Semantic Networks to Support Learning

Philippe A. Martin
Eurecom, research center in communications systems, Sophia-Antipolis, France
and adjunct researcher of the School of I.C.T. at Griffith Uni, Australia

Abstract. This article illustrates Conceptual Graph networks representing the content of courses to help students understand, relate, compare, memorize and retrieve many of their concepts. It shows that the ontology of WebKB-2 and its FL notation could be exploited by lecturers to create normalized representations in a scalable way and relatively quick way. They also permit the students to complement these representations, thus providing lecturers with ways to test the students' understanding and analytical skills. Very strong mechanisms supporting semantic checking, cooperation support and normalization need to be implemented for the approach to be successful. Current semantic wikis and knowledge servers (WebKB-2 included) are far from fulfilling such constraints.

Keywords: knowledge representation/sharing/retrieval/learning/evaluation

1 Introduction

Most of Semantic Learning Web projects [1, 2] and all Learning Object related standards or practices [3, 4, 5, 6] exploit *simple* meta-data : concept types/instances or mere keywords are manually or automatically associated to learning materials or students' user profiles. In more fine-grained approaches, semantic networks are used for representing the content of the course and/or the knowledge learnt by the students. Some of these networks are fully formal and very difficult to create, e.g., those of the Halo project [7] intended to solve some chemistry test questions automatically. Other semantic networks are mostly informal (and manually or automatically created), as in projects using Concept Maps [8, 9], or their ISO version, Topic Maps [10, 11].

In [12], the authors detail some problems with these networks and with the different kinds of approach currently used for indexing, representing, organizing, sharing and retrieving information (e.g., document retrieval approaches, fully formal or mostly informal approaches, approaches based on the mostly independently creation of (semi-)formal resources). First, they are insufficient for precision-oriented information retrieval and learning support. Second, they cannot be made much more precise, efficient or scalable, since they do not permit to create a normalized, formal, expressive, easy-to update network of concepts/statements semantically related to other concepts/statements (for example, by relations of specialization, argumentation, instrumentation, correction, authorship, spatial/temporal location and modality). However, the authors of [12] also provide solutions to the above cited problems. This is done using the KB server WebKB-2 [13] as an example (a KB server permits Web users to update one or several shared knowledge bases, and/or allow them to exchange knowledge between the KBs of the users). First, a normalized semantic network can be cooperatively and incrementally created by Web users. Second, protocols and replication mechanisms permit to remove *implicit* redundancies and inconsistencies within such a network as well as between networks (thus, it does not matter which knowledge base a user updates or queries first: the advantages of distribution and centralization are combined and there is only one "virtual" network).

Although already mostly implemented and having many advantages in the medium and long term, the proposed knowledge sharing approach suffers from two problems common to all precision-oriented knowledge acquisition/retrieval approaches, that is, approaches where the semantic network has to be (semi-)formal and displayed to the users: (i) people need to learn how to read such networks or knowledge representations, and (ii) entering knowledge representations requires much more intellectual rigor than writing informal sentences. The unwillingness of most people to learn new notations (e.g., musical notations, mathematical notations and programming languages) is well known. Furthermore, most people have not heard about knowledge representation languages nor about the usefulness of learning one. Yet, the author believes that his approach has some future with (at first) researchers, teachers and students since (i) the need of using very small learning objects is now well recognized by the e-learning research community [3, 5], (ii) the economy of time and resources brought by the use of truly re-usable learning objects will be understood by more and more e-learning/university teachers and administrators, (iii) more and more teachers are involved in e-learning, (iv) it is part of the roles of teachers and researchers to (re-)present knowledge in explicit and detailed ways, (v) the approach permits a better evaluation of the knowledge and analytic skill of the students than less precision-oriented approaches, and (vi) providing the semantic organization of the content of teaching materials (instead or in addition to these materials) help students find, compare and memorize the information scattered in these materials. This last point was recognized by many of the students after they had learnt how to read the semantic networks prepared for them.

