Overlapping and Competing Ontologies

Jakub Mácha

The Wittgenstein Archives at the University of Bergen
Sydnesplassen 12
N-5020 Bergen, Norway
macha@mail.muni.cz

Rune J. Falch

The Wittgenstein Archives at the University of Bergen
Sydnesplassen 12
N-5020 Bergen, Norway
rune.falch@fof.uib.no

Alois Pichler

The Wittgenstein Archives at the University of Bergen
Sydnesplassen 12
N-5020 Bergen, Norway
alois.pichler@fof.uib.no

ABSTRACT

The aim of this project is to investigate different mappings of documents in humanities into ontologies, i.e. the mutual relations between these mappings: agreement, differences, disagreement. First, we are going to focus on possible sources of differences and disagreement. Second, we want to indicate possible resolutions of differences and disagreement within ontology frameworks or/and employing topic maps. Third, we want to focus on overlapping and conflicting mappings based on mappings in different natural languages. Here, the starting point is not a single document, but various translations of a single document. Fourth, in the last part of the project, we want to provide an experimental justification of achieved results. Drawing on previous projects ([12], [16]), we want to produce a case study of ontologies of Ludwig Wittgenstein's texts based on different interpretative strategies and different translations.

Categories and Subject Descriptors

H.3.4 [**Semantic Web**], I.2.4 [**Ontologies**], D.2.2 [**OWL**].

Keywords

Ontologies, Agreement, Disagreement, Matching, Alignment, Conflict, Ludwig Wittgenstein.

1. INTRODUCTION

The main purpose of building ontologies is to make the source domain easily retrievable, surveyable, combinable, and especially susceptible to automatic processing. We are interested in text ontologies as opposed to more general approaches to cover up and map whole fields and disciplines like philosophy, sociology or political science. Although there are ontologies and mappings generated automatically via natural language processing techniques (e.g., Autonomy or Leximancer), texts in humanities cannot be fully interpreted by an external lexicon only. The nature of our field of inquiry demands that largely we manually build our ontologies and thus rely on human input.

There is an old ideal of text representation without interpretation (*recensio sine interpretatione*). However, texts in humanities and philosophical texts especially typically have different and competing or even contradicting interpretations. This approach has been radicalized by Nietzsche and by recent hermeneutic,

post-modern and analytic philosophy (e.g., Gadamer, Derrida, Quine, Davidson). According to these approaches, every attempt to read and understand a text is biased by previous knowledge and self-knowledge of the interpreter.

Some authors have suggested that ontologies of philosophy do not have to engage with or resolve philosophical disputes [8]. This may be true of general ontologies, but our project focuses on text ontologies. We want to employ mappings from text documents to ontologies in order to cope with issues of inconsistency, incompatibility and contradiction within a text, within its interpretation and between different ontologies.

2. EARLIER WORK BY THE WITTGENSTEIN ARCHIVES

This project follows upon earlier work by the Wittgenstein Archives at the University of Bergen (WAB). The current participation of WAB in the EU projects Agora [1] and DM2E [4] is largely about making primary and secondary sources in Ludwig Wittgenstein scholarship digitally available to the research community. The need for easy navigation and annotation sees WAB move progressively towards the Semantic Web as a technical solution for both the making available and working with these sources.

As part of this work WAB has developed a *Wittgenstein ontology* ([15], [16]), the first version of which was developed in the EU Discovery project in the period 2006-2009 and further developed in later projects. The ontology's classes, subclasses, instances and relations/properties have been chosen with a main view to browsing Wittgenstein's writings and their relations both "internally" and "externally". "Internal" relations include for example references of a Wittgenstein document to another Wittgenstein document or text genetic relations between two Wittgenstein documents; "external" relations include Wittgenstein's references to persons, works by others, but also references of a secondary source to another secondary source or again to a Wittgenstein document. The ontology also includes philosophical subject terms and even phrasings of "points", expressing philosophical claims.

