SAMBO - A System for Aligning and Merging Biomedical Ontologies*

Patrick Lambrix and He Tan
Department of Computer and Information Science
Linköpings universitet, Sweden
{patla,hetan}@ida.liu.se

Abstract

Due to the recent explosion of the amount of on-line accessible biomedical data and tools, finding and retrieving the relevant information is not an easy task. The vision of a Semantic Web for life sciences alleviates these difficulties. A key technology for the Semantic Web are ontologies. In recent years many biomedical ontologies have been developed and many of these ontologies contain overlapping information. To be able to use multiple ontologies they have to be aligned or merged. In this paper we propose a framework for aligning and merging ontologies. Further, we developed a system for aligning and merging biomedical ontologies (SAMBO) based on this framework. The framework is also a first step towards a general framework that can be used for comparative evaluations of alignment strategies and their combinations. In this paper we evaluated different strategies and their combinations in terms of quality and processing time and compared SAMBO with two other systems.

keywords: ontologies, alignment, merging, biomedical informatics

This is a pre-print version of an article in the Journal of Web Semantics 4(3):196-206, doi: 10.1016/j.websem.2006.05.003.

---

*The home page for SAMBO is http://www.ida.liu.se/∼iislab/projects/SAMBO/.

1 Introduction

Researchers in various areas, e.g. medicine, agriculture and environmental sciences, use biomedical data sources and tools to answer different research questions or to solve various tasks [3], for instance, in drug discovery or in research on the influence of environmental factors on human health and diseases. During recent years an enormous amount of biomedical data has been generated. These data are spread in a large number of autonomous data sources that are often publicly available on the Web. There are also numerous tools available on the Web. Due to this recent explosion of the amount of on-line accessible data and tools, finding the relevant sources and retrieving the relevant information is not an easy task. Further, often information from different sources needs to be integrated. The vision of a Semantic Web for life sciences alleviates these difficulties [38, 19]. A key technology for the Semantic Web are ontologies. The Semantic Web can be seen as an extension of the current Web in which information is given a well-defined meaning by annotating Web content with ontology terms.

Intuitively, ontologies (e.g. [18, 14]) can be seen as defining the basic terms and relations of a domain of interest, as well as the rules for combining these terms and relations. Ontologies are used for communication between people and organizations by providing a common terminology over a domain. They provide the basis for interoperability between systems. They can be used for making the content in information sources explicit and serve as an index to a repository of information. Further, they can be used as a basis for integration of information sources and as a query model for information sources. They also support a clear separation of domain knowledge from application-based knowledge as well as validation of data sources. The benefits of using ontologies include reuse, sharing and portability of knowledge across platforms, and improved documentation, maintenance, and reliability. Overall, ontologies lead to a better understanding of a field and to more effective and efficient handling of information in that field. In the field of bioinformatics, for instance, the work on ontologies is recognized as essential in some of the grand challenges of genomics research [3] and there is much international research cooperation for the development of ontologies (e.g. the Gene Ontology (GO) [13] and Open Biomedical Ontologies (OBO) [32] efforts) and the use of ontologies for the Semantic Web (e.g. the EU Network of Excellence REWERSE Working Group A2 [38]).

Many ontologies have already been developed and many of these ontologies contain overlapping information. In figure 1, for instance, we see two small pieces from two ontologies where terms in the two ontologies are equivalent (bold face). Often we would therefore want to be able to use multiple ontologies. For instance, companies may want to use community standard ontologies and use them together with company-specific ontologies. Applications may need to use ontologies from different areas or from different views on one area. Ontology builders may want to use already existing ontologies as the basis for the creation of new ontologies by extending the existing ontologies or by combining knowledge from different smaller ontologies. Further, different data sources in the same domain may have annotated their data with different but similar ontologies. In each of these cases it is important to know the relationships between the terms in the different ontologies. It has been realized that this is a major issue and some organizations have started to deal with it. For instance, the organization for Standards and Ontologies for Functional Genomics (SOFG) [42] developed the SOFG Anatomy Entry List which defines cross species anatomical terms relevant to functional genomics and which can be used as an entry point to anatomical ontologies. In a similar spirit [41] defines a number of high-level relations in biomedical ontologies to promote interoperability of ontologies. In the remainder of this paper we say that we align two ontologies when we define the relationships between terms in the different ontologies. We merge two ontologies when we, based on the alignment relationships between the ontologies, create a new ontology containing the knowledge included in the source ontologies.

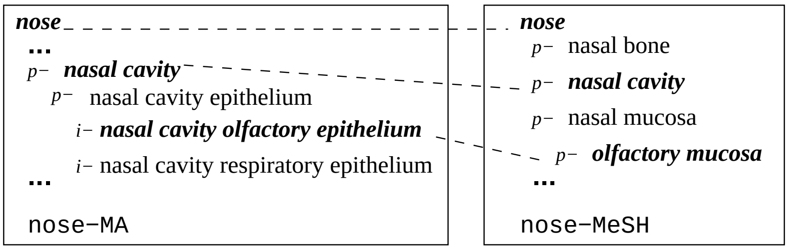


Figure 1: Example of overlapping ontologies.

In this paper we tackle the problem of aligning and merging biomedical ontologies. Our contribution is three-fold: we present a framework for aligning and merging ontologies, develop an ontology alignment and merging system based on the framework and evaluate different alignment strategies and their combinations. The first contribution is presented in section 3. We identified different types of alignment strategies and show how these strategies can be integrated in one framework. Most of the current alignment and merging systems can be seen as instantiations of our framework. Further, we developed a system for aligning and merging biomedical ontologies (SAMBO) according to this framework (section 4). Within this system we have implemented some already existing alignment strategies as well as some new strategies. Although the framework and the SAMBO architecture are domain independent, we have focused on strategies that are applicable to the types of ontologies that are currently available in the biomedical domain. We evaluated different alignment strategies and their combinations in terms of quality and processing time using several biomedical ontologies. We also compared SAMBO with two other systems. The results are discussed in section 5. Related work is discussed in section 6 and the paper concludes in section 7. In the next section we provide some background on biomedical ontologies.

2 Biomedical ontologies

Ontologies differ regarding the kind of information they can represent. From a knowledge representation point of view ontologies can have the following components (e.g. [18, 43]). Concepts represent sets or classes of entities in a domain. Instances represent the actual entities. Instances are, however, often not represented in ontologies. Further, there are many types of relations. Finally, axioms represent facts that are always true in the topic area of the ontology. These can be such things as domain restrictions, cardinality restrictions or disjointness restrictions. Depending on which of the components are represented and the kind of information that can be represented, we can distinguish between different kinds of ontologies such as controlled vocabularies, taxonomies, thesauri, data models, frame-based ontologies and knowledge-based ontologies. These different types of ontologies can be represented in a spectrum of representation formalisms ranging from very informal to strictly formal. For instance, some of the most expressive representation formalisms in use for ontologies are description logic-based languages such as OWL [34].

Biomedical ontologies (e.g. [18]) have been around for a while and their use has grown drastically since data source builders concerned with developing systems for different (model) organisms joined to create the Gene Ontology Consortium [13] in 1998. The research in biomedical ontologies is now also recognized as essential in some of the grand challenges of genomics research [3]. Further, the field has matured enough to develop standardization efforts. An example of this is the organization of the first conference on Standards and Ontologies for Functional Genomics in 2002 and the development of the SOFG resource on ontologies [42]. There exist ontologies that have reached the status of de facto standard and are being used extensively for annotation of data sources. Also, OBO was started as an umbrella web address for ontologies for use within the biomedical domain. Many biomedical ontologies are already available via OBO. There are also many overlapping ontologies available in the field. Most biomedical ontologies are vocabularies or taxonomies.

