Degree-of-Interest Visualization for Ontology Exploration

Peter Hüsken and Jürgen Ziegler

Department of Computer Science, Faculty of Engineering, University Duisburg-Essen,
47048 Duisburg, Germany
{huesken, ziegler}@interactivesystems.info,
WWW home page: http://live.interactivesystems.info

Abstract. In recent years, improvements in semantic web technologies
have given us new expressive description languages for modeling knowl-
edge domains — the so called ontologies. Nevertheless, ontology editors
lack of easy and intuitive user interfaces, so that the exploration and
creation of ontologies is often too difficult to be efficient. In this short
paper, we introduce a new tree widget which utilizes sophisticated visu-
alization and interaction features for ontology exploration and editing as
a work in progress study. Due to space limitations we concentrate here
on the aspect of ontology browsing.

1 Visualization and Interaction Issues for Ontologies

Collaborative development of ontologies is becoming an important activity in
various scientific and professional communities. In the context of the Ontoverse
project[1] we are trying to develop more intuitive user interfaces as well as col-
laboration support for this task. In this paper, we present a novel visualization
technique for ontology development.

Existing ontology tools, such as Protege [1] concentrate on visualizing the
basic class hierarchy, which can be seen as the structural backbone of any on-
tology by presenting different kinds of tree views. Graph or network visualiza-
tions tend to be less informative when showing ontologies with several thousand
nodes [2]. Our efforts are inspired by the family of *Focus+Context* techniques,
applying fisheye perspectives that have been introduced by Furnas [3]. Nodes
are automatically displayed or elided according to the user's computed degree-
of-interest (DOI) as explained in section 2. In this regard, Card et al. [2] describe
the application of the DOI concept for tree layouts as *logical filtering* of nodes.
Their implementation of the Focus+Context tree is completely visible without
the need to scroll. Showing the whole information structure, leads to distorted
visualizations where many nodes are typically too small to read their labels.

In our approach, we also use a distortion-based rendering of the tree while
keeping the node size at a readable level. As a consequence, scrolling is needed.

We aim at minimizing the need to scroll however, by providing smart ways to hide parts of the tree that are not relevant in the current context. To provide Focus+Context, the currently selected concept is rendered with a larger size. Further we apply a multifocal approach to highlight also other concepts that the user is probably interested in — those could be all concepts that are directly related to the selected concept by OWL object properties [5]. Object properties are represented as directed lines beside the concept hierarchy, which connect the selected concept with other concepts in the ranges of the selected concept's object properties. These so-called *PropertyLines* are shown on the right side of fig. 1. Here *BioinformaticsTask* is connected to the concepts *ComputationalMethod* and *Program* with additional straight lines beside the tree illustration. In this way, the user gets an impression about the semantic interrelations as being part of the ontology.
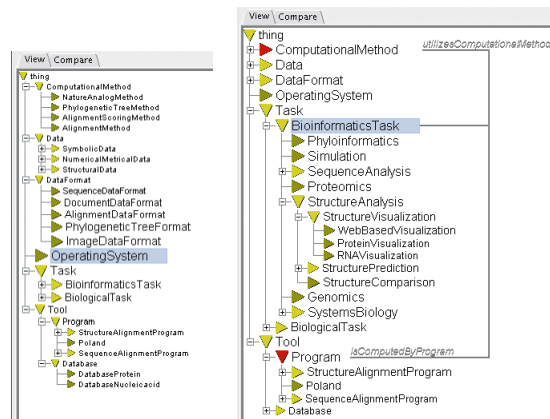

Fig. 1. SmartTree with Focus+Context (left) and Property-Lines (right)

As an additional means to reduce the complexity of the visualization, we have implemented an interactive function called *Condense & Explode*: After selecting the line, connecting subtrees with a common parent node, every subtree outside the focus is faded out (condensed), representing all hidden subtrees by an elision symbol. Clicking on the line once more will show the full tree again. In this way we can better utilize vertical space to provide the user with a more suitable overview of the current relevant parts of an ontology. Future versions of implementation will apply those interaction functions automatically by making use of the DOI concept (see section 2).

