POMELo: A PML Online Editor

Alvaro Graves

Tetherless World Constellation
Department of Cognitive Sciences
Rensselaer Polytechnic Institute
Troy, NY 12180
gravea3@rpi.edu

Abstract. This paper introduces POMELo, a simple, web-based PML (Proof Markup Language) editor. The objective of POMELo is to allow users to create, edit, validate and export provenance information in the form of PML documents. This application was developed with provenance novices in mind, making it usable in various settings, from educational to scientific. Since this is a web-based application, users do not need to install or run any software aside from a normal web browser, which simplifies its adoption and makes it more attractive for inexperienced users.

Keywords: PML, Web Application, Graphical Editor, Validation.

1 Introduction

There is an increasing acknowledgement of the importance of provenance information in different areas, from Semantic Web and eScience[15] to Ancient Art market[11]. Currently, there are powerful tools[4][12] that allow experts to edit and manage complex graphs describing provenance information. However, we see a need for simpler tools that can be used for educational purposes as well as being used by experts. Furthermore, most of the currently available applications related to provenance require installation of special software and libraries, making it difficult for novice users to create, visualize, edit and export provenance data.

In the past few years, application development for the web has become widespread. Nowadays, it is possible to write documents, read emails and edit spreadsheets using web applications. Web applications have multiple benefits: On one hand, the Web provides an open, interoperable platform where users can access resources from different locations. Moreover, these "web apps" usually do not require the user to install third party software, simplifying their adoption. These factors inspired the development of POMELo: A simple application that allow users to view and edit provenance information without the need for writing PML directly or installing software.

The rest of this paper is organized as follows: Section 2 provides a brief description of PML; Section 3 presents the related work; Section 4 discusses available

D.L. McGuinness, J.R. Michaelis, and L. Moreau (Eds.): IPAW 2010, LNCS 6378, pp. 91–97, 2010.

features in POMELo; Section 5 gives examples of situations where POMELo can be useful; Section 6 describe future work and potential extension for POMELo; Finally, Section 7 shows the conclusions.

2 Proof Markup Language (PML)

The Proof Markup Language[13] is an interlingua designed for sharing explanations and representing provenance knowledge over the Web. PML was designed modularly through three different ontologies.

2.1 PML-P: Provenance

The first component in PML is the provenance ontology, which aims to provide a set of primitives to represent provenance information. Thus it is possible to use PML-P for describing pieces of information, as well as their sources, languages, and formats in which they are expressed.

2.2 PML-J: Justification

The justification component provides primitives for explaining the steps taken in a process to draw a conclusion. PML-J allows users to express inference steps, assumptions, assertions, as well as sets of rules and engines used to obtain a specific conclusion.

2.3 PML-T: Trust

The Trust ontology allows users to indicate their trust in other users as well as their beliefs in assertions. This information can be expressed using numerical values (usually between 0 and 1). Is important to note that the method or algorithm for establishing a specific value is left to the user.

3 Related Work

Over the years, several initiatives have appeared to visualize and deal with provenance information. The Inference Web Browser (IWBrowser)[3] is a web-based tool for visualizing PML documents. This tool allows users to visualize PML proofs and show the results in english or in a graphical tree structure. *Provenance Explorer*[1] is another tool focused on visualizing provenance information expressed in OWL[14] (using an extension of ABC ontology[10]) and SWRL[8] and supports inference using Algernon[2] as an inference engine. *VisTrails* was developed by Freire et al.[7] and allows users to keep track of provenance information from workflows as well as data, and visualize it graphically. *VisTrails* was developed with scientists as the target audience and provides a rich infrastructure, including a workflow creator, a repository and a programming API. *Probe-It!*[4] is a Java-based application for visualizing PML documents obtained from inference engines. Users can see the justification behind a certain conclusion in a graphical way as well as explore data through different views (results, justification, provenance).

