Causality and the ontology of disease

Robert J. Rovetto [*] and Riichiro Mizoguchi
Research Center for Service Science, School of Knowledge Science, Japan Advanced Institute of Science and Technology, Nomi, Japan
E-mails: rrovetto@buffalo.edu, ontologos@yahoo.com; mizo@jaist.ac.jp

Abstract. The goal of this paper is two-fold: first, to emphasize causality in disease ontology and knowledge representation, presenting a general and cursory discussion of causality and causal chains; and second, to clarify and develop the River Flow Model of Diseases (RFM). The RFM is an ontological account of disease, representing the causal structure of pathology. It applies general knowledge of causality using the concept of causal chains. The river analogy of disease is explained, formal descriptions are offered, and the RFM disease definition is refined by describing causal chains in terms of causal relations. The definition is updated to coincide with the actual RFM classes found in its upper-level ontology, YAMATO, which brings to light some challenges in developing both YAMATO and the RFM. The RFM is also discussed in relation to another ontological account of disease. Strategies are offered toward interoperability between these theories.

Keywords: Causality, causation, causal chain, causal inference, ontology, formal ontology, applied ontology, biomedical ontology, biocuration, ontology of disease, disease causation, knowledge representation and reasoning, informatics

Accepted by: Mark Musen

> *"If a disease entity is fully understood, its description covers not only signs and symptoms in the course but also, particularly, the causal structure "lying behind" them: etiopathogenesis".*
>
> (Hucklenbroich, 2014)

1. Introduction

Disease is a central focus in biological and medical communities, a challenge for ontological analysis, and a concern of the utmost importance for all persons. Forming a general description, if not a definition, of *disease* that (i) accurately reflects biomedical science, (ii) is comprehensible, (iii) is computable, and (iv) has attained widespread consensus, is a controversial and challenging task. Exploring different accounts and perspectives on disease, or the notion thereof, may provide insights not shed by any single view. For example, the ontological accounts of disease discussed here are alternatives to one another, but may represent different aspects of the phenomena.

This is not to say that there should be a single definition, for it may do more harm than good (Hesslow, 1993). As a general term or concept, *disease* has a natural degree of vagueness that arguably provides a level of flexibility needed in biomedical research and practice. The absence of a univocal formal definition does not prevent clinicians from successfully doing their work and achieving the goal of curing and preventing illness. We can, in any case, agree that the concept of disease involves that of pathology, abnormality, dysfunctionality, deviation from health, and otherwise harmful conditions and processes in organisms.

---

[*]Corresponding author: Robert J. Rovetto, New York. E-mails: rrovetto@buffalo.edu, ontologos@yahoo.com. Alumnus: University at Buffalo, The State University of New York (2011); University of Maryland (2007), United States of America.

Within biomedical informatics disciplines, there is an apparent need to manage disease-related biological data, establish clear vocabularies therein, and form computational models that accurately reflect clinical and general biomedical knowledge. Ontological engineering may support clinicians and researchers of disease toward those goals. Commonly cited potential benefits include: reduction of ambiguity; data sharing; annotation; semantic interoperability; automated reasoning; and clinical decision support. There is also potential to advance the education and pedagogy of disease etiology and development. Consider disease ontology and disease data visualization applications, for instance. Philosophical ontology and analysis contribute by providing concept analysis and development, logical rigor and general insights into understanding the pathological phenomena (or our conceptualizations of them) generically referred to as 'disease'.

The River Flow Model of Diseases (RFM for short), originally introduced in Mizoguchi et al. (2011), is an ontological theory of disease. The RFM has been developed using the HOZO (2013) ontology editor. RFM classes are part of YAMATO: Yet Another More Advanced Top-level Ontology (Mizoguchi, 2010), which is part of the Japan Medical Ontology Development Project for Advanced Clinical Information Systems (Imai et al., 2009; Clinical Medical Ontology, 2013). Both ontological models are under development and no claims to completeness are made. The RFM is a causal account of disease for use in biomedical ontology and clinical settings, wherein medical practitioners, informaticians, data modelers, and ontologists organize, retrieve, and reason over the growing (Mihăilă et al., 2013) sea of biomedical data. Kozaki et al. (2015) describe an application of the RFM, an online graphical user interface (Disease Ontology, 2013; Disease Chain LD, 2014) displaying the causal structure of various diseases in terms of causal chains. The RFM aims to provide a general account that is consistent with clinicians explicit and intuitive beliefs on causality by employing the concept of *causal chains*, and *causal structure* classes.

Mizoguchi et al. (2011) define *disease* as *a dependent entity constituted of causal chains of clinical disorders*. We refine this definition in such a way that disease causal chains signify causally-related pathological states constituting the disease. We also provide an RFM disease definition that is more in line with the existing implementation: sequences of causally-linked processes form the causal structures of disease, and this expresses the idea of disease causal chains. The RFM therefore emphasizes the *causal aspect* of disease entities, an aspect that is not explicitly addressed or formalized in other existing ontological accounts of disease.

That *causality*, causal reasoning, and causal notions have central roles in biology, medicine (Rothman & Greenland, 1998), biomedical informatics (Kleinberg & Hripcsak, 2011), and in the philosophy of these scientific disciplines (Marcum, 2011; Dominic, 2009) is apparent. Medicine has "had a long history of trying to establish whether there is a causal link between a pathogen and disease [. . .], between the environment and disease [. . .], or between an adverse event and a drug" (Kleinberg & Hripcsak, 2011, p. 4). Clinicians seek vehemently to find the cures, causes and origins – the etiology and pathogenesis – of diseases. Definitions of specific diseases are themselves formulated in terms of what changes they *cause* in a patient. *Causal inference*, which involves the identification and discovery of causal relationships, is a central concern for epidemiology (Susser, 1991; Olsen, 2003). *Causal relationships* are sought after, to explain, control, and predict natural phenomena (Kim, 2004, p. 207). In addition to *prediction* and *explanation* (Kleinberg & Hripcsak, 2011, p. 3), a sought-after benefit of automated methods for causal inference in informatics is to assist researchers in the development of new hypotheses (Rothman & Greenland, 1998, p. 15). In short, *"one of the core concerns of all branches of medicine is causality"* (Kleinberg & Hripcsak, 2011, p. 1, italics added) and the identification of causal relationships between the relevant sorts of entities.

Given this, and given that causal thought and language are fundamental, widespread, practical, and explanatory, employing causal notions or categories in an ontological account of disease is a step in the right direction toward comprehensible, useful and scientifically accurate descriptions and formal representations. This approach provides an intuitive and accurate way of conceptualizing and describing disease entities. It reflects and appeals to our knowledge and experience of causality in the world; and the fundamentality and centrality of causal reasoning (Lagnado, 2011). "Causality is something we must reason with constantly in life" (Kleinberg & Hripcsak, 2011, p. 3).

This paper stresses the importance of representing causality in ontological models of disease. In so doing, we aim to contribute to the discussion of causality in ontology development. Within this context, we clarify and develop the RFM. An updated *disease* definition is presented so that it is more consistent with RFM classes as presently found in its upper ontology, YAMATO (classes which are not mentioned in the original RFM definition). This brings to light some areas for further development in YAMATO and the RFM. Addressing these points in detail, however, is beyond the scope of this paper. For example, although causal feedback mechanisms are essential parts of any causal description of pathology, they are not discussed here (rather, a general focus is unidirectional causal occurrences).

The paper is divided as follows. Section 2 introduces top-level categories adopted by the RFM/ YAMATO. Section 3 presents a philosophical discussion of causality and causal chains independent of any particular theory, but focusing on disease causation. Section 4 explains the RFM river analogy. Section 5 unpacks the RFM disease definition. A revised definition is presented, whereby a disease is described as interrelated causal chains of pathological entities in an organism. We also demonstrate similarities between the RFM and another ontological theory of disease, discussing strategies for interoperability between the two. Section 6 offers two generic (theory-independent) definitions of disease. Finally, concluding remarks are in Section 7.

The word 'entity' is used to encompass all that can or does exist at least at the particular level: objects (and their parts), processes, properties, qualities, attributes, features, functions, systems, states, etc. 'Biological entity' and 'pathological entity' are similarly used, but constrained to the domains of biology and pathology. 'Pathological', like 'abnormal' is undefined, but associated with (potential) harm, dysfunction, poor health, etc. Pathological entities, then, include: disorders; pathological states, processes, features; etc. Double quotes are used for quotations from references, single quotes for individual words or phrases. Key phrases may be either italicized or in bold. Capitalized, italicized and bold relational terms mark (existing or candidate) ontological predicates (relations). Bold, camel-cased category terms mark (existing or candidate) ontological categories.

2. Background concepts

Some widely used philosophical categories and distinctions, including the universal–particular distinction (Simon et al., 2006) are adopted. Similarity and generality have historically been distinguished from particularity, the former being described using the notion of universals (categories, kinds, types, classes), the latter with particulars (individuals, tokens, instances). Heart disease, as a general, universal or repeatable entity, is distinguished from a specific case of heart disease affecting an individual human being.

The domain-neutral ontological categories, **Endurant** (or **Continuant**) and **Perdurant** (or **Occurrent**), originate from philosophical inquiry, and have been applied as formal classes in ontologies. In accordance with both ordinary philosophical understanding (Loux, 2006), and some contemporary foun-

dational ontologies such as the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE, 2013) and the Basic Formal Ontology (BFO, 2013), endurants/continuants are object-like entities that persist self-identically through time while undergoing change, are the bearers of properties (**Dependent Continuants/Entities**), and are said to *Participate In* perdurants/occurrents. Perdurants are processual entities that occur or unfold through time. Paradigm examples of each category include table-top material objects, and processes and events, respectively. Formal ontological relations such as ***Participation*** (Grenon & Smith, 2004; BFO 2.0, 2013; Masolo et al., 2003) relate continuant and occurrent entities, thereby taking into account both ontological perspectives.[1] *Participation* in YAMATO relates an object to a process or event. More specifically, in *enacting* (Mizoguchi et al., 2011), a process, an object is participating in it.[2] Participation is used in the sense of DOLCE which is approximately equivalent to that of the Relation Ontology and BFO.

These ontological perspectives are considered insufficient by themselves to accurately represent the world. Neither is ontologically prior to the other. Rather, each is dependent on the other (Galton & Mizoguchi, 2009, p. 72). Object and process – or object-like and process-like aspects – co-exist (Mizoguchi, 2004) and are existentially dependent on one another. They are **mutually interdependent**.

Using distinctions from (Galton & Mizoguchi, 2009), a given object is described as being the "interface between internal and external processes". This description is insightful because it communicates our dynamic and changing world, and emphasizes the interconnectedness and mutual interdependence between what are called **Objects** (Endurants) and **Processes** (Perdurants). Biological entities, such as organisms, are excellent examples of this interdependence, change with stability, and unity (Lillie, 1940). It is an obvious biological reality that organisms can behave in various ways, are affected by processes external to them, and have a variety of processes constantly taking place within their bodies (many of which sustain their existence).

The next section is a non-exhaustive[3] introduction to the philosophy of causality for foundational comprehension. It provides a philosophical and theoretical basis for (candidate) causal categories in ontological theories, but is not necessarily specific to the RFM/YAMATO. RFM/YAMATO treatments are, however, mentioned where appropriate.

