e-Health – For Continuity of Care
C. Lovis et al. (Eds.)

doi:10.3233/978-1-61499-432-9-1048

Does SNOMED CT post-coordination scale?

Daniel KARLSSON[a,1], Mikael NYSTRÖM[a], and Ronald CORNET[a, b]
[a] *Department of Biomedical Engineering, Linköping University, Sweden*
[b] *Department of Medical Informatics, Academic Medical Center – University of Amsterdam, The Netherlands*

Abstract. SNOMED CT is a compositional terminology. Construction of post-coordinated expressions allows users to specify new meaning by referencing existing SNOMED CT concepts. The use of post-coordinated expressions in information systems requires special software, a reasoner, to give the exact relations between post-coordinated expressions and existing SNOMED CT content. Thus, the performance characteristics of reasoners are important for implementation of post-coordination in information systems. This study aims to test how reasoners perform when a large number of post-coordinated expressions are added to SNOMED CT. The time needed to classify an ontology consisting of SNOMED CT plus an increasing number of post-coordinated expressions is measured. The best performing reasoner in this test classifies SNOMED CT plus 1 million post-coordinated expressions in 42 seconds. The time to classify grows a little less than quadratic as the size of the ontology increases. In conclusion, classification time is not a problem using current reasoners and current SNOMED CT releases even if a large number of post-coordinated expressions are added.

Keywords. Post-coordination, SNOMED CT

Introduction

SNOMED CT is a compositional reference terminology which allows the construction of expressions to represent meaning beyond what has been given a unique concept identifier by IHTSDO, the owner of SNOMED CT, or by a national release center or by a local extension [1]. This composition is often called post-coordination and is considered necessary in order to achieve coverage in practical tests [2-4].

Organizations with large-scale implementations of SNOMED CT in clinical practice like Kaiser Permanente and Hospital Italiano de Buenos Aires have seen the need to extend SNOMED CT in order to cover clinical use cases and to meet interface terminology needs [5, 6]. Both organizations have created extensions to SNOMED CT in the magnitude of tens of thousands of concepts.

Enumerating all sensible combinations of SNOMED CT concepts is not possible due to the combinatorial explosion [e.g. 7]. For example, pain may be qualified by severity (7 severities), pain character (149), body site (25851 sites), and course (31). This results in some 800 million possible, although maybe not always sensible, ways to express different kinds of pain. Thus, while post-coordination seems to be necessary

[1] Corresponding Author.

for coverage of the clinical domain the large amount of possible combinations poses a potential problem for implementers of SNOMED CT, and the means to manage post-coordination in a scalable manner have to be made available to implementers.

Expression repositories may provide a solution [1, 7]. Expression repositories are systems which cache results of some of the time-consuming tasks involved in the use of post-coordinated expressions in information systems. As the systems only provide a cache, expression repositories require no manual curation. An example of such a time-consuming task is classification of post-coordinated expressions, i.e. putting the expressions into the correct positions in the subclass (aka. IsA) hierarchy by identifying the correct and only the correct parents and children. Expression repositories rely on software called reasoners doing the work of classification. The scalability of the expression repository is thus dependent on the scalability of the reasoner software.

Previous studies have benchmarked and compared reasoners for classification of large-scale ontologies [8, 9], and when new reasoners are presented they are usually benchmarked towards existing reasoners [e.g. 10, 11]. Apart from classification performance, reasoners differ in the description logic (DL) being supported, soundness and completeness, available APIs etc. The previous studies we found often estimate performance for classifying the international release of SNOMED CT as well as some other well-known ontologies. Previous studies have also examined the statistical average-case complexity of reasoning as the size of the ontology grows, for example Ian Horrocks' Masters Thesis [12], but we are unaware of any studies where the performance of reasoners for SNOMED CT with massive post-coordination is assessed.

The aim of this study is to assess the feasibility of the use of post-coordination in practice by testing the scalability of using current reasoners to classify SNOMED CT together with a repeated addition of post-coordinated content in an as realistic as possible manner.

1. Methods

To assess how SNOMED CT classification time changes when post-coordinated expressions are added, a set of post-coordinated expressions was created as follows. For each role group of a clinical finding with finding site and associated morphology, the direct descendants of the specified finding site and associated morphology were retrieved, and a cross product of the retrieved values was generated. For example, acute interstitial nephritis is defined with one role group containing associated morphology = acute inflammation, which has 14 direct descendants, and finding site = structure of interstitial tissue of kidney, which has 3 direct descendants. This results in 3x14=42 refined post-coordinated expressions. Applying this method to all clinical findings resulted in over 7 million expressions, which were then randomly ordered. The generated expressions were regarded as fully defined, and no test was performed to determine whether the generated post-coordinated expression equalled any pre-coordinated SNOMED CT concept.

