A Semi-Automated Approach for Anatomical Ontology Mapping

Peter Petrov[1], Milko Krachunov[1] and Dimitar Vassilev[2,*]

[1]Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski", 5 James Bourchier Blvd., 1164 Sofia, Bulgaria, http://fmi.uni-sofia.bg/

[2]Bioinformatics group, Agro Bio Institute, 8 Dragan Tsankov Blvd., 1164 Sofia, Bulgaria, http://abi.bg/

Summary

This paper presents a study in the domain of semi-automated and fully-automated ontology mapping. A process for inferring additional cross-ontology links within the domain of anatomical ontologies is presented and evaluated on pairs from three model organisms. The results of experiments performed with various external knowledge sources and scoring schemes are discussed.

1 Introduction

Ontologies have recently gained popularity because they help with interoperability, information/knowledge sharing and knowledge reuse. Information sources, ontologies in particular, are often heterogeneous even when they originate in the same problem domain. In order to enable compatibility between such ontologies and to integrate the knowledge from multiple such information sources, one needs to build mappings between them. These mappings establish the semantic correspondence between concepts and relations in different ontologies.

As it is noted in [1] there are some terminological differences pertaining to the integration of ontologies within the ontology mapping/merging/matching (OM) community. Those terminological differences are mostly between the terminology adopted by J. de Bruijn et al. in [2] on one side and by J. Euzenat and P. Shvaiko in [3] on the other. In our works, we adopt the terminology used in [2]. In the sense of [2], *ontology mapping* is the process of taking two input ontologies and generating semantic links between their concepts/terms. The generated links are not part of the two input ontologies; they are stored separately from them. Two other terms are related to ontology mapping: *ontology aligning* and *ontology merging*.

Ontology aligning [2] can be viewed as an automatic or semi-automatic ontology mapping; it denotes the process of discovery of cross-ontology links by a computer program. Again, these links are stored separately of the two input ontologies. Ontology merging [2] is the ultimate goal when integrating/mediating two input ontologies; it consists of taking two input ontologies and generating an output ontology that unifies the knowledge contained in them. It is usually a process which follows those of aligning and mapping, and which utilizes the intermediate results produced by them. During this process, some term pairs from different ontologies are

*To whom correspondence should be addressed. Email: jim6329@gmail.com

merged into single nodes of the output ontology, while the other terms of the input ontologies are left unchanged in the output ontology.

This paper discusses a set of procedures pertaining to the processes of automatic aligning and mapping of species-specific anatomical ontologies by utilization of different knowledge sources.

2 Problem description

Given a pair of anatomical ontologies of different species (model organisms), e.g. *mouse* and *zebrafish*, the goal is to establish semantic links between the terms of the two such that: (i) these links are of one of the following types; R_1**=synonymy,** R_2**=hypernymy,** R_3**=hyponymy,** R_4**=holonymy,** R_5**=meronymy**, and (ii) each of these links has some *degree of certainty* or *degree of confidence* or *confidence score* which is a real number in the interval $[0, 1]$. The semantic relation types R_k that we refer to here are well-known and are widely utilized in the areas of linguistics, knowledge representation and ontology engineering; that's why we don't provide any formal or informal definitions for them here.

The two input ontologies are represented in the form of OBO files. OBO stands for "Open Biomedical Ontology" and denotes an *ontology language* or an *ontology file format* [4] for defining ontologies used mostly in the biomedical domain. Nowadays OBO is adopted by the GO project [4, 5], the OBO Foundry initiative [6], and other communities.

For the discovery of those semantic links we utilize more general external ontologies and other vocabularies, such as UMLS [7, 8], FMA [9, 10] and WordNet [11, 12, 13]. We additionally attempt to utilize the structure of the ontologies to enrich and reaffirm the predicted semantic links.

