



FREIE UNIVERSITÄT BOZEN
LIBERA UNIVERSITÀ DI BOLZANO
FREE UNIVERSITY OF BOZEN - BOLZANO



Faculty of Computer Science, Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy
Tel: +39 04710 16000, fax: +39 04710 16009, <http://www.inf.unibz.it/krdb/>

KRDB Research Centre Technical Report:

Constraints for representing transforming entities in bio-ontologies

C. Maria Keet

Affiliation	KRDB Research Centre, Free University of Bozen-Bolzano, Italy
Corresponding author	Maria Keet keet@inf.unibz.it
Keywords	transformation_of relation, OntoClean, bio-ontologies, Relation Ontology, ontology development
Number	KRDB09-2
Date	April 22, 2009
URL	http://www.inf.unibz.it/krdb/pub/

©KRDB Research Centre

This work may not be copied or reproduced in whole or part for any commercial purpose. Permission to copy in whole or part without payment of fee is granted for non-profit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of the KRDB Research Centre, Free University of Bozen-Bolzano, Italy; an acknowledgement of the authors and individual contributors to the work; all applicable portions of this copyright notice. Copying, reproducing, or republishing for any other purpose shall require a licence with payment of fee to the KRDB Research Centre.

Constraints for representing transforming entities in bio-ontologies

C. Maria Keet

KRDB Research Centre, Free University of Bozen-Bolzano, Italy
keet@inf.unibz.it

Abstract. Things change—develop, mature morph—but not everything in the same way. Representing this knowledge in biomedical ontologies faces issues on three fronts: what the category of the participating objects are, which type of relations they involve, and where constraints should be added. More precise distinctions can be made by using OntoClean’s properties and a novel status property that is generalised from formal temporal conceptual data modeling. Criteria are identified, formulated in 17 additional constraints, and assessed on applicability for representing transformations more accurately. This enables developers of bio(medical) ontologies to represent and relate entities more precisely, such as monocyte & macrophage and healthy & unhealthy organs.

1 Introduction

Much effort has been invested in development of biomedical ontologies such as the Gene Ontology, Foundational Model of Anatomy, Cell Cycle, SNOMED and so forth [1–3]. This started from preliminary categorizations and vocabularies of biological entities that primarily focussed on endurants. With the maturing of bio-ontologies, ontology development tools, and ontology languages, this scope is broadening to temporal aspects. For instance, the Relation Ontology (RO) [4] contains the hitherto underused *transformation_of* and *derives_from* relations and more evolution constraints have been identified for temporal conceptual modelling [5]. To be able to say one thing transforms into another, one has to identify the transforming entity x , the transformed entity y , and that the entity preserves its identity irrespective of the transformation while instantiating distinct classes at distinct points in time. What kind of entities are x and y ; phased sortals, roles, or merely different states? How should one deal with the temporality to achieve implementable knowledge bases that can handle representations of, and reasoning over, transforming entities? Likely candidates are the categories phased sortal, role, and state of an entity, but each one carries assumptions that reveal implicit semantics of what constitutes a transformation. Put differently, by considering the relata of the transformation relation, one can make explicit the processes of various constrained transformations.

The aim of this paper is to characterise constraints on changing entities more precisely in a way so that it can aid a domain ontology developer to introduce

or improve the representation of transforming entities, using an approach that is sufficiently generic to be extensible also to other types of change and to represent as much as possible about temporal aspects in the, to date, a-temporal commonly used ontology languages. We will commence with an analysis from an ontology development perspective, including the RO’s *transformation_of* relation from biomedical ontologies [4] and basic OntoClean notions, to formulate a set of basic constraints (section 2). In section 3, we extract implicit ontological commitments from advances in temporal knowledge representation, which enables us to define additional constraints on the entities and the process of transformation. The resulting 17 constraints are assessed on their usefulness through a re-examination of the classical example of the transformation from caterpillar to butterfly and by assessing some recent problems in biomedical ontology development, such as representing the changes of blood cells, in section 4. We discuss and conclude in sections 5 and 6.

