C-SAW—contextual semantic alignment of ontologies: using negative semantic reinforcement

Brendan Tierney
Technological University Dublin, brendan.tierney@tudublin.ie

Mike Jackson
Birmingham City University

Recommended Citation

Tierney, B. & Jackson, M. (2008) C-SAW – Contextual Semantic Alignment of Ontologies – Using Negative Semantic Reinforcement, *2008 ACM symposium on Applied computing*,DOI:10.1145/1363686.1364242

C-SAW – Contextual Semantic Alignment of Ontologies – Using Negative Semantic Reinforcement

Brendan Tierney
School of Computing
Dublin Institute of Technology
Kevin St., Dublin 8, Ireland
brendan.tierney@dit.ie

Mike Jackson
Business School
Birmingham City University
Birmingham, England
mike.jackson@bcu.ac.uk

ABSTRACT

Understanding the meaning of each term in an ontology is essential for successfully integrating and aligning ontologies. Much ontology integration research to date is focused on syntactic, structural and semantic matching where the actual meaning of the concepts is disregarded. The C-SAW approach to ontology alignment is based on the Contextualizing the concepts by using a set of Semantic Alignment Words (C-SAW). The C-SAW approach is enhanced by Negative Semantic Reinforcement (NSR), where additional semantic meaning can be added to the set of Semantic Alignment Words, by considering words which are unrelated to the concept.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Content Analysis and Indexing, Information Search and Retrieval.

Keywords: Ontology alignment, integration, contextual meaning, similarity measures

1. INTRODUCTION

We focus on presenting the contextualisation of ontologies using a set of Semantic Alignment Words (C-SAW). The context of each concept in an ontology is provided by way of a set of associated words that, when combined, give a specific contextual meaning for the concept. To achieve this, some dictionary/thesaurus is required. We use WordNet for this. C-SAW is enriched with Negative Semantic Reinforcement (NSR) where we enrich the SAW with words that emphasise the semantics of the concept and helps to rule out alternative concepts where the name of the concept has many interpretations.

2. ONTOLOGIES & CONTEXT

Ontologies are shared models of a domain that encode a view which is common to a set of different parties [5]. Ontologies have been proposed as a possible solution for the knowledge sharing and reuse problem, by providing a formal mechanism for defining the semantics of data [4]. The most commonly discussed integration and alignment approaches for ontologies are linguistic, statistical, structural and semantic methods. In the view of the authors, a number of the methods associated with semantic integration and alignment are really structural [7], [6], [8]. There are only a small number of methods that deal with the actual semantics of the concepts in relation to their context [1], [2], [9]. Contexts are local models encoding a party's subjective view of a domain [3]. Ontologies are shared models of some domain encoding a view which is common to a set of different parties [3]. Contexts are local (where local is intended to imply not shared) models encoding a party's view of a domain [5]. A locally implemented ontology is contextualized, i.e. is based on the local environment. The semantics of this contextualized ontology is implemented by allowing the ontology engineer to use a common dictionary. In our case we use WordNet. This allows the ontology to use WordNet to identify the different senses, to define the meaning of the concepts. This is achieved specifying a set of words (SAW) that gives the meaning/semantics of the concept, based on using the WordNet noun hypernym senses.

3. RELATED WORK

There are a number of approaches which attempt to work with the meaning of the concepts. Whilst some of these are classified in the literature as semantic approaches, the authors believe that few really consider the meaning of the concepts in the ontology. One approach using WordNet to determine the similarity of concepts is by [1] where they scan the sysnet of WordNet for the concepts label to find similarity. They calculate a gloss-overlap score between the two concepts by looking up both concepts in WordNet's taxonomy and comparing the synsets. Their approach can involve a large search space involving a large amount of complex processing and matching, with a considerable amount of this being wasted processing. It can also generate a significant number of incorrect matches and does not take into account the actual meaning intended, only possible similarities. This approach is compared the word-pairs given by Wu & Palmer [9].

4. CONTEXTUAL SEMANTIC ALIGNMENT

The context for each concept in an ontology is added by way of a set of associated words that, when combined, gives a specific contextual meaning for the concept. This gives us our set of Semantic Alignment Words (SAW). In our approach we use WordNet as the dictionary, allowing the ontology to use the structure of WordNet to identify the different senses to define the meaning of the concepts based on using the noun hypernym senses. The ontology engineer interprets each concept and uses the WordNet noun hpernym senses to define a set of words that gives the concept meaning (W_1…Wn) i.e. its meaning for the given context. The Contextual Semantic Distance function (CSD) is then used to measure the degree of similarity of concepts in different ontology by using the following function.

$$CSD(O_1.C_1,O_2.C_2) = \frac{O_1.C_1(W1\ldots Wn) \cap O_2.C_2(W_1\ldots Wn)}{O_1.C_1(W_1\ldots Wn) \cup O_2.C_2(W_1\ldots Wn)}$$

Where $0 \le CSD \le 1$

If $CSD = 1$ Then

$O_1.C_1 = O_2.C_2$

Else If $CSD \ge T$ then (T = Threshold)

