Lost in Translation? Empirical Analysis of Mapping Compositions for Large Ontologies

Anna Tordai,[1] Amir Ghazvinian,[2] Jacco van Ossenbruggen,[1] Mark A. Musen,[2] Natalya F. Noy[2]

[1] VU University Amsterdam, Amsterdam, The Netherlands
{atordai,jrvosse}@few.vu.nl
[2] Stanford University, Stanford, CA 94305, US
{amirg,musen,noy}@stanford.edu

Abstract. When three or more ontologies have been aligned, longer chains of mapped concepts start to appear. In this paper, we empirically study the nature of these composite mappings, focusing on chains of (near) equivalence links of length two. We ask human experts to evaluate samples of composite mappings, taken from large real life data sets. Based on these evaluations, we analyze the features of mappings produced by composition in three different domains (bio-medicine, cultural heritage, and library subject headings), among ontologies in multiple languages (English, Dutch, German, and French), and using existing mappings that were created by different methods (lexical and instance-based methods). We examine the quality of the composite mappings relative to the quality of the input mappings and analyze how characteristics of the input mappings and the ontologies influence the composition.

1 Introduction

Researchers typically study ontology alignments in the context of a single source and target ontology. As more and more of such alignments are being created and published, however, longer chains of equivalent or otherwise related concepts start to emerge in our data sets. In this paper, we analyze the quality of a subset of such chains, focusing on short chains of equivalence and near equivalence links. Most of us have clear intuitions about the properties of such chains. For example, equivalence relations such as `owl:sameAs` and `skos:exactMatch`, are defined as being transitive, so it should be safe to assume that if term A is equivalent to B, and B is equivalent to C, then A should also be equivalent to C. We will test this hypothesis empirically by determining to what extent such transitivity actually holds in our data sets, and if not, what is going wrong Furthermore, for relations such as `skos:closeMatch`, which are not defined as being transitive, we might ask how often chains of these relations turn out to be transitive after all.

We use the notion of a mapping as defined in [15]. Given a mapping from A to B and from B to C, where concepts A, B and C are part of three different ontologies, we call the mapping from A to C a *composite mapping*. Although mapping composition is related to the use as background knowledge where concept B would be part of the background ontology [2], we do not predefine ontologies as a source of background knowledge. We analyze the properties of such composite mappings on real life data sets, addressing the following two research questions:

- What is the quality of composite mappings relative to the quality of input mappings?
- Does the quality of composite mappings depend on other characteristics of the input mappings or ontologies?

In order to answer these research questions, we study composite mappings for ontologies in different domains, using input mappings generated in different ways (Section 3.1). We analyzed the precision of composite mappings by sampling them and having human experts verify the samples (Section 3.3). In some cases, we already had pre-existing alignments for the sets of ontologies for which we analyze composite mappings. In these cases, we compared the precision of the composed mappings with the precision of existing mappings. We then analyzed our results (Section 5) and made observations regarding the quality and quantity of composed mappings, trying to identify reasons for correct and incorrect mapping compositions based on characteristics of the data and the input mappings.

The main contribution of this paper is a large-scale empirical analysis of the nature of composite mappings given varied sets of input ontologies and mappings.

2 Related Work

Researchers in the area of database schema matching have studied mapping composition extensively [1, 14, 4]. However, these researchers have focused on the use of mapping composition to perform query transformation rather than for generating mappings.

In ontology matching, Euzenat [6] discusses mapping composition in a theoretical paper on algebras of relations as a means for validating existing mappings and creating new mappings. This work considers composition through equivalence mappings to be a trivial case because the result is an equivalence relation, and because we can assume that equivalence is transitive. In practice, however, automatically generated mappings are usually similarity mappings at best, and therefore the composition of such mappings is not trivial. We look at such automatically generated mappings and analyze results of composition to find out whether they are interesting or truly lost in translation.

Researchers have already developed a plethora of tools for generating mappings and compared their performance at the OAEI. These off-the-shelf tools, such as ASMOV [12], RiMOM [22], Falcon-AO [11], and DSSim [16] perform well on OAEI benchmarks and on certain specialized tracks. However, the results of the 2009 library track showed that current tools largely fail on extremely large vocabularies and vocabularies that use multiple languages [7].