2 Analysis from an ontology development perspective

2.1 Preliminaries

Underspecification in the Relation Ontology. Let us recollect the definition of the *transformation_of* relation in the RO [4]: *C transformation_of C₁ = [definition] C and C₁ for all c, t, if Cct, then there is some t₁ such that C₁ct₁, and t₁ earlier t, and there is no t₂ such that Cct₂ and C₁ct₂.* This definition reveals two issues. First, it is ignorant of the distinction between the cases of unidirectional transformations versus where some instance of C₁ may, after transforming into C, transform back into C₁; e.g., the transformations of erythrocytes (red blood cells) into echinocytes and back again, and of a healthy organ into a non-healthy organ and back again. Second, and more important for developing logic-based ontologies, the RO definition does not say how the entities undergoing transformation are able to change and yet keep their identity, other than through the use of the same variable ‘c’. This under-specification can lead to unintended models of the theory; e.g., a given particular remains unchanged, but either C or C₁ changes due to increased understanding or correcting a representation error in the domain ontology. Clearly, this does not meet the *intention* of the authors of the above definition. To exclude such unintended models, we have to make explicit in the definition that the individual changes somehow. Let *a* instantiate universals C_s and C_t—with *s* for source and *t* for target the source is transformed into—at the two different times, then we have to assert that ‘enough’ properties are shared by *a* ∈ C_s and *a* ∈ C_t so that they can be identified as the same individual—in shorthand notation: *a_s =_i a_t*—while other properties π₁ . . . π_n are lost or gained so that after the transformation, *a* instantiates a different universal. A basis for a generic definition might then be:

Definition 1 (C_t transformation_of C_s). *Let C_t be the target and C_s the source universal, x, y range over instances and t₀, . . . , t_n range over points in time, then C_t(x) transformation_of C_s(y) iff for all x, there exist y, t₀, . . . , t_n,*

if $C_t(x, t_0)$, then there is some t_1 such that $C_s(y, t_1)$, $t_1 < t_0$, x and y have the same identity criterion ($x =_i y$), x and y differ in at least one other property π_i , and there does not exist a t_2 such that $C_t(x, t_2)$ and $C_s(y, t_2)$.

In the next sections, we elicit additional constraints to model and achieve the intended behaviour of *transformation_of* in ontologies and knowledge bases.

Basic notions from OntoClean. The main proposal in ontology to represent the entities involved in transformations is that of *phased sortals*. Postulating that C_t and C_s are phased sortals (or another category), brings forward constraints on the participating entities of the transformation relation that, in turn, may affect constraints on the relation itself, and certainly influences ontology development. To arrive at the point where we can specify unambiguously its how and where, we first summarise a section of [6, 7] that form the basis of the OntoClean methodology¹ [12]. The relevant aspects for the current scope are the notion of identity, the meta-property rigidity, and identity criteria. Identity of an instance, and by extension also the universal it instantiates, focuses on the problems of distinguishing an instance of one class from other instances of that class by means of a characteristic property, which is unique for that whole instance, and is a relation that each thing has to itself and to nothing else; consequently, the property is essential to the universal or concept that it instantiates. The property *rigidity* is based on this notion of essential property, two of which are relevant in the current scope [6].

Definition 2 (+R). A rigid property ϕ is a property that is essential to all its instances, i.e., $\forall x\phi(x) \rightarrow \Box\phi(x)$

Definition 3 (~R). An anti-rigid property ϕ is a property that is not essential to all its instances, i.e., $\forall x\phi(x) \rightarrow \neg\Box\phi(x)$

For instance, the properties being a *Patient* and being a *Caterpillar* are anti-rigid and being a *Person* and being a *Herbivore* are rigid that may subsume *Patient* and *Caterpillar*, respectively. Objects can keep their identity through the changes it may undergo during its lifetime, thereby exhibiting different properties at different times; thus, we consider *diachronic* identity (cf. synchronic identity) which is about establishing that two entities are the same at two different points in time. Setting aside *how* we can identify instances—it is philosophically difficult [9], but one always follow common practice and use identifiers in the knowledge

¹ There are some refinements regarding identity criteria, sortal concepts, ‘types’ of identity, and OntoClean [8–10], but (i) they do not affect the principal method of categorisation (in Fig.1, below) and (ii) they go beyond immediate applicability to domain ontology development, because both the usage of the guiding principles of the present taxonomy of properties [6] for OntoClean and adding formal definitions to a universal in an OWL ontology (as opposed to relying on identifiers) would already be a considerable step forward. Moreover, the extension toward OntoClean 2.0 [10] with temporal aspects fit well with the temporal language we build upon in section 3 (as has been demonstrated by [11] for rigidity and part-whole relations).