During the period of his e-learning fellowship [14], the author represented the content of three courses given by three different lecturers at Griffith Uni. Section 2 shows an extracts of one semantic network, with some additions made by the students. Indeed, as part of his/her homework, each student was asked to add at least twenty relations to the networks. The semantic content of these additions were evaluated by the author (did they make sense? were they interesting?). The conclusion draws some lessons of this experiment. Thus, this article does not repeat but truly complements a previous article [13]: indeed, this new article does not describe the approach or features of the WebKB server, nor does it compare them to other approach or features, but it presents the result of their use in a teaching context.

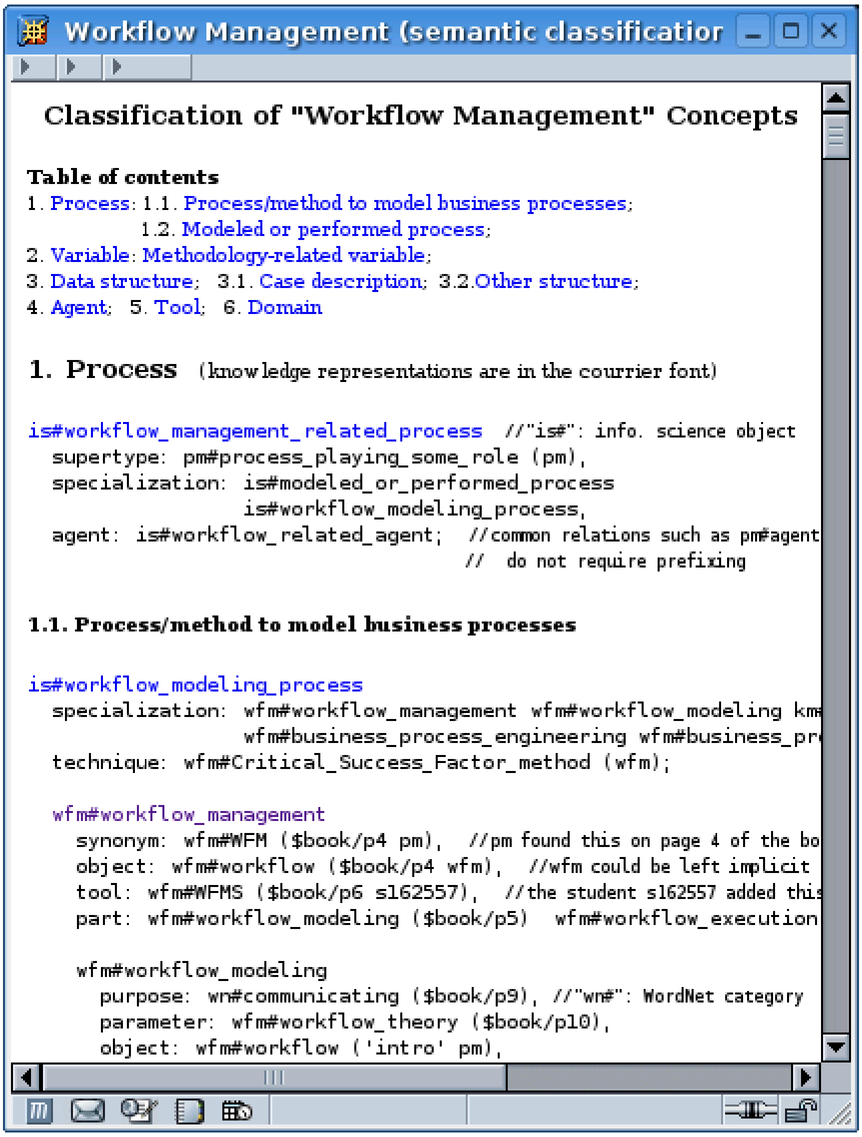
2 Presentation of a Semantic Network

The input files containing the initial knowledge representations for the three courses are accessible from http://www.webkb.org/kb/it/. These input files were loaded into (i.e., executed by) WebKB-2 and hence their formal objects (concepts or statements) became part of the unique global semantic network that can be queried, browsed and complemented by any Web user via WebKB-2 (http://www.webkb.org). The students were given the URL of WebKB-2 and the URLs of the input files for their courses. As shown in Figure 1, within each file the formal representations are included within sections and indented. This indentation most often reflects the specialization relations existing between the represented objects. The FL (For-Links) notation [15] used in these files is the most concise possible formal notation that is as expressive as RDF+OWL. It is similar to N3 but has a more regular structure. FL was derived by the author from CGLF. It permits to pack much more information into a certain amount of space than other notations, especially graphic notations, and hence reduces the needs for scrolling or browsing. This permits people to see many relations between the formal objects, and hence better compare and understand these objects. In the figures, no cardinalities are explicitly associated to the relations between the objects. Thus, each statement in these figures follow the generic schema "CONCEPT1 RELATION1: CONCEPT2 CONCEPT3, RELATION2: CONCEPT4, ...;". Such a statement should be read: "any CONCEPT1 may have for RELATION1 one or many CONCEPT2, and may have for RELATION1 one or many CONCEPT3, and may have for RELATION2 one or many CONCEPT4, ...". Some comments within the figures explain how the creators of each object (here, relation or concept/relation type) are made unambiguous; please note the example of the relation added by the student "s162557".

Very few relation types were required for representing the three courses in a precise and normalized way. Most of these types were: subtype, instance, specialization, part (physical_part or subtask), technique, tool, definition, annotation, use, purpose, rationale, role, origin, example, advantage, disadvantage, argument, objection, requirement, agent, object, input, output, parameter, attribute, characteristic, support and url. (This list is ordered topically, not by frequency of occurrence). This list is small compared to all the basic relations that can be found in top-level ontologies or that would potentially be needed if long and diverse natural language texts had to be represented. This shows that the above list includes many of the most important (i.e., primitive and common) relation types.