3 Formalization of the problem

In mathematical terms, each of the two input anatomical ontologies can be represented, in whole or in part, as a directed acyclic graph (DAG) with a coloring function on its edges. The colors represent the relations defined within the input ontologies (e.g. ***is_a*** and ***part_of***) which we call ***inner-ontology*** relations. Typically, there are other inner-ontology relations except those two, but those additional relations usually pertain to the development of the particular organism and not just to its anatomy. Examples include ***start_stage, end_stage*** and ***develops_from.*** Practically, we don't deal with them as we are mainly concerned with the adult or gross anatomy of an organism, not with its growth and development. However, should the need arise, this model can be easily generalized to cover these relations as well.

We will use the following notation for the graph model:

$$\begin{aligned} O_1 : & DAG_1 = (V_1, E_1); \\ & F_1 : E_1 \to C = \{c_1, c_2, \ldots, c_n\} \\ O_2 : & DAG_2 = (V_2, E_2); \\ & F_2 : E_2 \to C = \{c_1, c_2, \ldots, c_n\} \end{aligned} \tag{1}$$

Here O_1 and O_2 are the two anatomical ontologies; DAG_1 and DAG_2 are their corresponding ***directed acyclic graphs***; V_1 and V_2 are the ***sets of terms*** of the ontologies (each with an identifier and a name); E_1 and E_2 are are the ***relations*** defined within the ontologies; F_1 and F_2 are the ***edge-coloring functions***. The relations ***is_a*** (specialization) and ***part_of*** (aggregation) are the two typical examples of inner-ontology relations defined within the ontologies O_1 and O_2. Since we deal only with those two, we can assume that within our notation $n = 2$, $c_1 =$ *is_a* and $c_2 =$ *part_of*. Thus if for example, $u_1 =$ *'brain'* $\in V_1$, $u_2 =$ *'central nervous system'* $\in V_1$, then there usually exists an edge e between u_1 and u_2 such that $F_1(e) =$ *part_of* because the *brain* is a part of the *central nervous system*, and anatomical ontologies, regardless of the subject organism, usually declare this fact explicitly.

We will also discuss external knowledge sources, both biomedical and general-purpose ones, that contain anatomical terms and relations (***is_a***, ***part_of*** or others). Three concrete external knowledge sources have been used for the purposes of this work. These are $T_1 =$ *UMLS*, $T_2 =$ *FMA*, $T_3 =$ *WordNet*. UMLS and FMA are biomedical knowledge sources, while WordNet is a general purpose knowledge source. Formally stated, each of these knowledge sources $T_s, s \in \{1, 2, 3\}$ contains the following information:

- **Terms.** These is the set terms M_s defined within the external knowledge source T_s:

$$M_s = \{t_{s1}, t_{s2}, \ldots, t_{sm_s}\} \tag{2}$$

  Here $t_{sk} = (id_{sk}; name_{sk})$; id_{sk} is the identifier within T_s of the term t_{sk}; $name_{sk}$ is the textual name of the term t_{sk}; m_s is the number of terms in the knowledge source T_s.

- **Relations.** These are the sets of ***is_a*** and ***part_of*** relations defined within the external knowledge source T_s:

$$\begin{aligned} R'_{T_s} &= R^{is\_a}_{T_s} \subseteq M_s \times M_s \\ R''_{T_s} &= R^{part\_of}_{T_s} \subseteq M_s \times M_s \end{aligned} \tag{3}$$

Just like with the input ontologies, more relations can be found in each external knowledge source, but we utilize only these two. They are also the only two present in the intersection of all external knowledge sources and the ontologies; they are common in unspecialized vocabularies.

Each knowledge source $src = T_s, s \in \{1, 2, 3\}$ is assigned a fixed score $f(src)$ which denotes its precision in predicting synonymy and parent-child (***is_a, part_of***) relations between terms of the two input ontologies. Details on the evaluation of the precision of these three knowledge sources can be found in [14].

