Browsing Causal Chains in a Disease Ontology

Kouji KOZAKI[1], Hiroko KOU[1], Yuki Yamagata[1], Takeshi IMAI[2], Kazuhiko OHE[2]
and Riichiro MIZOGUCHI

[1]The Institute of Scientific and Industrial Research, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047 Japan
[2]Department of Medical Informatics, Graduate School of Medicine, The University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan
{miz, kozaki, kou, yamagata}@ei.sanken.osaka-u.ac.jp
{ken, kohe}@hcc.h.u-tokyo.ac.jp

Abstract. In order to realize sophisticated medical information systems, many medical ontologies have been developed. We proposed a definition of disease based on River Flow Model which captures a disease as a causal chain of clinical disorders. We also developed a disease ontology based on the model. It includes definitions of more than 6,000 diseases with 17,000 causal relationships. This demonstration summarizes the disease ontology and a browsing system for causal chains defined in it.

Keywords: disease ontology, causal chain, ontology visualization

1 Introduction

In these days, medical information systems store huge amount of data. Semantic technologies are expected to contribute to effective use of them, and many medical ontologies such as OGMS [1], DOID [2], and IDO [3] have been developed for realizing sophisticated medical information systems. They mainly focus on the ontological definition of disease with related properties. The authors proposed a definition of a disease involving capturing a disease as a causal chain of clinical disorders and a computational model called *River Flow Model of Disease* [4, 5]. Based on the model, we developed a disease ontology which includes definitions of about 6,000 diseases with causal relations between 17,000 clinical disorders (abnormal state). This demonstration shows a system to browse causal chains defined in the disease ontology.

This article is organized as follows. The next section overviews the *River Flow Model of Disease* discussed in [4, 5]. Section 3 summarizes developments of the disease ontology and a browsing system for the disease ontology. Finally, we present concluding remarks together with future work.

2 River flow model of disease

After it begins to exist, a typical disease, as a dependent continuant, enacts extending, branching, and fading processes before it disappears. Thanks to these processes, a

disease can be identified as a continuant that is an enactor of those processes. Such an entity (a disease) can change according to its phase while keeping its identity. On the basis of this observation, we defined a disease as:

Definition 1: Disease [4]
　A disease is 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.

　　Note that, although any disease has dynamic flows of the propagation of causality as its internal processes, it is the enactor of its external processes, such as branching and extending its causal chain of disorders. When we collect individual causal chains belonging to a particular disease type (class), we are able to find a common causal chain (partial chain) that appears in all of the instance chains. By generalizing such a partial chain, we obtain the notion of a core causal chain of a disease as follows:

Definition 2: Core causal chain of a disease
　A sub-chain of the causal chain of a disease whose instances are included in all the individual chains of all instances of a particular disease type. It corresponds to the essential property of a disease type.

　　Definition 2 provides a necessary and sufficient condition for determining the disease type to which a given causal chain of clinical disorders belongs. That is, when an individual causal chain of clinical disorders includes instances of the core causal chain of a particular disease type, it belongs to the disease type. We can thus define such a disease type that includes all possible variations of physical chains of clinical disorders observed for patients who contract the disease. According to a standard definition of subsumption, we can introduce an *is-a* relation between diseases using the chain-inclusion relationship between causal chains.

Definition 3: *Is-a* relation between diseases
　Disease A is a supertype of disease B if the core causal chain of disease A is included in that of disease B. The inclusion of nodes (clinical disorders) is judged by taking an *is-a* relation between the nodes into account, as well as sameness of the nodes[4].

　　Definition 3 helps us systematically capture necessary and sufficient conditions of a particular disease which roughly corresponds to the so-called "main pathological/ etiological conditions". Fig. 1 shows the main types of diabetes constituted by corresponding types of causal chains. Assume, for example, that (non-latent) diabetes and type-I diabetes are respec-
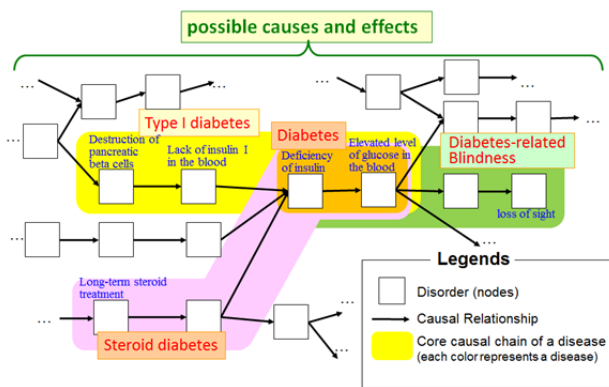

Fig. 1 Types of diabetes constituted of causal chains.
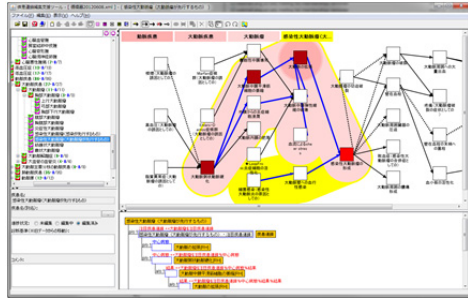
Fig. 2 A visual editing tool for causal
chains to define disease concepts.

