

Ontobull and BFOConvert: Web-based programs to support automatic ontology conversion

Edison Ong¹, Zuoshuang Xiang¹, Jie Zheng², Barry Smith³, Yongqun He^{1*}

¹University of Michigan, Ann Arbor, Michigan, USA; ²University of Pennsylvania, Philadelphia, PA, USA; ³University at Buffalo, Buffalo, NY, USA.

ABSTRACT

When a widely reused ontology appears in a new version which is not compatible with older versions, the ontologies reusing it need to be updated accordingly. Ontobull (<http://ontobull.hegroup.org>) has been developed to automatically update ontologies with new term IRI(s) and associated metadata to take account of such version changes. To use the Ontobull web interface a user is required to (i) upload one or more ontology OWL source files; (ii) input an ontology term IRI mapping; and (where needed) (iii) provide update settings for ontology headers and XML namespace IDs. Using this information, the backend Ontobull Java program automatically updates the OWL ontology files with desired term IRIs and ontology metadata. The Ontobull subprogram BFOConvert supports the conversion of an ontology that imports a previous version of BFO. A use case is provided to demonstrate the features of Ontobull and BFOConvert.

1 INTRODUCTION

Biological ontologies are sets of computer- and human-interpretable terms and relations that represent entities in the biological world and how they relate to each other. Hundreds of ontologies have been developed. Over 150 biomedical ontologies have been developed following the Open Biological and Biomedical Ontologies (OBO) Foundry principles (Smith et al., 2007). NCBO BioPortal (Whetzel et al., 2011) currently comprises more than 400 ontologies including both OBO and non-OBO ontologies.

The Basic Formal Ontology (BFO) (Grenon and Smith, 2004; Smith et al., 2007) has been used as the top ontology by some 175 other ontologies (<http://ifomis.uni-saarland.de/bfo/users>), with the goal of facilitating interoperability among these ontologies. Besides BFO, OBO Foundry ontologies use common relations defined in the OBO Relation Ontology (RO) (<https://github.com/oborel/obo-relations>). For example, the information Artifact Ontology (IAO) and the Ontology for Biomedical Investigations (OBI) (Brinkman et al., 2010) both import BFO and RO.

One challenge in reusing external ontologies is that once the reused source ontology is updated with modified IRIs, it is difficult to update the target ontologies. Although these reference ontologies are designed to be maximally stable, they do sometimes still undergo major changes. For example, BFO 1.0 and 1.1 terms do not follow OBO Foundry ID policy (<http://obofoundry.org/id-policy.html>). As a result, BFO has been updated from BFO 1.0/1.1 to 2.0.

To support automatic conversion of ontologies aligned with BFO, we previously developed the BFOConvert software. In light of the fact that the IRI conversion issue exists also in other cases, we have now expanded the BFOConvert program to create Ontobull, a program with greater flexibility and a more user-friendly web interface.

In addition to ontology version updating, Ontobull can be for other purposes, including term IRI updating and updating of ontology metadata such as ontology header (e.g., *owl:imports*) and XML namespaces (xmlns) information (<https://www.w3.org/TR/owl-guide/>). The *owl:imports* attribute in ontology header specifies an OWL file to import into current ontology. The xmlns attribute qualifies element and attribute names of the ontology.

In this software demonstration, we will introduce the tool design and use cases of Ontobull and BFOConvert.

2 ONTOBULL SOFTWARE DESIGN

The Ontobull web program provides a user-friendly interface for input of an ontology file and conversion settings. The Ontobull Java program, developed using the Spring Model-view-controller (MVC) framework and Thymeleaf template engine, processes the user's requests and generates an Ontobull OWL output file. The BFOConvert program is a subprogram of Ontobull specifically targeted conversions involving a move from use of BFO 1.1 to BFO 2.0.

The screenshot shows the Ontobull web interface with three main steps:

- Step 1: Please upload an OWL or ZIP input file as your target OWL files (RDF/XML format) file:** A "Choose File" button is circled in red. Below it, example files like `JAO.zip` and `OPL.owl` are listed.
- Step 2: Provide IRI mapping information:** A table for mapping old IRIs to new IRIs. The first row has "Old IRI" `http://www.obofoundry.org/roto.owl#part_of` and "New IRI" `http://purl.obolibrary.org/obo/BFO_0000050`. Both fields are circled in red. Below the table, an "Alternatively, please upload a custom IRI Mapping file" section has a "Choose File" button.
- Step 3 (Optional): Provide ontology header cleanup information:** A table for header cleanup. The first row has "Action" `add`, "Type" `xmlns`, and "IRI" `http://purl.obolibrary.org/obo/`. The second row has "Action" `remove`, "Type" `xmlns`, and "IRI" `http://www.ifomis.org/bfo/1.1#`. The third row has "Action" `replace`, "Type" `import`, and "IRI" `http://www.ifomis.org/bfo/1.1/http://purl.obolibrary.org/obo/bfo/2014-05-03/classes-only.owl`. The "replace" row is circled in red. Below the table, an "Alternatively, please upload a custom Header Cleanup file" section has a "Choose File" button.

Fig. 1. Ontobull web interface. The circled settings are used in the use case demonstration described next.