Using this notation, we can describe the set of predictions (represented by 4-tuples) that we seek to generate in our algorithmic procedures as following:

$$D = \{(v_{1k}, v_{2k}, r_k, s_k) | k = 1, 2, \ldots, |D|\} \tag{4}$$

where $v_{1k} \in V_1$, $r_k \in \{R_1, R_2, R_3, R_4, R_5\}$, and $s_k \in (0, 1]$. Here, for each k, v_{1k} is a term from the input ontology O_1, v_{2k} is a term from the input ontology O_2, r_k is an automatically predicted cross-ontology relation from one of the five types defined in the previous section, and s_k is a real number denoting the confidence score of the prediction that the terms v_{1k} and

v_{2k} are related/linked by a cross-ontology link of the type r_k. Requiring that $s_k \in (0, 1]$ we basically imply that the set D which we seek, is in fact a set of *cross-ontology predictions* or a *set of predicted cross-ontology links* between O_1 and O_2 in which each score is probability-based. Given certainty in the input and validity of the external knowledge source evaluation, the final score calculated in our scoring procedures can be used to model the probability that the corresponding prediction is actually true.

4 Algorithmic procedures

Three algorithmic procedures are applied to the graph structures described in the previous section to discover links for our set D that is being sought. These three procedures are described in more detail in [15], here a brief summary is given.

Within the first procedure called ***direct matching*** (DM), the two input ontologies are scanned for identity matches between the names of their terms. If $t_1 \in V_1$ and $t_2 \in V_2$ have the same names, they are marked as synonyms (R_1). The cross-ontology links discovered this way are assigned the highest possible score of 1.0 as these predictions come from information contained entirely in the two input ontologies and make the external knowledge sources redundant, although we have conducted experiments with lower score values.

During the second procedure called ***source matching predictions*** (SMP), we use the relations in the external knowledge sources together with the identity matches between term names of the two input ontologies and the knowledge sources to build a graph structure which aligns each of the two input ontologies to each knowledge source. The model contains a set of semantic links (of the types R_k, $k = 1, 2, \ldots, 5$ that were defined above) between the two input ontologies on the one side, and the three external knowledge sources on the other side. Then a set of logical rules is applied, and conclusions are drawn for the semantic relations that exist between terms $t_1 \in V_1$ and $t_2 \in V_2$ of the two input ontologies. The following rules are applied at this stage:

- *Rule (A).* If two terms $t_1 \in V_1$ and $t_2 \in V_2$ have been detected as synonyms of the same term $t \in T_s$, then t_1 and t_2 are marked as predicted cross-ontology synonyms of each other.
- *Rule (B).* If $t_j \in V_j$ has been detected as a synonym of $t \in T_s, s = 1, 2, 3$, and if the term $t_{3-j} \in V_{3-j}$ has been detected as an ***is_a/part_of*** child/parent of t, then t_j is marked as predicted a cross-ontology ***is_a/part_of*** child/parent of t_{3-j} (here $j = 1$ or 2 and respectively $3 - j = 2$ or 1).

Rule (A), when applied, finds the synonymy relations (i.e. the relations of type R_1) between terms from the two input ontologies. Rule (B) is a composite version of four separate rules (two options for ***is_a/part_of*** and two options for child/parent makes four options in total). These four rules which originate from rule (B), when applied, find the hypernymy, hyponymy, holonymy, and meronymy relations (i.e. the relations of types R_2, R_3, R_4, R_5) between terms of the two input ontologies. All links predicted through SMP are given a score $f(src)$ where src is the knowledge source implying these predictions.

Finally, we run a procedure that we denote as the ***child matching predictions*** (CMP) procedure. This one tries to find R_1, R_2, R_3, R_4 and R_5 links between terms of the two input ontologies,

$t_1 \in V_1$ and $t_2 \in V_2$, for which no links have been predicted either by DM or by SMP. The CMP approach is to consider patterns of cross-ontology connectivity (found by DM and SMP) between $t_1 \in V_1$ (parent term 1), $t_2 \in V_2$ (parent term 2), and the inner-ontology children of the two terms t_1 and t_2. Three separate patterns of connectivity are considered by CMP:

(i) $t_1 \in V_1 \leftarrow t_{ch1} \in V_1 \leftrightarrow t_{ch2} \in V_2 \rightarrow t_2 \in V_2$ (we call this a **U-pattern**)

