Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Technology 00 (2013) 000–000

Procedia Technology

www.elsevier.com/locate/procedia

CENTERIS 2013 - Conference on ENTERprise Information Systems / HCIST 2013 - International Conference on Health and Social Care Information Systems and Technologies

Ontology based clinical practice justification in natural language

David Mendes[a]*, Irene Pimenta Rodrigues[a], Carlos Baeta[b]

[a]*First affiliation, Address, City and Postcode, Country*
[b]*Second affiliation, Address, City and Postcode, Country*

Abstract

One of the most important contributions that any decision support system can make to achieve wide acceptance among any community is to be able to justify its own suggestions. When dealing with highly technical and scientifically advanced practitioners like medical doctors or any other related clinical workers, the ability to justify itself using the domain specialist usual terminology and technicalities is imperative. In this article we demonstrate the use of an ontological framework as inferencing basis for automatic sound clinical suggestions providing. Our work has two main contributions, consolidating the use of OGCP (Ontology for General Clinical Practice) as foundation and providing controlled English justifications of the extracted suggestions. We found that clinical practitioners feel as acceptable the Attempto Controlled English justifications generated from the knowledge base.

Keywords: OGCP; OGMS; CPR; NLP; Clinical Concepts Extraction; Ontological Realism; Natural Language Justification; ACE

* Corresponding author. Tel.: +351 96 3564161.
E-mail address: dmendes@uevora.pt.

1. Introduction

The development of AI[†] tools like CDSS[‡] have to rely upon strong KR[§] techniques that are currently pro-eminent in the research community. In the Biomedical domain at large and specifically in the healthcare sub-domain major contributions have surfaced recently in the area of ontological representation of healthcare providing [1] and medicine in general [2].

We are now using this ground for automatic reasoning and provide adequate justifications in natural/technical language to logically inferred conclusions. These justifications aim at getting good acceptance by the clinicians.

We started our research trying to figure out how to create a reasoning framework that could provide a picture of the usual practice of any MD. This was an attempt to deliver CDSS systems based in Semantic Web technologies and state of the art automatic reasoning tools. We went to explore HL7 Messaging as information source for our KR efforts [3] where we present a detailed state of the art. After thorough evaluation of ontologies state-of-the-art in the healthcare domain we started our exploration of ***CPR***[**] [4] and its enrichment from the widest available semi-structured corpora that are text reports [5]. We felt, however, that the structure was feeble to support some theoretical foundations of clinical thinking and we thought of adopting the, yet novel, but deep rooted in the philosophical and medical community ***OGMS***[††] [6] using ***DO***[‡‡] [7] for the concept linkage. With a generated knowledge base available we try to figure out some inferred conclusions and present them in ***ACE***[§§] technical/natural jargon that was found acceptable as justification by clinical peers.

