], Ingvar Johansson[2]**

[1] Department of Medical Informatics, Freiburg University Hospital, Germany

[2] Institute for Formal Ontology and Medical Information Science
(IFOMIS), Saarland University, Germany

We defend the fundamental ontological-pragmatic principle that often where there are continua in reality, science is forced to make partly fiat terminological delimitations. In particular, this principle applies when it comes to describing biological organisms, their parts, properties, and relations. Human-made fiat delimitations are indispensable at the level of both individuals and the natural kinds which they instantiate. The kinds of pragmatically based 'fiatness' that we describe can create incompatibilities and lack of interoperability even between properly designed ontologies, if not appropriately taken care of.

## 1. Introduction

In the life science-based informatics disciplines the need for terminological support in the treatment of data is widely acknowledged, and controlled vocabularies, terminological coding systems and thesauri have a long tradition in biology and medicine. These consist of standardized terms and relations, interpreted variously as referring to word meanings, concepts, universals, or to classes of entities in the world, which are used in the coding and annotation of data for many different kinds of purposes. Among the life sciences there is a broad consensus about the existence, nature, and functioning of basic entities such as organs, tissues, cells, molecules, and

genes. This consensus embraces the existence of both macroscopic and microscopic entities as well as many facts about their properties and relations. That proteins consist of amino acids, that cells have membranes, or that green plants perform photosynthesis, are today, just common presuppositions of many empirical scientific inquiries.

In addition to the controlled vocabularies and clinical terminologies, in the recent years there have been devised an increasing number of biomedical *ontologies*. What we say in the sequel will affect how the term '*ontology*' ought to be defined, but in the meanwhile let us say that an ontology is something like: *a representational artifact whose representational units are intended to designate classes or universals in reality and their interrelations.*[1,2] We will avoid another sense that we frequently find in informatics contexts, namely that an "*ontology is a system that describes concepts and the relationships between them*".[3] The main reason behind our choice is that in our opinion biomedical ontologies can be most useful when they attempt to give a taxonomic account of the biomedical reality that exists independently of the ontology itself and that is described for example in textbooks accepted by the biomedical research community.

Biomedical ontologies strive for a representation of biomedical reality that is not subject to the arbitrariness of individual perceptions. The biomedical ontologies currently in use, however, serve specific application requirements – for example in the context of health information systems – and this means that they often need to represent putative and heavily context-dependent facts. In several ontologies under development we do register an increasing approximation to the end of non-arbitrary representation, of which the Foundational Model of Anatomy[4,5] is the most important; and some other ontologies of the Open Biomedical Ontologies (OBO) Foundry platform might be mentioned in this connection also.[6]

A sign of the growing move toward standardization of ontologies is the increasing use of the Web Ontology Language OWL[7] as language of formalization, as well as proposals that recommend a parsimonious set of semantically clear-cut relations.[1] So far, however, the degree of convergence and interoperability of ontologies are still quite disappointing, and the terms standardly used to describe biological facts still do not work together sufficiently well. One major reason for this is, we will argue, the very nature of the entities in the biological domain. In particular, it has to do with the phenomenon of *continuity*. Any enterprise that aims at creating an ontology of biological entities will have to face the challenge of drawing boundaries in and across many of the continua that biological reality contains.

Let us illustrate what we mean by an example that brings in both physics and everyday life. In physics we find continuous scales for qualities such as length and mass (weight), but in everyday communication such qualities are treated as if there is a least discrete unit which applies to each measurement of length. When weighting items on a scale, for instance, we seldom go below the discrete units of '1 mg', '2 mg', etc., so that what is continuous is treated as if it were discrete.

We observe an analogous phenomenon in clinical medicine in relation to quantities such as body temperature and blood pressure and to many other quantities measured in clinical laboratories. Often, where continua exist in both theory and reality, practical medical considerations suggest the transformation of these continua into discontinuities. One reason for this derives from the need to make clear-cut decisions as to whether to treat or not to treat a given patient. Between any two points on a continuous scale there is always a third point, but between treatment and non-treatment there is nothing. This is the general background for the so-called 'cut-off values' that are used when continuous variables are mapped on to discrete ones, whereby both the range and the number of divisions such as *'hypotensive'*, *'normotensive'*, and *'hypertensive'* become a matter of practical convention.

Moreover, due to the stochastic nature of the outcome of many empirical investigations, misassignments of individual medical cases are unavoidable; this is especially true of cases where values are close to a cut-off point. Let us use the *concentration of glucose in blood* as an illustration. This concentration is defined as the ratio between the mass or molecular weight of glucose in a portion of blood and the volume of this portion at a given temperature. It varies across time and between blood portions, but both in theory and in reality there is at every instant for every portion of blood exactly one true value that represents the given concentration. In principle, there is a continuum of such values. Nonetheless, physicians work with a bipartition between 'normal' and 'abnormal' concentrations. This division is based on statistical knowledge about the association between 'glucose concentration' and what is regarded – on the basis of consequences for the well-being of the subject in question – as the normal or abnormal functional status of the corresponding organism.

Figure 1 depicts results of empirical investigations involving measures of glucose concentration. On the y-axis we find the number of people observed, and on the x-axis values for measured glucose concentrations. On the continuous scale for glucose concentration a cut-off

point that distinguishes between the discrete binary values 'normal' and 'abnormal concentrations' is inserted, in part on the basis of fiat.[8] The medical ideal is a correlation which is such that all normal conditions result in normal test values, and vice versa; and that all abnormal conditions result in abnormal test values, and vice versa. This ideal is hardly ever met, and it is definitely not met in the example depicted in Figure 1. This means that some persons with normal test values for glucose concentration are nonetheless in an abnormal condition; such results are called 'false negatives' (FN in the figure). It also means that some persons with abnormal test values for glucose concentration are nonetheless in a normal condition; such results are called 'false positives' (FP).