3. General causal notions

The literature on causation is extensive[4] and spans various disciplines from philosophy to biology and medicine. A number of general theories have been offered but with little agreement (Losee, 2011; Sosa & Tooley, 2004). In epidemiology, for example, 'cause' is defined in a number of ways, each of which corresponds to a philosophical approach to causation (Weed & Parascandola, 2001, p. 906). Drawing largely from philosophical inquiry, this section provides a philosophical context for causality in the ontology of disease, the formal representations thereof, and introduces related concepts.

Causality/Causation is generally understood as the idea or the mind-independent metaphysical principle of cause-and-effect. It reflects the fact that phenomena in the world have causes, and that what are

---

[1]These categories reflect the traditional metaphysical perspectives of Endurantism and Perdurantism (Lowe, 2002, p. 49), philosophical theories on persistence that classify entities in the world as essentially either kinds of objects or kinds of processes.

[2]Borgo and Mizoguchi (2014) present a first-order formalization of ***Enactment*** in YAMATO.

[3]Counterfactual dependence, for example, is often associated with causation, but not discussed here.

[4] See Schaffer (2008) and Causation Annotated Bibliography (2008) for a sample.

called 'cause' and 'effect' are related in some non-arbitrary, non-random, and lawful manner. As such, causality involves recurrence and regularity – not mere accidental regularities, but those of laws of nature (Swinburne, 1997, p. 80). Causality mirrors the foundational idea that the world is comprehensible and structured. It is related to **explanation**, **reasons**, **necessity** (Mellor, 1995, p. 3), and **prediction**, the latter of which is notably significant in medicine.

The causal relation. "As ordinarily understood, the causal relation is a binary relation relating causes to their effect" (Kim, 1973), but given the causal complexity of any event, the arity may very well be greater. We therefore do not restrict the relation to two arguments. Described as a relation, it obtains between at least two entities: cause, and effect (the causal relata). 'Causation relation' and 'causal relation' are used synonymously. In philosophy, formal ontology, and knowledge representation, an n-ary predicate represents a general causation relation, and can take the following form (Table 1), where the domain and range of the causal predicate represent causal relata.

A causation relation is typically understood as having the following formal properties (Table 2), both of which are consistent with, but do not exhaust, causal scenarios at biological scales. Both are translatable into a logical formalism, such as first-order predicate logic. Here, they are implicitly universally generalized ($\forall$).

In what follows '$\rightarrow$' signifies implication, '$\neg$' negation, '$\wedge$' conjunction, '$\exists$' the existential quantifier (read: there exists), '$<$' temporal precedence is undefined here but can be interpreted along the lines of Allen's Interval Calculus (Allen, 1983), and 'A#' enumerates causal axioms.

If one desires an irreflexive causal relation, communicating the idea that the relata do not cause themselves, then we have (A3): $\neg causes(c, c)$.

Philosophical theories posit general or type-level causation holding between domain-neutral categories, such as Event, Process, Property, State and Fact. They also posit instance- or token-level causation between instances of these categories. However, there is no philosophical agreement on "how to relate the type and token levels" (Kleinberg & Hripcsak, 2011, p. 3). Given that YAMATO classes quantify over particulars (Borgo & Mizoguchi, 2014), at least a token-level causal account is to be formulated for the RFM.

Table 1

Causation predicate, its inverse, and candidate domain and range

Causation relation w/relata	Inverse relation	Domain and range
c causes e	*e is_caused_by c*	Events, processes, objects, properties, qualities,
causes(c, e)	*is_caused_by(e, c)*	states, facts, . . .
c causes e at t	*e is_caused_by c at t*	
causes(c, e, t)	*is_caused_by(e, c, t)*	

Notes: Italicized terms are predicates. '*c*' and '*e*' mark the domain and range. '*t*' marks temporal instants or intervals, indicating a temporally-indexed version of the predicate.

Table 2

Common formal properties of a causation relation

Formal property	Natural language description	Logical notation
Causal asymmetry (A1)	Effects do not cause their causes. If *c* causes *e*, then it is not the case that *e* causes *c*.	$causes(c, e) \rightarrow \neg causes(e, c)$
Temporal asymmetry (A2)	Effects do not temporally precede their causes. If *c* causes *e*, then it is not the case that *e* temporally precedes *c*.	$causes(c, e) \rightarrow \neg causes(e < c)$

Generally speaking, one can formulate distinct theories in which domain-specific entities of different ontological categories are candidate **causal relata** for causal relations holding at varying scales. One (or more) disorder; pathogen; foreign body; pathological process, state, feature; or dysfunctioning organism system (most being sorts of Pathological Entities) may be the cause of some other pathological entity, sign or symptom.

The question of the type of causal relata remains open in the causation literature. However, relata *can* be restricted based on the situation and needs. For computational applications they very likely *should*: constraints on the causal relata may be necessary to avoid inconsistent or haphazard use of putative causation predicates (some of which may not actually be causal). Relata can be restricted by appealing to causal intuitions, or by focusing on causation in a specific domain or context. Causal expressions and descriptions in biomedicine, for instance, provide cues to restrict the relata to certain sorts of entity. At a much more general level, the preponderance of causal talk in which events take the role of causal relata make it natural to view **Events** and/or **Processes** as causes/effects.

If the domain and range of a causal predicate is, indeed, constrained to **Perdurants/Occurrents**, then without additional qualification (and in the context of biology), Biomedical Occurrents (broadly construed) are the causal relata of a causal relation. This is the case for the RFM's current treatment of causation: it holds between occurrents.

Change and production.   Causes are said to "produce", "result in" or "bring about" effects, each of which has a similar meaning. Causes produce their effects. Effects, however, do not necessarily produce their causes.

Change is implied in production and causation. That something caused something else involves one or more changes, but exactly what kind of change varies. *Prevention*, *reversal* and *manipulation* (Whitbeck, 1977) of causal conditions are related to etiopathogenesis in that they involve critical, often life-saving, changes. Preventing or reversing, for instance, involves artificial changes to intervene with the conditions for, or the existence of, some illness. These changes often include a biological or otherwise physical mechanism or process. As such, a Process-theory of causation may be applied to an ontological causal account of disease.

Connection and dependence.   There is a connection between cause and effect such that (a) they do not merely occur in regular succession or temporal sequence; the connection is neither mere association nor mere correlation, and (b) the effect is dependent on the cause in some manner. Distinguishing causation from correlation has historically been an important topic for epidemiology. In what follows, 'causally linked' and 'causally (inter)connected' are synonymous: if entities are ***Causally-linked***, then they stand in causal relations to one another.

Necessitation and determination.   Causality is often associated with determination and necessity, typically *physical necessity*. The latter is in turn related to laws of nature. The cause is described as necessitating and/or determining the effect. However, this is not to say that given some cause the effect is non-stoppable (von Watcher, 2010), and it is also not to deny the existence of free will. In terms of biomedical reality this expresses the following uncontroversial fact. Without (i) intervention by clinicians (or others), and given (ii) certain conditions and laws, specific occurrences will happen. Determination or necessity, so involved in causal scenarios, may be a kind of *lawful production* (Bunge, 1979, p. 19), or perhaps as a type of determination.

Furthermore, biomedical investigation into disease informs us that there exist various **contributing causal factors**. Both prevention of and interference with disease are possible by manipulating one or more causal factors. This means that deterministic necessary and sufficient condition (in the logical

sense) accounts of disease are too strict and inflexible. Multiple and possibly interacting conditions can be: necessary, sufficient, but also neither (Weed & Parascandola, 2001, p. 907) for the occurrence of the disease in question. An ontological account of disease should take this into account.

The foregoing causal considerations, among others, should be taken into account when representing disease causation and ontology. A generic bio-causal account should also address the **teleological** aspect involved with causality and biota. Finally, axioms (A1) through (A3) are applicable in formally representing disease causation. The next subsection presents a brief and general interlude on the category of Pattern (here undefined) because it is commonly associated with that of Causality. Note that the disease accounts discussed below do not presently include the category of Pattern. This discussion merely presents an important general concept related to that of causality, and should therefore inform general ontological accounts of illness.

3.1. On pattern, causal-chains and disease

Given that causation involves recurrence and regularity, it is reasonable to believe that the category of **Pattern** is applicable. Both causal chains and patterns require parts (broadly construed) that stand in certain relationships to one another. Patterns and their instances take various forms. Clinicians and researchers, for example, observe recurrent patterns among disease entities, including their effects, signs and symptoms. Some can be characterized as *patterns of*: bio-causal-chains, bio-causal-processes, or biological processes that are causally-linked.

Biological activity and the stability of entire living systems have been described as being dependent upon the interdependent and interconnected character of separate processes and events (Lillie, 1940, p. 331). That is, the repeatability of *"complex chains and networks of causation . . . the regular features of causation are thus responsible for the stable side of organic life"* (Lillie, 1940, p. 318, italics added). Patterns in nature exemplify this stability in life. And given that causation is typically held to relate occurrences in the world, *processual* entities – biological and abiotic – are major factors in the stability of life (a point we will see expressed in the RFM).

Various, if only partial, pattern-based descriptions of disease types are conceivable: disease as ***exhibiting***, being ***constituted by/of***, or at the class-level being a ***subtype of*** one or more patterns of pathological changes. A **pattern of change**, which is a pattern of occurrences or processual entities, is more generally a unity of changes that stand in certain spatio-temporal–causal relations to one another, the whole of which – the pattern – is exhibited time and again by distinct particulars. Disease, as a pattern of change(s), involves recurrent causally-related pathological processes (and their continuant participants). The basic idea is that bio-causal relations unify (in some sense) disease entities into a whole, e.g., a pattern, one or more interconnected biophysical causal chains, etc. Hucklenbroich (2014) describes a disease entity (type or pattern) as "a pattern of pathological behavior of the organism" and asserts: "disease entities are always *patterns*, i.e., *sets* of *different* courses". We will briefly return to this idea in Section 6.

> *"Disease entities are discovered (as a rule) in clinical practice, by observing or recognizing patterns or typical traits in the disease course of individual cases [. . .] by virtue of their pathological manifestations"*.
>
> (Hucklenbroich, 2014)[5]

---

[5]Hucklenbroich considers *disease entity* to be a type (universal), but not necessarily a disposition.

We should note, however, that one can question whether there is mind-independent unification be-tween all the parts/elements of a pattern (and causal chain) that would justify the characterization as *a whole*. One can also question whether disease is a natural kind or universal. The final subsection of this excursion into causality presents a philosophical analysis of the general idea of causal chains.

3.2. Causal chains

The ordinary idea of a causal chain (C-C for short) and its application in discourse conveys a connected and ordered sequence of entities, and is often metaphorical. C-C's are typically conceived of as some occurrence causing another, which then causes another, and so on over time. Visual representations of causal chains often take the form of unidirectional arrows connecting alphanumeric characters or shapes. The latter, also called 'nodes', represent causal relata, whereas arrows represent causal relations. A C-C is not a mere sequence or succession of occurrences just as the relationship between cause and effect is neither mere correlation nor regular constant conjunction. Rather, the connection is a stronger relationship between that which the links in the chain represent (the causal relata). In a sense, C-Cs are *relational entities*: without causal relations among relata a causal chain would cease to be what it I; their existence requires certain relational arrangements obtain.

The ontology of causal chains includes: (1) the C-C as a whole; (2) the links (relata) in the chain as the parts of the whole (in some sense of parthood); and (3) the causal relationship(s) between relata. The latter may be a generic causation relation or more specific causal relations.

