Ontology Mapping Tools, Methods and Approaches – Analytical Survey

Tatyana Ivanova[1] and Valentina Terzieva[2]

[1] College of Energetics and Electronics, TU-Sofia,
8 Kl. Ohridski Blvd., Sofia 1000, Bulgaria
[2] Institute of Computer and Communication Systems,
Acad. G. Bonchev Str. Bl. 2, 1113 Sofia, Bulgaria
[1] tiv72@abv.bg, [2] valia@isdip.bas.bg

Abstract. In this paper we present the results of browsing, analyzing and comparing many ontology mapping tools, approaches and methods. We extract and classify valuable parameters for strict and unambiguous tool or method description. Every mapping tool, algorithm or approach must have such a description, practically usable for both human and software agents and sufficient for easy checking if it suitable or not for a given task. We will use our classifications for developing ontology, conceptualizing all valuable metadata for semantic machine-processable mapping tools description.

Keywords: ontology mapping tools, ontology mapping approaches, ontology alignment evaluation, ontology alignment

1 Introduction

Almost any application that uses multiple ontologies must establish semantic mappings among them. Mapping is also needed to ensure interoperability among different applications in many domains, including e-commerce, knowledge management, e-learning, information extraction, bioinformatics, (semantic) web services, multiagent systems etc. A grand number of ontology mapping methods and tools have been developed so far, including early mapping investigations [1], [2], [3], or modern ones, such as [4] and [5]. As each one of these tools is appropriate for some tacks or domains and give poor results (or may not work at all) for others, to choose the right mapping tool one have to explore thousands of pages, containing tool descriptions and test in his own risk some of tools. The aim of this paper is to analyze and classify ontology mapping tools and methods and extract its valuable properties, needed for strict and unambiguous machine-processable description, usable in automatic tool selection or for making dynamic goal – directed recommendations.

2 Research Methodology

We use many sources of information about mapping tools and methods: mapping projects Web sites, scientific papers from digital libraries, scientist opinions (from blogs, forums and other web 2.0 applications), Ontology Alignment Evaluation Initiative (OAEI) materials [5], tool's documentation. We obtain these resources using keyword based web search (google, yahoo, bing), folowing citations in scientific papers or surveys, browsing project sites or downloading tools extracting and reading tools documentation. We browse, analyze and make a brief comparision of at about 90 ontology mapping tools. Our aim is to extract basic classification dimensions, make structured classification of terminology, used in ontology mapping domain and collect metadata for semantic description of ontology mapping tools, methods, approaches.

3 Ontology Alignment Tool Characteristics

There are many independent dimensions along which approaches, algorithms or tools can be examined, classified or selected. After analyzing several surveys [1], [2], [3], 4], tool documentations and many other represented in the last few years related to mapping papers we propose the following main dimensions for comprehensive description and classification of ontology matching tools (table 1):

Table 1. A part of the classification of the important ontology mapping tool characteristics.

Tool dimension	sub dimension	second sub dimensions	S/P
Input	Size	Number of ontologies	P
		Ontology size	S
	category	Dbschema, ontology, thesaurus	S
	formality level	Informal, Semi-formal, formal	P
	Input natural languages	One language, multilanguage	P
	Input representation language	One, several	P
	Input ontology type	task , domain, upper, application	
Output	Output type	For software, For human	S
	Matching Cardinality	Global, local	P
	Execution Completeness	Subjective, injective, partial, full	
Usage	Application type Area	Application domain	P
		Knowledge management	S
	Application place of usage	Local, network, web	P
	Application domain Area	one	P
		Multi domains	S
	User type	Human, software	S
Adaptation ability	Domains, applications	(list) P, or classification (S)	
	Tasks, usage	(list) P, or classification (S)	
Evaluation features	benchmark	P	
	Tested parameters	S	

The tool's input characteristics: size; The characteristics of the matching process, which describes the matching approaches, methods and algorithms themselves; The output of the tool (output type, matching cardinality, execution completeness[6]); The usage characteristics of the matching tool (different situations where the tool have been used: for various approaches, application areas, etc.; Matching strategy; Matching quality; Tool code characteristics; Tool vendor and support characteristics; Documentation characteristics, and cost characteristics. Some of the upper levels of tool characteristics classification is shown in table 1. In this table we don't include characteristics, related to the used algorithms, as we will discuss them latter. The valuable subdimensions are listed in, column 2 from the left, and second level subdimensions - in column 3. “P” stands for property, and “S” – for subdimension.

