

Visualizing Consensus with Online Ontologies to Support Quality in Ontology Development

Mathieu d'Aquin, Enrico Motta

Knowledge Media Institute, The Open University, Milton Keynes, UK
{m.daquin, e.motta}@open.ac.uk

Abstract. Representing a consensus is one of the most important of the qualities of an ontology. Generally, consensus is understood in ontology engineering as an accurate representation of a shared view or model within an ontology. However, we can also show that, while developing an ontology, checking for the agreement of other existing ontologies that have been published online can be very valuable. Indeed, relying on the Watson Semantic Web search engine and on existing measures of agreement, consensus and controversy in ontologies, we develop a visualization method that allows to emphasize ‘areas’ of an ontology which are in disagreement with online ontologies, as well as elements touching on controversial statements, for which a clear consensus has not been established. Integrating this visualization into the ontology development environment allows for the ontology engineer to quickly identify elements of the ontology which quality might need to be improved, as they relate to such generally disagreed or controversial areas.

1 Introduction

In a previous paper, we presented a set of measures allowing to check if an ontology agreed or disagreed with a given statement [1]. These measures could flexibly represent situations where gradual agreement and disagreement were needed, as well as where no actual ‘opinion’ about the statement was expressed in the ontology. From such measures, we derived notions of global agreement and disagreement in a collection of ontologies, and showed using the Watson Semantic Web search engine (<http://watson.kmi.open.ac.uk>) that interesting quality information can emerge from checking such global agreement and disagreement in online ontologies, in search for consensus (either positive, with high global agreement on the statement, or negative, with high global disagreement). This also allows one to identify statements for which there is not clear cut division between agreement and disagreement, therefore having a low absolute value of consensus, leading to the notion of ‘controversy’ in online ontologies.

In this paper, we consider the practical application of this fundamental work to guide the ontology engineer towards better quality ontologies. Indeed, giving a direct, easy to explore and interpret way to visualize areas of ontologies with negative consensus can help the ontology engineer in discovering mistakes and revising the ontology towards the commonly admitted view. Conversely, showing areas of positive consensus can reinforce the developer of the ontology into a particular design choice. In a possibly more subtle way, through identifying areas of controversy in the ontology, such a visualization puts emphasis on the elements

for which more careful validation should be applied, and a specific design choice might have to be more firmly established.

We describe how such a visualization has been integrated in the ontology development environment of the NeOn Toolkit (<http://neon-toolkit.org>), allowing the ontology engineers to test their ontologies for consensus and controversy, as they are being developed. In Section 2, we give a quick summary of the measures of agreement, disagreement, consensus and controversy already described in [1]. We then, in Section 3, detail the visualization we devised from applying these measures on an ontology graph. Section 4 gives more detail on the tool itself and shows several examples clearly demonstrating how such a tool can be used to improve ontology quality in ontology development.

2 Measuring Consensus with Respect to Online Ontologies

Ontologies are knowledge artifacts built within the communities that rely on them, meaning that they represent consensual representations inside these communities. However, when considering, the set of ontologies distributed on the Web, many different ontologies can cover the same domain, while being built by and for different communities. They can therefore represent different opinions about particular objects and entities, which can be in agreement or in disagreement with given statements on these entities.

In [1] we defined two basic measures for assessing agreement and disagreement of an ontology O with a statement $s = \langle \textit{subject}, \textit{relation}, \textit{object} \rangle$ (due to space limitations, giving the complete definitions of these measure is not feasible here. We encourage the interested reader to refer to the original paper):

$$\begin{aligned} \textit{agreement}(O, s) &\rightarrow [0..1] \\ \textit{disagreement}(O, s) &\rightarrow [0..1] \end{aligned}$$

Two distinct measures were used for agreement and disagreement so that an ontology can, at the same time and to certain extents, agree and disagree with a statement. These two measures have to be interpreted together to indicate the particular belief expressed by the ontology O regarding the statement s . For example, if $\textit{agreement}(O, s) = 1$ and $\textit{disagreement}(O, s) = 0$, it means that O fully agrees with s and conversely if $\textit{agreement}(O, s) = 0$ and $\textit{disagreement}(O, s) = 1$, it fully disagrees with s . Now, agreement and disagreement can vary between 0 and 1, meaning that O can only partially agree or disagree with s and sometimes both, when $\textit{agreement}(O, s) > 0$ and $\textit{disagreement}(O, s) > 0$. Finally, another case is when $\textit{agreement}(O, s) = 0$ and $\textit{disagreement}(O, s) = 0$. This means that O neither agrees nor disagrees with s , as it does not express any belief regarding the relation encoded by s .

The actual values returned for both measures, when different from 0 and 1, are not very important. They correspond to different levels of dis/agreement and only an order between pre-defined levels is needed to interpret them. The values used and the ways to compute them are given in [1].

