

Modeling Logical Definitions in Biomedical Ontologies by Reusing Ontology Design Patterns

Mirna El Ghosh¹, Fethi Ghazouani¹, Benjamin Birene¹, Elise Akan¹, Jean Charlet^{1,3} and Ferdinand Dhombres^{1,2}

¹INSERM, Sorbonne Université, Univ. Sorbonne Paris-Nord, LIMICS, Paris, France

²Médecine Sorbonne Université, GRC-26, Service de Médecine Fœtale, AP-HP, Hôpital Armand Trousseau, Paris, France

³AP-HP/DRCI, Paris, France

Abstract

Logical definitions are addressed in biomedical ontologies such as the Human Phenotype Ontology (HPO) to allow cross-species mapping by means of automated semantic reasoning. These definitions aim to associate ontological terms within an ontology to external species-neutral ontological resources. However, applying logical definitions manually in ontologies under development is a challenging issue. Thus, approaches supporting extensible ontology development and pattern-based ontology design are tackled to reuse logical definitions as Ontology Design Patterns (ODPs) and apply them in the context of the SUOG project. ODPs are reusable modeling solutions used to facilitate ontology development.

Keywords

logical definitions, ontology reuse, Ontology Design Patterns, biomedical ontologies, HPO

1. Introduction and Motivation

Logical definitions in biomedical ontologies such as HPO¹ [1] and MPO² [2] allow cross-species mapping, using automated semantic reasoning. Besides, they support quality control [3] and classifications (is-a/subclass relationships inferences). These definitions aim to associate terms within an ontology with terms in external species-neutral ontological resources such as PATO³ an ontology of phenotypic qualities. For example, consider the following logical definition of the HPO term Immunodeficiency (HP:0002721).

```
'has part' some ('decreased rate'  
                and ('inheres in' some 'immune response')  
                and ('has modifier' some 'abnormal'))
```

International Conference on Biomedical Ontologies 2021, September 16–18, 2021, Bozen-Bolzano, Italy

✉ mirna.el-ghosh@inserm.fr (M. E. Ghosh); fethi.ghazouani@inserm.fr (F. Ghazouani);

benjamin.birene@etu.sorbonne-universite.fr (B. Birene); elise.akan@sorbonne-universite.fr (E. Akan);


jean.charlet@sorbonne-universite.fr (J. Charlet); ferdinand.dhombres@inserm.fr (F. Dhombres)

ORCID 0000-0001-6341-3847 (M. E. Ghosh); 0000-0002-0830-1954 (F. Ghazouani); 0000-0000-0000-0000 (B. Birene);

0000-0000-0000-0000 (E. Akan); 0000-0002-7966-9203 (J. Charlet); 0000-0003-3246-8727 (F. Dhombres)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

¹Human Phenotype Ontology, <https://hpo.jax.org/app/>

²Mammalian Phenotype Ontology, <http://www.obofoundry.org/ontology/mp.html>

³Phenotype And Trait Ontology, <https://github.com/pato-ontology/pato/>

Immunodeficiency refers to the *failure of the immune system to protect the body adequately from infection, due to the absence or insufficiency of some component process or substance*. In its logical definition, Immunodeficiency is defined as being equivalent to the intersection of all classes of things that are “a rate which is lower relative to the normal” (decreased rate), “deviate from the normal or average” (abnormal), and inhering in the “immune system” using the term immune response from Gene Ontology. The logical definition uses relations such as *has part*, *inheres in*, and *has modifier* reused from logically well-formed ontologies [4] such as BFO⁴ and RO⁵.

Encoding logical definitions manually in ontologies under development is a challenging task. Thus, approaches supporting extensible ontology development (e.g., MIREOT [5]) and pattern-based ontology engineering approaches (e.g., eXtreme Design (XD) [6]) are tackled. Such approaches bring solutions to reuse and apply generic logical definitions as ontology patterns. While extensible ontology development permits the extraction of ontology subsets for term reuse and semantic alignment, pattern-based approaches aim to model new parts of an ontology using Ontology Design Patterns (ODPs)⁶. ODPs are defined as reusable modeling solutions to frequently occurring ontology design problems [7]. In the biomedical domain, ODPs are encouraging to capture common modeling situations, help facilitate ontology development and avoid common mistakes [8].

