MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190184

Expanding Evolutionary Terminology Auditing with Historic Formal and Linguistic Intensions: A Case Study in SNOMED CT

Werner Ceusters[a] and Sarah Mullin[a]

[a] *Department of Biomedical Informatics, University at Buffalo, Buffalo NY, USA*

Abstract

A method is described to use SNOMED CT's history mechanism as a means to compute how the formal and linguistic intensions of its concepts change over versions. As a result of this, it is demonstrated that the intended principle of concept permanence is not always adhered to. It is shown that the evolution of formal intensions can be monitored fully automatically and that the proposed procedure includes a method to suggest missing subsumers in a concept's transitive closure set by identifying mistakes that have been made in the past. Changes in linguistic intensions were found to be much more labor-intensive to identify. It is suggested that this could be improved if the history mechanism would come with more detailed motivations for change than the current and insufficiently used annotation to the effect that a fully specified name 'fails to comply with the current editorial guidance'.

Keywords:

Systematized Nomenclature of Medicine; Algorithms; Semantics

Introduction

SNOMED CT is a biomedical terminology which is anchored in an ontology of 'concepts' which are defined and related to each other using a combination of formal logic and editorial rules as specified in SNOMED CT's concept model [10, p24]. Fundamental in SNOMED CT is that each of its concepts comes with a unique identifier (SCTID) and is intended to have a unique meaning. What this meaning is intended to be, is conveyed formally by means of description logic-based assertions and informally by means of human-readable 'terms' which are called '*fully specified names*' (FSN) and which are connected to the concepts by means of '*descriptions'* [10, p12]. Each FSN is since 2003 formed by what we henceforth call a '*term proper'* and a '*semantic tag*' which indicates to which subhierarchy of SNOMED CT the concept belongs. For example, the concept with SCTID '87612001' has the FSN 'Blood (substance)' where 'Blood' is the term proper and 'substance' is the semantic tag. Concept 87612001 is an example of a concept for which the term leaves not much room for misinterpretation of the intended meaning. Its formal definition, however, specifies the intended meaning only partially: blood is a subtype of 'Body fluid (substance)' and of 'Blood material (substance)'. Semantic tags also disambiguate concepts for which the term proper is homonymous. This is for instance the case for the term 'hematoma' which comes in two flavors: 'Hematoma (morphologic abnormality)' (SCTID: 35566002) and 'Hematoma (disorder)' (SCTID: 385494008) which is formally defined in terms of the former.

SNOMED CT is since its inception in 2002 updated twice annually, for instance to include missing relevant content [7] or to remove content that was erroneous because of editorial mistakes [9] or because of mismatches in intended meaning between formal definitions and terms [11]. Also changes in the concept model itself require updating which impacts both the formal components and the FSNs. A unique feature of SNOMED CT is that it comes with a history mechanism involving certain formal metadata components that – to some degree – describe what and when changes have been made, what the reasons for these changes were (for example inconsistency with editorial rules), and how impacted components relate to each other after the change (for example what concepts, if any, replace inactivated concepts) [5].

One central principle that is intended to be maintained over versions is code or concept permanence [6]: '*Once assigned a meaning, a code must not change its meaning. Refinements, due to changes in the state of knowledge, may lead to inactivation of codes from SNOMED CT. An inactivated code may be replaced by a new, more precisely defined code*' [9, p203]. The objectives of the work presented here are twofold. The first one is to assess the extent to which the principle of concept permanence is adhered to and whether adherence to this principle can be quantified by resorting to SNOMED CT's history mechanism. The 2nd one is to find methods using this quantification to improve on prior efforts in Evolutionary Terminology Auditing which attempts to find mistakes in the last version of an ontology on the basis of errors made in the past [2]. Our hypothesis is that the stability of a concept's position in the hierarchy over distinct versions and the formal representation of reasons for change [3] contribute positively to quantification while changes in the concept model and in the FSNs contribute negatively [1].