base—the instances and its corresponding universal have identity criteria (IC), which are both necessary and sufficient for identity [6, 7]; properties carrying an IC are called *sortals* and only +R properties are sortal properties. In addition, considering we need sortal properties for ICs, we can distinguish between a rigid property ϕ that either carries the necessary IC Γ or it carries the sufficient IC Γ (see also [6] definitions 5 and 6) and if $\forall x(\phi(x) \rightarrow \psi(x))$ and Γ is a necessary IC for ψ , then ϕ replaces ψ (the necessary condition of the subsumee ϕ refines the necessary condition of its subsumer ψ). Now we can introduce two other definitions from [6], which are important for phased sortals.

Definition 4 (+I). *A property that is not rigid carries an IC Γ iff it is subsumed by a rigid property carrying Γ .*

Definition 5 (+O). *A property ϕ supplies an IC Γ iff i) it is rigid; ii) it carries Γ ; and iii) Γ is not carried by all the properties subsuming ϕ .*

Thus, an +O property brings in its own identity criterion as opposed to just carrying it; conversely, properties that do not carry identity or do not supply identity are marked with -I and -O, respectively. Fig.1 summarises a classification of property kinds based on the R, O, and I meta-properties.

+O	+I	+R	+D	Type	Sortal	
			-D			
-O	+I	+R	+D	Quasi-Type		
			-D			
-O	+I	\sim R	+D	Material role		
-O	+I	\sim R	-D	Phased sortal		
-O	+I	\sim R	+D	Mixin		
			-D			
-O	-I	+R	+D	Category		Non-Sortal
			-D			
-O	-I	\sim R	+D	Formal role		
			-D			
-O	-I	\sim R	+D	Attribution		
			-D			

Fig. 1. Ontological properties [6]. +R: rigid; \sim R anti-rigid and \neg R semi-rigid, +O own identity supplied, +I identity carried, and +D dependent.

2.2 Characterising the transforming entities

Given the classification (Fig.1), this translates into the following set of basic constraints CT1-CT5 for phased sortals, where C_p is a *parent* of the focal concept or universal (C_s , C_t , etc). CT2 ensures that the entity that changes local IC when going from one ‘phase’ to another, still can be identified as the same entity (thanks to property inheritance from C_p). CT3 can be derived from CT1, CT2, and Definition 4.

- (CT1) A phased sortal does *not* supply an IC, i.e., -O
- (CT2) A phased sortal *must* be subsumed by C_p that has +O
- (CT3) A phased sortal carries an IC, i.e., +I
- (CT4) A phased sortal is a sortal
- (CT5) A phased sortal is anti-rigid, i.e., $\sim R$

It is now straightforward to demonstrate that *if the continuants C_t and C_s of the transformation_of relation are both categorised as phased sortals, then:*

- (CT6) C_t and C_s both must be subsumed by C_p

As we shall see later with the examples, CT6 is particularly important. Observe that CT6 does *not* state that C_p has to be the *direct* common subsumer of C_t and C_s . It immediately follows from Definition 5, the classification (Fig.1), and CT6 that C_p must be a type (CT7), because it is the only kind that has +O. Further, phased sortals *together* with the *transformation_of* relation cover the implicit requirement that phased sortals never can occur ‘alone’, because in order to phase, one needs *at least 2* phased sortals (CT8).

- (CT7) C_p must be a type (+O+I+R)
- (CT8) Each type that subsumes phased sortals, which are related through the *transformation_of* relation, must subsume *at least two* phased sortals

It is possible that C_t and C_s represent universals for different ‘states’ of a common subsumer C_p , provided the states are distinct enough. However, how does state differ from phase? Intuitively, phases succeed one another and have no circularity, whereas states (e.g., *sensu DOLCE* [13]) permit going back- and forward between alternating states. Let us take the latter assumption, then *if C_t transformation_of C_s and C_t and C_s are categorised as states, then the following constraints must hold:*

- (CT9) C_t and C_s must carry identity (+I)
- (CT10) If C_t is a transformation of C_s , then it is possible, but not necessary, that at a later point in time C_t transforms back into C_s
- (CT11) C_t and C_s have meta-properties that are either $\sim R$ or +R