In the SUOG (Smart Ultrasound in Obstetrics and Gynecology) project⁷, an ontology-based decision support system for complex ultrasound diagnosis in obstetrics and gynecology is intended [9]. The SUOG ontology, which is under development, distinguishes two main sub-ontologies: (1) *disorders* that describe the *pregnancy state* concepts (e.g., alagille syndrome, limb dysostosis, cerebral midline anomaly, congenital anomaly of truncal valve, etc.) and *prenatal findings* (e.g. abnormal atrial arrangement finding, absent right superior caval vein finding, cerebral arteriovenous malformation finding, etc.), that *suggest* one or more disorders. Aiming to enrich the SUOG ontology semantically, logical definitions are required to define findings and disorders concepts in terms of other more elementary (atomic) concepts [3]. This work proposes combining extensible ontology development and pattern-based approaches to reuse (from HPO) shared and validated logical definitions as ODPs and adapt and apply them in the SUOG ontology. The remainder of this paper is organized as follows. Section 2 outlines the related work. In section 3, the proposed approach is presented. Section 4 discusses applying the approach in the SUOG ontology. Finally, section 5 concludes the paper.

2. Related Works

This section outlines briefly the eXtensible Ontology Development (XOD) strategy [10] and Dead Simple Ontology Design Patterns (DOS-DPs) [11] as related works. XOD is based mainly on two principles: *ontology term reuse* from existing reliable ontologies that are commonly used by the ontology community and *ODP usage* for new term generation and existing term editing. For ontology reuse, XOD suggests applying MIREOT strategy [5] as a commonly known

⁴Basic Formal Ontology, <http://www.obofoundry.org/ontology/bfo.html>

⁵Relation Ontology, <http://www.obofoundry.org/ontology/ro.html>

⁶<http://ontologydesignpatterns.org>

⁷<https://www.suog.org/>

approach. Meanwhile, for ODP usage, the definition of an ODP-based strategy is required. DOS-DP is a pattern-based ontology development practice used to manage the generation of logical definitions in HPO. This approach contributed to developing common patterns valuable for phenotype ontologies which can be applied to a whole branch of an ontology at once. DOS-DPs, which are encoded using JSON⁸, are intended for ontology editors with limited computational expertise. Each pattern is composed of core specification fields such as classes, relations, and *vars* that range over OWL classes. Thus, DOS-DPs rely mainly on *developing* ontology patterns to define ontology terms. Meanwhile, in the SUOG ontology, to not reinvent the wheel, we seek to *reuse* shared and validated ontology patterns for modeling logical definitions.

3. Proposed Approach: Reusing Logical Definitions as ODPs

Our main objective is to define an approach permitting to *reuse* (from HPO) and *apply* (in the SUOG ontology) logical definitions based on ODPs. Thus, combining pattern-based and extensible ontology development approaches is proposed. Therefore, we tackled pattern-oriented ontology design methodologies such as the eXtreme Design (XD) [6]. XD describes strategies for *selecting ODPs for reuse purposes*. ODPs are classified into different types such as, *Presentation*, *Reasoning*, *Content*, and *Structural*. We are interested in *Content* ODPs that solve modeling issues regarding ontology content, either in the general or a specific domain of the study [12]. This study explores *Content* ODPs in the specific domain which is the modeling of logical definitions. Inspired by XOD [10] and based on XD [6], the following steps are defined.

1. *Ontology requirements*: describe the main requirements related to logical definitions in the context of the ontology under development.
2. *Competency questions (CQs)*: translate the ontology requirements into natural or formal language (e.g., SPARQL) questions. CQs will be used further to validate the ontology parts concerned with logical definitions.
3. *Pattern selection*: aims to select ontology patterns related to logical definitions from ontological resources (HPO in our case). More specifically, common patterns representing the “blueprint of logical definitions” [1] and matching the defined competency questions are selected and not the actual definition of terms.
4. *Pattern reuse*: proposes to reuse the selected pattern(s) in the target ontology. Different reuse operations are identified [7]. We are interested in: *import* of ontology pattern as a “building block”, *specialization* that creates an ODP version to specialize concepts and relations conceptually, and *composition* that combines different ODPs to solve the design problem. [6].
5. *Ontology reuse*: to apply the selected patterns, ontology reuse is required. It is prescribed as *content-based* reuse of ontological concepts, or classes, and relations, or properties, from existing validated biomedical ontologies called *source ontologies* [13] (e.g., PATO, BFO, and RO). For implementing the *content-based* reuse, MIREOT [5] can be applied using automated tools such as ROBOT [14] and OntoFox [15]. MIREOT proposes using the minimal information of an external ontology term that is of direct interest to a target

⁸<http://json-schema.org/>

ontology. Thus, when a class is reused, an ontology module that contains the class unique identifier, superclasses, and annotations is built automatically [13] and imported to the ontology under development.

6. *Pattern verification and integration*: aims to verify that the reused pattern covers the requirements. Furthermore, the pattern is integrated in the ontology module [12].

These steps are performed iteratively for each selected pattern under the supervision and validation of domain experts.

