Comparison between ontology distances (preliminary results)

Jérôme David, Jérôme Euzenat

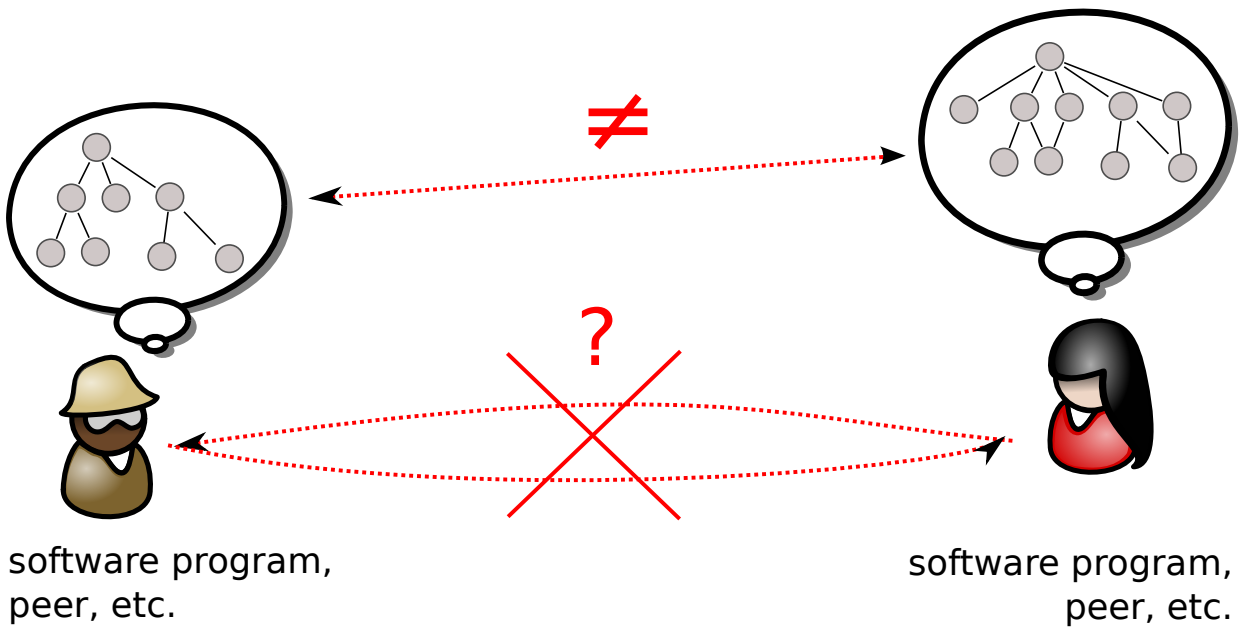
ISWC 2008

Context

A distributed and heterogenous environment: the semantic Web

- Several ontologies on the same domain

Problem: how to facilitate the communication between programs using different ontologies ?