$O_1.C_1 \approx O_2.C_2$

Other wise

$O_1.C_1 \ne O_2.C_2$

5. NEGATIVE SEMANTIC REINFORCEMENT (NSR)

C-SAW uses words to specify the meaning of a concept, but these, or a combination of them, can have multiple meanings leading to ambiguity in the SAW. The C-SAW approach is enhanced by Negative Semantic Reinforcement (NSR). This involves the specification of an additional set of words that can say that the concept is not something else, improving semantic representation. The list of negative words can only be constructed from the other noun hypernym senses - allowing the ontology engineer to remove ambiguity between the senses. Negative Semantic Reinforcement is included in our *CSD* function;

$$CSD = \frac{O_1.C_1(W_1\ldots Wn) \cap O_2.C_2(W_1\ldots Wn) - O_1.C_1(\neg W_1\ldots\neg Wm) \cap O_2.C_2(\neg W_1\ldots\neg Wm)}{O_1.C_1(W_1\ldots Wn) \cup O_2.C_2(W_1\ldots Wn)}$$

6. EXPERIMENTATION & EVALUATION

Experimentation involved a comparison with the results given by [1] and [9]. Table 1 illustrates the differences.

Table 1. Comparison of C-SAW with other approaches

Word Pair		Gloss [1]	W&P [9]	C-SAW
Car	Automobile	1	1	Y
Gem	Jewel	1	1	Context Dependent
Journey	Voyage	1	0.91	Y
Boy	Lad	1	0.92	Y
Coast	Shore	1	0.89	Y
Asylum	Madhouse	1	0.94	Y
Magician	Wizard	1	1	Y
Midday	Noon	1	1	Y
Furnace	Stove	1	0.46	Y
Food	Fruit	0.14	0.73	Context Dependent
Bird	Cock	1	0.93	Context Dependent
Bird	Crane	0.22	0.82	Context Dependent
Tool	Implement	1	0.93	Context Dependent
Brother	Monk	1	0.93	Context Dependent
Crane	Implement	0.09	0.67	Context Dependent
Lad	Brother	0.33	0.67	Context Dependent
Journey	Car	0.17	0	N
Monk	Oracle	0.06	0.53	N
Food	Rooster	0.14	0	N
Coast	Hill	0.75	0.6	N
Forest	Graveyard	0.17	0	N
Monk	Slave	0.13	0.67	N
Coast	Forest	0.25	0.25	N
Lad	Wizard	0.13	0.67	N
Chord	Smile	0	0.38	N
Glass	Magician	0	0.18	N
Noon	String	0.07	0	N
Rooster	Voyage	0.849	0.82	N

'*Context Dependent*' in C-SAW illustrates this approach does not necessarily give the same result as [1] and [20], depending on the context in which the words can be defined. *Context Dependent* occurs in 29% of the cases. C-SAW identifies two mismatches (7% of cases), giving a total mismatch of 36% of the cases. The reason for the difference is that C-SAW would take account of the context for which the meaning of the word-pairs are defined. This lends contextual meaning to the word-pairs and hence C-SAW calculates the *CSD*. When C-SAW is implemented fully, (using relevant domain knowledge to create the SAW for each concept and incorporating Negative Semantic Reinforcement (NSR)), 29% of mismatches are eliminated by becoming matches or not-matches. C-SAW with NSR provides a more reliable approach compared to [1] and [9] as they have no way of being able to specify how the differences in the semantics of the concepts are resolved.

7. CONCLUSION

The C-SAW approach adds the meaning of each concept that is based on a set of Semantic Alignment Words (SAW). The existence of this set of words for each concept allows for greater automation in the semantic integration of ontologies and the repeatability of this process. The addition of a SAW for each concept, means that there is less ambiguity in the meaning, understanding and how the data should be manipulated. Negative Semantic Reinforcement (NSR) improves the clarity of concepts and accuracy of alignment/matching. The comparison of our approach (C-SAW) compared to [1] and [9] shows that they produce a large number of possible incorrect matches between the concepts. In our approach we expect a more reliable result as we will know, depending of the context, if the concepts are an actual match.

8. REFERENCES

[1] Achananuparp, P., Han, H., Semantically Enhanced User Modeling, Symposium on Applied Computing, March, 2007.

[2] Beneventano, D., Bergamaschi, S., Guerra, F., Vincini, M., *Synthesizing an Integrated Ontology*, IEEE Internet Computing, September-October, 2003

[3] Bouquet, P., Giunchiglia, F., Harmelen, F., Serafini, L., Stuckenschmidt, H., Contextualising Ontologies, Journal of Web Semantics, 2004, Vol. 1 No. 4.

[4] Fensel, D., Ontologies: A silver bullet for Knowledge Management and Electronic Commerce, Springer-Verlag, Berlin, 2003.

[5] Ghidini, C., Giunchiglia, F., Local Models Semantics, or Contextual Reasoning = Locality + Compatibility, Artificial Intelligence, 127 2, 2001 221-259.

[6] Giunchiglia, F., Shvaiko, P., Semantic Matching, The Knowledge Engineering Review Journal, Vol. 18(3), pp 265-280, 2003)

[7] Harmelen, F., Ontology Mapping: A way out of the Medical Tower of Babel, AIME, 2005, pp1-4.

[8] Klien. M., Combining and Relating Ontologies: An analysis of problems and solutions, Workshop on Ontologies and Information Sharing, IJCAI'01, 2001.

[9] Wu, Z. and Palmer, M., Verbs semantics and lexical selection. In Proceedings of the 32nd Annual Meeting on Association For Computational Linguistics (Las Cruces, New Mexico, June 27 - 30, 1994), 133-138.