4. Application in the SUOG Ontology

This section presents preliminary work on applying the proposed approach (Section 3) for defining prenatal findings in the SUOG ontology logically.

1. *Ontology requirements*: in SUOG, prenatal findings are “signs” identified using echographic mechanisms. They are classified into different categories regarding the affected global anatomical structure. Examples of findings categories are fetal abdomen finding, fetal brain finding, fetal heart finding, etc. Following HPO logical definitions of phenotype abnormalities, findings can be recognized as signs having some abnormal qualities. Thus, the main requirement in SUOG is to associate quality-oriented logical definitions to prenatal findings categories and their subclasses. A quality is defined in PATO as a *dependent entity that inheres in a bearer by virtue of how the bearer is related to other entities*. Moreover, to differentiate prenatal findings from other types of findings (e.g., adult findings), an additional requirement is specified to define the human life cycle stage at which the finding’s existence begins or appears.
2. *Competency questions*: examples of CQs describing the ontology requirements in SUOG represented in natural language are: (CQ1) What are the main qualities related to prenatal findings? (CQ2) How are these qualities defined? (CQ3) What are the basic anatomical structures that the defined qualities inhere in? (CQ4) What are the specific modifiers associated to the defined qualities? (CQ5) At which age do the findings existence appear?
3. *Pattern selection*: different definitions patterns in HPO matched our competency questions. Thus, a selection of blueprint logical definition patterns is performed such as the following *existence* (*P1*) and *quality* (*P2*) patterns. While *P2* describes quality-oriented definitions, *P1* aims to define the stage at which manifestation, or existence, of findings, starts. As in DOS-DPs [11], these patterns are composed of basic categories (classes and relations) and variables *var* that span across OWL classes.

```
P1: 'existence starts during' some var
P2: 'has part' some ('quality'
                    and ('inheres in' some var)
                    and ('has modifier' some var))
```

4. *Pattern reuse*: since all prenatal findings in SUOG are commonly existent at the *fetal stage*, *P1* is required to define the different categories of prenatal finding. In the current work, *P1* and *P2* are applied by composition 107 times to define prenatal finding, prenatal findings categories (e.g., fetal brain finding, fetal heart finding, etc.) and subcategories

(e.g., cerebellum finding, cortex finding, and 4th ventricle finding are subcategories of fetal brain finding). In the following, prenatal finding and fetal brain finding are defined.

```
'prenatal finding' Equivalent To 'has part' some ('quality'  
and ('inheres in' some 'anatomical structure')  
and ('has modifier' some 'abnormal'))  
and 'existence starts during' some 'Fetal stage'
```

```
'fetal brain finding' Equivalent To 'has part' some ('quality'  
and ('inheres in' some 'brain')  
and ('has modifier' some 'abnormal'))  
and 'existence starts during' some 'Fetal stage'
```

Besides, the pattern specialization operation is feasible in the SUOG ontology for defining more specific findings (e.g., enlarged 4th ventricle and dilated 4th ventricle are specific findings of 4th ventricle finding). In the following, an example of defining enlarged 4th ventricle is presented by specializing some concepts of the pattern *P2*.

```
'enlarged 4th ventricle' Equivalent To 'has part' some ('increased quality'  
and ('inheres in' some 'fourth ventricle')  
and ('has modifier' some 'abnormal'))  
and 'existence starts during' some 'Fetal stage'
```

5. *Ontology reuse*: to apply the selected patterns, there is a need to reuse concepts such as quality, brain, fourth ventricle, and Fetal stage, and relations such as *has part*, *inheres in*, and *has modifier*. The ontological resources for reuse purposes are defined by fetal ultrasound experts. Figure 1 depicts an example of ontology modules (e.g., anatomical entity, Onset, quality) reused from UBERON⁹, HPO, and PATO using Ontofox [15]. These modules are imported in the SUOG ontology using owl:import.
6. *Pattern verification and integration*: the associated logical definitions are verified against the CQs, validated by the domain experts, and integrated in SUOG. Figure 2 depicts the logical definition of fetal brain finding.

5. Conclusion

Modeling logical definitions is a promising research field to enrich biomedical ontologies semantically. At the early stages of ontology development, logical definitions associate the terms to external validated ontological resources. In this work, combining pattern-based and extensible ontology development is proposed to select and reuse logical definitions as ODPs. In the SUOG ontology, ODPs reused from HPO are adapted to define prenatal findings logically. The preliminary results are encouraging, 35 findings categories and 72 subcategories are defined by reusing a quality-oriented HPO pattern. In further works, we will accomplish the definitions of findings subcategories and specific classes. ROBOT [14] will be applied for ontology reuse. Besides, patterns adapted to define pregnancy disorders will be considered. In this regard, unlike findings, which are based on *qualities*, disorders will be based on *dispositions* (*BFO:disposition*).