4 Features of POMELo

POMELo was based on "RDF Editor"[9] and the ARC2 library[16] for parsing RDF/XML. Implementation of POMELo was done using PHP, Javascript, AJAX and JQuery. A screenshot of POMELo can be seen in Figure 2. We now describe the most important features available in POMELo.

Graphical Visualization: One of the main features in POMELo is the use of graphical visualization to manage PML: Each entity is shown as a rectangle (green for resources and blank nodes, blue for literals) while the predicates are described as labeled edges. These nodes and edges can be rearranged for better display. Finally, hovering the mouse over a resource, it is possible to obtain its type (when available). An example can be seen in Figure 1.

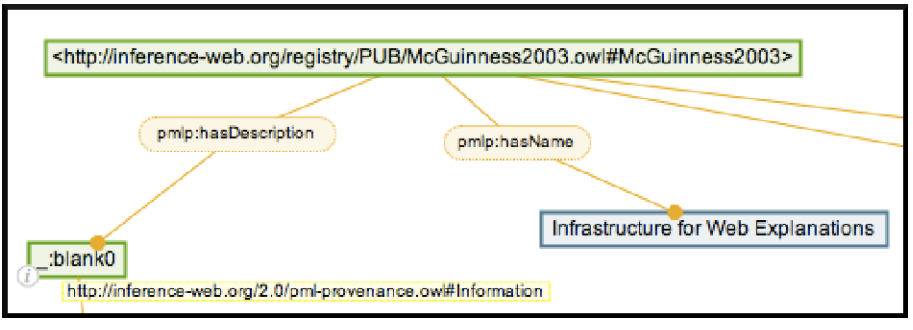

Fig. 1. Graphical representation of PML

Editing: Users can edit and curate provenance data. They can add Information Sources, Agents, Inference steps and other entities from PML based on four menus available on the left panel (one for each component of PML plus Literals). They can include them as resources or blank nodes. It is also possible to link different resources using predicates available from PML-P (PML's provenance ontology), PML-J (justification ontology) and PML-T (trust ontology).

Import and Export: POMELo allow users to import PML documents in the RDF/XML format. Users need to provide the URL where the PML document is located and POMELo will retrieve, parse and display it graphically. In the same way, it is possible to export the current PML model into RDF/XML.

Validation: POMELo can also validate PML graphs against the Inference Web PML Validation Service[5]. This service allows users to verify if their PML document is valid by making 10 evaluations, from loading of the data to issues related to typing of the resources. The validations may return warnings, errors or fatal errors. In turn, the user can make modifications to their model and perform follow-up validation. In this way, users can create valid PML documents using an integrated platform.