This reduces C_t and C_s to being either types, quasi-types, material roles, or phased sortals. They cannot be roles, because an instance can have more than one role at the same time (e.g., *Patient* and *Employee*), thereby violating Definition 1. That is, toles are ‘played’ by an entity, which means that they are externally dependent (+D) on some other entity (see also [6, 14]). For instance, a particular organism x can play the role of both predator and prey, which it can do *at the same point in time* and move in and out of such a role, but it does neither *transform* from one into the other nor entails that x must play at least two roles. Therefore, C_t and C_s cannot be categorised as roles (unless we remove from the definition the constraint “and there is no t_2 such that $C_t(x, t_2)$ and $C_s(y, t_2)$ ”). C_s and C_t cannot be types either, because then one cannot ensure a common IC for diachronic identity. They can be both quasi-types (-O+I+R),

which is a kind that enables one to group entities based on properties that *do not affect identity*, such as being a Herbivore. In order to prevent C_t and C_s to be also types or roles, we add CT12. Further, we make explicit an assumption that transformation of phased sortals is unidirectional (CT13).

- (CT12) If C_t and C_s are categorised as states, they are neither both types nor both roles
- (CT13) If C_t is a transformation of C_s , C_t and C_s are phased sortals, then it is *not* possible that at a later point in time C_t is a transformation of C_s , i.e., C_t does not transform back

Thus, based on foundational notions of Ontology, CT1-CT13 offers a more precise categorisation for the relation of the *transformation_of*, as well as their position in a taxonomy.

3 Constraints from temporal conceptual modelling

The challenge of how to represent changing entities has been investigated also from formal and engineering perspectives, most notably with formal temporal conceptual data modelling. Artale et al [5] have formalised the well-known core elements of temporal databases in \mathcal{DLR}_{US} , a temporal description logic that is an expressive fragment of the first order temporal logic $L^{\{\text{Since, Until}\}}$, and a corresponding \mathcal{ER}_{VT} temporal conceptual data modelling language that extends EER. However, because it is intended for conceptual data modelling, they do not take into account the kind of classes like phased sortal—hence, nor the possible consequences on the logical implications—and they include *evolution constraints* other than transformation, such as dynamic extension and generation², which address more ways of change than is currently deemed necessary for (bio-)ontologies. Therefore, we only consider their notion of status classes here. *Status* is associated to a class to log the evolving status of membership of each object in the class and the relation between the statuses. The four possible statuses are scheduled, active (for backwards compatibility with an a-temporal class), suspended, and disabled (based on [16], see Fig.2), with corresponding formalised constraints, e.g., that an object that is disabled can never become active again and existence persist until it becomes disabled [5]. In this setting, one can assert, e.g., that *Caterpillar* is temporally related to *Butterfly*: when at t_0 object $o \in \text{Caterpillar}$ (and $o \in \text{Scheduled-Butterfly}$) starts transforming into

² For instance, *generation* deals with objects that have different object identifiers (different individuals) and emergence of new instances starting from a *set* of instances of the same type: in *production* the source objects survive the generation process and during *transformation* all instances involved in the process are consumed. This notion of ‘transformation’ has other intended semantics than the definition of *transformation_of*, although it is not excluded based on the examples in [4], in particular regarding development. The concept of generation might be applicable to RO’s *derives_from* relation [15, 4].

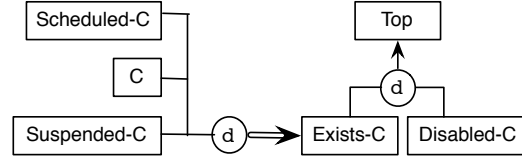


Fig. 2. EER graphical rendering of the \mathcal{DLR}_{US} axioms for status classes; **Top** represents the universe of discourse, arrow denotes subtype, and the encircled **d** means disjoint (Source: [5]).

an instance of *Butterfly*, then we have immediately after (\oplus in \mathcal{DLR}_{US}) transformation at t_1 (with $t_0 < t_1$) that $o \in \text{Disabled-Caterpillar}$ and $o \in \text{Butterfly}$. Thus, status allows us to add additional constraints to C_t and C_s .

Status property S and additional constraints. Given the formal semantics of status classes [5], we introduce the property S, analogous to OntoClean’s meta-properties in the previous section and, moreover, to emphasize the core notions of the four status classes.

Definition 6 (+S). A property ϕ has status active at time t iff $\phi(x)$ holds at time t .