2. Proposed system

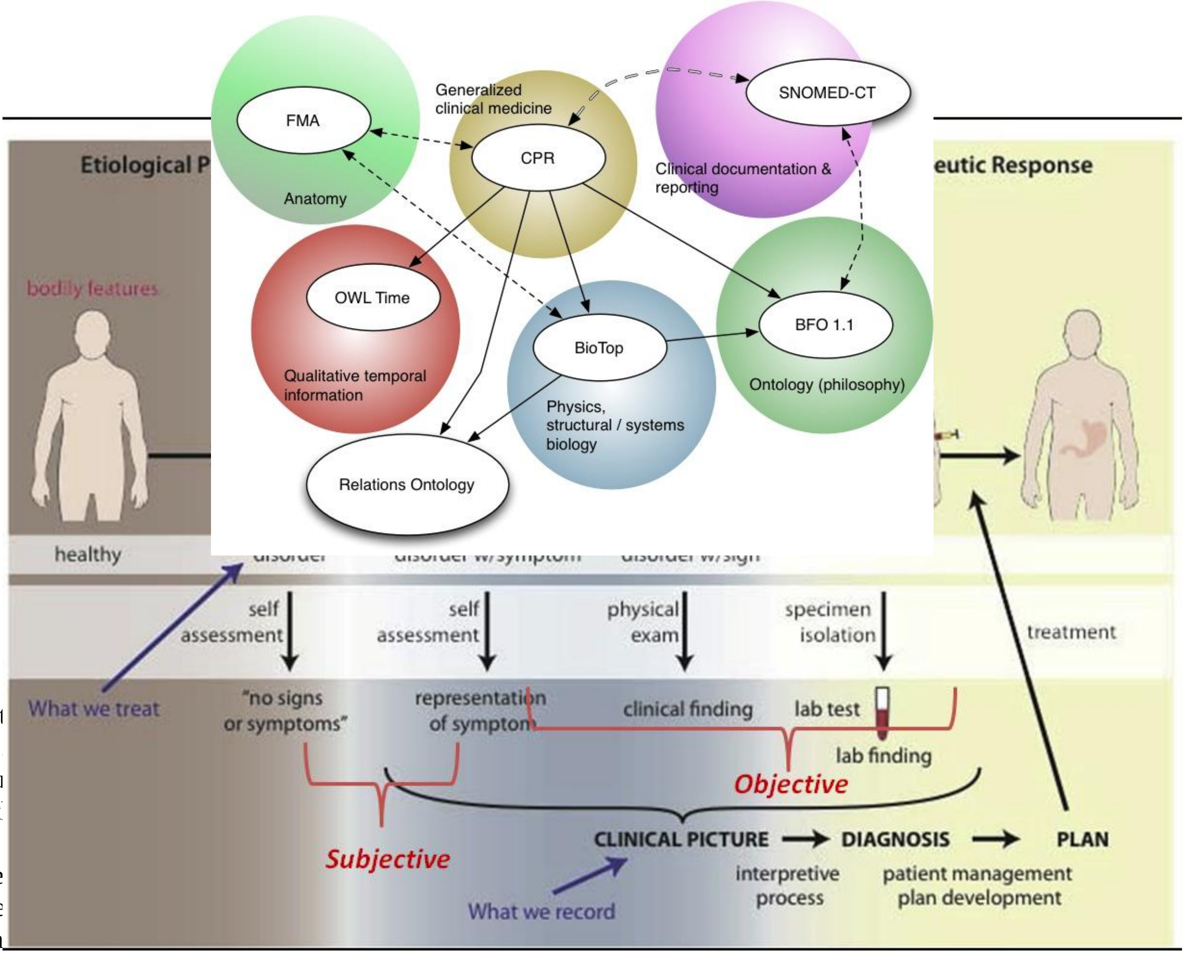
We follow the philosophic approach of **Ontological Realism** [8, 9] to extend the **OGMS** with **CPR** and **DO** and its foundational ontologies as shown in Fig. 1 into the **OGCP** [10].

With **OGCP** in place we populate into a Clinical Practice KB[***] as introduced in our previous work [3] thus rendering the framework for QA[†††] in the represented domain.

---

† Artificial Intelligence
‡ Clinical Decision Support System
§ Knowledge Representation
** Computer based Patient Record Ontology
†† Ontology for General Medical Science
‡‡ Disease Ontology
§§ Attempto Controlled English
*** Knowledge Base
††† Question Answering

init
5,
reu
ort
of
ide
Pie
ma



We made an effort of trimming and pruning of the ***OGMS*** and ***CPR*** complementing in accordance to our team of cardiologists to better accommodate their needs expressed in the reports we sampled. That included some "gardening" to include: ***SO*** the Symptom Ontology, ***VSO*** the Vital Signs Ontology and others all of them accord to OBO Foundry principles.

In order to align the clinical concepts in the various ontologies present, an effort was needed to amalgamate them according to a sound theory of disease and that's why we incorporate the ***DO*** that was expressly built with this purpose in mind [13].

