From: *Proceedings of Bio-Ontologies Workshop*, Detroit, June 24 (ISMB 2005)

On the proper treatment of pathologies in biomedical ontologies

Barry Smith[1,2*] *and Anand Kumar*[1]

[1]*IFOMIS, University of Saarbrücken, Germany;* [2]*Department of Philosophy, University at Buffalo, USA*

ABSTRACT

Motivation: In previous work on biomedical ontologies we showed how the provision of formal definitions for relations such as *is_a* and *part_of* can support new types of automated reasoning about biomedical phenomena. We here extend this approach to the *transformation_of* characteristic of pathologies.

1 INTRODUCTION

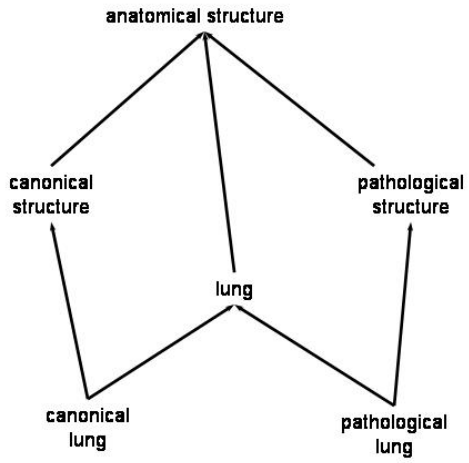
Pathological entities exist in the biological domain at various levels of granularity, from cell components to whole populations. Such entities involve in every case processes – for example, the patho-physiological pathways involved in the etiology of a range of different types of diseases – which cross granular levels. The proper treatment of pathologies within a formal-ontological framework thus demands a facility for simultaneous representation of entities existing at different levels of granularity. We here present the outlines of such a framework, against the background of our previous work on the treatment of relations, time, change and granularity in biomedical ontologies especially within the context of the OBO Relation Ontology [1].

We shall confine our focus here on pathological continuant entities, which include neoplasms, blisters, punctures, fractures, portions of pus, portions of amyloid but also the bearers of such entities (a wounded knee, a carcinomatous colon, a fractured tibia, and so on).

To say that such entities are *continuants* is to say that they endure as self-identical through time even while undergoing a variety of different sorts of changes. This is in contrast to occurrents (events, processes, happenings), which unfold themselves through time in successive temporal phases or stages which share no parts in common.

Some pathological continuants are subclasses of the class *anatomical structure*, a designation we take over from the Foundational Model of Anatomy (FMA) [2]. Such pathological structures are classified as pathological relative to others classified as 'normal' or 'canonical'. We here leave open the question of what it is in virtue of which a given instance is to be classified as normal. We shall assume that both are subclasses of the same parent *anatomical structure*, whose instances include both pathological and normal structures, along the lines indicated in Figure 1.

Figure 1. A simple classification-schema for pathological structures



We then have, for example:

pathological colonic mucosal cell is_a pathological structure

canonical colonic mucosal cell is_a canonical structure

colon with pathological features is_a pathological structure

normal colonic mucosal cell is_a colonic mucosal cell

pathological colonic mucosal cell part_of colon

Here *is_a* and *part_of* are defined as follows, using variables We use *A*, *B*, *C* ... to range over classes (universals, types) of pathological continuants, *c*, *c′*, ... to range over the instances of such classes, and *t*, *t′*, ... to range of instants of time:

A is_a B =def. for all *c*, *t*, if *c* **instance_of** *A* **at** *t* then *c* **instance_of** *B* **at** *t*.

A part_of B =def. for all *c*, *t*, if *c* **instance_of** *A* **at** *t* then there is some *c′* such that: *c′* **instance_of** *A* **at** *t* and *c* **part_of** *c′* **at** *t*,

where '**part_of**' is the instance-level part relation (between, for example, this particular cell and this particular lung). The reference to *times* in these definitions is designed to do justice to the fact that one and the same entity can instantiate different classes and gain and lose parts in the course of time. Note also the all-some structure of the definition of *part_of*, which is characteristic of almost all relations between classes of the sort treated by biomedical ontologies.

Some (but not all) kinds of pathological structures are such that their instances are in every case transformations of ca-

* Corresponding author

nonical structures of entities of a given kind existing earlier. We define *transformation_of* as a relation between continuant classes:

A transformation_of B =def. for all *t* and all *c*, if *c* is an instance of *A* at *t*, then there is an earlier time *t′* at which *c* is an instance of *B*, and for no *t*, *c* is *c* an instance of both *A* and *B* at *t*.

