A Semantic Web Pragmatic Approach to Develop Clinical Ontologies, and Thus Semantic Interoperability, Based in HL7 v2.XML Messaging

David Mendes and Irene Rodrigues

Universidade de Évora

Abstract. The ISO/HL7 27931:2009 standard intends to establish a global interoperability framework for Healthcare applications. However, being a messaging related protocol, it lacks a semantic foundation for interoperability at a machine treatable level has intended through the Semantic Web. There is no alignment between the HL7 V2.xml message payloads and a meaning service like a suitable ontology. Careful application of Semantic Web tools and concepts can ease extremely the path to the fundamental concept of Shared Semantics. In this paper the Semantic Web and Artificial Intelligence tools and techniques that allow aligned ontology population are presented and their applicability discussed. We present the coverage of HL7 RIM inadequacy for ontology mapping and how to circumvent it, NLP techniques for semi automated ontology population and discuss the current trends about knowledge representation and reasoning that concur to the proposed achievement.

1 Introduction

A pragmatic approach is presented in order to identify the different issues faced and for each one of them we discuss the possible and feasible solutions according to the State-of-the-Art in the Semantic Web and Artificial Intelligence science fields. Paramount interest arrived due to the very recent acknowledgment of the clinical practice encoding communities about the possibilities of redirecting efforts to capture the "meaning of data" instead of coding directed to a particular purpose like reimbursement or government funding and reporting as introduced by Cimino in [20].

Although the most significant amount of work in ontology enrichment and population has been done in the Biomedicine research area as illustrated by [2], taking into account the considerations introduced in [1,5] and more recently illuminated by the developments in technology and tooling as referred in [2] we introduce here the proposal of taking advantage of standardization of messaging in EHR[1] to develop the tooling to finally evolve into "evidence based harmonization" in ontology development meant mainly for clinical practice. The completeness and full coverage of ISO/HL7[2] 27931:2009 Standard will allow solutions that do not fall short in particular fields of the different medical specialties. To accomplish a successful work the resulting ontologies have to achieve the sort of user-friendliness, reliability, cost-effectiveness, and breadth of coverage that is necessary to ensure extensive usage as introduced by Smith in [11].

[1] Electronic Health Record.

[2] Health Level 7.

M.M. Cruz-Cunha et al. (Eds.): CENTERIS 2011, Part III, CCIS 221, pp. 297–306, 2011.

Several factors have to be judiciously handled using all the latest trends in technological and scientific development, among these are the proper selection of what ontologies shall be used for learning/enrichment and all the pragmatic aspects that may render broad usage of the resulting automatically produced knowledge. For all of these we suggest what we feel are the most promising, or already proved on the field, techniques that will lead us to the above explained desiderata.

2 Ontology Population in Health

The amount of Clinical data digitally preserved in EHRs is colossal, ever increasing and numerous problems have to be devised and solved as reviewed by Meystre et al. [1] and Liu et al. [13]. Most of the clinical data is in text form coming either from typing entry, transcription from dictation or from speech recognition applications. Accurate coding is necessary for comparability, auditability and last but not least important, accountability. We will figure out a "picture of Healthcare provisioning" through clear identification of the meaning of the available data and not only by the capability of cataloging and codifying that huge amount of data.

2.1 From Clinical Text Information Extraction to Ontology Population

Ontology population/enrichment is performed through Information Extraction from the clinical texts embedded in the messages. IE[3] is a specialized sub-domain of NLP[4] that returns pieces of information from text analysis, unlike IR[5] that returns documents. Facts extracted from documents must refer to a common agreed upon meaning as expressed in some ontology to function as a knowledge enhancement tool. As illustrated in the review by Meystre et al. [1] complemented by the review in [13] many IE methodologies are already thoroughly presented and discussed, and all those considerations shall be taken into proper account in the present work. Aligning the extracted information in form of Clinical Concepts and its relationships in Clinical Practice directed ontologies involves classification to some specific ontology (or a network of them) using several NLP techniques. These tasks form a pipeline of NER[6] , WSD[7] [21] , CRR[8] [22,23,24] , DR[9] , EAV[10] [25], and finally clinical concept matching being these concepts the 'cognitive constructs' introduced by Cornet et al. in [26]. The ontology to be improved will then be refined and developed from some first ontologies in the Biomedical domain. For this purpose different valuable approaches from the symbolic, statistic and hybrid approaches reviewed by Liu et al. in [13] will be discussed ahead in Section 4 and we present here the problems involved in tagging the information so that it will be usable for the ontology enrichment.