Methods

Since SNOMED CT is an ontology that does not explicitly adhere to a view based on Ontological Realism [12], the meaning of a SNOMED CT concept can be thought of as what is conveyed by means of three aspects: (1) a *linguistic intension* as conveyed through its *label(s)*, (2) a *formal intension*, i.e. the properties implied by it as exhibited, for instance, by means of the formal relations it holds with other concepts and (3) an *extension*, i.e. the collection of data elements in, for instance, electronic medical record systems annotated with the concept [13]. Whether two concepts have the same meaning can then be determined by applying appropriate similarity functions to each of the three aspects followed by an assessment of whether the similarities are sufficiently high. For systems like SNOMED CT that maintain explicit identity over versions, concept permanence – or the opposite: *concept drift* [13] – can then be

Table 1 – Transitive closure history for SCTID:10001005 with most recent FSN: 'Bacterial sepsis (disorder)'

	Historic Formal Intension (HFI)					Historic subsumers' SCTIDs and most recent FSNs	
Ref.	**S1**	**S2**	**S3**	**S4**	**S5**		
(1)	1111	111111111	11	111111111	1111111111	91302008	Sepsis (disorder)
(2)	1111	111111111	11	111111111	1111111111	87628006	Bacterial infectious disease (disorder)
(3)	1111	111111111	11	111111111	1111111111	64572001	Disease (disorder)
(4)	1111	111111111	11	111111111	1111111111	40733004	Infectious disease (disorder)
(5)	1111	111111111	11	111111111	!!!!!!!!!!	105592009	Septicemia (disorder)
(6)	----	111111111	11	111111111	1111111111	404684003	Clinical finding (finding)
(7)	----	---------	11	111111111	0000000000	431950004	Bloodstream finding (finding)
(8)	----	---------	11	111111111	0000000000	431193003	Infection of bloodstream (disorder)
(9)	----	---------	11	111111111	0000000000	301811001	Bacterial infection by site (disorder)
(10)	----	---------	11	111111111	0000000000	301810000	Infection by site (disorder)
(11)	----	---------	11	111111111	0000000000	123946008	Disorder by body site (disorder)
(12)	----	---------	11	111111111	0000000000	118234003	Finding by site (finding)
(13)	----	---------	--	111111111	0000000000	434156008	Infectious agent in bloodstream (finding)
(14)	----	---------	--	---------	1111111111	238149007	Systemic inflammatory response syndrome (disorder)
(15)	----	---------	--	---------	1111111111	128139000	Inflammatory disorder (disorder)

Legend: '**Ref.**': reference. **S1**, … **S5**: stable history segments.

Table 2 – Historic signatures of some example concepts and their Fully Specified Names

Concept ID Ref.	Historic Signature	Fully Specified Name (FSN)	RS	RV	S
373981005	-111111111111111111111111111111111				
(1)	-111111111111111111111111111111000	Triazolam 0.125mg tablet (product)	S1	V3	0
(2)	-------------------------------100	Product containing triazolam 0.125 mg/1 each oral tablet (clinical drug)	S1	V2	0
(3)	--------------------------------10	Product containing only triazolam 0.125 mg/1 each oral tablet (clinical drug)	S1	V3	0
(4)	---------------------------------1	Product containing precisely triazolam 125 microgram/1 each conventional release oral tablet (clinical drug)			0
103497003	1111111111111111111111111111111111				
(5)	1111110000000000000000000000000000	Streptococcus penumoniae 3 (organism)			1
(6)	------1000000000000000000000000000	Streptococcus penumoniae serotype 3 (organism)			1
(7)	-------11111111111111111111111110	Streptococcus pneumoniae serotype 3 (organism)		V3	0
(8)	---------------------------------1	Streptococcus pneumoniae Danish serotype 3 (organism)			0
290537002	1111111111111111111111111111111111				
(9)	1111100011111111111111111111111111	Accidental prilocaine poisoning (disorder)			1
(10)	-----11100000000000000000000000000	Accidental prilocaine poisoning (event)			1
120713006	1111111111111111111111111111111111				
(11)	1111111111111111000000000000000000	Parainfluenza virus II antibody (substance)			0
(12)	----------------111111111111111111	Human parainfluenza virus 2 antibody (substance)			0

