UNIVERSITÄT LEIPZIG

wdilab

ATOM

Salvatore Raunich, Erhard Rahm

Automatic Target-driven Ontology Merging

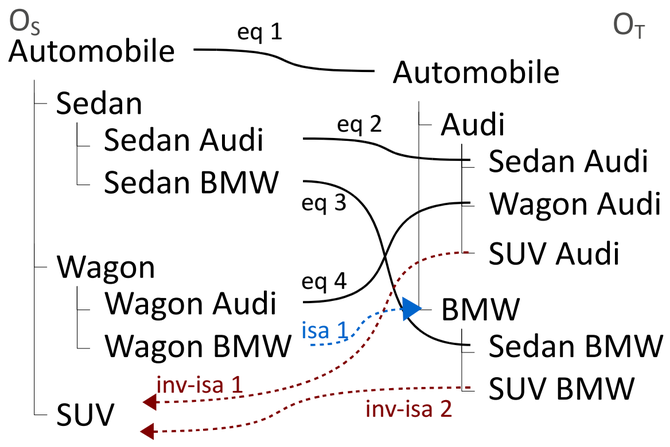
Problem: Reducing the Semantic Overlap and Preserving the Target Ontology

- Common merge algorithms keep both input ontologies in the merge result, merging equivalent concepts into a common concept.
- Preserving both input views in the merge result lead to a semantic overlap (redundancy) and reduced understandability of the merge result.
- Multiple inheritance can be introduced in the result if multiple paths to the merged concepts are generated.
- The ATOM approach gives preference to the target ontology (2nd input):
  only the target concepts and structure are preserved while source concepts and relationships are dropped if they would introduce redundancy in the merge result.

The Automobile Example

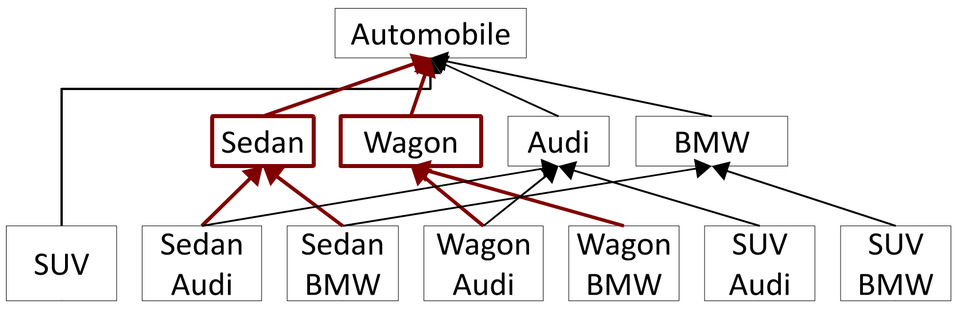
Example Input:

- Taxonomies O_S and O_T which classify vehicles in different ways.
- A match mapping, expressed as a set of correspondences between O_S and O_T .



Full Merge Solution

- Preserves all information from the input ontologies, in particular all concepts and relationships.
- Introduces multiple paths to merged concepts.



ATOM Solution with only Equivalence Mapping

- Preserves the target ontology.
- Reduces part of the semantic overlap in full merge solution.
- No multiple inheritance.

Automobile*
- Audi
  - Sedan Audi*
  - Wagon Audi*
  - SUV Audi
- BMW
  - Sedan BMW*
  - SUV BMW
- Wagon
  - Wagon BMW
- SUV

ATOM Solution with Extended Mapping

- Preserves the target ontology.
- Reduces the semantic overlap utilizing an enhanced input mapping.