The large ontology of WebKB-2 [16] is a transformation of WordNet into a genuine lexical ontology and its extension with many top-level ontologies. Using FL and this ontology, it was not too difficult to categorize all the important concepts and represent all the important facts (relationships between concepts) contained in the source learning materials of the three courses. This representation by extension of a large shared ontology eases knowledge retrieval, re-use and understanding.

Although using a KB server such as WebKB-2 is unavoidable to allow the representation, querying and cooperative updating of a large semantic network, the author found that a structured document editor (SDE; for example Amaya - the W3C Web browser - or any other XML editor) would have been a useful intermediary or complementary tool: (i) the manual creation of the representations would have been much easier if the source documents had been organized via a SDE instead of Word or Powerpoint, (ii) the manual exploitation of the input files would have been simpler with a SDE since for example some sections could have been temporarily hidden, and (iii) despite its predefined document schemas and semantic un-awareness, a SDE could also guide beginners in the creation of files and representations similar to the FL representations illustrated below.



Workflow Management (semantic classificatior

Classification of "Workflow Management" Concepts

Table of contents

1. Process (knowledge representations are in the courrier font)

```
is#workflow_management_related_process  //"is#": info. science object
  supertype: pm#process_playing_some_role (pm),
  specialization: is#modeled_or_performed_process
                  is#workflow_modeling_process,
  agent: is#workflow_related_agent;  //common relations such as pm#agent
                                     //  do not require prefixing
```

1.1. Process/method to model business processes

```
is#workflow_modeling_process
  specialization: wfm#workflow_management wfm#workflow_modeling km
                  wfm#business_process_engineering wfm#business_pr
  technique: wfm#Critical_Success_Factor_method (wfm);

  wfm#workflow_management
    synonym: wfm#WFM ($book/p4 pm),  //pm found this on page 4 of the bo
    object: wfm#workflow ($book/p4 wfm),  //wfm could be left implicit
    tool: wfm#WFMS ($book/p6 s162557),  //the student s162557 added thi
    part: wfm#workflow_modeling ($book/p5)  wfm#workflow_execution

    wfm#workflow_modeling
      purpose: wn#communicating ($book/p9), //"wn#": WordNet category
      parameter: wfm#workflow_theory ($book/p10),
      object: wfm#workflow ('intro' pm),
```

Fig. 1. Extract from a file representing statements from a book in Workflow Management (here referred to by the variable $book).