The position of the cut-off value in the figure is by no means wholly determined by biological reality itself; nor, by the way, is it a reflection of purely mathematical considerations of symmetry. It is partially chosen by fiat, and reflects a *chosen* trade-off between the specificity (percentage of abnormal conditions that come out as abnormal) and sensitivity (percentage of normal conditions that come out as normal) of the test procedure in question. With the chosen cut-off value, the 'specificity' and the 'sensitivity' of the test become equal. If the cut-off value is placed where the Bell Curve for abnormal conditions ends on the x-axis, then the specificity
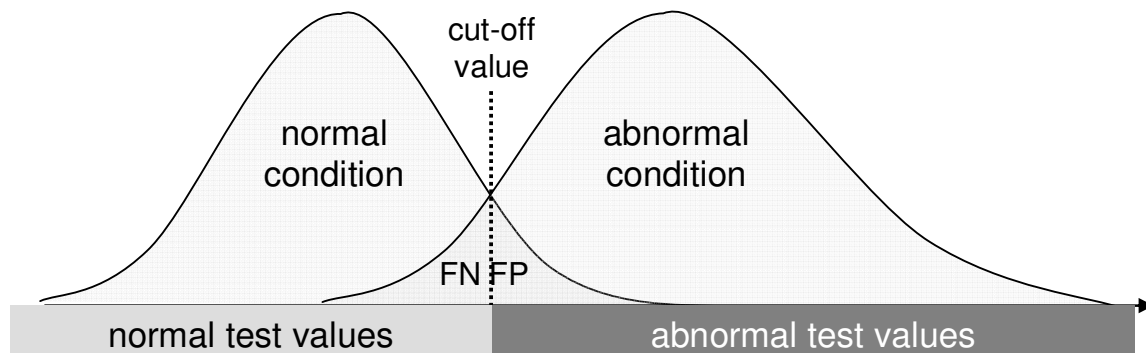


Figure 1. Continuum of test values mutated into a discrete bipartition between what is normal and abnormal. Errors of measurement are neglected in this example

would have been perfect (= 1), but the sensitivity would be rather low.

At present, the conventional cut-off value that distinguishes normal (*euglycemia*) and abnormal concentration (*hyperglycemia*) is set to 110 mg-glucose/dl-blood after fasting. This choice partly depends on observations, partly on clinical pragmatics. From a pragmatic point of

4

view, one has to answer this question: is it more acceptable to treat healthy individuals that need no treatment than not to treat diseased individuals that do need treatment? The answer reflects no real discontinuity in reality but only a practical one. If ontologies are to be useful, however, they ought to take account of such practical discontinuity.

The example given is merely one of a large number of cases in medicine where theoretically continuous quantitative values are transformed into discreta. In fact, this phenomenon extends to nearly every facet of biological reality. We will examine some further cases below, seeking to demonstrate that, *because of the existence of continua, every useful ontology for living organisms has to contain divisions that rest on pragmatic considerations and that are thus partially fiat in nature*.

## 2. Continua and Classification of Natural Kinds and Parts

We first need to make a clear distinction between entities that can exist in many places simultaneously and entities that can only be at one place at a time. This distinction is implicitly used in most descriptions of reality. If a person is said to have a glucose concentration of 114 mg/dl, it is taken for granted that there can be other persons who have exactly the same concentration, but that the person spoken of and his concentration cannot at that moment be also at some other place in addition to where they are. Similarly, whenever a nurse reports that 'patient in bed 3 has a fever of $38.9^{o}C$', it is taken for granted that in the world there can be other patients, other beds, other instances of fever, and other instances of the temperature $38.9^{o}C$, but that the patient spoken of and his fever cannot at that moment be at any other place. We will call the former entities '*repeatables*' (others call them 'universals' or 'types') and the latter '*particulars*' ('instances', 'tokens', 'individuals', or 'spatiotemporally located entities'). All particulars that instantiate the same repeatable will be said to belong to the same *class* (sometimes called the 'extension' of the repeatable in question). A class is thus a collection of particulars of one or several repeatables.

We have deliberately chosen the term 'repeatable' instead of the more traditional 'universal', because the latter still bears associations deriving from the non-evolutionary view of nature still prevalent at the time it was coined. Aristotle held that all natural kinds of living entities were instantiated at every point of time. Natural kinds, in this sense, were seen not only as repeatables, but also as existing 'universally'. To be a dinosaur is to instantiate a *repeatable*,

5

but it is *not* to instantiate something that has a *universal existence* in this sense. We have avoided the terminology of types and tokens since we wish to avoid connotations deriving from the fact that this terminology was coined to make the repeatable–non-repeatable distinction in relation to linguistic entities such as words and sentences.

According to *realism* with respect to repeatables, there are repeatables that exist both independently of and in dependence on human minds, languages, and acts of cognition. According to *conceptualism*, repeatables can exist *only* in minds, uses of languages or acts of cognition, i.e., all repeatables are 'concepts', 'words', or 'terms'. (For simplicity's sake, we leave conceptualism's more extreme cousin *nominalism* out of account.) Most realist philosophers have rested content with arguing that their general position must be true, i.e., that on pain of absurdity one has to believe that there are at least *some* mind-independent repeatables (otherwise for example the very activity of scientific investigation would be rendered incoherent). Conceptualists surely must go wrong somewhere, they argue, since there were surely fevers and determinate temperatures in the world before the corresponding concepts were invented. In this we agree, but we see a need to take realism further if it is to be relevant to the needs of contemporary biomedical informatics and related disciplines. To this end we need to lay stress on the fact that realism posits not only repeatables that are independent of language and cognition, but also repeatables that are dependent on these factors, and we need to investigate the relations that may exist between the former and the latter.

Each language-independent repeatable, and many language-dependent repeatables, delineate a corresponding class of particulars. But such delineations can be of different sorts. Consider atoms and molecules. Excluding particles under extreme pressures and temperatures, there is no doubt that the class of atoms is completely distinct from the class of molecules. There are no intermediary entities that are *x%* atoms and *(100 – x)%* molecules. Even more, to the best of our knowledge, all atoms of a given isotope (defined by atomic mass) are strictly qualitatively identical; they do not exhibit any individual distinguishing features. The members of such a class could thus not be grouped together in various different ways depending on what features one chooses to focus on.

What we have said about isotopes applies to molecules, too, and even to large bio-molecules. What kind of protein a molecule is, is unambiguously determined by its amino acid sequence and its spatial configuration. However, when we come to larger associations of

6

biomedical matter, things are less neat. Let us consider two gene sequences that differ in exactly one nucleotide, so-called SNPs (single nucleotide polymorphisms): Should we regard them as instances of the same sequence (repeatable) or not? Here, some element of human delimitation must come into play. Indeed, we must decide how many mutations are needed before a gene of one sort becomes a gene of another sort. This number may be high if the functional impact of the mutations is small, or low if even a single variation already inflicts a significant change of biological properties. However, it is nonetheless true that all of these 'urelements' of life have a definite non-stochastic composition with a limited and discrete number of varieties.

