Argumentation-Based Ontology Maintenance

Adrian Groza, Raluca Mechno

Abstract

The study proposes a method for ontology update based on the large amount of semi-structured data available from the semantic wikis. An argumentation reasoning process evaluates the contradictory pieces of knowledge posted by different users in order to adjust the axioms of an ontology or to decide each individual to what concept it belongs.

1 Introduction

The potential of combining Web 2.0 with Web 3.0 is advocated in literature [1]. At the moment, we are at the beginning of developing the social computing science [7]. In this line, the current study applies the active social machine behind semantic wikis to the hard task of ontology maintenance.

Ontologies are continuously confronted to the evolution problem. Our goal here is to develop a method for ontology maintenance by exploiting the work done in argumentation theory. An argumentation debate aims at increasing or decreasing the acceptability of a controversial standpoint for the listener or reader, by conveying a set of arguments which support or attack the issue in hand.

The paper presents an argumentation framework, which provides straightforward tool to reason about arguments posted by usual users. The proposal of this framework keeps the abstraction from the logic used to represent knowledge inside arguments while specifying a logic scheme to give some structure to arguments [12].

2 Technical Instrumentation

2.1 Description Logic

Description logics\(^1\) (DL) are a family of formal knowledge representation languages. They are more expressive than propositional logic but have more efficient decision problems than first-order predicate logic. DLs are used in Artificial Intelligence for formal reasoning on the concepts of an application domain, having an important role for defining, integrating, and maintaining ontologies.

In the description logic \(\mathcal{ALC}\), concepts are built using the set of constructors formed by negation, conjunction, disjunction, value restriction, and existential restriction, as table 1 bears out. Here, \(C\) and \(D\) represent concept descriptions, whilst \(r\) a role name. The semantics is defined based on an interpretation \(I = (\Delta^I, \cdot^I)\), where the domain \(\Delta^I\) of \(I\) contains a non-empty set of individuals, whilst the interpretation function \(\cdot^I\) maps each concept name \(C\) to a set of individuals \(C^I \in \Delta^I\) and each role \(r\) to a binary relation \(r^I \in \Delta^I \times \Delta^I\). The last column of table 1 illustrates the extension of \(\cdot^I\) to arbitrary concepts.

The syntax of a member of the description logic family is characterized by its recursive definition, in

\(^{1}\)http://en.wikipedia.org/wiki/Description_logic
which the constructors that can be used to form concept terms are stated. Some constructors are related to logical constructors in first-order logic (FOL) such as intersection or conjunction of concepts, union or disjunction of concepts, negation or complement of concepts, universal restriction and existential restriction. Other constructors have no corresponding construction in FOL including restrictions on roles for example, inverse, transitivity and functionality.

**Definition 1**
A concept \( C \) is satisfiable if there exists an interpretation \( i \) such that \( C^I \neq \emptyset \). The concept \( D \) subsumes the concept \( C \) (\( C \sqsubseteq D \)) if \( C^I \subseteq D^I \) for all interpretations \( I \).

**Definition 2**
An ABox is a finite set of concept assertions \( C(a) \) or role assertions \( r(a,b) \), where \( C \) represents a concept, \( r \) a role, and \( a \) and \( b \) are two instances. Usually, the unique names assumption holds within the same ABox. A TBox is a finite set of terminological axioms of the form \( C \equiv D \) or \( C \sqsubseteq D \).

### 2.2 Semantic Wikis

Semantic wikis provide users the capability to annotate their text with specific concepts and roles from a set of imported ontologies, in order to be processed against semantic queries. Among the available semantic wikis, such as DBpedia [4], ACEWiki [10], or OntoWiki [2], we drive our attention towards Semantic Media Wiki (SMW), due to its success in terms of number of users.

Semantic MediaWiki (SMW) is an extension of MediaWiki, the wiki application best known for powering Wikipedia, that helps to search, organise, tag, browse, evaluate, and share the wiki’s content. While traditional wikis contain only text which computers can neither understand nor evaluate, SMW adds semantic annotations that allow a wiki to function as a collaborative database.

To create the template, used for the argumentation scheme, it was used the "special" page, Special: CreateTemplate, defined by SpecialForms\(^3\), an extension of SMW. With this extension it was created also the properties and categories. Another extensions: Halo could be used to import the ontology, ExportRDF to export the data in RDF format.