Mapping composition has some parallels to the use of background knowledge by mapping tools. Tools such as SAMBO [13] and ASMOV use background knowledge (UMLS Metathesaurus, WordNet) to improve the quality of mappings. When mapping two domain ontologies, these tools either use existing mappings from these domain ontologies to some background source, such as UMLS or WordNet, or create these mappings "on the fly" through lexical comparison or other means. The tools then use

Set	Domain	Ontologies	Language	Ontology size
BioPortal	Biomedicine	151 ontologies from BioPortal	English	Ranging from under 100 concepts to 380K concepts Mean size=17,805 (SD= 61,614) Total concepts: 2,688,609
CH	Cultural Heritage	Getty's Art and Architecture Thesaurus (AAT)	English and Dutch	27,077 concepts with English and Dutch labels
		Princeton WordNet	English	115,424 synsets with 203,147 English labels
		Cornetto	Dutch	70,370 synsets and 103,762 Dutch labels
Library	General	Library of Congress Subject Headings (LCSH)	English	339,612 concepts
		Rameau	French	157,287 concepts
		SWD	German	163,175 concepts

Table 1. Sets of ontologies that we used in mapping composition and their characteristics.

these mappings to a single source of background knowledge for creating mappings for the domain ontologies. This method is related to mapping composition because we use a mapping to a third ontology or vocabulary. In this sense, in mapping composition *any* ontology becomes a source of background knowledge.

The COMA [5] and COMA++ [3] tools combine several matching techniques including composition of mappings. The evaluation of the tools demonstrated the effectiveness of mapping composition without going into a more detailed analysis of the results.

3 Materials and Methods

In this section, we describe the ontologies and existing mappings that we used for mapping composition (Section 3.1), the method for creating compositions and its complexity (Section 3.2), and our methodology for assessing the precision of the composed mappings (Section 3.3).

3.1 Data: Ontologies and Input Mappings

In order to get a comprehensive analysis of mapping composition under different conditions, we considered three sets of ontologies and mappings. We have ontologies in three different domains: biomedicine, cultural heritage and library subject headings (Table 1). The terms in these ontologies have labels in four languages: English, Dutch, German and French, and the input mappings we use for composition were generated using two types of methods: lexical method, and instance-based method (Table 2).

Set	Method for mapping generation	Number of mappings used for composition	Precision
BioPortal	Lexical comparison of preferred names and synonyms	575,642 mappings	0.94
		459,941 Preferred–Preferred	0.99
		115,701 Preferred–Synonym	0.76
CH	Lexical comparison of labels	6,914 AAT–Cornetto	0.88
		4,592 AAT–WordNet	0.82
		3,144 Cornetto-WordNet	0.95
Library	Instance-based	2,242 LCSH–Rameau	0.95
		2,334 SWD–LCSH	0.54
		685 Rameau–SWD	0.72

Table 2. Input mappings that we used for mapping composition for the three sets of ontologies.

Our first set of ontologies came from BioPortal [17], a Web-based repository of biomedical ontologies. At the time we collected the data, BioPortal contained 151 ontologies with more than 2.5 million concepts among them. We generated mappings between these ontologies using simple lexical comparisons of preferred names and synonyms after normalization [9, 8].

The second set of mappings links three large vocabularies in the cultural-heritage domain: Getty's Art and Architecture Thesaurus (AAT[3], extended with Dutch labels from AATNed[4]), Princeton WordNet[5] version 2.0 and Cornetto,[6] a WordNet-like lexical resource for Dutch. We generated mappings between AAT and WordNet, and between AAT and Cornetto using simple lexical comparison [19]. The Cornetto project [20] created mappings between Cornetto and different versions of WordNet using a combination of manual and automatic methods.

Finally, we used a set of ontologies and mappings from the Library track in the OAEI 2009. This set contains three lists of subject headings for describing content of books: the Library of Congress Subject Headings (LCSH); Rameau, a list used by the French National Library; and the Subject Heading Authority File (SWD), which is used by the German National Library. Each list contains from 150,000 to 300,000 concepts.