A Java™ program was developed to gradually add the randomly generated post-coordinated expressions and measure time needed for classification by various reasoners. Two tests were performed. Firstly, 10,000 expressions were added at a time after which classification time was measured. This was repeated until either 1,000,000 post-coordinated expressions had been added or a total time of 4 hours had passed. To reduce the random effects in the test, for example due to other processes running on the

system, the full test was run 3 times and the minimum time from the three test runs was used for analysis. The second test aimed at further exploring the potential for large-scale post-coordination. By adding 1,000,000 post-coordinated expressions seven times or until the reasoner or underlying system would fail in the given system environment.

The time span measured was the time from just before starting classification until the time just after finishing classification.

Classification time measurements for the tested reasoners were plotted on a log-log-scale diagram with ontology size on the x-axis and classification time on the y-axis. To determine the order of the polynomial describing classification time, a least-square-method linear estimation was applied to log(time) versus log(size) data to estimate the slope of the logarithmic curve.

The Java program was executed on a Linux system running a 3.9.6 64-bit kernel and OpenJDK 1.7.0_45. The system had two Intel Xeon processors with 2x4 processor cores and 96 GB internal memory. A default of 25.9 GB was assigned to the test process. Linear estimation and graph plotting was done with the R statistical computing environment.

In these tests the 2013-07-31 International Release of SNOMED CT was used. The release was translated to OWL RDF/XML using the OWL-transformation Perl script provided with the International Release. The OWL ontology contained 297,315 class expression axioms (229,247 subclass axioms, 68068 equivalent-class axioms). The resulting file was loaded into the Java application using OWL API version 3.4.9.

In order to automate creation of post-coordinated expression test cases, the SNOMED CT Release Format 2 files were imported into a MySQL 5.5 database.

Reasoners that provide an OWL API were selected for the experiment: ELK [10], HermiT [13], and Snorocket [11] (see Table 1). The reasoners allow reasoning with different (or slightly different) subsets of DL, use different reasoning algorithms, and represent different stages of the development of reasoning tools. The ELK and Snorocket reasoners allow parallelization of the reasoning task. ELK supports incremental reasoning but this feature was not used when comparing to other reasoners. The primary purpose of this study was not to compare reasoners performance but to assess and describe how SNOMED CT classification behaves with increasingly sized ontologies, using different OWL reasoners for increased test validity.

Table 1. Reasoners used in the test

Reasoner	Version	DL
ELK 0.4.1 [10]	Consequence reasoning	EL+
Snorocket 2.1.1 [11]	Completion rules	EL+
HermIT 1.3.8 [13]	Hyper-tableau	SHOIQ+

2. Results

The time for classification of the SNOMED CT release without additional post-coordinated content ranged from approximately 3.6 seconds for the more recent ELK reasoner to over 30 minutes (1,900 s) for the HermiT reasoner (see Table 2). When adding 100,000 expressions classification time increased 67 % for ELK and 180 % for Snorocket. Only the ELK reasoner managed to complete the test within 4 hours taking 42 seconds to classify SNOMED CT plus 1,000,000 post-coordinated expressions (see

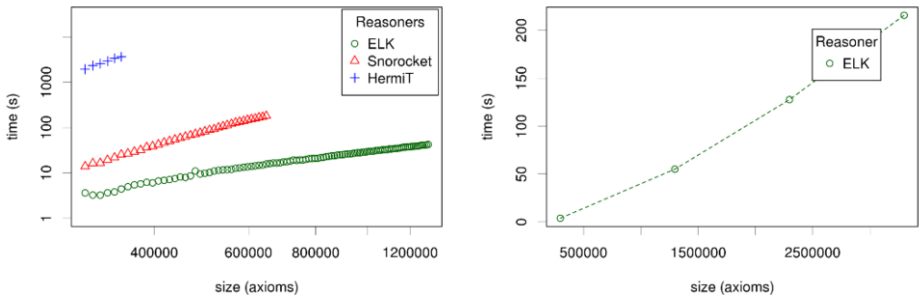

Figure 1. Classification time vs. size of ontology, log axes (left), large additions (right).

Figure 1). The Snorocket reasoner managed to classify about 650,000 concepts in about 3 minutes (180 s) before reaching the 4 hour test limit.

The slope of the curve when a logarithmic transformation had been applied ranged between 1.6 and 4.0 (see Figure 1 and Table 2, Standard Error in parentheses). Thus in the best case classification time grows less than quadratic with the size of the ontology.

Table 2. Classification time results (seconds)

Reasoner	SNOMED CT only	S. CT+100,000 post-coordinated	Number of iterations/4h	Order of growth (S.E.)
ELK 0.4.1 [10]	3.6	6.0	101	1.6 (0.019)
Snorocket 2.1.1 [11]	14	39	36	3.3 (0.026)
HermIT 1.3.8 [13]	1900	-	6	4.0 (0.19)

In the second test, when adding even larger amounts of post-coordinated expressions to the ontology, only the best performing reasoned, ELK, was used. The application managed to classify 3.3 million concepts, i.e. the SNOMED CT release plus 3 million post-coordinated expressions, in about 3 and a half minutes (220 s) before reaching a state were almost all time was used for Java garbage collection.