The Disease Ontology is a community driven, open source ontology that is designed to link disparate datasets through disease concepts. It's provided a computable structure of inheritable, environmental and infectious origins of human disease to facilitate the connection of genetic data, clinical data, and symptoms through the lens of human disease [14]. The DO semantically integrates disease and medical vocabularies through extensive cross mapping and integration of MeSH, ICD, NCI's thesaurus, SNOMED CT and OMIM [15] disease-specific terms and identifiers. It represents a comprehensive knowledge base of 8043 inherited, developmental and acquired human diseases.

3. Ontological Framework

The rendered ontology framework reveals it's soundness for supporting the previously named concept of "clinical thinking" as pictured in [16].

We developed a simple pragmatic approach to the representation of disease and diagnostic as illustrated in the referred article by Scheuermann.

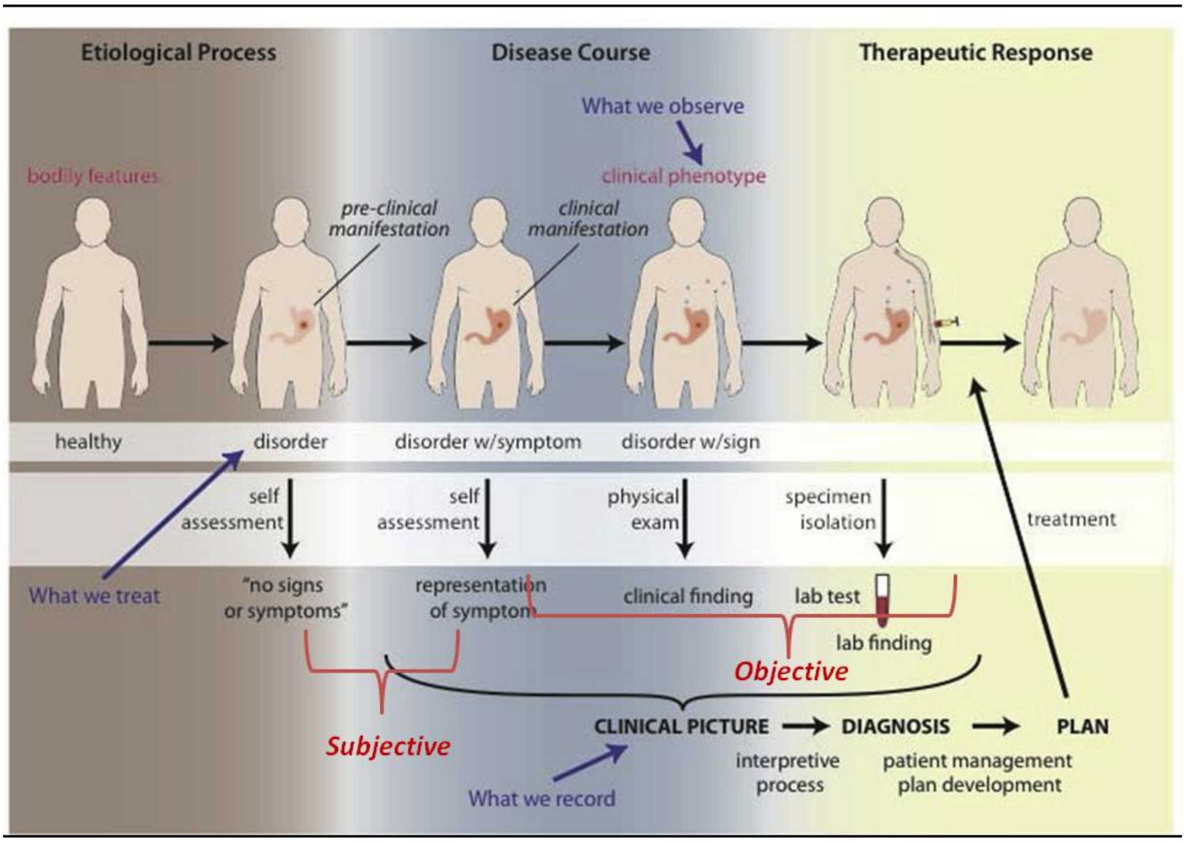


Fig. 2
SOAP Points Insertion

The text for any particular encounter (actually for any Clinical Episode) may be collected in the form suitable for processing into the Ontology framework using some **NLP pragmatics**. Populating the ***OGCP*** the

``Clinical Picture" is completed and thus our KB is available for validation and further logical inferencing. The semantic representation is done using pragmatic interpretation as defined in our fellow researcher at CENTRIA[‡‡‡] Dora Melo's article [17].