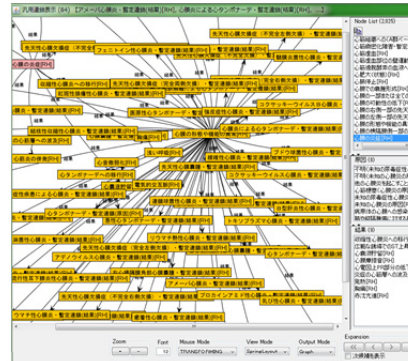Fig. 3 A browsing tool for causal chains.

tively defined as *<deficiency of insulin → elevated level of glucose in the blood>* and
*<destruction of pancreatic beta cells → lack of insulin I in the blood → deficiency of
insulin → elevated level of glucose in the blood>*. Then, we get *<type-I diabetes is-a
(non-latent) diabetes>* according to Definition 3.

3   Development of a disease ontology and a browsing tool for it

Based on the model discussed in the previous section, we developed a disease
ontology. Its design policy and conceptual structures of the ontology were decided
through repeated discussions by ontology engineer and medical expert. We defined its
upper level concepts (classes) such as clinical disorders (abnormal states), causal
chains, causal relationships (cause and effect), etc. based on YAMATO[1]. Disease
concepts were defined sub-class of these concepts by clinicians in 12 special fields.
Although the disease ontology are developed using Hozo[2], we developed a visual
editing tool for it so that clinicians can easily edit the definition of disease concepts.
Fig. 2 shows its user interface. It visualizes causal chains defined in a selected disease
as directed graph like Fig.1. In the graph, nodes represent clinical disorders and links
represents causal relationships between them. When users edit the graph, it automati-
cally translated into ontology in Hozo's format. The ontology could be exported in
OWL formats thanks to the export functions of Hozo.

Note here that each clinician defined disease concepts in his/her special field
without knowing how other diseases were defined in other filed by others. After they
finished defining disease concepts, we collected all causal relationships from all dis-
ease concepts defined in the 12 special fields. Then, we combined causal chains
which included the same clinical disorder. As the result, we obtained causal chains
which include about 17,000 clinical disorders defined in 6,000 diseases. They repre-
sent possible causal chains in human body.

In order to browsing these causal chains, we developed a browsing tool (Fig. 3).
It visualizes causal chains defined in the disease ontology and the user can browse

---

[1] http://www.ei.sanken.osaka-u.ac.jp/hozo/onto_library/upperOnto.htm

[2] http://www.hozo.jp/

them through some functions such as searching, tracing, changing layout, zooming etc. Although it is implemented as a client application using Hozo's ontology API, we plan to develop web services version of it. We also consider publishing the disease ontology as Linked Open Data with SPARQL endpoint to get their causal chains.

Currently, we focus on definitions of disease in order to provide a basic knowledge for medial information systems without considering particular applications. When we use the disease ontology for a specific application, we will consider making some extension the ontology and browsing system according to the purpose.

4. Concluding Remarks

We developed a disease ontology based on River Flow Model and a browsing tool for causal chains defined in it. Because the ontology is based on ontological consideration of causal chains, it could capture characteristics of diseases appropriately. The definition of disease as causal could be also very friendly to clinicians since it is similar to their understanding of disease in practice. Moreover, it could include richer information about causal relationships in disease than other disease ontologies or medical terminologies such as SNOMED-CT. Currently we are refining the ontology through reviewing definitions of disease concepts. We are also organizing definitions of clinical disorders into an abnormality ontology based on YAMATO. After these refinement processes, the ontology could become more systematized knowledge. Other future works includes development of a web service for browsing causal chains and publishing the disease ontology as Linked Open Data.

The demonstration is available at the URL: http://www.hozo.jp/demo/

Acknowledgement

References

1. Scheuermann, R. H., Ceusters, W., and Smith, B. (2009) *Toward an Ontological Treatment of Disease and Diagnosis. P*roc. of the 2009 AMIA Summit on Translational Bioinformatics, 116-120, San Francisco, CA.
2. Osborne, J. D., et al. (2009) *Annotating the human genome with Disease Ontology.* BMC Genomics 10(1):S6.
3. Cowell, L. G. and Smith, B (2010) *Infectious Disease Ontology.* Infectious Disease Informatics, Chapter 19, Sintchenko V., 373-395.
4. Mizoguchi, R., et al. (2011) *River Flow Model of Diseases,* Proc. of ICBO2011, 63-70, Buffalo, USA.
5. Kozaki, K., et al. (2012) *Identity Tracking of a Disease as a Causal Chain,* Proc. of ICBO2012