

Where GO is going and what it means for ontology extension

Catia Pesquita and Francisco M. Couto

Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa
Campo Grande, Lisboa, Portugal

Abstract. Developing and maintaining a biomedical ontology is a time and effort-consuming task, given the dynamic and expanding nature of biomedical knowledge. This is a relevant issue for very large ontologies which cover a broad domain, for smaller ontologies maintained by a small team and also for domains where being able to perform quick updates is critical (e.g. epidemiology).

The first step in the process of extending an ontology is identifying the areas of the ontology that need to be changed - change capturing. In this paper we propose that this process can be semi-automated by exploring ontology information. This would be a valuable tool to support ontology developers in ontology extension, easing their burden.

In order to accomplish this, we have developed a framework for analysing the extension of ontologies, to create a general panorama of ontology extension processes that can guide the development of change capturing techniques. We have applied it to the analysis of the extension of the Gene Ontology and uncovered some of the underlying tendencies in its extension. Building upon the results of this analysis and a set of guidelines for ontology change capturing, we then investigated the feasibility of predicting which classes of the ontology will be extended in a future version.

Finally, we discuss the obtained results and identify the main challenges and future directions for the budding area of ontology extension prediction.

Key words: Biomedical ontologies, ontology extension, ontology refinement, ontology enrichment, prediction of ontology extension

1 Acknowledgements

This work was supported by FCT, through the Multiannual Funding Programme and the PhD grant SFRH/BD/42481/2007. The authors also wish to thank the European Commission for the financial support of the EPIWORK project under the Seventh Framework Programme (Grant #231807)