The enrichment process must always maintain the entailments provided by the base (gold-standard) ontologies and so can never lead to inconsistency. We use a round-trip, debug and repair, building method to populate/enhance the OGCP then. For any new instance the validation is performed and new possible inferred facts generated if consistency is yet valid. These new facts are candidates for NLP justifications generation. The main objective of the system is to provide accurate answers to questions posed by users and, in our proposal these answers are clinically valid because the generation method guarantees that.

QA is, however, only one of the interesting features of our work that is enhanced by the adequate justification to be evidently useful for practitioners. To develop justifications from DL[§§§] arguments inferred from consequence based reasoners [18,19] we based our work in [20] to study and compare the justificatory structure to those present in the NCBO BioPortal addressed in the mentioned article.

The results so far are in the realm of **'ontology verbalization'**, the generated explanations are still in a **controlled natural language (CNL)** fashion. The obtained results seem to be adequate enough for the users to find them believable and thus the justifications stand in our controlled clinical setting. We use the verbalization tooling [21] to present the justifications in an acceptable manner. The foundational techniques were introduced in [22]. For the verbalization to function properly all the restrictions of content are guaranteed in the process of ontology (Knowledge base) enrichment from SOAP reports. For instance, *all names are English words and individuals are singular proper names (preferably capitalized) named classes are denoted by singular countable nouns and (object) properties by transitive verbs in their lemma form (i.e. infinitive form)* [22]. The decision of what inferred knowledge is then presented with its justifications to the user is a task handled by the DC[****] using the developed pragmatics introduced in the above referred article [17].

4. Conclusion

We are developing a knowledge representation infrastructure enabling the usage of highly optimized distributed consequence based reasoners that are referred in literature only in 2011. With these very recent developments it's finally possible to validate the enormous knowledge bases that are created by automatically populating a proposed ontology ***OGCP*** that relies on extensive, and very solid, foundations like **SNOMED-CT** and **FMA** among others. Logical inferencing and clinical facts entailment that is possible through this capability is an interesting contribution to the application of Artificial Intelligence to healthcare. We introduce clinical decision support systems (CDSS) that are based on such a breakthrough technique. We further argue that it is imperative, for the broad acceptance of these tooling by the medical community, that their inferences are justified using controlled natural language and adequate terminology.

[‡‡‡] Centre for Artificial Intelligence (CENTRIA) is a research centre from the Faculty of Science and Technology of the New University of Lisbon (FCT/UNL)

[§§§] Description Logic

[****] Discourse Controller

Acknowledgements

We acknowledge IIFA - Instituto de Investigação e Formação Avançada of Universidade de Évora for the Bento de Jesus Caraça scholarship that is awarded to the corresponding author and CENTRIA – Center for Research in Artificial Intelligence of Faculdade de Ciências e Tecnologia of Universidade Nova de Lisboa for its continuous financial support of our work.

