Measuring the Structural Preservation of Semantic Hierarchy Alignments

Cliff A Joslyn, Patrick Paulson, and Amanda White

Pacific Northwest National Laboratory, Richland, WA, USA
`{cjoslyn,patrick.paulson,amanda.white}@pnl.gov`

Abstract. We present a method to measure the amount of structural distortion carried by an alignment between two taxonomic cores of ontologies represented as semantic hierarchies. We present our formalism based in metric order theory. We then illustrate the results of such an analysis on the Anatomy track of the 2008 Ontology Alignment Evaluation Initiative (OAEI).

Key words: Ontology alignment; lattice theory; order theory.

1 Introduction

Since top-down, monolithic development of unitary ontologies is at best difficult, and at worst undesirable, ontology alignment is increasingly seen as a critical Semantic Web technology [4, 17]. Although many semantic relations can be present in ontologies, they tend to be dominated by their taxonomic cores; that is, subsumptive inheritance (`is-a`) and/or meronomic compositional (`part-of`) class hierarchies. Thus techniques which address the specific nature of these structures as *semantic hierarchies* are critical for ontology management tasks.

An alignment is modeled as a mapping (single- or multi-valued) between two semantic hierarchies, taking concepts from one into another. Depending on the relative size, structure, and domains of the two hierarchies, their quality, and the size and quality of the alignment, different properties of the alignment might hold. It might be that that mapping is partial in one direction or the other; it may be concentrated in one portion or another of each hierarchy; may takes nodes which are "close together" in one hierarchy into nodes which are "far apart" in the other; and may take nodes in a particular structural relationship (e.g. parent-child or sibling) into the same or a different such structural relationship. Knowledge of such properties is valuable for the ontology designer and aligner, an important adjunct to visual inspection of large ontologies and alignments.

One straightforward example of this reasoning is to say that if the two semantic hierarchies were intended to model the same domain, then an alignment mapping should be structure-preserving, taking pairs of nodes which are close together in one structure into pairs which are also close together in the other, and similarly for pairs of nodes which are far apart. To the extent that this is not the case, this could indicate a problem with either one ontology, the other, the alignment mapping, or some combination of these structures. Even when semantic or pragmatic criteria dictate that it is appropriate for a mapping to violate structural preservation, it is still valuable to be able to *measure* and *quantify* the amount of structural preservation or distortion which an alignment introduces.

This is true both after the alignment has been produced, and also *while* the alignment is being produced, for example in an interactive environment such as the Protege tool PROMPT [17].

We describe an algorithmic approach to the measurement of the extent to which an ontology alignment preserves the structural properties of the two ontologies. We use **order theory** (the formal theory of hierarchy represented by ordered sets and lattices [3]) to model taxonomies as semantic hierarchies on sets of nodes P, where nodes $a \in P$ are ontology concepts related by transitive edges such as subsumption ("`is-a`") or composition ("`part-of`"). These in turn are represented as finite, bounded, partially ordered sets (posets) $\mathcal{P} = \langle P, \leq \rangle$, where the relation $\leq$ is one (or a union) of these transitive link types. Such ordered structures are not, in general, trees, nor even lattices, but can be rich in multiple inheritance and lack unique least common subsumers between nodes.

We demonstrate our approach by analyzing the alignments of the Anatomy track of the 2008 Ontology Alignment Evaluation Initiative (OAEI) campaign (http://oaei.ontologymatching.org/2008/anatomy). We compare the precision and recall results of the OAEI against our discrepancy measures, as well as analyze the highest discrepancy nodes and alignment links.

Prior work in both ontology alignment in general, and graph matching in knowledge systems (e.g. [7]), is voluminous, and order theory is used in many areas of computer science outside of knowledge systems. But there is relatively little in the ontology literature about measuring structural relations in ontologies, and we've been able to find nothing in the specific use of a lattice theoretical approach to hierarchy mapping and measurement. Kalfoglou and Schorlemmer [12] have an approach to order morphisms similar to ours; and some researchers [5, 17] take a structure mapping approach, but do so as a graph theory problem, not using hierarchy theory. Although He and Xiaoyong [8] recognize the need to work in order theory, they don't actually do so.

The algebraic relations among class extents and intents used by a number of researchers (e.g. [14, 15]) do point to metric properties similar to ours. But while these have implications for an order-theoretical approach, they are not themselves explicitly order-theoretical. The closest correlate to our order metric approach is in the use of "semantic similarity" measures [1]. Still, these are generally used *within* a particular lexical or bio-ontology, and have only been used to a small extent [19] as an adjunct to the alignment problem. Some of our work [10] marries structural similarities with our order metrics. We are actively working [18] to identify how our order metrics are actually foundational to semantic similarities, and generate them as a special case.

An early desription of this concept has been previously reported in a poster [11].