[3] Information Extraction.

[4] Natural Language Processing.

[5] Information Retrieval.

[6] Named Entity Recognition.

[7] Word Sense Disambiguation.

[8] Co-Reference Resolution.

[9] Discourse Reasoning.

[10] Extraction of Attributes and Values.

2.2 Named Entities Disambiguation (NED)

GSO[11] will provide us with the controlled vocabularies that can unambiguously co-relate the found term with its due meaning and simultaneously aligning in the direction of a suitable CRR contributing to the desired Concept Acquisition. The fine selection of the GSO will entail the quality of the approval/rejection option for every singular case.

2.3 The Corpora and Its Size Relevancy

The size of the corpora itself is an open issue. Several recent papers[13] question the value of using an over-sized amount of text. In our particular case that of clinical notes resident in HL7 messages. We believe that the proper dimensionality of the corpora will be self adjusted by the factor of rejection attained in the pre-processing operations of our proposal. That is, if during the spell checking, document structure harmonization, tokenization, de-identification, term pruning, word sense disambiguation, named entities disambiguation and semantic concept choice, for instance, tasks no "high valued disambiguation" is achieved then that particular case will get into the rejected corpora and so the refined corpora will only have those items that provide real learning potential.

2.4 Semantic Similarity and/or Patient and Clinical Distance between Cases

Semantic distance is based on weighted path length between concepts. A particular application here is to classify the proximity between our refined corpora messages for the purposes of clustering, indexing and context insertion for classification. In the general case of using semantic methods for text analysis there are some generally available, proved and used on the field. They vary mainly around two different approaches based in linear algebra or probabilistic modeling like the Principal Component Analysis -PCA [9], Vector Space Model -VSM [10], Latent Semantic Analysis -LSA [7], Probabilistic LSA -PLSA [6] and Latent Dirichlet Allocation -LDA [8]. A distinguishable characteristic of our sub-domain of interest resides in the complexity of finding semantic similarity between two terms hence we propose the use of the method built upon SNOMED CT presented by Batet et al. in 2010 [15] which essentially provides independence from the semantic similarity search and the underlying working methods also carefully reviewed in the referred work. For use in a coherent strategy of developing our evidence based population the major concern is not about which method is more appropriate but to develop an interface for our "ontology aligned population" that every chosen method shall adhere to.

2.5 De-identification Issues

This is an extremely important duty because all clinical data has to be cleansed of the possibility of re-identifying in many of the purposes that may be of interest in our work. In the U.S. de-identification itself is due to be in accordance to a specific standard, namely the so-called "Safe Harbor" by the HIPAA[12] that implies the proper

[11] Gold Standard Ontologies.

[12] Health Insurance Portability and Accountability Act.

anonimization of 18 patient identifiers including names, all geographical subdivisions smaller than a state, all elements of dates related to the individual, identifying numbers like phone, fax, social security, medical record, health plan , accounts, certificate or license, vehicle identification, device identification or serial numbers, e-mail addresses, URLs, IP Addresses, Biometric Identifiers, full face photographs and any other uniquely identifying numbers or codes.

Two possibilities can be of concern, whether we are directing our pre-processing labors to populate aggregate ontology information and then it seems adequate to have the kind of care suggested by the US Government and similar identifying removal practices must be enforced or our work is directed to other useful endeavors like EHR enrichment through automated reasoning and decision support aids in the clinical ground and then the identity must be removed but the record tagged for follow up purposes. For instance to correlate diagnostic findings to exams and to therapy applied later.

3 Automated Ontology Population

IE typically requires some "pre-processing" such as spell checking, document structure analysis, sentence splitting, tokenization, WSD, part-of-speech tagging, and some form of parsing namely for identification of strings representing quantities or abbreviations, as in laboratory results for instance. The telegraphical form that is common among clinicians also poses some constraints to the usual NLP techniques used in other fields. Contextual features like negation, temporality, and event subject identification are crucial for accurate interpretation of the extracted information, most work however has been developed so far, as presented by Demner-Fushman et al. in [12].