⁹<http://www.obofoundry.org/ontology/uberon.html>

This decision is grounded on the assumption of representing *disorders as material basis of dispositions realized in pathological processes* [16]. Moreover, this work will support the ongoing HPO evolution to cover the fetal phenotype.

Figure 1: Example of ontology modules imported into the SUOG ontology (represented in Protégé¹⁰).

Figure 2: The logical definition of fetal brain finding in the SUOG ontology (represented in Protégé).

Acknowledgments

This project is funded by the EIT-Health Innovation program, selected as part of the bp2020#20062.

References

- [1] S. Köhler, et al., Expansion of the human phenotype ontology (hpo) knowledge base and resources, *Nucleic Acids Research* 47 (2019). doi:10.1093/nar/gky1105.
- [2] C. L. Smith, C.-A. W. Goldsmith, J. T. Eppig, The mammalian phenotype ontology as a tool for annotating, analyzing and comparing phenotypic information, *Genome Biology* 6 (2005). doi:https://doi.org/10.1186/gb-2004-6-1-r7.
- [3] S. Köhler, S. Bauer, C. J. Mungall, G. Carletti, C. L. Smith, P. Schofield, G. V. Gkoutos, P. N. Robinson, Improving ontologies by automatic reasoning and evaluation of logical definitions, *BMC Bioinformatics* 12 (2011).
- [4] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. J. Goldberg, K. Eilbeck, A. Ireland, C. J. Mungall, N. Leontis, P. Rocca-Serra, A. Ruttenberg, S.-A. Sansone, R. H. Scheuermann, N. Shah, P. L. Whetzel, S. Lewis, The obo foundry: coordinated evolution of ontologies to support biomedical data integration, *Nat Biotechnol* 25 (2007). doi:10.1038/nbt1346.
- [5] M. Courtot, F. Gibson, A. L. Lister, J. Malone, D. Schobe, R. R. Brinkman, A. Ruttenberg, Mireot: the minimum information to reference an external ontology term, *Nature Precedings* (2011). doi:10.1038/npre.2009.3576.
- [6] E. Blomqvist, K. Hammar, V. Presutti, Engineering ontologies with patterns: The extreme design methodology, in: *Ontology Engineering with Ontology Design Patterns*, IOS Press, 2016, pp. 23–50. doi:http://dx.doi.org/10.3233/978-1-61499-676-7-23.
- [7] A. Gangemi, V. Presutti, *Ontology design patterns*, in: *Handbook on Ontologies*, 2009.
- [8] J. M. Mortensen, M. Horridge, M. A. Musen, N. F. Noy, Applications of ontology design patterns in biomedical ontologies, in: *AMIA*, 2012, pp. 643–652.
- [9] F. Dhombres, P. Maurice, L. Guilbaud, L. Franchinard, B. Dias, J. Charlet, E. Blondiaux, B. Khoshnood, D. Jurkovic, E. Jauniaux, J.-M. Jouannic, A novel intelligent scan assistant system for early pregnancy diagnosis by ultrasound: Clinical decision support system evaluation study, *JOURNAL OF MEDICAL INTERNET RESEARCH* 21 (2019). doi:10.2196/14286.
- [10] Y. He, Z. Xiang, Y. Lin, J. A. Overton, E. Ong, The extensible ontology development (xod) principles and tool implementation to support ontology interoperability, *Journal of Biomedical Semantics* 9 (2019). doi:10.1186/s13326-017-0169-2.
- [11] D. Osumi-Sutherland, M. Courtot, J. P. Balhoff, C. Mungall, Dead simple owl design patterns, *J. Biomed. Semantics* 8 (2017).
- [12] K. Hammar, *Content Ontology Design Patterns: Qualities, Methods, and Tools*, Ph.D. thesis, Linköping University, 2017.
- [13] C. Ochs, Y. Perl, J. Geller, S. Arabandi, T. Tudorache, M. A. Musen, An empirical analysis of ontology reuse in bioportal, *Journal of Biomedical Informatics* 71 (2017) 165–177.
- [14] R. C. Jackson, J. P. Balhoff, E. Douglass, N. L. Harris, C. J. Mungall, J. A. Overton, Robot: A tool for automating ontology workflows, *BMC Bioinformatics* 20 (2019).
- [15] Z. Xiang, M. Courtot, R. R. Brinkman, A. Ruttenberg, Y. He, Ontofox: web-based support for ontology reuse, *BMC Research Notes* 3 (2010).
- [16] R. H. Scheuermann, W. Ceusters, B. Smith, Toward an ontological treatment of disease and diagnosis, in: *AMIA Summit Transl. Bioinforma*, 2009, pp. 116–120.