## 2.1. *Classification of Continuous Repeatables*

### *Example: Cell Types*

The picture changes radically when we reach the level of highly complex unities such as cells and organelles. While the comparison of particular molecules of the same type will reveal no structurally distinctive features, it is very improbable (even though logically possible) that we will ever observe individual cells which are qualitatively identical and have exactly the same number and arrangement of parts. Every particular cell exhibits innumerable individual traits. Cells are highly adaptive and may gradually change the relative expressions of their numerous features over time; during the formation of an embryo such changes take place as reactions to both external stimuli and in the course of internally generated abnormal growths. Particular cells may accordingly be ordered on a variety of continuous scales which resembles the situation depicted in Figure 1 with respect to qualities.
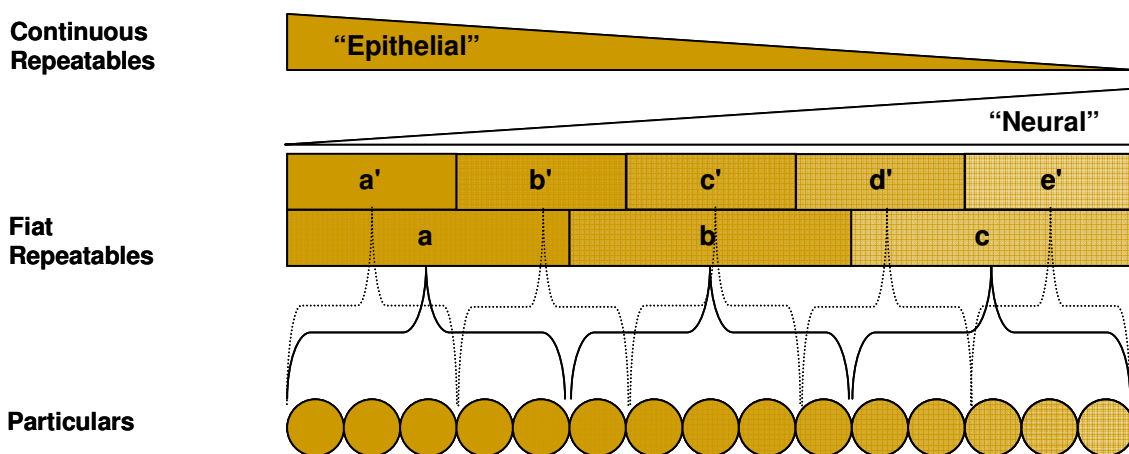


Figure 2. Continuum between two types of cells and assignment of particular cells

to fiat repeatables (*a, b, c* vs. *a´, b´, c´, d´, e´*), which partition this continuum

Classifying a particular as being an instance of a certain repeatable is here not a matter of judging some single parameter; gradients of several parameters have to be taken into account simultaneously, and new techniques have evolved that bring in a broad range of morphological, chemical, and genetic properties. Some of these properties are discrete (e.g., number of nucleoli), some are continuous (e.g., enzymatic activity). Consequently, the repeatables advanced for consideration in such a domain always depend on a consensual decision on the part of the relevant research community: which factors should be taken into account, and how should the cells be partitioned? For the sake of diagnostic clarity it may be adequate to partition the developmental or transitional continuum of cells into discrete types as in Figure 2. Two possible ways of inserting fiat boundaries have been marked (which may characterize diverging viewpoints within a research community). The upper level gives us three classes, *a*, *b*, and *c* (e.g., epithelial, neuroepithelial, and neural cells); the lower level gives us five classes, *a′* to *e′*. The terms 'epithelial cell', 'neuroepithelial cell', and 'neural cell' reflect repeatables which are in part a matter of fiat; if these fiat delimitations are not made the terms become vague. Their application presupposes both the existence of the language-independent repeatable *cell* and the continuum depicted in the Figure. Clearly, different partitions of a continuum cannot in general be associated with each other in any one-to-one mapping. For example, a cell classified as a *b*-cell (in Figure 2) is either a *b′*-, *c′*-, or *d′*-cell on the lower classification scheme, and a cell classified as a *d′*-cell is either a *b*-cell or a *c*-cell on the upper classification scheme.

## 2.2. *Classification of Analogous Biological Structures*

*Example: "Stomach" across Species*

Not all biological natural kinds or objects face the sort of classification problems we have just exposed in relation to cells. The classical anatomical organs make up one such group. There are no intermediary organ types between human livers, stomachs, and hearts. However, the picture changes when we approach both the *ontogenesis* of organs and the question of what organs are *analogous* from the point of view of inter-species variability. What is considered to be

8

one specific kind of organ as defined in relation to a single species may become quite problematic when other species are considered. The stomach affords us with a good example of this problem (see Figure 3).
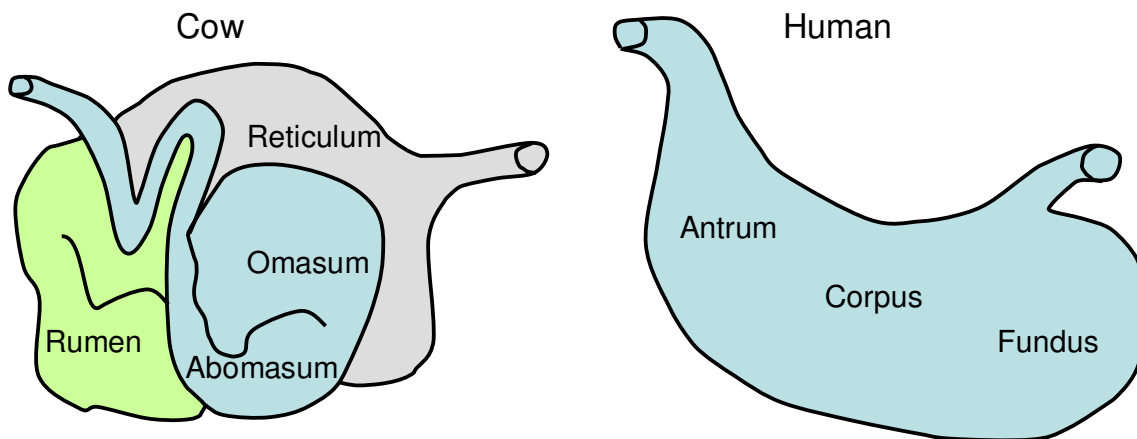


Figure 3. Cross-species example: Stomach

What would be an adequate criterion for distinguishing stomachs from non-stomach intestinal segments in, say, chicken embryos or adult flies? In the latter cases, there are no discontinuities by means of which an enlarged segment of a digestive canal could naturally be considered a stomach. So far, the problem is similar to the cell classification problem; but the 'stomach problem' takes on another dimension when we try to come up with descriptions which apply to ruminant organisms such as cows and sheep. It is commonly accepted that where most other species have only one stomach, cows have four distinct compartments: reticulum, rumen, omasum and abomasum. While the last is similar in function to the stomach of non-ruminants, the first three are important mainly for fermentation. The ontological trilemma here is whether: (i) the whole four-compartment system should be considered *one* stomach, (ii) each compartment should count as a stomach in its own right, or (iii) only the abomasum should so count. Biological reality in itself seems not to give us an answer. Every delimitation we select is in part a fiat delimitation. If we refuse to make such a delimitation, the term 'stomach' remains unacceptably vague.

