Exploring Complexity in Health: An Interdisciplinary Systems Approach
A. Hoerbst et al. (Eds.)

doi:10.3233/978-1-61499-678-1-431

Representing ICD-11 JLMMS Using IHTSDO Representation Formalisms

Marzouk MAMOU[a,1], Alan RECTOR[d], Stefan SCHULZ[e], James CAMPBELL[c], Harold SOLBRIG[f] and Jean-Marie RODRIGUES[a,b]
[a] *INSERM U1142, LIMICS, Paris, France*
[b] *Department of Public Health and Medical Informatics, Univ.Jean Monnet of Saint Etienne, France*
[c] *Department of Internal Medicine University of Nebraska Medical Center Omaha USA*
[d] *University of Manchester, United Kingdom*
[e] *IMI, Medical University of Graz, Austria*
[f] *Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, 55905, USA*

Abstract. It is investigated whether the content of the Joint Linearization for Mortality and Morbidity Statistics of the 11th ICD revision can be semantically represented by formalisms acting on the clinical terminology SNOMED CT, *viz.* the IHTSDO Compositional Grammar (CG) and the Expression Constraint Language (ECL). Whereas CG provides a composition syntax for building coordinated SNOMED CT expressions, ECL provides a powerful query mechanism. Both formalisms can be leveraged to guarantee inter-operation between an ontology-based terminology like SNOMED CT and a statistical classification like ICD, characterized by single hierarchies and exhaustive, mutually exclusive classes. We test the feasibility of the method on the circulatory chapter of ICD-11 JLMMS.

Keywords. Ontology Alignment; ICD-11; SNOMED CT

1. Introduction

The semantic interoperability between the most widely used health care terminologies is becoming more critical when documentation specialists, epidemiologists, health care administrators, and health services researchers are identifying more and more use cases where SNOMED ST (SCT) is used in parallel to the WHO International Classification of Diseases, generally ICD-10. ICD the leading standard for mortality and morbidity statistics is also used in other contexts like healthcare documentation and billing within national modifications and extensions.

ICD-11 JLMMS [1] is the final output of the eleventh ICD revision, intended to replace ICD-10, currently used at a global scale.SCT is an international clinical terminology standard, developed and maintained by the IHTSDO, which aims to cover

[1] Corresponding Author: marzouk.mamou@outlook.com

the whole field of health care by codes, terms and logical formalisms, in order to represent the entire health care process in detail [2, 3, 4].

The parallel use of SCT and ICD is addressed by an agreement between WHO and IHTSDO, who decided in 2010 to harmonize between the multi-component architecture of ICD-11 [5] and SCT [6, 7, 8].They put in place a Joint Advisory Group (JAG)in order to lay the grounds for semantic interoperability between ICD-11 and SCT by a semantic alignment method, based on 1998 foundation works [9, 10]and ontology design principled developed since then [11, 12, 13, 14].

In this paper, we present the application of two methods for the ICD – SCT alignment process, *viz.* the SCT compositional grammar (CG) and the new IHTSDO expression constraint language (ECL). Limitations and results are discussed.

2. Materials and Methods

We used the WHO browser for the circulatory chapter of the ICD-11 JLMMS [1] together with IHTSDO tools for SNOMED CT [15, 16].JLMMS was extracted from the ICD-11 foundation component [8] and meets the criteria of WHO classifications, *viz.*mutually exclusive classes, single hierarchies and exhaustiveness. Exhaustiveness is warranted by residual categories "Other" and "Unspecified", known from past ICD releases. The single hierarchy principle is enforced by the organization of ICD-11 into non-overlapping chapters, which requires exclusions rules at different hierarchical level in each chapter. E.g., the chapter "circulatory system" excludes infections, neoplasms, endocrine and congenital diseases called "developmental", which have their own chapters.

For this study, we focused on the circulatory chapter of the JLMMS, from which we excluded the arrhythmia sub-chapter, in which only few SCT concepts are fully defined. Two formalisms were tested for their ability to represent the meaning of ICD-11 classes:

- The SNOMED CT Compositional Grammar (CG), a composition syntax developed for assembling coordinated SNOMED CT expressions. Such SNOMED CT expression may include only the definition, without reference to the concept being defined [15]. For example, the concept "194727002 |Non-rheumatic mitral valve stenosis (disorder)|" can be defined as a CG expression as being 'equivalent to' "79619009 |Mitral valve stenosis (disorder)| + 274097009 |Non-rheumatic heart valve disorder (disorder)|".
- The SNOMED CT Expression Constraint Language (ECL) is a computable rule language that allows creating sets of clinical meanings represented by either pre-coordinated or post-coordinated expressions joined by conjunction, disjunction or exclusion. Each focus concept in an Expression Constraint is either a concept reference or a wildcard, and is normally preceded by either a constraint operator or a *memberOf* function. Each attribute consists of the attribute name (optionally preceded by a cardinality, reverse flag and/or attribute operator) together with the value of the attribute [16]. The same example of "194727002 |Non-rheumatic mitral valve stenosis (disorder)|" can be written in ECL with taking into account the JLMMS's exclusion of "postprocedural Mitral valve stenosis" as : <<60573004 |Aortic valve stenosis (disorder)|+ <<274097009 |Non-rheumatic heart valve disorder (disorder)|MINUS <703223000 |Postprocedural aortic valve stenosis