The information collected from template, based on Semantic MediaWiki, relates to the ontology language [17]. In MediaWiki the method used for entering information into a wiki is wikitext, which is transformed into HTML pages. For the interrelation between pages, hyperlinks are used. All the defined pages are classified in namespaces, which cannot be defined by wiki users. Every page can have one or more categories.

Like in MediaWiki, semantic data in Semantic MediaWiki is structured by pages and every page corresponds to an ontology entity. Different ontology entities can be represented by namespaces: individuals are represented by the majority of the pages, classes by categories in MediaWiki, which classify individuals an also create subcategories, properties are relationships between two individuals or an individual and a data value and types for distinguish different kinds of properties. Categories are available in MediaWiki, but properties and types are introduces by SemanticMediaWiki. The elements of wikitext are easy to interpret: triple quotes “’...’” are used for bold text, the text within brackets square [[...]] is transformed into link. For example [[England]] is just a link which do not carry any machine-understandable semantics yet. For instance, asserting a property called capital of to London can be done by writing [[capital of :: England]], where property capital of has value England.

\(^2\)http://semantic-mediawiki.org

\(^3\)http://www.mediawiki.org/wiki/Extension:Semantic_Forms
Argumentation-Based Ontology Maintenance

Owl individual | Semantic MediaWiki
---|---
owl:Class | article in namespace Category
owl:ObjectProperty | article in namespace Relation
owl:DatatypeProperty | article in namespace Attribute

Statement about element page | Syntax in wiki-source of page
---|---
objectProperty | [[propertyName::objectArticle]]
attributeProperty | [[propertyName:=valueString]]
rdf:type className | [[Category:className]] (on article pages)
rdfs:subClassOf className | [[Category:className]] (on category pages)

Figure 1: Description and relationships between individuals

3 System Architecture

3.1 Engineering the Restaurant ontology

To model the interaction between arguments, an ontology made in Protege\(^4\) is used. In the development process of "Restaurant" ontology several engineering steps were followed. The first step is to determine the scope, by creating some competency questions like: Which is the best restaurant from Cluj? What features should be taken into account, when you choose an restaurant? What kind of kitchen does one prefer? Is it close to work? Exists room for smoking? How long does it take to be served?.

In the second step several knowledge repositories have been considered for re-use: the restaurant.owl\(^5\), time.owl\(^6\) and food.owl\(^7\) ontologies. In the third step relevant terms are identified: Cuisine, Asian Cuisine, African Cuisine, location, kitchen, Tusa, Shangai, has time of service, has location. The nouns form are the basis for the class names and the verb phrases form the basis for property names. The classes are created in the fourth step: SmokingRoom, Restaurant, Location, Cuisine, African. In the fifth five, a property named hasServiceTime was created and some object properties: hasCuisine, hasAddress, isCuisine, hasSmokingRoom. Finally, the instances for the classes were added and the ontology was checked for anomalies.

\(^4\)http://protege.stanford.edu/
\(^5\)http://gaia.fdi.ucm.es/ontologies/restaurant.owl
\(^6\)http://www.w3.org/TR/owl-time/
\(^7\)http://krono.act.uji.es/Links/ontologies/food.owl/view
Figure 3: Part of the restaurant ontology.

According to axiom 1, in figure 3, \(tusa\) is an instance of class \(Restaurant\). The instance \(tusa\) has the name "TusaSRL" based on axiom 2. In axiom 3 the class \(Restaurant\) is included in the result of intersection of classes \(SmokingRoom\), \(Location\) and \(Cuisine\). Axiom 4 describes that the class \(Restaurant\) must have instances of classes \(Name\), \(Address\), \(Cuisine\) and \(ServiceTime\). Class \(Restaurant\) is a subclass of \(Restaurant\) (axiom 5). Axiom 6 define class \(Cuisine\), which is included in the result of union of classes \(Asian\), \(African\), \(European\) and \(CuisineStyle\).

The inverse role of \(isKitchen\) shown by axiom 7 is \(hasKitchen\).

### 3.2 Argument Representation

For the abstraction of the debate, we use the theoretical model of Walton based on argumentation schemes [18]. Argument schemes encapsulate common patterns of human reasoning such as: argument from position to know, argument from evidence, argument from sign, etc. Argumentation schemes are defined by the following items: a name, a set of premises \(A_i\), a conclusion \(C\). Figure 5 details these attributes of Argument from position to know scheme.