2   Order Theory for Semantic Hierarchy Alignment

We represent semantic hierarchies as bounded, partially ordered sets (posets) $\mathcal{P} = \langle P, \leq \rangle$ [3], where P is a finite set of ontology nodes, and $\leq \subseteq P^2$ is a reflexive, anti-symmetric, and transitive binary relation such as subsumption ("`is-a`")

or composition ("part-of"). In ontology analysis, semantic hierarchies are typically Directed Acyclic Graphs (DAGs) [9] which are top-bounded, have a moderate amount of multiple inheritance, and branch downward very strongly. Each such structure uniquely determines a poset $\mathcal{P}$ by taking its transitive closure and including a bottom bound $0 \in P$ such that $\forall a \in P, 0 \leq a$.

For two taxonomies $\mathcal{P} := \langle P, \leq \rangle, \mathcal{P}' := \langle P', \leq' \rangle$, an **alignment relation** $F \subseteq P \times P'$ is a collection of pairs $\boldsymbol{f} = \langle a, a' \rangle \in F$, indicating that the node $a \in P$ on the "left" side is mapped or aligned to the node $a' \in P'$ on the "right" side. F determines a domain and codomain

$$Q := \{a \in P, \exists a' \in P', \langle a, a' \rangle \in F\} \subseteq P, \quad Q' := \{a' \in P', \exists a \in P, \langle a, a' \rangle \in F\} \subseteq P',$$

We call the $\boldsymbol{f} \in F$ **links**, the $a \in Q$ the **left anchors** and the $a' \in Q'$ the **right anchors**. Let $m := |Q|, m' := |Q'|$, and $N := |F| \leq mm'$.

Fig. 1 shows a small alignment. We have left anchors $Q = \{B, E, G\}, m = 3$; right anchors $Q' = \{I, J, K\}, m' = 3$; and $N = 4$ with links $F = \{\boldsymbol{f}_1 = \langle B, J \rangle, \boldsymbol{f}_2 = \langle B, I \rangle, \boldsymbol{f}_3 = \langle E, I \rangle, \boldsymbol{f}_4 = \langle G, K \rangle\}$.
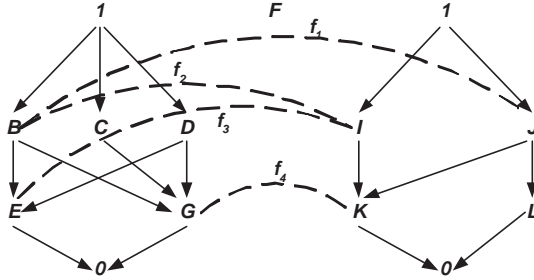
Fig. 1. An example of two semantic hierarchies and an alignment relation.

Let d be a metric on $\mathcal{P}$ and $\mathcal{P}'$. For links $\boldsymbol{f} = \langle a, a' \rangle, \boldsymbol{g} = \langle b, b' \rangle \in F$ to participate well in a good structural mapping between $\mathcal{P}$ and $\mathcal{P}'$, we want the metric relations between the $a, b \in Q$ to be the same as their corresponding $a', b' \in Q'$, so that $|\bar{d}(a, b) - \bar{d}'(a', b')|$ is small. In our example, F takes both B and E, which are somewhat distant in $\mathcal{P}$, to the single node I in $\mathcal{P}'$, so that there is no distance between them on the right. This is not preferred.

We now consider our metric d needed to compare the distances $d(a, b), d(a', b')$ between pairs of nodes $a, b \in P$ on one side of an alignment and their images $a', b' \in P$ on another. The knowledge systems literature has focused on **semantic similarities** [1] to perform a similar function, which are available when $\mathcal{P}$ is equipped with a probabilistic weighting function $p \colon P \to [0, 1]$, with $\sum_{a \in P} p(a) = 1$. p can be derived, for example, from the frequency with which terms appear in documents (for the case of the Wordnet [6] thesaurus), or which genes are annotated to bio-ontology nodes (in the case of the Gene Ontology [13]).

Our purpose is more general, since we may not have such a weighting function available, and semantic similarities are not required to be metrics satisfying the triangle inequality. In seeking out the proper mathematical grounding, we turn to **order metrics** [16, 18] which can use, but do not require, a quantitative

weighting, and always yield a metric. For details about order metrics built from isotone and antitone lower and upper semimodular functions on ordered sets, see [18]. In this work, we use the **upper and lower cardinality-based distances**

$$d_u(a,b) = |\uparrow a| + |\uparrow b| - 2 \max_{c \in a \vee b} |\uparrow c|, \qquad d_l(a,b) = |\downarrow a| + |\downarrow b| - 2 \max_{c \in a \wedge b} |\downarrow c|,$$

where for a node $a \in P$, its **upset** $\uparrow a := \{b \geq a\}$ and **downset** $\downarrow a := \{b \leq a\}$ are all its ancestors and successors respectively, so that $|\uparrow a|, |\downarrow a|$ are the number of ancestors and successors. The generalized join and meet are

$$a \vee b := \mathrm{Min}(\uparrow a \cap \uparrow b) \subseteq P, \qquad a \wedge b := \mathrm{Max}(\downarrow a \cap \downarrow b) \subseteq P,$$

where for a set of nodes $Q \subseteq P$ the **upper bounds** and **lower bounds** are

$$\mathrm{Min}(Q) := \{a \in Q : \nexists b \in Q, b < a\} \subseteq P, \mathrm{Max}(Q) := \{a \in Q : \nexists b \in Q, b > a\} \subseteq P.$$