One challenge of such an ontology lies in its being representative (or not), in its being accepted (or not) by the research community. On the one hand, an ontology like this may be general enough and sufficiently modest in its number of elements in order for it to be accepted and agreed upon by the research community. On the other hand, the individual scholar may not agree with all classes, relations and instances, or may even wish to add or suggest new ones (cf. Pichler's discussion in [14]). This raises a number of questions which we wish to pursue in this new project. A particular challenge is posed by the "content" side of the ontology: its terms for philosophical subjects and their organization (instances of Issue), and also the inclusion of philosophical claims (instances of Point). Philosophical claims, be they Wittgenstein's own or the Wittgenstein scholar's, may contradict each other, as may also different organizations of the philosophical subjects (for

examples see [9]). In response to such incompatibilities, the Wittgenstein Archives has introduced the class Perspective which assigns the contradicting organizations / statements to different Perspectives. This permits obtaining a valid ontology, while at the same time acknowledging the authentic tensions which exist in Wittgenstein's philosophy itself as well as in interpretations of it. It will be one of the project's objectives to research how such contradictions and tensions can be extracted from the ontology, though in the ontology itself they are "perspectivized" and thus not contradictions in a proper sense.

3. CHALLENGES

The project has to face the following questions and challenges:

Technical resources (like encoding methods, markup languages etc.) may differ not only in quality and comprehensiveness, but also in the sorts of biases they impose. One of the main preconditions of our approach is to make the user aware of the possibility of being "biased" by the ontology implemented [6].

Ontologies are typically hierarchical structures and they are based on hierarchical markup languages such as OWL. (OWL is better equipped for handling more complex structures than SKOS or RDFS.) The nature of the texts we want to focus on is often seen as non-hierarchical (multi-dimensional, circular etc.; see [9]). This discrepancy between hierarchical markup structure and non-hierarchical texts may influence and challenge the usability of the ontology (cf. [5]). Are topic maps therefore more appropriate for texts in the humanities? What are the advantages and disadvantages of one over the other?

Our backbone ontology consists of classes for sources (Source), persons (Person) and subjects (Subject). The Subject class is the most relevant here and includes subclasses for philosophical subject terms (Issue), philosophical claims (Point) and different interpretations (Perspective). This part of the ontology should be conceived and developed as constant and independent of interpretational differences on the content side. It should thus be "philosophically neutral", i.e. it should be neutral of any particular interpretation that it may include. This is, however, our next challenge: are these backbone elements definitional (as presupposed in [8]) or are they themselves subject to philosophical debates?

WAB's Wittgenstein ontology includes a very great number of RDF triples which are built on top of the backbone ontology and which aim at relating different instances of the hierarchical structure to each other. Implementing cross-ontology relations disrupts its strictly hierarchical structure. We want to investigate what impact the implementation of cross-ontology relations has on the usefulness of the ontology with regard to its suitability to queries and to automatic processing. We would like to propose several kinds of classification of cross-ontology relations in order to preserve the full functionality of the ontology.

4. GOALS OF THE PROJECT

The project has four main goals:

4.1 Semantic Web: methods and methodologies for building ontologies and topic maps

Hierarchical RDF/OWL ontologies on the one hand and topic maps on the other are often mentioned as two alternatives for the Semantic Web [13].

The term ontology has been applied in different ways, but the core meaning within computer science is an agreed model for describing the world that consists of a set of types, properties, and relationship types. It can be used to model a domain and supports reasoning, logical inferencing and similar techniques about concepts in polynomial time. Ontologies can be processed by Description logics (DLs), which are decidable fragments of first-order logic.

Topic maps can be expressed by ontology based languages like RDF/OWL. However, they allow a higher level of abstraction and n-ary relations (hypergraphs). Topic maps are not hierarchical structures. This may lead to an exponential increase of the processing time. There are, however, results indicating that a realistic processing performance is possible [19].

Our first goal is to evaluate whether and to what extent ontologies are more suitable tools for mapping philosophical texts. What are their advantages and disadvantages over topic maps?

4.2 Differences in mappings

The starting point of our considerations is a text document unambiguously identified by its URI. In principle, it may be any text document. In this project, however, the main focus will be on documents in humanities, especially in philosophy. The last part of the project will look at the writings of the Austrian-British philosopher Ludwig Wittgenstein.