Both C-Cs and sequences involve a following of one entity after another; they are ordered and have direction. The **directionality** of causality, and therefore of C-Cs, in the world is implicitly understood as the direction of time; that is, the asymmetry of time entails a direction. In epidemiology, direction is considered an essential causal property that indicates "consequential change" of and **from** cause to effect (Susser, 1991, p. 639).

By contrast, a literal chain taken end for end does not have a preferred direction. It exhibits direc-tional symmetry such that opposing spatial paths can be traced along it. This is due to the compositional structure of the chain: being a material (and thus spatially-extended) object. A C-C is representative of a spatio-temporal path from the beginning of the causal sequence to the end. Many causal chains, then, are lawfully connected, ordered and directed sequences of entities. For example, **disease causal chains** (DCC) are directed toward pathology and can be interconnected. Although intuitions on causal chains would have them be almost exclusively sequential, representing **simultaneous causation** (Huemer & Kovitz, 2003) is not uncommon. Furthermore, biomedicine requires the understanding of **causal feed-back mechanisms** and **cyclical** causation. We restrict the focus here to unidirectional and simultaneous causation.

Each link (node) in a C-C may be viewed as the cause/effect for the subsequent/preceding link.[6] The adding of links signifies: that some temporally preceding, overlapping, or simultaneous link (relatum) has changed or has caused the newly added link. A change has occurred in the world in virtue of the causal interaction, or perhaps the *causal powers* of the relata (plus any other relevant conditions).

These links need not be interpreted as representing discrete (rather than continuous) causal relata. Just as the links in an actual chain overlap one another through the holes they form, so can the relata overlap. A causal scenario may involve continuous, simultaneously occurring, causally-connected or overlapping processes with vague (if any) spatio-temporal boundaries. This is often the case in biology where crisp

---

[6]As such, Cause and Effect categories may be modeled as **Roles** played by causal relata.

spatio-temporal boundaries are difficult to discern, and where complex causal interactions takes place over time.

Toward formalizing a generic Causal Chain class (C-C FOL), a first-order formalization is as follows. For the moment, assume relata are processual entities, ignore temporal arguments, and keep the parent class of Causal Chain open. Lower case letters signify individuals, bold terms classes, and italicized terms relations. 'instance_of' signifies the instantiation relation between classes and their instances. (C-C FOL) says that instances of causal chains have occurrent parts that cause one another in some manner, sequence, etc.

(C-C FOL) $instance\_of(c, \textbf{Causal Chain}) \rightarrow \exists c1 \ldots cn:$
   (i) $instance\_of(c1, \textbf{Occurrent}) \wedge \ldots instance\_of(cn, \textbf{Occurrent})$
   (ii) $causes(c1, c2) \wedge causes(c2, c3) \wedge \ldots cn \wedge$
   (iii) $has\_causal\_part(c, c1 \ldots cn).$

To the extent that the relata are parts of a C-C, the mereological relation of ***Parthood*** is applicable, but not without complication. If we assume parts of a whole must be of the same ontological category as the whole, then the following holds. An Occurrent C-C is a group of causally-linked occurrent parts. Additional identity criteria for this class may include **temporal contiguity** of the processes, whereby $c1$ *meets* or *overlaps* $c2$ (Allen, 1983).[7] This can be added as a fourth conjunct (iv) to C-C FOL. The important point is that the relata ($c1$ through cn) stand in causal relations, or otherwise involve causal processes. Note that the RFM considers the smallest causal chain to be a single entity (e.g. disorder, abnormal state, etc.). This limiting case can arguably be ontologically justified by considering the single entity to have some causal power or influence, but we will not explore this here.

Now, if we do not assume parts are the same type as the whole, then another relation is needed. If $c1 \ldots cn$ and **Causal Chain** are mutually disjoint, or have different identity criteria, then a species of the formal ontological relation of ***Constitution*** (or a new relation) can relate them. It allows entities of distinct ontological categories to coincide in some manner.

For example, assuming occurrent relata: a continuant C-C is ***Constituted by*** causally-connected Occurrents. This may clash with the intuition that C-Cs are processual, but it is how the RFM models DCCs. Although this is an area for further development and clarification, at present, the RFM asserts a constitution relation to form Dependent Continuant diseases from C-Cs of Occurrents. Formal properties (A1), (A2) and (A3) from Table 2 are also are adopted by the RFM for its causal relation. In what follows, a general causation relation (***causes*** and ***is_caused_by***) is primitive, with the above content acting as an elucidation.

Finally, the concept of disease C-Cs appears similar to that of biological pathways. **Biological pathways** are series of molecular interactions causing specific biological processes. The representation of the *possible* reactions of an organism is "grouped into causal chains to form pathways" (Joshi-Tope et al., 2005). Likewise, Kozaki et al. (2015) describe how their browsing interface visually portrays potential disease courses (reactions, effects, etc.).

A representation of the full causal complexity inherent in biology and medicine may be a directed web or network of pathways and nodes, but for medical practitioners selectively focusing on specific biological interactions will likely be the task of the day. This focus is not uncommon in causal attribution and causal determination in general. After all, "a 'complete' causal account of an effect would be

---

[7]Depending on the account, what actually overlaps may be causal processes, a temporal part, an interval, etc.

unmanageable or even impossible ever to achieve. The causal chain, or, to use a more holistic metaphor, the causal web, may extend indefinitely" (Raisanen et al., 2006, p. 294).

Like Causation, the general category of Causal Chain is broad, and applicable to most, if not all, domains and levels of granularity. A C-C class and one or more causation predicates (relation terms) can therefore take a place in the class and relation hierarchy of domain-neutral ontologies. **Causal Chain**, for instance, may be a subclass of **Causal Structure** (broadly construed). Specialized categories and relations, with domain-specific causal relata, can be included in domain ontologies as needed. Let us move on from this general presentation of causality to begin our formal discussion of the RFM by clarifying its central analogy.

4. The river analogy

A disease is like a river. So the analogy goes, but how are they similar? The most general similarity according to the River Flow Model of Diseases is that both entities are of the same ontological category: Continuant. The central characterizing feature of rivers, however, is distinctly Occurrent: flowing waters. Consider the Heraclitean (Graham, 2011) proverb that one cannot step twice into the same river. It communicates the reality of change in and of the world, and uses a conspicuously dynamic natural phenomenon to do so. A more precise interpretation of this proverb conveys the idea that a river (anything whatsoever) is constantly changing, or involves change of one sort or another while retaining its distinctiveness. A yet stronger interpretation is that a river maintains its identity for but an instant . . . or not at all.

Yet we often speak of individual rivers. We are able to visit and later refer to the same river – or the unfolding geographic processes we call 'rivers' – time and again. To the extent that the general term 'river' and the corresponding concept reflects a natural kind, and is indispensible in hydrology, fluvial geomorphology and related scientific disciplines, it reflects an ontological category for scientific domain ontologies.

Whether rivers are more accurately described as a process, a continuant, or otherwise, the phenomenon in question seemingly exhibits some constancy and stability such that we identify persisting conditions and properties. We believe that **this** river – say, the Tedori River in Ishikawa Prefecture, Japan – is the same particular river at a later moment. Another interpretation of the thought-provoking proverb is that a river is an example of interconnected and contrary (but not contradictory) elements – *constancy* and *change*, or *same* and *different* (Cohen, 2006) – reflecting dual aspects, or a unity of opposites, in the nature of the entity in question.

This last interpretation is consistent with the ideas adopted by the RFM, e.g., that an object (here, a river) is "a point of stability" between or in virtue of processes (Galton & Mizoguchi, 2009, p. 25). Both communicate the **mutual interdependence** of the processual and object-like aspects of existents. As it is, we implicitly conceive of things in the world as persisting while undergoing various sorts of change over time, and as therefore having a form of unity.

We identify the physical boundaries of rivers while being cognizant of the fact that those boundaries objectively change as flowing waters continuously shape and carve their way into the surrounding terrain. Rivers "exhibit tremendous variability in the quantity, timing, and temporal patterns of river flow, and this profoundly influences their physical, chemical, and biological condition" (Allan & Castillo, 2007, p. 31). This constant change in the measurable properties of depth, shape, temperature, content (sediment, nutrients, biota), and current velocity (Allan & Castillo, 2007; Hebert & Ontario, 2008) is observable over time. Hence, rivers are described as persisting in time through changes in these, and other,

properties. The RFM, therefore, ontologically categorizes rivers, like waterfalls (Galton & Mizoguchi, 2009), as **continuants**.

The *flowing* of a river is an example of a process: it occurs through time, and according to traditional philosophical accounts, is temporally extended. Drawing again from (Galton & Mizoguchi, 2009), the water constituting the river is flowing, but not the river (conceived as a whole). The whole is character-ized as undergoing **external processes** that occur at the same scale as the whole. Any process at a finer granularity is an **internal process** enacted by parts (portions of water) of the river (Mizoguchi et al., 2011). When rivers extend or branch into new geographic regions, they are *enacting* or *participating in* external branching processes. According to the RFM, the same applies to diseases.

Diseases, like rivers, are continuants in the RFM. A disease is held to persist as one and the same entity while affecting an organism over time. Personal physicians describe a patient's disease in continuant-like terms, tracking an instance of a disease over time while observing changing signs and symptoms.

There is also a distinct process-like character to diseases. They develop over time, harmfully affect body parts, functions and systems, all of which are amenable to study and measurement just like the features of rivers. These pathological changes are processual parts composing the course of the disease. Just as a river affects its surroundings, such as the surfaces of the host terrain it flows over and through, so a disease affects the host organism.

The external–internal process distinction applied to rivers is also applied to disease. The *spreading* of a disease is an external process enacted by the whole. Internal disease processes at different biological scales are those (pathological) sub-processes occurring to, or by, parts of the organism. Examples of internal processes relative to the entire disease include growing blood clots, decreasing motor control, and onset of blindness. A disease causal chain has such processes as parts.

Constraints on the properties rivers and diseases have are placed, in part, by the surrounding environ-ment (Allan & Castillo, 2007, p. 32) – the wider ecosystem and the organism, respectively. The direction and flow of water, for example, is determined by certain conditions and laws of nature. In an analogous way, diseases develop in a biologically principled manner with a myriad of potential causally efficacious conditions and effects.

The changes that occur during the course of each disease instance vary from patient to patient. That is, the development of a disease takes different symptomatic and processual paths. They can spread in a mul-titude of ways, affecting different areas of the body. These variations consist of different causal chains. Distinct patients may exhibit different physical manifestations (signs) and experience different symp-toms just as distinct rivers uniquely affect their surrounding environments. Rivers change size, shape, and exhibit branching topological properties. So the analogy goes, a disease itself branches (spreads) into different parts of the body.

A disease type may manifest in different ways – in different disease courses of causal chains – but the idea is that instances of that type will nevertheless exhibit *similarity*, namely certain **core** features, patterns and causal processes. In short, just as rivers exhibit patterns of various sorts (topological, pro-cessual, etc.) so do diseases. "Diseases follow a characteristic course" (Whitbeck, 1977, p. 625), while exhibiting variation in their instantiations. This is represented by the RFM Core Causal Chain class discussed below.

