Hybrid Alignment Strategy for Anatomical Ontologies Results of the 2007 Ontology Alignment Contest

Songmao Zhang[1], Olivier Bodenreider[2]

[1] Institute of Mathematics, Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing, P. R. China
[2] U.S. National Library of Medicine, NIH, Bethesda, Maryland, USA
[1] smzhang@math.ac.cn, [2] olivier@nlm.nih.gov

Abstract. An ontology is a formal representation of a domain modeling the entities in the domain and their relations. When a domain is represented by multiple ontologies, there is a need for creating mappings among these ontologies in order to facilitate the integration of data annotated with and reasoning across these ontologies. The objective of this paper is to present our experience in aligning two medium-size anatomical ontologies and to reflect on some of the issues and challenges encountered along the way. The anatomical ontologies under investigation are the Adult Mouse Anatomy (MA) and the anatomy subset of the NCI Thesaurus (NCI). We also use the Foundational Model of Anatomy as a reference ontology. We present a hybrid alignment strategy for anatomical entities, combining direct and indirect alignment techniques, both supported by the NLM Anatomy Ontology Alignment System (AOAS). Overall, the hybrid strategy combining direct and indirect alignment techniques identified 1,338 matches between MA and NCI, accounting for about 49% of all MA concepts and 41% of all NCI concepts. 1,007 matches are shared by both alignments, leaving 277 matches specific to the direct alignment and 54 specific to the indirect alignment.

1   Presentation of the system

An ontology is a formal representation of a domain modeling the entities in the domain and their relations. Many domains, including anatomy, are represented by multiple ontologies, with variable overlap among them. There is a need for creating mappings among these ontologies in order to facilitate the integration of data annotated with and reasoning across these ontologies [1].

1.1   Introduction

Over the past five years, as part of the Medical Ontology Research project at the U.S. National Library of Medicine, we have developed domain knowledge-based techniques for aligning large anatomical ontologies, with the objective of exploring approaches to aligning representations of anatomy differing in formalism, structure, and domain coverage [2]. In this effort, we aligned the two anatomical ontologies

under investigation in the 2007 OAEI campaign, namely, the Adult Mouse Anatomy (MA) [3] and the anatomy subset of the NCI Thesaurus (NCI) [4]. More precisely, we explored two distinct approaches to aligning these two ontologies. The first approach is a direct alignment realized with the NLM Anatomy Ontology Alignment System (AOAS) [5]. The second approach is an indirect alignment through a reference ontology of anatomy: the Foundational Model of Anatomy (FMA) [6]. In the current study, we then combine these two alignment techniques into a hybrid strategy. To our knowledge, the combination of direct and indirect approaches represents a novel strategy for aligning ontologies.

1.2   Hybrid alignment strategy

The hybrid strategy for aligning anatomical entities combines direct and indirect alignment techniques, both supported by the NLM Anatomy Ontology Alignment System (AOAS).

Direct alignment

The direct alignment technique consists in the identification of one-to-one concept mappings between the MA and NCI using lexical resemblance between concept names, followed by the validation of the mappings through shared hierarchical paths among concepts across ontologies. A brief presentation of the method is given below. The interested reader is referred to [2] for further details.

Identifying matches lexically. The lexical alignment compares two ontologies at the term level, by exact match and after normalization. This process makes the source and target terms potentially compatible by eliminating such inessential differences as inflection, case, hyphen, and word-order variation. Both preferred terms and synonyms in the two ontologies are used in the alignment process. Moreover, synonymy in the Unified Medical Language System (UMLS) Metathesaurus is used to identify additional matches. For example, *Profunda femoris artery* in MA and *Deep femoral artery* in NCI, although lexically different, are considered as a match because they name the same anatomical concept in the UMLS. Our method does not address partial lexical matches.

Validating matches structurally. In order to facilitate the comparison of relations across ontologies, the structural alignment first normalizes the hierarchical relations among concepts, including IS-A and PART-OF. Missing inverse relations are complemented as necessary. Inference rules are used to generate a partitive relation between a specialized part and the whole or between a part and a more generic whole. Once all relations are represented consistently, the structural alignment is applied on the matches resulting from the lexical alignment in order to identify similar relations to other matches across ontologies (i.e., shared hierarchical paths). For example, the matching concepts *Forelimb* in MA and *Upper extremity* in NCI exhibit similar relations to other matches in the two ontologies, including *Limb* (through IS-A), *Arm* and *Hand* (through HAS-PART) across ontologies. Such structural similarity is used as positive evidence for the alignment. Instead of similar relations, one match may

exhibit relations to other matches in opposite directions in the two ontologies. Such relations suggest a structural conflict across ontologies. For example, in MA *Pericardial cavity* is in HAS-PART relationship to *Pericardium*, while in NCI *Pericardial cavity* is defined as part of *Pericardium*. These conflicts are used as negative evidence for the alignment, indicating the semantic incompatibility between concepts across ontologies despite their lexical resemblance. In some cases, no evidence (positive or negative) is found to support or reject the match.