We need to normalize distance to the size of the structure, so that we are measuring the relative proportion of the overall structure two nodes are apart, or in other words, what proportion of their potential maximum distance. These **normalized upper and lower distances** are

$$\bar{d}_u(a,b) := \frac{d_u(a,b)}{|P|-1} \in [0,1], \quad \bar{d}_l(a,b) := \frac{d_l(a,b)}{|P|-1} \in [0,1].$$

Considering the difference between upper and lower distance, it may at first appear to be more natural to use upper distance, since we're then "looking upwards" towards the top bound $1 \in P$ which almost always exists in the given structure. Moreover, it is sometimes the case that the upper distance $d_u(a,b)$ is the same as the minimum (undirected) path length between a and b (a favorite graph metric), but this is only required to be true when $\mathcal{P}$ is an upper-bounded tree: in general, path length and these metrics are unrelated.

When $\mathcal{P}$ is top-bounded and strongly down-branching (as in our cases), then it is preferable to use lower distance (this is possible because we always provide a lower bound $0 \in P$). One reason for this is that since semantic hierarchies are much more strongly down-branching than up-branching, up-sets are typically very small and narrow, frequently single chains; where down-sets are large, branching structures. Additionally, this allows siblings deep in the hierarchy to be closer together than siblings high in the hierarchy (this will be demonstrated below). This is considered valuable, for example, where e.g. "mammal" and "reptile" are considered farther apart than "horse" and "goat".

In Fig. 1, to calculate the lower distance $d_l(B,C)$, we have $|\downarrow B| = 4, |\downarrow C| = 3, B \wedge C = \{G, 0\}$, $\max_{c \in B \wedge C} |\downarrow c| = \max(1,2)$, so that $d_l(B,C) = 4 + 3 - 2 \times 2 = 3$. Finally, we have $|P| = 7$, so that $\bar{d}_l(B,C) = 1/2$. Table 1 shows distances $d_l(a,b)$ on the left in $\mathcal{P}$, and Table 2 shows distances $d_l(a',b')$ on the right in $\mathcal{P}'$. $|P| = 6$ and $|P'| = 5$, yielding Tables 3 and 4 showing the relative distances. Note that siblings high in the structure are farther apart than those lower, for example $\bar{d}_l(B,C) = 0.50, \bar{d}_l(E,G) = 0.33$, and $\bar{d}_l(I,J) = 0.60, \bar{d}_l(K,L) = 0.40$. Contrast this with the similar relative *upper* distances, shown in Table 5, where siblings lower in the structure are further apart.

$d_l(a,b)$	1	B	C	D	E	G	0
1	0	3	4	3	5	5	6
B	3	0	3	4	2	2	3
C	4	3	0	3	3	1	2
D	3	4	3	0	2	2	3
E	5	2	3	2	0	2	1
G	5	2	1	2	2	0	1
0	6	3	2	3	1	1	0

Table 1. Left lower distances $d_l(a,b)$

$d_l(a',b')$	1	I	J	K	L	0
1	0	3	2	4	4	5
I	3	0	3	1	3	2
J	2	3	0	2	2	3
K	4	1	2	0	2	1
L	4	3	2	2	0	1
0	5	2	3	1	1	0

Table 2. Right lower distances $d_l(a',b')$.

$\bar{d}_l(a,b)$	1	B	C	D	E	G	0
1	0.00	0.50	0.67	0.50	0.83	0.83	1.00
B	0.50	0.00	0.50	0.67	0.33	0.33	0.50
C	0.67	0.50	0.00	0.50	0.50	0.17	0.33
D	0.50	0.67	0.50	0.00	0.33	0.33	0.50
E	0.83	0.33	0.50	0.33	0.00	0.33	0.17
G	0.83	0.33	0.17	0.33	0.33	0.00	0.17
0	1.00	0.50	0.33	0.50	0.17	0.17	0.00

Table 3. Left lower relative distances $\bar{d}_l(a,b)$

$\bar{d}_l(a',b')$	1	I	J	K	L	0
1	0.00	0.60	0.40	0.80	0.80	1.00
I	0.60	0.00	0.60	0.20	0.60	0.40
J	0.40	0.60	0.00	0.40	0.40	0.60
K	0.80	0.20	0.40	0.00	0.40	0.20
L	0.80	0.60	0.40	0.40	0.00	0.20
0	1.00	0.40	0.60	0.20	0.20	0.00

Table 4. Right lower relative distances $\bar{d}_l(a',b')$

Let d be a metric used in both $\mathcal{P},\mathcal{P}'$, in our case, the lower distance d_l. Then the **link discrepancy** is given by $\delta(\boldsymbol{f},\boldsymbol{g}) := |\bar{d}(a,b) - \bar{d}(a',b')|$, and the **distance discrepancy induced by F between $\mathcal{P}$ and $\mathcal{P}'$ given d** is

$$D(F) := \frac{\sum_{\boldsymbol{f},\boldsymbol{g} \in F} \delta(\boldsymbol{f},\boldsymbol{g})}{\binom{N}{2}}.$$

$D \in [0,1]$, with $D = 0$ iff F is completely distance preserving, and $D = 1$ if F is maximally distance distorting, e.g. mapping diameters to equality, and neighbors and children to diameters. Table 7 shows the discrepancies δ comparing links against each other, yielding total distance discrepancy $D(F) = 0.26$.