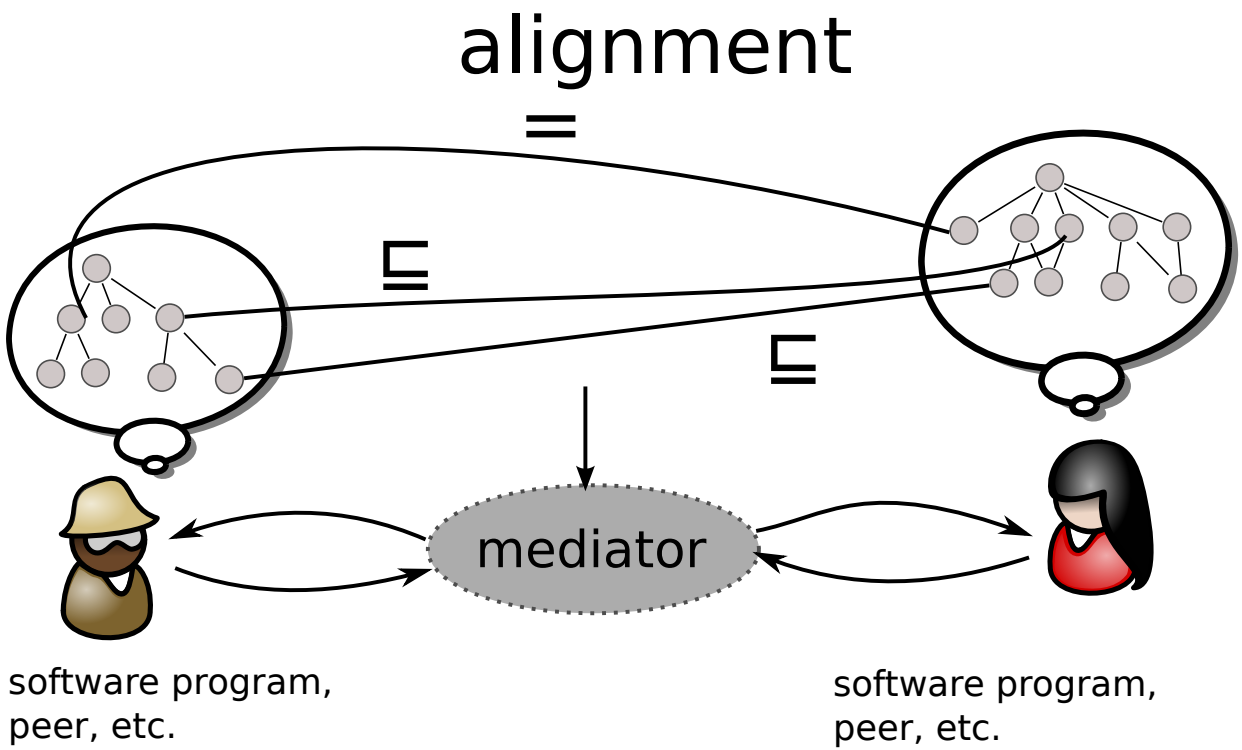


Context

A distributed and heterogenous environment: the semantic Web

- Several ontologies on the same domain

Problem: how to facilitate the communication between programs using different ontologies ?



Context

How to deal with many ontologies at Web scale ?
A solution: Distances between ontologies

- In peer-to-peer systems: find a peer using a close ontology
- In ontology engineering: find related ontologies for facilitating ontologies reuse
- etc.

Objectives:

- Review some distances between ontologies
  - distances issued from ontology matching
- Make a preliminary evaluation of these measures
  - accuracy vs. speed

