Incoherence as a Basis for Measuring the Quality of Ontology Mappings

Christian Meilicke & Heiner Stuckenschmidt
University of Mannheim, KR & KM Group
{christian,heiner}@informatik.uni-mannheim.de
Motivation

- Measuring the quality of an automatically generated alignment $M$ is in most cases based on a comparison with a reference alignment (gold standard)
  - To compute e.g. precision and recall

- PROBLEMS:
  1. Even though an alignment has acceptable precision and recall, internal logical problems might hinder a sensible use
  2. Reference alignment are often not available
     That's why we need matching systems!

- IDEA: Measure logical aspects (incoherence) as a
  - complement to classical evaluation strategies and as
  - alternative to classical measures in absence of a reference alignment
Measuring Incoherence

Outline

• Definition: Incoherence of an alignment

An objection and a problem

• The objection: only useful in specific application scenario
• The problem: from \{true, false\} to \([0,1]\)
  – Impact based measures
  – Measures based on revision effort

Implications

• Truth and Coherence: A simple proposition
  – How to make use of this proposition
• Future Work
**Definition (Merged Ontology).** The merged ontology of $O_1$ and $O_2$ connected via $M$ referred to as $O_1 \cup_{Mt} O_2$ is defined as

$$O_1 \cup_{Mt} O_2 = O_1 \cup O_2 \cup \{ \text{t}(c) \mid c \in M \}$$

where $t$ is a translation function that maps correspondences to axioms.

**Definition (Natural DL-Translation).** The natural translation $t_n$ is defined as a function that maps a correspondence to the accordant DL axiom. E.g. $t_n(<1\#e, 2\#e', \sqsubseteq, 0.788>) = 1\#e \sqsubseteq 2\#e'$

**Remark:** Choice of the translation function leaves some room for different semantics.
Definition: Incoherence

• Similar to the incoherence of an ontology, incoherence of an alignment can be defined as follows:

Definition (Incoherency of an alignment). An alignment $M$ between $O_1$ and $O_2$ is incoherent due to translation function $t$ iff there exists a concept $i\#C$ with $i \in \{1,2\}$ such that:

1. $i\#C$ is satisfiable in $O_i$ and
2. $i\#C$ is unsatisfiable in $O_1 \cup_M O_2$

Short reminder: An ontology is inconsistent if there exists no model.

An ontology is incoherent iff there exists an unsatisfiable concept.
Objection: It's only about Merging

- Definition is based on merging two ontologies, but there are many different application scenarios
  - Query answering/rewriting
  - Instance migration
  - ...  

- None of these application scenarios require merging of ontologies!

That's true, but incoherences will nevertheless often result in problems in these scenarios!
Counterexample: Instance migration

\( O_1 \)

\[ \text{Animal} \]
\[ \text{Insect} \]
\[ \text{RedWoodAnt} \]

\( O_2 \)

\[ \text{Plant} \]
\[ \text{WoodPlant} \]

(1) \( O_1 \# \text{RedWoodAnt} \subseteq O_2 \# \text{WoodPlant} \)

(2) \( O_1 \# \text{Animal} = O_2 \# \text{Animal} \)

\( O_2 \) is **inconsistent** after instance migration!
Problem: \( \{0, 1\} \rightarrow [0, 1] \) ?

Measuring Incoherence
Impact based measures
(derived from the field of ontology debugging)

• **Unsatisfiability Measure.** Count the number of unsatisfiable concepts in $O_1 \cup_M O_2$ that have not been unsatisfiable in $O_1$ resp. $O_2$

• Concepts becoming unsatisfiable are understood as negative *impact* of the alignment

\[
m^{t \text{sat}}(O_1, O_2, M) = \frac{| \text{Unsatisfiable concepts in } O_1 \cup_M O_2 \text{ satisfiable in } O_1 \text{ resp. } O_2 |}{| \text{Concepts satisfiable in } O_1 \text{ and } O_2 |}
\]

• Problem: A merged unsatisfiable concept will make all its subconcepts unsatisfiable.
  - We might only be interested in counting the root unsatisfiable concepts (see paper for *Root Unsatisfiability Measure*)
Measures based on revision effort

(based on our previous work)

- **Maximum Cardinality Measure.** Count the minimum number of correspondences that have to be removed to arrive at a coherent subset

  - The number of correspondences which have to be removed is understood as the *effort of revising* the alignment

  \[
  m_{\text{card}}(O_1, O_2, M) = \frac{|M - M'|}{|M|}
  \]

  where \(M' \subseteq M\) is a coherent alignment and there exists no \(M'' \subseteq M\) with \(|M''| > |M'|\) such that \(M''\) is coherent.

- **Variant of this measure is the Maximum Trust Measure**
  - Revision effort measured with respect to total of confidence values of removed correspondences (see paper)
Complexity Considerations

- **Unsatisfiability Measure**
  - Classify the merged ontology and count unsatisfiable concepts

- **Maximum Cardinality Measure**
  - Requires lots of reasoning in the merged ontology
  - Requires to solve the hitting set problem \((\text{NP–complete!})\)
  - First implementation works for alignments between ontologies up to several hundred concepts
  - Will not be directly applicable for large matching problems, but approximation is possible
Proposition (Upper bound for precision). Let $M$ be an alignment and let $R$ be a reference alignment between $O_1$ and $O_2$. Further let $R$ be coherent due to translation function $t$. Then we have

$$\text{precision}(M, R) \leq 1 - m^t_{\text{card}}(O_1, O_2, M).$$

$M^* = M \cap R$ (by definition)

Definition of precision

$M^*$ is a coherent subset of $M$ and $M'$ is the largest coherent subset of $M$

Note: There is a small error in the equation presented in the paper, don't be confused.
How to use this proposition?

- Example 1: Several matchers have been applied on the same problem
  - Each matcher generated an alignment. Which one should we choose?
  - Upper-Bound Proposition cannot be used to decide this question!
  - **BUT**: It might help us to decide *which one we should not choose!*

- Example 2: A matcher is applied to a matching problem of a new/unknown domain (experience missing), that requires a precision of e.g. at least 0.9
  - Which threshold should be used?
  - Compute upper bound for precision stepwise increasing threshold, provides useful information about threshold
Future Work

- Experiments
  - How useful is the upper bound of precision?
  - Different coherence characteristic for different matching systems?
  - ...
- Is there a interdependence between coherence and recall?
- Support different „distributed semantics“ (=different translation functions), for example DDL
  - In principle possible as long as chosen semantics provides a translation into DL
- Support matching datatypedproperties on objectproperties
  - Natural translation does not support this, we already implemented a weaker translation
Thanks for your attention, questions?