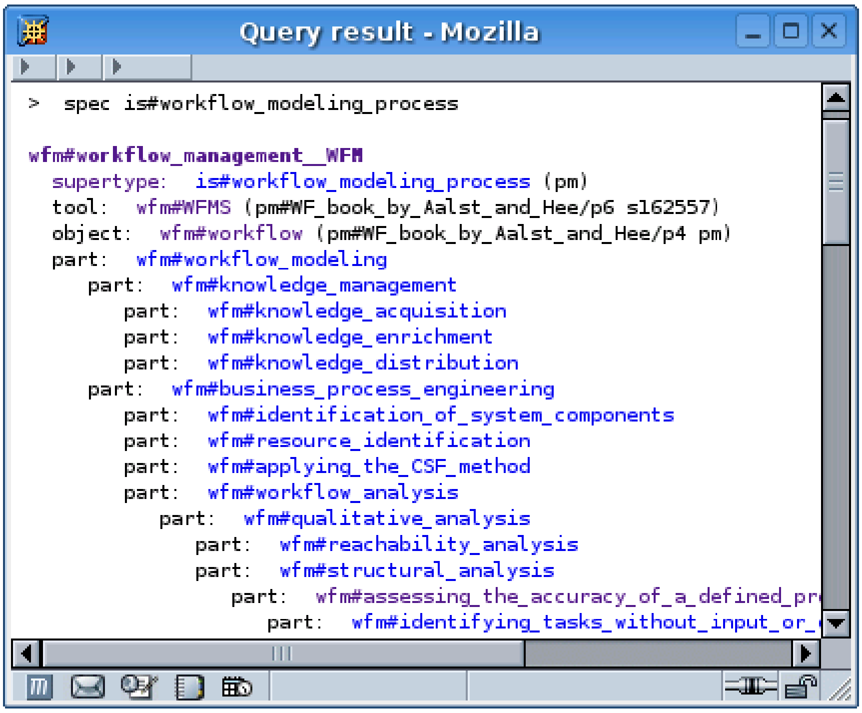


Fig. 2. Command to display the ***spec***ializations of a type, followed by its first result (wfm#workflow_management, along with some of its related objects; here, an informal format looking like FL is used for the display).

The approach includes those of argumentation-based collaboration tools (e.g., [17]) but also allows (i) more expressiveness when required (e.g., relations on relations), (ii) the exploitation of the recording of votes and object creators for filtering or evaluation purposes, and (iii) a more normalized representation of knowledge [12].

WebKB-2 was used to create the semantic networks. Unfortunately, the students of the WFM and Multimedia courses had to use a classic wiki instead of WebKB-2 for entering *new* statements because (i) the implementation of the graphic interfaces and parsing of some new features of FL was not at a sufficiently advanced stage at that time, and (ii) no time was allowed for training the students to use FL in a correct way (nor for giving them any real introduction to "knowledge representation"; the students were only shown how to read the representations and to avoid some ontological non-senses). The outcomes of the use of a wiki was that, except for some rare students, most of the additions by the students contained lexical errors (for example, typos or badly formed identifiers), syntactic problems (this is understandable), ontological problems (meaningless relationships, redundancies, inconsistencies) and indentation problems. In [18], a detailed list of errors made by the students of the WFM course in their first "semantically structured learning journals" was given.