The pathological colon mucosal cell can be a transformation either of the canonical colon mucosal cell or of its precursor, depending on whether the pathology is hereditary or acquired. Because *transformation_of* is transitive, however, we can assert quite generally:

pathological colon mucosal cell transformation_of canonical colon mucosal cell precursor

given that in every case:

colon mucosal cell transformation_of colon mucosal cell precursor

The temporal relationships between canonical entities and entities with pathological features have not been sufficiently addressed in ontologies thus far, and even developmental ontologies utilizing the methodology of stages have preferred not to incorporate a formal machinery for dealing explicitly with times [3]. Thus most of the instances of colonic mucosal cells with pathological features are a transformation of instances of normal colonic mucosal cell. In some cases, the former is a transformation of a precursor entity of the latter. Transtemporal relations of this sort are not recorded in the National Cancer Institute Thesaurus, where for example no relations are asserted between the two classes *abnormal cell* and *normal cell*, not even that they have a common parent: *cell* [4]. Transformation relations are also absent in the SNOMED CT terminology [5]. A relation which we do find in SNOMED CT, however, is that of location, for example in:

lung cyst finding_site lung structure

Better, however, would be to eliminate the epistemological connotations of '*finding_site*' by using a location relation such as GALEN's *locus* [6] or OBO's *located_in* [1]:

A located_in B =def. for all *c*, *t*, if *c* **instance of** *A* **at** *t* then there is some c_1 such that: C_1c_1t and *c* **located_in** c_1 **at** *t*.

PathBase [7] provides a subsumption hierarchy for various pathological processes. It has

endoplastic reticulum defect is-a subcellular defect

This relation can thus be used with the colon cell assertions above to generate for example:

pathological colon mucosalcell with endoplastic reticulum defect is-a pathological colon mucosal cell with subcellular defect

Further implications which can be drawn are:

endoplastic reticulum defect located-in endoplastic reticulum

fpathological colon mucosalcell has_level_of_granularity cell.

where levels of granularity can be inferred from an *is_a* hierarchy such as that of the FMA, e.g. from the fact that the colon is an organ we can infer:

colon has_level_of_granularity organ

and thus also that both *canonical and pathological colon have this same level of granularity.*

2 CANCER STAGING

We can use the framework to capture some of the information contained in systems for cancer staging such as the TNM (for: Tumour, Node, Metastasis) system, which is used to classify pathological states into specific categories important for carcinoma management. The T2 stage, for example, is defined as: *carcinoma has invaded the muscularis mucosa of the colon wall* and the T1 stage as: *carcinoma has invaded the mucosa*. N1 designates a stage with one to four lymph nodes, M1 a stage where a metastasis is present in a non-contiguous part of the body. We can then assert that a pathological entity of the type *carcinoma in colon* at stage T2N1M1 must be a transformation of either a T1N1M1 or a T2N0M1 structure. If a carcinoma is a transformation from T1N1M1 to T2N1M1 then there has occurred a process of the type *muscularis mucosa invasion*. If there is a transformation from T2N0M1 to T2N1M1 then this implies that the last process to take place was one of lymph node metastasis. If there is a transformation from T2N1M0 to T2N1M1, then this implies that the last process to take place was one of metastasis to a non-contiguous body region.

3 CONCLUSION

We have sketched a formal approach to class-class relations in the realm of pathologies that is designed to support new types of cross-granular reasoning and also reasoning about entities which exist at different points in time, for example in the domain of cancer staging.

REFERENCES

[1] Smith B, Ceusters W, Klagges BER, Kohler J, Kumar A, Lomax J, Mungall CJ, Neuhaus F, Rector AL, Rosse C. Relations in Biomedical Ontologies. *Genome Biology*, 2005;6:R46.

[2] Rosse C, Mejino JLV Jr: A reference ontology for bioinformatics: The Foundational Model of Anatomy. *J Biomed Informatics*, 2003; 36:478-500.

[3] Aitken S: Formalising concepts of species, sex and developmental stage in anatomical ontologies, *Bioinformatics*, in press.

[4]Ceusters W, Smith B, Goldberg L: A terminological and ontological analysis of the NCI Thesaurus, *Meth Inform Medicine*, in press.

[5] http://nciterms.nci.nih.gov/

[6] http://www.opengalen.org/

[7] http://eulep.anat.cam.ac.uk/Pathology_Ontology/MPATH-dynamic.php