In this part of the project, we want to deal with the functions of interpretation in ontology work. What are the sources of differences in mappings from a document to an ontology? We will in particular consider closely the following three sources of differences: (1) text understanding and interpretative strategy, (2) markup language, (3) nature of the text. In more detail:

(1) A source of different mappings of a document can be that its creators have different levels of text understanding and different background understandings; they may focus on different aspects of the text; or they may use different interpretative strategies. The source of different mappings lies in the understanding of their creators or in their approaches to the text. Ontologies can be perspicuous presentations of text understanding and may help to resolve disagreement. This resolution may result into a single ontology.

(2) If encoding, markup and ontology mapping are interpretational (as argued in [2] and [10]) different mappings may originate in decisions that have to be made during the ontology engineering. We want to investigate how particular markup languages can affect ontology mappings.

(3) Different mappings may be brought about by the nature of the text itself. It may be the case that a text is deliberately constructed so that it encourages different mappings. This may be the case with poetry and literary texts more generally, but also of certain philosophical texts like Nietzsche's or Wittgenstein's. If so, we should consider how ontologies can be successfully applied for an analysis of these texts.

With regard to different mappings of a *single document*, we want to suggest the following draft definitions for agreement, overlapping, difference, disagreement and related terms. The presence of logical contradiction is not the only type of disagreement in ontologies [3]. Disagreement in ontologies may arise on two levels: There might be disagreement in entities (instances of Issue) or disagreement in statements (instances of Point). To cope with disagreement in entities is the problem of *ontology matching* [18]. Any agreement or disagreement in statements presupposes (at least partial) agreement in entities. If two ontologies completely disagree in entities, they cover different domains and there cannot be any agreement or disagreement.

Definition 1. Let O be a backbone ontology. A *Perspective* P is a class containing subclasses which each represent a specific grouping or filtering of the instances of Issue and Point [16]. A Perspective must be *closed*: Each Point of P has to mention only Issues from P; and each Issue of P has to be mentioned in some Point of P. In a formal way:

For Every Point [x] and for every Issue [i]: if [x]assertsPerspective[P] and [x]discussesIssue[i], then [i]assertsPerspective[P].

For Every Point [x] and for every Issue [i]: if [i]assertsPerspective[P] and [x]discussesIssue[i], then [x]assertsPerspective[P].

Definition 2. Let P_1 and P_2 be Perspectives. Their *Issues-alignment* is a set of RDF triples of the form $\langle i_1, \text{matches}, i_2 \rangle$ where $[i_1]$assertsPerspective$[P_1]$ and $[i_2]$assertsPerspective$[P_2]$. An Issues-alignment is *complete/partial* if it contains all/some Issues from P_1 and P_2. If there is a partial Issues-alignment between Perspectives P_1 and P_2, we can say they are *overlapping*.

Definition 3. Let P_1 and P_2 be Perspectives and let IA be their partial Issues-alignment. Let x_1, x_2 be points (in the form $\langle s_1, r_1, o_1 \rangle$ and $\langle s_2, r_2, o_2 \rangle$) such that $[x_1]$assertsPerspective$[P_1]$ and $[x_2]$assertsPerspective$[P_2]$. Points x_1, x_2 are *matching* in IA if IA contains all triples $\langle s_1, \text{matches}, s_2 \rangle$, $\langle r_1, \text{matches}, r_2 \rangle$ and $\langle o_1, \text{matches}, o_2 \rangle$.

Definition 4. Let P_1 and P_2 be Perspectives and let IA be their partial Issues-alignment. Their *Points-alignment* is a set of matching Points from these Perspectives. The Points-alignment is *complete/partial* if it contains all/some Points from P_1 and P_2.

Definition 5. Perspectives P_1 and P_2 are in a *complete/partial agreement* if there is their complete/partial Points-alignment.

Definition 6. Perspectives P_1 and P_2 are *different* if there is not any complete Points-alignment between them.

Definition 7. Perspectives P_1 and P_2 are *conflicting* if there is an Issue-alignment IA between them and there are matching Points x_1, x_2 such that $[x_1]$assertsPerspective$[P_1]$ and $[x_2]$deniesPerspective$[P_2]$.

Equipped with these definitions we are able to formulate the following questions and goals:

4.2.1 Ontology matching

- How do we detect agreement, overlapping, difference between two or more ontologies? It is one thing to detect disagreement or overlap, and it other thing to provide an explanation of these facts, i.e., to find out where exactly the ontologies disagree or overlap ([3]: 152).