The ontologies that we use in our evaluations are GO ontologies, Signal-Ontology (SigO) [47], Medical Subject Headings (MeSH) [26] and the Anatomical Dictionary for the Adult Mouse (MA) [16]. The GO Consortium is a joint project which goal is to produce a structured, precisely defined, common and dynamic controlled vocabulary that describes the roles of genes and proteins in all organisms. Currently, there are three independent ontologies publicly available over the Internet: biological process, molecular function and cellular component. The GO ontologies are a de facto standard and many different bio-data sources are today annotated with GO terms. The terms in GO are arranged as nodes in a directed acyclic graph, where multiple inheritance is allowed. The purpose of the SigO project is to extract common features of cell signaling in the model organisms, try to understand what cell signaling is and how cell signaling systems can be modeled. SigO is based on the knowledge of the Cell Signaling Networks data source [46] and treats complex knowledge of living cells such as pathways, networks and causal relationships among molecules. The ontology consists of a flow diagram of signal transduction and a conceptual hierarchy of biochemical attributes of signaling molecules. MeSH is a controlled vocabulary produced by the American National Library of Medicine and is used for indexing, cataloging, and searching for biomedical and health-related information and documents. It consists of sets of terms naming descriptors in a hierarchical structure. These descriptors are organized in 15 categories, such as the category for anatomic terms, which is the category we use. The purpose of MA is to provide an ontology for annotating and integrating different types of data pertinent to anatomy. It is based on the Mouse Embryo Anatomy Nomenclature Database [1] and will be integrated with the Anatomical Dictionary for Mouse Development to generate an anatomy ontology covering the entire lifespan of the laboratory mouse. The ontology contains more than 2400 anatomical terms. They are structured as directed acyclic graphs across *is-a* and *part-of* relationships. The hierarchy of the ontology is organized in both spatial and functional ways. We have chosen these ontologies because there is substantial overlap between GO and SigO, and between

MeSH and MA, respectively. Further, there is a lot of interest and research in the areas of pathways and anatomy.

3 Ontology alignment and merging framework

Ontology alignment and merging is recognized as an important step in ontology engineering that needs more extensive research (e.g. [33]). Currently, there exist a number of ontology alignment systems that support the user to find inter-ontology relationships. Some of these systems are also ontology merging systems.

In this section we present a framework [22] for aligning and merging ontologies. The current systems that use the computation of similarity values between terms in the source ontologies[1], can be seen as instantiations of our framework.

3.1 Framework

Figure 2 shows a general strategy for aligning two ontologies based on the computation of similarity values between terms in the source ontologies. An alignment algorithm receives as input two source ontologies. The algorithm can include several matchers. The matchers can implement strategies based on linguistic matching, structure-based strategies, constraint-based approaches, instance-based strategies, strategies that use auxiliary information or a combination of these. Each matcher utilizes knowledge from one or multiple sources. The matchers calculate similarities between the terms from the different source ontologies. Alignment suggestions are then determined by combining and filtering the results generated by one or more matchers. By using different matchers and combining and filtering the results in different ways we obtain different alignment strategies. The suggestions are then presented to the user who accepts or rejects them. The acceptance and rejection of a suggestion may influence further suggestions. Further, a conflict checker is used to avoid conflicts introduced by the alignment relationships. The output of the alignment algorithm is a set of alignment relationships between terms from the source ontologies.

Figure 3 shows a simple merging algorithm. A new ontology is computed from the source ontologies and their identified alignment. The checker is

---

[1] There are also some systems that use other approaches such as FCA-Merge [44], HCONE [17], IF-Map [51] and S-Match [15].

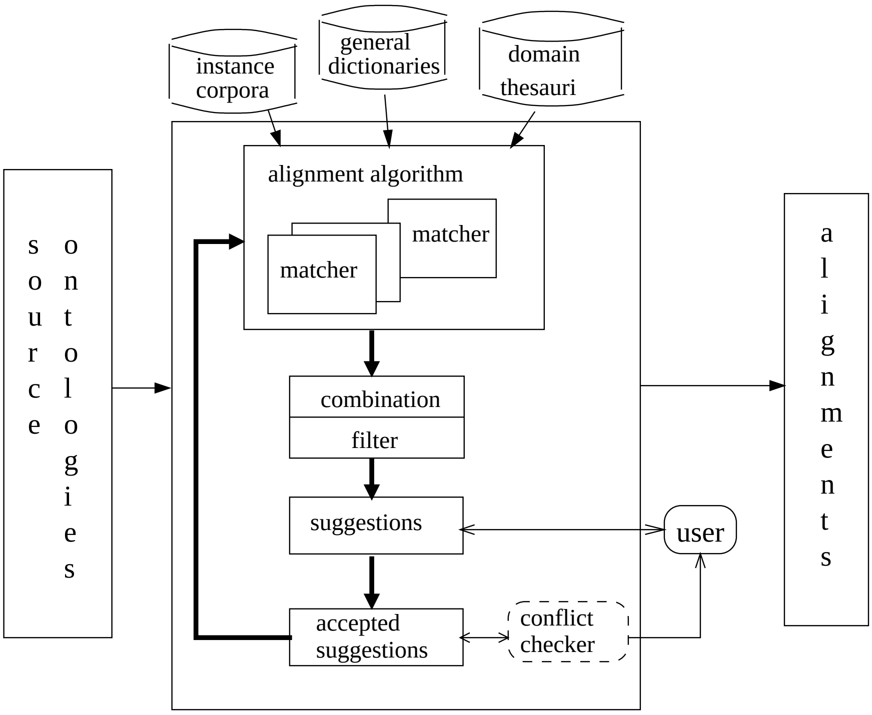


Figure 2: Alignment strategy.

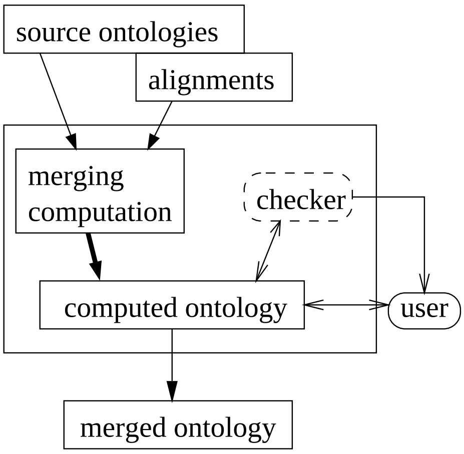


Figure 3: Merging algorithm.

used to avoid conflicts as well as to detect unsatisfiable concepts and, if so desired by the user, to remove redundancy.

3.2 Strategies

The matchers use different strategies to calculate similarities between the terms from the different source ontologies. They use different kinds of knowledge that can be exploited during the alignment process to enhance the effectiveness and efficiency.[2] Some of the approaches use information inherent in the ontologies. Other approaches require the use of external sources. We describe the types of strategies that are used by current ontology alignment systems and in section 6 we give an overview of the different types of knowledge used per system.

- *Strategies based on linguistic matching.* These approaches make use of textual descriptions of the concepts and relations such as names, synonyms and definitions. The similarity measure between concepts is based on comparisons of the textual descriptions. Simple string matching approaches and information retrieval approaches (e.g. based on frequency counting) may be used. Most systems use this kind of strategies.