We used the mappings that Wang and colleagues [21] created using instance-based matching based on books that were classified using terms from more than one vocabulary. This method for generating mappings ranks the resulting mappings according to confidence level. Although there are a total of almost 2 million mappings, over 90% of them have confidence measure lower than 0.1. For the purpose of composing mappings, we selected only those mappings that had a confidence measure greater than 0.7.

[3] http://www.getty.edu/research/conducting_research/vocabularies/aat/
[4] http://www.aat-ned.nl/
[5] http://wordnet.princeton.edu/
[6] http://www2.let.vu.nl/oz/cornetto/index.html

We estimate the precision of these mappings by evaluating samples manually. These mappings involve fewer than 1.5% of the concepts in the vocabularies.

In the cultural heritage and OAEI library track the number of input mappings is significantly lower than in the BioPortal case, as our aim was to select high-quality mappings. We chose a representative subset in order to analyze the properties of mapping composition.

3.2 Computing Mapping Composition

In this paper, we consider only composition of two mappings. The BioPortal compositions were computed using a relational database, and the cultural heritage and OAEI library track composition algorithms were written in SWI-Prolog.[7].

Since we had detailed information on the source of the input mappings for all BioPortal ontologies, we further analyzed the composed mappings for BioPortal to understand better how characteristics of input mappings affect the number and precision of composed mappings. To perform such analysis, we broke down the composed mappings into groups, based on the types of input mappings that contributed to the composition. There are two types of input mappings (see Table 2): Preferred–Preferred and Preferred–Synonym mappings. We do not include Synonym–Synonym mappings in our input because they have low precision(0.36). Different combinations of the input mappings produce six groups of composed mappings which are displayed in Figure 1.

For instance, Figure 1A illustrates the case where we composes a mapping from a preferred name for the concept C_1 to a preferred name for C_2 with a mapping from the preferred name for C_2 to the preferred name of C_3. We refer to this case as PPP. Note that this composition produces a subset of the Preferred–Preferred mappings between O_1 and O_3. PSP mappings (Figure 1B) also produce a subset of the Preferred–Preferred mappings. Similarly, PPS mappings (Figure 1C) and SPS mappings (Figure 1D) produce subsets of the Preferred–Synonym and Synonym–Synonym mappings between O_1 and O_3, respectively. We analyze these subsets and compare their precisions to those of the original Preferred–Synonym and Synonym–Synonym mappings that were generated directly by comparing O_1 and O_3. Figure 1E and F illustrate the other two cases, $PSPS$ and $PSSP$, which produce mappings that we cannot obtain by comparing preferred names and synonyms directly.

3.3 Sampling and evaluation

In order to evaluate the precision of the composed mappings as well as the precision of input mappings (see Table 2), we sampled the mappings and evaluated the samples manually. Because of the scale of our data—with hundreds of thousands of mappings to verify—evaluating all the mappings manually was not feasible. Furthermore, because of the size of the ontologies themselves, creating a complete set of mappings so that we can evaluate recall was not feasible either. In addition, the recall of mapping composition is necessarily limited by the recall of the input mappings used for composition. Thus, we focus in this evaluation on estimating only the precision of the composed mappings.

[7] http://www.swi-prolog.org/

For BioPortal mappings, we used stratified sampling [10] to select mappings for manual evaluation. Among the BioPortal ontologies, there is a large number of ontology pairs that have only one or two composed mappings between them. At the same time, there are pairs of ontologies that have thousands of mappings between them. Therefore, we constructed the strata to ensure that our samples include mappings between ontology pairs with only a few mappings between them, as well as mappings between ontology pairs with thousands of mappings, and clusters in between. We sampled a total of 2350 mappings from the different BioPortal mappings sets. Our sample sizes ranged from 210 to 400 mappings per set depending on the number of original mappings.

In the case studies involving cultural heritage and library subject headings, we evaluated manually all mapping sets containing fewer than 500 mappings and took samples of 100 mappings from larger sets. We sampled the total of approximately 1,000 mappings from these sets.