# 3. Continua and Spatiotemporal Identification

Living systems are made of physical matter, and for any comprehensive life science ontology, a representation of this physical basis is of paramount importance. Mereotopological considerations, i.e., the formal analysis of parts, spaces, boundaries, connections, etc., play an important role in this regard, and have received intensive discussions in the medical informatics community in recent years.[10,14-26] In this section, we will show that many of the terms used for this purpose designate biological repeatables whose instances have a partially fiat spatial or spatiotemporal boundary; otherwise, to mention this opportunity again, the terms remain vague.

## 3.1.    Spatial Demarcation of Biological Objects

### Example: The boundaries of the kidney

Many biological objects, including bodily organs or cell organelles, have no complete bona fide spatial boundary, and we can use the human kidney as an illustration of this fact (Figure 4). How is the kidney delimited from other anatomical parts? At first glance, this might seem to be a trivial affair. The parenchyma, which is the functional tissue of the kidney, has a clearly defined bona fide boundary with the capsule, the outer part of the kidney. Therefore, we should include the kidney parenchyma in the kidney as part.

But let us next direct our view to the kidney's in- and outgoing vessels and ducts: the renal artery, the renal vein, and the ureter. These transport body fluids and form a system of confluences and ramifications which link the kidney to its neighboring structures. Therefore we have good reason to include all those vessels into the arterious and venous trees (i.e., the whole system of ingoing and outgoing blood vessels). Vessels inside an organ, however, are also part of this organ; it would be counterintuitive to regard parenchyma and duct structures as being spatially disconnected. They have to be regarded as structures that are part of both the kidney and the arterial or venous tree, respectively. But where then should the boundary of the kidney be drawn?
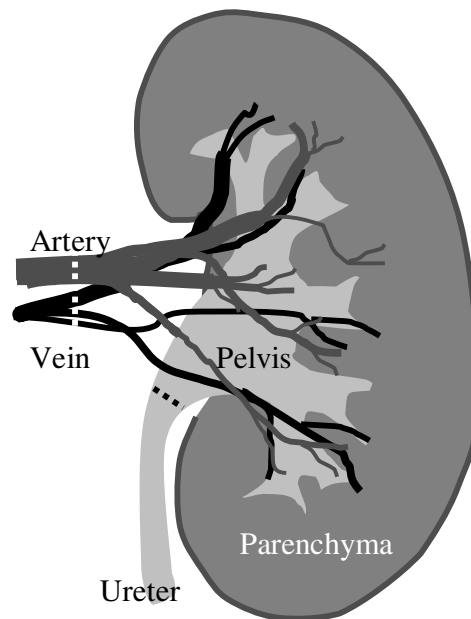
Figure 4. Kidney with fiat (dotted lines) and bona fide (solid) boundaries

We may resort to constructs such as the convex hull, or to a more local demarcation at the renal hilus (i.e., of the depression where the vessels and the ureter leave the organ). But in either case we have to insert boundaries by fiat[8,27] which cut somewhere through the ureter and the blood vessels. Yet while either result would fit well with a surgeon's view of the kidney, it fits less well with the view of an embryologist, who uses ontogenetic arguments in order to claim that the renal pelvis is not a part of the kidney but rather of the urinary tract. This is not an isolated problem, on the contrary. Nearly all seemingly well-delimited parts of biological organisms exhibit somewhere continuous spatial transitions into neighboring anatomical parts which can be viewed in different ways by different disciplines. Therefore, they have to be delineated partly by fiat before they can be spoken of in a well defined way.

## 3.2. Continua in the Demarcation of Biological States and Occurrents

### Example: Rheumatoid Arthritis

Complex biological phenomena are constituted by a number of different observable and unobservable states, events, and processes in various parts of the organism. The course of a disease is a unity of such entities that is characterized by abnormal functioning. The disease itself

11

is a dysfunction of some part of the body, whereby a number of associated observable states, events and processes in various parts of the body are regarded as symptoms of and as caused by the disease. In both cases, fiat delineation is involved; both the course of a disease and the symptom picture are delineated partly by fiat.

| Criterion | Definition |
|---|---|
| 1. Morning stiffness | Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement |
| 2. Arthritis of 3 or more joint areas | At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints |
| 3. Arthritis of hand joints | At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint |
| 4. Symmetric arthritis | Simultaneous involvement of the same joint areas (as defined in 2) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry) |
| 5. Rheumatoid nodules | Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxtaarticular regions, observed by a physician |
| 6. Serum rheumatoid factor | Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in < 5% of normal control subjects |
| 7. Radiographic changes | Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify) |

Table 1: Criteria for the classification of acute arthritis or rheumatoid arthritis[28]

Table 1 provides a list of criteria for the classification of acute arthritis or rheumatoid arthritis.[28] This table is accompanied by a set of classification rules. Accordingly, a patient shall be said to have rheumatoid arthritis if he/she has satisfied at least four of these seven criteria. Criteria 1 through 4 must have been present for at least six weeks.

The presence of a disease in an organism is frequently dissociated from the presence of any given bodily property. Rather it is dependent on several properties that may need to be weighted and combined in various ways as illustrated in Table 1. Such combinations have to be

agreed upon if medical practice and research is to run smoothly, refraining as much as possible from the use of vague terms. Effective diagnosis needs clear inclusion and exclusion criteria for clinical trials and therapeutic recommendations, and such criteria contain an essential admixture of pragmatically based decisions.

# 4. Continua and Classification of Relations

In discussions of biomedical ontology it has been argued that ontologists should stick to a limited set of relation types, for which a foundational status is claimed; see in particular the recently proposed OBO relation ontology.[1] The underlying assumption is that one's choice of relations when building an ontology should be as far as possible outside of any subjective or pragmatic judgment. But if (i) there is a continuum of relations, i.e., there is an infinite number of relations (as repeatables), and (ii) this continuum cannot be mapped on a continuous scale, then some pragmatics has to enter the scene. We observe, however, that this fact does not turn *prototypical* relations subjective. We will present three cases of relation continua: first, the relation *being located in* in relation to cavities and other hollow spaces; second, the relations *being part of* and *being contained in*; finally, the relations *being adjacent to* and *being physically connected to*.