Indirect alignment

An alternative to aligning MA and NCI concepts directly consists in the identification of mappings through a reference ontology, here, the Foundational Model of Anatomy (FMA). In practice, the following method was used for automatically deriving a mapping between MA and NCI from the two direct alignments MA-FMA and NCI-FMA. When a FMA concept C_F is aligned with both a MA concept ({MA: C_M, FMA: C_F}) and a NCI concept ({NCI: C_N, FMA: C_F}), the concepts C_M and C_N are automatically aligned ({MA: C_M, NCI: C_N}).

For example, as shown in **Figure 1**, the direct alignment MA-FMA identifies the match {MA: *Forelimb*, FMA: *Upper limb* (synonym: *Forelimb*)}, which is supported by positive evidence. The direct alignment NCI-FMA identifies the match {NCI: *Upper extremity*, FMA: *Upper limb* (synonym: *Upper extremity*)}, also supported by positive evidence. Therefore, the match {MA: *Forelimb*, NCI: *Upper extremity*} is derived automatically, through the FMA concept *Upper limb*, supported by positive structural evidence in both direct alignments.

The direct alignment method between MA and FMA (and between NCI and FMA) follows the same steps described above for the direct alignment between MA and NCI. Additional knowledge augmentation techniques are used to acquire and normalize relations from the FMA, including the extraction of partitive reified relations from terms (e.g., <*Heel*, PART-OF, *Foot*> was derived from <*Heel*, IS-A, *Subdivision of foot*>).

In the indirect alignment through a reference ontology, where an indirect match (e.g., between C_M and C_N) results from two direct matches to the reference (e.g., between C_M and C_F and between C_N and C_F), the evidence for the indirect match is the combination of the evidence for the two direct matches to the reference. Intuitively, the strongest mappings correspond to those cases supported by positive evidence in both direct alignments. Weaker mappings are identified when positive evidence is found in only one of the two direct alignments and no evidence is found in the other. The weakest mappings correspond to cases where no evidence is found in either direct alignment. Finally, mappings exhibiting negative evidence in either direct alignment are rejected.
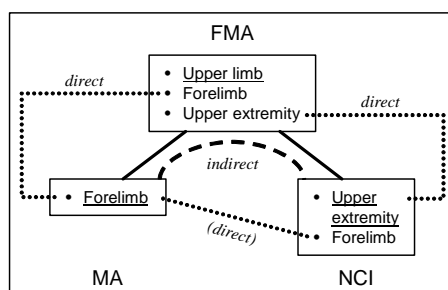
Figure 1. Indirect MA-NCI alignment through FMA

Combining direct and indirect alignment techniques

Although most mappings are identified by both direct and indirect alignments, we showed that each technique also produced specific mappings. Therefore, we propose to combine the mappings yielded by the two alignment techniques. In practice, a mapping can be identified by either technique or by both techniques. Moreover, the degrees of confidence attached to each mapping in the direct and indirect alignments can also be combined, resulting in a composite score reflecting the overall degree of confidence in the mapping.

In the direct alignment, direct matches can be supported by positive structural evidence, negative structural evidence, or no evidence. In the indirect alignment, as mentioned earlier, the degree of confidence in the matches goes from strong positive evidence (in both direct alignments), weak positive evidence (in only one direct alignment), no evidence (in either alignment). In all cases, the presence of negative evidence in either alignment causes the mapping to be rejected.

Intuitively, for the combined alignment strategy, the strongest mappings correspond to matches supported by positive evidence in both direct and indirect alignments. The presence of negative evidence in either alignment causes the mapping to be rejected. Otherwise, an additive model is used to combine degrees of evidence. The credit given to those matches specific to either direct or indirect alignment corresponds to one half of what they would receive if confirmed in the other alignment. The numeric scores reported in the results for the various combinations are summarized in **Table 1**.

