Levels of organisation in ontology verbalisation

Sandra Williams, Allan Third and Richard Power
Computing Department, The Open University
Walton Hall, Milton Keynes, MK7 6AA, U.K.
`s.h.williams@open.ac.uk a.third@open.ac.uk r.power@open.ac.uk`

Abstract

The SWAT TOOLS ontology verbaliser generates a hierarchically organised hypertext designed for easy comprehension and navigation. The document structure, inspired by encyclopedias and glossaries, is organised at a number of levels. At the top level, a heading is generated for every concept in the ontology; at the next level, each entry is subdivided into logically-based headings like 'Definition' and 'Examples'; at the next, sentences are aggregated when they have parts in common; at the lowest level, phrases are hyperlinked to concept headings. One consequence of this organisation is that some statements are *repeated* because they are relevant to more than one entry; this means that the text is longer than one in which statements are simply listed. This trade-off between organisation and brevity is investigated in a user study.

1 Introduction

Since OWL (Web Ontology Language) became the standard language for the semantic web in 2004,[1] several research groups have developed systems, known as 'verbalisers', for generating Controlled English from OWL ontologies (Kaljurand and Fuchs, 2007; Dolbear et al., 2007; Schwitter and Tilbrook, 2004; Funk et al., 2007). The resulting texts may contain linguistic errors, especially when the lexicon is inferred from identifier names (as in SWAT TOOLS) rather than handcrafted, but

they are still easier to understand than formal languages (Kuhn, 2010). We describe here a generic verbaliser[2] (applicable to any OWL-DL ontology with English identifiers/labels) which delivers its output (e.g., figure 1) in the form of an organised hypertext,[3] akin to an online encyclopedia or glossary, and investigate whether this extra organisation makes the text easier to understand and navigate.
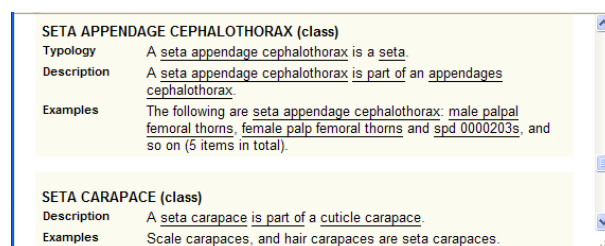

Figure 1: A section of SWAT TOOLS encyclopedia-style output, length 7746 words or 25 A4 pages, generated from an ontology about spider anatomy.

Elsewhere (Stevens et al., 2010; Stevens et al., 2011) we reported two evaluations with bioinformatics staff in which the glossary-style verbalisation was judged effective, for instance in detecting errors in the knowledge; however, these studies were not designed to test the value of organisation, through comparison with an unorganised list of statements. It might seem obvious that organisation will help: indeed, teachers of reading comprehension are required to ensure that

---

[1]http://www.w3.org/2004/01/sws-pressrelease

[2]SWAT TOOLS available at http://swat.open.ac.uk/tools/

[3]The hypertext in figure 1 was generated from spider.owl from owl.cs.manchester.ac.uk/repository/

their students are aware of the use of textual features such as headings and subheadings in locating information (Steeds, 2001); and it comes as no surprise that reading strategies differ between a text structured as an encyclopedia and one organised as, say, a poem (Hanauer, 1998). However, organisation actually has one potential disadvantage, in that statements may be relevant in more than one context, and must thus be repeated. We describe here a study which checks this point by pitting organisation against brevity for a task requiring accurate retrieval of information from a text.

2 Verbalising statements

The input to a verbaliser is a set of OWL statements describing individuals, classes or properties; the simplest output therefore consists of a set of sentences, one per statement, as illustrated in figure 2, where the sentence 'A cribellar spigot is part of a cribellum' has been generated from the following statement:[4]

```
subClassOf(class(#SPD_0000276),
  objectSomeValuesFrom(
    objectProperty(#part_of),
    class(#SPD_0000115)))
```

The basic verbalisation method in SWAT TOOLS has been described in detail elsewhere (Stevens et al., 2011). Briefly, the OWL/XML input is transcoded to Prolog,[5] using the format illustrated in the example just given; then a lexicon for realising atomic terms (individuals, classes or properties) is inferred from their identifier names or labels; finally, a sentence is generated from each statement using a Definite Clause Grammar (Clocksin and Mellish, 1987) covering almost all of OWL-DL, using wording influenced by earlier work on controlled languages (Schwitter et al., 2008; Kaljurand and Fuchs, 2007; Dolbear et al., 2007).