## 4.1.   The Location Relation and Immaterial Parts

### Example: Biological Surfaces

Biological objects are replete with cavities (called 'immaterial anatomical entities'[23]). All organs are permeated by a tight maze of blood vessels. Larger hollow spaces are constituted by the tubular structures of the digestive and genitourinary tracts as well as by the ventricles of the central nervous system. Glands such as the pancreas or the salivary glands are marked by a tree system of excretory ducts, and the lungs by the airways of the bronchial tree. Specification of the locations of material objects within cavities and hollow spaces is therefore of the utmost importance when describing biological phenomena. We here use the relational predicate '*located_in*' in its most general sense: a material or immaterial physical object $a$ is located in some object $b$ if and only if the region covered by $a$ is totally included by the region covered by $b$[1,25].
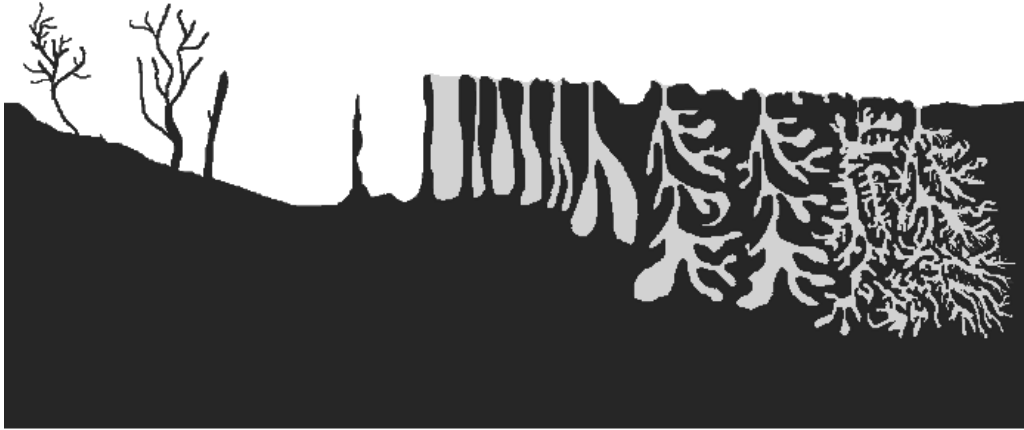
Figure 5. A prototypical biological surface showing the degree of variation of surface structures, e.g., in the intestinal tract.

This notion of 'being located in' is fundamental in those situations where the dislocation of biological material needs to be described. Since an object inside a hollow space that is located in turn within a material object (its host[29]) is generally itself considered to be inside the latter object, it is possible to view hollow spaces as standing in a special kind of parthood relation to their hosts.[20,25]

If the spatial extension of biological objects were strictly limited to their solid parts, an aspirated foreign body in an airway would have to be seen as lying outside the lung, and the blood cells in a heart ventricle as outside the heart. This obviously contradicts common-sense notions of anatomy. It is, however, subject to argument what criterion we should use in order to regard a certain spatial region[30] as 'being inside' a solid object. Using the convex hull in formulating such a criterion is hardly acceptable.[31] An earplug inside the outer ear is not thereby inside our body. A geometrical definition could be that a cavity is 'inside' its host if it is 'deeper than broader' (a criterion which would include the shaded regions in Figure 5). But it has problems because there are scenarios where individual differences between one organism and another are encountered, and because of the elasticity or mobility of the anatomical entity under scrutiny. Regions may change their shape: consider the space between hairs or intestinal villi (hairlike projections on mucous membranes), as well as the space within a closed hand. That is,

14

in relation to non-material objects, there can be a continuum of degrees of the relation *being spatially included*. The following situations describe four rough marks on this continuum:

1.  fully enclosed macroscopic spaces, which do not communicate with the exterior space (e.g., a cyst);

2.  almost fully enclosed spaces – the communicating region is very small compared to the extension of the space, and enclosed solids cannot move out without the help of physical forces (e.g., the ducts of excretory glands);

3.  spaces with an opening that allows enclosed solids to move out under the influence of external forces such as gravity (e.g., the outer ear);

4.  fully open spaces that are unable to withhold any solid object, but where small particles remain because of molecular adhesion to the host (e.g., the space between fingers).
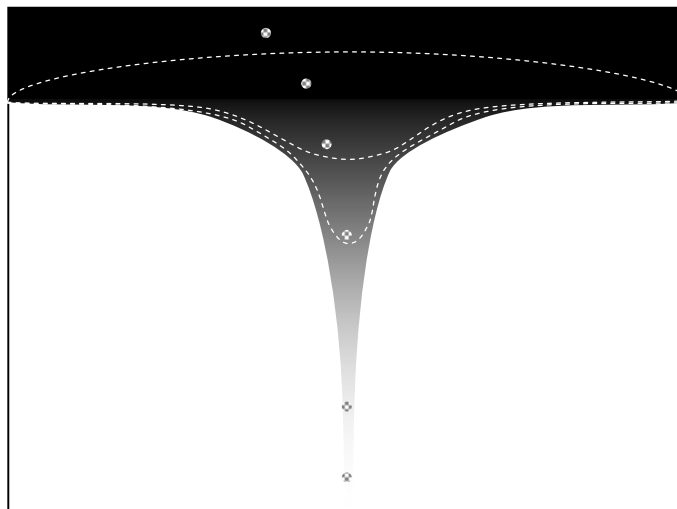


Figure 6. The continuum between internal and external location

Objective criteria may be added to refine this scheme, such as the relative size of the enclosed object and its physical state of aggregation. Figure 6 shows a typical design of an opening in a biological organism. A true continuum of possible boundaries gives rise to a continuum of relations between a prototypical 'being located outside' to 'being located inside'. According to the chosen boundary, the same objects (depicted as bullets) can be considered to be located inside or outside. Any simple bipartition here has to rely on some kind of fiat decision.

This does not mean that the effected delineation is completely arbitrary; only that the decision is based on reasons that do not stem from biological reality alone.