5. Formal ontological accounts of disease

Existing ontological theories have put forth distinct accounts of disease, the River Flow Model being one of them. In this section, its disease definition (Mizoguchi et al., 2011) is unpacked and developed

further using the preceding analysis. We first explain the original natural language (NL) definition of the RFM Disease class. Second, we discuss additional classes found in the RFM implementation, but not mentioned in the NL definition. We change some to more clearly reflect the concept of causal chains. We do not provide formal definitions for the C-C subclasses, however. The difference between the original definition as found in RFM publications and its current implementation is an important discrepancy to address because ideally they should cohere. The changes we present are a step toward greater consistency. Third, a revised definition is presented to better match the actual classes in the YAMATO hierarchy, of which RFM classes are a part. Fourth, the RFM is compared with the Ontology of General Medical Science's (OGMS, 2014) disease account. We offer potential strategies for their interoperability.

5.1. The River Flow Model disease definition

The RFM is neutral with respect to the cornucopia of theories of causation. It focuses on knowledge and intuitions of causality at the mesoscopic scale of daily occurrences, as understood in biology and at the relevant biological scales. If we accept that "[. . .] in medical contexts causality talk has the same character as everyday causality talk" (Johansson & Lynoe, 2008, p. 127), and that this mirrors (to some degree) the causal structure of the world, then we can apply causal notions to finer biological scales.

A causation relation (***causes***) is a primitive in the theory, but adopts axioms (A1)–(A3). The **is-caused-by** relation (found in the RFM implementation) is the inverse. We define RFM causal chains in terms of these primitive causal relations, while including others. More specialized causal relations may be formulated as needed. Bio-causal relations, for instance, can be asserted based on the biological scale at which the C-C in question occurs. On this note, consider that homeostasis at one scale may, itself, be due to causal chains at another (finer or coarser) scale.

We unpack the original natural language disease definition (RFM-NL) by discussing the relevant ontological categories as they appear therein.

Disease (RFM-NL) A dependent continuant constituted of one or more causal chains of clinical disorders appearing in a human body and initiated by at least one disorder (Mizoguchi et al., 2011, 2012).

The supposed ontological classes mentioned in this definition are: **Dependent Continuant**, **Clinical Disorder**, **Human Body**, **Causal Chain**; ***Constitutes***, and ***Initiates***, respectively. At face value, a first-order formalization of RFM-NL is:

(NL-FOL)   *instance_of*(*x*, **Disease**) = *def.*
             instance_of(*x*, *Dependent Continuant*) ∧
             instance_of(*y*, *Causal Chain*) ∧ *constituted_of*(*x*, *y*) ∧
             instance_of(*z*, *Disorder*) ∧
             instance_of(*a*, *Human Body*) ∧ *located_in*(*z*, *a*) ∧ *initiated_by*(*y*, *z*).

The actual formal implementation of the RFM, however, utilizes classes not mentioned in the NL definition (and vice versa). RFM-NL is at least acceptable as a shorthand, expressing the main idea of the theory, for the more complicated implementation.

As a subclass of Continuant, a **Dependent Continuant** (YAMATO: **Dependent Entity**)[8] retains its identity over time, and is existentially dependent on some other entity(s). Examples include, but are

---

[8]Equivalent to BFO: Dependent Continuant (version 1.1).

not limited to properties, qualities, functions and roles. Without a particular organism, an instance of a disease (a dependent continuant) would not exist. The formal ontological relation of ***Inherence*** is used in foundational ontologies such as DOLCE (Masolo et al., 2003) and BFO, and relates objects (property-bearers) to their features. Generalizing to any organism, a disease is dependent on an **Organism** and/or the organism parts (the bearers of the disease).

To say a disease is ***Constituted of*** causal chains of clinical disorders is to say that a C-C class groups the causally-linked entities into a unified whole. "The course of a disease is a unity of such entities that is characterized by abnormal functioning" (Schulz & Johansson, 2007), and related pathological processes causing that dysfunction. ***Constitution*** relates potentially compresent or coincident entities of disjoint categories. It specifically relates disease (continuant) to causal chains (occurrents) (Fig. 1). **Constitution** is understood in the sense of DOLCE as an irreflexive relation, but with the caveat that the domain and range may need adjusting to include different ontological classes. The RFM uses a constitution (or similar) relation in part because the continuant and occurrent classes spatio-temporally coincide. Referents of the latter helping to sustain that of the former.

'**Causal chain**' has been defined as: *"a chain of entities linked by causality"* (Mizoguchi et al., 2011, p. 64). A more recent definition is: *"a chain of entities linked by causal relation[s]"* (Mizoguchi et al., 2012). As stated, this allows for a plurality of entities to compose C-Cs. At the instance level, however, RFM C-Cs include only Occurrent relata, focused primarily on YAMATO Abnormal (Pathological) States. We assert axiom (A4a–b), which expresses (ii) and (iii) from (C-C FOL), and with temporal arguments omitted. These axioms apply for the is-caused-by and has-effect-on relations as well. **Changes in states** (or changes in values of qualities) are YAMATO: **Processes**. Entire processes, not their parts or internal processes, stand in the ***causes*** relation to one another.[9]

(A4a) $causes(c1, c2) \rightarrow instance\_of(c1, Occurrent) \land instance\_of(c2, Occurrent)$;

(A4b) $causes(c1, c2) \rightarrow instance\_of(c1, Abnormal/Pathological\ State) \land$
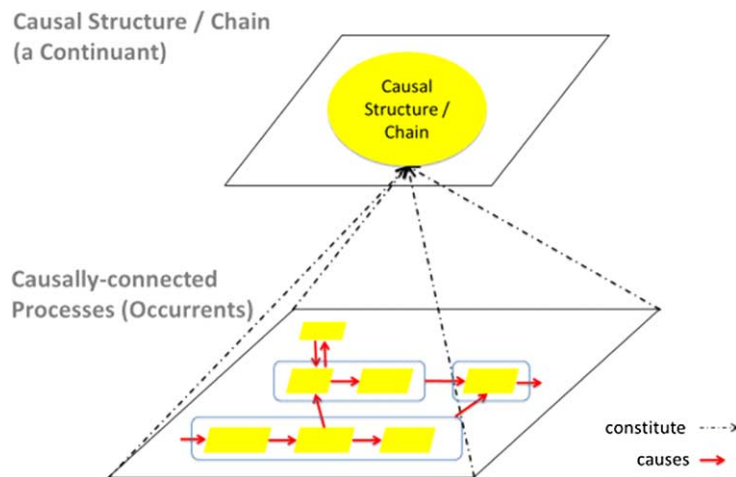$instance\_of(c2, Abnormal/Pathological\ State)$.

Fig. 1. The constitution of the RFM causal structure class and causally-linked occurrents. The rounded rectangles signify that, depending on the disease, each causal sequence of occurrents can be represented as a particular causal chain. (Colors are visible in the online version of the article; http://dx.doi.org/10.3233/AO-150147.)

---

[9]The parts of a whole process stand in the *causally-contributes* relation from Borgo and Mizoguchi (2014).
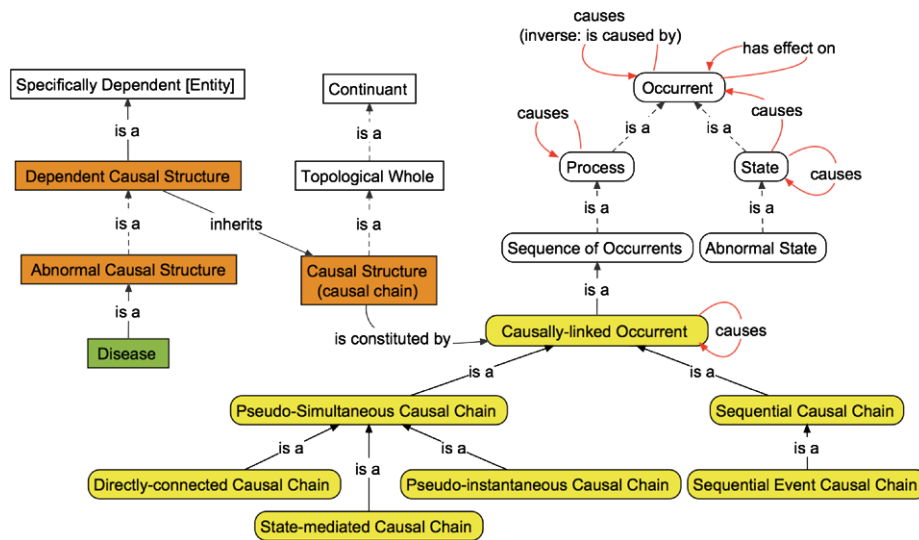
Fig. 2. RFM classes and relations in YAMATO. Solid and dashed *is a* arrows signify direct and indirect class subsumption, respectively. Red arrows indicate primitive *causes* and *has effect on* relations. Not all potential causal relations between occurrents are shown. Boxes are continuants, ovals occurrents. Orange marks continuant Causal Structure classes. Yellow marks the newly added/changed Causal Chain classes. (Colors are visible in the online version of the article; http://dx.doi.org/10.3233/ AO-150147.)

With the category of Occurrents as the most general relata restriction, a number of Occurrent subclasses (and domain-range combinations thereof) thereby can stand in causal relations to one another (as Fig. 2 partially portrays). The YAMATO: **Sequence of Occurrent** (SO) class (discussed below) when qualified by the causal aspect, is the essence of the RFM concept of causal chains of a disease. To capture this, we add the **Causally-linked Occurrent** (CLO) subclass to SO.

We now present the additional categories and relations that are part of the RFM implementation. Although Causal Chain, Disorder and *initiates* from (RFM-NL) are not presently found as individual classes/relations in the implementation files (under development), we hereby change some pre-existing classes to better express the RFM concept of C-Cs. Within the YAMATO class hierarchy, the RFM has more than one causal class (Fig. 2).

Causal Structure (CS) is a type of Continuant and is ***Constituted of*** Occurrents, specifically types of Sequence of Occurrents, such as CLO. Due to the fact that the RFM models different causal scenarios (e.g. simultaneous, sequential, etc.), 'sequence' is not intended to exclusively connote linear or sequential causation. For CS and SO to exemplify the idea of a causal chain, the occurrents composing the SO must be causally-linked. As explained in Section 3, without causal relations – or an otherwise explicit statement of causality – between occurrents a SO cannot be asserted as causal. For this reason we added the **Causally-linked Occurrent** class to the YAMATO hierarchy. It expresses the most general occurrent class of RFM causal chains. Combined in this way, we have the RFM/YAMATO conception of causal chains.

The RFM uses different kinds of Causally-linked Occurrents to distinguish causal scenarios. For example, the ordinary conception of causal chains is one of sequential happenings. Our **Sequential Causal Chain** class captures this scenario. Some causal scenarios may be held to exhibit **simultaneous causation** in which cause and effect occur simultaneously. Strictly speaking, absolute simultaneity may be precluded by modern physics, but given that an approximation is sufficient for contemporary biology

and daily life, the RFM allows for simultaneous causation. The class of **Pseudo-Simultaneous Causal Chain** has thereby been added (changed from YAMATOs 'pseudo-instantaneous') as a subclass of CLO to capture the different simultaneous causal occurrences the RFM seeks to represent. More specifically, if distinct changes occur simultaneously and without mediation, they are described as *directly connected* processes, or as **Directly-connected Causal Chains**. Consider an example.

A growing blood clot – a process with a clot as a participant – simultaneously causes the reduction of the cross section of the respective blood vessel, which simultaneously causes a reduction in oxygen supply to some organ. If left untreated, organ death will occur. At least three distinct processes occur simultaneously.[10] This is but one causal scenario to be considered in a causal account of disease. The RFM treats the (existence of the) blood clot as a state causing other states.