$\bar{d}_u(a,b)$	1	B	C	D	E	G	0
1	0.00	0.17	0.17	0.17	0.50	0.67	1.00
B	0.17	0.00	0.33	0.33	0.33	0.50	0.83
C	0.17	0.33	0.00	0.33	0.67	0.50	0.83
D	0.17	0.33	0.33	0.00	0.33	0.50	0.83
E	0.50	0.33	0.67	0.33	0.00	0.83	0.50
G	0.67	0.50	0.50	0.50	0.83	0.00	0.33
0	1.00	0.83	0.83	0.83	0.50	0.33	0.00

Table 5. Left upper relative distances $\bar{d}_u(a,b)$

$\bar{l}(a,b)$	1	B	C	D	E	G	0
1	0.00	0.33	0.33	0.33	0.67	0.67	1.00
B	0.33	0.00	0.67	0.67	0.33	0.33	0.67
C	0.33	0.67	0.00	0.67	1.00	0.33	0.67
D	0.33	0.67	0.67	0.00	0.33	0.33	0.67
E	0.67	0.33	1.00	0.33	0.00	0.67	0.33
G	0.67	0.33	0.33	0.33	0.67	0.00	0.33
0	1.00	0.67	0.67	0.67	0.33	0.33	0.00

Table 6. Normalized minimum undirected path length.

We wish to understand the contribution which particular links and anchors make to the overall discrepancy. So we aggregate discrepancies over links $\boldsymbol{f},\boldsymbol{g} \in F$, normalized by the number of links; and over left and right anchors $a \in Q, a' \in Q'$, normalized by the number of left and right anchors respectively (results for our example are shown in Tables 8 and 9):

	$f_1 = \langle B, J\rangle$	$f_2 = \langle B, I\rangle$	$f_3 = \langle E, I\rangle$	$f_4 = \langle G, K\rangle$
$f_1 = \langle B, J\rangle$	0.00	0.60	0.27	0.07
$f_2 = \langle B, I\rangle$	0.60	0.00	0.33	0.13
$f_3 = \langle E, I\rangle$	0.27	0.33	0.00	0.13
$f_4 = \langle G, K\rangle$	0.07	0.13	0.13	0.00

Table 7. Distance discrepancy $\delta(f_i, f_j)$.

$$D(f) := \frac{\sum_{g \in F} \delta(f, g)}{N - 1}, \; D(a) := \frac{\sum_{\langle a, a'\rangle \in F} D(\langle a, a'\rangle)}{m}, \; D(a') := \frac{\sum_{\langle a, a'\rangle \in F} D(\langle a, a'\rangle)}{m'}.$$

Because we use lower distance, links high in the structure are further apart, for example $\delta(\langle B, I\rangle, \langle B, J\rangle) = 0.60$, since the identical pair $\langle B, B\rangle$ which are zero apart are taken to the nodes $\langle I, J\rangle$ high in the structure; while $\delta(f_1, f_4) = 0.07$, since $\langle B, G\rangle$ are almost as close on the left as $\langle J, K\rangle$ on the right. The link $f_2 = \langle B, I\rangle$ is the greatest contributor to distance discrepancies, as are its anchors $B \in P, I \in P'$. This result is slightly counterintuitive, but instructive. Considering *link* comparisons in Fig. 1: comparing f_1 to f_3, for example, the *differences* in the distances between their left and right anchors is smaller than the similar difference comparing the left and right anchors f_2 and f_3.

3   Analysis of the 2008 OAEI Anatomy Track

We now describe the application of this alignment evaluation technology against the Anatomy track of the 2008 OAEI campaign [2]. In the OAEI, one or more "gold standard" reference alignments are developed (in part, by hand) between pairs of ontologies. The community is challenged to submit alignments, and their quality is measured by calculating the precision, recall, and F-score of the matches between nodes made by the submitted alignments against those matches made by the reference alignment. We calculated distance discrepancies for the alignments in the challenge track, including the references. We compared the discrepancy scores of the submitted alignments to each other, and to the references, and correlated the precision and recall results of the submitted alignments against their discrepancies.

	$D(f_i)$
$f_1 = \langle B, J\rangle$	0.31
$f_2 = \langle B, I\rangle$	0.35
$f_3 = \langle E, I\rangle$	0.24
$f_4 = \langle G, K\rangle$	0.11

Table 8. Aggregate distance discrepancy by link $D(f)$.

	I	J	K	$D(a)$
B	0.35	0.31		0.22
E	0.24			0.08
G			0.11	0.04
$D(a')$	0.20	0.10	0.04	

Table 9. Aggregate distance discrepancy by anchor $D(a), D(a')$.