Legend: '**Ref.**' = reference. '**RS**' = annotated in a SNOMED CT reference set ('S1': SCTID-900000000000490003: '*Description inactivation indicator attribute value reference set*'). '**RV**': Value associated with the description annotated in RS ('V2': SCTID-900000000000485001: '*a component that contains a technical error*'; 'V3': SCTID-723277005: '*A component that fails to comply with the current editorial guidance*'). '**S**': Semantic similarity function result ('0' = not considered semantically equivalent, '1' = 'semantically equivalent').

evaluated by determining (1) whether changes in the concept's labels lead to different interpretations, (2) whether the formal intension is kept constant, and (3) whether the same sorts of data elements are annotated irrespective of version.

In absence of data annotated by means of different versions of SNOMED CT, we focused our efforts on linguistic and formal intensions. Using the history information from the July 2018 version of SNOMED CT International, we computed for all concepts, all descriptions specifying an FSN containing any semantic tag introduced since 2003 and all relations a *historic signature* as a 34-character string, using one character for each version since January 2002. A character '1' in position *n* indicates the presence of that component in version *Vn* (*n* ranging from 1 to 34), '-' absence of the component prior to its introduction, and '0' absence of the component because of its deactivation. For relationships between two concepts, the extra character '!' was used to indicate that a relationship does not hold anymore because of deactivation of the target concept.

For the formal intension aspect of a concept's meaning, we followed a suggestion from [13] and defined the *formal intension of a concept in Vn* (FIn) as the collection of all subsumers in its transitive closure in *Vn* and *the historic formal intension of a concept* (HFI) as the union of any subsumer ever encountered in some version. We further defined the *rigid formal properties set of a concept* (RFP) as the subset of transitive closure subsumers that is present in all versions in which the concept is active. Finally, we defined a *stable history segment of a concept* (SHSC) as a segment of the concept's history during which the FIn remains constant over successive versions.

Consistent with our proposal advanced in [4], we identified two types of *suspicious events* for historic signatures of concept-subsumer pairs. Concepts C1 and C2 form a concept-subsumer pair if and only if C2 is in the historic transitive closure set of C1. A historic signature for a concept-subsumer pair is *suspiciously gapped* for any transitive closure subsumer which becomes reactivated after having been deactivated: during the gap, the subsumer was thus unjustifiably absent ('missing'). A concept-subsumer pair is *suspiciously annulled in S* whenever there is a stable history segment S for which (1) the historic signature of that concept-subsumer pair contains only '0's and (2) when there is at least one other concept-subsumer pair for that concept of which the historic signature contains only '!' in S: this might be an indication that the subsumers marked by '0' are missing from the transitive closure set because of the removal of any property marked by '!'. To clarify these definitions, Table 1 provides an example of a concept with a HFI of 15, and an RFP of 4. Its HFI has 5 stable segments of which S5 is suspicious; this is because the historic signature of the properties (7) … (13) are suspiciously annulled in S5 due to the historic signature of property (5) in S5 containing '!'. This suggests that the subsumers marked by '0' are missing, i.e. unjustifiably absent, because of the – most likely justified – removal of the duplicate concept 'septicemia' (marked by '!') which was subsumed by these concepts.