Table 1. Degree of confidence for various combinations of support in direct and indirect alignments

		Indirect alignment				
		Strong positive evidence	Weak positive evidence	No evidence	Negative evidence	No match
Direct alignment	Positive evidence	1.0	.875	.75	0	.50
	No evidence	.75	.625	.50	0	.25
	Negative evidence	0	0	0	0	0
	No match	.50	.375	.25	0	

1.3 Adaptations made for the evaluation

In previous alignment experiments, we purposely stayed away form any particular formalism in order to avoid overfitting tools and techniques to a given formalism. Since our alignment approach is not designed to specifically take advantage of OWL, we simply extracted concept names and <concept, relationship, concept> triples from the class definitions in OWL. In addition to the files provided for MA and NCI (in OWL DL), we used the file provided last year in the OAEI 2006 campaign for the FMA (in OWL Full) as our reference ontology.

In practice, we used **rdf:ID** to identify concepts, **rdf:label** to acquire concept names, **oboInOwl:hasRelatedSynonym** to acquire synonyms, and **rdfs:subClassOf** to acquire taxonomic relations. The various partitive relations represented in the FMA (e.g., *part_of*, *constitutional_part_of*, *regional_part_of*) were acquired using the corresponding properties and merged – for alignment purposes – with *UNDEFINED_part_of* in MA and NCI. The other properties (e.g., **oboInOwl:Definition**) were not used in the alignment.

1.5 Link to the set of provided alignments (in align format)

The result of our alignment for the 'anatomy' data set is available at: http://mor.nlm.nih.gov/pubs/supp/2007-oaei-sz/Zhang&Bodenreider.rdf in the format specified by the OAEI organizers at: http://oaei.ontologymatching.org/2007/.

2 Results

The only data set for which we report results is 'anatomy'. Overall, the hybrid strategy combining direct and indirect alignment techniques identified 1,338 matches between MA and NCI, accounting for about 49% of all MA concepts and 41% of all NCI concepts.

Acquiring concept names and relations. The main characteristics of the data sets under investigation are listed in **Table 2**, including the number of classes, concept names, and types of partitive relationships. The number of *IS-A* and partitive relations extracted from the OWL file and generated by complementation, augmentation and inference is shown in **Table 3**. Not surprisingly, in the three ontologies, a majority of relations come from inference, which performs similarly to a transitive closure of the hierarchical relations. Also listed in **Table 3** is the small number of relations removed from the FMA because they create cycles.

Table 2. Main characteristics of the three ontologies

#	MA	NCI	FMA
Concepts	2,738	3,298	72,560
Normalization of preferred terms	3,024	3,586	98,942
Synonyms	324	1,796	44,597
Normalization of synonyms	370	1,947	60,051
Part-of relationships	1	1	7

Table 3. Number of relations in the three ontologies

Types of relations	MA	NCI	FMA
Explicitly represented is-a	2,857	3,761	72,560
Explicitly represented partitive relations	1,631	1,662	101,161
Explicitly represented associative relations	0	0	48,804
Complemented inverse-isa	2,857	3,761	72,560
Complemented partitive relations	1,631	1,662	3,561
Complemented associative relations	0	0	11,697
Removed because of cycles	0	0	- 40
Augmented	0	0	169,378
Inferred	23,504	35,092	5,169,034
Total	32,480	45,938	5,648,715

Three direct alignments. Results for three direct alignments are summarized in **Table 4**. The alignment NCI-FMA yielded the largest number of matches (2,314) and MA-NCI the smallest (1,284). A very small number of conflicts (matches exhibiting negative evidence) was identified in each direct alignment. In the three direct alignments, a vast majority of the matches (> 96%) was supported by positive structural evidence. No evidence (positive or negative) was found for 2-4% of the matches in three direct alignments. For example, although *Elbow joint* has relations to other matches in both MA (e.g., *PART-OF Forelimb*) and NCI (e.g., *PART-OF Skeletal system*), none of these relations are shared.

Table 4. Three direct alignments

	MA - NCI 1,284 matches	MA - FMA 1,562 matches	NCI - FMA 2,314 matches
No evidence	25 (1.9%)	49 (3.1%)	85 (3.7%)
Positive evidence	1,254 (97.7%)	1,507 (96.5%)	2,215 (95.7%)
Negative evidence	5 (0.4%)	6 (0.4%)	14 (0.6%)

Indirect alignment MA-NCI through FMA. 1,061 matches between MA and NCI were automatically derived from the 1,562 matches in the direct alignment MA-FMA and the 2,314 matches in NCI-FMA. 1,008 of them (95%) received positive structural evidence in both direct alignments MA-FMA and NCI-FMA, 15 (1.4%) received negative evidence in one of the two direct alignments, and 38 (3.6%) received no evidence in at least one of the two direct alignments.