## 4.2.  The Continuum between Parthood and Containment

*Example: The continuous incorporation of foreign material by a cell*
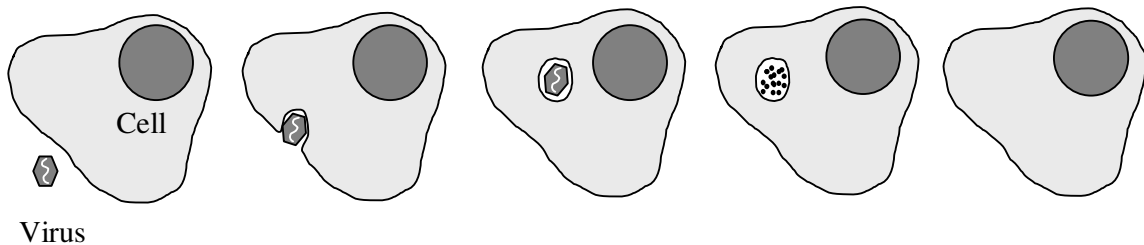
Cell

Virus

Figure 7. A virus being ingested by a macrophage.[32]

When do virus components become part of the macrophage?

Using the relation *located_in*, all the parts of a biologically complex whole are also located in that whole, but not everything that is located in the whole is thereby also a part. Rather it is only contained therein. For instance, my mitral valve is part of my heart; therefore it is also located in my heart. But a bolus of food that is located in my stomach is not a part of my stomach. Rather it is a case of *containment*. Containment is location-within without parthood.[1,25]

There are, clearly, prototypical cases of both *being part of* and *being contained in*; but there are also cases where biological reality does not give us any criterion by means of which we can draw a sharp line between parthood and containment. Look at Figure 7. Are the metabolites of a virus ingested by a killer cell a part of that cell? Is a water molecule in a cell a part of that cell even when it has just entered the cell? When do components of ingested food become part of the body? In all these cases, we have a continuous transition from prototypical containment through less prototypical containment and parthood relations to prototypical parthood. If one works with only two relations of parthood and containment, and does not want the terms to be vague, then a fiat boundary has to be drawn between them. According to the reigning theory[33],

16

mitochondria evolved from aerobic bacteria and came to live inside eukaryotic cells as endosymbionts. Are mitochondria parts of, or only contained in, eukaryotic cells? What about metastases, transfused blood, kidney transplants, artificial tissues, tooth fillings, prostheses, and heart pacemakers? Where in all these cases should the line be drawn between parthood and containment? The paper[25] provides some heuristic rules for drawing these distinctions by introducing criteria for deciding where a locative relationship can be refined to a mereological one and where it cannot. The application of these criteria, however, remains in part a pragmatic matter.

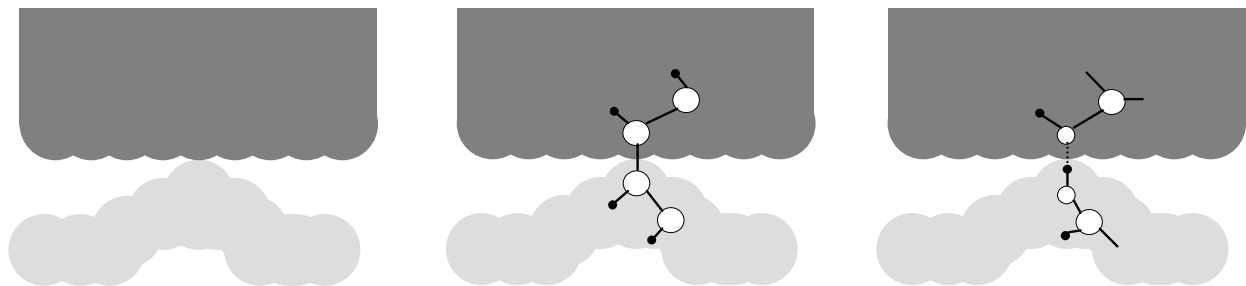## 4.3.  Adjacency and Physical Connection

*Example: Chemical Bonds*



Figure 8. Chemical bonds.

Left: abutting without bond, center: covalent bond, right: hydrogen bond.

Another continuum of relations exists between the mere spatial adjacency of two material objects and degrees of physical connectedness between such objects. When discussing this issue we must be aware of the ambiguities of the terms involved. The term 'connected' is in our and in its everyday sense more concrete than in mathematical topology, where two spaces are said to be connected if they share at least a single point. Similarly, 'adjacency' has several meanings. Sometimes two objects are said to be adjacent merely because they are represented as adjacent on a map or appear to be adjacent on an image, as for instance, where the liver and kidney abut on some coarse-grained radiological images. In reality, however, they are divided by the diaphragm. We will exclude all such cases here.

For our purposes, 'adjacency (between two objects)' might be defined as spatial contact without any other force or bond than that of the inevitable gravitational force between the objects. And it may be looked upon as a zero point to which different degrees of 'physical connectedness' can be related. Now, curiously enough, physicists and chemists define different kinds of *physical* connectedness by means of different kinds of *chemical* bonds.[34] Physical connectedness between objects, according to this terminology, requires chemical connectedness between atomic parts. Atoms can be connected by a number of different kinds of chemical bonds. The ones commonly distinguished are (in decreasing strength): covalent bonds, ionic bonds, metallic bonds, the bonds characteristic of hydrogen bonding, and *van der Waals* bonding. But more kinds of bonds (as well as continuous transitions between them[35]) can be distinguished, and our point is that these kinds can be regarded as representing a continuum of 'strongness of physical connectedness'.

Covalent bonds involve the sharing of valence electrons by two atoms; these atoms, therefore, are parts of the same molecule (Following the currently accepted theory, based on quantum mechanics, the covalent bond is understood as atomic orbitals of the interacting atoms forming a hybrid molecular orbital[36]). If these atoms are parts of two different physical objects, *a* and *b,* then the molecule they are part of has to be part of or to be identical with some overlapping and physically self-connected object, *c.* Such cases cannot possibly be called cases of adjacency. But where there is only a hydrogen bond between molecules, e.g., between two DNA strands, then we are closer to adjacency. Looser or tighter bonds between macromolecules are fundamental to many biochemical processes. Even non-covalent bonds can form quite stable complexes, such as antigen/antibody associations (see Figure 9) that are sufficient to completely change the physico-chemical properties of their compounds. For instance, such a reaction can instantaneously precipitate blood components in the case of transfusion of incompatible blood. In spite of the absence of covalent bonds in antigen/antibody complexes, most immunologists seem to support the view that the components of such complexes are physically connected and not merely adjacent.
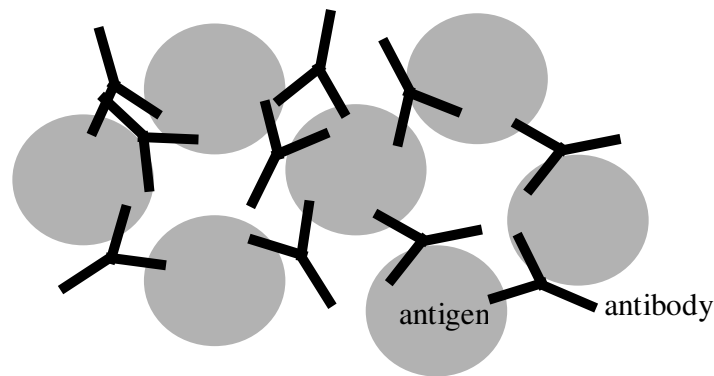
Figure 9. An antigen/antibody complex, forming a stable molecular
aggregate based on a multitude of weak hydrogen bonds.