Variables used in the analyses thereafter are the sizes of HFI and RFP as well as their ratio, the lifetime of a concept, and its number of suspicious events. A random number generator was used to select two random samples of each 100 concepts that were active since the first version. The samples were manually inspected for possible missing subsumers. The 1st sample was drawn from concepts which were marked as having a suspiciously gapped subsumer but not as having a suspiciously annulled subsumer. The 2nd sample consists of concepts which have at least one suspiciously annulled subsumer but are not marked as having a suspiciously gapped subsumer. Decisions for whether a subsumer is truly missing were based on SNOMED CT's editorial guidelines [9].

For the linguistic intension of a concept's meaning, we collected the historic signatures of all its FSNs over time and indicated for each change from one FSN to the next one whether a reason for the change was specified in one or other reference set distributed as part of SNOMED CT's metadata components (Table 2). Changes were syntactically qualified as having occurred in the term proper, in the semantic tag (ST), or in both. Changes from one ST to another were semantically qualified as being different, thus suggesting a distinct linguistic intension for the concept under scrutiny. To identify whether syntactic changes in the term proper for FSNs with the same ST would qualify as constituting a semantic change as well, we implemented a simple rule-based string transformation algorithm based on 99 rules. This algorithm processes each FSN in the history of a concept by iterating over a manually constructed knowledgebase sanctioning the substitution of certain character sequences (case insensitive). If at the end of the process an identical string is obtained for some FSNs, then these FSNs are considered semantically equivalent. It takes advantage of the fact that FSNs and subsequent changes thereof follow certain patterns. The example in Table 3 works for any HLA-*X*, e.g. HLA-Cw2, HLA-DQw8. Possible non-intended changes as in 'Chlamydia' → 'Cmydia' are innocent for our purposes as they would happen in each FSN of that concept. But obviously, it renders this algorithm inappropriate for computing the semantic similarity of distinct concepts on the basis of their linguistic intensions. A random sample of 200 concepts exhibiting at least one FSN change for which the algorithm failed to conclude semantic similarity was manually inspected for verification.

Changes in semantic tags (ST) were further analyzed by computing transition probabilities from one ST to another ST, and by performing agglomerative hierarchical clustering on larger trajectories and including activation and deactivation, for example finding → event → inactive, or substance →product → medicinal product. The result was assessed using the Ward (minimization of residual variance), average (averages of distances), and complete (minimization of diameter of each new group) methods from R cluster.

Table 3 – String transformation algorithm example

Search string	Replacement	Rule
"human leukocyte antigen"	→ " "	R1
"antigen"	→ " "	R2
"hla"	→ " "	R3
"-"	→ ""	R4
" "	→ ""	R5
String transformation sequence		**Rule**
'hla-dr8 antigen'		
'hla-dr8 '		R2
'-dr8 ''		R3
'dr8 '		R4
'dr8'		R5
'human leukocyte antigen hla-dr8 antigen'		
' hla-dr8 antigen'		R1
' hla-dr8 '		R2
' -dr8 '		R3
' dr8 '		R4
'dr8'		R5
'human leukocyte antigen dr8'		
' dr8'		R1
'dr8'		R5

Results

Our analysis involved 403,360 concepts that were active for at least one version, 340,639 (84.45%) of which are still active in the July 2018 version (Table 4).

The size of the historic formal intensions of concepts ranged from 2 (204 concepts) to maximally 152 (1 concept, most likely not the one most frequently found in an EHR: SCTID:35057008 - *Nonvenomous insect bite of penis with infection (disorder)*). The number of stable history segments ranged from 1 (7,961 concepts) to 25 (15 concepts). Only 61,001 concepts of all concepts (15.13%) exemplified a rigid formal property set (RFP) constituting 100% of its historic formal intension, while 51,936 concepts (15.25%) do so for all currently active concepts. 39,771 (=340,639-300,868, 11.68%) active concepts have at least one suspiciously annulled formal property. 7,583 concepts (403,360-395,777) have at least one suspiciously gapped subsumer. 2,706 concepts exhibited both.