In the dependent entity portion of the YAMATO, **Dependent Causal Structure** (DCS) is a direct subclass of **Specifically Dependent** [**Entity/Continuant**],[11] a subclass of YAMATO: **Dependent Entity**. From a philosophical perspective, these sorts of entity are existentially dependent on specific individuals. A particular disease C-C is *specifically existentially dependent* on a particular organism, organism parts, and causal relata.

Disease is subclass of **Abnormal Causal Structure** (ACS), which is a subclass of DCS, a dependent entity/continuant. Disease, then, is a pathological causal structure. Although the implementation is complex, and in need of further definitions, there is nothing necessarily unintuitive about this particular idea.

The implementation – including OWL2 (OWL, 2014)[12] – indicates DCS is the most general class *inheriting* a CS. DCS can be understood as a derived class from, or defined by referring to, CS. As presented therein, the ***inherits*** relation holds between a DCS and CS, and allows the DCS to inherit the properties of CS without being the class subsumption relation. 'Class *X* inherits *Y*' means: Any property *P* obtaining for an instance/individual *y* of *Y* obtains for any instance *x* of *X*. Given that DCS (and therefore ACS) inherits the properties of CS, and CS is *constituted* of SO, the original idea that disease is a dependent continuant constituted of causal chains of disorders is captured. This is a complicated implementation, but there are, of course, simpler ways to formally implement this idea.

Clinical Disorders (or simply Disorders) are associated with dysfunctional or otherwise pathological functioning in organisms. There is no actual Disorder class in YAMATO. Rather, 'clinical disorders' in RFM-NL was intended to be broadly construed. The 'disorder' definition from (Scheuermann et al., 2009) is applicable if substituting 'components' with 'entities'. This was intended to permit different conceptualizations of disorder to be captured by RFM C-Cs. Disorders from Scheuermann et al. (2009) and OGMS, for instance, are represented as occurrents, specifically **Abnormal** (or Pathological) **States** (AS), a type of YAMATO: **State** (Yamagata et al., 2013, 2014).

'State' has been defined as one or more time-indexed properties (Yamagata et al., 2013) that may or may not change over time. It is a snapshot of a process. A state category, like a C-C category, abstracts or groups together certain realities to be represented as a single whole. Thus, bio-causal systems can be described and symbolically represented as units.

---

[10]Further analysis may lead to replacing a *simultaneous causation* predicate with another potentially more precise, perhaps non-causal one. In the given scenario, it may depend on the asserted ontological category of cross sections (Rovetto, 2013), and the relation may be one of non-causal determination for example. Generally speaking, it depends on the scenario and the relata.

[11]Equivalent to BFO: (Specifically) Dependent Continuant (BFO, 2013).

[12]OWL files by Kouji Kozaki and Riichiro Mizoguchi: http://rfm.hozo.jp/ontology/RFM.owl.

Disease C-Cs are ***initiated by*** some disorder. This means that the first relatum of a disease C-C is a disorder (AS). Any temporal reading of "initiated by" in the definition should include a causal association because temporal connection and overlap between occurrences does not imply causal connection.

If external agents are included in a wider causal description of etiopathogenesis, they effectively cause (if only partially) the initiating disorder. For example, invading infectious foreign bodies, pathogens, harmful doses of radiation, or some combined set of factors, cause the formation of a disorder (or some other pathological entity). This external (set of) *contributing causal factor(s)* serves as the cause of the first link of a disease C-C, but is not included in the RFM C-C (which is strictly ***located within*** the organism).

"Appearing" expresses *location*.[13] Disorders and disease C-Cs are localized within organisms. They are ***Located In*** and/or ***Occur Within*** it. External causal factors are not considered *part of* the disease. Given that the concept of a disease C-C is not necessarily unique to human beings, generalize the definition by replacing "human body" with "organism".

As pathological entities affect an organism, a disease spreads and the C-C changes. Depending on how the disease course unfolds – that is, what organism parts, properties and systems are affected and where – the causal chain may branch or fork . . . just like a river. **Abnormal Causal Structure**, the direct superclass of **Disease**, expresses this deviation from health. Figure 2 displays the central RFM classes in YAMATO as found in the January 22, 2015 version of the implementation file. The primitive causation relation is added in red.

Note that for dynamic scenarios, subclasses can be asserted with a temporally-indexed ternary *is_a* predicate: $is\_a(A, B, t) = def. \forall x, t[instance\_of(x, A, t) \rightarrow instance\_of(x, B, t)]$. The *has_effect_on* predicate has Occurrent arguments. We differentiate it from *causes* in an intuitive fashion: x might have an effect on y, but it does not cause y. It captures the idea of contributing to some occurrence, or of bringing about some change. This relation, like that of 'causes' (or other causal relations) and 'inherits', requires further development.

Consistency in ontology development is important to, among other things, avoid confusion. In order to have a NL definition that coheres with the implementation (and vice versa), we present two revised RFM disease definitions. They are stated in genus–species–differentia form in order to foster a computable formalization. The first revised definition takes RFM-NL and includes locational relations, Abnormal/Pathological State and the parent class of Disease.

> **Disease** (RFM-A) An **Abnormal Causal Structure** *constituted_of* one or more **Causal Chains** of **Abnormal** (**Pathological**) **States** *located* (or *occurring*) in an **Organism**, and *initiated by* at least one **Abnormal State**.

A FOL formalization of RFM-A is similar to RFM-FOL, but with the obvious substitutions. (i) + (ii) from (C-C FOL) express the causally-linked occurrents, and (iii) expresses the composition of causal chains.

The second definition closely matches the YAMATO hierarchy and implementation by including the classes and relations actually used in the implementation. As such this definition is more easily used as a computable definition. Using abbreviated class names, we have the following.

> **Disease** (RFM-B) An **ACS** that *inherits* a **CS**, which is *constituted_of* some **CLO**, and **CS** is *located* (or *occurs*) *in* an **Organism**, and **CLO** *has_part* (or is *initiated by*) at least one **AS**.

---

[13]See Smith et al. (2005) for a proposed formal definition.

Let us note two points. First, logical definitions for each class removes the need for explicitly including all the relations in these definitional variations. For example, an axiom describing ACS[14] removes the need to mention *inherits*; and "and initiated by ..." can be omitted since a formal definition of CS (should) includes a condition indicating a limiting case of one state. Second, a potential problem with RFM-B is this. If a 'location' predicate implies a dependency of the CS upon the organism, then this may contradict the RFM definition/classification of CS. Further development is in order.

An important class neither mentioned in the definition nor found in the current implementation (yet described at length in publications) is **Core Causal Chain** (Kozaki et al., 2012) (CCC). A disease type will often involve the same kinds of organs, bodily systems, processes and features. The course of a disease with its patterns of signs and symptoms vary from instance to instance while manifesting certain core patterns and changes.

To the degree that there are essential characteristics of disease, the CCCs of a disease are those causal chains *essential* to every instance of the disease type. They include the pathology common to a disease. If biomedical science discovers that all instances of a disease type involve some specific dysfunctional change, they should be represented using this class.

For example, cases (instance) of non-latent type 1 diabetes will involve elevated glucose and a deficiency of insulin, which give rise to a common set of effects on certain kinds of biological entity (blood vessels, kidneys, eyes). CCC serves to help identify a clinical case as an instance of a disease, distinguishing it from others. They are therefore necessary conditions, and minimum identity criteria for diseases. CCCs have Occurrents as parts, and are therefore parts of the Causally-linked Sequences of Occurrents. To express this idea we assert (A5).

(A5) $instance\_of(x, \textit{Causally-linked Occurrent}) \rightarrow \exists y [instance\_of(y, CCC) \wedge part\_of(y, x)]$.

Disease causal chains represent the causal structure of pathology through the full temporal range of a disease. They express the causal relationships of past to presently ongoing pathological changes. For example, RFM's **Historical Causal Chain** (Mizoguchi et al., 2012, p. 3) represents causally-linked **completed** events.[15] The onset and early stages of a disease, the course of disease progression, and any subsequent decline in the disease (the healing stages) can be represented as parts of one or more causal chains. Table 3 summarizes the RFM concept of disease causal chain from object and process ontology perspectives.

The question 'What is the causal structure of disease?' can be partially answered with a description of the relevant causally-connected pathological entities (i.e. disease causal chains). An exhaustive causal representation of a disease that reflects the changes affecting a patient over time – the overall etiopathogenesis of the disease – will, needless to say, be highly intricate. Contributing causal factors, (quasi)simultaneous causation, reciprocal causation, and bio-causal feedback mechanisms will be involved. In short, considering only sequential/linear causal scenarios does not do justice to biological reality, which is more interrelated and continuous (Schulz & Johansson, 2007). Desiderata for any causal account of disease includes the ability to represent this complexity. Toward that end, systems biology may inform representations of this bio-causal complexity.

> *"The elaboration of causal chains and networks ... is the business of all those areas of medical science which are concerned with elucidating the causation of disease"*.
>
> (Renton, 1994, p. 82, italics added)

---

[14]$instance\_of(x, ACS) \rightarrow \exists y [instance\_of(y, CS) \wedge inherits(x, y)]$.

[15]YAMATO: Events (but not Processes) are completed temporally extended wholes.

Table 3

Disease Causal Chains in Continuant and Occurrent terms

	Corresponding domain-specific entity/relationship	
	From an object-like perspective	From a processual perspective
(Disease) Causal chain(s), Causal structure(s):	A whole constituting a disease.	The disease developing over time.
		The causally structured disease course composed of pathological changes.
		Causally (inter)connected pathological occurrents.
Causal relata (represented by nodes in chain):	Biological continuants. Pathological continuants. Foreign bodies.	Processes. RFM Abnormal States.
Relationship(s) b/w relata	Causation relation, a sub-relation thereof, or a domain-specific bio-causal relation	

Notes: In combining the two ontological perspectives, some biological continuant, such as an organ, *participates in* pathological processes. Distinct causal relations can obtain between these entities (a subject for research).

5.2. OGMS and unstated causality

Other attempts at a general account of disease include the Ontology of General Medical Science (OGMS, 2014). It draws ideas from Scheuermann et al. (2009), and uses the BFO as its top-level ontology. It distinguishes disease from disease courses, a distinction reflecting the two ontological perspectives adopted by its top-level. OGMS: **Disease** is a subclass of BFO: **Disposition**, which is an entity that is realized through some **Process**. Diseases as dispositions are potentially realizable via **Pathological Processes**, and have some OGMS: **Disorder** as their physical basis. At the time of this writing, the OGMS disease definition was (Disease-OGMS): A disposition (i) to undergo pathological processes that (ii) exists in an organism because of one or more disorders in that organism.

The realizable nature of dispositions is intended to account for the fact that a disease may not manifest itself, i.e., without the organism exhibiting observable signs and symptoms, at all times it exists in the organism. A disease, then, is conceived as something that is potentially manifested via pathological processes. Once manifested, it causes other disorders. Causality in OGMS is unstated, implicit or stated indirectly. OGMS: **Etiological Process** and **Pathological Bodily Process**, for instance, have been described as "producing" abnormal bodily features and disorders.