3.1   The Anatomy Track Ontologies and Alignments

The OAEI-2008 Anatomy track was selected due to its sufficient size, moderate amount of complexity and multiple inheritance, and publicly available partial reference alignment. It included the 2744 classes of the Adult Mouse Anatomy

(MA, http://www.informatics.jax.org/searches/AMA_form.shtml) and the portion of the 3304 classes from the NCI Thesaurus (NCIT)[1] describing human anatomy.

The full reference alignment was provided by the Mouse Genome Informatics group at the Jackson Laboratory. The partial reference alignment had 988 links, derived from 934 purely lexical matches and 54 additional links from the full reference (http://oaei.ontologymatching.org/2008/results/anatomy). There were multiple tasks in the anatomy track, and we focused on Task 1, which was to maximize the F-score of an alignment. We also focused on the nine submitted alignments with code names shown in Table 10.

A statistical analysis of MA and NCIT shows structures which are somewhat similar in size (2744 vs. 3304 nodes, respectively), "edge density" (1.04 vs. 1.14 links/edge), and "leaf density" (82.3% vs. 79.6% of the nodes being leaves). But MA is dramatically shorter, with a height (maximum chain length from the top 1 to bottom 0) of 8 compared to 14. NCIT is more complex, with dramatically more multiple inheritance (4.0% vs. 13.2% of nodes with more than one parent).

3.2   Discrepancy Measurement Results

Table 10 lists basic statistics for all alignments, and also shows the number of anchors and links, the discrepancies, and (for the submitted alignments) the precision P, recall R, and F-score $2PR/(P+R)$. Fig. 2 shows distance discrepancies with the number of anchors and links on the right Y axis; Fig. 3 shows them with P, R and F-score on the right Y axis.

Alignment	m	m'	N	$D(F)$	P	R	F-score
reference_partial	986	984	988	0.08%			
reference_full	1501	1504	1523	0.11%			
aflood	1186	1186	1186	0.11%	87.4%	68.2%	76.6%
AROMA	1062	1062	1062	0.36%	80.3%	56.0%	66.0%
ASMOV	1261	1261	1261	0.09%	78.7%	65.2%	71.3%
DSSim	1170	1086	1545	2.02%	61.6%	62.4%	62.0%
Lily	1324	1324	1324	0.13%	79.6%	69.3%	74.1%
RiMOM	1205	1205	1205	0.10%	92.9%	73.5%	82.1%
SAMBO	1465	1465	1465	0.09%	86.9%	83.6%	85.2%
SAMBOdtf	1527	1527	1527	0.10%	83.1%	83.3%	83.2%
TaxoMap	2533	1279	2533	1.40%	46.0%	76.4%	57.4%

Table 10. Discrepancy results for anatomy track alignments. $D(F)$ = lower distance discrepancy; N = # links; P = precision; R = recall; F = F-score.

Generally, discrepancies are low, especially for the two reference alignments, except for AROMA, DSSIM, and Taxomap. These are also the worst performers, and DSSIM and Taxomap have the biggest difference between number of anchors and links. Fig. 4 shows distance discrepancy $D(F)$ against F-score. Significant discrepancy is an indication of poor F-score, and conversely high F-score requires effectively no discrepancy: Pearson correlation between D and F-score $-.780$.

Table 11 shows the top nine links by aggregate discrepancy for the partial and full reference alignments, and the two worst-scoring alignment by both F-score and discrepancy. As illustrated in Table 9, aggregation of discrepancy by anchor is most valuable when the alignment F is not very one-to-one. This is the