The template that was created for the argumentation scheme is described in figure 6. The user completes the following fields: \(Premise\) represents the premise of the argument, \(ValueOfCredibilityOfPremise\) represents the certainty of the premise, \(Conclusion\) is the text that represents the conclusion, made based on premise, \(ValueOfCredibilityOfConclusion\) represents the value that will be computed with equation 1, \(ValueOfCredibilityOfUser\) represents value of credibility of user, calculated with equation 3, \(User\) is the name of the user that complete the fields of the template.

\[
\text{credibility}(c) = \min(v_i) \ast \alpha_f \ast \mu_u
\]

where \(v_i\) represents the values of the premises, \(i\) could have values between 1 and 5, user can choose how many premises will use, the template illustrated in figure 7 have just one premise, if the premise is an existing argument, the value taken will be the value of credibility of that argument, otherwise, the value taken will be 1, \(\alpha_f\) represent the value of the strength with which it is sustained the premise, this value is computed with equation 2 and \(\mu_u\) is the value of credibility of the user, calculated based on 3.

\[
\alpha_f = \sigma_v \ast (\rho_t - \rho_f) / \rho_t
\]

where \(v_t\) represents all the arguments posted by the user \(u\) and \(\mu_f\) all the arguments posted by the user \(u\), but was not valid.
Figure 4: Graphical representation of the restaurant ontology.

\[ A_1: E \text{ asserts that } A \text{ is known to be true.} \]
\[ A_2: E \text{ is an expert in domain } D. \]
\[ C: A \text{ may (plausibly) be taken to be true.} \]

Figure 5: Using argumentation scheme for structured arguments.

```{Argumentation
|Premise=
|ValueOfCredibilityOfPremise=
|Conclusion=
|ValueOfCredibilityOfConclusion=
|User=
|ValueOfCredibilityOfUser=
}
```

Figure 6: Semantic template for the argumentation scheme.
4 Running Scenario

This section presents a usage scenario and the application results.

4.1 Knowledge elicitation

The user can choose the number of premises, for the argumentation scheme that wants to sustain. We suppose that some users posted the following arguments, about the ontology restaurant.owl:

```
{{Argumentation
|Premise= Monday, 4PM, sustain quick service at tusa
|ValueOfCredibilityOfPremise= 0.28
|Conclusion=quick service at tusa
|ValueOfCredibilityOfConclusion= 0.084
|User= u1
|ValueOfCredibilityOfUser= 0.3
}}
```

The user $u_1$ with value of credibility 0.3, obtained based on equation 3, where $\mu_t=10$ and $\mu_f=6$, uses the word "sustain", which has the value 0.7, that Monday at restaurant Tusa, on 4PM it is always quick service. The force with which the premise is sustained is computed with equation 2 where, value of the word from tree is 0.7, $\rho_t = 10$ and $\rho_f = 6$, then the final value is 0.28, so the value of credibility of conclusion computed with equation 1 will be 0.084.

```
{{Argumentation
|Premise= Monday, 4PM, attack quick service
|ValueOfCredibilityOfPremise= 1
|Conclusion=not quick service
|ValueOfCredibilityOfConclusion= 0.020
|User= u2
|ValueOfCredibilityOfUser= 0.17
}}
```

The user $u_2$ with value of credibility 0.17, obtained based on equation 3, where $\mu_t=17$ and $\mu_f=7$, uses the word "attack", which has the value 0.15, that Monday at restaurant Tusa, on 4PM it is always quick service. The force with which the premise is sustained is computed with equation 2 where, value of the word from tree is 0.7, $\rho_t = 6$ and $\rho_f = 1$, then the final value is 0.83, so the value of credibility of conclusion computed with equation 1 will be 0.020.

```
{{Argumentation
|Premise= quick service at tusa, agree, high quality
|ValueOfCredibilityOfPremise= 0.084
|Conclusion=high quality at tusa
|ValueOfCredibilityOfConclusion= 0.102
|User= u3
|ValueOfCredibilityOfUser= 0.42
}}
```
Argumentation-Based Ontology Maintenance

