Center for Cognitive Science

University at Buffalo, State University of New York

Wednesday, October 19, 2005

280 Park Hall
North Campus
2:00 pm –4:00 pm

Werner Ceusters, M.D.

European Centre for Ontological Research
Saarland University, Saarbrücken, Germany
(Sponsored by the Department of Philosophy, UB)

"Computational Linguistics for Referent Tracking in Electronic Healthcare Records: a research agenda"

In the last two decades we have witnessed considerable efforts directed towards making electronic healthcare records (EHRs) comparable and interoperable through advances in record architectures and (bio)medical terminologies and coding systems. Deep semantic issues in general, and ontology in particular, have received some interest from the research communities. However, with the exception of work on so-called 'controlled vocabularies', ontology has thus far played little role in work on standardization. The prime focus has been rather the rapid population of terminologies at the level of fine detail with the purpose of making them available for "clinical coding". The rationale for clinical coding is the belief that codes will make it possible to associate precise meanings to the terms used in expressing patient data in a way that can be interpreted by software for further processing for purposes such as statistic analysis, billing, reimbursement, automated decision support, and triggering alerts. However, many aberrations in the biomedical coding and classification systems and terminologies that are used today, prevent such further processing to be done in a reliable way. This is because the terms or codes contained in the latter are used simply as an alternative to what would otherwise have been registered by means of general terms in natural language. By picking a code from such a system and then registering that code in an EHR, one refers generically to some instance of the class represented by the code. It is still left at best only partially specified which particular instance ("referent") is intended in concrete reality.

It is here that referent tracking comes into play. The goal of referent tracking is to create an ever-growing pool of data relating to concrete entities in reality. In the context of Electronic Healthcare Records (EHRs) the relevant concrete entities are not only particular patients but also their body parts, diseases, therapies, lesions, and so forth insofar as these are relevant to their diagnosis and treatment. Within a referent tracking system, all such entities are referred to explicitly, something which cannot be achieved when only the codes from concept-based systems are used.

A particular challenge for the referent tracking paradigm in the context of EHRs is to minimize the amount of overhead that users would experience when entering data in the records. Ideally, they should be able to continue to work in the same way as before, either by writing clinical narrative or by working with biomedical terminologies, while software running in the background should replace generic codes with entity identifiers where applicable. This can be viewed as a modified version of referent resolution which is known to be a very hard problem in computational linguistics.

Werner Ceusters studied medicine ('77-'84) neuropsychiatry ('84-90'), informatics ('88-'90) and knowledge engineering ('91-'93). He started a series of international research projects in medical natural language processing under the Third, Fourth and Fifth Research Frameworks of the European Commission through his R&D company Office Line Engineering nv. Since then, he has also been active in standardisation bodies related to medical terminology such as CEN/TC251/WG2 and ISO/TC215/WG3. In April 1998, he started a new company - Language & Computing nv (L&C) - to exploit the results of his research. He left L&C in 2004 and created together with Prof. Barry Smith the European Centre for Ontological Research, his main interest being now applying and testing a new theoretically-grounded approach to ontological engineering.

1. Ceusters W. and Smith B. *Referent Tracking in Electronic Healthcare Records.* (Download draft). Accepted for MIE 2005, Geneva, 28-31 Augustus 2005.
2. Ceusters W, Smith B. *Strategies for Referent Tracking in Electronic Health Records.* (Download draft). Proceedings of IMIA WG6 Conference on "Ontology and Biomedical Informatics". Rome, Italy, 29 April - 2 May 2005. (in press).
3. Smith B, Ceusters W. *An Ontology-Based Methodology for the Migration of Biomedical Terminologies to Electronic Health Records.* (Download draft) Accepted for AMIA 2005, October 22-26, Washington DC.

Refreshments will be available; Open to the public!