4 Ontology Alignment Approaches Characteristics

We present the simple mapping approaches terminology classification and the methods combination related terminology. We also analyze and classify mapping evaluation terminology, but don’t present this classification here because of the restricted space. We classify ontology mapping approaches according to subsequent dimensions (table 2): Automation level, Type of mapped ontology elements, Kind of mapped relations, Mapping cardinality, Used External information source types, Mapping metric types, Mapping aim, Tack-dependency, Domain dependency, mapped Ontology types, used during mapping Relation types, Ontology size, expressiveness, Terminology language, globality, general matching directions, strategies, Mapping discovery base, Mapping representation (output), quality (left column of table 2). The valuable subdimensions are shown in column 1 from the left. Because of the restricted space, in column 2 only little part of terminology, related to every subdimension, is shown, and in the right column some algorithms or tools, classified to corresponding dimensions are listed.

5 Discussion, Conclusions and Future Work

Before selecting the best approach, method or tool for concrete application the comprehensive exploration of grand number of variants is needed. Manual exploration is difficult, time consuming tack and is not suitable for ordinary users as well as not applicable in the cases of automatic Multistrategy or Multiagent mapping in dynamic environment. The exploration of several hundred of textual pages, describing the last research in this area would cost months working of professionals, and all this information is not processable for software agents. So, the clear, short, structured and machine-processable explicit description of valuable characteristics of the ontology mapping tools, algorithms and approaches is needed. The main requirements to this description are: Easy usage by people (domain experts or usual users); Easy readable and processable from software agents and web services; Comprehensive, explicit and clear description of all the characteristics, valuable for choosing in every possible practical situation. Computer programs should make context-aware recommendations to the users in choosing the right tools, or make automatic dynamic choosing of needed mapping services.

Currently, we are working on development of ontology, conceptualizing all listed in the tables dimensions, listed or omitted because of the restricted space subdimensions, it properties and inter relations. The tools, methods and algorithms are individuals in this ontology. Our aim is to develop well structured and comprehensive terminology classification. In this ontology we also will include contextual information about every concept (mapping tool or algorithm describing dimension): synonyms, abbreviations, related words and corresponding relations. Terminology richness and completeness will guarantee successful usage of this conceptualization for tools or methods comparison. Our classification is based to the exploration and analysis of all the ontology mapping

Table 2. A part of the classification of the important ontology mapping algorithm characteristics.