Sentences are ordered according to the alphabetical order of their underlying OWL statements: i.e., sentences generated from `ClassAssertion` statements will come before those generated from `SubClassOf` statements.

3 Document structuring

The highest levels of organisation, illustrated in figure 1, are headings and subheadings. Subheadings are inspired mainly by Berzlanovich et al.'s (2008) 'information oriented' discourse labels (name, definition, description, etc.) from their analysis of the discourse structure of encyclopedia articles; and also by Aristotle's genus-differentia descriptions.[6]

Lower levels of organisation were also influenced by Berzlanovich et al. (2008), whose investigation of lower-level lexical cohesion in encyclopedia entries highlighted the high incidence of hypernymic lexical cohesion.

3.1 Headings

The top level of organisation is an alphabetical series of headings corresponding to atomic terms in the ontology (i.e., individuals, classes, or object/data properties), taken directly from the lexicon that the system infers from the ontology's identifier names or labels. Where singular and plural forms have been inferred, the singular is used, as illustrated by 'SETA CEPHALOTHORAX (class)' in figure 1.

An OWL statement is selected for inclusion under a heading if the class, property or individual that the heading refers to occurs as a top-level argument in the statement.[7] Inevitably, sentences that apply to more than one group are duplicated, e.g., 'a seta appendage cephalothorax is a seta' is added to entries for both SETA APPENDAGE CEPHALOTHORAX and SETA.

3.2 Subheadings

The second level of organisation is a set of subheadings. Within each entry, statements are organised into sub-groups according to their logical type. Subheadings are always generated in a fixed order (Definition, Taxonomy, Description, Distinctions, Examples) similar to that found in encyclopedia entries (Berzlanovich et al., 2008). For classes, `EquivalentClasses` statements in which the atomic class is the first

---

[4]The 'non-semantic' identifiers #SPD_0000276 and #SPD_0000276 are annotated with English labels elsewhere in the ontology.

[5]The verbaliser is implemented in SWI Prolog

---

[6]en.wikipedia.org/wiki/Genus_differentia_definition

[7]Theoretically, this could mean that some statements are omitted altogether because their top-level arguments are non-atomic, but in practice such statements are almost never found (Power and Third, 2010).

argument occur under the definition subheading, the taxonomy is the superclass (from an OWL `SubClassOf` statement), descriptive statements correspond to the OWL functor `SubClassOf`, distinctions to `DisjointClasses`, and examples to the individuals belonging to the class. For individuals the class is given first (from an OWL `ClassAssertion` statement), followed by descriptions typically corresponding to `ObjectPropertyAssertion`. For properties, the descriptive statements specify the domain and range, and features such as functionality and transitivity, and examples are provided by statements about individuals or classes in which the property is used.

3.3 Aggregating and truncating

A third level of organisation occurs when statements with identical structures and one identical argument are aggregated; see Williams and Power (2010) for more details. For some ontologies, this process can lead to very long lists of subclasses or individuals, so under the 'Examples' subheading where these occur we truncate them to a predefined maximum length and add the phrase 'and so on (N items in total)'. Figure 1 shows an example of aggregation and truncation in the sentence 'The following are seta appendage cephalothorax: male palpal femoral thorns, female palp femoral thorns and spd 0000203s, and so on (5 items in total)'. An obvious refinement would be to add a facility to view the entire list, if desired.

3.4 Hypertext links

The final and lowest level of organisation occurs when hyperlinks are introduced for each phrase corresponding to a class, individual or property; these link to the headings of their entries.

3.5 Related systems

To our knowledge, SWAT TOOLS takes document structuring further than other domain-*independent* ontology verbalisers. We are aware of only one other domain-independent system that attempts document structuring, ACE (Kaljurand and Fuchs, 2007). ACE lists statements under class, individual and property headings; and it inserts hyperlinks; but it has no intermediate levels of organisation.