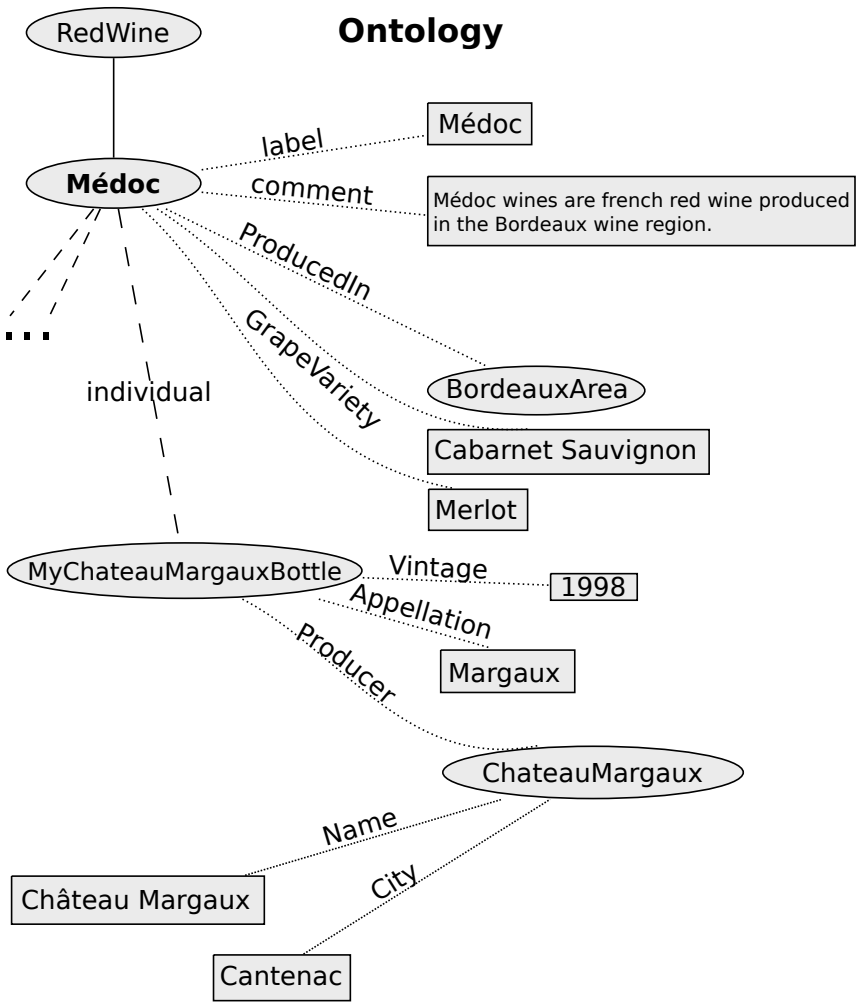
Various kinds of ontology distances

Two kinds of ontology distances evaluated:

- Vectorial distances: ontologies are viewed as bags of terms
  1. which weight scheme to use ?
  2. which vector distance to use ?

- Entity-based distances: the distance between ontologies is function of the distance between their entities
  1. entity distances: structural or/and lexical ?
  2. aggregating entity distance values: which collection distance to use ?

Vectorial ontology distances



Ontology vector

wine	3
red	2
médoc	2
french	1
bordeaux	1
cabarnet	1
sauvignon	1
merlot	...
margaux	
chateau	
cantenac	
1998	
grape	
appelation	
producer	
...	

Possible weights

Boolean

Frequency

TF.IDF

Vector measures

Cosine
Jaccard (Tanimoto index)

Entity distances

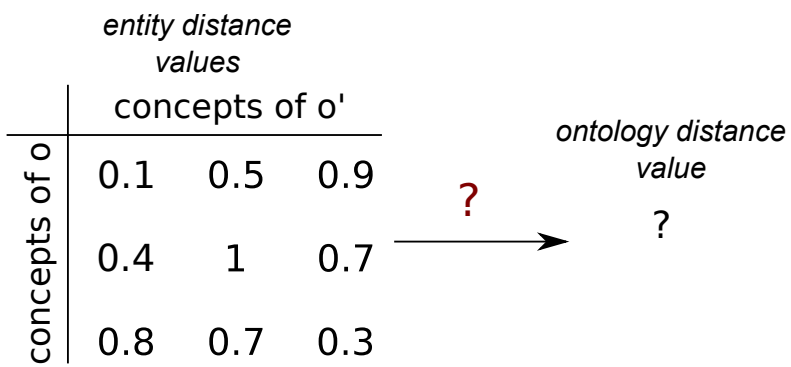
Label-based distance:

- String-distance between entity annotations : Jaro-Winkler
- Minimum Weight pairing + arithmetic mean

Structural distances:

- OLA similarity
  - Structural iterative similarity for OWL ontologies
- Triple-based iterative similarity
  - Structural iterative similarity for RDF graphs

Problem: How to aggregate entity measures values ?



Collection measures

Goal: compute ontology measure from concept measure values.
Collection measures used

- Average linkage : the arithmetic mean of concept measure values.

	concepts of o'		
concepts of o	0.1	0.5	0.9
	0.4	1	0.7
	0.8	0.7	0.3

mean = 0.5

Collection measures

Goal: compute ontology measure from concept measure values.
Collection measures used

- Average linkage : the arithmetic mean of concept measure values.
- Hausdorff :
$\Delta_{Hausdorff}(o, o') = \max(\max_{e \in o} \min_{e' \in o'} \delta_K(e, e'), \max_{e' \in o'} \min_{e \in o} \delta_K(e, e'))$

concepts of o \ concepts of o'			
	0.1	0.5	0.9
	0.4	1	0.7
	0.8	0.7	0.3

Collection measures

Goal: compute ontology measure from concept measure values.
Collection measures used