Combining direct and indirect alignments. Of the 1,284 matches obtained by direct alignment and 1,061 matches derived from the indirect alignment through the FMA, 1,007 matches are shared by both alignments, leaving 277 matches specific to the direct alignment and 54 specific to the indirect alignment. The repartition of the matches with respect to the degree of confidence is presented in **Figure 2**.
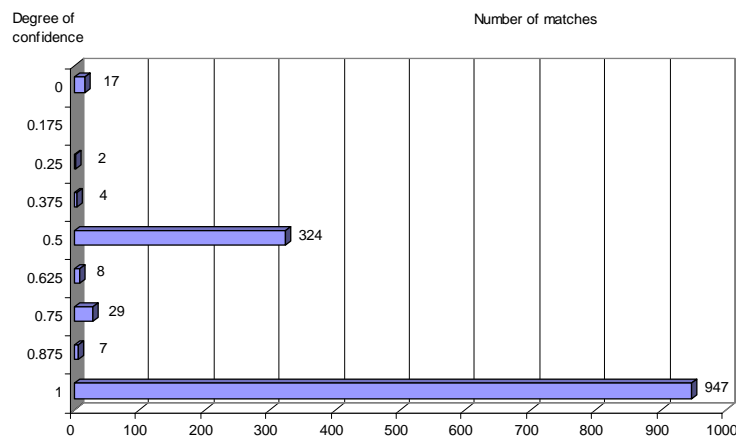

Figure 2. Repartition of the matches with respect to the degree of confidence

Matches in common. Of the 1,007 matches common to both alignments, 947 (94%) received positive evidence in the direct alignment and strong positive evidence in the indirect alignment. 15 matches (1.5%) received negative evidence in at least one of the alignments and were therefore eliminated. All but one of the remaining matches received positive evidence in one alignment and no evidence in the other. One match did not receive any evidence in either alignment. For example, the match {MA: *Forelimb*, NCI: *Upper Extremity*}is common to both alignments and receives positive evidence in the direct alignment and strong positive evidence in the indirect alignment.

Matches specific to the direct alignment. The direct alignment yielded 277 matches that were not identified by the indirect alignment. All but one received positive evidence and one received no evidence. For example, the match {MA: *tibial artery*, NCI: *Tibial Artery*} is specific to the direct alignment and receives positive evidence.

Matches specific to the indirect alignment. The indirect alignment yielded 54 matches that were not identified by the direct alignment. A majority of them (51 or 94%) received positive evidence (weak in 4 cases) and one received no evidence. Two matches received negative evidence and were eliminated. For example, the match {MA: *ovary follicle*, NCI: *Ovarian Follicle*} is specific to the indirect alignment (through the FMA concept *Follicle of ovary*) and receives strong positive evidence.

3   General comments

3.1   Comments on the results

The objective of combining several approaches is to increase recall and precision over each technique used in isolation. We note a significant increase in recall compared to the indirect alignment (1,338/1,061 or +26%) and a more modest increase compared to the direct alignment (1,338/1,284 or +4%). From a qualitative perspective, although most matches are common to both approaches, each technique contributes a number of specific matches.

In terms of precision, the use of two independent techniques represents a form of cross-validation of the matches. While insufficient for a completely automated mapping strategy, this validation method can help establish a degree of confidence in the matches, based on corroborated findings across alignment techniques. This degree of confidence can be used to guide the effort of manually curating the matches, by targeting those matches identified with a lesser confidence.

The scoring strategy presented in this paper (**Table 1**) also offers a more precise framework for estimating precision. In fact, while most alignment systems use some kind of arbitrary threshold over a continuous score between 0 and 1, the scheme we use for scoring the quality of the mappings is totally transparent and mathematically sound. Unlike most systems, the scoring system based on our hybrid approach attaches an explanation to each score (e.g., .75 consistently means that the lexical match is supported by positive evidence in one alignment, but uncorroborated in the other). The scoring scheme is also mathematically sound as it simply averages the scores from both direct and indirect alignments, except in the case where negative evidence is found, in which case the matches are excluded.

3.2   Discussion on the way to improve the proposed system

The strengths and weaknesses of our system have been analyzed in previous papers [7]. The major difference with other systems is that we take advantage of domain knowledge throughout the mapping process. For example, we use specific tools and resources, including normalization techniques developed for biomedical terms and synonyms from the Unified Medical Language System. We also developed techniques specific to the anatomical ontologies under investigation in order to represent explicitly relations implicitly present in these ontologies. These additional synonyms and relations increase the chances of identifying matches both at the lexical and structural level.

In a recent analysis of the matches produced by several systems for the alignment between FMA and GALEN in the 2006 OAEI campaign [8], we noted a small number of false negatives in our system. Alignment systems such as FALCON [9] and PRIOR [10] are based on an information retrieval paradigm and accommodate partial matches. Thus, they tend to handle gracefully the kinds of spelling variation (and sometimes misspelling) encountered in some anatomical ontologies. In contrast, with a stricter, domain-specific model of lexical similarity, our system has a better precision, but is more sensitive to missing synonyms and misspellings.