Manual inspection of the samples for possible missing subsumers revealed that 83 of the 100 concepts with the suspicious gap criterium and 91 of those selected on the basis of suspicious annulment of a subsumer did, in our opinion, miss at least one subsumer. Some examples are provided in Table 5.

The number of concepts involved in changes in linguistic intensions, separated in semantic tag changes and term proper changes, are displayed in Table 6. Only 19% of these changes were found to be documented by means of a reference set. Of the remaining 81%, 91% could be eliminated through our term transformation algorithm, thereby still leaving over 48,000 term changes to be manually inspected.

Table 4 - Descriptive statistics for Historic Formal Intension related variables

Descriptive Statistic	*HFI*	*SHSC*	*ACTIVE*	*LIFETIME*	*RFP*	*RFP%*	*Susp. Ann.*	*Susp. gapped*
Mean	22.708	6.249	0.845	25.836	9.975	49.935	1.384	0.381
Standard Error	0.028	0.006	0.001	0.018	0.016	0.048	0.009	0.027
Median	17	5	1	34	7	43	0	0
Mode	7	2	1	34	4	100	0	0
Standard Deviation	17.784	3.861	0.362	11.417	10.026	30.703	5.627	17.164
Kurtosis	2.247	0.577	1.615	-0.609	6.597	-1.091	311.509	63,995.461
Skewness	1.428	0.952	-1.901	-0.999	2.256	0.371	13.736	202.515
Minimum (Min)	1	1	0	1	0	0	0	0
Maximum (Max)	152	25	1	34	125	100	326	6,643
Confidence Level(95.0%)	0.055	0.012	0.001	0.035	0.031	0.095	0.017	0.053
N concepts with Max	1	15	**340,639**	224,447	1	**61,001**	1	1
N concepts with Min	204	7,961	**62,721**	9,689	9,241	9,241	**300,868**	**395,777**
N Active concepts with Max	1	15	340,639	224,447	1	**51,936**	1	1
N Active concepts with Min	180	7,961	0	6,505	8,274	8,274	238,147	333,066

Table 5 – Examples of missing (i.e. once present, but deleted) subsumers in the transitive closure of SNOMED CT concepts

88425004: Congenital anomaly of nervous system (disorder)
- 299735001: Neurological lesion (finding)
- 102957003: Neurological finding (finding)

87290003: Congenital anomaly of head (disorder)
- 204223000: Ear, face and neck congenital anomalies (disorder)

83502000: Operation on tendon sheath (procedure)
- 118667007: Procedure on skeletal muscular system (procedure)

11381005: Acne (disorder)
- 95320005: Disorder of skin (disorder)
- 80659006: Disorder of skin and/or subcutaneous tissue (disorder)
- 106076001: Skin finding (finding)
- 301857004: Finding of body region (finding)
- 19660004: Disorder of soft tissue (disorder)

19838004: In-vitro immunologic test (procedure)
- 103693007: Diagnostic procedure (procedure)
- 362961001: Procedure by intent (procedure)
- 127789004: Laboratory procedure categorized by method (procedure)

230179001: Chronic viral encephalitis (disorder)
- 102957003: Neurological finding (finding)

116316008: Finding of foot region (finding)
- 250171008: Clinical history and observation findings (finding)

118835007: Procedure on ileum (procedure)
- 174035000: Lower gastrointestinal procedure (procedure)

29857009: Chest pain (finding)
- 250171008: Clinical history and observation findings (finding)
- 118222006: General finding of observation of patient (finding)
- 250171008: Clinical history and observation findings (finding)

10002003: Resection of stomach fundus (procedure)
- 38829003: Partial excision (procedure)

116175006: Proximal subtotal gastrectomy (procedure)
- 38829003: Partial excision (procedure)

287812001: Repair of stomach and/or duodenum (procedure)
- 118821005: Procedure on digestive organ (procedure)
- 118717007: Procedure on organ (procedure)