- How do we represent alignment between two or more ontologies? How do we make alignment accessible for users?

- The matching method can be improved by background knowledge (e.g., a search engine, specific corpora, linked open data, involvement of users, social and collaborative matching) which is not part of the final ontology.

- User involvement can take place on two levels: first, in the sense of developing different text ontologies; second in the process of their matching. This matching process can be ex-

tended into the form of "social matching" (through explicit arguing and/or voting).[1]

4.2.2 Disagreement in Points

As to the possible sources of differences in mappings, we may pose the following questions:

- Could there be conflicting mappings originating in different interpretative strategies / in the markup language / in the nature of the text?

- If so, could these conflicts originating in different interpretative strategies / in the markup language / in the nature of the text be reconciled within a single ontology or even within a single Perspective?

- How do we detect conflicting ontologies (given their alignment report)?

- How do we reconcile conflicting ontologies? Can ontology work help resolve differences in interpretations?

- Could conflicting ontologies be merged into one single topic map?

4.3 Overlapping and conflicting mappings based on differences in natural languages

In this part of the project, the starting point is not a single document, but various translations of a single document. We want to focus on two key questions:

Do translations bias ontologies?

In the previous section, we mentioned three possible sources of differences in mappings. Mappings from different documents into a single ontology may be captured as Perspectives. If so, we may employ the framework sketched above in order to investigate where different translations yield different or even conflicting Perspectives.

Can ontology resolve differences in interpretations caused by different translations?

A translation may be considered successful if the ontology of the translation is isomorphic to the ontology of the original text. We can, thereby, employ the framework sketched above to evaluate translations.

4.4 Experimental part: A case study of ontologies of Wittgenstein's texts

(1) Different interpretations: In this part we want to apply results from the previous parts of the project to some specific instances from Wittgenstein's works and different interpretations of them. Here, a number of questions are raised: Which ontologies can be said to be present in different seminal interpretations of Wittgenstein?[2] How can these ontologies be "extracted" and articulated; how do these ontologies relate to or differ from each other; where do different interpretations entail different ontologies; and how do we treat these differences in the WAB ontology ([15], [16]) and in the setup of browsing/annotation tools like *SWickyNotes* [11] and *Pundit* [17]? Here it might be interesting and fruitful to invite scholars to actively participate in producing such articulations of their ontologies. Another possibility would be that we ourselves

---

[1] These suggestions are discussed in [18]: §§9–13.

[2] E.g. Hallett's, Baker's, Hacker's and von Savigny's commentary works, Glock's Wittgenstein dictionary, Diamond's and Conant's "resolute" readings a.o.

try to extrapolate and articulate the ontologies and then present them to and discuss with the respective scholars. In both cases there is a positive need for close collaboration with the research community.

(2) Translations: Here we want to look at translations of Wittgenstein's works in order to see if/how different translations might give rise to different interpretations (and thus possibly also different ontologies). We will deal with translations from German into English, from English to German,[3] from German to Czech and also from German to Norwegian. An example of a translation which may give rise to interpretational differences would be the way the German word "Satz" has been variously rendered in English as both "sentence" and "proposition". Another example would be to look at different translations of Wittgenstein's *Tractatus*: E.g. the German word "Sachverhalt" is translated as "atomic fact" in the Ogden translation and as "state of affairs" in the Pears/McGuinness translation,[4] as "stav věcí" [state of things] in the Czech translation and as "saksforhold" in the Norwegian. A third possibility is to look at the two English translations of *Culture and Value* which is interesting in that one and the same translator has applied different strategies in his two translations (cf. [7]: 187–200). For the interpretational and ontological aspects we could furthermore investigate whether/how commentaries and interpretative works differ according to the translation used, and ultimately how these impinge upon the ontology work.[5]

5. CONCLUSION

Our project wants to investigate knowledge-representational challenges to ontology work stemming from different mappings in humanities scholarship. We use Wittgenstein texts and Wittgenstein studies as our test bed. The project as such is propaedeutic, but may ultimately see the implementation of its results in adequate applications to the domain of Wittgenstein scholarship, but also other fields in philosophy or the humanities more generally.