Human experts evaluated the samples using the evaluation tool used in [18] for the cultural heritage and Library track data, and a similar tool for the BioPortal data, and categorized each mapping into one of six categories: *exact match*, *close match*, *broader match*, *narrower match*, *related match*, or *incorrect*. For measuring precision, we considered only exact and close matches as correct. A detailed analysis of the broader, narrower and related matches is out of scope for this paper but we plan to perform it in the future. We measured agreement using Cohen's kappa on subsets of samples between raters, finding substantial agreement for BioPortal (0.72) and cultural heritage evaluation (0.70) and almost perfect agreement with the manually evaluated mappings used in the OAEI library track (0.85).

4 Results

In this section, we present the precision of mapping composition for the three sets of ontologies in our study. We discuss these results in Section 5.

4.1 Results: Biomedical Ontologies

Figure 2A shows the results for the overall precision of composed mappings. Using 575,642 input mappings with precision 0.94, we generated 599,625 composed map-

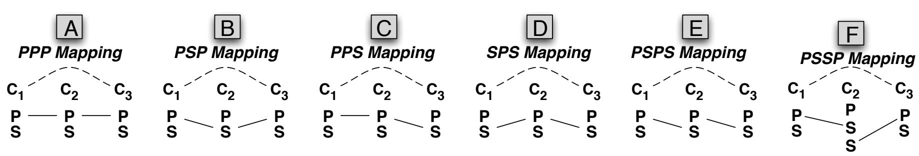

Fig. 1. Methods for composing mappings between concepts in three different ontologies ($C_1 \in O_1$, $C_2 \in O_2$, $C_3 \in O_3$) using mappings between preferred labels (P) and synonyms (S). Figure A illustrates the PPP mappings: a composition of a mapping from a preferred name of C_1 to preferred name of C_2 with the mapping between preferred names of C_2 and C_3. Figure B illustrates PSP mappings: a match of C_1 preferred name to C_2 synonym with a match of C_2 synonym to C_3 preferred name. Figures C-F illustrate the remaining possible cases.

pings with a precision of 0.92. Figures 2B, 2C, and 2D show the precision of composition for different cases from Figure 1. We group these cases by the sets of input mappings that they used. Composing Preferred–Synonym mappings, which had a precision of 0.76, yielded 147,438 composed mappings with precision 0.84. Other combinations (Figures 2C and 2D) resulted in sets of composed mappings with precisions similar to the precisions of the input mappings.

Figure 3 provides additional information on the precision of the individual cases. The two cases that resulted in the subset of what we could have obtained directly by comparing preferred names lexically (PPP and PSP), provided mappings with the highest precision, 0.99. The SPS mappings constitute a subset of the Synonym–Synonym mappings for O_1 and O_3. We did not use these types of mappings as input mappings because they have very low precision, 0.36. However, using mapping composition to identify a subset of Synonym–Synonym mappings almost doubles the precision of these mappings, bringing it up to 0.6.

Additionally, using composition, we identified mappings without lexical similarity in their preferred names or synonyms ($PSPS$ and $PSSP$ mappings). Such mappings can be identified by composition through a concept with lexical similarity to both mapped concepts. These two cases produced 50,353 new mappings with the precision of 0.68. For example, we found a $PSSP$ mapping between the concept CRANIAL_SKELETON from the Amphibian gross anatomy ontology and SKULL from the Foundational Model of Anatomy. These two concepts each map to the concept CRANIUM from the Teleost anatomy and development ontology, which has the synonyms CRANIAL SKELETON and SKULL.

4.2 Results: Cultural Heritage

Figure 4A shows the results of mapping composition for the cultural heritage domain. The precision of composed mappings is at least 0.8 in all three cases, with the number of mappings identified through composition ranging from 263 to 1,774. In fact, the composed mappings between Cornetto and WordNet have a precision of 0.9.

Because we have lexical mappings available for this set, we can compare the composed mappings to the lexical ones, and analyze how many non-lexical mappings we generate by composing lexical mappings.

Upon closer examination of the mappings, we found that 134 (30%) of the composed mappings between AAT and WordNet have little or no lexical similarity. For example, through composition we mapped TOBACONNISTS' SHOP to TOBACCO SHOP and WATCHMEN to GUARD. Similarly, we found 110 non-lexical mappings between AAT and Cornetto, such as BADKLEDING to BADKOSTUUM, both of which mean "bathing suit" in Dutch. This subset of composed mappings not including lexical similarity has a precision of 0.56, which is lower than the precision of composed mappings in general.