## 5. Conclusions

We have now presented eight examples (normal-abnormal glucose concentrations, epithelial-neuroepithelial-neural cells, stomachs in different animals, the boundary of the kidney, the delineation of rheumatoid arthritis, the location of immaterial parts, parthood-containment relations, and connection-adjacency relations) where the corresponding entities have to be delineated partially by fiat. The reason is that, although there is in biological reality a continuum of more specific subtypes, there are obvious pragmatic obstacles to the creation of a term for each and every such type even if this would be theoretically possible. The resultant fiat character of our delimitations of biological kinds is no more curious than the fact that we turn continuous scales into scales with minimal discrete units when we measure reality using measuring instruments. The other way to turn a continuous scale into a finite number of slots is to designate the slots by means of vague or fuzzy terms. In small groups this might suffice, but in large groups with anonymous members it might easily lead to misunderstandings.

The phenomenon of continua is a challenge that any ontological theory has to address, and it has far-reaching consequences for the practice of constructing and maintaining bio-ontologies. It makes clear that two properly designed domain representations may nonetheless be non-interoperable because the two groups of domain experts involved may rely on different fiat delimitations of terms and classes. There are in such situations two different ways in which biomedical informaticians can try to create or construct interoperability:

19

(1) they can try to convince the domain experts that these have to come together and create consensus about where the fiat delimitations should be drawn, or

(2) they can try to create translation algorithms by means of which different fiat divisions of the same part of biological reality can be compared.

Since the problems connected with the first alternative are fairly obvious, we will confine ourselves to some brief remarks in relation to the second. The traditional translation algorithms in physics and chemistry (such as the conversion formulas for inches and centimeters and for Fahrenheit and Celsius measures of temperature) produce one-to-one mappings, but nothing of the sort is to be expected in biology. Let us return to our simple cell example in section 2.1. From Figure 2 three kinds of information about the relationship between the upper and the lower classification schemes can be gathered (only the first was mentioned earlier):

- disjunctive relationships: 'a *b*-cell is either a *b´*-, *c´*-, or *d´*-cell', 'a *d´*-cell is either a *b*-cell or a *c*-cell', and so on

- inclusive relationships: 'an *a´*-cell is always an *a*-cell', 'a *c´*-cell is always a *b*-cell', and 'an *e´*-cell is always a *c*-cell'

- stochastic relationships: on the assumption that all the lower-most kinds of cells are equally probable, one get relationships such as 'the probability that an *a*-cell is an *a´*-cell is 3/5', 'the probability that an *a*-cell is a *b´*-cell is 2/5', 'the probability that a *b´*-cell is an *a*-cell is 2/3', and so on.

It is impossible to tell a priori whether any algorithms useful for the information sciences can be constructed on the basis of relationships such as these. Here, the proof of the pudding is in the construction of real-world applications.

Summing up, the phenomenon of pragmatically based 'fiatness' has to be dealt with in any ontological account of living systems, but to what extent existing incompatibilities and lack of interoperability between properly designed ontologies can be reduced is partially a matter of human inventiveness. Thinking, however, that they can be *completely* taken away must probably be deemed a dream.

As a last comment, we want to emphasize that we have written our paper from the realist perspective that biologists normally take for granted.[37] Nonetheless, we are firmly convinced that

in whatever way conceptualists try to make their case that repeatables (universals) can only exist in language, they will have to introduce a distinction that in some way or other comes near to our distinction between (a) continua in biological reality and (b) pragmatically useful fiat and finite discontinuous delimitations of such continua. In so far as this leads to structurally similar solutions, the gap between the realist and the conceptualist perspectives can to this extent be bridged for the practical purposes of biomedical ontology engineering.

**References**

1. B. Smith, W. Ceusters, B. Klagges, J. Köhler, A. Kumar, J. Lomax, C. Mungall, F. Neuhaus, A. Rector and C. Rosse, "Relations in Biomedical Ontologies". *Genome Biology*, 6(5), 2005.
2. B. Smith, W. Kusnierczyk, D. Schober and W. Ceusters, "Towards a Reference Terminology for Ontology Research and Biomedical Domain", KR-MED 2006. *Biomedical Ontologies in Action. Second International Workshop of the AMIA Working Group on Biomedical Knowledge Representation.* Baltimore, Nov 8, 2006.
3. P. G. Baker, C. A. Goble, S. Bechhofer, N. W. Paton, R. Stevens and A. Brass, "An Ontology for Bioinformatics Applications", *Bioinformatics*, 15(6), 510–520, 1999.
4. C. Rosse and J. L. V. Mejino, "A Reference Ontology for Bioinformatics: the Foundational Model of Anatomy", *Journal of Biomedical Informatics*, 36:478–500, 2003.
5. C. Rosse, L. G. Shapiro and J. F. Brinkley, "The DIGITAL ANATOMIST Foundational Model: Principles for Defining and Structuring its Concept Domain", In: C. G. Chute, editor, *Proceedings of the 1998 AMIA Annual Fall Symposium.* Philadelphia, PA:Hanley & Belfus, 1998, 820–824.
6. Open Biological Ontologies (OBO) [http://obo.sourceforge.net], 2007; see also http://obofoundry.org. Last accessed March 6th, 2007.
7. I. Horrocks, P. F. Patel-Schneider and F. van Harmelen, "From SHIQ and RDF to OWL: The making of a Web Ontology Language", *Journal of Web Semantics*, 1(1), 7–26, 2003.
8. B. Smith, "Fiat Objects", *Topoi* 20(2), 2001, 131–148.
9. B. Smith, "Beyond Concepts, or: Ontology as Reality Representation", In: A. C. Varzi and L. Vieu, editors, *Formal Ontology in Information Systems. Proceedings of the 3rd International*. Amsterdam: IOS Press, 2004, 73–84.
10. O. Bodenreider, B. Smith and A. Burgun, "The Ontology-Epistemology Divide: A Case Study in Medical Terminology", In: A. C. Varzi and L. Vieu, editors, *Formal Ontology in Information*

*Systems. Proceedings of the 3rd International Conference,* Amsterdam: IOS Press, 2004, 185–195.