Regarding domain-*dependent* systems, most of them aggregate statements and generate referring expressions (Bontcheva and Wilks, 2004; Dongilli, 2008; Galanis and Androutsopoulos, 2007; Hielkema, 2009; Liang et al., 2011). Only one attempts further discourse structuring: Laing et al.'s system for verbalising medical ontologies organises text according to rhetorical structure.

4 Evaluation

The evaluation study reported here focusses on the following question: Does the organisation just described help people to understand and navigate a text in spite of its longer length? This is addressed through a navigation task in which people were asked to locate information in either an organised text or an unorganised one and then give a judgement on how difficult the information was to find. The study design is between-subjects in two independent groups. Participants were 57 members of the ACL special interest groups SIGGEN[8] and SIGdial[9].

4.1 Materials and method

The texts were generated from an ontology about spider anatomy.[3] One group saw the encyclopedia-styled version illustrated in figure 1, henceforth the 'organised text'; the other saw the same information as a list of sentences[10] as shown in figure 2 ('*un*organised text'). At 7746 words (25 A4-sized pages), the organised text is much longer than the unorganised one (4803 words, 9-pages) mainly because of duplicated information and headings (as explained in section 3). To render the unorganised text's appearance as similar as possible to the organised one, spaces were introduced every fourth line with blocks of text placed on a taupe-coloured background identical to that of the entries in the organised text.

The same five navigation and information location tasks (table 1) were used for both groups. The survey was administered via SurveyMonkey[11] in which each navigation question was followed

---

[8]www.siggen.org
[9]www.sigdial.org
[10]Sentences were ordered alphabetically by their underlying OWL statements as described in section 2.
[11]www.surveymonkey.com

A seta appendage cephalothorax is a seta.

A seta appendage cephalothorax is part of an appendages cephalothorax.
A spd 0000275 is part of a median trachea.
A cribellar spigot is part of a cribellum.

A PLS spigot unspecific is a spigot.
A PLS spigot unspecific is part of a PLS spinning field.
A seta anal tubercle is part of an anal tubercle.

A PMS PC base is a spigot base.
A PMS PC base is part of a PMS PC.
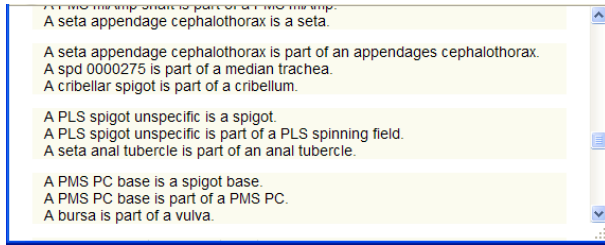A bursa is part of a vulva.

Figure 2: A section of SWAT TOOLS sentence-list-style output, 4,803 words or 9 A4 pages, the 'unorganised text' of our study generated from the spider anatomy ontology.

by a judgement 'How difficult was it to find the information?' on a 5-point scale ('Very Easy' to 'Very Hard').

Regarding search for information to answer the questions, both texts were viewed on-line and of course could be navigated with the usual 'Find' menu items, CTRL-F key sequence, and so on. To determine whether textual features such as headings, subheadings and hyperlinks had been used, subjects with the organised text were asked whether the following features had helped them to search for information: (i) heading, (ii) typology subheading, (iii) description subheading, (iv) examples subheading, (v) alphabetical ordering of entries,(vi) hyperlinks within entries, and (vii) totals for number of items in lists (section 3.3). Subjects given the unorganised text answered instead seven questions about techniques used for navigation, e.g., 'Did you use scrolling?'.

No.	Questions
Q1	What is a tarsus?
Q2	Name 3 kinds of spigot shaft.
Q3	What is a palp?
Q4	Name 2 kinds of silk cable.
Q5	How many kinds of seta appendage cephalothorax are there in total?

Table 1: Questions for the navigation tasks

Lastly, we had chosen the spider anatomy ontology because we hoped that the subject would be unfamiliar to participants causing them to consult the text (as we instructed) rather than using their own general knowledge to answer the questions. The final question in the survey asked about familiarity with spider anatomy.