1. \text{hasServiceTime}(\text{tusa}, \text{Medium})
2. \text{hasServiceTime}(\text{tusa}, \text{Medium}) \sqsubseteq \forall \text{hasServiceTime}(\text{Monday} \sqcup \text{Tuesday} \\
\sqcup \text{Wednesday} \sqcup \text{Thursday} \sqcup \text{Friday} \sqcup \text{Saturday} \sqcup \text{Sunday})
3. \text{hasServiceTime.} \text{Monday} \sqsubseteq (\forall \text{hasServiceTime}(\text{MI} \sqcap \text{MII} \sqcap \text{MIII}))
4. \text{Monday} \sqsubseteq (\exists \text{hasServiceTime.} \text{Quick})
5. \text{MI} \sqsubseteq (\exists \text{hasServiceTime.} \text{Quick})
6. \text{MII} \sqsubseteq (\exists \text{hasServiceTime.} \text{Medium})
7. \text{MIII} \sqsubseteq (\exists \text{hasServiceTime.} \text{Quick})

Figure 8: Axioms for the initial ontology

The user \text{u}_3 with value of credibility 0,42, obtained based on equation 3, where \(\mu_t=7\) and \(\mu_f=4\), uses the word \textit{agree}, which has the value 0,5, that at restaurant \textit{tusa}, it is \textit{highquality} based on the valid argument, \textit{quickservice at tusa}. The force with which the premise is sustained is computed with equation 2 where, value of the word from tree is 0,5, \(\rho_t = 15\) and \(\rho_f = 6\), then the final value is 0,30, so the value of credibility of conclusion computed with equation 1 will be 0,102.

\{
{Argumentation
|Premise= Sunday, 6PM, decline quick service at tusa
|ValueOfCredibilityOfPremise= 0,57
|Conclusion=not quick service at tusa
|ValueOfCredibilityOfConclusion= 0,15
|User= u4
|ValueOfCredibilityOfUser= 0,41 }
\}

The user \text{u}_4 with value of credibility 0,41, obtained based on equation 3, where \(\mu_t=17\) and \(\mu_f=10\), uses the word \textit{decline}, which has the value 0,65, that \textit{Sunday} at restaurant \textit{tusa}, on 6PM it is \textit{quickservice}. The force with which the premise is sustained is computed with equation 2 where, value of the word from tree is 0,65, \(\rho_t = 7\) and \(\rho_f = 3\), then the final value is 0,57, so the value of credibility of conclusion computed with equation 1 will be 0,15.

4.2 Ontology Enrichment

This subsection illustrates how the arguments posted affects the ontology or not. Consider part of the initial ontology illustrated in figure 8.

By the axiom 1, figure 8, the service time for restaurant \textit{tusa} is \textit{Medium}. The role \textit{hasServiceTime} of restaurant \textit{tusa} is composed of union of values of \textit{hasServiceTime} for all days from the week. Based on axiom 3, the value of \textit{hasServiceTime} for day \textit{Monday} results from the intersection of values of \textit{hasServiceTime} for all 3 parts of the day, where \textit{MI} is between 8AM and 12AM, \textit{MII} between 12AM and 3PM and \textit{MIII} between 3PM and 9PM.

According to axiom 4, the value of \textit{hasServiceTime} for \textit{Monday} is \textit{Quick}. This value results from the comparison of the values obtained in those 3 parts of the day. The value of the role \textit{hasServiceTime} for part \textit{MI} is initial \textit{Quick} (axiom 5). The value of the role \textit{hasServiceTime} for part \textit{MII} is initial \textit{Medium} (axiom 6). The value of the role \textit{hasServiceTime} for part \textit{MIII} is initial \textit{Medium} (axiom 7).

The arguments are taken into account if the value of conclusion of the argument is bigger then the actual value of the proper part. The arguments which are not taken into account will be saved and will be proposed for improvement. If other users will sustain the arguments proposed for improvements, the value of the sustained argument will be updated, by adding the current value. To support the decision of change, ontology engineers will use metrics for ontology evaluation and validation.

The value of credibility of conclusion for the argument proposed by the user \textit{u1} is taken into account by the ontology engineers. This argument will change the value of the \textit{hasServiceTime} for \textit{MII}, because the initial value for \textit{MII} is \textit{Medium} and it is different from the conclusion of the argument. The axiom from the new ontology become:

\[
\text{MII} \sqsubseteq (\exists \text{hasServiceTime.} \text{Quick})
\]
The value of the role hasServiceTime for part MII will be Quick. In this case, this result will not affect the value of hasServiceTime of day Monday, for restaurant Tusa because it is already Quick.

5 Discussions and Related Work

Argumentation is a vital aspect of intelligent behaviour by humans. Consider diverse professionals such as politicians, journalists, clinicians, scientists, and administrators, who all need to collate and analyse information looking for pros and cons for consequences of importance when attempting to understand problems and make decisions [8].