(ii) $t_1 \in V_1 \leftarrow t_{ch2} \in V_2 \leftrightarrow t_{ch1} \in V_1 \rightarrow t_2 \in V_2$ (we call this an **X-pattern**)

(iii) $t_1 \in V_1 \leftarrow t_{ch1} \in V_1 \rightarrow t_2 \in V_2$ or
$t_1 \in V_1 \leftarrow t_{ch2} \in V_2 \rightarrow t_2 \in V_2$ (we call this a **V-pattern**)

In this notation, the $\rightarrow$ and $\leftarrow$ arrows denote sets of non-CMP parent-child links (the arrows always point from child to parent), these are asymmetrical links; the $\leftrightarrow$ arrows denote sets of non-CMP synonymy links, these are symmetrical links. Each occurrence of any of these patterns between t_1 and t_2 (the two parent terms) we call a ***pattern instance***. It must be noted that all arrows within a pattern instance represent either ***is_a*** or ***part_of*** links (i.e. we don't allow mixing these two within a single pattern instance).

Based on these patterns of connectivity, new cross-ontology links (CMP links) are introduced (one CMP link per pattern instance) between t_1 and t_2. We call these links ***individual CMP links***. To assign scores to the individual CMP links, the concepts ***score of a set of non-CMP links between two terms*** and ***score of a pattern instance*** (or ***score of an individual CMP link***) are defined below. Also, we introduce two functions, conjunction $Conj$ and disjunction $Disj$, with $N \geq 2$ parameters each, which, provided that probabilities of N events are given (p_1, p_2, $\ldots$, p_N), define the probabilities of (i) all these events occurring at the same time ($Conj$), and (ii) at least one of these events occurring ($Disj$). We denote the $Conj$ and $Disj$ functions as ***accumulation functions*** as they accumulate scores of non-CMP links to produce a score for an individual CMP link. Finally, all individual CMP links between t_1 and t_2 are aggregated through what we call an ***aggregation function*** (which can be e.g. the $\max$ of $N \geq 1$ numbers). Next, we define in some more details the concepts which we just introduced in relation to CMP.

Definition 1 ($Conj$)**.** The $Conj$ is a function which takes N arguments each in $[0, 1]$ and returns a result in $[0, 1]$. It should accumulate scores linked by a conjunctive association of events that are all necessary to occur. We discuss a possible implementation for it below.

Definition 2 ($Disj$)**.** The $Disj$ is a function which takes N arguments each in $[0, 1]$ and returns a result in $[0, 1]$. It should accumulate scores linked by a disjunctive association of events that reaffirm each other. We discuss possible implementations for it below.

Definition 3 (score of a non-CMP link)**.** The score of a non-CMP link between any two terms (which could be from the same ontology or not) is defined as

$$score(s_{ij}) = \begin{cases} I, & \text{if } s_{ij} \text{is an IO link} \\ D, & \text{if } s_{ij} \text{is a DM link} \\ f(src), & \text{if } s_{ij} \text{is a SMP link which came from the source} \\ & src \in \{\textit{UMLS}, \textit{FMA}, \textit{WordNet}\} \end{cases} \quad (5)$$

Definition 4 (score of a set of non-CMP links)**.** The score of a set of links (score of an evidence set) is defined as

$$score(\overline{S_i}) = Disj_{j=1}^{m}(score(s_{ij})), \tag{6}$$

where $Disj$ is the function from Definition 2, s_{ij} are links which are non-CMP (i.e. either IO or DM or SMP), and the $Disj$ is taken over all non-CMP links taking part in the evidence set $\overline{S_i}$.

Definition 5 (score of an individual CMP link)**.** The score of an individual CMP link e is defined as

$$score(e) = p \cdot Conj_{i=1}^{n}(score(\overline{S_i})), \tag{7}$$

where $p \in [0, 1]$ is a penalty constant accounting for the uncertainty of CMP; $Conj$ is the function from Definition 1; and $Conj$ is taken over all evidence sets that take part in the pattern instance, which the link e originates from.