Definition 7 (-S). If a property ϕ has status scheduled at time t then $\phi(x)$ holds at some time t_0 , for $t_0 > t$.

Definition 8 (\sim S). If a property ϕ has status suspended at time t then $\phi(x)$ holds at some time t_0 , with $t_0 < t$.

Definition 9 (\neg S). A property ϕ has status disabled at time t iff ϕ holds at some time t_0 , with $t_0 < t$, and for all t' , such that $t' \geq t$, $\phi(x)$ does not hold.

This enables us to specify the following constraints for the *transformation_of* relation and its relata, when the instance *cannot* transform back:

(CT14) C_s has +S at the time of transformation and \neg S after transformation

(CT15) C_t has -S at the time of transformation and +S after transformation

If the entity *can* transform back, then CT14 and CT15 have to be replaced with:

(CT14') C_s has +S at the time of transformation and either \neg S or \sim S after transformation

(CT15') C_t has either -S or \sim S at the time of transformation and +S after transformation

Thus, theoretical contributions in formal conceptual modelling for temporal databases have the ontological commitment for CT14' and CT15', whereas

transformation_of was ignorant about this aspect. One could argue S is ontologically awkward for it hints toward intentionality, but representing transformations does consider that anyway, and therefore it is preferable to make this explicit. In addition, bio-ontologies are, or should be, represented in some formal ontology language; using S then provides a useful mechanism to represent precisely and explicitly such implicit assumptions about the subject domain, thereby pushing for closer analysis by ontology developers as well as ontological investigation into intentional aspects in a temporal setting. In addition, it offers a way to deal with some underlying ideas about temporal aspects yet representing it in a commonly used a-temporal ontology language such as OWL.

CT14 and CT15 versus CT14' and CT15' leads to two distinct sets of constraints on the position of classes in a taxonomy. Let $\forall x(\phi(x) \rightarrow \psi(x))$, henceforth abbreviated as $\phi \rightarrow \psi$, a class' property S indicated with superscript ϕ^+ , ϕ^\sim and so forth, and 'in the past' denoted with a "◊", then we can adjust Artale et al's constrains for CT14' & CT15' as shown in (1-5).

$$\phi^+ \rightarrow \psi^+ \tag{1}$$

$$\phi^\sim \rightarrow \psi^\sim \vee \psi^+ \tag{2}$$

$$\phi^- \rightarrow \psi^- \vee \psi^\sim \vee \psi^+ \tag{3}$$

$$\phi^- \rightarrow \neg\psi^- \tag{4}$$

$$\psi^- \wedge \diamond\phi^+ \rightarrow \phi^- \tag{5}$$

When transformation back is not permitted, then (1), (4), (5), and (6) hold; that is, *suspended* is not permitted, such that (2) is not applicable and ψ^\sim is removed from (3) and therefore replaced by (6).

$$\phi^- \rightarrow \psi^- \vee \psi^+ \tag{6}$$

Combining this with the meta-properties of phased sortals, then the subsumption constraints (1) & (4), CT6, CT14 & CT15 imply C_p^+ , because always one of the phased sortals subsumed by C_p is active³. Regarding permitting suspension, $\sim S$, then C_p^+ is also implied, because of (1), (2), (4), CT14' & CT15'.

4 Typical biomedical examples re-examined

In this section we assess if and how the proposed constraints suffice for modelling transformations in biology and biomedicine, using monocytes/macrophages and pathological transformations as examples.

Monocyte/macrophage: alternating states or phased sortals? Multiple cell types in the human body, most notably those involved in hematopoiesis,

³ C_p^- due to C^- and C_s^- is not possible even if we attempt to represent knowledge about extinct species: by virtue of being extinct, there is nothing to transform; note that corpses transforming into fossils involves other types that are not disabled.

can be pluri-potent stem cells, progenitor cells that differentiate, mature, and change in other ways, i.e., they seem to transform and are candidates for being categorised as phased sortals. However, looking at the details of the cells and processes, they are neither necessarily transformations nor necessarily phased sortals. For instance, progenitor cells undergo *cell division and differentiation* up to their final types, which is a non-deterministic process up to the point that the cells cease to be progenitor cells. But what about those final types? Here, we analyse the deterministic transformation from monocyte to macrophage, which, curiously, is also called in one compound term “monocyte/macrophage” or “monocyte-macrophage” in recent scientific literature (e.g. [17]) as if we have a cellular version of the Morning/Evening Star. We summarise the facts about monocytes and macrophages first and subsequently discuss the ontology development considerations.