Automobile*
- Audi
  - Sedan Audi*
  - Wagon Audi*
  - SUV Audi + S1
- BMW
  - Sedan BMW*
  - SUV BMW + S2
  - Wagon BMW
- SUV (others)

ATOM Features

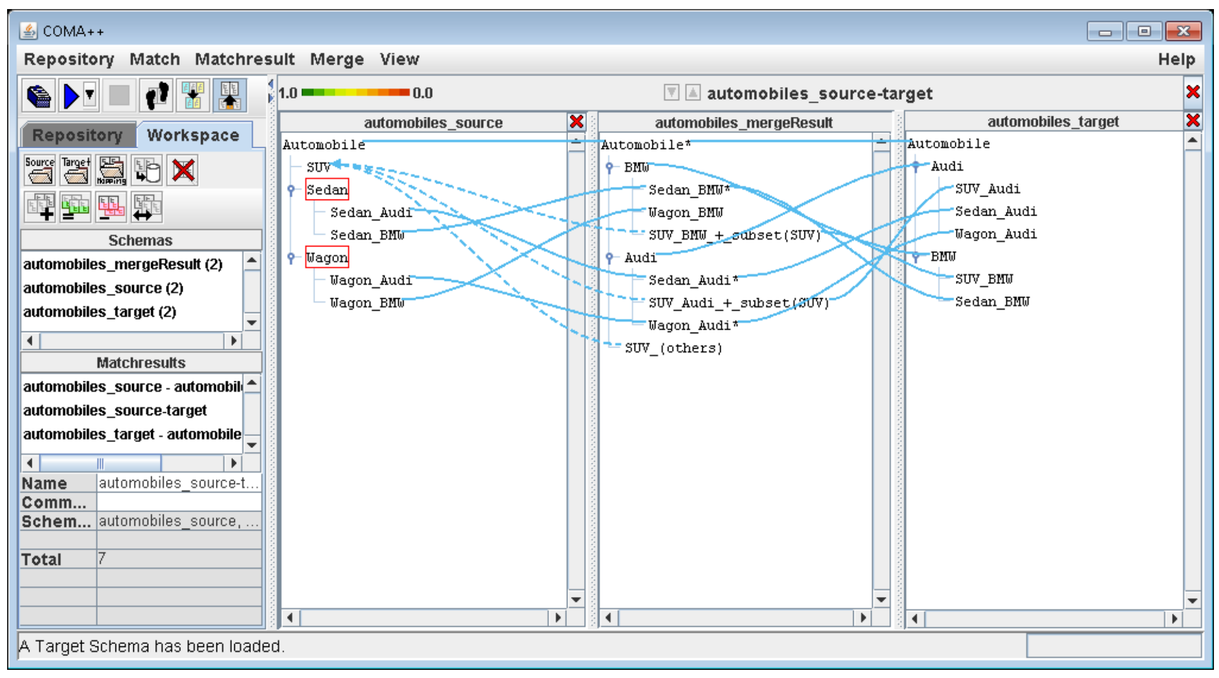
- New target-driven algorithm. For example, the catalog of a new online shop may be merged into the catalog of a price comparison portal.
- Reduction of the semantic overlap in the merge result for improved understandability.
- Use of match (equivalence) mapping between input ontologies.
- Optional utilization of enhanced input mappings containing equivalence, is-a and inverse is-a relationships between concepts of the input ontologies.
- High effectiveness in experimental results: ATOM has very good performance also with real and large taxonomies.
- Automatic generation of equivalence mappings between the input taxonomies and the merged taxonomy., e.g. to support instance migration.
- User interaction to influence or refine the merge result.

Experimental Results

Merge Example		Anatomy		eBay Catalog	
		Mouse	NCI	v94	v93
Input size	Concepts	2,700	3,300	21,000	23,500
	Leaf paths	2,300	2,600	18,400	19,700
# correspondences		~1,000		~20,200	
FULL solution	Concepts	5,100		23,400	
	Leaf paths	12,900		21,600	
ATOM solution	Concepts	5,000		23,300	
	Leaf paths	6,900		20,400	
Execution time		1 second		7 seconds	

ATOM Integration in COMA++

- COMA++ supports semi-automatic generation of input mappings [1].
- Equivalence, is-a and inverse-isa correspondences are supported by the system.
- Both the ATOM and a full merge solution can be produced and inspected by the user.



[1] D. Aumüller, H. Do, S. Massmann, and E. Rahm. Schema and Ontology Matching with COMA++. ACM SIGMOD 2005.

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

UNTERNEHMEN REGION
Die BMBF-Innovationsinitiative Neue Länder

More info:
http://wdilab.uni-leipzig.de

Gedruckt im Universitätsrechenzentrum Leipzig