* To whom correspondence should be addressed: yongqunh@umich.edu

3 ONTOBULL WEB INTERFACE

The use of the Ontobull web interface (Fig. 1) involves three steps:

Upload an ontology source file: An OWL file or a zip file including an OWL file(s) can be selected and uploaded. The zip file format is typically used when an ontology includes many unmerged OWL files. In this case, Ontobull can automatically process all the OWL files in succession.

Define ontology term IRI mapping: The user is required to provide at least one pair consisting of an original ontology term IRI and a new IRI that will be used to replace it. Multiple mapping pairs can be specified using the dynamic web user-interface. Alternatively, a user can choose a file that includes all the IRI mapping pairs organized in a simple tab-delimited format.

Provide an ontology header and XML namespace updating setting (optional): This setting is used to remove, add, or replace xmlns or *owl:imports* contents in the OWL output file.

Use cases are shown in what follows.

4 ONTOBULL USE CASE DEMONSTRATION

Fig. 2 illustrates how Ontobull is used to convert the Parasite Life Cycle ontology (OPL) which uses terms from the Relation Ontology from an old version to a new version. OPL is an ontology designed to provide a consistent representation of lifecycle stages in parasitic organisms (Parikh et al., 2012). The “OPL.owl” highlighted in the first circle in Fig. 1, above, represents the OWL file of the original version of this ontology. The second highlighted circle in Fig. 1 represents the setting for an IRI mapping. Fig. 2A represents the results of the conversion of the IRI for the RO expression ‘part of’. The third highlighted circle in Fig. 1 represents one header cleanup option, illustrated in Fig. 2B.

A more detailed tutorial and more use case demos are available on the Ontobull tutorial website at: <http://ontobull.hegroup.org/tutorial>. It is noted that BFOConvert has been used to support the conversion of many ontologies that used old versions of BFO and RO.

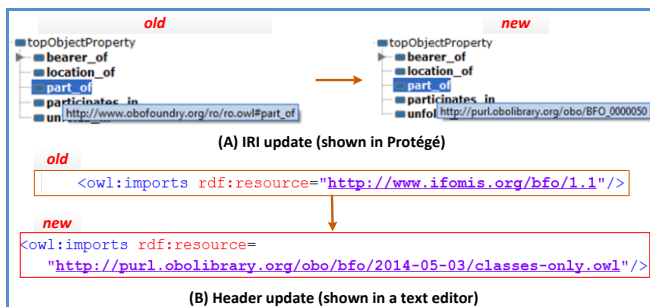


Fig. 2. Ontobull usage demo. The input (OPL.owl) and settings are shown in Fig. 1. (A) The RO object property ‘part of’ IRI updated. The results are displayed using the Protégé-OWL editor. (B) Header updated. The source file updates are shown.

Ontobull can also be used to integrate multiple OWL files. The OWL ontology files may have different IRIs referring to the same entities. For integration, we need to replace these different IRIs with common ones across all salient files. For example, the OWL representation of Beta cell genomics data (<http://www.betacell.org/gbco/>) integrates information from two resources that include terms assigned with different IRIs. Ontobull was used to update IRIs to enable this integration.

5 BFOCONVERT SUBPROGRAM

The BFOConvert subprogram is now available at: <http://ontobull.hegroup.org/bfoconvert>. It provides a pre-designed mapping file according to the current stage of BFO/RO development.

6 SOURCE CODE AND LICENSE

The Ontobull source code is openly available at: <https://github.com/OntoZoo/Ontobull>. The Ontobull source code license is Apache License 2.0.

7 SUMMARY

With ever increasing needs for ontology updating, the web-based Ontobull and BFOConvert programs provide a timely platform for automatic ontology conversion.

ACKNOWLEDGEMENTS

This research is supported by an NIH R01 grant (1R01AI081062). We acknowledge the assistance of James A. Overton and Alan Ruttenberg in the development of BFOConvert.

REFERENCES

- Brinkman, R.R., Courtot, M., Derom, D., Fostel, J.M., He, Y., Lord, P., Malone, J., Parkinson, H., Peters, B., Rocca-Serra, P., Ruttenberg, A., Sansone, S.A., Soldatova, L.N., Stoeckert, C.J., Jr., Turner, J.A., Zheng, J., and Consortium, O.B.I. (2010). Modeling biomedical experimental processes with OBI. *J Biomed Semantics* 1 Suppl 1, S7.
- Grenon, P., and Smith, B. (2004). SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition and Computation* 4, 69-103.
- Parikh, P.P., Zheng, J., Logan-Klumpler, F., Stoeckert, C.J., Jr., Louis, C., Topalis, P., Protasio, A.V., Sheth, A.P., Carrington, M., Berriman, M., and Sahoo, S.S. (2012). The Ontology for Parasite Lifecycle (OPL): towards a consistent vocabulary of lifecycle stages in parasitic organisms. *J Biomed Semantics* 3, 5.
- Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L.J., Eilbeck, K., Ireland, A., Mungall, C.J., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.A., Scheuermann, R.H., Shah, N., Whetzel, P.L., and Lewis, S. (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol* 25, 1251-1255.
- Whetzel, P.L., Noy, N.F., Shah, N.H., Alexander, P.R., Nyulas, C., Tudorache, T., and Musen, M.A. (2011). BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. *Nucleic Acids Res* 39, W541-545.