Our full automation proposal includes two harmonizing steps with the currently available techniques and services that qualify for the considered mission. The first is using the above picked harmonization GSO to provide alignment. The second step is to use the available CORE[13] subset of the UMLS[14] Methathesaurus to further simplify and certify our terminology. It is possible in a loosely way to query data remotely via Web Services using the API available in the UTS[15] , a service of the U.S. NLM[16] , to validate against the referred CORE Problem List and Route of Administration Subsets of SNOMED CT[17] . The software needed to accomplish this, as all of the work presented here, has a Loosely Coupled Architecture based in Web Services and certified, auditable messaging as enforced in the ISO/HL7 standard.

4 Clinical Practice Ontology Population vs. General Ontologies

Relations specifically associated to Biomedicine or Clinical Practice retain knowledge associated with the clinical domain. Apart from relations such as is-a and part-of, biomedical ontologies also contain domain specific relations such as has-location,

[13] Clinical Observations Recording and Encoding.

[14] Unified Medical Language System.

[15] UMLS Terminology Services.

[16] National Library of Medicine.

[17] Systematized Nomenclature of Medicine - Clinical Terms.

has-manifestation or clinically-associated-with. These relations are, however, nothing but that. That is, relations. And this turns them semantically transparent, no specific domain knowledge differentiates these relations from any other given the appropriate definition (cardinality, direction, object, datatype and annotation properties) which for proper computability purposes can be achieved with the adequate OWL DL[18] representation. Currently several tools exist for bi-directional converting which can automatically transform OBO[19] ontologies into the OWL-based format used by the Semantic Web namely OWL DL [3]. The problem of defining what are the ontologies that should be considered as adequate for proper enrichment will be discussed ahead in Section 4.1.

Being standardized in 2009 the language of choice, and consequently the associated tooling, is OWL2. OWL2 addresses key expressive and computational limitations of OWL. By adding new constructs to the language, OWL2 more directly supports medical applications. For example, so called "role chains" allow ontologists to express the connection between spatial relations and part-whole relations, e.g., if a fracture is located on a bone which is part of a leg, that fracture is a fracture of that leg.

4.1 Adequate Ontologies for Harmonization

The formation of the possible list of Ontologies shall take in consideration the steps suggested in the Ontology Engineering area with the developments and tools introduced in recent years. What is a 'good' ontology to use as a GSO for "evidence based harmonization"?

Items to be evaluated are usage, application performance, data coverage, corpus fit and reasoning adequacy for instance, with quality criteria as accuracy, adaptability, clarity, completeness, computational efficiency, conciseness, consistency and organizational fitness. Tools and methodologies that perform this categorization like OntoClean [27] are available. Ontologies are to be gathered in a Clustered Network and it seems advisable to use foundational Ontologies covering: Anatomy like the FMA[20] , the foundries from OBO like Biological Process[21] , Adverse Event Reporting[22] , Human disease[23], Infectious Disease[24] , Symptom[25] and time ontologies like DAML or SUMO for instance. The NeOn[26] toolkit is the reference implementation of the NeOn architecture that entails support for ontology engineering and management, complete ontology lifecycle, different ontology languages (OWL2 or F-Logic[19]) and support for networked ontologies (modules, mappings). It fits naturally in a Java enterprise environment with extensions through plugins and Web Services. Manipulating all the proposed eco-system through Web Service interfacing is the suggested architecture.

[18] Web Ontology Language Description Logic.

[19] Open Biological and Biomedical Ontologies - http://www.obofoundry.org/

[20] Foundational Model of Anatomy - http://sig.biostr.washington.edu/projects/fm/

[21] http://www.obofoundry.org/cgi-bin/detail.cgi?id=biological process

[22] http://www.obofoundry.org/cgi-bin/detail.cgi?id=AERO

[23] http://www.obofoundry.org/cgi-bin/detail.cgi?id=disease ontology

[24] http://www.obofoundry.org/cgi-bin/detail.cgi?id=infectious disease ontology

[25] http://www.obofoundry.org/cgi-bin/detail.cgi?id=gemina symptom

[26] http://www.neonproject.org/nw/Welcome to the NeOn Project.