As discussed in Section 3, production is causative. Etiological and pathological processes stand in causal relations[16] to various processes, but perhaps also to abnormal bodily features and disorders (depending on the relata constraints). A *causes* relation may be asserted as holding between pathological bodily processes, and a *production* relation between, say, an OGMS process and Disorder. There are various possibilities. Figure 3 displays part of OGMS and these candidate placements of causal relations.

One or more causation predicates are also potential additions to, say, the Relation Ontology (RO, 2014), which is used by OGMS and BFO. Mihăilă et al. (2013, p. 3) note that although general physical causation is of interest in biocuration efforts, terms such as 'cause' rarely appear in domain ontologies. Terms such as 'regulation' and 'stimulation' are used instead. Both the comment annotation for RO: **preceded_by**, and the description of the **Regulates** relation for the Gene Ontology (GO, 2013), implicitly support the idea that domain-specific relations, such as *Regulation*, are subtypes of a broader causal relation. These causative terms communicate causal influence, and are therefore candidate domain-specific

---

[16]*Production* is a potential sub-relation for a general causal relation, but would need to be explicitly differentiated.
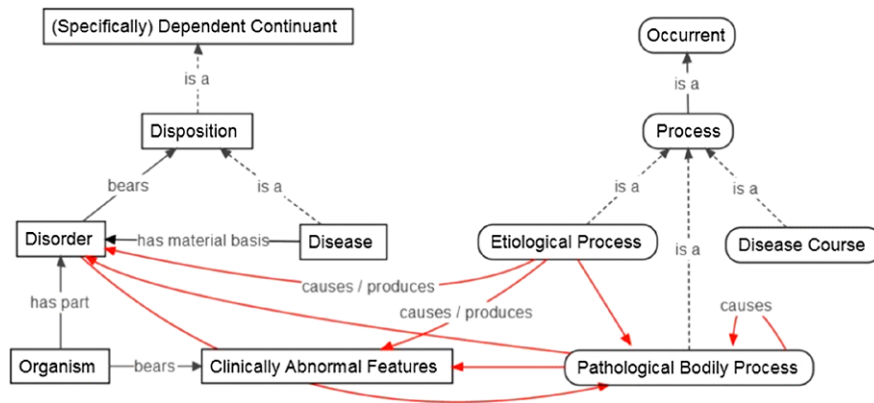
Fig. 3. OGMS classes with suggested areas for causal relations (red). Rectangles and rounded rectangles represent continuant and occurrents, respectively. (The colors are visible in the online version of the article; http://dx.doi.org/10.3233/AO-150147.)

causal relations. Moreover, the RO appears to be considering some general causal predicates such as 'causally related to' and 'causally upstream of',[17] which are best treated as domain-neutral relations.

Before we discuss strategies for interoperability between the RFM to OGMS, there is one concern to mention. Although dispositions capture the idea of disease manifestation/realization, categorizing all disease as disposition appears false. Medical practitioners work with concrete biological realities, and are naturally concerned with domain-specific phenomena. Far from being the sort of entity that biologists study, dispositions (at least in a pre-theoretic sense) are intuitively of a more intangible and perhaps inaccessible nature. How are dispositions detected when they are not manifested, for instance. It may appear too detached (or abstract) from the concrete biomedical reality practitioners encounter on a daily basis. Moreover, medical discourse using 'disposition' or 'predisposition' does not necessarily indicate reference to the posited corresponding ontological categories, and does not necessarily reflect widespread (or even individual) belief in diseases *as* (pre)dispositions. That is, dispositional talk in biomedicine does not imply or express a general scientific (or personal) view of disease as being a type of disposition. Some disease may be accurately characterized as dispositions, but not necessarily all. As such, clinicians may find a dispositional account of disease to be counterintuitive or contrary to their scientific understanding of the phenomena called 'disease'. A revision to OGMS may be in order.

That being said the RFM generally aims for interoperability with other ontological accounts of disease, while putting forth an alternative and autonomous account (a point we will return to in Section 5.5). If individual disease accounts succeed in providing representations of some, but not all, aspects of disease, then an application of one to the other may yield a more thorough representation of disease.

5.3. Conceptual overlap with OGMS and BFO

There are a few points of overlap between the RFM (and thus YAMATO) and OGMS (and BFO). As noted in Section 2, the River Flow Model of Diseases shares upper-level categories and distinctions with BFO and DOLCE. It specifically shares the classification of disease as a (Specifically) Dependent Continuant with OGMS (Mizoguchi et al., 2011, p. 64). A partial one-to-one mapping is therefore feasible. Another similarity is with respect to causal notions.

---

[17]See: http://obo-relations.googlecode.com/svn/trunk/src/ontology/ro.obo.

The RFM explicitly applies causal concepts to disease ontology development by using classes that can be formalized with one or more causation predicates. A formal account of cause–effect pairs and causation in DOLCE is found in Lehmann et al. (2004). The BFO discusses types of causal unity in BFO 2.0 (2013), but not causation or causal relations. Each ontological theory seems to share an understanding of the general idea of **causal unity**.

For example, according to the RFM and YAMATO, some internal processes of a solid material object help sustain (or maintain) its existence. Internal interactions and molecular bonds maintain its structural integrity. This is on conceptual par with the second kind of causal unity described in BFO 2.0 (2013), namely "causal unity via internal physical forces", as well as Topological Unity as described by DOLCE (Guarino & Welty, 2000).

Generally speaking, causal unity helps make something a physical, but at least a material, object. As Bird (2014, pp. 24–25) states, a physical object is more than just the mereological sum or physical proximity of its parts; "some kind of causal unity among those parts is also required". Material objects are those composed of some portion of matter. Physical objects are, more broadly, spatio-temporal wholes.

The unity of disease causal chains is partly in virtue of the causal connectedness of the interacting, dysfunctional parts, processes, and the surrounding environments. According to the RFM, a disease, like a river, is an example of this unity.

Looking more closely at OGMS classes we see another similarity. OGMS: **Etiological Process**, is defined as "a process in an organism that leads to a subsequent disorder" (Scheuermann et al., 2009). The definition implicitly expresses the causality at work through "leads to". "Subsequent" can be interpreted as connoting the temporal order of a causal or causal-like relation and the occurrent relata. Finally, "etiological" has a causal meaning by definition.

5.4. Strategies toward harmonization between disease accounts

Although OGMS defines Disorder as a Material Entity (a BFO: Continuant), the RFM would represent them as Abnormal States in causal chains (Causally-linked Occurrents). These states can be interpreted as the **Qualities** of OGMS: **Disorders**, perhaps a Quality Pattern. A problem, not addressed here, is that Disorders are Continuants, but YAMATO: States are Occurrents.

Consider another idea. If OGMS persists with a dispositional classification, its Disease class can be described as an Abnormal Causal Disposition, which would be similar (perhaps equivalent) to some RFM classes.

In any event, at least two OGMS classes are candidate causal relata in RFM causal chains. The class **Etiological Process** represents either (a) a process causing the first link in the chain, or (b) a proper part (i.e., the first, or some other, link) of the chain. OGMS: **Pathological Process** serves as relata in *any* part of a disease C-C. Indeed, the processes mentioned in "undergo pathological processes" from the OGMS definition are causally efficacious and therefore easily accepted as C-C processes. It is not contradictory for OGMS to assert its pathological process class as causally-connected. Additionally, if the BFO (or OGMS) adds a (disease) causal chain/structure class, then it will be feasible to include them in the OGMS Disease definition.

Some Disorders may be asserted as ***participating in*** the processes that are parts of a RFM Causally-linked Occurrent or subclass C-C. A disease definition along the lines of D1 expresses the idea (Fig. 4).

(D1) **Disease** *is_a* **Dependent Entity** (e.g., **Dependent Continuant**) that is *constituted_by* **Disorders** *participating_in* causally-linked **Pathological Processes** (RFM **Causally-linked Occurrents**) *located_in* and *occurring_in* an **Organism**.
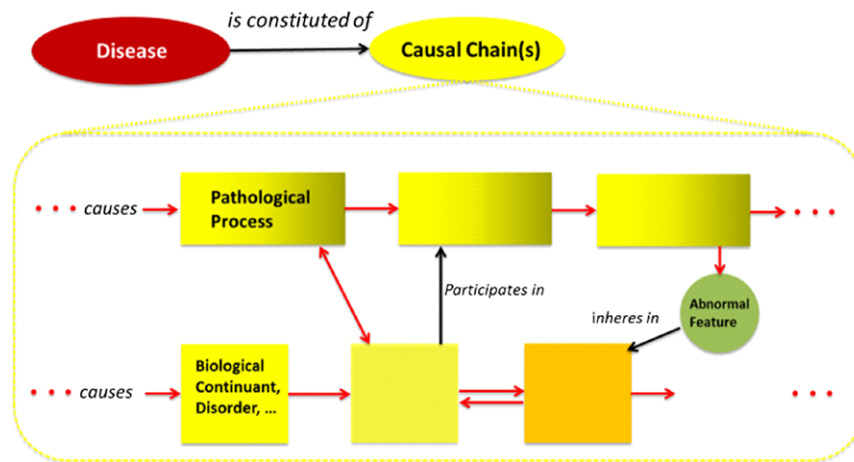
Fig. 4. The causal structure of a generic disease causal chain. The dotted yellow lines from the Causal Chain node represent an expansion/zooming. Squares are causally-linked biological continuants. Rectangles are causally-linked processes. Red arrows signify one or more causal relations. Two opposing arrows indicate *mutual causal interaction* (not necessarily simultaneous, but perhaps causal feedback mechanisms). These represent different ways to model causal chains. Statements about continuants causing others may be defined in terms of their participation in causal processes, for instance. (The colors are visible in the online version of the article; http://dx.doi.org/10.3233/AO-150147.)

The top part of the yellow rounded-rectangle displays a disease causal chain in terms of pathological processes, with the lower part in terms of causally-linked biological continuants. If one seeks to describe a DCC using only processes, then the top part is appropriate. Together, top and lower portions demonstrate the combining of process-like and object-like ontological perspectives (Table 3). If processes presumably play the preferred role of causal relata.

At least part of an OGMS: **Disease Course** is a causally-connected sequence of processes, i.e. a C-C of pathological processes. An RFM C-C class, then, groups causally-linked occurrent parts of the disease course. Since RFM **Processes** are state-changes, OGMS: **Pathological Processes** can ideally (c) be described in terms of RFM **Abnormal States**, and therefore play the role of causal relata in C-Cs. In this way, we have conceivable options (a)–(c) toward applying RFM concepts to at least one other disease account. A challenge that requires further consideration is in theory-specific upper-level ontological distinctions, specifically the treatments of Processes and States.

To partially address any incompatibility between YAMATO: Process and BFO: Process, consider this. If either (or both) BFO or YAMATO generalize their Process (or Occurrent classes), then they may be in a better place toward harmonization. Alternatively, if either (or both) disease account assert a Pathological Occurrent superclass (to represent types of disease occurrents) in their disease models, then an equivalence between the respective classes appears viable. This class can, for instance, subsume RFM: Abnormal State as well as OGMS: Pathological (Bodily) Process. Classes from other biomedical ontologies, such as the Gene Ontology, the Infectious Disease Ontology (IDO, 2013), and the Disease Ontology (DO, 2013) should be applicable in a similar fashion. Further research is in order.

To the extent that the RFM concentrates on the instance-level,[18] another idea relating the two theories is as follows. When a disease disposition is realized at time t, the processes beginning the disease course[19] are the first links in a disease causal chain at t. That is, the **Realization Processes** of OGMS

---

[18]This would be consistent with philosophical accounts holding causation to be of the token, not type, variety.