---

[1] http://ncicb.nci.nih.gov/NCICB/infrastructure/cacore_overview/vocabulary
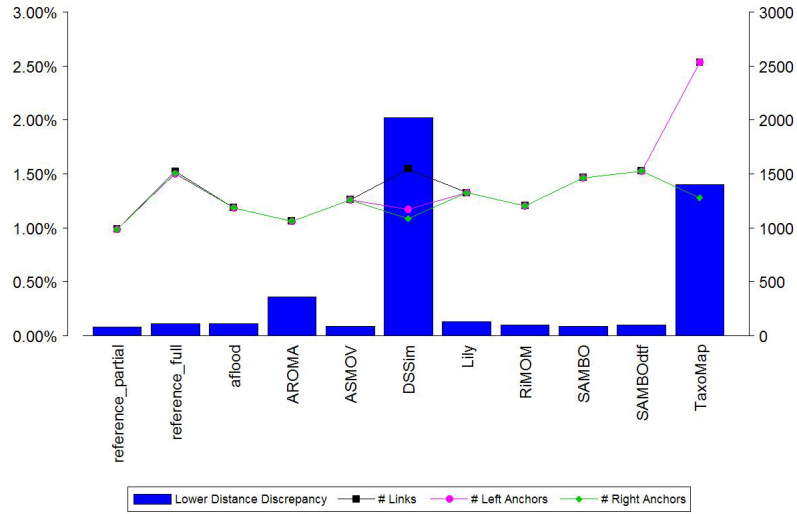
Fig. 2. Discrepancy against number of anchors and links.
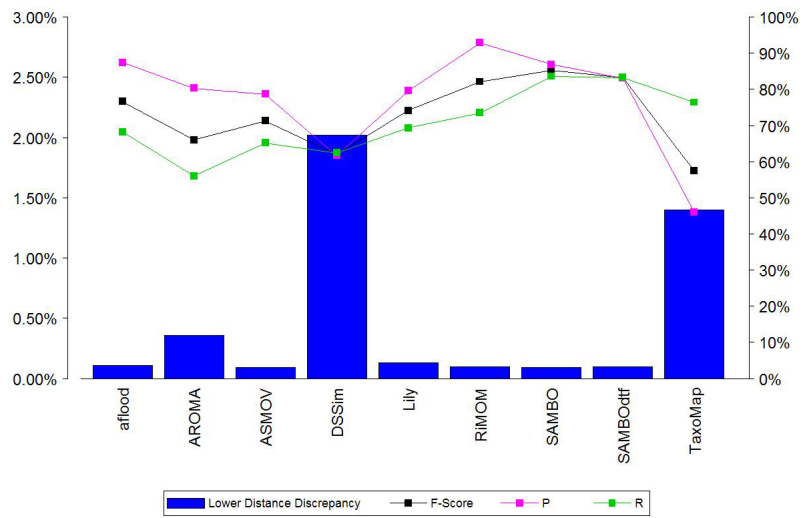
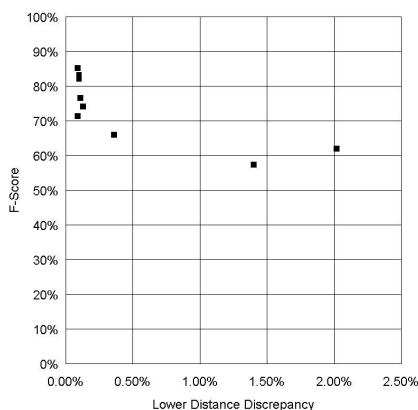**Fig. 3.** Discrepancy against precision, recall, and *F*-score.

Fig. 4. Anatomy track alignments, distance discrepancy against F-score.

case with the two reference alignments, so Table 13 shows the top nine aggregate discrepancies by left- and right-anchors for DSSIM and Taxomap.

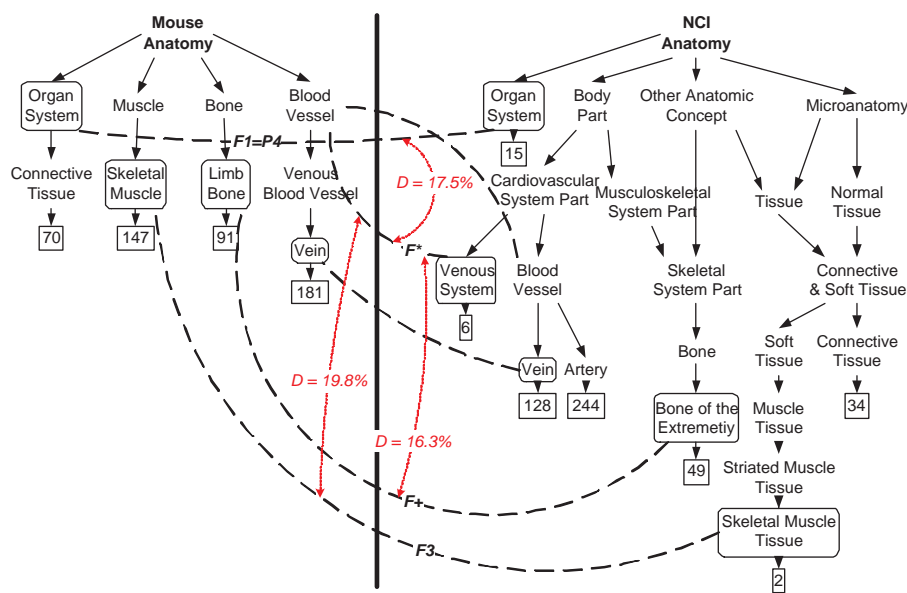| Partial Reference			Full Reference		
D	**MA**	**NCIT**	D	**MA**	**NCIT**
3.05%	organ system	Organ_System	17.30%	blood vessel	Blood_Vessel
2.81%	blood vessel	Blood_Vessel	6.78%	venous blood vessel	Venous_System
2.73%	vein	Vein	5.32%	skeletal muscle	Skeletal_Muscle_Tissue
1.53%	connective tissue	Connective_Tissue	3.07%	organ system	Organ_System
1.38%	bone	Bone	2.74%	vein	Vein
1.00%	artery	Artery	1.86%	limb bone	Bone_of_the_Extremity
0.97%	foot bone	Foot_Bone	1.68%	vertebra	Vertebra
0.71%	lymphoid tissue	Lymphoid_Tissue	1.57%	head/neck muscle	Head_and_Neck_Muscle
0.68%	ligament	Ligament	1.55%	connective tissue	Connective_Tissue
0.67%	muscle	Muscle	1.39%	bone	Bone
DSSIM			**Taxomap**		
D	**MA**	**NCIT**	D	**MA**	**NCIT**
62.82%	joint	Body_Part	11.72%	tail blood vessel	Blood_Vessel
15.66%	cardiovascular system	Cardiovascular_System_Part	11.72%	foot blood vessel	Blood_Vessel
13.02%	capillary	Blood_Vessel	11.72%	neck blood vessel	Blood_Vessel
11.04%	bone	Loose_Connective_Tissue	11.72%	head blood vessel	Blood_Vessel
9.84%	perineal artery	Perineal_Artery	11.72%	lung blood vessel	Blood_Vessel
9.84%	ethmoidal artery	Artery	11.72%	upper leg blood vessel	Blood_Vessel
8.87%	brachial artery	Brachial_Artery_Branch	11.72%	lower leg blood vessel	Blood_Vessel
8.84%	celiac artery	Artery	11.72%	pelvis blood vessel	Blood_Vessel
8.82%	radial artery	Artery	11.72%	abdomen blood vessel	Blood_Vessel