(disorder)|. This means that this expression constraint is satisfied only by the set of "60573004 |Aortic valve stenosis (disorder)|" (or descendant) associated with "274097009|Non-rheumatic heart valve disorder (disorder)|" (or descendant), and with exclusion of "703223000 |Postprocedural aortic valve stenosis".

The first step is to match the JLMMS classes with the IHTSDO Short Normal Form (SNF) CG of a SCT concept with the same lexical description (see Tab. 1).If there is a full match with this SCT concept, the SNF is taken as CG representation. When ICD-11 classes do not have a full match with SCT FSN, we develop whenever possible a logical representation formed by two or more existing CG expressions. When only a part of the ICD-11 class representation by existing CG expressions is possible, the representation is extended using a CG expression. When not even a partial meaning of an ICD-11 class can be expressed, logical expression in accordance with CG is developed. The second step uses the ECL allowing representing inclusions and exclusions of JLMMS classes.

Table 1. The types of correspondence between the ICD-11 JLMMS classes and SCT compositional grammar.

Match type and meaning	Action	Compositional grammar
Full match (M).	Take the SNF representation of the same SCT concept	The Short Normal Form (SNF) which exists.
No full match, but pre-coordination possible (O/A).	Create a pre-coordinate representation by merging two or more existing SNF representations	A new pre-coordinated SNF expression.
Partial match, no pre-coordination possible (O/E).	Create a logical post-coordinate expression with an existing SNF representation plus a new representation according to CG.	A new post-coordinated SNF expression.
No match (O/R).	Create a logical expression according to CG	A new logical expression respecting the compositional grammar.

3. Results

The results of the first step have been published elsewhere [17].We presents here some examples showing alternatively the CG and the ECL representation of ICD-11JLMMS classes. The first column shows the JLMMS classes with M meaning that there is SCT concept with the same name. The second column lists the inclusions and exclusions of the JLMMS class as they were at the last access to the JLMMS browser [1]. The third column is the CG [15] representation of the same lexical SCT concept and the fourth column is the ECL representation of the ICD-11JLMMS class [16].

The first line shows how ECL allows the representation of the JLMMS inclusions with the operator AND. The second line shows how the ECL allows the representation of the JLMMS exclusions with the operator MINUS. The third line shows how the ECL allows the representation JLMMS class inclusions and exclusions with operators AND and MINUS.

Table 2. Examples of Compositional Grammar and Expression Constraint Language of ICD-11 JLMMS

ICD-11 JLMMS class	JLMMS exclusions and inclusions rules	Short Normal Form compositional grammar (CG)	Expression Constraint Language (ECL)
Hypertensive renal disease (M)	*Inclusions :* -Chronic kidney disease due to hypertension. - Arteriosclerosis of kidney. - Arteriosclerotic nephritis. - (Chronic)(interstitial). hypertensive nephropathy. - Nephrosclerosis. - Glomerular diseases due to hypertension. - Unspecified contracted kidney due to hypertension.	90708001 |Kidney disease (disorder)|: 47429007 |Associated with (attribute)|= 38341003 |Hypertensive disorder, systemic arterial (disorder)|	<<90708001 |Kidney disease (disorder)|AND (<<32916005 |Nephrosclerosis (disorder)|AND709978007 |Contracted kidney (disorder)|) : 42752001 |Due to (attribute)|= <<38341003 |Hypertensive disorder, systemic arterial (disorder)|
Mitral valve stenosis (M)	*Exclusion :* Mitral stenosis with regurgitation.	11851006 |Mitral valve disorder (disorder)|: { 363698007 |Finding site (attribute)|= 91134007 |Mitral valve structure (body structure)|, 16676008 |Associated morphology (attribute)|= 415582006 |Stenosis (morphologic abnormality)|}	<<11851006 |Mitral valve disorder (disorder)|MINUS <<194726006 |Mitral stenosis with insufficiency (disorder)|: { 363698007 |Finding site (attribute)|= <<91134007 |Mitral valve structure (body structure)|, 116676008 |Associated morphology (attribute)|= <<415582006 |Stenosis (morphologic abnormality)|MINUS 708027006 |Valvular stenosis with valvular insufficiency (morphologic abnormality)|}
Coronary artery aneurysm (M)	*Inclusion :* Coronary arteriovenous fistula, acquired. *Exclusion :* Congenital coronary artery aneurysm.	301433005 |Aneurysm of artery of trunk (disorder)|: { 116676008 |Associated morphology (attribute)|= 85659009 |Aneurysm (morphologic abnormality)|, 363698007 |Finding site (attribute)|= 41801008 |Coronary artery structure (body structure)|}	(<<301433005 |Aneurysm of artery of trunk (disorder)|AND253720000 |Congenital coronary arteriovenous fistula (disorder)|) MINUS 204378009 |Congenital coronary aneurysm (disorder)|: { 116676008 |Associated morphology (attribute)|= <<85659009 |Aneurysm (morphologic abnormality)|, 363698007 |Finding site (attribute)|= <<41801008 |Coronary artery structure (body structure)|}