- *Structure-based strategies.* These approaches use the structure of the ontologies to provide suggestions. Typically, a graph structure over the concepts is provided through is-a, part-of or other relations. The similarity of concepts is based on their environment. An environment can be defined in different ways. For instance, using the is-a relation (e.g. [21]) an environment could be defined using the parents (or ancestors) and the children (or descendants) of a concept. Some approaches also use other relations (e.g. [30]).

- *Constraint-based approaches.* In this case the axioms are used to provide suggestions. For instance, knowing that the range and domain of two relations are the same, may be an indication that there is a relationship between the relations. Similarly, when two concepts are both disjoint with a third concept, we may have a similarity between the first two concepts. On their own these approaches may not be

---

[2] Also the approaches that are not based on the computation of similarity values may use these types of knowledge.

sufficient to provide high quality suggestions, but they may complement other approaches to reduce the number of irrelevant suggestions. Constraint-based approaches are currently used by only a few systems.

- *Instance-based strategies.* In some cases instances are available directly or can be obtained. For instance, the entries in biological data sources that are annotated with GO terms, can be seen as instances for these GO terms. When instances are available, they may be used in defining similarities between concepts.
- *Use of auxiliary information.* Dictionaries and thesauri representing general or domain knowledge, or intermediate ontologies may be used to enhance the alignment process. They provide external resources to interpret the intended meaning of the concepts and relations in an ontology (e.g. [28]). Also information about previously aligned or merged ontologies may be used. Many systems use auxiliary information.
- *Combining different approaches.* The different approaches use different strategies to compute similarity between concepts. Therefore, a combined approach may give better results. Although most systems combine different approaches, not much research is done on the applicability and performance of these combinations.

4 SAMBO

SAMBO is an ontology alignment and merging tool developed according to the framework described in section 3. Regarding the strategies for the alignment process our work has focused on strategies that are applicable to the types of ontologies that are currently available in the biomedical domain.

4.1 System

The current implementation supports ontologies in OWL format. This means that ontologies may need to be translated to OWL format (see e.g. the test cases in section 5). The system separates the process into two steps: aligning relations and aligning concepts. The second step can be started after the first step is finished. In the suggestion mode several kinds of matchers can be used and combined. Figure 4 shows how different matchers can be chosen and weights can be assigned to these matchers. Filtering is performed using a threshold value. The pairs of terms with a similarity value above this value are shown to the user as alignment suggestions. An example alignment suggestion is given in figure 5. The system displays information (definition/identifier, synonyms, relations) about the source ontology terms in the suggestion. For each of the alignment suggestions the user can decide whether the terms are equivalent, whether there is an is-a relation between the terms, or whether the suggestion should be rejected. If the user decides that the terms are equivalent, a new name for the term can be given as well. Upon an action of the user, the suggestion list is updated. If the user rejects a suggestion where two different terms have the same name, she is required to rename at least one of the terms (figure 6). At each point in time during the alignment process the user can view the ontologies represented in trees with the information on which actions have been performed, and she can check how many suggestions still need to be processed. Figure 7 shows the remaining suggestions for a particular alignment process. A similar list can be obtained to view the previously accepted alignment suggestions. In addition to the suggestion mode, the system also has a manual mode in which the user can view the ontologies and manually align terms (figure 8). The source ontologies are illustrated using is-a and part-of hierarchies (*i* and *p* icons, respectively). The user can choose terms from the ontologies and then specify an alignment operation. Previously aligned terms are identified by different icons. For instance, the *M* icons in front of 'nasal_cavity' in the two ontologies in figure 8 show that these were aligned using an equivalence relationship. There is also a search functionality to find specific terms more easily in the hierarchy. The suggestion and manual modes can be interleaved. The suggestion mode can also be repeated several times, and take into account the previously performed operations.

After the user accomplishes the alignment process, the system receives the final alignment list and can be asked to create the new ontology. The system merges the terms in the alignment list, computes the consequences, makes the additional changes that follow from the operations, and finally copies the other terms to the new ontology. Furthermore, SAMBO uses a DIG description logic reasoner (e.g. Racer [39], FaCT [11]) to provide a number of reasoning services. The user can ask the system whether the new ontology is consistent and can ask for information about unsatisfiable concepts and cycles in the ontology.

4.2 Matchers

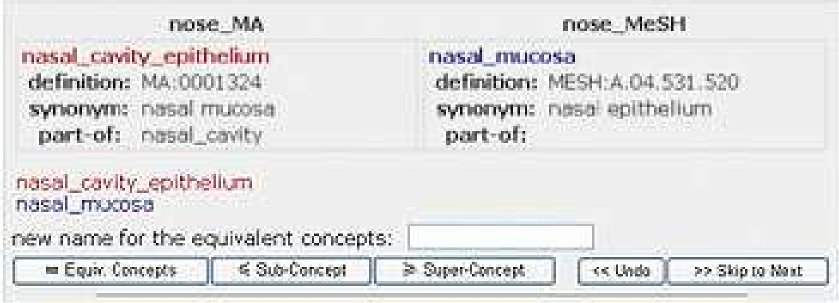
We experimented with the combination of already existing strategies as well as some newly implemented strategies. All matchers compute similarity values between 0 and 1.

Align Concept in nose_MA and nose_MeSH

matchers:
1.0 ☑ Terminology
1.0 ☐ WordNet Plus
1.0 ☐ UMLS
1.2 ☑ Bayes
1.0 ☐ Hierarchy

threshold: 0.6

Figure 4: Combination and filtering.



nose_MA
nasal_cavity_epithelium
definition: MA:0001324
synonym: nasal mucosa
part-of: nasal_cavity

nose_MeSH
nasal_mucosa
definition: MESH:A.04.531.520
synonym: nasal epithelium
part-of:

nasal_cavity_epithelium
nasal_mucosa
new name for the equivalent concepts:

Figure 5: Alignment suggestion.

nose_MA
nasal_cavity
definition: MA:0000284
synonym:
part-of: nose

nose_MeSH
nasal_cavity
definition: MESH:A.04.531.449
synonym:
part-of:

New name for
○ nasal_cavity
● nasal_cavity
nasal_cavity_MeSH

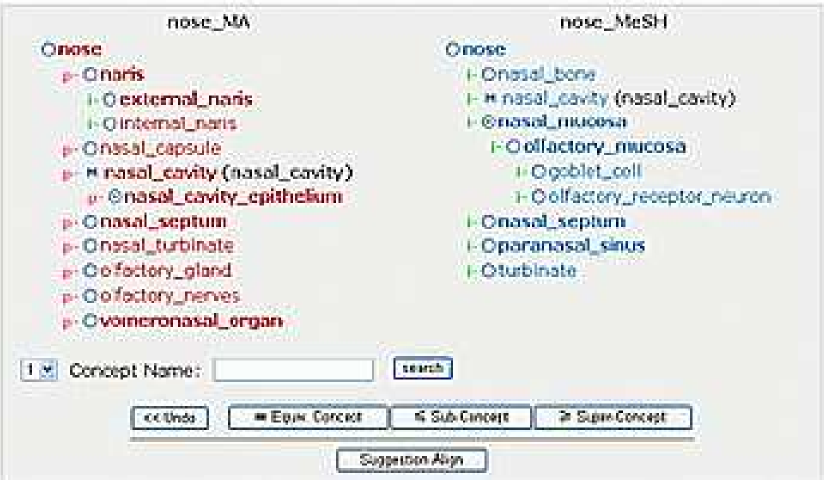
Figure 6: New name required.