Definition 6 (aggregation function)**.** Let K be the number of all individual CMP links drawn between $t_1 \in V_1$ and $t_2 \in V_2$. An ***aggregation function*** is a known function F_{agg} which takes the scores of all these K individual CMP links and produces a single number $score_{CMP} \in [0, 1]$, which we call the score of the *aggregated* (or *final*) ***CMP link*** drawn between t_1 and t_2.

As a final result this aggregated CMP link is drawn between any two terms t_1 and t_2 for which at least one pattern (of the three types **X**, **U**, **V**) is found and then the score of this link is calculated.

5 Scoring functions

A natural candidate to compute the accumulated score of predictions found in several references or procedures is the law for addition of probabilities. The law for multiplication of probabilities can be used along with it to calculate the score of evidence based on multiple necessary conditions in CMP. The simplest way to apply these laws would be to assume that the different predictions are independent, which would give us a set of the functions $Conj$, $Disj$ and F_{agg} that constitute our first scoring scheme.

Scheme #1 (“simple”)

(1a) $Conj(s_1, s_2) = s_1 \cdot s_2$
$Conj(s_1, s_2, \ldots, s_N) = Conj(Conj(s_1, s_2, \ldots, s_{N-1}), s_N)$

(1b) $Disj(s_1, s_2) = s_1 + s_2 - s_1 \cdot s_2$
$Disj(s_1, s_2, \ldots, s_N) = Disj(Disj(s_1, s_2, \ldots, s_{N-1}), s_N)$

(1c) $F_{agg}(s_1, s_2, \ldots, s_N) = \max(s_1, s_2, \ldots, s_N)$

Since the assumption of independence is least applicable to our CMP procedure, during the aggregation we have chosen not to use the same function to aggregate scores coming from it, but to pick the maximum of the scores instead. Thus, our choice for $Disj$ is the formula for the

probability of a union of independent events, our choice for F_{agg} – a union of events completely dependent on each other, and our choice for $Conj$ – an intersection of events. Indeed, if A and B are events, then:

- $P(A \cup B) = P(A) + P(B) - P(AB) = P(A) + P(B) - P(A)P(B)$, if A and B are independent
- $P(A \cup B) = P(A) + P(B) - p(A) = P(B) = max(P(A), P(B))$, if B is necessary for A (A is fully dependent on B).
- $P(AB) = P(A)P(B|A) = P(A)P(B)$, if A and B are independent

During the validation testing described in the next section, we found that this proposed scoring scheme is empirically good enough, and making additional improvements to it doesn't necessarily lead to significantly better results. In spite of that, we have proposed and tested two additional schemes to address some questions that our assumptions of independence or dependence might create.

The simplest modification we can make to scheme #1 is to assume independence between all individual CMP links as well. This would answer one very natural question – if CMP finds a great number of shared children between two nodes, shouldn't the probability that they are synonyms also increase significantly? If one shared child is a rare event in the data, having multiple children should indeed be even rarer. This new assumption gives us the following set of functions for our second scoring scheme:

Scheme #2 ("staircase")

(2a) $Conj(s_1, s_2) = s_1 \cdot s_2$
$Conj(s_1, s_2, \ldots, s_N) = Conj(Conj(s_1, s_2, \ldots, s_{N-1}), s_N)$

(2b) $Disj(s_1, s_2) = s_1 + s_2 - s_1 \cdot s_2$
$Disj(s_1, s_2, \ldots, s_N) = Disj(Disj(s_1, s_2, \ldots, s_{N-1}), s_N)$

(2c) $F_{agg}(s_1, s_2, \ldots, s_N) = Disj(s_1, s_2, \ldots, s_N)$

These simple proposals, however, ignore some more detailed statistical features of the knowledge source data, such as the degree of dependence, leaving some questions concerning the validity and precision unanswered:

- Because of the dependence between the individual CMP links, the aggregated CMP score in scoring scheme #2 could grow too big in cases with a large number of incorrect links, leading to a lot of false negatives. Such score growth might work better when restricted.
- Some knowledge sources might be highly correlated with each other, e.g. in case they borrowed data from the same third-party. As an example, a direct match would lead to a match in any knowledge source that knows about the existence of the terms.
- The dependence in each case can be different, requiring a parameter to account for it.