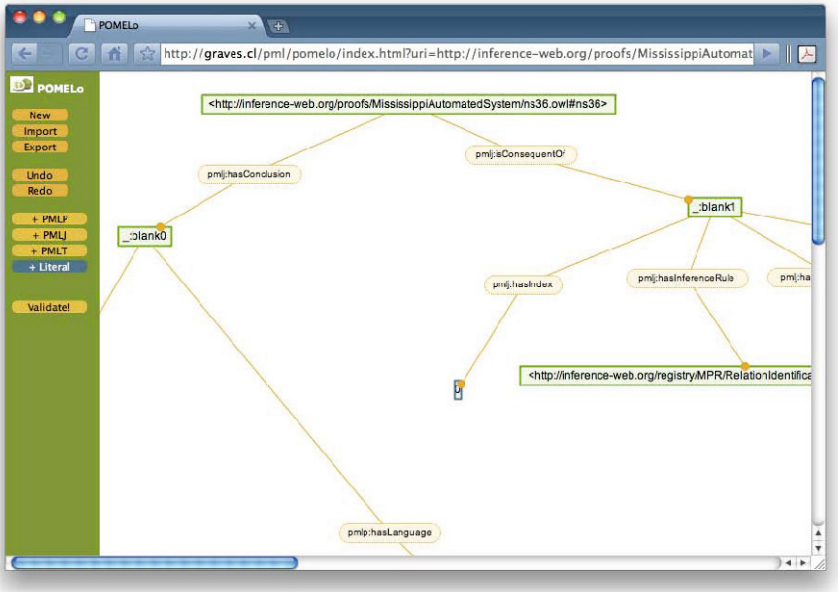

Fig. 2. Example of a PML document loaded into POMELo

5 Examples of Use

POMELo can be used for several purposes: First, it can help with curating small portions of a bigger provenance model, where a user may not be interested in the whole model but only a small part. He may load that portion into POMELo, study it, make edits and finally export it again.

It is also possible to use POMELo for educational purposes. People interested in studying provenance and in particular PML may not want to install additional software. To use POMELo, only a web browser and Internet connection is required.

5.1 Example 1: Visualizing PML

Professor *A* is giving a lecture on using provenance (in particular on how to express provenance information using PML). For that purpose, Professor *A* prepared several examples as documents in RDF/XML format available on the Web. In turn, students can use POMELo to visualize these documents.

5.2 Example 2: Editing and Validating PML

A scientist *B* receives a URI from scientist *C* containing provenance information about an experiment in PML. *B* opens the document using POMELo for visualizing it. After a few minutes studying it, she decides to add more information

to the PML graph (for example, adding a new antecedent to justify the default value in a sensor). In order to confirm that she has not made any mistake editing the provenance information, *B* validates the current PML graph using the "Validate!" button. POMELo indicates that the PML information is valid, but also report a warning because of not using a more specific type for the Information node. Figure 3 shows an example of a successful validation in POMELo. Finally she exports her work in RDF/XML format.

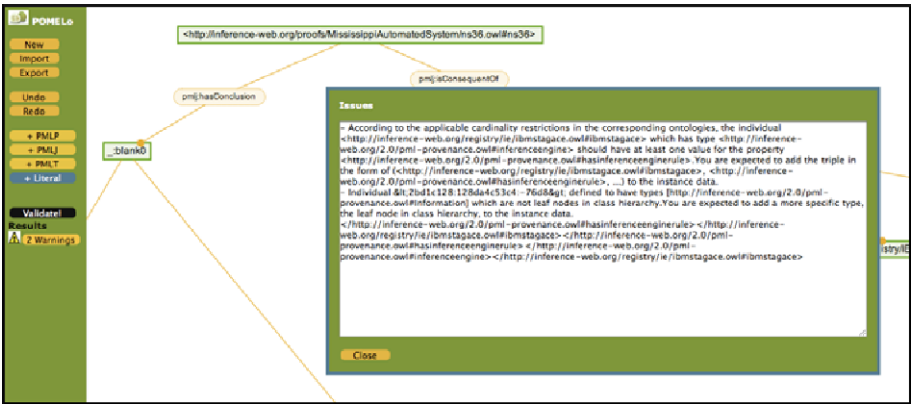

Fig. 3. POMELo validates a PML document using the PML validation service. The dialog box indicates warnings related to this document.

6 Future Work

There are several improvements that can be made to POMELo: First, the default layout was developed based on the assumption that trees will be the most likely structure found in PML documents (i.e., a documents describes the provenance of one thing). While this is useful, we have found certain situations where PML documents refers to several entities, making the tree layout inappropriate. To solve this, we are already working in integrating graphviz[6] as a "layout manager" that can coordinate how POMELo should render the nodes depending on the graph structure of the PML document. Another interesting feature would be to cluster specific parts of the PML document and visualize them as one node. Due the nature of PML, the use of blank nodes and other different levels intermediate types of nodes may overload the screen (e.g., when describing a list of authors). Thus, it would be desirable to allow users to group several recurrent sub-structures in a PML document into one node, alleviating from an overload of nodes on the screen.