References

[1] Ogbuji, C. (2011). A framework ontology for computer-based patient record systems. In Proceedings of the ICBO: International Conference on Biomedical Ontology, pages 217-223, Buffalo, NY, USA.
[2] Brochhausen, M., Burgun, A., Ceusters, W., Hasman, A., Leong, T., Musen, M., Oliveira, J., Peleg, M., Rector, A., Schulz, S., et al. (2011). Discussion of "biomedical ontologies: Toward scientific debate". Methods of information in medicine, 50(3), 217.
[3] Mendes, D. and Rodrigues, I. P. (2011). A semantic web pragmatic approach to develop clinical ontologies, and thus semantic interoperability, based in hl7 v2.xml messaging. In HCist 2011 - Proceedings of the International Workshop on Health and Social Care Information Systems and Technologies. Springer-Verlag - book of the CCIS series (Communications in Computer and Information Science).
[4] W3C (2009). Computer-based patient record ontology. Available Online: http://code.google.com/p/cprontology.
[5] Mendes, D. and Rodrigues, I. P. (2012). Advances to semantic interoperability through cpr ontology extracting from soap framework reports. electronic Journal of Health Informatics.
[6] OGMS (2010). Ontology for general medical science. Available Online: http://code.google.com/p/ogms.
[7] DO (2010). Disease ontology. Available Online: http://diseaseontology.sourceforge.net/.
[8] Smith, B., Ceusters, W., Klagges, B., Khler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A. L., and Rosse, C. (2005). Relations in biomedical ontologies. Genome Biol, 6(5), R46.
[9] Smith, B. and Ceusters, W. (2010). Ontological realism: A methodology for coordinated evolution of scientific ontologies. Applied ontology, 5(3-4), 139-188.
[10] David Mendes and Irene Pimenta Rodrigues. (2013). Ontology for general clinical practice. Available Online: http://www.semanticweb.org/david/ontologies/2013/0/OGCP.
[11] Ghazvinian, A., Noy, N., Musen, M., et al. (2011). How orthogonal are the obo foundry ontologies. J Biomed Semantics, 2(Suppl. 2).
[12] BFO (2012). Basic formal ontology 2.0. Available Online: http://code.google.com/p/bfo.
[13] Schriml, L., Arze, C., Nadendla, S., Chang, Y., Mazaitis, M., Felix, V., Feng, G., and Kibbe, W. (2012). Disease ontology: a backbone for disease semantic integration. Nucleic Acids Research, 40(D1), D940{D946.
[14] WIKI-DO (2012). Disease ontology wiki page. Available Online: http://dowiki.nubic.northwestern.edu/do-wiki/index.php/Main Page.
[15] OMIM (2012). Omim, online mendelian inheritance in man. Available Online: http://www.ncbi.nlm.nih.gov/omim.
[16] Scheuermann, R. H., Ceusters, W., and Smith, B. (2009). Toward an ontological treatment of disease and diagnosis. In 2009 AMIA Summit on Translational Bioinformatics, pages 116{120, San Francisco, CA.
[17] Melo, D., Rodrigues, I., and Nogueira, V. (2012). Work out the semantic web search: The cooperative way. Advances in Artificial Intelligence, 2012.
[18] Kazakov, Y. (2009). Consequence-driven reasoning for horn SHIQ ontologies. In C. Boutilier, editor, IJCAI, pages 2040-2045. IJCAI Distinguished Paper Award Winner.
[19] Kazakov, Y., Krötzsch, M., and Simancancik, F. (2011). Concurrent classification of EL ontologies. In Proceedings of the 10th international conference on The semantic web - Volume Part I, ISWC'11, pages 305-320, Berlin, Heidelberg. Springer-Verlag.
[20] Bail, S., Horridge, M., Parsia, B., and Sattler, U. (2011). The justificatory structure of the NCBO Bioportal ontologies. In L. Aroyo, C. Welty, H. Alani, J. Taylor, A. Bernstein, L. Kagal, N. F. Noy, and E. Blömqvist, editors, International Semantic Web Conference (1), volume 7031 of Lecture Notes in Computer Science, page 6782. Springer.
[21] Kaljurand, K. (2010). Owl verbalizer: making machine-readable knowledge also human-readable. Available Online: https://code.google.com/p/owlverbalizer/.
[22] Kaljurand, K. and Fuchs, N. E. (2007). Verbalizing owl in Attempto controlled English. Proceedings of OWLED07.