For enrichment. The same considerations introduced above for the selection of the collection of ontologies to use as GSO for the pre-processing and harmonizing proceedings may be used to pick the Ontologies that are to be learned/enriched. Naturally the first possible subjects for automated enrichment are some of the OBO foundries themselves like the Ontology for General Medical Science[27] or Ontology of Medically Related Social Entities[28] just to mention two evident candidates. We should be bold enough, however, to ascertain that the pinnacle of the possibilities of the current proposal shall be the capability of gathering a "virtual picture" of the clinicians activity. That is, a photograph of a MD activity, the evaluable and comparable performance of a Service, a Hospital or a Health System at large for instance.

4.2 Shared Semantics and Ontology Harmonization through Modeling around HL7, Its Intentions and Its Flaws

In the 2009 edition of the HL7[29] Version 3 complete suite of specifications some salient features have been focused and the most important as what relates to this work are: (1) A focus on semantic interoperability by specifying that information be presented in a complete clinical context that assures that the sending and receiving systems share the meaning (semantics) of the information being exchanged; (2) Model-based specifications that provide consistent representation of data laterally across the various HL7 domains of interest and longitudinally over time as new requirements arise and new fields of clinical endeavor are addressed. This has proved to be the most far sighted motivation particularity as it enabled the interaction and harmonization within BRIDG[30] [17]; (3) Technology-neutral standards that allow HL7 and the implementers of HL7 standards to take advantage, at any point in time, of the latest and most effective implementation technologies available like the latest trend in developing loosely coupled architectures for integration based in SOA[31] ; (4) A development methodology and metamodel that assures consistent development and the ability to store and manipulate the specifications in robust data repositories rather than as word-processing documents.

A significant amount of problems still are fattening the above bill of fair intentions, mainly in its application to reality:

4.3 HL7 a ill defined Standard?

HL7 is adopted by Oracle as basis for its Electronic Health Record technology; supported by IBM, GE and most major vendors and users like the US DoD VA. In HL7 V2 the realization of the messaging task allows ad hoc interpretations of the standard by each sending or receiving institution. Then vendor products never properly interoperate, and always require mapping software. The solution to this problem is the HL7 RIM or Reference Information Model that was touted as a world standard for exchange of information between clinical information systems. The V3 solution was to

[27] http://www.obofoundry.org/cgi-bin/detail.cgi?id=OGMS

[28] http://www.obofoundry.org/cgi-bin/detail.cgi?id=omrse

[29] www.hl7.org

[30] http://www.bridgmodel.org/

[31] Service Oriented Architecture.

remove optionality by having the RIM serve as a master model of all health information, from blood banks to Electronic Health Records to clinical genomics [18] and to be the standard of choice for countries and their initiatives to create national EHR and EHR data exchange standards.

Yet, despite the claim of being "credible, clear, comprehensive, concise, and consistent", as well as "universally applicable" and "extremely stable", the huge efforts themselves undermined several problems that surfaced through the development of the practicalities of implementations.

Questions arose mainly regarding documentation (1) that is divided into 7,573 files, subject to frequent revisions and very difficult to understand marked by sloppy and unexplained use of terms such as 'act', 'Act', 'Acts', 'action', 'ActClass' 'Act-instance', 'Act-object'; scope (2) since the class structure is built upon only two main classes Act and Entity basic categories cannot be agreed upon for common phenomena because the inheritance from the upper classes can be discussed upon. In RIM there is no distinction between an activity and its documentation, an Act is the document about an Act that is, by definition, an intentional action (!) and finally; implementation problems (3) since it had difficulties growing to embrace the technological developments occurred since it was adopted as early as 1997.

4.4 Clinical Notes Acquisition from V2.XML Protocol

As early as 2003, the American National Standards Institute approved the HL7 Version 2 XML Encoding Syntax informally known as HL7 V2.xml. HL7's Version 2.xml messaging standard is the workhorse of electronic data exchange in the clinical domain and arguably the most widely implemented standard for Healthcare in the world. The V2.xml defines the Extensible Markup Language (XML) encoding rules for traditional HL7 Version 2 message content. There have been seven releases of the Version 2.x Standard to date. HL7 Version 2 was also recently selected by the U.S Office of the National Coordinator for Health Information Technology as part of its initial set of standards, implementation specifications and certification criteria for EHR technology. Version 2.5 was also published as an international standard by ISO in June 2009 as the ISO/HL7 27931:2009 standard that is the subject of this works proposal as the departure point for knowledge acquisition. Acquisition from clinical notes is possible now by using the Web Services exposed in EHRs or flowing through Hubs like Mirth[32] or "traveling" in the different Regional, National or Supra-national HIE[33] networks currently under strong global dissemination.