Another direction we have considered is to allow users to collaboratively visualize and edit a PML document: This could make easier for students to understand provenance and in particular to use PML. Also, the ability for exporting

in different formats and publishing documents from POMELo, would add value from the educational perspective as well as from the professional point of view. Finally, the possibility of uploading PML documents directly from POMELo to a central repository would make the whole process of creating, editing and publishing provenance information more fluid and easy for users.

7 Conclusions

In this paper, we presented POMELo, a web application that allows users to work with provenance information using PML. The aim of this application is to allow users to visualize, modify and export provenance information in PML. One of the goals of POMELo is to serve as a tool that can be adopted by expert as well as novice users. Since it is based on the Web, POMELo does not need special libraries, but only a web browser and an Internet connection. We show the functionalities of POMELo, including its integration with PML Validator that allow users to create, verify and correct their model in a simple, unified interface. A demo of POMELo can be seen at `http://graves.cl/pml/pomelo`.

Acknowledgement

Thanks to James Michaelis for helping in the revision of this paper.

References

1. Cheung, K., Hunter, J.: Provenance Explorer–Customized Provenance Views Using Semantic Inferencing. In: Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L.M. (eds.) ISWC 2006. LNCS, vol. 4273, pp. 215–227. Springer, Heidelberg (2006)
2. Crawford, J.M., Kuipers, B.J.: Algernona tractable system for knowledge-representation. ACM SIGART Bulletin 2(3), 44 (1991)
3. Da Silva, P., Sutcliffe, G., Chang, C., Ding, L., Del Rio, N., McGuinness, D.: Presenting TSTP proofs with inference web tools. In: CEUR Workshop Proceedings, Citeseer, vol. 373, pp. 81–93.
4. Del Rio, N., da Silva, P.: Probe-it! visualization support for provenance. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Paragios, N., Tanveer, S.-M., Ju, T., Liu, Z., Coquillart, S., Cruz-Neira, C., Müller, T., Malzbender, T. (eds.) ISVC 2007, Part II. LNCS, vol. 4842, pp. 732–741. Springer, Heidelberg (2007)
5. Ding, L., Tao, J., McGuinness, D.: An initial investigation on evaluating semantic web instance data (2008)
6. Ellson, J., Gansner, E., Koutsofios, L., North, S., Woodhull, G.: Graphviz, open source graph drawing tools. In: Graph Drawing, pp. 594–597. Springer, Heidelberg (2001)
7. Freire, J., Silva, C., Callahan, S., Santos, E., Scheidegger, C., Vo, H.: Managing rapidly-evolving scientific workflows. In: Provenance and Annotation of Data, pp. 10–18

8. Horrocks, I., Patel-Schneider, P., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL: A semantic web rule language combining OWL and RuleML. W3C Member submission 21 (2004)
9. Krsulovic, E.: Rdf editor (2009), http://code.google.com/p/tesis-e/
10. Lagoze, C., Hunter, J.: The ABC ontology and model. Journal of Digital Information 2(2), 77 (2001)
11. Levine, J.: et al. The importance of provenance documentation in the market for ancient art and artifacts: The future of the market depend on documenting the past. DePaul J. Art Tech. & Intell. Prop. L. 19, 219–421 (2009)
12. McGuinness, D., da Silva, P., Chang, C.: IWBase: Provenance metadata infrastructure for explaining and trusting answers from the web. Technical report, Citeseer
13. McGuinness, D., Ding, L., da Silva, P., Chang, C.: PML 2: A modular explanation interlingua. In: Proceedings of AAAI, vol. 7 (2007)
14. McGuinness, D., Van Harmelen, F., et al.: OWL web ontology language overview. W3C recommendation, February10 (2004)
15. Miles, S., Groth, P., Deelman, E., Vahi, K., Mehta, G., Moreau, L.: Provenance: The bridge between experiments and data. Computing in Science and Engineering 10(3), 38 (2008)
16. Nowack, B.: Arc, rdf classes for php (2006), http://arc.semsol.org/