Between Cornetto and WordNet, 1,208 of the 1,774 composed mappings are listed as "near equal synonym" mappings in the original Cornetto-WordNet mappings of the Cornetto project. These are not the same as the "equal synonym" mappings used as input mappings for other compositions. Another 448 of the composed mappings are entirely new and have an average precision of 0.7.

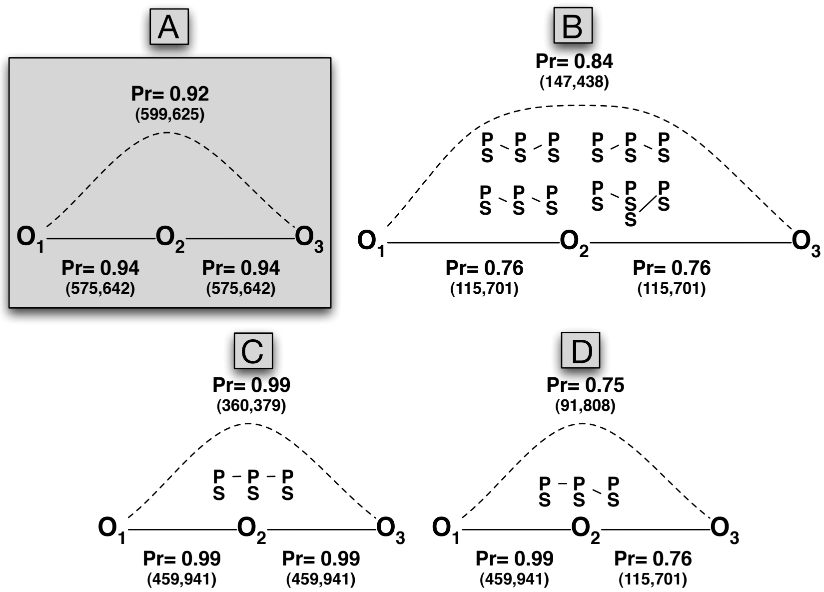

Fig. 2. Mapping composition results for BioPortal ontologies. O_1, O_2 and O_3 represent any three ontologies linked through mappings from Bioportal. Figure A (the shaded diagram) shows the overall precision of the input mappings and their number in parentheses on the solid lines. It shows the precision of composed mappings and their number above the dotted line. Figures B, C, and D provide details for the precision of composed mappings, grouped by the precision of input mappings. Figure B contains the mappings that used only Preferred–Synonym mappings as input; Figure C contains the mappings that composed Preferred–Preferred mappings; and Figure D provides the data for the composition of Preferred–Preferred mappings and Preferred–Synonym mappings.

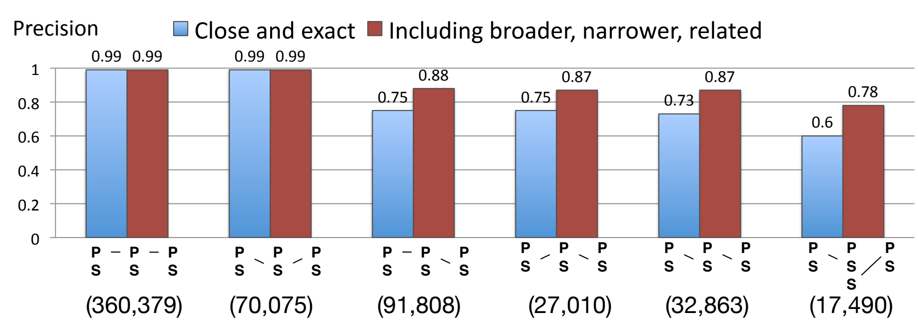

Fig. 3. Mapping composition results for BioPortal ontologies. The bar graph shows precision for composed mappings. The (lighter) left bar shows precision of exact and close matches, the (darker) right bar shows the precision if we include broader, narrower, and related matches. Numbers in parentheses indicate the total number of mappings

4.3 Results: The OAEI Library Track

Figure 4B shows the results of mapping composition using the library subject headings mappings. Precision of the composed mappings is higher than 0.74 and the number