From such measures, an interesting information can be derived based on exploiting the collection of ontologies in Watson, i.e., the level to which particular

statements are globally agreed with in online ontologies, or in other terms, the level of consensus on a statement. Conversely, a related information concerns the level of controversy on the statement, i.e., whether there is a clear cut between agreement and disagreement. A normalized mean is used to measure the global agreement and disagreement of a statement st in a set of ontologies R (being here the collection of ontologies in Watson, see details in [1]). From these two measures, consensus is defined as having a high level of certainty on whether ontologies in R agree or disagree with st : There is a high level of (positive consensus) if the overall agreement about this statement is high and the overall disagreement is low. Thus, the measure of consensus is computed in a set of ontologies R upon a statement st as follows:

$$consensus(st, R) = agreement(st, R) - disagreement(st, R)$$

It is important to notice here that consensus can either be positive, when there is a high level of consensus on agreeing with the statement, or negative, when ontologies in R generally disagree with st .

The notion of controversy is considered to be the inverse from the one of consensus: there is a high level of controversy on a given statement when there is no clear cut between agreement and disagreement, i.e. there is a low level of consensus. The measure of controversy in a set of ontologies R upon a statement st is straightforwardly defined as:

$$controversy(st, R) = 1 - |consensus(st, R)|$$

Looking at examples presented in Table 1, we can see that some statements have a very high level of agreement, and a very low level of disagreement, leading to a high positive consensus and a low controversy level. For the 3 last statements however, there is a higher level of controversy. The last one is by far the most disagreed with, having a high level of negative consensus.

Table 1. Consensus and controversy on statements about *SeaFood* class in Watson.

Statement	Consensus	Controversy
$\langle SeaFood, subclassOf, EdibleThing \rangle$	1.0	0.0
$\langle ShellFish, subclassOf, SeaFood \rangle$	0.89	0.109
$\langle Fish, subclassOf, SeaFood \rangle$	0.875	0.125
$\langle SeaFood, disjointWith, Fruit \rangle$	0.75	0.25
$\langle Meat, disjointWith, SeaFood \rangle$	0.53	0.46
$\langle SeaFood, subclassOf, Meat \rangle$	-0.719	0.281

3 Visualizing Consensus in an Ontology

Our goal here is to use the measures presented above to produce an overview of an ontology, showing areas of positive and negative consensus, as well as areas of controversy. The basic idea is to display the ontology as a graph, with statements relating entities with each other, and colors indicating the values of

the measures. Each statement st in the graph is an edge, which is blue if the consensus is on agreement (i.e., $consensus(st, R) > 0$) and red if the consensus is on disagreement. A special case to take into account is when no ontology actually contains any relation that could be compared to the statement st (i.e., the two entities related by the statement do not appear in other ontologies). In such a case, where we can only express a neutral opinion, the edge is green.

The level of consensus, and corollary, the level of controversy, is represented in this graph through the brightness of the edges. In other words, the color level for each edge is directly proportional to the value of $consensus(st, R)$. If the value is 1, the edge will be represented with a pure, bright blue. If the consensus value is -1 , it will be pure, bright red. If the value is 0, the edge will be represented black, representing the highest level of controversy. Accordingly, any value between -1 and 0 shows as an accordingly dark red, and values between 0 and 1 are more or less blue (see the examples next section).

Nodes in the graph represent entities that are linked through the statements. The basic idea here is that the color of each node should represent an aggregated view of the statements that relate to it. Each component of the color (red, green or blue) is therefore the average of the corresponding component of the colors of the ingoing and outgoing edges of the node. In this way, an entity mostly related through positive edges will be blue, implying that there is a general, positive consensus on the definition of this entity. Similarly, if there is a high level of consensus on the statements linking the entities, the corresponding nodes will be colored brightly, while many controversies in these statements will make them darker. In this way, we expect to be able to quickly identify areas of negative consensus, looking for nodes and edges colored red, as well as controversial elements, by looking at dark zones in the graph.

4 Implementation, Tests and Examples

We implemented the visualization described above as a plugin for the NeOn Toolkit and relying on the Java API of the Watson system to access online ontologies. It works very simply by showing the graph of the currently selected ontology, and re-computing it every time the selection changes. Integrating with an ontology development environment is very useful as elements discovered can be directly rectified, with the changes being propagated to the graph. Also, one interesting element is that the tool, through the Watson API, has access to the ontologies used to compute the consensus, so that a user can always inspect them, and obtain explanation and different views on the representation of the considered entities.

Also, the complexity of computing the consensus level of a statement being high, cache mechanisms are included to avoid having to re-compute this value more than once for each statement. We tested this tool on a number of example ontologies, in order to show how it could help their developers identify and improve some elements.

The AKT Portal Ontology is designed to represent a research domain, including classes such as organizations, universities, researchers, or publications. As can be seen from Figure 1, this ontology is well covered by online ontologies

and generally, there is a high level of positive consensus about its content, even if some areas have brighter blues than others. We can however also identify high levels of controversy in the dark area towards the middle right part of the graph. This part corresponds to the “employees” branch where the taxonomy can be questionable. In particular, it contains a statement that a “visiting researcher” is an “affiliated person” with which online ontologies tend to disagree.