Table 6 – Changes in Fully Specified Names

		Concepts with Semantic Tag changes				
		0	**1**	**2**	**3**	**Total**
Concepts with Term Proper changes	**0**	333,712	33,697	3,833	56	371,298
	1	23,631	3,026	468	19	27,144
	2	1,748	1,704	1,186	0	4,638
	3	46	145	82	0	273
	4	1	1	5	0	7
Total		359,138	38,573	5,574	75	403,360

Table 7 – Examples of changes in linguistic intension without concept deactivation

SCTID: 374142001	
1	Product containing miglitol 25 mg/1 each oral tablet (clinical drug)
2	Product containing only miglitol 25 mg/1 each oral tablet (clinical drug)
3	Product containing precisely miglitol 25 milligram/1 each conventional release oral tablet (clinical drug)
SCTID: 100191000119105	
1	Acquired asymmetry of prostate (finding)
2	Asymmetry of prostate (finding)
SCTID: 102549009	
1	Night cramps (finding)
2	Cramp in lower leg associated with rest (finding)
SCTID: 106109006	
1	Number of previous abortions (finding)
2	Number of previous induced termination of pregnancy (finding)
SCTID: 302828001	
1	Syringoma (disorder)
2	Syringoma of skin (disorder)

In our sample of 200 concepts with at least one such non-documented change, we discovered 15 concepts with an FSN change exhibiting a clear shift in meaning. Some examples are shown in Table 7. Hierarchical clustering revealed statistically significant (1) that findings typically transition to events, (2) that multiple semantic tags (context dependent category, finding, procedure, regime/therapy, and disorder) transition to situation semantic tags, and (3) that substance and product have transitioned to medicinal product form, clinical drug, or became significantly more inactive.

Discussion

Computing the historic formal intension of all concepts in SNOMED CT requires a thorough understanding of the meta-components, but is algorithmically straightforward. As to the question of what formal properties should be included in it – direct subsumers only, stated relationships separate from inferred ones, the complete transitive closure with or without all associative relationships – there is no agreed upon answer [13]. Our preference for using the full transitive closure set made it possible to identify for those concepts whose formal intension changes under that criterion, subsumers that were possibly inadvertently removed as a consequence of rightfully removing some subsumed concept. While many assessments are straightforward, a problem for the evaluation, however, is that not enough textual definitions are provided for terms and concepts. What is, for instance, the scope of 'partial' in a subsumer? A 'resection of stomach fundus' (Table 5), whether complete or partial to the fundus is for sure partial for the stomach. Also the use of 'and' and 'or' is problematic. SNOMED CT's editorial guide comes in here handy, but it seems that the application thereof by SNOMED CT's authors is not rigorously followed.

The same holds for evaluating FSN changes that are suspicious for changes in the linguistic intension (Table 7). We can't imagine that clinicians who used in an earlier version a concept of the form 'Product containing X', would consider that equivalent to 'Product containing *only* X' and 'Product containing *precisely* X'. Over 300 change-sequences of this sort have been made in 2018 despite deactivation of the concepts involved seems to have been the more logical choice in light of concept permanence. Finding such meaning changes turned out not to be straightforward precisely because the reason for change mechanism is insufficiently used. Our algorithm for comparing linguistic intensions can for sure be improved, but more practical would it be if SNOMED CT would include a much more detailed list of reasons for change, and why not, a formal representation of all those conditions which make a component follow – or not – the 'current editorial guidelines'. This includes changes related to the concept model itself. Our findings related to the transitions involving semantic tags are consistent with those obtained via another methodology in [1]. It is in the first place an incomplete anchoring of the semantic tags into the formal hierarchy that poses a problem.

A limitation of the work presented here is that more manual analysis of discrepancies found is required in order to produce clear cut precision and recall values for our proposed algorithm. Also more experimenting with alternatives for historic formal intension computation is needed. Finally, it is worth exploring which missing subsumption relations detected through our effort are found as well through other methods [7; 8; 14].