- Average linkage : the arithmetic mean of concept measure values.
- Hausdorff :
$\Delta_{Hausdorff}(o, o') = \max(\max_{e \in o} \min_{e' \in o'} \delta_K(e, e'), \max_{e' \in o'} \min_{e \in o} \delta_K(e, e'))$

concepts of o \ concepts of o'			
	0.1	0.5	0.9
	0.4	1	0.7
	0.8	0.7	0.3

Collection measures

Goal: compute ontology measure from concept measure values.
Collection measures used

- Average linkage : the arithmetic mean of concept measure values.
- Hausdorff :
$\Delta_{Hausdorff}(o, o') = \max(\max_{e \in o} \min_{e' \in o'} \delta_K(e, e'), \max_{e' \in o'} \min_{e \in o} \delta_K(e, e'))$

concepts of o \ concepts of o'			
	0.1	0.5	0.9
	0.4	1	0.7
	0.8	0.7	0.3

Collection measures

Goal: compute ontology measure from concept measure values.
Collection measures used

- Average linkage : the arithmetic mean of concept measure values.
- Hausdorff :
  $\Delta_{Hausdorff}(o, o') = \max(\max_{e \in o} \min_{e' \in o'} \delta_K(e, e'), \max_{e' \in o'} \min_{e \in o} \delta_K(e, e'))$
- Minimum Weight Maximum Graph Matching distance:

	concepts of o'		
concepts of o	0.1	0.5	0.9
	0.4	1	0.7
	0.8	0.7	0.3

mean = 0.4

Selected Measures

12 measures evaluated:

3 vectorial measures:

- Jaccard: the most basic measure (proportion of common terms)
- Cosine + TF weights
- Cosine + TF·IDF weights

3 entity measures:

- EntityLexicalMeasure
- TripleBasedEntitySim
- OLAEntitySim

→ 9 entity based measures

combined with 3 collections measures:

- AverageLinkage
- Hausdorff
- MWMGM

Benchmark suite

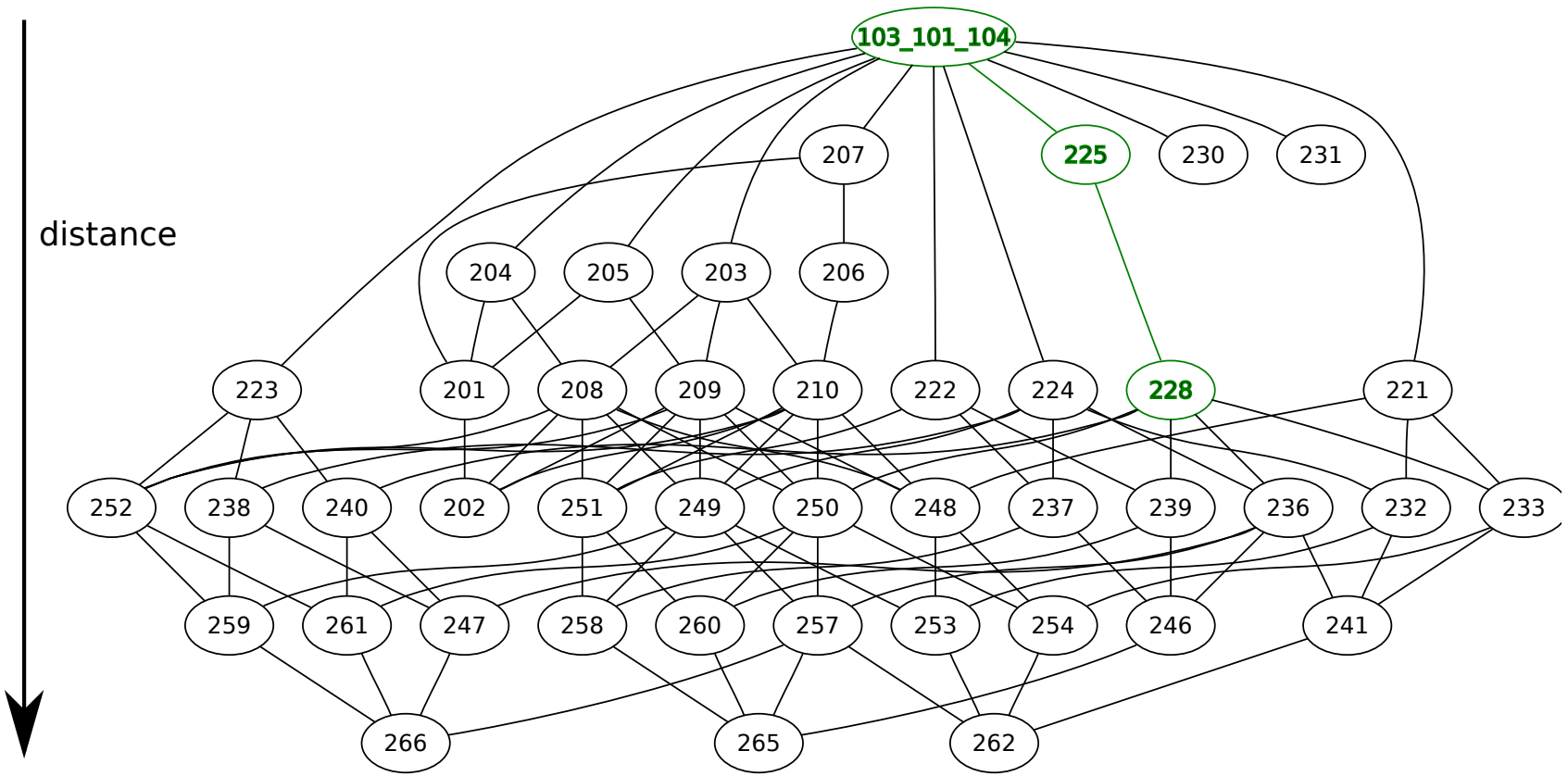
OAEI benchmark test set:

- 1 reference ontology 101 and altered ontologies
- 6 kinds of alterations:
  - **Name**: suppressed, randomized, synonyms, convention, other language
  - **Comment**: suppressed, other language
  - **Specialization hierarchy**: suppressed, expanded, flattened
  - **Instances**: suppressed
  - **Properties**: suppressed, discarded restrictions
  - **Classes**: split, flattened
- Order between alterations
  - Example on name:
    $\{$suppressed, randomized$\} < \{$ synonym, convention, other language $\}$
  - ontology with synonym names is closer to the reference one than ontology with no names

Benchmark suite

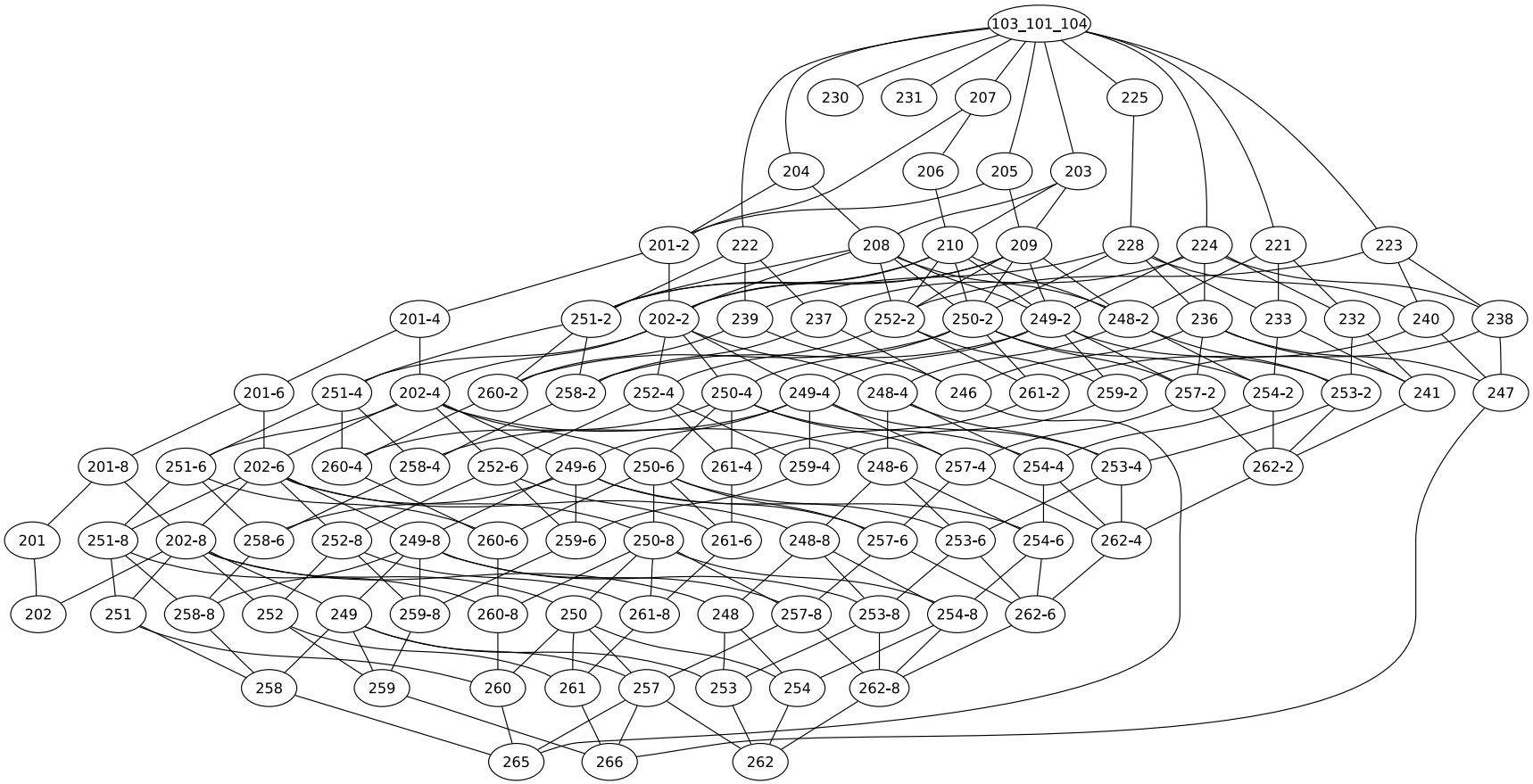
Induced proximity order:

101 is closer to 225 than 228



Benchmark suite

Induced proximity order (with partial alterations):



Tests

We performed **3 different tests**:

- Order comparison on benchmark test set
  - check if the similarity orders are compatible with the order induced by alterations
  - For each o, o' check if $101 < o < o' \Rightarrow \delta(101, o) < \delta(101, o')$
  - Test set is biased towards 1-1 matching: some measure favored

Tests

We performed **3 different tests**:

- Order comparison on benchmark test set
  - check if the similarity orders are compatible with the order induced by alterations
  - For each o, o' check if $101 < o < o' \Rightarrow \delta(101, o) < \delta(101, o')$
  - Test set is biased towards 1-1 matching: some measure favored

- Different cardinality matching
  - compare with related but different ontologies
  - compare with unrelated ontologies

Tests

We performed **3 different tests**:

- Order comparison on benchmark test set
  - check if the similarity orders are compatible with the order induced by alterations
  - For each o, o' check if $101 < o < o' \Rightarrow \delta(101, o) < \delta(101, o')$
  - Test set is biased towards 1-1 matching: some measure favored
- Different cardinality matching
  - compare with related but different ontologies
  - compare with unrelated ontologies
- Time consumption test
  - compare the time consumption of evaluated measures

Order comparison

Proportion of o, o' satisfying the induced order:

Measure	Tests Passed (ratio)	
	Original	Enhanced
MWMGM (EntityLexicalMeasure)	0.53	0.72
Hausdorff (EntityLexicalMeasure)	NaN	NaN
AverageLinkage (EntityLexicalMeasure)	0.44	0.31
MWMGM (OLAEntitySim)	0.75	0.78
Hausdorff (OLAEntitySim)	0.75	0.65
AverageLinkage (OLAEntitySim)	0.79	0.74
MWMGM (TripleBasedEntitySim)	0.86	0.92
Hausdorff (TripleBasedEntitySim)	0.86	0.89
AverageLinkage (TripleBasedEntitySim)	0.82	0.91
CosineVM (TF)	0.82	0.92
CosineVM (TFIDF)	0.57	0.81
JaccardVM (TF)	0.71	0.87

- best entity measure : TripleBasedEntitySim

Order comparison

Proportion of o, o' satisfying the induced order:

Measure	Tests Passed (ratio)	
	Original	Enhanced
MWMGM (EntityLexicalMeasure)	0.53	0.72
Hausdorff (EntityLexicalMeasure)	NaN	NaN
AverageLinkage (EntityLexicalMeasure)	0.44	0.31
MWMGM (OLAEntitySim)	0.75	0.78
Hausdorff (OLAEntitySim)	0.75	0.65
AverageLinkage (OLAEntitySim)	0.79	0.74
MWMGM (TripleBasedEntitySim)	0.86	0.92
Hausdorff (TripleBasedEntitySim)	0.86	0.89
AverageLinkage (TripleBasedEntitySim)	0.82	0.91
CosineVM (TF)	0.82	0.92
CosineVM (TFIDF)	0.57	0.81
JaccardVM (TF)	0.71	0.87

- best entity measure : TripleBasedEntitySim
- best collection measure : MWMGM

Order comparison

Proportion of o, o' satisfying the induced order:

Measure	Tests Passed (ratio)	
	Original	Enhanced
MWMGM (EntityLexicalMeasure)	0.53	0.72
Hausdorff (EntityLexicalMeasure)	NaN	NaN
AverageLinkage (EntityLexicalMeasure)	0.44	0.31
MWMGM (OLAEntitySim)	0.75	0.78
Hausdorff (OLAEntitySim)	0.75	0.65
AverageLinkage (OLAEntitySim)	0.79	0.74
MWMGM (TripleBasedEntitySim)	0.86	0.92
Hausdorff (TripleBasedEntitySim)	0.86	0.89
AverageLinkage (TripleBasedEntitySim)	0.82	0.91
CosineVM (TF)	0.82	0.92
CosineVM (TFIDF)	0.57	0.81
JaccardVM (TF)	0.71	0.87

- best entity measure : TripleBasedEntitySim
- best collection measure : MWMGM
- best vectorial measure : Cosine with TF weights

Order comparison

Proportion of o, o' satisfying the induced order:

Measure	Tests Passed (ratio)	
	Original	Enhanced
MWMGM (EntityLexicalMeasure)	0.53	0.72
Hausdorff (EntityLexicalMeasure)	NaN	NaN
AverageLinkage (EntityLexicalMeasure)	0.44	0.31
MWMGM (OLAEntitySim)	0.75	0.78
Hausdorff (OLAEntitySim)	0.75	0.65
AverageLinkage (OLAEntitySim)	0.79	0.74
MWMGM (TripleBasedEntitySim)	0.86	0.92
Hausdorff (TripleBasedEntitySim)	0.86	0.89
AverageLinkage (TripleBasedEntitySim)	0.82	0.91
CosineVM (TF)	0.82	0.92
CosineVM (TFIDF)	0.57	0.81
JaccardVM (TF)	0.71	0.87

- best entity measure : TripleBasedEntitySim
- best collection measure : MWMGM
- best vectorial measure : Cosine with TF weights

- Best results: TripleBasedEntitySim with MWMGM
- Cosine with TF weights: basic measure but works well
- Hausdorff works only with TripleBasedEntitySim

Is MWMGM favored by the 1-1 nature of this test ?