To resolve this, we are also proposing a third generalized scoring scheme which has a parameter α which is intended to limit the growth of $Disj$. We have also confirmed that substituting this α with the correlation (for generic probabilistic datasets in which it is known), the proposed formula for $Disj$ in our third scoring scheme yields the actual probability.

Scheme #3 ("hybrid")

(3a) $Conj(s_1, s_2) = s_1 \cdot s_2$
$Conj(s_1, s_2, \ldots, s_N) = Conj(Conj(s_1, s_2, \ldots, s_{N-1}), s_N)$

(3b) $Disj(s_1, s_2) = \alpha(s_1 + s_2 - s_1 \cdot s_2) + (1 - \alpha)\max(s_1, s_2)$
$Disj(s_1, s_2, \ldots, s_N) = Disj(Disj(s_1, s_2, \ldots, s_{N-1}), s_N)$

(3c) $F_{agg}(s_1, s_2, \ldots, s_N) = Disj(s_1, s_2, \ldots, s_N)$

The parameter α for a pair of scores is a function of the two sets of sources that lead to the prediction of each score. This means that its introduction adds not one, but numerous parameters to our procedure. Furthermore, these parameters represent the correlations which are not easily obtainable even for pairs of sources, let alone pairs of sets that could contain multiple sources each. This makes the use of the third scoring scheme difficult. For our testing, we selected initial values for α that seemed reasonable, and used trial and error to improve the score until we had found optimal values.

Once a prediction has its final score calculated by the scoring scheme, a threshold is needed to separate the correct predictions from the incorrect ones. For our datasets, all the scores ended up in two easily identifiable groups which could be separated easily through either statistical methods or clustering algorithms. Once the statistical properties of the two sets were available, we simply picked the threshold using the average of the means shifted with the standard deviations of the two sets: $(\mu_c - v_c + \mu_e + v_e)/2$

6 Results and discussion

For validation of the quality of the predictions generated by the procedures and of the scores produced by the scoring schemes, three anatomical ontologies from the OBO foundry were used – the mouse anatomy, the zebrafish anatomy and the xenopus anatomy. All predictions generated by the presented algorithmic procedures were curated one by one by an expert in human and animal anatomy, and the decisions of the curator were used for evaluating the calculated scores when using the respective threshold.

We conducted tests with CMP both enabled and disabled to show that it increases the score quality and the number of correctly identified predictions. The three score tables summarize the results, detailing the number of those correctly accepted or rejected by the scoring threshold, as well as the number of false negatives and false positives (in the"Incorrectly" columns). The hybrid scoring scheme #3 is placed between #1 and #2 as its aggregation function is a linear combination of their two aggregation functions.

Mouse – Zebrafish

	Without CMP				With CMP			
	Accepted		Rejected		Accepted		Rejected	
	Correctly	Incorrectly	Correctly	Incorrectly	Correctly	Incorrectly	Correctly	Incorrectly
#1	231	**5**	6	28	**257**	7	346	14
#3	223	**5**	6	36	**257**	7	347	14
#2	231	**5**	6	28	256	31	323	15

Xenopus – Mouse

	Without CMP				With CMP			
	Accepted		Rejected		Accepted		Rejected	
	Correctly	Incorrectly	Correctly	Incorrectly	Correctly	Incorrectly	Correctly	Incorrectly
#1	196	**3**	9	22	216	9	309	14
#3	196	**3**	9	22	216	9	309	14
#2	196	**3**	9	22	**219**	32	287	11

Xenopus – Zebrafish

	Without CMP				With CMP			
	Accepted		Rejected		Accepted		Rejected	
	Correctly	Incorrectly	Correctly	Incorrectly	Correctly	Incorrectly	Correctly	Incorrectly
#1	393	**23**	21	16	405	36	311	14
#3	393	**23**	21	16	405	36	311	14
#2	393	**23**	21	16	**406**	61	285	13

As it is seen, the use of CMP increases the number of correct predictions which is achieved at the cost of a small increase in the number of false positives. The majority of incorrect predictions added by CMP indicates that common children alone are not a good indicator for synonymy with the utilized knowledge sources, which can be a suggestion that these sources are relatively complete. In spite of that, CMP helps further validate the rest of the predictions made during the DM and SMP procedures by providing additional evidence for them.