In the partonomic tree of the Foundational Model of Anatomy [3], Monocyte is part of Blood partOf Hematopoietic system partOf Hemolymphoid system, whereas Macrophage is not part of anything—in reality, the cells dock into tissues to transform into macrophages, and are considered part of the ImmuneSystem [18]. In the FMA taxonomy, Monocyte is subsumed by NongranularLeukocyte that is subsumed by Leukocyte, whereas Macrophage is directly subsumed by Leukocyte [3]. Monocytes have as progenitor monoblasts and promonocytes, are considered to be the end stage of the differentiation, and they can change into different types of macrophages, such as microglia (macrophages located in the brain), so that CT14 & CT15 hold. A particular monocyte transforms into a macrophage only when it leaves the blood stream into tissue, changing from “APC, circulating” to “APC, tissue resident”, respectively [19] (APC = Antigen Presenting Cell); thus, macrophages are alternative functional states of monocytes, and they are both considered ‘adult’ cells, where [20] add that “*monocytes represent the circulating macrophage population and should be considered fully functional for their location, changing phenotype* in response to factors encountered in specific tissue after migration” (emphasis added). In fact, many research articles use the label “monocyte/macrophage” and thereby do admit to *one* universal. A phenotypic difference is that macrophages also stimulates T helper cells [19] whereas both perform phagocytose of foreign organisms. The processes of change involves, among others, physiological response (chemical & DNA) on mechanical pressure [17], levels of CCR5 that is the main co-receptor for M-tropic (macrophage tropic phenotypes) viruses [21], thus induced by an ‘external source’, and response on the chemical platelet-derived α -chemokine platelet factor 4 (PF4) [22]. In ontology development practice, three possible modelling options have been proposed to represent the transformation of monocyte to macrophage: as alternate states, as phased sortals, or by means of some other relationship. We assess each option on how it best captures the observed characteristics of monocytes and macrophages.

Option 1: Monocyte and macrophage are phased sortals subsumed by C_p . The problem with this approach is that there is no known suitable C_p that subsumes them both, i.e., violating CT6, CT7, CT8. If they were to be subsumed

by Phagocyte (a type of cell), then neutrophil (also a phagocyte) has to be excluded, for which we need at least one new subtype for Phagocyte. However, this new concept, as far as known, does not have a corresponding universal in reality. There are multiple similar cases with hemal cells, which, if decided upon nevertheless, require additions of many more *concepts* that do not have a (known) corresponding universal in reality.

Option 2: Monocyte and macrophage are different states that instances of C_p can have, because of the alternative functional states in the adult stage where each one is best adapted to its own environment: one residing in a fluid with the optimal shape of a sphere, the other in tissue with docking extensions on the cell surface, whilst remaining the same individual cell (cf. multi-cellular life cycles) instantiating one universal throughout the changes. Classifying the two states as distinct classes based on these differences is pursued by the Physiome project [19]. Both are APCs and phagocytes, which are candidates for the $=_i$. Ascertaining $\pi_1 \dots \pi_n$, they could be location and shape (attributions -I~R) or the property of stimulation of T helper cells for macrophages, or their differences in activated genes. This approach satisfies CT1, CT3, CT5, CT9, CT10, CT11, CT12, CT14 & CT15.

Option 3: Take Monocyte and Macrophage as universals, place them somewhere in the taxonomy of cell types and create a new relationship type R to relate them. This does not guarantee diachronic identity of the instances and is therefore inappropriate.

Clearly, with the current available information about the mechanisms of transformation from monocyte to macrophage, option 2 offers the more accurate representation.

Pathological transformations. Smith et al [4] include pathological transformations, which complicates both the permissible relata and the *transformation_of* relation itself for two reasons. First, it is ambiguous if the pathological entity may transform back to its healthy form and if this should be under the assumption that nature takes its course or if it also permits medical intervention. Second, it is not the case that for all instances that are transformed into a pathological entity they either all can or all cannot transform back; e.g., 5 out of 100 lung cancer patients survive and keep their lungs. It is true that for all carcinomatous lungs there must have been healthy lungs, but the inverse does not hold for all instances. Moreover, given that there is no common subsumer C_p for the healthy and carcinomatous lungs, the relata cannot be phased sortals. This leads to the combination of constraints where CT9-CT12 hold together with either CT14 & CT15 for healthy & pathological entities of *non-curable* diseases or CT14' & CT15' for *curable* diseases.