Different cardinality matching

Objective: Ensure that MWMGM is still the best measure when we use different cardinality ontologies

Test set:

- A reference ontology: 101
- Highly altered ontologies: 2xx
- Different but similar ontologies: 3xx
- Irrelevant ontologies: conference test set (could share some vocabulary with benchmark ontologies)

Expected results:

$$\delta(101, 2xx) < \delta(101, 3xx) < \delta(101, \mathit{conference})$$

Different cardinality matching

Results: MWMGM still performs best, vectorial measures fail when ontologies are lexically altered

Main misorderings:

- MWMGM
  - only one altered (250) with similar (3*xx*) and conference ontologies
  - some similar (3*xx*) with conference ontologies
- Hausdorff :
  - with TripleBasedEntitySim: only 250 with similar (3*xx*) and a conference ontology
  - with OLAEntitySim: 228, 248, 250, 251, 252 with 304 and conference ontologies
  - with LexicalEntitySim: does not work
- AverageLinkage
  - highly altered (2*xx*) and similar (3*xx*) ontologies
- Vectorial measures
  - many altered (248, 250, 251, 252) ontologies with similar (3*xx*) ontologies

Time consumption

Are measures useable in real-time applications ?

- Vectorial measures are more time efficient

Time consumption on benchmark test:

Measure	Total time (s)	Average time per similarity value (s)
MWMGM(EntityLexicalMeasure)	558	0.46
MWMGM (OLAEntitySim)	39 074	31.9
MWMGM (TripleBasedEntitySim)	7 950	6.49
Hausdorff (EntityLexicalMeasure)	451	0.37
Hausdorff (OLAEntitySim)	38 912	31.76
Hausdorff (TripleBasedEntitySim)	7 410	6.05
AverageLinkage (EntityLexicalMeasure)	444	0.36
AverageLinkage (OLAEntitySim)	38 995	31.83
AverageLinkage (TripleBasedEntitySim)	7 671	6.26
CosineVM (TF)	101	0.08
CosineVM (TFIDF)	102	0.08
JaccardVM (TF)	101	0.08

Time consumption

Are measures useable in real-time applications ?

- Vectorial measures are more time efficient
- Not significant differences between MWMGM (N^3), Hausdorff and AverageLinkage(N^2)

Time consumption on benchmark test:

Measure	Total time (s)	Average time per similarity value (s)
MWMGM(EntityLexicalMeasure)	558	0.46
MWMGM (OLAEntitySim)	39 074	31.9
MWMGM (TripleBasedEntitySim)	7 950	6.49
Hausdorff (EntityLexicalMeasure)	451	0.37
Hausdorff (OLAEntitySim)	38 912	31.76
Hausdorff (TripleBasedEntitySim)	7 410	6.05
AverageLinkage (EntityLexicalMeasure)	444	0.36
AverageLinkage (OLAEntitySim)	38 995	31.83
AverageLinkage (TripleBasedEntitySim)	7 671	6.26
CosineVM (TF)	101	0.08
CosineVM (TFIDF)	102	0.08
JaccardVM (TF)	101	0.08

Time consumption

Are measures useable in real-time applications ?

- Vectorial measures are more time efficient
- Not significant differences between MWMGM (N^3), Hausdorff and AverageLinkage(N^2)
- Structural measures are not really useable in real-time applications

Time consumption on benchmark test:

Measure	Total time (s)	Average time per similarity value (s)
MWMGM(EntityLexicalMeasure)	558	0.46
MWMGM (OLAEntitySim)	39 074	31.9
MWMGM (TripleBasedEntitySim)	7 950	6.49
Hausdorff (EntityLexicalMeasure)	451	0.37
Hausdorff (OLAEntitySim)	38 912	31.76
Hausdorff (TripleBasedEntitySim)	7 410	6.05
AverageLinkage (EntityLexicalMeasure)	444	0.36
AverageLinkage (OLAEntitySim)	38 995	31.83
AverageLinkage (TripleBasedEntitySim)	7 671	6.26
CosineVM (TF)	101	0.08
CosineVM (TFIDF)	102	0.08
JaccardVM (TF)	101	0.08

Conclusion

- Why use complex measures when basic measures performs well ?
  - for discriminating highly related ontologies
- More precise test set is needed
- Characterize confidence intervals
  - Is there a threshold from which returned values are always correct ?
- Other ontology distances measures: alignment based