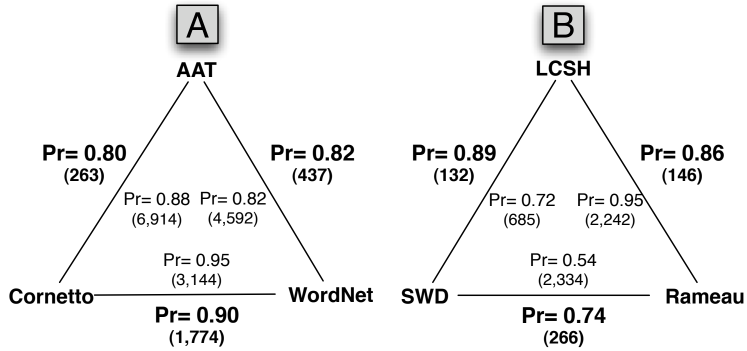

Fig. 4. Mapping composition results for cultural heritage domain (Figure A) and OAEI library track (Figure B). The numbers in bold outside the triangle show the precision and the number of composed mappings in parentheses. The numbers inside the triangle show the precision and number of input mappings

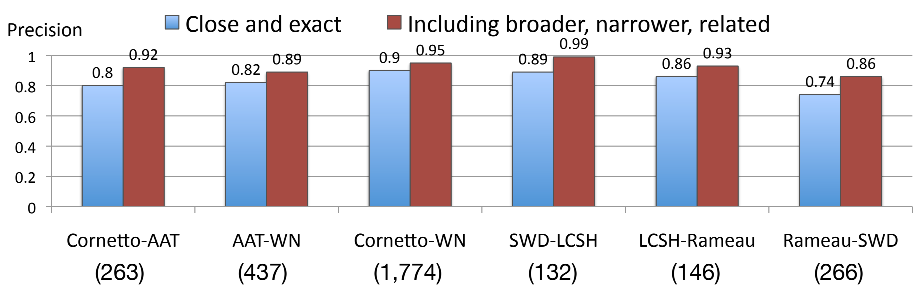

Fig. 5. Mapping composition results for cultural heritage and library-track ontologies. The bar graph shows precision for composed mappings. The (lighter) left bar shows precision of exact and close matches, the (darker) right bar shows the precision if we include broader, narrower, and related matches. Numbers in parentheses indicate the total number of mappings

of generated mappings ranges from 132 between the Subject Heading Authority File (SWD) and the Library of Congress Subject Headings (LCSH) and 266 between SWD and Rameau (a list used by the French National Library).

In two cases—mappings between SWD and LCSH and mappings between Rameau and SWD—the composed mappings actually had higher precision than the input mappings.

In this case, we also compared the composed mappings to the input mappings. We found that, of the 132 mappings between SWD and LCSH, 13 (10%) mappings did not overlap with any of the original instance-based mappings, including those that had a confidence measure lower than 0.7. In other words, for these 13 mappings, there were no instances (books) available. For LCSH and Rameau, we found 8 (5%) such "new" mappings, and for Rameau and SWD, 65 (24%) mappings. The high number of new composed mappings between Rameau and SWD is due to the low number of instances available for creating the original mappings. However, the precision of these subsets is lower: 0.37 between LCSH and Rameau, 0.54 between Rameau and SWD, and 0.92 between SWD and LCSH.

4.4 Broader, Narrower, and Related Mappings

When evaluating the composed mappings, we have also recorded whether each mapping represented a narrower, broader, or related mapping, rather than a close or exact match. Figures 3 and 5 show the increase in precision of composed mappings if we also count broader, narrower, and related mappings as correct. Figure 5 includes the data for both the cultural-heritage and the library-track case. The increase in precision in both of these cases is less dramatic than for the biomedical ontologies. In this case, the average increase in precision was 11%, whereas for BioPortal ontologies the average increase was 14%, with the most significant increase (30%) in the $PSSP$ case.

5 Discussion

In this paper, we have presented the results of our analysis of mapping composition in three different domains. Our results show that the quality of composed mappings depend on the ontology characteristics, and the content and quality of the input mappings.