🟥 nose_MA 🟦 nose_MeSH

(nasal_cavity_epithelium, nasal_mucosa)
(nasal_cavity_olfactory_epithelium, olfactory_mucosa)
(nasal_septum, nasal_septum)
(nose, nose)
(vomeronasal_organ, vomeronasal_organ)

5 Remaining Suggestions, Previously Aligned Concepts

Figure 7: Information about the remaining suggestions.



nose_MA:

- nose
  - naris
    - external_naris
    - internal_naris
  - nasal_capsule
  - nasal_cavity (nasal_cavity)
    - nasal_cavity_epithelium
  - nasal_septum
  - nasal_turbinate
  - olfactory_gland
  - olfactory_nerves
  - vomeronasal_organ

nose_MeSH:

- nose
  - nasal_bone
  - nasal_cavity (nasal_cavity)
  - nasal_mucosa
    - olfactory_mucosa
      - goblet_cell
      - olfactory_receptor_neuron
  - nasal_septum
  - paranasal_sinus
  - turbinate

1 Concept Name:

Figure 8: Manual mode.

Terminological matcher. The terminological matcher contains matching algorithms based on the textual descriptions (names and synonyms) of concepts and relations. In the current implementation, the matcher includes two approximate string matching algorithms, n-gram and edit distance, and a linguistic algorithm. An n-gram is a set of n consecutive characters extracted from a string. Similar strings will have a high proportion of n-grams in common. Edit distance is defined as the number of deletions, insertions, or substitutions required to transform one string into the other. The greater the edit distance, the more different the strings are. The linguistic algorithm computes the similarity of the terms by comparing the lists of words of which the terms are composed. Similar terms have a high proportion of words in common in the lists. A porter stemming algorithm is employed to each word. Further, a general thesaurus, WordNet [50], can be used to enhance the similarity measure by looking up the hypernym relationships of the pairs of words in WordNet. These matchers were evaluated in [21] using MeSH anatomy (ca 1400 terms) and MA (ca 2350 terms). The terminological matcher outputs similarity values by combining the results from these three algorithms using a weighted sum. If the weights are chosen carefully, this combination can overcome the weaknesses of the individual algorithms [21]. In our experiments we used the weights 0.37, 0.37 and 0.26 for the linguistic algorithm, edit distance and n-gram, respectively.

Structural matcher. The *structural matcher* is an iterative algorithm based on the *is-a* and *part-of* hierarchies of the ontologies. The algorithm requires as input a list of alignment relationships and similarity values and can therefore not be used in isolation. The intuition behind the algorithm is that if two concepts lie in similar positions with respect to is-a or part-of hierarchies relative to already aligned concepts in the two ontologies, then they are likely to be similar as well. For each pair of concepts (C_1,C_2) in the original list of alignment relationships the structural matcher augments the original similarity value for pairs of concepts (C'_1,C'_2) such that C'_1 and C'_2 are equivalent to, are in an is-a relationship with, or participate in a part-of relationship with C_1 and C_2, respectively. The augmentation depends on the relationship and on the distance between the concepts in the is-a and part-of hierarchies. The augmentation diminishes with respect to distance. The new similarity value can also not exceed 1. In our experiments we used a maximal distance of 2 and the effect on ancestors is lower than the effect on descendants.

Use of domain knowledge. Another strategy is to use domain knowledge. We utilize the Metathesaurus in the Unified Medical Language System (UMLS) [49] which contains more than 100 biomedical and health-related vocabularies. The Metathesaurus is organized using concepts. The concepts may have synonyms which are the terms in the different vocabularies in the Metathesaurus that have the same intended meaning. The similarity of two terms in the source ontologies is determined by their relationship in UMLS. In our experiments we use the UMLS Knowledge Source Server to query the Metathesaurus with source ontology terms. As a result we obtain concepts that have the source ontology term as their synonym. We assign a similarity value of 1 for exact matches of query results for the two terms, 0.6 if the source ontology terms are synonyms of the same concept and 0 otherwise.

Learning matcher. We also included a *learning matcher*. The matcher makes use of life science literature that is related to the concepts in the ontologies. It is based on the intuition that a similarity measure between concepts in different ontologies can be defined based on the probability that documents about one concept are also about the other concept and vice versa. The strategy contains the following basic steps. (i) For each ontology that we want to align we generate a corpus of PubMed abstracts. PubMed [37] is a service of the National Library of Medicine that includes over 15 millions citations from MEDLINE [27] and other biomedical journals. In our implementation we generated a corpus of maximally 100 PubMed abstracts per concept using the programming utilities [40] provided by the retrieval system Entrez [7]. (ii) For each ontology a document classifier is generated. This classifier returns for a given document the concept that is most closely related to the document. To generate a classifier the corpus of abstracts associated to the classifier's ontology is used. In our algorithm we use a naive Bayes classification algorithm.[3] (iii) Documents of one ontology are classified by the document classifier of the other ontology and visa versa. (iv) A similarity measure between concepts in the different ontologies is computed by using the results of step (iii). The similarity is computed as

$$lsim(C_1, C_2) = \frac{n_{NBC2}(C_1, C_2) + n_{NBC1}(C_2, C_1)}{n_D(C_1) + n_D(C_2)}$$

where $n_D(C)$ is the number of abstracts originally associated with C, and $n_{NBCx}(C_p, C_q)$ is the number of abstracts associated with C_p that are also related to C_q as found by classifier $NBCx$ related to ontology x. More details about this algorithm as well as some extensions can be found in [48].

Combinations. The user is given the choice to employ one or several matchers during the alignment process. The suggestions can be determined

---

[3] The implementation of the naive Bayes classifier is based on the code available at http://www.cs.utexas.edu/users/mooney/ir-course/.

based on the similarity value from one matcher, or the combination of the similarity values measured by several matchers using weights,

$$sim(C_1, C_2) = (\sum_{k=1}^{n} w_k * sim_k(C_1, C_2)) / \sum_{k=1}^{n} w_k,$$

where sim_k and w_k represent the similarities and weights, respectively, for the different matchers (combination in figure 2).

5 Evaluations

In our evaluation we have focused on several aspects. We compare different matchers and different combinations of matchers using different thresholds. We compare them with respect to the quality of the suggestions they generate and the time they take to generate the suggestions. Further, we compare one of the SAMBO matchers with similar matchers in two other systems (PROMPT [31] and FOAM [12]) with respect to the quality of the suggestions. We have chosen our test cases such that they are based on biomedical ontologies and that they include ontologies with different granularity. In the remainder of this section we describe our test cases and our evaluation results.

5.1 Test cases

We created five test cases based on two groups of biomedical ontologies. For the first two cases we use a part of a GO ontology together with a part of SigO. Each case was chosen in such a way that there was an overlap between the GO part and the SigO part. The first case, *behavior* (B), contains 57 terms from GO and 10 terms from SigO. The second case, *immune defense* (ID), contains 73 terms from GO and 17 terms from SigO. The granularity of GO is higher than the granularity of SigO for these topics.