5 Ontology Learning and Enrichment

The Biomedical Research Integrated Domain Group (BRIDG) Model is a collaborative effort engaging stakeholders from the CDISC[34] , the HL7 RCRIM TC[35], the NCI[36] and

[32] http://www.mirthcorp.com/

[33] Heath Information Exchange.

[34] Clinical Data Interchange Standards Consortium.

[35] Regulated Clinical Research Information Management Technical Committee.

[36] National Cancer Institute.

its caBIG®[37] , and the US FDA[38] . The BRIDG model is an instance of a DAM[39]. The goal of the BRIDG Model is to produce a shared view of the dynamic and static semantics for the domain of protocol-driven research and its associated regulatory artifacts. The BRIDG Model represents biomedical/clinical research. It was developed to provide an overarching model that could readily be understood by domain experts and would provide the basis for harmonization among standards within the clinical research domain and between biomedical/clinical research and Healthcare.

A DAM is a conceptual model used to depict the behavioral and static semantics of a domain of interest. A domain analysis model is used as reference material in development of information system interoperability specifications as well as design specifications of information system components. The preferred language for expression of a domain analysis model is UML[40] . A shared view of the various data structures and processes that define the BRIDG Model's domain-of-interest is essential in achieving the larger goal of semantic interoperability (SI) namely between systems (computable semantic interoperability (CSI)). Through the explicit definitions of shared semantics CSI is possible both within the BRIDG domain of-interest and between the BRIDG domain and other 'intersecting' domains (e.g. Public Health, Healthcare, etc.).

The goal of defining and representing the shared semantics (aka "meaning") of the BRIDG Model's domain-of-interest is achieved through the gathering and documenting the various business processes (dynamic semantics), data structures (static semantics), and relationships (static and dynamic semantics) that collectively are required to support CSI. The first formal release of BRIDG was published in June 2007. The BRIDG Model does bear a certain resemblance to the HL7 RIM. However, the overarching goal of the BRIDG Model is to represent domain-specific semantics in an implementation-independent fashion that is understandable to domain experts. This will deal with the problems illustrated in 2006 in [18] and is currently well addressed by the current 3.0.3 model version.

For our work to be fully contained we suggest an expansion of the work about Categorial Structure introduced in [14] into the ISO/HL7 27931:2009 but with the concrete OWL DL representation extracted from the BRIDG 3.0.3 Model.

6 Conclusion

We try to illustrate the possibility of taking advantage of the recent standardization and harmonization efforts and investigation in the related fields to seriously improve the capacity of ontology enrichment by automating the knowledge acquisition in the Health domain. Our proposal is based at one point in using the contents of the XML messages normalized to achieve data interoperability among health information systems and in the other end we suggest the use of the shared semantics model, fundamental to achieve broad acceptance and usage of the developed/enriched ontologies, recently developed

[37] Cancer Biomedical Informatics Grid.

[38] Food and Drug Administration.

[39] Domain Analysis Model.

[40] Unifed Modeling Language.

by BRIDG. With these two focal points in mind we present and discuss which particularities are the more steep to handle and the recent contributions to their pragmatic resolution for the specific work in the knowledge sub-domain of Healthcare.