Table 11. Top nine aggregate link distance discrepancies $D(\boldsymbol{f})$ for four alignments.

Fig. 5 shows a selection of nodes from MA and NCIT, and anchors and links from both the partial and full reference alignments. Numbers below terminal nodes indicate the total number of nodes below them. The top three link discrepancies are shown in Table 12, with labels referring to particular links in Fig. 5. We can see that the biggest discrepancies are between links which take nodes high in MA to nodes low in NCIT. But in fact, our method does not count vertical ranks, but rather the order metrics focus on the numbers of common nodes below the corresponding pairs of anchors.

δ					MA	NCIT
		f			**g**	
		MA	NCIT		MA	NCIT
19.8%	F*	blood vessel	Venous System	F3	skeletal muscle	Skeletal_Muscle_Tissue
17.6%	F*	blood vessel	Venous System	F1=P4	organ system	Organ_System
16.3%	F*	blood vessel	Venous System	F+	limb bone	Bone_of_the_Extremity

Table 12. Highest discrepancies between link pairs shown in Fig. 5, full reference alignment.

Comparing alignments now, while both reference alignments had low discrepancy, the full alignment was generally more discrepant, perhaps through the addition of non-lexical matching links like ⟨venous blood vessel, Venous_System⟩. In DSSIM, clearly the link ⟨joint, Body_Part⟩ is most discrepant. This is because while both Joint and Body_Part are relatively near the tops of MA and NCIT respectively, Joint covers only 21 nodes, while Body_Part covers 2137. This forces that link to be far from all the others, and reveals directly a dramatic difference in structure between the two ontologies. This is then reflected in a very high anchor aggregate score $D(\text{Body\_Part}) = 5.44$. Finally, for Taxomap, we see many links to the NCIT node Blood_Vessel, yielding another high anchor discrepancy of $D(\text{Blood\_Vessel}) = 9.48$. In both cases, the discrepancy measures can point directly to anomolous mappings of high significance.

Fig. 5. Selection of nodes from MA and NCIT, and anchors and links from both the partial and full reference alignments. Table 12 provide details on comparisions of high discrepancy links.

4   Conclusions and Further Work

The results presented here are the first serious attempt to apply this technology to alignment analysis, and are partial and preliminary. Results here may be dependent on the particular properties of the Anatomy track. While a further analysis relating alignment quality to discrepancy awaits, it is suggestive that a discrepancy analysis can reveal to the aligner and ontology designer aspects of their structures not clear from visual inspection. Nor is a robust order theoretical technology limited to discrepancy measures: we can see above that other considerations such as the degree to which alignments are many-to-many, vertical rank structure, degree of multiple inheritance, and a range of other topics in interaction with discrepancies awaits much more serious consideration.

DSSIM				Taxomap			
MA		NCIT		MA		NCIT	
D	a	D	a'	D	a	D	a'
0.830	joint	5.440	Body Part	0.117	tail blood vessel	9.483	Blood Vessel
0.207	cardiovascular system	2.836	Artery	0.117	foot blood vessel	4.093	Muscle
0.172	capillary	1.893	Vein	0.117	neck blood vessel	3.418	Vein
0.165	skeletal muscle	0.985	Bone	0.117	head blood vessel	3.356	Artery
0.146	bone	0.614	Blood Vessel	0.117	lung blood vessel	1.413	Bone
0.130	perineal artery	0.273	Loose Connective Tissue	0.117	upper leg blood vessel	1.215	Connective Tissue
0.130	ethmoidal artery	0.257	Skeletal Muscle Tissue	0.117	lower leg blood vessel	0.719	Other Anatomic Concept
0.117	brachial artery	0.225	Muscle	0.117	abdomen blood vessel	0.649	Skeletal System Part
0.117	celiac artery	0.223	Cardiovascular System Part	0.117	pelvis blood vessel	0.521	Respiratory System

Table 13. Top nine aggregate anchor distance discrepancies $D(a)$ for DSSIM and Taxomap alignments.