The other cases are taken from two biomedical ontologies that are available from OBO: MeSH (anatomy category) and MA. The two ontologies cover a similar subject domain, anatomy, and are developed independently. The three cases used in our test are: *nose* (containing 15 terms from MeSH and 18 terms from MA), *ear* (containing 39 terms from MeSH and 77 terms from MA), and *eye* (containing 45 terms from MeSH and 112 terms from MA). We translated the ontologies from the GO flat file format to OWL retaining identifiers, names, synonyms, definitions and is-a and part-of relations. The synonyms were transformed into equivalence statements.

Domain experts were asked to analyze the cases and provide alignment relationships based on equivalence and is-a relations. The domain experts were not experts in formal ontologies. We checked the consistency of the received alignments. As the test cases are relatively small, we assume that the experts considered the reasoning implications of the is-a relations. Therefore, in our evaluations we have used the ontologies and the alignment relationships from the experts as they were provided to us.

5.2 Comparison of matchers

In table 1 we present information about the suggestions generated by the individual matchers: terminological (Term), terminological using WordNet (TermWN), algorithm using domain knowledge UMLS (Dom), and learning (Learn). The cases are given in the first column. The second column represents the number of expected suggestions provided by domain experts. In our evaluation we consider only expected suggestions related to equivalence of terms or is-a relations between terms. For instance, in the ear case, there are 27 alignments that are specified by domain experts. This is the minimal set of suggestions that matchers are expected to generate for a perfect recall. This set does not include the inferred suggestions. Inferred suggestions will be inferred by the merging algorithm and we therefore count them neither as correct nor as wrong suggestions. An example of an inferred suggestion is that incus is-a ear ossicle. In this case we know that auditory bone (MA) is equivalent to ear ossicle (MeSH), and incus is-a auditory bone in MA. Then the system should derive that incus is-a ear ossicle. The third column represents the threshold value. Pairs with a similarity value higher than the threshold are suggestions. The other columns present the results of the different algorithms. The four numbers in the cells represent the number of suggestions provided by the matcher, the number of correct suggestions, the number of wrong suggestions and the number of inferred suggestions, respectively. For instance, the learning matcher (last column) for the case B and threshold 0.4 generates 4 suggestions of which 2 suggestions are correct, 1 suggestion is wrong and 1 suggestion is inferred. The structural matcher (Struct) requires a set of already identified alignments as input, and thus there are no results for the structural matcher in table 1.

We note that these results refer to the initial list of suggestions that the matchers compute. A system like SAMBO will update the suggestion list upon actions of the user. Therefore, some initial suggestions may not be shown to the user.

The quality of the suggestions for Term and TermWN differs significantly

Case	ES	*Th*	Term	TermWN	Dom	Learn
B	4	0.4	58/4/22/32	58/4/22/32	4/4/0/0	4/2/1/1
		0.5	35/4/13/18	35/4/13/18	4/4/0/0	2/2/0/0
		0.6	13/4/4/5	13/4/4/5	4/4/0/0	2/2/0/0
		0.7	6/4/0/2	6/4/0/2	4/4/0/0	2/2/0/0
		0.8	4/4/0/0	4/4/0/0	4/4/0/0	1/1/0/0
ID	8	0.4	96/7/66/23	96/7/66/23	4/4/0/0	9/6/3/0
		0.5	49/7/25/17	49/7/25/17	4/4/0/0	5/5/0/0
		0.6	15/5/4/6	16/5/5/6	4/4/0/0	2/2/0/0
		0.7	7/5/2/0	7/5/2/0	4/4/0/0	1/1/0/0
		0.8	6/4/2/0	6/4/2/0	4/4/0/0	0/0/0/0
nose	7	0.4	47/7/36/4	48/7/37/4	7/7/0/0	6/5/1/0
		0.5	27/7/17/3	28/7/18/3	7/7/0/0	6/5/1/0
		0.6	7/6/1/0	8/6/2/0	7/7/0/0	5/5/0/0
		0.7	6/6/0/0	6/6/0/0	6/6/0/0	5/5/0/0
		0.8	6/6/0/0	6/6/0/0	6/6/0/0	3/3/0/0
ear	27	0.4	147/26/104/17	155/26/110/19	26/23/2/1	18/16/2/0
		0.5	92/26/58/8	99/26/65/8	26/23/2/1	15/14/1/0
		0.6	47/26/19/2	47/26/19/2	26/23/2/1	12/11/1/0
		0.7	33/25/8/0	34/26/8/0	24/22/2/0	11/10/1/0
		0.8	26/24/2/0	28/25/3/0	24/22/2/0	3/3/0/0
eye	27	0.4	130/26/95/9	135/26/100/9	22/21/1/0	25/18/7/0
		0.5	72/23/42/7	74/23/44/7	22/21/1/0	18/17/1/0
		0.6	33/22/10/1	33/22/10/1	22/21/1/0	14/14/0/0
		0.7	24/21/3/0	24/21/3/0	19/18/1/0	10/10/0/0
		0.8	19/18/1/0	22/20/2/0	19/18/1/0	3/3/0/0

Table 1: Comparison of matchers: quality of the suggestions.

for different thresholds. The precision[4] diminishes fast when the threshold becomes lower, e.g. for the case B the precision goes down from 1 to 0.67, 0.31, 0.12 and 0.07. In our test cases for the threshold 0.8 the quality of the results is good. Most correct suggestions are kept while the wrong and inferred suggestions are filtered out. The terminological matcher can give suggestions where the names of terms are slightly different, e.g. (stapes, stape). As the test ontologies contain a large number of synonyms, also suggestions where the names of terms are completely different can be found, e.g. (inner ear, labyrinth), where inner ear has labyrinth as synonym. By using a general dictionary (WordNet), TermWN finds suggestions such as (perilymphatic channel, cochlear aqueduct) where cochlear aqueduct has perilymphatic duct as synonym, and duct is a synonym of channel in WordNet. On the other hand, since endothelium is a kind of epithelium in WordNet, it generates a wrong suggestion (corneal endothelium, corneal epithelium).

As the similarity values set by Dom can only be 1, 0.6 and 0, we obtain good results for the threshold 0.6. The matcher finds suggestions of which the terms have completely different names and synonyms, or have no synonyms at all, e.g. (external acoustic meatus, ear canal). The matcher works for some terms with slightly different names, e.g. (optic disc, optic disk), which are mapped to the concept optic disc in UMLS, but does not work for others, e.g. (stapes, stape), which are mapped to different concepts in UMLS.

The quality of the suggestions for Learn varies in the different cases in this evaluation. The recall of the results goes down significantly when the threshold becomes higher, e.g. in the ID case the recall goes down from 0.75 to 0.625, 0.25, 0.125 and 0. Learn can in most test cases be outperformed by the other matchers (by choosing appropriate thresholds) except in the ID case. In this case it avoids the wrong suggestions with slightly different names, such as (B cell activation, T Cell Activation), and also finds the suggestion (natural killer cell activation, Natural Killer Cell Response), which is not found by Dom. The quality of the suggestions from the learning matcher depends on the generated corpora of PubMed abstracts. The fact that the retrieval of the documents for each term does not consider their synonyms, may reduce the quality of their associated documents, e.g. the suggestion (nasal cavity epithelium, nasal mucosa), where nasal cavity epithelium has nasal

---

[4] We use precision as it is usually defined in information retrieval, i.e. the number of correct suggestions divided by the number of suggestions. As noted before, inferred suggestions are counted neither as correct nor wrong. Similarly, recall is defined as the number of correct suggestions divided by the total number of correct suggestions, in this case the expected suggestions.