References

1. Meystre, S.M., Savova, G.K., Kipper-Schuler, K.C., Hurdle, J.F.: Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research (2008)
2. Smith, B., Brochhausen, M.: Establishing and Harmonizing Ontologies in an Interdisciplinary Health Care and Clinical Research Environment (2008)
3. Obo-Owl RESTful Conversion API, `http://www.berkeleybop.org/obo-conv.cgi`
4. Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, J.-F., HuaData, L.: Mining in Healthcare and Biomedicine: A Survey of the Literature. Journal of Medical Systems (2010)
5. Spasic, I., Ananiadou, S., McNaught, J., Kumar, A.: Text mining and ontologies in Biomedicine: Making sense of raw text. Brief Bioinform. 6(3), 239–251 (2005)
6. Hofmann, T.: Probabilistic latent semantic analysis. In: Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (1999)
7. Deerwester, S., Dumais, S., Landauer, T., Furnas, G., Harshman, R.: Indexing by latent semantic analysis. Journal of the American Society of Information Science 41(6), 391–407 (1990)
8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet Allocation. Journal of Machine Learning, Research 3, 993–1022 (2003)
9. Jollife, I.T.: Principal component analysis. In: Everitt, B.S., Howell, D.C. (eds.) Encyclopedia of Statistics in Behavioral Science, pp. 1580–1584. John Wiley and Sons Ltd., New York (2005)
10. Salton, G., Wong, A., Yang, C.S.: A Vector Space Model for Automatic Indexing. Communications of the ACM 18(11), 613–620 (1975)
11. Smith, B., Brochhausen, M.: Putting biomedical ontologies to work. Methods of Information in Medicine 49(2), 135–140 (2010), doi:10.3414/ME9302
12. Demner-Fushman, D., Mork, J.G., Shooshan, S.E., Aronson, A.R.: UMLS content views appropriate for NLP processing of the biomedical literature vs. clinical text. Journal of Biomedical Informatics 43, 587–594 (2010), doi:10.1016/j.jbi.2010.02.005
13. Liu, K., Hogan, W.R., Crowley, R.S.: Natural Language Processing methods and systems for biomedical ontology learning. Journal of Biomedical Informatics 44, 163–179 (2011), doi:10.1016/j.jbi.2010.07.006
14. Rodrigues, J.M., Kumar, A., Bousquet, C.: Using the CEN / ISO Standard for Categorial Structure to Harmonize the Development of WHO International Terminologies. Medical Informatics (Icd), 255–260 (2009), doi:10.3233/978-1-60750-044-5-255
15. Batet, M., Sanchez, D., Valls, A.: An ontology-based measure to compute semantic similarity in biomedicine. Journal of Biomedical Informatics 44, 118–125 (2010), doi:10.1016/j.jbi.2010.09.002
16. HL7 Health Level Seven ® International, `http://www.hl7.org`
17. The Biomedical Research Integrated Domain Group, `http://www.bridgmodel.org`
18. Smith, B., Ceusters, W.: HL7 RIM: An Incoherent Standard. Medical Informatics, 133–138 (August 2006)

19. Kifer, M., Lausen, G., Wu, J.: Logical foundations of object-oriented and frame based languages. Journal of the ACM 42(4), 741–843 (1995), doi:10.1145/210332.210335
20. Cimino, J.J.: High-quality, Standard, Controlled Healthcare Terminologies Come of Age. Methods of Information in Medicine 50(2), 101–104 (2011), retrieved `http://www.ncbi.nlm.nih.gov/pubmed/21416108`
21. Navigli, R., Velardi, P.: Structural semantic interconnections: a knowledge-based approach to word sense disambiguation. IEEE Trans. Pattern Anal. Mach. Intel (PAMI) 27, 1075–1086 (2005)
22. Poesio, M., Vieira, R., Teufel, S.: Resolving bridging references in unrestricted text. In: Proceedings of the ACL Workshop on Operational Factors in Robust Anaphora Resolution, pp. 1–6 (1997)
23. Soon, W.M., Ng, H.T., Lim, D.C.Y.: A machine learning approach to co-reference resolution of noun phrases. Comput. Linguist. 27, 521–544 (2001)
24. Ng, V., Cardie, C.: Improving machine learning approaches to co-reference resolution. In: Proceedings of the 40th Annual Meeting of the ACL. ACL, Philadelphia (2001)
25. Friedman, C., Borlawsky, T., Shagina, L., Xing, H.R., Lussier, Y.A.: Bio-ontology and text: bridging the modeling gap. Bioinformatics 22, 2421–2429 (2006)
26. Cornet, R., De Keizer, N.F., Abu-Hanna, A.: A framework for characterizing terminological systems. Methods Inf. Med. 45, 253–266 (2006)
27. Guarino, N., Welty, C.: Identity, Unity, and Individuality: Towards a formal toolkit for ontological analysis. In: Horn, W. (ed.) Proceedings of ECAI 2000, pp. 219–223. IOS Press, Berlin (2000)