11. J. J. Cimino and B. Smith, "Introduction: International Medical Informatics Association Working Group 6 and the 2005 Rome Conference". *Journal of Biomedical Informatics*, 39(3), 249–251, 2006.

12. R. J. Brachman, "On the Epistemological Status of Semantic Networks", In: N. V. Findler, editor, *Associative Networks: Representation and Use of Knowledge by Computers*, New York: Academic Press, 1979, 3–50.

13. A. McCray, T. Yang and R. Patil, "The UMLS Semantic Network", In: L. C. Kingsland III, editor, *Proceedings of the 13th Annual Symposium on Computer Applications in Medical Care*, New York, N.Y.: IEEE Computer Society Press, 1989, 503–507.

14. A. L. Rector, A. Gangemi, E. Galeazzi, A. J. Glowinski and A. Rossi Mori, "The GALEN Model Schemata for Anatomy: Towards a Re-usable Application-Independent Model of Medical Concepts", In: P. Barahona, M. Veloso and J. Bryant, editors, *Proceedings of the 12th Conference of the European Federation for Medical Informatics*, Amsterdam: IOS Press, 1994, 229–233.

15. P. J. Neal, L. G. Shapiro and C. Rosse, "The DIGITAL ANATOMIST Structural Abstraction: A Scheme for the Spatial Description of Anatomical Entities", In: C. G. Chute, editor, *Proceedings of the 1998 AMIA Annual Fall Symposium*. Philadelphia, PA: Hanley & Belfus, 1998, 423–427.

16. C. Rosse, J. L. V. Mejino, B. R. Modayur, R. Jakobovits, K. P. Hinshaw and J. F. Brinkley, "Motivation and Organizational Principles for Anatomical Knowledge Representation: The DIGITAL ANATOMIST Symbolic Knowledge Base", *Journal of the American Medical Informatics Association*, 5(1), 17–40, 1998.

17. J. L. V. Mejino and C. Rosse, "Conceptualization of Anatomical Spatial Entities in the DIGITAL ANATOMIST Foundational Model", In: N. M. Lorenzi, editor, *Proceedings of the 1999 Annual Symposium of the American Medical Informatics Association*, Philadelphia, PA: Hanley & Belfus, 1999, 112–116.

18. J. Rogers and A. L. Rector, "GALEN's Model of Parts and Wholes: Experience and Comparisons". In: J. M. Overhage, editor, *Proceedings of the Annual Symposium of the American Medical Informatics Association*, Philadelphia, PA: Hanley & Belfus, 2000, 714–718.

19. R. Schubert, K. Priesmeyer, H. C. Wulf and K. H. Höhne, "VOXEL-MAN (WEB), Basistechnologie zur modellbasierten multimedialen Repräsentation von komplexen räumlichen Strukturen", *Künstliche Intelligenz*, 14(1), 44–47, 2000.

20. S. Schulz and U. Hahn, "Mereotopological Reasoning about Parts and (W)holes in Bio-Ontologies", In: C. Welty and B. Smith, editors, *Formal Ontology in Information Systems. Collected Papers from the 2nd International FOIS Conference*, New York, NY: ACM Press, 2001, 210–221.

21. J. L. V. Mejino, A. V. Agoncillo, K. L. Rickard and C. Rosse, "Representing Complexity in Part-whole Relationships within the Foundational Model of Anatomy". In: M. A. Musen, editor, *Proceedings of the 2003 Annual Symposium of the American Medical Informatics Association*, Philadelphia, PA: Hanley & Belfus, 2003, 71–75.

22. S. Schulz and U. Hahn, "Representing Natural Kinds by Spatial Inclusion and Containment", In: R. L. de Mántaras and L. Saitta, editors, *Proceedings of the 16th European Conference on Artificial Intelligence*, Amsterdam: IOS Press, 2004, 403–407.

23. B. Smith, J. L. V. Mejino, S. Schulz and C. Rosse. "Anatomical Information Science", In: *COSIT 2005: Spatial Information Theory. Foundations of Geographic Information Science*, New York: Springer, 2005, 149–164.

24. S. Schulz and U. Hahn, "Part-Whole Representation and Reasoning in Biomedical Ontologies", *Artificial Intelligence in Medicine*, 34(3), 2005.

25. S. Schulz, A. Kumar and T. Bittner, "Biomedical Ontologies: What Part-of Is and Isn't", *Journal of Biomedical Informatics*, 39(3), 2006, 350-361.

26. A. Rector, J. Rogers and T. Bittner, "Granularity Scale and Collectivity: When Size does and doesn't Matter", *Journal of Biomedical Informatics*, 39(3), 2006, 333-349.

27. B. Smith and A. C. Varzi, "Fiat and Bona Fide Boundaries", *Philosophy and Phenomenological Research*, 60(2), 401–420, 2000.

28. F. C. Arnett, S. M. Edworthy, D. A. Bloch, D. J. McShane, J. F. Fries, N. S. Cooper, L. A. Healey, S. R. Kaplan, M. H. Liang, H. S. Luthra, et al., "The American Rheumatism Association 1987 Revised Criteria for the Classification of Rheumatoid Arthritis", *Arthritis and Rheumatism.*, 3(31), 315–324, 1988.

29. R. Casati and A. C. Varzi, *Holes and Other Superficialities*, Cambridge, MA: MIT Press, 1994.

30. M. Donnelly, "Relative Places". In: A. C. Varzi and L. Vieu, editors, *Formal Ontology in Information Systems. Proceedings of the 3rd International Conference*. Amsterdam: IOS Press, 2004, 249–260.

31. S. Schulz and U. Hahn, "Spatial Location and its Relevance for Terminological Inferences in Bio-Ontologies". *BMC Bioinformatics*, 2007, accepted for publication.

32. D. A. Randell and A. G. Cohn, "Modelling Topological and Metrical Properties in Physical Processes", In: R. J. Brachman, H. J. Levesque, R. Reiter, editors, *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*. Toronto: Morgan Kaufmann 1989, 357-368.

33. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell*, New York: Garland Science, 2002.

34. L. Pauling, "The Nature of the Chemical Bond". Ithaca: Cornell University Press, 1960.

35. N. C. Norman, "Periodicity and the s- and p-Block Elements". Oxford University Press, 2006.

36. J. Daintith, *Oxford Dictionary of Chemistry*. New York: Oxford University Press, 2004.

37. I. Johansson, "Bioinformatics and Biological Reality", *Journal of Biomedical Informatics*, 39(3), 2006, 274-287.