4.2 Results

Table 2 shows that despite the drop in performance on question 5 (and question 1 for the unorganised text group), both groups were relatively successful in locating information. However, difficulty judgements differed significantly between groups (see table 3), with the group using the organised text judging the tasks much easier. This preference was confirmed by a non-parametric independent samples Mann-Whitney U test over all judgements ($p < 0.0001$). Results for questions about usage of specific organisational features (answered only by the group that viewed the organised text) are given in table 4. None of the participants claimed to be expert in spider anatomy.

n	Text	Q1	Q2	Q3	Q4	Q5
28	Unorganised	20	27	26	27	19
29	Organised	28	27	26	27	25

Table 2: Results for numbers of correct answers (n = number of participants per group)

Q	n	Text	VE	Easy	Neith	Hard	VH
Q1	28	Unorg	2	6	7	11	2
	29	Org	10	12	3	4	0
Q2	28	Unorg	3	13	4	6	2
	29	Org	11	10	6	2	0
Q3	28	Unorg	1	11	6	5	5
	29	Org	11	13	4	1	0
Q4	28	Unorg	6	17	2	2	1
	29	Org	12	13	3	1	0
Q5	28	Unorg	0	1	12	10	4
	29	Org	10	11	4	3	1

Table 3: Results for difficulty judgements (Q = question number, n = number of participants per group, Unorg = unorganised text, Org = organised text, VE = Very easy, Neith = neither hard or easy, VH = Very hard).

5 Discussion

In our earlier user studies (Stevens et al., 2010; Stevens et al., 2011), experts in bioinformatics assessed technical descriptions corresponding to glossary entries, with statements linked by aggregation but not grouped by logical subheadings. The consensus was that these were understandable, and useful to developers as a means of checking accuracy. However, various criticisms were made,

Organisational Feature	Used by
Headings	17
Typology subheadings	17
Description subheadings	20
Examples subheadings	22
Alphabetical ordering	20
Hyperlinks	19
Totals for items in truncated lists	21

Table 4: Results for usage of organisational features (number of people who used them out of 29 participants who viewed the organised text).

the main theme being that natural English should be priveleged over fidelity to OWL semantics.

The SWAT TOOLS system evaluated here incorporates some stylistic changes proposed in the earlier study, and retains the aggregation feature which combines several statements into a single sentence (thus potentially reducing fidelity to the underlying OWL); it also adds grouping by subheadings. None of this organisation is directly encoded in an OWL ontology: it represents rather a further move towards making the verbalisation more natural and humanlike.

From our study comparing organised and unorganised texts, two main points emerge: first, we find no evidence that people viewing the organised text perform a navigation task more accurately; second, people viewing the organised texts found the task easier. One explanation for these findings would be that people do whatever is necessary to achieve a desired level of performance, so that when provided with superior tools they achieve roughly the same result but with less effort.[12]

The drop in performance by the unorganised text group on question 1 might have been due to unfamiliarity with a sentence-list type of text (all participants answered question 1 first since questions were always presented in the same order). Improvements on later questions could have been the result of a learning effect with this group. The near-perfect performance of the organised text group on the first questions demonstrates the benefit of viewing a familiar genre. A drop in performance by the unorganised text group on question 5, 'How many kinds of seta appendage cephalothorax

are there in total?' was expected since it is a harder question that requires a search of the entire unorganised text whilst simultaneously counting instances. It is not clear why four people in the organised text group failed to get the correct answer to question 5 since it merely required them to understand the text '5 items in total' under 'Examples' (see figure 1).

Regarding the analysis of different organisational features, the overall response was that all these features were considered useful by a majority of users, although none of them stood out as particularly important.

6   Conclusion

We assume that most users prefer an ontology verbalisation that is worded and organised like a naturally occurring text of the appropriate genre — i.e., an encyclopedia or technical glossary. One possible objection is that such a text provides a loose rendering of OWL semantics, introducing organisational principles that are not present in the original code; however, as evidenced by the earlier studies we have cited, this attitude is not taken even by OWL specialists. A second possible objection is that the organised text is necessarily longer than a bare list of sentences; this point is tested in the study reported here, which suggests that organisation makes the texts easier to use, with no loss of performance. In future work we intend to look more closely at how the texts are used in retrieval tasks, and to obtain accurate measures of time differences.

References

I. Berzlanovich, M. Egg, and G. Redeker. 2008. Coherence structure and lexical cohesion in expository

---

[12]In this case responses for organised texts should be faster, a point we intend to check in future work.

and persuasive texts. In *A. Benz, P. Kuhnlein and M. Stede (Eds.), Proceedings of the Workshop Constraints in Discourse III, Potsdam, Germany*, pages 19–26.