[19]Either simultaneous with t or at $t + 1$.

disease dispositions **are** processes (causal relata) in RFM disease causal chains. In other words, C-Cs are the manifestation(s) of supposed disease dispositions. C-C classes can therefore be understood as involving an abstraction from instances of causally-connected occurrences in the world.

For example, if Arterial Hypertension is a disposition, then its realization process would stand in certain causal relations to other parts of the disease course directed to heart disease. If treated, the hypertension is not manifested via some elevated blood pressure process, i.e., the respective C-C or causal process does not occur. If untreated, there is high blood pressure: the C-C occurs.

The problem with this example is that if Arterial Hypertension simply is high blood pressure (not a measurement of it) or that which causes it, then instances of hypertension are continuously ongoing, abnormal processes of blood flow. There may be no observable symptoms, but these processes still occur *at some scale* in the organism. They remain causally efficacious upon the organism. Thus, there is always an ongoing process, not necessarily a disposition.

5.5. RFM as a non-dispositional account of disease

The River Flow Model does not require disposition classes to define disease, but allows for them. Moreover, talk of dispositions is unnecessary for *some* diseases. The characterization of disease as being a type of disposition may merely be a summary notion for (unknown or unobserved) causal chains, causal networks and causal processes **at some scale or level of granularity**. Whenever a disorder or disease is present in an organism, there is some causally influential disruption, interruption, dysfunctionality, pathology or abnormality present. There is some disordered organ, biological system, and/or some ongoing pathological process (as with hypertensive blood flows) occurring at some scale … regardless of whether there are resultant observable changes. The current state of medical science may not afford detection or measurement of certain subtle, complex or fine-grain disruptions, processes or interactions.

One or more disease C-Cs at one or more scales is constantly influencing the organism. When they cause specific pathological effects at the same (or more coarse-grained) levels, signs and symptoms become observable to the physician and patient. In short, some pathology may always be present.

Consider the blood clot example. The clot decreases the cross section of the blood vessel. The result is an abnormally narrow channel for blood to flow through. Rather than a dispositional account according to which the disorder affords the existence of a disposition, we simply state the causal realities. The clot – the existence of which already establishes some pathology and the partial causal basis for subsequent pathological entities – causes certain abnormal processes, e.g., abnormal blood flow processes. It disrupts what otherwise is normal blood flow. The dysfunctional blood flow processes continuously unfold for the duration of the clot's existence regardless of the degree of deviation from healthy blood flow. They merely do not cause certain signs or symptoms at some level.

OGMS would describe this scenario as one of an unrealized disposition, but this may be unnecessary for some diseases because talk of *unrealized dispositions* could simply be a shorthand reference to the existent pathology at some finer biological scale, or to some as yet unknown phenomenon or causal interaction. What is actually happening is a constantly present/occurring pathology at different levels of description, not all of which are necessarily known or understood. The clot and altered blood flow may eventually cause some pathological condition[20] at another scale that is observable (signs) or that does

---

[20]*Condition* is used broadly, but may be interpreted in the sense of Schulz et al. (2011) where we read: "By introducing the disjunctive class bt:PathologicalEntity (sct:Condition) we are able to represent diseases without specifying the ontological category, and to relate them to anatomical objects and spaces".

satisfy some clinical threshold. Having a disease without exhibiting signs, or without having observable pathological processes, does not necessarily mean the organism bears an unrealized (disease) disposition. If the preceding line of reasoning is correct, then not all diseases are types of disposition.

Regardless of the classification or ontological description of disease, ultimately what is important is the degree to which the ontological characterization can help support the prevention and curing of illness/disease.

6. Discussion

Two general points about disease causation and ontology are easily accepted. (i) The pathogenesis of diseases is comprehensible, lawfully unfolding in time according to knowable principles, and thus has discoverable causes. This affords the manifestation and perception of recurrent but variable, discoverable and determinable patterns, which help form putative disease kinds. (ii) Some dysfunctional or disordered organism parts and processes are causally connected. They stand in certain causal relations (causal . . . dependence, necessitation, determination, etc.) to one another. They are interconnected such that changes in or *disruptions to* some body part, process, property, state or system, affects others.

Clinicians and biomedical researchers study not merely disordered organs and their characteristics, but also the dysfunctional or pathological processes occurring in the patient. Biomedical specialists may also be unconvinced by having a specific entity (e.g., disorder), classified/characterized in a specific way, take an exclusive role in disease representation. A generic disease C-C or causal structure class, then, communicates the causal interconnectedness of some pathological entities, their causal efficacy and the way they temporally unfold. In other words, it expresses the causal aspect of pathology.

With this in mind, more inclusive causal and pattern-based definitions of disease can be formed by using higher-level categories such as Pathological Entity, Condition (as in footnote [20]), and so on. For example, a broader disease definition that allows for any pathological entity playing a causal role in the disease development is this: (**Disease-Generic**) Disease is a dependent entity constituted by causal chains of pathological entities located (and/or occurring) in an organism.

This definition retains the intuitions that disease is existentially dependent and causally efficacious without committing to a more specific classification (e.g., pathological causal structure, disposition, process, etc.). It encompasses RFM and OGMS conceptions of disease, while not necessarily requiring all the relations used in the RFM conception of disease (similar to Fig. 4). A problem with this formulation, however, is that it may be too broad. In replacing 'Pathological Entities' with 'Pathological Occurrents' (and defining it appropriately), one constrains the relata while allowing for both Abnormal States and Pathological Process classes. It also allows for domain-specific specializations, e.g. biological occurrents, mental occurrents, etc.

In Section 3.1 the category of Pattern was mentioned as being related to causation, and thus as being potentially applicable to the representation of disease causation. To briefly explore one of the pattern-based characterizations, consider this generic disease definition (the terms of which are undefined). (**Disease-Pattern**) Disease may be conceived as either a sort of dependent entity constituted by patterns of causally-connected pathological occurrents, or a sort of pathological pattern of causally-connected occurrents. Continuant disorders or pathological continuants would be the participants in the respective disease processes.

To the extent that Patterns and Dispositions are potentially realizable entities, we may also assert the following translation or equivalence. A disease disposition is equivalent to a disease causal pattern. The realization of the former is then equivalent to that of the latter.

To speculate on an application, consider this. Envision a multi-scale graphical simulation in which the causal influence of a disease is visualized at different scales, and superimposed on a simulation of the organism body. Both typical and atypical (but possible) disease courses can be represented through all temporal modalities: past, present and future. The usefulness of such a system is as a pedagogical tool for students of biology and medicine, and a tool medical researchers and practitioners. It would help in visualizing how various diseases commonly unfold in time, and the unique permutations of effects and disease courses. It could stimulate comprehension of: factors believed to be causally-relevant for the onset, development, and remediation of a disease; specific effects of disease entities, the primary changes characterizing a disease, and specific causal relationships involved.

The preceding discussion has offered some ideas toward better incorporating the causal aspect of etiopathogenesis in ontological accounts of disease. As a final, very important, point, note that the descriptions of disease discussed in this paper do not mention the causal roles played by other important factors in pathology development. Examples include: diet, mental states, stress-levels, lifestyle, social conditions, etc. These are essential factors to represent in any ontological account of disease because they are causally related to (often pivotally so) susceptibility to, onset, progression and curing of disease.

7. Conclusion

We have revisited the river, and begun to navigate the currents of causality in disease ontology. The River Flow Model of Diseases (RFM) is an ontological account representing the causal structure of pathology. It differs from other ontological theories in that it explicitly takes causation into account. Diseases, such as Diabetes, are described as being constituted of causal chains of pathological occurrents in an organism. Disease causal structures are modeled as wholes that are constituted by abnormal states and processes (causal relata) composing causally-linked occurrents (causal chains) that may temporally unfold in a sequential or simultaneous manner.

This paper provided a general discussion of causality and causal chains vis-à-vis disease ontology; and a clarification and development of the RFM, including a disease definition that more closely coincides with the formal implementation of the theory. By providing formal axioms, we have begun a first-order formalization of the RFM. We also demonstrated the commonalities between the RFM/YAMATO, OGMS and the BFO, and discussed avenues for their harmonization.

The RFM is in accord with some "desirable features" for disease ontologies suggested by Bodenreider and Burgun (2009). For example, although OWL has been unnecessary for applications of the RFM in Japan, OWL translations of the RFM are under development. This is in accord with the feature of having a standard and friendly format. In reusing existing ontological category terms, the RFM is on course toward interoperability with other ontologies, another suggested feature.

Further development of the RFM should include: clear natural language and formal definitions of YAMATO classes, curation efforts for greater comprehensibility; additional development of RFM causal relations; a complete first- or higher-order formalization; and applications to other disease ontologies. Specific research topics include systems biology and biological pathways, the former of which overlaps with the RFM in that both appear to have holistic approaches. Representing pathological causal feedback mechanisms and cyclical causality is also highly relevant, and should be included within the scope of the RFM (and other disease accounts).

Criteria for the success of the RFM, like other ontological theories of disease, include: faithfulness in representing disease entities; consistency and coherence; applications to biomedical data and specific diseases, and the degree to which it assists medical practitioners and researchers.

Just as neighboring rivers are likely to flow into the same ocean, so the "flow" (goal) of theories of disease must be toward supporting the ocean of research on preventing and curing disease. In other words, a shared purpose among medical, informatics and ontology communities involved in the study, management, and representation of biomedical data and disease entities should be to heal.

Acknowledgements

References

Allan, D.J. & Castillo, M.M. (2007). *Stream Ecology: Structure and Function of Running Waters* (2nd edn., p. 31). Springer.

Allen, J.F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, *26*(11), 832–843. doi:10.1145/182.358434. Available at: http://doi.acm.org/10.1145/182.358434.

Basic Formal Ontology (BFO) (2013). Available at: http://www.ifomis.org/bfo (Accessed 25 September 2013).

Basic Formal Ontology 2.0 (2013). Textual Specification (BFO 2.0) (User's Manual Draft). Available at: http://bfo.googlecode.com/svn/trunk/docs/bfo2-reference/BFO2-Reference.docx (Accessed 25 September 2013).

Bird, A. (2014). The Metaphysics of Natural Kinds, Unpublished manuscript. Available at: http://eis.bris.ac.uk/~plajb/research/inprogress/Metaphysics_Natural_Kinds.pdf (Accessed 21 August 2014).

Bodenreider, O. & Burgun, A. (2009). Towards desiderata for an ontology of disease. In *Proceedings of the First International Conference on Biomedical Ontology* (pp. 39–42).

Borgo, S. & Mizoguchi, R. (2014). A first-order formalization of event, object, process and role in YAMATO. In *Proceedings of the 8th International Conference on Formal Ontology in Information Systems (FOIS2014)* (pp. 79–92).

Bunge, M. (1979). *Causality and Modern Science*. New York: Dover Publications, Inc.

Causation Annotated Bibliography (2008). *AHRC Metaphysics of Science Project*. UK: Univ. of Bristol. Available at: http://www.bris.ac.uk/metaphysicsofscience/bibliographies/causation/causationbibliography.doc (Accessed 28 June 2015).

Clinical Medical Ontology (2013). Department of Medical Informatics, The Graduate School of Medicine, The University of Tokyo. Available at: http://www.m.u-tokyo.ac.jp/medinfo/medont2010-2012proj/ (Accessed 25 September 2013).