Case	Term	TermWN	Dom	Learn
B	0.6	10.2	39.1	354.1
ID	2.9	35.8	47.5	421.2
nose	0.5	7.7	45	401.1
ear	3.8	35.3	124	667.5
eye	7.8	60.7	195.6	1299.1

Table 2: Comparison of matchers: time for computation of suggestions (in seconds).

mucosa as synonym, received the similarity value 1 in the other matchers but only 0.04 in Learn. Another factor is that for some terms only few PubMed abstracts are retrieved.

For these test ontologies we also observed that in most cases (except ID with thresholds 0.4 and 0.5, and nose with threshold 0.6) the set of correct suggestions provided by Learn was a sub-set of or equal to the set of correct suggestions provided by Dom, which in its turn was a sub-set of or equal to the set of correct suggestions provided by TermWN. However, TermWN also gives the highest number of wrong suggestions.

We also evaluated the time it takes for these algorithms to compute the suggestions. As the system responds instantaneously during the user interaction phase, we have focused on the time needed for the generation of the suggestions and putting them in the suggestion list. The results are presented in table 2. For TermWN we used a local installation of WordNet. We accessed UMLS for Dom via Internet. For Learn we generated the PubMed corpora beforehand. The time thus covers the time for learning the classifier and the time for computing the similarity values. We used a SUN Ultra 5_10 Sparc workstation for these tests.

Table 3 shows the *new* suggestions generated by the structural matcher based on the alignment results given by the other matchers. We used a threshold of 0.5. In this experiment the structural matcher did not give any new suggestions for Dom and Learn. For TermWN we did not receive any new correct suggestions. For the settings in the evaluation (size of effect of the structure on the similarity and maximal distance) the maximal augmentations of the similarities are as shown in table 4. The first number in each cell represent the maximal augmentation. However, these augmentations are always associated to previously accepted suggestions. The second number represents the highest augmentation for pairs of concepts that had

Case	ES	TermWN+Struct	Dom+Struct	Learn+Struct
B	4	3/0/0/3	0/0/0/0	0/0/0/0
ID	8	2/0/2/0	0/0/0/0	0/0/0/0
nose	7	1/0/1/0	0/0/0/0	0/0/0/0
ear	27	8/0/6/2	0/0/0/0	0/0/0/0
eye	27	6/0/6/0	0/0/0/0	0/0/0/0

Table 3: Structural matcher - extra suggestions.

Case	ES	TermWN+Struct	Dom+Struct	Learn+Struct
B	4	0.08/0.02	0.16/0.04	0.08/0.02
ID	8	0.13/0.02	0.22/0.17	0.11/0.05
nose	7	0.12/0.05	0.24/0.10	0.11/0.09
ear	27	0.14/0.08	0.29/0.18	0.11/0.06
eye	27	0.19/0.05	0.38/0.17	0.16/0.05

Table 4: Structural matcher - highest augmentation of similarity values.

an original similarity below 0.5. In the test cases this is too low to find new correct suggestions. This may be explained by the fact that some missing suggestions concern concepts in completely different positions in the two hierarchies. For other missing suggestions the concepts have a common ancestor or common descendants, but the ancestor or descendants are too distant for the similarity values to be influenced much. We also note that for B and nose we could not expect to find new correct suggestions for Dom and TermWN. Further, the missing correct suggestions for Dom had an original similarity value of 0, and therefore we did not expect Dom+Struct to provide new correct suggestions.

Tables 5, 6 and 7 present the quality of the suggestions considering the combination of the different matchers. We do not include the structural matcher because of its poor quality. We have evaluated combinations of matchers using different weights (1 and 1.2) for the matchers. The threshold value in this evaluation is 0.5. Table 5 represents pairs of matchers, while tables 6 and 7 represent the combination of three matchers. In our tests we found no significant difference for the different weight assignments in the combinations of the matchers. With respect to the correct suggestions the combinations of matchers do not find new results compared to the matcher in the combination that found the most correct suggestions. All

Case	ES	*Weights*	TermWN+Dom	TermWN+Learn	Dom+Learn
B	4	1.0, 1.0	4/4/0/0	6/4/0/2	4/4/0/0
		1.2, 1.0	4/4/0/0	6/4/0/2	4/4/0/0
		1.0, 1.2	4/4/0/0	6/4/0/2	4/4/0/0
ID	8	1.0, 1.0	4/4/0/0	8/7/1/0	4/4/0/0
		1.2, 1.0	4/4/0/0	9/7/2/0	4/4/0/0
		1.0, 1.2	4/4/0/0	7/6/1/0	4/4/0/0
nose	7	1.0, 1.0	7/7/0/0	8/7/1/0	7/7/0/0
		1.2, 1.0	7/7/0/0	9/7/2/0	7/7/0/0
		1.0, 1.2	7/7/0/0	7/6/1/0	6/6/0/0
ear	27	1.0, 1.0	25/23/2/0	27/22/5/0	24/22/2/0
		1.2, 1.0	26/24/2/0	30/24/6/0	24/22/2/0
		1.0, 1.2	25/23/2/0	23/18/5/0	20/18/2/0
eye	27	1.0, 1.0	22/21/1/0	24/21/3/0	20/19/1/0
		1.2, 1.0	22/21/1/0	24/21/3/0	20/19/1/0
		1.0, 1.2	22/21/1/0	24/21/3/0	20/19/1/0

Table 5: Combinations of pairs of matchers.

correct suggestions that are found by the combinations of matchers were also found by TermWN. As Dom only provides 1, 0.6 and 0 as similarity values, it tends to remove suggestions for which it has assigned the value 0. As Dom is based on domain knowledge this usually has the effect that wrong suggestions are removed. However, in some cases, where the domain knowledge is incomplete, it may also remove correct suggestions. Learn has a similar effect when combined with TermWN. For the test ontologies and the threshold 0.5, the best results are obtained when using TermWN (providing many correct suggestions) combined with at least one of Dom and Learn (removing most wrong suggestions).

An advantage of using a system like SAMBO is that one can experiment with different (combinations of) strategies and different (combinations of) types of ontologies. For instance, our evaluation gives an indication about what (combinations of) strategies may work well for aligning ontologies with similar properties as our test ontologies. However, when choosing a strategy other factors may also play a role. For instance, the combination strategy is more time consuming than the strategy using only the terminological matcher.

Case	ES	1, 1, 1	1.2, 1, 1	1, 1.2, 1
B	4	4/4/0/0	4/4/0/0	4/4/0/0
ID	8	4/4/0/0	4/4/0/0	4/4/0/0
nose	7	7/7/0/0	7/7/0/0	7/7/0/0
ear	27	24/22/2/0	25/23/2/0	24/22/2/0
eye	27	21/20/1/0	21/20/1/0	21/20/1/0

Table 6: Combination of matchers TermWN, Dom and Learn - 1.

Case	ES	1, 1, 1.2	1.2, 1.2, 1	1.2, 1, 1.2	1, 1.2, 1.2
B	4	4/4/0/0	4/4/0/0	4/4/0/0	4/4/0/0
ID	8	4/4/0/0	4/4/0/0	4/4/0/0	4/4/0/0
nose	7	7/7/0/0	7/7/0/0	7/7/0/0	7/7/0/0
ear	27	24/22/2/0	25/23/2/0	24/22/2/0	24/22/2/0
eye	27	21/20/1/0	21/20/1/0	21/20/1/0	21/20/1/0

Table 7: Combination of matchers TermWN, Dom and Learn - 2.