2 User Adaptation

For estimating the degree of user interest in a node, we can monitor the observable user behaviour and parameters such as the history of concept selection. In

order to estimate the *Degree of Interest* value (DOI) of every concept representation, we have to distinguish two basic factors:

1. *A Priori Importance (API).* Independent of any application context the API value of a concept representation depends on the ontology structure and is constant as long as the given ontology remains unchanged. API values have to be updated after a new ontology version has been released. We identified two simple rules for estimating the API value of concept respresentations:
    - The importance of a concept is the higher the more instances of the concept exist, because the frequency of assertions is a strong indication for an importance independent of any user.
    - The importance of a concept is the higher the more object properties use the concept as part of their domain or range, because we suppose that a concept with many object properties will be used frequently in the concept's related assertions. Again, a high number of assertions probably denote importance of the concept in question.
2. The *Distance D(x, SP)* between concept x and the concept SP that has the focus (also called *Selection Point*). A larger distance means that concept x is less important. Different types of distances can be applied. To produce the fisheye effect, which is essential to Focus+Context techniques, the geometric distance D_{vis} is used. It is usally defined as the number of concepts between SP and x inside the rendered tree layout. The effect is that the user gets a better overview of those concepts being close to the selected concept. It is also possible to use a taxonomic distance D_{tax} as one instance of semantic distance presented by Rada et al. [6]. To sum up the idea, two assumptions are made: (a) Concepts are similar (less distant) when the shortest path between them in the taxonomy is brief and (b) concepts at upper levels of the concept hierarchy are more general and are semantically less similar than concepts at lower levels. A combined type of distance is used by Card et al. [2]. They assume that a node's importance decreases intrinsically with its path distance to the hierarchy root together with the geometric distance to the focus node(s).

The DOI value for concept x with respect to the concept in focus can be calculated by function F:

$$DOI(x, SP) = F(API(x), D(x, SP)) \tag{1}$$

Equation 2 shows an example of F. The additional factor β relatively weights the distance D ($\beta = 0$: only $API(x)$ is considered, $\beta = \infty$: only $D(x, SP)$ is relevant, $\beta = 1$: both arguments are equally weighted). The arguments API and D range from zero to one. As a consequence, the function F_β has the same range.

$$F_\beta(x, y) = \frac{(\beta^2+1) \cdot x \cdot y}{\beta^2 \cdot x + y} \tag{2}$$

The explorations of interactive visualizations can be seen as an iterative process, because the user plans the next steps based on attained information. As a consequence the sequence of concept selections has to be followed:

$$DOI_i(x, SP_i) = \begin{cases} API(x) & : i = 0 \\ F(DOI_{i-1}(x, SP_{i-1}), D(x, SP_i)) & : i > 0 \end{cases} \qquad (3)$$

Applying user adaption in this way, fast distant calculations are required, otherwise the SmartTree's performance is affected adversely.

3 Summary and Outlook

In this paper we have introduced a customized tree widget for ontology exploration with new interactive functions. User adaptation can be realized by monitoring the sequence of user interactions, so that those concepts get a higher DOI value that have been selected recently.

We are currently working on the implementation of *Semantic Zooming* [2]: As the display is zoomed in and nodes are expanded past a certain threshold their content changes. The larger the concept representations are scaled the more information items will be shown inside them. This information might contain the date, the concept has been added and by whom that has been done.

References

1. Grosso, W. E., Eriksson, H., Fergerson, R. W., Gennari, J., Tu, S. W., and Musen, M. A.: Knowledge modeling at the millenium (the design and evoulution of protege-2000). Proceedings of the 12th Knowledge Acquisition Workshop (KAW'99) (1999)
2. Card, S. K., Nation D.: Degree-of-Interest Trees — A Component of an Attention-Reactive User Interface International Conference on Advanced Visual Interfaces (2002)
3. Furnas, G. W.: The FISHEYE view — a new look at structured files. Readings in Information Visualization: Using Vision to Think (1981) 312–330
4. Munzner, T., Guimbretiere, F., Tasiran, S., Zhang, L., and Zhou, Y.: TreeJuxta-poser: Scalable Tree Comparision using Focus+Context with Guaranteed Visibility Proc. SIGGRAPH (2003), ACM Transactions on Graphics 22(3) 453–462
5. Smith, M. K., Welty, C., McGuiness, D. L.: OWL Web Ontology Language Guide. W3C Recommendation (2004), http://www.w3.org/TR/owl-guide/
6. Rada, R., Mili, H., Bicknell, E., Blettner, M.: Development and application of a metric on semantic nets. IEEE Transactions on Systems, Man and Cybernetics (1989) 17–30