4. Discussion

On the one hand, a number of SNF CG expressions are incomplete or primitive. E.g., "401303003 |Acute ST segment elevation myocardial infarction (disorder)|" and "401314000 |Acute non-ST segment elevation myocardial infarction (disorder)|"have the same CG representation; it is necessary to complete the representation by adding "has_definitional_manifestation" 76388001|ST segment elevation. On the other hand by introducing AND, OR and MINUS operators, ECL allows a representation of inclusions and exclusions of the JLMMS classes, which is an essential characteristics of WHO classifications. Other and Unspecified classes cannot be represented by the

compositional grammar but by the expression constraint language. Unspecified classes are considered equivalent to their parent class and other by the Parent class CG and ECL MINUS.

5. Conclusion

Most classes in the JLMMS can best be represented with the two IHTSDO tools compositional grammar and expression constraint language much more effectively than with the Foundation Component (FC) [8].It seems therefore, that the methods initiated by the work reported in this paper can contribute to improving the interoperability between the two main worldwide health terminologies ICD-11 and SNOMED CT, despite their different structures, details and uses cases.

Acknowledgments

This work has benefited from the provision of JLMMS by WHO, and the method of semantic alignment developed by the JAG (WHO / IHTSDO).

References

[1] The World Health Organization. The International Classification of Diseases 11th revision JLMMS. http://apps.who.int/classifications/ICD-11/browse/l-m/en#/ (last access 23 Feb 2016).
[2] IHTSDO.SNOMED CT browser : http://browser.ihtsdotools.org/ (last access 23 Feb 2016).
[3] IHTSDO. SNOMED CT Starter Guide; Version : 2014-12-02, 12.
[4] IHTSDO. SNOMED CT Technical Implementation Guide; Jan 2015 : 52-465, 560.
[5] J.M. Rodrigues et al. Using the CEN/ISO standard for categorial structure to harmonise the development of WHO international terminologies, *Studies in Health Technology and Informatics 2009*; 150 : 255-259.
[6] J.M. Rodrigues et al. Sharing Ontology between ICD-11 and SNOMEDCT will enable seamless re-use and semantic interoperability. *Studies in Health Technology and Informatics 2013*; 192 : 343-346.
[7] J.M. Rodrigues et al. ICD-11 and SNOMED CT Common Ontology : Circulatory System. *Studies in health technology and informatics*, 2014;205 : 1043-1046.
[8] J.M. Rodrigues et al.Semantic Alignment between ICD-11 and SNOMED CT, *Proceedings MEDINFO 2015*, 790-794.
[9] B. Smith. Applied Ontology: A New Discipline is Born, *Philosophy Today*, 1998; 29(12) : 5-6.
[10] A. Rector et al. Practical development of re-usable terminologies : GALEN-IN-USE and the GALEN Organisation, *International Journal of Medical Informatics 1998*; 48 : 71-84.
[11] M. Ashburner et al. Gene Ontology: tool for the unification of biology. *Nature Genetics 2000*; 25:29,1-5.
[12] S. Schulz et al. Consolidating SNOMED CT's ontological commitment, *Applied Ontology 2011*; 6 : 1-11.
[13] OBO foundry principles: http://obofoundry.org/ontology/ma.html (last accessed 18 Jan2016).
[14] W3C OWL Working Group : OWL 2 Web Ontology Language Document Overview 2012, http://www.w3.org/TR/owl2-overview/ (last accessed 18 Jan2016).
[15] IHTSDO SNOMEDCT .Compositional Grammar.http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_CompositionalGrammarSpecification (last accessed 23 Feb 2016).
[16] IHTSDOSNOMEDCT languages: Expression Constraint language http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_ExpressionConstraintLanguageSpecificationAndGuide_Current-en-US_INT_20150820.pdf?ok (last accessed 23 Feb 2016).
[17] M. Mamou et al. ICD-11 (JLMMS) ICD-11 (JLMMS) and SCT Inter-Operation, *Proceedings E-HEALTH 2016, 267-272.*