Approach (dimension)	Approach name	Approach description (synonyms, related terms)	Algorithms /tools
Type of mapped ontology elements	Instance –based (Extensional)	Bayes Classifier, relaxation labeling, Jaccard coefficient	iMapper, SAMBO, ASMOV, OKKAM
	Schema – based	QGram, graph-matching, SVMs	ASMOV, CIDER
Kind of mapped relation	syntactic	equivalence relation with certain level of plausibility or confidence	Edit distance N-gram, Corpus
	structural	analyzing how entities are related together, graph-matching	H-match, Anchor Flood, OLA
	semantic	equivalence (=); more general (⊃) specification (⊂); mismatch (!);(∩)	Gloss-based Anchor Flood, Lily
Mapping *cardinality*	One to one	Element-level	Many
	One to many	Element-level	MAFRA, SKAT
	Many to many	structure-level, semantic bridging ontology	structural semantics MAFRA
External information source	no	Quik ontology mapping, information retrieval	Anchoor-flood, Prior+, X-SOM
	thesaurus	Linguistic, TF/IDF, QGram	H-Match,ASCO,ASMOV
	ontology	Domain ontology	KitAMO,API
	IR	TF/IDF, feature vector	PRIOR
	learning	Mashine-learning	GLUE
	Patterns	Hearst pattern based	PANKOW
	user	interactive	AnchorPrompt
Mapping metric type	fuzzy	Fuzzy Conceptual Graphs, Jaccard's coefficient	Monge-Elkan, SLIM (algorithms)
	deterministic	weighted average of measurements	ASMOV, MapPSO
	probabilistic	Bayesian networks (BN) Dempster-Shafer theory	GLUE, OMEN, Onto-Mapper, DSsim
Mapping aim	All possible	Mapping terminology systems	many
	Minimal [12]	Use debugging heuristics	MinSMatch
	one (The best)	ontological context	CIDER
Ontology type	global	Domain-independent	VerbNet-Cyc
	domain	e.g. using logical deduction	CTXMATCH
	Global -local	Multi-strategy learning	MOMIS, OIS
Ontology size	Large [10]	divide-and-conquer approach Large scale Ontology Matching	Anchor-Flood, Gmo, Lily, MapPSO
	small	Generic Ontology Matching	Lily
	mixed	Combination of matchers	Anchor-Flood
expressiveness	expressive	diagnostic reasoning	ASMOV, KOSIMap
	lightweight	supervised machine learning	myOntology [8]
Terminology language	Multilingual [7]	corpus-based, lexicon-based	SOCOM[11]
	One language	many	many
general matching directions	complete mapping	many	many
	partial	Partitioning algorithms	Bmo, V-Doc, Gmo
	oriented alignment	PAP (Partition, Anchor, Partition)	TaxoMap
	interactive	User-interface building	OMIE, PROMPT
	community-driven	public alignment reuse	myOntology [8]
	multi-ontology	Schutze's automatic WSD	FOAM, PROMPT
Type of used techniques	linguistic	relational learning algorithms, string similarity, SMOA	KOSIMap
	constraint-based	based on keys and relationships	Combined with others
	deduction	Extensional, DL reasoning	AROMA, MAFRA
	Rule-based	association rule	AROMA
Strategies	Static	Semantic, instance, schema-based	GLUE, AUTOMS
	dynamic	Alignment Strategy Recommendation Algorithm	KitAMO, Falcon-AO, APFEL
	Query-driven	Learning-based	FOAM, APFEL
Mapping quality [9]	approaches	Iterative pruning and validation, Debugging Mappings	ASMOV, KOSIMap
	metrics	Precision, recall, efficiency	many

algorithms and tools, tested in OAEI, many surveys and other publications, presenting tools or algorithms. Our ontology will be useful in various tools or algorithms exploration, comparison, recommendation and selection tacks, and will be easily extensible with characteristics of further newly developed ones. Further we will develop recommendation agent that will help in finding the best tool or approach for solving well described mapping problems by comparing the semantic descriptions of tools or methods, stored in our ontology with the application needs.

References

1. Kalfoglou, Y., Schorlemmer, M., : Ontology mapping: the state of the art. The Knowledge Engineering Review, 18(1):1–31, (2003)
2. Euzenat, J., :D2.2.3: State of the art on ontology alignment. http://starlab.vub.ac.be/research/projects/knowledgeweb/kweb-223.pdf (2004)
3. Choi, N., et al, :A Survey on Ontology Mapping. http://www.sigmod.org/sigmod/record/issues/0609/p34-article-song.pdf (2006)
4. Godugula, S.,: Survey of Ontology Mapping Techniques. (2008)
5. Ontology Alignment Evaluation Initiative site, http://oaei.ontologymatching.org/
6. Euzenat,J.,et al.,:D1.2.2.2.1:Case-based recommendation of matching tools and techniqu-es.http://knowledgeweb.semanticweb.org/semanticportal/deliverables/D1.2.2.2.1.pdf(2007)
7. Trojahn, C., Quaresma, P., Vieira R.: A Framework for Multilingual Ontology Mapping . In: Proceedings of LREC, 1034- 1037 (2008)
8. Siorpaes, K., Hepp, M.,: myOntology: The Marriage of Ontology Engineering and Collective Intelligence. In: Bridging the Gep between Semantic Web and Web 2.0 (2007)
9. Hollink, L.,et al.,: Two Variations on Ontology Alignment Evaluation: Methodological Issues. http://www.eswc2008.org/final-pdfs-for-web-site/oa-2.pdf (2008)
10. Hu, W., Qu ,Y., Cheng, G ., : Matching large ontologies: A divide-and-conquer approach, Data & Knowledge Engineering, (2008)
11. Fu, B., et al., : Cross-Lingual Ontology Mapping and Its Use on the Multilingual Semantic Web, WWW 2010, April 26-30, (2010)
12. Giunchiglia F., et al.,: Computing minimal mappings. In: Proc. of the 4th Ontology Matching Workshop. (2009)