The reversible transformation from erythrocytes (red blood cells) to unhealthy echinocytes concerns shape change of the cell membrane morphing into other energy state optimums influenced by both biochemistry (including toxins, pH) and mechanics [23]. This is similar to monocyte and macrophages, except that the transformation is reversible.

5 Discussion

Regarding the *transformation_of* relation, the basic questions to answer for representing the universals in the subject domain in a domain ontology are: what transforms, how does it transform, and why does it transform? We focussed on the first two questions. Definition 1 in conjunction with constraints CT1-CT15' facilitate the analysis of transforming entities as well as the transformations themselves. For requirements on the participating endurants, the constraints for phased sortals to be subsumed by a type may be too restrictive for representing transforming biological universals most accurately for the main reasons that often either no such C_p is known to exist in reality that readily fits this requirement or it does exist but the criterion that supplies identity is difficult to establish; put differently, states and quasi-types seem to be more adequate for the participating continuants. Nevertheless, phased sortals remain useful kinds for representing stages in organism's life cycle as well as transformations of progenitor cells to differentiated cells.

The meta-property status S helps in understanding and representing more precisely *how* the instances transform, thanks to the constraints it imposes both on status change during transformation and the universals' position in the taxonomy. For the biomedical domain, this is particularly useful for transformations involving pathological entities.

Based on the few examples analysed, one cannot generalise this up to pre-defined sets of combinations of constraints that hold universally for the *transformation_of* relation and its relata, although several combinations and exclusions of constraints can be given:

- i. Phased sortals, unidirectional transformation: CT1-CT8, CT13, CT14, CT15;
- ii. States (including quasi-types), unidirectional transformation: CT1-CT9, CT11-CT15;
- iii. States (including quasi-types), transformation back is possible: CT1-CT13, CT14', CT15';
- iv. Pathological transformations, terminal disease: see constraints point ii, permit status change from -S directly into \neg S;
- v. Pathological transformations, reversal possible: see constraints point iii, permit status change from -S directly into \neg S.

Regarding pathological transformations and status change in points iv. and v., this can be easily added computationally with a database trigger or included in the default setting for a new term (representation of a universal or class), or be by default omitted to keep backward compatibility with other ontologies of diseases and pathological entities. Concerning point v, one can make a further distinction to define two semantically distinct sub-types of the transformation relation, which is to distinguish between self-healing transformations and those that require medical intervention, i.e., distinguishing between *natural* transformations and *human-mediated* transformations.

To actually use the different constraints and options for representing the *transformation_of* relation, be it as guideline or hard constraints, it will be helpful to make the distinctions explicit also in the ontology development soft-

ware by bringing representation choices of the meta-properties to the foreground, which could be added to OntoClean.

6 Conclusions

The current *transformation_of* relation as defined in the Relation Ontology does address an important aspect of relations among biological entities, but it is too flexible to enable representing the finer-grained distinctions between *what* transforms and *how*. Using the some core ideas of the OntoClean approach together with the notion of a status property that was generalised from temporal conceptual data modeling, we specified 17 constraints that constrain the *transformation_of* relation and its relata, which concern both the permissible categories of the relata and directionality of the relation. These were grouped into several combinations to facilitate more precise representations of transformations in biology and biomedicine. The proposed constraints were assessed with several typical examples from the biology and biomedicine domain. The here presented results enable developers of biomedical ontologies to better represent and relate entities such as monocyte & macrophage, caterpillar & butterfly, and their analogues, and thereby improve their ontology. Currently, we are investigating implications of the interactions between OntoClean's property kinds, the status meta-property, and temporal constraints in \mathcal{DLR}_{US} , and in future works we may consider more complex time modeling, such as with GFO [24].

Acknowledgments. The authors would like to thank Alessandro Artale and Barry Smith for comments on an earlier draft; responsibility for any errors that may be found in this paper lies with the author.