As mentioned before, our alignment strategy does not take advantage of the textual definitions available for most concepts in NCI. Textual definitions are not systematically present in anatomical ontologies and require natural language processing to be meaningfully interpreted. For these two reasons, we did not use them, nor do we have any plans to do so in the future.

3.3   Comments on the OAEI 2007 procedure

Unlike the FMA and GALEN in the 2006 OAEI campaign, provided in OWL Full after conversion from their native environments (Protégé-frames for the FMA and GRAIL for GALEN), MA and NCI are provided this year in OWL DL, which is the native format for NCI and an easy conversion from the Open Biological Ontology (OBO) format for MA. This simpler formalism, along with the smaller size and limited complexity of MA and NCI compared to FMA and GALEN, made the 2007 OAEI campaign for anatomy more interesting as it made it possible for the participants to focus on alignment issues.

Moreover, there exists a gold standard mapping between MA and NCI developed by biologists at the Jackson Laboratory. This will be the first time in the young history of the OAEI campaign that such a reference mapping is provided for anatomical ontologies. Although limited to equivalent concepts, this mapping will make the evaluation much more meaningful as we demonstrated that mere cross-validation among alignment systems was not adequate for evaluation purposes [8].

3.4 Comments on the OAEI 2007 test cases

Every year, the anatomy test case is one of the most challenging test cases in the OAEI campaign, generally because of the sheer size of the ontologies to be aligned, as well as the level of domain expertise required. Nonetheless, the anatomy task has attracted a growing number of participants over the years. However, we have shown that the performance of most participating systems, usually general-purpose, domain-independent alignment systems applied to the anatomy task, is generally mediocre [8]. Beside absolute performance, it is interesting to see the progress made by some systems (e.g., FALCON), due in part to the stimulation generated by the OAEI campaign and enabled by the availability of large test datasets.

4 Conclusion

Anatomy is central to biomedicine and a key element to translational medicine, i.e., the effective exchange of information between the "bench" (basic research) and the "bedside" (clinical practice). While aligning large anatomical ontologies remains challenging, because of the sheer size of these resources and the need for domain knowledge, the progress made in the past years is encouraging. By providing a forum for comparing methods and datasets for evaluation purposes, the OAEI campaign has contributed significantly to this progress.

References

1. Euzenat, J., Shvaiko, P.: Ontology matching. Springer, New York (2007)
2. Zhang, S., Bodenreider, O.: Experience in aligning anatomical ontologies. International Journal on Semantic Web and Information Systems **3** (2007) 1-26
3. Hayamizu, T.F., Mangan, M., Corradi, J.P., Kadin, J.A., Ringwald, M.: The Adult Mouse Anatomical Dictionary: a tool for annotating and integrating data. Genome Biology **6** (2005) R29
4. Sioutos, N., Coronado, S.d., Haber, M.W., Hartel, F.W., Shaiu, W.L., Wright, L.W.: NCI Thesaurus: A semantic model integrating cancer-related clinical and molecular information. Journal of Biomedical Informatics **40** (2007) 30-43

5. Bodenreider, O., Hayamizu, T.F., Ringwald, M., de Coronado, S., Zhang, S.: Of mice and men: Aligning mouse and human anatomies. Proc AMIA Symp (2005) 61-65

6. Zhang, S., Bodenreider, O.: Alignment of multiple ontologies of anatomy: Deriving indirect mappings from direct mappings to a reference. Proc AMIA Symp (2005) 864-868

7. Zhang, S., Mork, P., Bodenreider, O.: Lessons learned from aligning two representations of anatomy. In: Hahn, U., Schulz, S., Cornet, R. (eds.): Proceedings of the First International Workshop on Formal Biomedical Knowledge Representation (KR-MED 2004) (2004) 102-108

8. Zhang, S., Bodenreider, O.: Lessons learned from cross-validating alignments between large anatomical ontologies. Medinfo (2007) 822-826

9. Hu, W., Cheng, G., Zheng, D., Zhong, X., Qu, Y.: The results of Falcon-AO in the OAEI 2006 Campaign. Proceedings of the Ontology Alignment Evaluation Initiative 2006 Campaign (OAEI 2006) (2006) 118-125

10. Mao, M., Peng, Y.: PRIOR system: Results for OAEI 2006. Proceedings of the Ontology Alignment Evaluation Initiative 2006 Campaign (OAEI 2006) (2006) 165-172