The syntax used for displaying the semantic network was a big issue for the students, although curiously one of them thought that "most of the notations were intuitive or well known". Controlled languages are not a solution since, like natural languages, they cannot display information in a sufficiently structured way; [15] presents Formalized English (a formal controlled language derived from the Conceptual Graph Linear Form) and compares it to several other notations. The use of FL with a good indentation leads to a structured display but which is apparently not explicit enough for beginners. Understanding the *structure and scope* of the described relations was the students' main problem. Although more space-consuming than FL, an interface based on structured elements (e.g., XML elements or embedded HTML tables) with specific background colors - and menus associated to each element - seems necessary for permitting beginners to immediately understand the structure and scope of the described relations - and complement them more easily. However, precise knowledge representations necessarily include elements such as cardinalities, quantifiers, sets or contexts, and therefore require the use of a special notation to express them and their scopes (structured elements are of no help for displaying such additional intertwined scopes). Using special notations for presenting information often has a lot of advantages. This is illustrated by the above survey synthesis itself since (i) a large table would have been impractical to display, and (ii) a list of tables (or worse, individual surveys) would have not permitted people to easily compare and understand the information.

3 Conclusion

WebKB-2 has various input-output formats and many presentation options but, as previously noted, an additional format exploiting structured document elements seems necessary. The full implementation of the interfaces and mechanisms permitting the users to cross-evaluate each other's statements also need to be completed urgently. Finally, it is essential to complement the cooperation protocols [12, 13] with much stronger mechanisms to detect inputs that are either semantically incorrect or potentially redundant/contradictory with already existing statements. On the other hand, enhancing the search and browsing methods is not urgent and no user model is required: displaying large amounts of well structured information as query/navigation results appears sufficient to let the users quickly find the information they want.

The temporary use of a wiki confirmed how inadequate wikis are for (i) letting people collaboratively build structured knowledge, and (ii) evaluating them doing so. Indeed, the ease-of-use of wikis does not compensate for their lack of semantic structure, semantic checking and cooperation protocols. Current semantic wikis are only timid advances toward the support of semantic structures/checking. Apart from OntoWiki [19] which includes the features of a frame-based system, most semantic wikis offer very little support for fine-grained systematic knowledge modelling. For example, within a page, Semantic MediaWiki [20] only allows to set semantic relations from/to the object represented by the page, and only in a rather hidden way within an unstructured text. No current semantic wiki has genuine cooperation protocols.

The goal of the author is the scalable cooperative building and cross-evaluation of structured knowledge. To achieve it he also aims for the efficient retrieval of this knowledge, its deep-learning and the evaluation of this deep-learning. The author has collected or designed and implemented the minimal components that a KB server should have to support that goal, for example, a large general ontology, expressive and concise notations, normalization techniques and cooperation protocols. The author does not believe that the complexity inherent to that goal can be hidden to the knowledge providers or readers. Instead of going for other goals permitting that complexity to be hidden, or instead of aiming a KB server at trained knowledge engineers only, the author has made the rare choice of trying to progressively bring people to use it. As explained in the introduction, these people will first have to be researchers, lecturers and students and, preferably, in knowledge engineering related domains. If the approach is successful, it will be progressively adopted by other communities.

The first tests of the author had to be done on courses unrelated to knowledge engineering. They confirm the urgency of implementing more features. Unlike data management tools, knowledge base management tools cannot come in small independent tools. Indeed, KB management tools must be full-featured to be adopted. Limiting their number of features to reduce their complexity is not a winning strategy [21], however tempting and popular it may be. This is especially true to achieve the constraint of "scalability", that is, to reduce future extension problems and keep guiding users as the knowledge base grows.

Acknowledgments. This article was made possible by an "e-learning" grant given by Griffith Uni. to the author and he would like to thank the senior lecturers that allowed this experiment to be conducted on their courses, and especially Dr Jun Jo.