Cohen, M.S. (2006). *Heraclitus, History of Ancient Philosophy*. Lecture Notes. Univ. of Washington. Available at: http://faculty.washington.edu/smcohen/320/heracli.htm (Accessed 25 September 2013).

Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (2013). Available at: http://www.loa.istc.cnr.it/old/DOLCE.html (Accessed 20 August 2015).

Disease Ontology (DO) (2013). Available at: http://disease-ontology.org/ (Accessed 20 December 2013).

Disease Ontology (Disease Chain LD) (2014). Available at: http://lodc.med-ontology.jp/ (Accessed 28 June 2015).

Dominic, M. (2009). Concepts of Disease and Health. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2009 edn.). Available at: http://plato.stanford.edu/archives/sum2009/entries/health-disease/ (Accessed 25 September 2013).

Galton, A. & Mizoguchi, R. (2009). The water falls but the waterfall does not fall – New perspectives on objects, processes and events. *Journal of Applied Ontology*, *4*(2), 71–107.

Gene Ontology (GO) (2013). Available at: http://www.geneontology.org/ (Accessed 25 September 2013).

Graham, D.W. (2011). Heraclitus. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2011 edn.). Available at: http://plato.stanford.edu/archives/sum2011/entries/heraclitus/ (Accessed 25 September 2013).

Grenon, P. & Smith, B. (2004). The cornucopia of formal-ontological relations. *Dialectica*, *58*, 279–296.

Guarino, N. & Welty, C. (2000). Ontological analysis of taxonomic relationships. In A. Laender and V. Storey (Eds.), *Proceedings of ER-2000: The International Conference on Conceptual Modeling*. Lecture Notes in Computer Science. Springer.

Hebert, P. & Ontario, B. (2008). Physical properties of rivers. Available at: http://www.eoearth.org/view/article/155233 (Accessed 25 September 2013).

Hesslow, G. (1993). Do we need a concept of disease? *Theoretical Medicine*, *14*, 1–14.

HOZO Ontology Editor (2013). Available at: http://www.hozo.jp/ (Accessed 20 August 2015).

Hucklenbroich, P. (2014). 'Disease Entity' as the Key Theoretical Concept of Medicine. *The Journal of Medicine and Philosophy*, *39*(6), 609–633. doi:10.1093/jmp/jhu040.

Huemer, M. & Kovitz, B. (2003). Causation as simultaneous and continuous. *The Philosophical Quarterly*, *53*(212), 556–565.

Imai, T., Kou, H., Zhou, J., Kozaki, K., Mizoguchi, R. & Ohe, K. (2009). Japan medical ontology development project for advanced clinical information systems. In *10th International HL7 Interoperability Conference, IHIC 2009* (presentation). Available at: http://www.hl7.jp/ihic2009/presentation/pr20.pdf (Accessed 25 September 2013).

Infectious Disease Ontology (IDO) (2013). Available at: http://infectiousdiseaseontology.org/page/Main_Page (Accessed 20 December 2013).

Johansson, I. & Lynoe, N. (2008). *Medicine and Philosophy – A Twenty-First Century Introduction*, Frankfurt: Ontos.

Joshi-Tope, G., et al. (2005). Reactome: A knowledgebase of biological pathways. *Nucleic Acids Research*, *33*(Suppl. 1), D428–D432. Available at: http://nar.oxfordjournals.org/content/33/suppl_1/D428.long (Accessed 22 July 2014).

Kim, J. (1973). Causation, nomic subsumption and the concept of event. *The Journal of Philosophy*, *70*(8), 217–236.

Kim, J. (2004). Causes and counterfactuals. In E. Sosa and M. Tooley (Eds.), *Causation*. Oxford Readings in Philosophy (pp. 205–207). New York: Oxford Univ. Press.

Kleinberg, S. & Hripcsak, G. (2011). A review of causal inference for biomedical informatics. *Journal of Biomedical Informatics*, *44*(6), 1102–1112.

Kozaki, K., Kou, H., Yamagata, Y., Imai, T., Ohe, K. & Mizoguchi, R. (2012). Browsing causal chains in a disease ontology. In *Poster & Demo Notes of the 11th International Semantic Web Conference*, Boston, USA. The Semantic Web Science Association (SWSA).

Kozaki, K., Yamagata, Y., Mizoguchi, R., Imai, T. & Ohe, K. (2015). Disease compass – A navigation system for disease knowledge based on ontology and linked data techniques. In *Proceedings of International Conference on Biomedical Ontology (ICBO2015)*.

Lagnado, D. (2011). Causal thinking. In P. McKay Illari, F. Russo and J. Williamson (Eds.), *Causality in the Sciences*. Oxford Univ. Press.

Lehmann, J., Borgo, S., Masolo, C. & Gangemi, A. (2004). Causality and causation in DOLCE. In A.C. Varzi and L. Vieu (Eds.), *Proceedings of the Third International Conference on Formal Ontology in Information Systems (FOIS2004)*. Amsterdam: IOS Press.

Lillie, R.S. (1940). Biological causation. *Philosophy of Science*, *7*(3), 314–336.

Losee, J. (2011). *Theories of Causality – From Antiquity to the Present*. New Brunswick: Transaction Publishers.

Loux, M.J. (2006). *Metaphysics: A Contemporary Introduction* (3rd edn.). New York: Routledge.

Lowe, E.J. (2002). *A Survey of Metaphysics*. New York: Oxford Univ. Press.

Marcum, J.A. (2011). Philosophy of Medicine. In *The Internet Encyclopedia of Philosophy*, ISSN: 2161-0002. Available at: http://www.iep.utm.edu/medicine/ (Accessed 25 September 2013).

Masolo, C., Borgo, S., Gangemi, A., Guarino, N. & Oltramari, A. (2003). *WonderWeb Deliverable D18 – Ontology Library* (final). IST Programme of the Commission of the European Communities, IST Project 2001-33052 WonderWeb: Ontology Infrastructure for the Semantic Web. Available at: http://wonderweb.man.ac.uk/deliverables/documents/D18.pdf (Accessed 31 July 2015).

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. & Schneider, L. (2003). WonderWeb Deliverable D17: The WonderWeb Library of Foundational Ontologies Preliminary Report.

Mellor, D.H. (1995). *The Facts of Causation*. New York: Routledge.

Mihăilă, C., Ohta, T., Pyysalo, S. & Ananiadou, S. (2013). BioCause: Annotating and analysing causality in the biomedical domain. *BMC Bioinformatics*, *14*, 2. doi:10.1186/1471-2105-14-2.

Mizoguchi, R. (2004). Tutorial on ontological engineering – Part 3: Advanced course of ontological engineering. *New Generation Computing*, *22*(2), 198–220.

Mizoguchi, R. (2010). YAMATO: Yet another more advanced top-level ontology. In *Proceedings of the Sixth Australasian Ontology Workshop* (pp. 1–16).

Mizoguchi, R., Kouji, K., Imai, T. & Ohe, K. (2012). Identity tracking of a disease as a causal chain. In *Proceedings of the 3rd International Conference on Biomedical Ontology (ICBO2012)*. Available at: http://ceur-ws.org/Vol-897/sessionJ-paper26.pdf (Accessed 27 July 2014).

Mizoguchi, R., Kozakil, K., Kou, H., Yamagata, Y., Imai, T., Waki, K. & Ohe, K. (2011). River Flow Model of diseases. In *Proceedings of International Conference on Biomedical Ontology* (pp. 63–70).

Olsen, J. (2003). What characterizes a useful concept of causation in epidemiology? *Journal of Epidemiology and Community Health*, *57*, 86–88.

Ontology for General Medical Science (OGMS) (2014). Available at: http://code.google.com/p/ogms/; http://bioportal.bioontology.org/ontologies/OGMS (Accessed 23 October 2014).

Raisanen, U., Bekkers, M.-J., Boddington, P. & Sarangi, S. (2006). The causation of disease – The practical and ethical consequences of competing explanations. *Medicine, Health Care and Philosophy*, *9*, 293–306. doi:10.1007/s11019-006-9007-5.

Relation Ontology (RO) (2014). Available at: https://github.com/oborel/obo-relations (Accessed 31 July 2015).

Renton, A. (1994). Epidemiology and causation: A realist view. *Journal of Epidemiology and Community Health*, *48*, 79–85.

Rothman, K. & Greenland, S. (1998). Causation and causal inference. In *Modern Epidemiology* (3rd edn.). Lippincott.

Rovetto, R. (2013). Shaping up: The phenotypic quality ontology and cross sections. In *Proceedings of the Second Interdisciplinary Workshop the Shape of Things (Shapes 2.0)* (Vol. 1007, pp. 125–136). Available at: http://ceur-ws.org/Vol-1007/paper10.pdf.

Schaffer, J. (2008). The metaphysics of causation. In E.N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2008 edn.). Available at: http://plato.stanford.edu/archives/fall2008/entries/causation-metaphysics/.

Scheuermann, R.H., Ceusters, W. & Smith, B. (2009). Toward an ontological treatment of disease and diagnosis. *Summit on Translational Bioinformatics*, *2009*, 116–120. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3041577/.

Schulz, S. & Johansson, I. (2007). Continua in biological systems. *Monist*, *90*(4).

Schulz, S., Spackman, K., James, A., Cocos, C. & Boeker, M. (2011). Scalable representations of diseases in biomedical ontologies. *Journal of Biomedical Semantics*, *2*(Suppl. 2), S6. Available at: http://www.jbiomedsem.com/content/2/S2/S6.

Simon, J., Dos Santos, M., Fielding, J. & Smith, B. (2006). Formal ontology for natural language processing and the integration of biomedical databases. *International Journal of Medical Informatics*, *75*, 224–231.

Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A.L. & Rosse, C. (2005). Relations in biomedical ontologies. *Genome Biology*, *6*, R46.

Sosa, E. & Tooley, M. (Eds.) (2004). *Causation*. Oxford Readings in Philosophy. New York: Oxford Univ. Press.

Susser, M. (1991). What is a cause and how do we know one? A grammar for pragmatic epidemiology. *American Journal of Epidemiology*, *133*(7), 635–648.

Swinburne, R. (1997). The irreducibility of causation. *Dialectica*, *51*(1), 79–92.

von Watcher, D. (2010). Roman Ingarden's theory of causation revised. *Polish Journal of Philosophy*, *4*, 182–196.

Web Ontology Language (OWL) (2014). Available at: http://www.w3.org/TR/owl-features/; http://www.w3.org/TR/owl2-overview/ (Accessed 6 October 2014).

Weed, D.L. & Parascandola, M. (2001). Causation in epidemiology. *Journal of Epidemiology and Community Health*, *55*, 905–912.

Whitbeck, C. (1977). Causation in medicine: The disease entity model. *Philosophy of Science*, *44*(7), 619–637.

Yamagata, Y., Kou, H., Kozaki, K. & Mizoguchi, R. (2013). Ontological modeling of interoperable abnormal states. In *Proceedings of JIST2012*. Lecture Notes in Computer Science (Vol. 7774, pp. 33–48). Heidelberg: Springer.

Yamagata, Y., Kozaki, K., Imai, T., Ohe, K. & Mizoguchi, R. (2014). An ontological modeling approach for abnormal states and its application in the medical domain. *Journal of Biomedical Semantics*, *5*, 23. Available at: http://www.jbiomedsem.com/content/5/1/23.