

NaturalOWL: Generating Texts from OWL Ontologies in Protégé and in Second Life

George Karakatsiotis and Dimitrios Galanis and Gerasimos Lampouras and Ion Androutsopoulos¹

Abstract. NaturalOWL is an open-source natural language generation engine written in Java. It produces descriptions of individuals (e.g., items for sale, museum exhibits) and classes (e.g., types of exhibits) in English and Greek from OWL DL ontologies. The ontologies must have been annotated in RDF with linguistic and user modeling resources. We demonstrate a plug-in for Protégé that can be used to produce these resources and to generate texts by invoking NaturalOWL. We also demonstrate how NaturalOWL can be used by robotic avatars in Second Life to describe the exhibits of virtual museums. NaturalOWL demonstrates the benefits of Natural Language Generation (NLG) on the Semantic Web. Organizations that need to publish information about objects, such as exhibits or products, can publish OWL ontologies instead of texts. NLG engines, embedded in browsers or Web servers, can then render the ontologies in multiple natural languages, whereas computer programs may access the ontologies directly.

1 INTRODUCTION

NaturalOWL [1] is an open-source natural language generation engine written in Java.² It produces descriptions of individuals (e.g., items for sale, museum exhibits) and classes (e.g., types of exhibits) in English and Greek from OWL DL ontologies. The ontologies must have been annotated in RDF with linguistic and user modeling resources. We demonstrate a plug-in for Protégé that can be used to produce these resources and to generate texts by invoking NaturalOWL. We also demonstrate NaturalOWL being used by a robotic guide in Second Life to describe the exhibits of a virtual museum.

NaturalOWL demonstrates the benefits of adopting Natural Language Generation (NLG) on the Semantic Web. Organizations that need to publish information about objects, such as exhibits or products, can publish OWL ontologies instead of texts. NLG engines, embedded in browsers or Web servers, can then render the ontologies in natural language, whereas computer programs may access the ontologies, in effect logical statements, directly. This way, information becomes fully accessible to both humans and computers, which is a major goal of the Semantic Web. Standardizing the linguistic and user modeling annotations of OWL ontologies would allow competing natural language generation engines to be used, in the same way that different Web browsers can be used with the same HTML source. A further advantage of NLG is that texts can be generated in multiple

natural languages from the same ontology; their content and expressions can also be tailored to the type of the user (e.g., child vs. adult) and to the user's interaction history (e.g., by avoiding repetitions of information, or by comparing to previous objects, as shown in Figures 1 and 2).



Figure 1. Generating texts in Second Life



Figure 2. Generating texts in Second Life

¹ Department of Informatics, Athens University of Economics and Business, Greece; see <http://pages.cs.aueb.gr/nlp/>. The third author is also affiliated with the Digital Curation Unit, Research Centre "Athena", Greece; see <http://www.dcu.gr/>. This was a system demonstration, 18th European Conference on Artificial Intelligence, Patras, Greece, 2008.

² NaturalOWL is available from <http://pages.cs.aueb.gr/nlp/>.