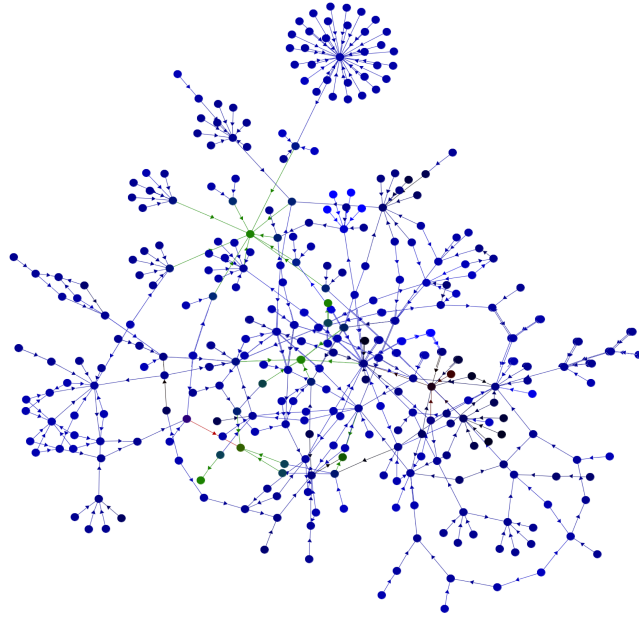


Fig. 1. Consensus visualization for the AKT Portal ontology.

The AKT Support Ontology is the upper level ontology for AKT portal, containing general classes such as “physical entities” and units of measures. Here again, with various levels, there is a general positive consensus on this ontology (Figure 2). However, one statement can clearly be identified as problematic: “duration” as a subclass of “physical quantity” (bottom right part of the graph, the dark red edge).

The Drama Ontology has been built locally to represent the area of classical greek dramas and their recent productions. As can be seen from Figure 2, as the domain is more specialized and less covered by online ontologies, fewer statements could be evaluated for consensus. However, even in this case, we can clearly identify elements of high positive consensus (e.g., “Company” is a subclass of “Organization”) and of clear negative consensus (e.g., “Archive” is a subclass of “Organization”, organization being the darker node located in the middle left area of the graph). The other area with many controversies (middle right part of the graph) is centered around the class person, where for example, statements such as “Character” subclass of “Person” appear clearly to be disagreed with, while others such as “Editor” being subclass of “Person” are shown darker, and so more controversial.

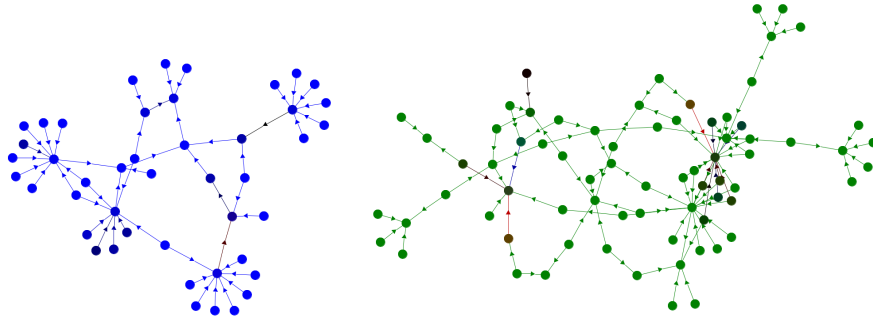


Fig. 2. Consensus visualization for the AKT Support ontology (left) and of the Drama ontology.

5 Conclusion

In this paper, we have presented a new visualization technique relying on measuring the consensus between statements and online ontologies, in order to support an ontology developer in identifying potentially problematic (disagreed with or controversial) areas to be improved in an ontology. Such an approach can be related to the many other approaches for ontology visualisation, that especially focus on summarising the knowledge ontologies contain from a given perspective (see e.g., [2, 3]). However, to the best of our knowledge, this is the first tool that integrates external ontologies to give an overview of the consensus and controversies present in an ontology. This tool is integrated with the ontology development environment and has been shown on different examples to be able to actually provide valuable visual overviews of the consensus levels in elements of the ontology.

Amongst the limitations of this approach is the complexity of the consensus measure, making the visualization hard to scale to larger ontologies. Cache mechanisms are implemented and approaches are envisaged to provide a gradual, incremental view of the ontology so that the user does not have to wait for the entire graph to compute. An improved interaction with the graph, allowing to more easily navigate it, identify important areas and obtain explanations for the measures is also part of our future work. Finally, we plan to test the tool as part of the building of new ontologies, checking and revising problematic parts as the ontology is being developed.

1. d'Aquin, M.: Formally measuring agreement and disagreement in ontologies. In: International Conference on Knowledge Capture - K-CAP. (2009)
2. Motta, E., Li, S.P.N., d'Aquin, M.: KC-Viz: A Novel Approach to Visualizing and Navigating Ontologies. In: Demo at The 17th International Conference on Knowledge Engineering and Knowledge Management by the Masses. (2010)
3. Khatchadourian, S., Consens, M.: Explod: Summary-based exploration of interlinking and rdf usage in the linked open data cloud. The Semantic Web: Research and Applications (2010) 272–287