Conclusions

Our results demonstrate that SNOMED CT's intended adherence to the criterion of concept permanence can be quantified but that, unfortunately, this criterion is not sufficiently applied. That changes in the concept model as expressed through changes in semantic tags have a negative effect on the automatic interpretation is re-confirmed: it cannot be formally computed whether for any given concept a semantic tag is changed because of a *local* mistake in the interpretation of that concept or a *global* change at the level of the concept model. Changes in linguistic intensions quantified on the basis of changes in the term proper of FSNs are detectable as well, but currently only with low estimated recall and precision.

Without doubt, our work demonstrates that SNOMED CT's history mechanism is a formidable resource from which valuable knowledge can be extracted to prevent mistakes in the future. It is our opinion that a mechanism like this should be standardly available in any ontology worth the name.

Acknowledgements

This work was supported in part by Clinical and Translational Science Award NIH 1 UL1 TR001412-01 from the National Institutes of Health and the Department of Veterans Affairs award #80307 'Identifying clinical and logical shortcomings in SNOMED CT', Projects #1144333 and #18114806.

References

[1] J.P. Bona and W. Ceusters, Mismatches between major subhierarchies and semantic tags in SNOMED CT, *J Biomed Inform* **81** (2018), 1-15.
[2] W. Ceusters, Applying Evolutionary Terminology Auditing to SNOMED CT, *AMIA Annu Symp Proc* **2010** (2010), 96-100.
[3] W. Ceusters, SNOMED CT's RF2: Is the future bright?, *Stud Health Technol Inform* **169** (2011), 829-833.
[4] W. Ceusters, SNOMED CT revisions and coded data repositories: when to upgrade?, *AMIA Annu Symp Proc* **2011** (2011), 197-206.
[5] W.M. Ceusters, K.A. Spackman, and B. Smith, Would SNOMED CT benefit from realism-based ontology evolution?, *AMIA Annu Symp Proc* (2007), 105-109.
[6] J.J. Cimino, Desiderata for controlled medical vocabularies in the twenty-first century., *Methods of Information in Medicine* **37** (1998), 394-403.
[7] L. Cui, O. Bodenreider, J. Shi, and G.Q. Zhang, Auditing SNOMED CT hierarchical relations based on lexical features of concepts in non-lattice subgraphs, *J Biomed Inform* **78** (2018), 177-184.
[8] L. Cui, W. Zhu, S. Tao, J.T. Case, O. Bodenreider, and G.Q. Zhang, Mining non-lattice subgraphs for detecting missing hierarchical relations and concepts in SNOMED CT, *J Am Med Inform Assoc* **24** (2017), 788-798.
[9] International Health Terminology Standards Development Organisation, SNOMED CT Editorial Guide, in, 2018.
[10] International Health Terminology Standards Development Organisation, SNOMED CT Starter Guide, in, 2018.
[11] J.M. Rodrigues, S. Schulz, B. Mizen, A. Rector, and S. Serir, Is the Application of SNOMED CT Concept Model sufficiently Quality Assured?, *AMIA Annu Symp Proc* **2017** (2017), 1488-1497.
[12] B. Smith and W. Ceusters, Ontological realism: A methodology for coordinated evolution of scientific ontologies, *Applied Ontology* **5** (2010), 139-188.
[13] S. Wang, S. Schlobach, and M. Klein, Concept drift and how to identify it, *Journal of Web Semantics* **9** (2011), 247-265.
[14] D. Wei, H. Helen Gu, Y. Perl, M. Halper, C. Ochs, G. Elhanan, and Y. Chen, Structural measures to track the evolution of SNOMED CT hierarchies, *J Biomed Inform* **57** (2015), 278-287.

Address for correspondence

Werner Ceusters, Department of Biomedical Informatics, 77 Goodell street, Bufallo NY – 14203, USA. Email: ceusters@buffalo.edu