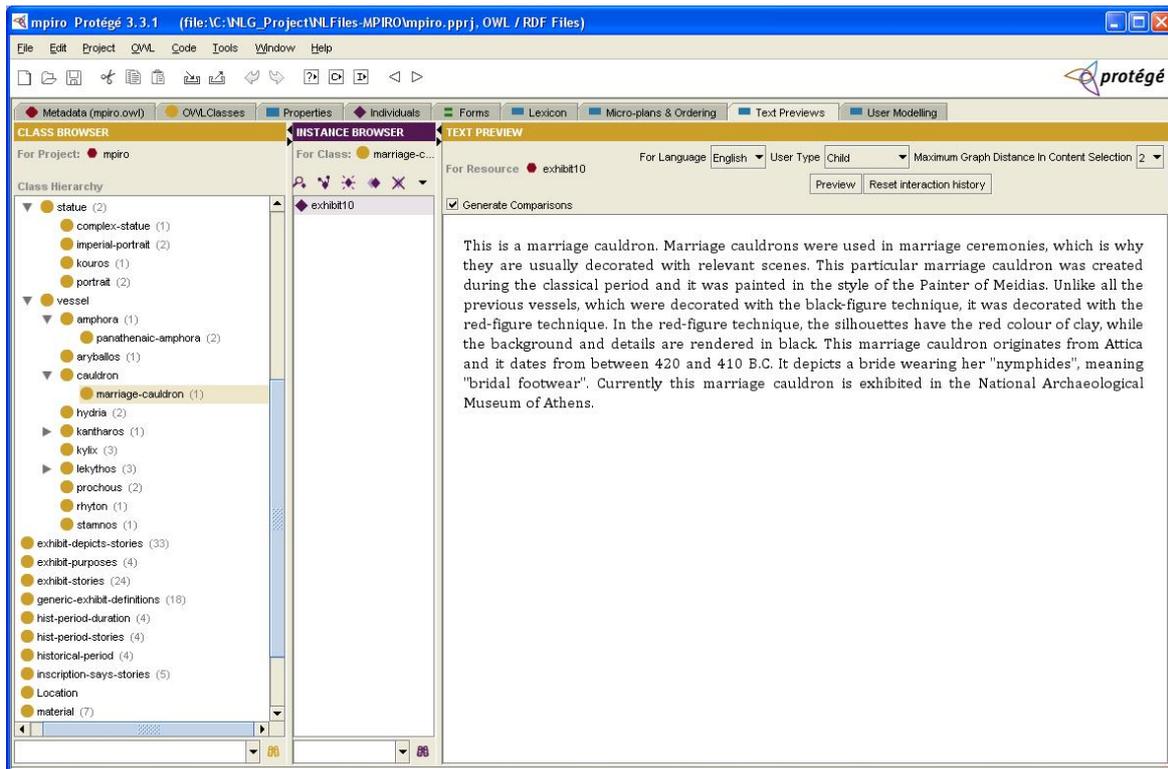


Figure 3. Generating texts in Protégé

2 USING NATURALOWL IN PROTEGE

In the first part of the demonstration, we present a plug-in for Protégé, which can be used to annotate OWL DL ontologies with the linguistic and user modeling resources required by NaturalOWL; the plug-in is based on M-PIRO's authoring tool [4, 2].³ The resources include: a domain-dependent lexicon, which associates classes and individuals of the ontology with nouns and proper names of the target natural languages; micro-plans, in effect patterns, which show how to express the properties of the ontology as sentences of the target languages; a partial order of properties, which allows the system to order the resulting sentences as a coherent text; interest scores, indicating how interesting the various logical facts of the ontology are to each type of user; parameters that control, for example, the desired length of the generated texts. The plug-in allows the user of Protégé to annotate any ontology with these resources. Textual descriptions of the ontology's classes and individuals can then be generated, as illustrated in Figure 3.

3 USING NATURALOWL IN SECOND LIFE

In the second part of the demonstration, we present a robotic guide that uses NaturalOWL to describe the exhibits of a virtual museum in Second Life [5]. The guide follows the visitors, and whenever they touch an exhibit, it produces a description of the exhibit, as illustrated in Figures 1 and 2.⁴

³ M-PIRO's authoring tool, now called ELEON [3], can also be used instead of Protégé; consult <http://www.iit.demokritos.gr/skel/>.

⁴ A video showing the robotic guide in action is available from <http://www.vimeo.com/801099>.

ACKNOWLEDGEMENTS

NaturalOWL was developed in the Greek project XENIOS, where it was embedded in a mobile robotic guide [6]. XENIOS was funded by the European Union and the Greek General Secretariat of Research and Technology; see <http://www.ics.forth.gr/xenios/>. NaturalOWL is now being extended in the INDIGO project, where it is combined with spoken dialogue management and facilities to mark up the generated texts with gestures, emotions, and facial expressions, again in mobile robotic guides. INDIGO is funded by the European Union; see <http://www.ics.forth.gr/indigo/>.

REFERENCES

- [1] I. Androutsopoulos, S. Kallonis, and V. Karkaletsis, 'Exploiting OWL ontologies in the multilingual generation of object descriptions', in *Proceedings of the 10th European Workshop on Natural Language Generation*, pp. 150–155, Aberdeen, UK, (2005).
- [2] I. Androutsopoulos, J. Oberlander, and V. Karkaletsis, 'Source authoring for multilingual generation of personalised object descriptions', *Natural Language Engineering*, **13**(3), 191–233, (2007).
- [3] D. Bilidas, M. Theologou, and V. Karkaletsis, 'Enriching OWL ontologies with linguistic and user-related annotations: the ELEON system', in *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence*, Patras, Greece, (2007).
- [4] A. Isard, J. Oberlander, I. Androutsopoulos, and C. Matheson, 'Speaking the users' languages', *IEEE Intelligent Systems*, **18**(1), 40–45, (2003).
- [5] J. Oberlander, G. Karakatsiotis, A. Isard, and I. Androutsopoulos, 'Building an adaptive museum gallery in Second Life', in *Proceedings of Museums and the Web*, Montreal, Quebec, Canada, (2008).
- [6] D. Vogiatzis, D. Galanis, V. Karkaletsis, I. Androutsopoulos, and C.D. Spyropoulos, 'A conversant robotic guide to art collections', in *Proceedings of the 2nd Workshop on Language Technology for Cultural Heritage Data, Language Resources and Evaluation Conference*, Marrakech, Morocco, (2008).