4 References

1. Stutt, A., Motta, E.: Semantic Learning Webs. Journal of Interactive Media in Education, Special Issue on the Educational Semantic Web, No. 10 (2004)
2. Devedzic, V.: Education and the Semantic Web. International Journal of Artificial Intelligence in Education, 14, 39--6 (2004)
3. Downes, S.: Learning Objects: Resources For Distance Education Worldwide. International Review of Research in Open and Distance Learning, Vol. 2, No.1 (2001)
4. IEEE LTSC: IEEE Learning Technology Standards Committee Glossary. IEEE P1484.3 GLOSSARY WORKING GROUP, draft standard (2001)
5. Hodgins, W.: Out of the past and into the future: Standards for technology enhanced learning. In: Ehlers, U., Pawlowski, J. (eds.). Handbook on Quality and Standardisation in E-Learning, pp. 309--327. Springer Berlin (2006)
6. Tane, J., Schmitz, C., Stumme, G., Staab, S., Studer, R.: The Courseware Watchdog: an Ontology-based tool for finding and organizing learning material. In: Fachtagung Mobiles Lernen und Forschen, Kassel (2003)
7. Friedland, N.S, Allen, P., Mathews, G., Witbrock, M., Baxter, D., Curtis, J., Shepard, B., Miraglia, P., Angele, J., Staab, S., Moench, E., Opperman, H., Wenke, D., Israel, D., Chaudhri, V., Porter, B., Barker, K., Fan, J., Chaw, S.Y., Yeh, P., Tecuci D., Clark, P.: Project Halo: Towards a Digital Aristotle. AI Magazine, 25(4), 29--48 (2004)
8. Novak J.D., Gowin D.B.: Learning How to Learn. Cambridge University Press (1984)

9. Novak. J.D.: Reflections on a Half Century of Thinking in Science Education and Research: Implications from a Twelve-year Longitudinal Study of Children's Learning. Canadian Journal of Science, Mathematics and Technology Education, 4(1), 23--41 (2004)
10. Scott, B., Johnson, Z.: Using Topic Maps as Part of Learning Design - Some History and a Case Study. In: Recent Research Developments in Learning Technologies, Badajoz, Spain, pp. 605--609. FORMATEX (2005)
11. Leung, J.: Concept Maps On Various Topics. Created in 2005 at http://www.fed.cuhk.edu.hk/~johnson/misconceptions/concept_map/concept_maps.html
12. Martin, P., Eboueya, M.: For the ultimate accessibility and re-usability. Handbook of Research on Learning Design and Learning Objects: Issues, Applications and Technologies, IGI Global, May 2008. http://www.webkb.org/doc/papers/LO_handbook07/
13. Martin, P.: Large-scale cooperatively-built heterogeneous KBs. In: ICCS 2001, LNAI 2120, pp. 231--244. Springer (2001)
14. Martin, P.: Griffith E-Learning Fellowship Report. Created in 2006 at http://www.webkb.org/doc/papers/GEL06/
15. Martin, P.: Knowledge representation in CGLF, CGIF, KIF, Frame-CG and Formalized-English. ICCS'02, LNAI 2393, pp. 77--91. Springer-Verlag (2002)
16. Martin, P.: Correction and Extension of WordNet 1.7. In: ICCS 2003, LNAI 2746, pp. 160-173. Springer Verlag (2003).
17. Uren, V., Buckingham Shum, S., Bachler, M., Li, G.: Sensemaking Tools for Understanding Research Literatures: Design, Implementation and User Evaluation. International Journal of Human Computer Studies, No 64 - 5, 420--445 (2006)
18. Martin, P.: Common errors in the learning journals. Created in 2006 at http://www.phmartin.info/wf/journalErrors.html
19. Auer, S., Dietzold, S., Riechert, T.: OntoWiki - A Tool for Social, Semantic Collaboration. In: ISWC 2006, LNCS 4273, pp. 736--749. Springer (2006).
20. Krötzsch, M., Vrandecic, D., Völkel, M., Haller, H., Studer, R.: Semantic Wikipedia. Journal of Web Semantics, 5, 251--261, September (2007)
21. Shipman, F.M., Marshall, C.C.: Formality considered harmful: experiences, emerging themes, and directions on the use of formal representations in interactive systems. Computer Supported Cooperative Work, No. 8, pp. 333--352 (1999).