The characteristics of the ontologies, such as the way they are implemented or the way preferred labels and synonyms are used, have a profound effect on composition. For example, in Medical Subject Headings (MeSH) concepts often have narrower terms as synonyms. The concept TREMORS in MeSH has a synonym NERVE TREMOR, which in reality is a narrower term, not a synonym. As a result, many of the composed mappings that involved MeSH terms were not close matches but rather broader or narrower mappings.

It is clear that the number of input mappings determines the number of composed mappings, but we see in our results that there are large variations in the number of input mappings in three cases studies. This is partly due to the size of the ontologies and partly because of the confidence level of the original input mappings which is a limiting factor for example in the Library track case study.

The content of the ontologies also influences the quality of the mapping compositions. When the content overlaps, meaning the domains of the ontologies are the same or very similar, the meaning of the concepts is also closer, and the composed mappings are likely to be equivalence mappings rather than broader, narrower or related mappings. In the cultural heritage case study Cornetto and WordNet are unlikely to cover art and architectural concepts, reducing the chance of creating equivalence compositions between AAT and Cornetto, and AAT between AAT and WordNet.

Finally, the quality of input mappings has a direct effect on the quality of mapping compositions. High quality input mappings tend to result in high quality mapping compositions. Intuitively, the precision of the compositions should be the product of the precisions of the input mappings. However, especially in the BioPortal data we find cases where the precision of compositions exceeds the precision of input mappings (Figure 3). We find similar cases in both the AAT and Library track case studies.

We also found that many of the composed mappings though not exact, or close matches nevertheless represent a semantic relationship such as broader, narrower or related (Figure 3 and 5). For example, the concept BLURRED VISION from the "Suggested Ontology for Pharmacogenomics" maps to composition to VISION ABNORMAL in the "WHO Adverse Reaction Terminology", forming a narrower relationship between the two concepts. This kind of semantic drift between concepts seems to arise often through mapping composition caused by ontology characteristics, or concepts deviating in meaning in different languages.

In our future work we plan to perform a more detailed evaluation of the content of the mappings to determine why the precision of the composed mappings exceeds the precision of the input mappings in certain cases. We also need to study the effect of semantic drift by analyzing the relationship between the type of equivalence relation in input mappings and compositions, and extend our scope to longer composition chains.

6 Conclusion

In this paper, we presented an empirical analysis of the quality of mapping composition for various use cases. We conclude that mapping composition produced mappings of comparable quality to the input mappings; precision of the composed mappings is not much worse than the original precision of mappings, and sometimes it is even better. Even when composing lexical mappings, in some cases we produced mappings that lexical methods would not produce. Finally, the quality and the number of composed mappings can be affected significantly by the characteristics of the ontologies themselves, the type of input mappings, the number and coverage of the input mappings. Our results confirm our intuitions on mapping composition. The contribution of this paper is that we have tested these intuitions empirically and validated them using methods described in literature.

Acknowledgments

Many thanks to Guus Schreiber and Bob Wielinga for their help and comments on this paper. We thank Antoine Isaac, Shenghui Wang and the Telplus project for providing the library track data and their mappings, and for the many explanations they provided. We also thank the Cornetto project for providing data and mappings and Babette Heyer for helping with the evaluation of the mappings. This work was supported in part by the National Center for Biomedical Ontology, under roadmap-initiative grant U54 HG004028 from the National Institutes of Health, and partly by the EC eContentplus programme via de EuropeanaConnect project.

References

1. K. Aberer, P. Cudré-Mauroux, and M. Hauswirth. A framework for semantic gossiping. *SIGMOD Rec.*, 31(4):48–53, 2002.
2. Z. Aleksovski, M. C. A. Klein, W. ten Kate, and F. van Harmelen. Matching unstructured vocabularies using a background ontology. In *EKAW*, pages 182–197, 2006.
3. D. Aumueller, H.-H. Do, S. Massmann, and E. Rahm. Schema and ontology matching with coma++. In *SIGMOD '05: Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pages 906–908. ACM, 2005.
4. P. A. Bernstein, T. J. Green, S. Melnik, and A. Nash. Implementing mapping composition. In *VLDB '06: Proceedings of the 32nd international conference on Very large data bases*, pages 55–66. VLDB Endowment, 2006.
5. H.-H. Do and E. Rahm. Coma: a system for flexible combination of schema matching approaches. In *VLDB '02: Proceedings of the 28th international conference on Very Large Data Bases*, pages 610–621. VLDB Endowment, 2002.