The use of the scheme #2 ("staircase") for CMP predictions increases the number of false positives without any improvement in the number of correct predictions, thus making it a poor choice for the examined ontologies and knowledge sources. The hybrid scoring scheme #3 with a sensible choice of the α parameters produces results that are nearly identical to the simple scoring scheme #1, showing that the simple scheme #1 is a good enough choice, and in some cases – better.

7 Conclusion

We presented an original algorithmic approach for predicting and scoring cross-ontology links within semi-automatic or automatic mapping of different species-specific anatomical ontologies. The predictions were individually checked by a curator and their input was used to confirm the validity of the scoring procedures during a possible fully-automated mapping. While

a semi-automated approach, in which the predictions are carefully checked by a curator, is still preferable, the low number of detected false positives show that a fully-automated approach is also viable. The procedures described briefly here and detailed in [15], and the scoring schemes introduced are utilized in the software program AnatOM [1, 16] developed as part of our work on mapping and merging of anatomical ontologies. The source code of an implementation of the discussed procedures is available at `https://launchpad.net/anatom`.

Acknowledgements

The authors would like to express their gratitude to the late professor Jack Leunissen of the Wageningen University for his inspiration and the collaboration in the conception and the development of this work.

References

[1] P. Petrov, N. Natchev, M. Krachounov, M. Nisheva, O. Kulev and D. Vassilev. Anatom – an intelligent software program for semi-automatic mapping and merging of anatomy ontologies. In *Integrated Systems and Grid Technologies, Sixth International Conference ISGT'2012*, pages 173–187. St. Kliment Ohridski University Press, Sofia, Bulgaria, 2012.

[2] J. de Bruijn et al. Ontology mediation, merging, and aligning. In J. Davies, R. Studer and P. Warren (editors), *Semantic Web Technologies*, pages 95–113. Wiley, 2006.

[3] J. Euzenat and P. Shvaiko. *Ontology Matching*. Springer, Heidelberg, 2007.

[4] J. Day-Richter. Obo flat file format specification, version 1.2, 2006. URL `http://www.geneontology.org/GO.format.obo-1_2.shtml`, [Online; accessed 14 February 2012].

[5] M. Ashburner et al. Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1):20–29, 2000.

[6] B. Smith et al. The OBO foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25:1251–1255, 2007.

[7] O. Bodenreider. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Research*, 32:267–270, 2004.

[8] Web site of the Unified Medical Language System (UMLS). URL `http://www.nlm.nih.gov/research/umls/`, [Online; accessed 1 October 2012].

[9] C. Rosse and J. Mejino. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *J. Biomed. Inform.*, 36(6):478–500, 2003.

[10] Web site of the Foundational Model of Anatomy (FMA). URL `http://sig.biostr.washington.edu/projects/fm/`, [Online; accessed 1 October 2012].

[11] G. Miller. A lexical database for English. *Communications of the ACM*, 38(11):39–41, 1995.

[12] C. Fellbaum. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA, 1998.

[13] Web site of the WordNet project. URL `http://wordnet.princeton.edu/`, [Online; accessed 1 October 2012].

[14] E. van Ophuizen and J. Leunissen. An evaluation of the performance of three semantic background knowledge sources in comparative anatomy. *J. Integrative Bioinformatics*, 7:124–130, 2010.

[15] P. Petrov, M. Krachunov, E. van Ophuizen and D. Vassilev. An algorithmic approach to inferring cross-ontology links while mapping anatomical ontologies. *Serdica Journal of Computing*, 6, 2012.

[16] P. Petrov, M. Krachunov, E. Todorovska and D. Vassilev. An intelligent system approach for integrating anatomical ontologies. *Biotechnology and Biotechnological Equipment*, 26(4):3173–3181, 2012.