As part of a broader infrastructure for the analytical management of ontologies and alignments, further development of these methods is required. Nonetheless, these results suggest that minimizing discrepancy may be related to alignment quality. Thus discrepancy may be an important adjunct to alignment evaluation, playing a role as an automatic pre-filter for hand-built alignments. Moreover, the detailed examination of how particular links and anchors participate with respect to discrepancy within an overall alignment should have high utility for knowledge managers and ontology engineers, revealing details of the nature and structure of the mappings being considered. Perhaps most exciting is the dual problem to that considered here: given an alignment F which is *a priori* believed to be of high quality, how can $D(F)$ be used to aid in the *design* of those ontologies? Some of the results above are very suggestive of these possibilities.

5   Acknowledgements

References

1. Butanitsky, A and Hirst, G: (2006) "Evaluating WordNet-based Measures of Lexical Semantic Relatedness", *Computational Linguistics*, v. **32**:1, pp. 13-47
2. Caracciolo, Caterina; Stuckenschmidt, Heiner; Svab, Ondrej; *et al.*: (2008) "First Results of the Ontology Alignment Evaluation Initiative 2008", in: *Proc. 3rd Int. Wshop. On Ontology Matching (OM2008)*,
3. Davey, BA and Priestly, HA: (1990) *Introduction to Lattices and Order*, Cambridge UP, Cambridge UK, 2nd Edition
4. Euzenat, J and Shvaiko, P: (2007) *Ontology Matching*, Springer-Verlag, Hiedelberg
5. Falconer, Sean M and Maslov, Dmitri: (2006) "Hierarchical Alignment of Weighted Directed Acyclic Graphs", `arVix:cs.DS/0606124v1`
6. Fellbaum, Christiane, ed.: (1998) *Wordnet: An Electronic Lexical Database*, MIT Press, Cambridge, MA
7. Feng, Y; Goldstone, RL; and Menkov, V: (2004) "ABSURDIST II: A Graph Matching Algorithm and its Application to Conceptual System Translation", in: *Proc. 7th Int. Florida AI Research Society Conference (FLAIRS 04)*, v. **2**, pp. 640-645
8. He, Hu and Xiaoyong, Du: (2007) "Ontology Hierarchies Matching by Lattices Alignment", in: *Proc. Ontology Matching 2007 (OM-2007), ISWC 2007*
9. Joslyn, Cliff: (2009) "Hierarchy Analysis of Knowledge Networks", in: *IJCAI Int. Wshop. on Graph Structures for Knowledge Representation and Reasoning*, in press
10. Joslyn, Cliff; Baddeley, Bob; Blake, Judith; *et al.*: (2009) "Automated Annotation-Based Bio-Ontology Alignment with Structural Validation", in: *Proc. Int. Conf. on Biomedical Ontology (ICBO 09)*
11. Joslyn, Cliff; Donaldson, Alex; and Paulson, Patrick: (2008) "Evaluating the Structural Quality of Semantic Hierarchy Alignments", poster at the *Int. Semantic Web Conf. (ISWC 08)*, `http://dblp.uni-trier.de/db/conf/semweb/iswc2008p.html#JoslynDP08`
12. Y Kalfoglou, M Schorlemmer: (2002) "IF-Map: An Ontology-Mapping Method based on Information-Flow Theory", Proc. 1st Int. Conf. Ontologies, Databases and Application of Semantics (ODBASE'02), Irvine, CA, USA
13. Lord, PW; Stevens, Robert; Brass, A; CA Goble: (2003) "Investigating Semantic Similarity Measures Across the Gene Ontology: the Relationship Between Sequence and Annotation", *Bioinformatics*, v. **10**, pp. 1275-1283
14. Maedche, Alexander and Staab, Steffen: (2002) "Measuring Similarity Between Ontologies", in: *Proc. 13th Int. Conf. Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, LNCS*, v. **2473**, pp. 251-263
15. Meilicke, Christian and Stuckenschmidt, Heiner: (2008) "Incoherence as a Basis for Measuring the Quality of Ontology Mappings", in: *Proc. 3rd Int. Wshop On Ontology Matching (OM2008)*, ed. Pavel Shvaiko et al., Karlsruhe
16. Monjardet, B: (1981) "Metrics on Partially Ordered Sets - A Survey", *Discrete Mathematics*, v. **35**, pp. 173-184
17. Noy, N and Musen, MA: (2003) "The PROMPT Suite: Interactive Tools for Ontology Merging and Mapping", *Int. J. Human-Computer Studies*, v. **59**, pp. 983-1024
18. Orum, C and Joslyn, CA: (2009) "Valuations and Metrics on Partially Ordered Sets", `http://arxiv.org/abs/0903.2679v1`, submitted
19. Sanfilippo, A; Posse, C; Gopalan, B; *et al.*: (2007) "Combining Hierarchical and Associative Gene Ontology Relations With Textual Evidence in Estimating Gene and Gene Product Similarity", *IEEE Trans. on Nanobioscience*, v. **6**:1, pp. 51-59
20. Schröder, Bernd SW: (2003) *Ordered Sets*, Birkhauser, Boston