6. J. Euzenat. Algebras of ontology alignment relations. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *International Semantic Web Conference*, volume 5318 of *Lecture Notes in Computer Science*, pages 387–402. Springer, 2008.
7. J. Euzenat, A. Ferrara, L. Hollink, A. Isaac, C. Joslyn, V. Malaisé, C. Meilicke, A. Nikolov, J. Pane, M. Sabou, F. Scharffe, P. Shvaiko, V. Spiliopoulos, H. Stuckenschmidt, O. Sváb-Zamazal, V. Svátek, C. T. dos Santos, G. A. Vouros, and S. Wang. Results of the ontology alignment evaluation initiative 2009. In *OM*, 2009.
8. A. Ghazvinian, N. F. Noy, C. Jonquet, N. Shah, and M. A. Musen. What four million mappings can tell you about two hundred ontologies. In *ISWC '09: Proceedings of the 8th International Semantic Web Conference*, pages 229–242. Springer-Verlag, 2009.
9. A. Ghazvinian, N. F. Noy, and M. A. Musen. Creating mappings for ontologies in biomedicine: Simple methods work. In *AMIA Annual Symposium (AMIA 2009)*, San Francisco, CA, 2009.
10. W. V. Hage, A. Isaac, and Z. Aleksovski. Sample evaluation of ontology-matching systems. *Proceedings of the ISWC workshop on Evaluation of Ontologies and Ontology-based tools*, pages 41–50, 2007.
11. W. Hu and Y. Qu. Falcon-AO: A practical ontology matching system. *Web Semant.*, 6(3):237–239, 2008.
12. Y. R. Jean-Mary, E. P. Shironoshita, and M. R. Kabuka. ASMOV: Results for OAEI 2009. In *OM*, 2009.
13. P. Lambrix, H. Tan, and Q. L. 0002. Sambo and sambodtf results for the ontology alignment evaluation initiative 2008. In *OM*, 2008.
14. J. Madhavan and A. Y. Halevy. Composing mappings among data sources. In *VLDB '2003: Proceedings of the 29th international conference on Very large data bases*, pages 572–583. VLDB Endowment, 2003.
15. A. Miles and S. Bechhofer. Skos simple knowledge organization system reference, August 2009.
16. M. Nagy, M. Vargas-Vera, and P. Stolarski. Dssim results for oaei 2009. In *OM*, 2009.
17. N. F. Noy, N. H. Shah, P. L. Whetzel, B. Dai, M. Dorf, N. Griffith, C. Jonquet, D. L. Rubin, M.-A. D. Storey, C. G. Chute, and M. A. Musen. Bioportal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Research*, 37(Web-Server-Issue):170–173, 2009.
18. A. Tordai, J. van Ossenbruggen, and G. Schreiber. Combining vocabulary alignment techniques. In *K-CAP '09: Proceedings of the fifth international conference on Knowledge capture*, pages 25–32, New York, NY, USA, 2009. ACM.
19. A. Tordai, J. van Ossenbruggen, G. Schreiber, and B. Wielinga. Aligning large SKOS-like vocabularies: Two case studies. In *The Semantic Web: Research and Applications*, volume 5554 of *Lecture Notes in Computer Science*, pages 198–212. Springer Berlin / Heidelberg, 2010.
20. P. Vossen, I. Maks, R. Segers, and H. van der Vliet. Integrating lexical units, synsets and ontology in the Cornetto database. In E. L. R. A. (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, 2008.
21. S. Wang, A. Isaac, B. A. C. Schopman, S. Schlobach, and L. van der Meij. Matching multilingual subject vocabularies. In M. Agosti, J. L. Borbinha, S. Kapidakis, C. Papatheodorou, and G. Tsakonas, editors, *ECDL*, volume 5714 of *Lecture Notes in Computer Science*, pages 125–137. Springer, 2009.
22. X. Zhang, Q. Zhong, F. Shi, J. Li, and J. Tang. Rimom results for oaei 2009. In *OM*, 2009.