5.3 Comparison of tools

We compare SAMBO with two freely available tools, Protégé [36] with PROMPT [31] and FOAM [12], regarding the quality of suggestions.

Protégé is a tool for creating, editing, browsing and maintaining ontologies. It also has a number of plug-ins, among which is PROMPT, which includes several interactive tools for ontology merging and aligning. For the evaluation Protégé 3.1 with PROMPT 2.4.8 was locally installed and we used the ontology merging tool iPROMPT [29] in the PROMPT suite in the comparison. When merging two ontologies iPROMPT creates a list of initial suggestions based on the underlying alignment algorithms. The suggestions can, for instance, be to merge two terms, or to copy a term to the new ontology. The user can then perform an operation by accepting one of the suggestions or creating her own suggestions. iPROMPT performs the operation and additional changes that follow from that operation. The list of suggestions is then updated and a list of conflicts and possible solutions to these conflicts is created. This is repeated until the new ontology is ready.

FOAM is a (semi-)automatic tool for aligning and merging two or more OWL ontologies. The current distribution of the local Java application is used in the evaluation. It is a command-line application. When merging

Case	ES	iPROMPT	FOAM	SAMBO
B	4	4/4/0/0	7/3/2/2	4/4/0/0
ID	8	6/4/2/0	8/5/3/0	6/4/2/0
nose	7	6/6/0/0	12/6/5/1	6/6/0/0
ear	27	34/24/6/4	33/24/6/3	26/24/2/0
eye	27	29/20/6/3	27/21/5/1	19/18/1/0

Table 8: Comparison of systems: quality of the suggestions.

ontologies in semi-automatic mode, FOAM proposes alignment suggestions and the user can accept or reject these suggestions. The output of the system after processing all the suggestions is the accepted list of alignments. We use the alignment algorithm on the data layer [5] in the comparison (settings: complete comparison, semi-automatic).

As we do not have complete information on the exact algorithms and settings of iPROMPT and FOAM, we decided to use only the terminological matcher Term with threshold 0.8 for SAMBO. We note that this is not the optimal threshold and that the quality of the suggestions also can be improved by a combination matcher with TermWN.

Table 8 gives the results of our evaluation. In the test cases the precision of SAMBO is always higher than or equal to the precision of the other two systems. In the ID case FOAM gives one more correct suggestion that is not found by the other systems. In the eye case the recall of FOAM is the highest among the systems.

6 Related work

There are two kinds of related work: tools and evaluations. Up to date only three comparative evaluations of systems for ontology alignment and merging have been performed. The EU OntoWeb project [33] evaluated the systems PROMPT [29] based on Protégé (with extension Anchor-PROMPT [30]), Chimaera [25] (described, not evaluated), FCA-Merge [44] and ODE-Merge. This evaluation focused on such things as functionality, interoperability and visualization, but did not include tests on the quality of the alignment. In [20, 21] PROMPT, Chimaera and a former version of SAMBO were evaluated in terms of the quality of the alignment as well as the time it takes to align ontologies with these tools. Further, an ontology alignment contest was held at EON-2004 [8]. QOM [6] (FOAM), OLA [9], SCM [10] and PROMPT participated. The main goal of the contest was to show how ontology alignment tools can be evaluated and a follow-up was planned for 2005.

There are other tools such as ArtGen [28], ASCO [24], GLUE [4], HCONE [17], IF-Map [51], iMapper [45], OntoMapper [35], and S-Match [15], but these have not appeared in comparative evaluation studies.

The current systems use different types of knowledge. Most systems use linguistic, structure-based and/or instance-based strategies. Also auxiliary information is used often. In most cases this auxiliary information is WordNet. Constraint-based approaches are not used much. An overview is given in table 9.

7 Conclusions

In this paper we have presented a general framework for aligning and merging ontologies, a system (SAMBO) and several evaluations. Most of the current alignment and merging systems can be seen as instantiations of our framework. We described SAMBO, a system that is developed according to the framework and that implements different strategies. Within this system we have implemented some already existing alignment strategies as well as some new strategies. Further, the framework and SAMBO can be used to experiment with combinations of strategies. This is a first step towards a general framework that can be used for comparative evaluations of alignment strategies. In this paper we experimented with different strategies and their combinations and showed results for well-known biomedical ontologies. We evaluated these strategies and their combinations in terms of quality of the suggestions and processing time. We also compared SAMBO with two other systems.

In the future we will extend our work in different ways. We started to create a toolkit (KitAMO) based on our framework and our experience with SAMBO [23]. The toolkit can be used for evaluating alignment strategies and their combinations using different types of ontologies. This will result in recommendations on which (combinations of) strategies are well suited for aligning which kinds of ontologies. Another track is to further develop SAMBO. We have already started to work on integrating an ontology visualization tool into SAMBO that will improve the way information is provided to the users, but it may also lead to the development of new alignment strategies. Further, we will improve existing matchers, develop new matchers and evaluate them. We also started work on improved filtering

	linguistic	structure	constraints	instances	auxiliary
ArtGen	name	parents, children		domain-specific documents	WordNet
ASCO	name, label, description	parents, children, siblings, path from root			WordNet
Chimaera	name	parents, children			
FCA-Merge	name			domain-specific documents	
FOAM	name, label	parents, children	equivalence		
GLUE	name	neighborhood		instances	
HCONE	name	parents, children			WordNet
IF-Map				instances	a reference ontology
iMapper		leaf, non-leaf, children, related node	domain, range	instances	WordNet
OntoMapper	name	parents, children		documents	
(Anchor-) PROMPT	name	direct graphs			
SAMBO	name, synonym	is-a and part-of, descendants and ancestors		domain-specific documents	WordNet, UMLS
S-Match	label	path from root	semantic relations codified in labels		WordNet

Table 9: Knowledge used by alignment systems.

mechanisms [2].

Acknowledgements

We thank Vaida Jakonienė for comments on the paper and SAMBO and Bo Servenius for discussions and comments on SAMBO. We also acknowledge the financial support of the Center for Industrial Information Technology, the Swedish Research Council, the Swedish national graduate school in computer science, and the EU Network of Excellence REWERSE (Sixth Framework Programme project 506779).

References

[1] J. Bard, M. Kaufman, C. Dubreuil, R. Brune, A. Burger, R. Baldock and D. Davidson, An internet-accessible database of mouse developmental anatomy based on a systematic nomenclature, Mechanisms of Development 74 (1998) 111-120.

[2] B. Chen, H. Tan and P. Lambrix, Structure-based filtering for ontology alignment, in: Proceedings of the IEEE WETICE Workshop on Semantic Technologies in Collaborative Applications, 2006.

[3] F. Collins, E. Green, A. Guttmacher and M. Guyer, A vision for the future of genomics research, Nature 422 (2003) 835-847.

[4] A. Doan, J. Madhavan, P. Domingos and A. Halevy, Ontology matching: A machine learning approach, in: S. Staab and R. Studer, eds., Handbook on Ontologies in Information Systems, Springer, 2003, pp. 397-416.

[5] M. Ehrig, P. Haase, N. Stojanovic and M. Hefke, Similarity for Ontologies - A Comprehensive Framework in: 13th European Conference on Information Systems, 2005.

[6] M. Ehrig and S. Staab, QOM - quick ontology mapping, in: Proceedings of 3rd International Semantic Web Conference, LNCS 3298, Springer, 2004, pp. 683-697.