6. REFERENCES

[1] AGORA Scholarly Open Access Research in European Philosophy. `http://www.project-agora.org/`.

[2] Burnard, L. 1995. What Is SGML and How Does It Help? *Computers and the Humanities* 29, 41-50. DOI= http://dx.doi.org/10.1007/BF01830315.

[3] d'Aquin, M. 2009. Formally Measuring Agreement and Disagreement in Ontologies. *K-CAP '09* (Redondo Beach, California, September 1 - 4, 2009). ACM, New York, 145-152. DOI= http://dx.doi.org/10.1145/1597735.1597761.

[4] DM2E Digitalized Manuscripts to Europeana, `http://dm2e.eu/`.

[5] Erbacher, C. 2011. Unser Denken bleibt gefragt: Web 3.0 und Wittgensteins Nachlass. In *Wissenschaftstheorie, Sprachkritik und Wittgenstein*, S. Windholz and W. Feigl, Eds. Ontos Verlag, Heusenstamm, 135-146.

[6] Falch, R., Erbacher, C., Pichler, A. 2013. Some observations on developments towards the Semantic Web for Wittgenstein scholarship. In *Papers of the 36th International Ludwig Wittgenstein Symposium* (Kirchberg am Wechsel, Austria, August 11 - 17, 2013). ALWS, Kirchberg am Wechsel.

[7] Gorlée, D. L. 2012. *Wittgenstein in Translation: Exploring Semiotic Signatures*. De Gruyter Mouton, Berlin/Boston.

[8] Grenon, P. and Smith B. 2011. Foundations of an Ontology of Philosophy. *Synthese* 182, 2, 185-204. DOI= http://dx.doi.org/10.1007/s11229-009-9658-x.

[9] Huitfeldt, C. 1995. Multi-Dimensional Texts in a One-Dimensional Medium. *Computers and the Humanities* 28, 4/5, 235-241. DOI= http://dx.doi.org/10.1007/BF01830270.

[10] Huitfeldt, C. 2004. Text Technology and Textual Criticism. *Linguistica Computazionale*, XX-XXI, 259–275.

[11] Morbidoni, C. and Nucci, M. 2010. SWickyNotes user guide. `http://www.swickynotes.org/docs/SWickyNotesStartingGuide.pdf`.

[12] Pasin, M. and Motta, E. 2008. PhiloSURFical: Browse Wittgenstein's World with the Semantic Web. In *Wittgenstein and the Philosophy of Information*. A. Pichler and H. Hrachovec, Eds. Ontos Verlag, Heusenstamm, 319-334.

[13] Pepper, S., Vitali, F., Garshol, L. M., Gessa, N. and Presuti, V. 2006. A Survey of RDF/Topic Maps Interoperability Proposals, *W3C Working Group Note*. 10 February 2006, `http://www.w3.org/TR/rdftmsurvey/`.

[14] Pichler, A. 2013. *The Wittgenstein Incubator and Swicky Notes*. Technical Report. University of Bergen. `http://www.slideshare.net/DM2E/berlin-16161631`.

[15] Pichler, A. 2013. *Wittgenstein Ontology*. Technical Report. University of Bergen. `http://wab.uib.no/wab_philospace.page`

[16] Pichler, A. and Zöllner-Weber, A. 2013. Sharing and Debating Wittgenstein by Using an Ontology. *Literary and Linguist Computing*, 1-8. DOI= http://dx.doi.org/10.1093/llc/fqt049.

[17] Pundit. `http://www.thepund.it/`.

[18] Shvaiko, P. and Euzenat, J. forthcoming. Ontology matching: state of the art and future challenges. *IEEE Transactions on Knowledge and Data Engineering*.

[19] Stefanova, S. and Risch, T. 2008. Viewing and Querying Topic Maps in terms of RDF. *SeMMA 2008*, 69-83 `http://ceur-ws.org/Vol-346/6.pdf`.

---

[3] Some of Wittgenstein's works were originally written in English.

[4] While "Tatsache" is translated as simply "fact" in both, as "fakt" in Czech and "kjennsgjerning" in Norwegian.

[5] We thank Claus Huifeldt and DH-CASE reviewers for their helpful and insightful comments. Jakub Mácha has been supported by project GACR P401/11/P174.