K. Bontcheva and Y. Wilks. 2004. Automatic report generation from ontologies: the MIAKT approach. In *Nineth International Conference on Applications of Natural Language to Information Systems (NLDB'2004)*, pages 214–225, Manchester, UK.

F. Clocksin and Chris Mellish. 1987. *Programming in Prolog*. Springer-Verlag, 3 edition.

Cathy Dolbear, Glen Hart, Katalin Kovacs, John Goodwin, and Sheng Zhou. 2007. The RABBIT Language: Description, Syntax and Conversion to OWL. Technical Report Technical Report, Ordnance Survey Research.

Paolo Dongilli. 2008. Natural language rendering of a conjunctive query. Technical Report Knowledge Representation Meets Databases (KRDB) Research Centre Technical Report: KRDB08-3, Free University of Bozen-Bolzano.

Adam Funk, Valentin Tablan, Kalina Bontcheva, Hamish Cunningham, Brian Davis, and Siegfried Handschuh. 2007. Clone: Controlled language for ontology editing. In *K. Aberer et al. (Eds.) Proceedings of ISWC/ASWC 2007*, pages 142–155. Springer-Verlag Berlin Heidelberg 2007.

Dimitrios Galanis and Ion Androutsopoulos. 2007. Generating multilingual descriptions from linguistically annotated owl ontologies: the naturalowl system. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 143–146, Saarbrücken, Germany, June. DFKI GmbH. Document D-07-01.

David Hanauer. 1998. The genre-specific hypothesis of reading: Reading poetry and encyclopedic items. *Poetics*, 26:63–80.

Feikje Hielkema. 2009. *Using Natural Language Generation to Provide Access to Semantic Metadata*. Ph.D. thesis, University of Aberdeen.

Kaarel Kaljurand and Norbert E. Fuchs. 2007. Verbalizing owl in attempto controlled english. In *Proceedings of Third International Workshop on OWL: Experiences and Directions, Innsbruck, Austria (6th–7th June 2007)*, volume 258.

Tobias Kuhn. 2010. An evaluation framework for controlled natural languages. In Norbert E. Fuchs, editor, *Proceedings of the Workshop on Controlled Natural Language (CNL 2009)*, volume 5972 of *Lecture Notes in Computer Science*, pages 1–20, Berlin / Heidelberg, Germany. Springer.

Shao Fen Liang, Donia Scott, Robert Stevens, and Alan Rector. 2011. Unlocking medical ontologies for non-ontology experts. In *Proceedings of BioNLP 2011 Workshop*, pages 174–181, Portland, Oregon, USA, June. Association for Computational Linguistics.

Richard Power and Allan Third. 2010. Expressing OWL axioms by English sentences: dubious in theory, feasible in practice. In *Proceedings of the 23rd International Conference on Computational Linguistics*.

R. Schwitter and M. Tilbrook. 2004. Controlled natural language meets the semantic web. In *Proceedings of the Australasian Language Technology Workshop*, pages 55–62, Macquarie University.

Rolf Schwitter, Kaarel Kaljurand, Anne Cregan, Catherine Dolbear, and Glen Hart. 2008. A comparison of three controlled natural languages for owl 1.1. In *OWL: Experiences and Directions (OWLED)*, page online.

Andrew Steeds. 2001. Adult literacy core curriculum including spoken communication. Technical report, Cambridge Training and Development Ltd. on behalf of The Basic Skills Agency. ISBN 1-85990-127-1.

Robert Stevens, James Malone, Sandra Williams, and Richard Power. 2010. Automating class definitions from owl to english. In *Proceedings of Bio-Ontologies 2010: Semantic Applications in Life Sciences, SIG at 18th Annual International conference on Intelligent Systems for Molecular Biology (ISMB 2010)*.

R. Stevens, J. Malone, S. Williams, R. Power, and A. Third. 2011. Automating generation of textual class definitions from owl to english. *Journal of Biomedical Semantics*, 2(Suppl 2):S5.

Sandra Williams and Richard Power. 2010. Grouping axioms for more coherent ontology descriptions. In *6th International Natural Language Generation Conference (INLG 2010)*.