3. Discussion

As post-coordination is necessary for practical implementation of SNOMED CT [2-6] the resulting terminology will be significantly larger than the 300,000 concepts in the international release. Although the exact size of the ontology of "SNOMED CT-in-action" is yet to be found through implementation experiences, the ability of systems to deal with the increased load of a larger ontology becomes crucial to successful implementation. The results from this study indicate that use of post-coordinated expressions in information systems does not pose a technical problem for at least some current OWL reasoners even if a large amount of post-coordinated content is added.

With efficient reasoners, expression repository systems can further improve effectiveness through various sorts of caching including syntactic comparison of expressions, transitive closure tables, branch numbering, etc. [1].

All post-coordinated expressions used in this test had the same form, i.e. clinical findings refined by finding site and associated morphology. The post-coordinated expressions in the test were created using only refinement of existing content and not by adding additional relationships from the Concept Model or by creating intersections of existing content. In practice, post-coordinated expressions will most probably be much more varied in both form and content. Whether it is harder or easier to classify

this kind of expressions in relation to the more varied ones most probably used in practice is currently unknown and is a subject for further investigation.

One way of making a more representative set of post-coordinated expressions would be to extrapolate from existing sets of post-coordinated content rather than existing pre-coordinated SNOMED CT concepts.

There are increasingly strong requests for making the DL language underlying SNOMED CT more expressive, including concrete domains, number restrictions, negation, and universal restriction. This will reduce the number of reasoners being able to fulfill the task. This is particularly problematic as reasoners optimized for EL reasoning such as ELK and Snorocket seem to be the top performers in this and other studies. Before accepting these requests for new language facilities the implications for, among other things, post-coordination must be carefully considered.

Does SNOMED CT post-coordination scale? Yes, it does. Post-coordination with SNOMED CT is possible even when a large amount of expressions is added. Caveats exist though. These results does no longer hold if new constructs are added to the chosen dialect of DL.

References

[1] IHTSDO, "Technical Implementation Guide, January 2014." [Online]. Available: http://www.ihtsdo.org/tig. [Accessed: 7-Feb-2014].

[2] G. Wade and S. T. Rosenbloom, "Experiences mapping a legacy interface terminology to SNOMED CT," *BMC Med Inform Decis Mak*, vol. 8 Suppl 1, p. S3, 2008.

[3] N. F. de Keizer, F. Bakhshi-Raiez, E. de Jonge, and R. Cornet, "Post-coordination in practice: evaluating compositional terminological system-based registration of ICU reasons for admission," *Int J Med Inform*, vol. 77, no. 12, pp. 828–835, Dec. 2008.

[4] J. R. Campbell, J. Xu, and K. W. Fung, "Can SNOMED CT fulfill the vision of a compositional terminology? Analyzing the use case for problem list," *AMIA Annu Symp Proc*, vol. 2011, pp. 181–188, 2011.

[5] R. H. Dolin, J. E. Mattison, S. Cohn, et al., "Kaiser Permanente's Convergent Medical Terminology," *Stud Health Technol Inform*, vol. 107, no. Pt 1, pp. 346–350, 2004.

[6] A. L. Osornio, D. Luna, M. L. Gambarte, A. Gomez, G. Reynoso, and F. G. B. de Quirós, "Creation of a local interface terminology to SNOMED CT," *Stud Health Technol Inform*, vol. 129, no. Pt 1, pp. 765–769, 2007.

[7] D. Karlsson, M. Nyström, and B. Kron, "An integrated Expression Repository EHR system," in *SNOMED CT Implementation Showcase 2012*, Stockholm, Sweden, 2012.

[8] K. Dentler, R. Cornet, A. ten Teije, and N. de Keizer, "Comparison of reasoners for large ontologies in the OWL 2 EL profile," *Semantic Web*, vol. 2, no. 2, pp. 71–87, Jan. 2011.

[9] R. S. Gonçalves, S. Bail, E. Jimenez-Ruiz, N. Matentzoglu, B. Parsia, B. Glimm, and Y. Kazakov, "OWL Reasoner Evaluation (ORE) Workshop 2013 Results: Short Report," in *2nd OWL Reasoner Evaluation Workshop (ORE 2013)*, p. 1.

[10] Y. Kazakov, M. Krötzsch, and F. Simancík, "Concurrent classification of EL ontologies," in *Proceedings of the 10th international conference on The semantic web - Volume Part I*, Berlin, Heidelberg, 2011, pp. 305–320.

[11] M. Lawley and C. Bousquet, "Fast classification in Protégé: Snorocket as an OWL 2 EL reasoner," in *Australasian Ontology Workshop 2010 (AOW 2010): Advances in Ontologies*, Adelaide, Australia, vol. 122, pp. 45–50.

[12] I. Horrocks, "A comparison of two terminological knowledge representation systems," Master's thesis, University of Manchester, 1995.

[13] B. Motik, R. Shearer, and I. Horrocks, "Hypertableau Reasoning for Description Logics," *J. Artif. Int. Res.*, vol. 36, no. 1, pp. 165-228, Sep. 2009.