[7] Entrez. http://www.ncbi.nlm.nih.gov/Database/index.html

[8] J. Euzenat, Introduction to the EON ontology alignment context, in: Proceedings of the 3rd International Workshop on the Evaluation of Ontology-based Tools, 2004.

[9] J. Euzenat, D. Loup, M. Touzani and P. Valtchev, Ontology alignment with OLA, in: Proceedings of the 3rd International Workshop on the Evaluation of Ontology-based Tools, 2004.

[10] T. Hoshiai, Y. Yamane, D. Nakamura and H. Tsuda, A semantic category matching approach to ontology alignment, in: Proceedings of the 3rd International Workshop on the Evaluation of Ontology-based Tools, 2004.

[11] FaCT. http://www.cs.man.ac.uk/∼horrocks/FaCT/

[12] FOAM. http://www.aifb.uni-karlsruhe.de/WBS/meh/foam/

[13] The Gene Ontology Consortium, Gene Ontology: tool for the unification of biology, Nature Genetics 25(1) (2000) 25-29. http://www.geneontology.org/.

[14] A. Gómez-Pérez, Ontological Engineering: A state of the Art, Expert Update 2(3) (1999) 33-43.

[15] F. Giunchiglia, P. Shvaiko and M. Yatskevich, S-Match: an algorithm and an implementation of semantic matching, in: Proceedings of the European Semantic Web Symposium, LNCS 3053, Springer, 2004, pp. 61-75.

[16] T. Hayamizu, M. Mangan, J. Corradi, J. Kadin and M. Ringwald, The Adult Mouse Anatomical Dictionary: a tool for annotating and integrating data, Genome Biology 6(3) (2005) R29.

[17] K. Kotis and G.A. Vouros, The HCONE Approach to Ontology Merging, in: Proceedings of the First European Semantic Web Symposium, LNCS 3053, Springer, 2004, pp. 137-151.

[18] P. Lambrix, Ontologies in Bioinformatics and Systems Biology, in: W. Dubitzky and F. Azuaje, eds., Artificial Intelligence Methods and Tools for Systems Biology, Springer, 2004, chapter 8, pp. 129-146.

[19] P. Lambrix, Towards a Semantic Web for Bioinformatics using Ontology-based Annotation, in: Proceedings of the 14th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises, 2005, pp. 3-7. Invited talk.

[20] P. Lambrix and A. Edberg, Evaluation of ontology merging tools in bioinformatics, in: Proceedings of the Pacific Symposium on Biocomputing, 8 (2003) 589-600.

[21] P. Lambrix and H. Tan, Merging DAML+OIL Ontologies, in: J. Barzdins and A. Caplinskas, eds., Databases and Information Systems - Selected Papers from the Sixth International Baltic Conference on Databases and Information Systems, IOS Press, 2005, pp. 249-258.

[22] P. Lambrix and H. Tan, A Framework for Aligning Ontologies, in: Proceedings of the Third International Workshop on Principles and Practice of Semantic Web Reasoning, LNCS 3703, Springer, 2005, pp. 17-31.

[23] P. Lambrix and H. Tan, A Tool for Evaluating Ontology Alignment Strategies, Journal on Data Semantics VIII (2006).

[24] B.T. Le, R. Dieng-Kuntz and F. Gandon, On ontology matching problem (for building a corporate semantic web in a multi-communities organization), in: Proceedings of the 6th International Conference on Enterprise Information Systems, 2004.

[25] D. McGuinness, R. Fikes, J. Rice and S. Wilder, An Environment for Merging and Testing Large Ontologies, in: Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning, 2000, pp. 483-493.

[26] Medical Subject Headings. http://www.nlm.nih.gov/mesh/

[27] MEDLINE. http://www.nlm.nih.gov/databases/databases_medline.html

[28] P. Mitra and G. Wiederhold, Resolving terminological heterogeneity in ontologies, in: Proceedings of the ECAI Workshop on Ontologies and Semantic Interoperability, 2002.

[29] N.F. Noy and M. Musen, PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment, in: Proceedings of the Seventeenth National Conference on Artificial Intelligence, 2000, pp. 450-455.

[30] N.F. Noy and M. Musen, Anchor-PROMPT: Using Non-Local Context for Semantic Matching, in: Proceedings of the IJCAI Workshop on Ontologies and Information Sharing, 2001, pp. 63-70.

[31] N.F. Noy and M.A. Musen, The PROMPT Suite: Interactive Tools For Ontology Merging And Mapping, International Journal of Human-Computer Studies 59(6) (2003) 983-1024.

[32] Open Biomedical Ontologies. http://obo.sourceforge.net/

[33] OntoWeb Consortium, Deliverables 1.3 (A survey on ontology tools) and 1.4 (A survey on methodologies for developing, maintaining, evaluating and reengineering ontologies), 2002. http://www.ontoweb.org

[34] OWL. http://www.w3.org/2004/OWL/

[35] S. Prasad, Y. Peng and T. Finin, Using Explicit Information To Map Between Two Ontologies, in: Proceedings of the AAMAS Workshop on Ontologies in Agent Systems, 2002.

[36] Protégé. http://protege.stanford.edu/index.html

[37] PubMed. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi

[38] REWERSE. EU Network of Excellence on Reasoning on the Web with Rules and Semantics. Deliverables of the A2 Working Group, 2004-2006. http://www.rewerse.net/

[39] Racer. http://www.cs.concordia.ca/~haarslev/racer/

[40] Sayers E, Wheeler D. Building Customized Data Pipelines Using the Entrez Programming Utilities (eUtils). *NCBI Coursework*, http://www.ncbi.nlm.nih.gov/

[41] B. Smith, W. Ceusters, B. Klagges, J. Köhler, A. Kumar, J. Lomax, C. Mungall, F. Neuhaus, A. Rector and C. Rosse, Relations in biomedical ontologies, Genome Biology 6 (2005) R46.

[42] SOFG, Standards and Ontologies for Functional Genomics, 2002. http://www.sofg.org/

[43] R. Stevens, C. Goble and S. Bechhofer, Ontology-based knowledge representation for bioinformatics, Briefings in Bioinformatics 1(4) (2000) 398-414.

[44] G. Stumme and A. Mädche, FCA-Merge: Bottom-up merging of ontologies, in: Proceedings of the 17th International Joint Conference on Artificial Intelligence, 2001, pp. 225-230.

[45] X.M. Su, S. Hakkarainen and T. Brasethvik, Semantic enrichment for improving systems interoperability, in: Proceedings of the 2004 ACM Symposium on Applied Computing, 2004, pp. 1634-1641.

[46] T. Takai-Igarashi, Y. Nadaoka and T. Kaminuma, A Database for Cell Signaling Networks, Journal of Computational Biology 5(4) (1998) 747-754.

[47] T. Takai-Igarashi and T. Takagi, SIGNAL-ONTOLOGY: Ontology for Cell Signalling, Genome Informatics 11 (2000) 440-441.

[48] H. Tan, V. Jakonienė, P. Lambrix, J. Aberg and N. Shahmehri, Alignment of biomedical ontologies using life science literature, in: Proceedings of the International Workshop on Knowledge Discovery in Life Science Literature, LNBI 3886, Springer, 2006, pp. 1-17.

[49] UMLS. http://www.nlm.nih.gov/research/umls/about_umls.html

[50] WordNet. http://wordnet.princeton.edu/

[51] Y. Kalfoglou and M. Schorlemmer, IF-Map: an ontology mapping method based on information flow theory, Journal of data semantics 1 (2003) 98-127.