References

1. Antezana, E., Tsiporkova, E., Mironov, V., Kuiper, M.: A cell-cycle knowledge integration framework. In: Proc. of DILS'06. Volume 4075 of LNBI., Springer Verlag (2006) 19–39
2. Gene Ontology Consortium, .: The Gene Ontology GO database and informatics resource. *Nucleic Acids Research* **32**(1) (2004) D258–D261
3. Rosse, C., Mejino Jr, J.L.V.: A reference ontology for biomedical informatics: the foundational model of anatomy. *J. of Biomedical Informatics* **36**(6) (2003) 478–500
4. Smith, B., et al.: Relations in biomedical ontologies. *Genome Biol.* **6** (2005) R46
5. Artale, A., Parent, C., Spaccapietra, S.: Evolving objects in temporal information systems. *Annals of Mathematics and Artificial Intelligence* **50**(1-2) (2007) 5–38
6. Guarino, N., Welty, C.: A formal ontology of properties. In Dieng, R., ed.: Proc. of EKAW '00. LNCS, Springer Verlag (2000)
7. Guarino, N., Welty, C.: Identity, unity, and individuality: towards a formal toolkit for ontological analysis. In: Proc. of ECAI-2000, IOS Press, Amsterdam (2000)
8. Carrara, M., Giaretta, P.: Identity criteria and sortal concepts. In: FOIS '01: Proceedings of the international conference on Formal Ontology in Information Systems, New York, NY, USA, ACM (2001) 234–243

9. Noonan, H.: Identity. In Zalta, E.N., ed.: The Stanford Encyclopedia of Philosophy. Fall 2008 edn. (2008) <http://plato.stanford.edu/archives/fall2008/entries/identity/>.
10. Welty, C., Andersen, W.: Towards ontoclean 2.0: a framework for rigidity. *Journal of Applied Ontology* **1**(1) (2005) 107–11
11. Artale, A., Guarino, N., Keet, C.M.: Formalising temporal constraints on part-whole relations. In: Proc. of KR'08, AAAI Press (2008) Sydney, Australia, September 16-19, 2008.
12. Guarino, N., Welty, C.: An overview of OntoClean. In Stab, S., Studer, R., eds.: *Handbook on ontologies*. Springer Verlag (2004) 151–159
13. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: *Ontology library*. WonderWeb Deliverable D18 (ver. 1.0, 31-12-2003). (2003)
14. Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N.: Social roles and their descriptions. In: Ninth International Conference on the Principles of Knowledge Representation and Reasoning (KR'04). (2004) Whistler, Canada, June 2004.
15. Brochhausen, M.: The *Derives_from* relation in biomedical ontologies. *Studies in Health Technology and Informatics* **124** (2006) 769–774
16. Spaccapietra, S., Parent, C., Zimanyi, E.: Modeling time from a conceptual perspective. In: Proceedings of CIKM98. (1998)
17. Yang, J.H., Sakamoto, H., Xu, E.C., Lee, R.: Biomechanical regulation of human monocyte/macrophage molecular function. *Am. J. Path.* **156** (2000) 1797–1804
18. Hoffbrand, A.V., Pettit, J.E.: *Atlas of Clinical Hematology*. 3 edn. Elsevier (2000)
19. Hunter, P., Borg, T.: Integration from proteins to organs: The physiome project. *Nature* **4**(3) (2003) 237–243
20. Stvrtinova, V., Jakubovsk, J., Hulin, I.: Inflammation and fever. In: *Pathophysiology: Principles of Diseases*. Academic Electronic Press (1995) <http://nic.sav.sk/logos/books/scientific/Inffever.html>.
21. Tuttle, D.L., Harrison, J.K., Anders, C., Sleasman, J.W., Goodenow, M.M.: Expression of ccr5 increases during monocyte differentiation and directly mediates macrophage susceptibility to infection by human immunodeficiency virus type 1. *Journal of Virology* **72**(6) (1998) 4962–4969
22. Scheuerer, B., et al.: The cxc-chemokine platelet factor 4 promotes monocyte survival and induces monocyte differentiation into macrophages. *Blood* **95**(4) (2000) 1158–1166
23. Mukhopadhyay, R., Lim, G., Wortis, M.: Echinocyte shapes: Bending, stretching, and shear determine spicule shape and spacing. *Biophysical Journal* **84**(4) (2002) 1756–1772
24. Herre, H., Heller, B.: Ontology of time and situoids in medical conceptual modeling. In: Proc. of AIME 2005. Volume 3581 of LNAI., Springer-Verlag (2005) 266–275