There are a lot of approaches for the argumentation like: real arguments [16], fuzzy argumentation [9], arguments in OWL [14]. In case of real arguments, they do not have enough explicitly presented premises for the entailment of the claim and the proponent of an argument encode an argument into a real argument by ignoring the common knowledge, and it allows a recipient of a real argument to decode it, also by ignoring the common knowledge. In our view the user is forced to complete all the fields in the created template, so there will not be empty fields about the common knowledge before he create arguments.

The fuzzy approach enriches the expressive power of the classical argumentation model by allowing to represent the relative strength of the attack relationships between arguments, as well as the degree to which arguments are accepted [9]. In our case, a method for computing the relative strength of the attack relationships between arguments and also the support relationships between them is proposed. The computation method is based on a word tree, in which each node has a degree of support (degree[0, 1]) attribute, assessed based on the degree of support of the sons and a historic attribute. The historic attribute represents the normalization of the arguments in which the current node was used.

The unambiguous and effective delivery of data and knowledge on the Web relies heavily on the correct representation and understanding of the associated contexts. However, the current way of encoding contexts of data and knowledge on the Web is largely ad hoc. Contexts are often embedded in the application programs or are implied by the application or community-specific agreements. This makes the linking and reusing of data and knowledge, and thus the integration of Web applications, a difficult problem. Therefore, building the architectural support for contexts is one of the major challenges for the Web, and in particular, for the Semantic Web [3].

Similar to our paper, ArgDF [11] is a framework which addresses the issue of argument representation in Semantic Web, aiming at developing the Argumentative Web, a large scale network of interconnected arguments created by human agents in a structured manner [15], based on the argument interchange format (AIF) ontology. The main aims of the AIF ontology are: i) to facilitate the development of (closed or open) multi-agent systems capable of argumentation based reasoning and interaction using a shared formalism; and ii) to facilitate data interchange among tools for argument manipulation and argument.

Design patterns are used in [13] to manage ontologies in an easier manner. ODPs are building blocks for ontology management representing small ontologies that can be extended and adapted to a specific application. An initial ontology is enriched based on ODPs in [5, 6]. The process is semi-automatic and has been implemented in two phases: i) element extraction uses an initial ontology in order to extract elements together with a confidence ii) patterns matching and ranking evaluates against ODPs the ontology elements previously extracted based on words metrics or using WordNet. The ontology is evaluated and enriched with the best new elements. In our case, the up to date knowledge is extracted from semantic wikis, where the decision to add an element or the change a axiom is taken after an argumentation process.

Several approaches for extracting concepts, instances and relationships exploit separately or integrate statistical methods, semantic repositories such as WordNet, natural language processing libraries such as OpenNLP, or lexicon-syntactic patterns in form of regular expressions [19]. Instead of NLP techniques we exploit here argumentation schemes for extracting structured information. When creating an application based on semantic knowledge it is necessary to guarantee that the considered ontology meets the application requirements. In this line, ontology evaluation is important in cases where the ontology is automatically populated from different resources that might not be homogeneous, leading to duplicate instances, or instances that are clustered according to their sources in the same ontology [20].
6 Conclusion

This research addresses the task of ontology maintenance by exploiting the large amount of structured information available in semantic wikis. The proposed solution makes use of argumentation theory, using as an argumentation platform Semantic MediaWiki, aiming at enacting the argumentative web as envisaged in [14].

Ongoing work regards empirical evaluation of the method by using several evaluation metrics. WordNet ontology can be exploited even more, and with the help of OpenNLP, relationships between concepts from the ontology or new domain concepts could be discovered even when the context of use causes word ambiguity.

References


**Acknowledgment:** We are grateful to the anonymous reviewers for their useful comments. This work was supported by the grant ID 170/672 from the National Research Council of Romanian Ministry of Education and Research and POSDRU/89/1.5/S/62557/EXCEL.

Groza Adrian  
Technical University of Cluj-Napoca  
Computer Science Department  
Memorandumului 28, Cluj-Napoca  
ROMANIA  
E-mail: adrian.groza@cs.utcluj.ro

Raluca Mechno  
Technical University of Cluj-Napoca  
Computer Science Department  
Memorandumului 28, Cluj-Napoca  
ROMANIA  
E-mail: raluca.mechno@student.utcluj.ro