

Development of A Web-based SNOMED Post-Coordinated Code Searching Tool

Cai Wu^{a, b}, Zhong Xie^a, Todd J. Johnson^a

^aSchool of Health Information Sciences, University of Texas Health Science Center, Houston, TX, USA

^bDepartment of Family and Community Medicine, Baylor College of Medicine, Houston, TX, USA

Abstract

Post-coordination is the ability to combine individual terms in a controlled vocabulary so as to code a concept that cannot be coded with a predefined term. Even though there are many pre-coordinated concepts that allow fairly complex concepts to be represented by a single term in SNOMED CT, there are always some clinical concepts that require post-coordination. This paper reports on the development of a Web-based tool for helping a coder construct post-coordinated SNOMED CT codes. The application allows the user to input the searching keyword or phrase with a selected modifier attribute to search the SNOMED code of the related qualifier and refiner for the concept. This feature improves the functionality and usability of SNOMED CT.

Keywords:

SNOMED CT, Post-coordination, Code, Web, Tool

Introduction

Structured and coded patient information is very important to improve the quality of health care and reduce medical errors. Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT) is a comprehensive and precise clinical reference terminology. The purpose of developing SNOMED CT is to provide a multipurpose health care terminology, to support improved quality of care, to support a reduction in health care costs, and to improve provider and patient access to information [1].

SNOMED CT contains about 344, 000 pre-coordinated concepts that allow fairly complex concepts to be represented by a single concept ID. It also permits the qualification or refinement of concepts to represent more detailed concepts by post-coordinated combinations of several concept IDs. A user may search for words or a phrase that can be represented by a rational combination of two or more concepts. For example, if the concept “severe headache” does not exist in a pre-coordinated form a search for the combination of the words “headache” and “severe” would ideally return the combination of two concepts, one for “headache” and the other for the qualifying concept “severe”.

There are three types of post-coordination: refinement, qualification and combination [1]. Qualification is a type of post-

coordination in which a concept is qualified by appending one or more qualifying characteristics. Refinement is a type of post-coordination which is very similar to qualification. Rather than adding a qualifier, refinement specifies a subtype of a defining characteristic associated with a concept. Combination is a type of post-coordination, which differs from qualification and refinement. Two (or occasionally more) concepts may be combined to represent a combined diagnosis or procedure [1].

Although the future releases of SNOMED CT will include new pre-coordinated concepts, there will always be some clinical concepts that require post-coordination. For example, “severe acute respiratory syndrome (SARS)” is a recently discovered disease, there is no pre-coordinated concept to represent it based on released SNOMED CT terminology.

Correctly coding the clinic terms with post-coordinated concepts represents several difficulties. It is not appropriate to refine a concept that is not refinable; qualifiers should only be used where the result of applying them results in a true subtype of the original concept; combination between the concepts should only be constructed for primitive concepts [1]. These and similar major modifications need to be handled in the ways to accurately retrieve and analyze the available information.

Currently there are a few available SNOMED code-searching tools to help coder to search the post-coordinated code. The CIC Look Up Engine - CLUE-5 - is an efficient Windows engine designed to ease implementation of SNOMED CT [2]. It shows all the concept structure through keyword searching including refiners and qualifiers, which is related to the searching concept. If the user is familiar with the SNOMED CT structure, and knows the functions of refiner and qualifier in the post-coordination, the information is very useful and can help the user to build the post-coordinated code. However, if a user is not familiar with the SNOMED CT structure, it is not easy to use the searching information to code the clinic concepts that require post-coordinated code. Thus it will be necessary and useful to develop an easy searching tool to assist with SNOMED CT post-coordination, which will improve the functionality and usability of SNOMED CT. This paper reports the development of a web-based SNOMED CT post-coordinated code-searching tool.

Methods

To assist with post-coordination, we propose a tool with the following characteristics

Recognize modifier from SNOMED CT terminology severe

SNOMED CT is a reference terminology with a semantic network of explicitly defined concepts. A semantic network is a network that defines the meaning of concepts in terms of relations to other concepts in the networks [3]. There are eleven modules in SNOMED, there are Topography, Morphology, Function, Disease and diagnosis, Procedures, Occupations, Living organisms, Chemicals, drugs and biological products, Physical agents, forces and activities, Social context, and General modifier, which can be conceived as distinct classifications [4]. Concepts within one module are related to each other using hierarchical relationships and concepts between different modules are related by non-hierarchical relationships [4]. Concepts, which represent objects, are non-vague; concepts from the General modifier module are vague but these are only used to compose new concepts [4]. Concepts of the various modules can be linked in order to compose new complex medical concepts.

There are a set of linkages, descriptors and qualifiers used to link or modify terms from each module in General linkage/modifiers module. If the linkages, descriptors and qualifiers that link with the key concepts can be determined, then those modifiers can be used for post-coordination by qualification or refinement.

Find Modifiers for Refiner or Qualifier

To assist with post-coordination, the application should allow a user to refine a concept. For example, the concept “total replacement of hip” could be refined by allowing the user to specify one of the subtypes of “hip prosthesis”. The application should also allow a user to qualify a concept. For example, the concept “tracheostomy” could be qualified by allowing user to specify a priority as “emergency”.

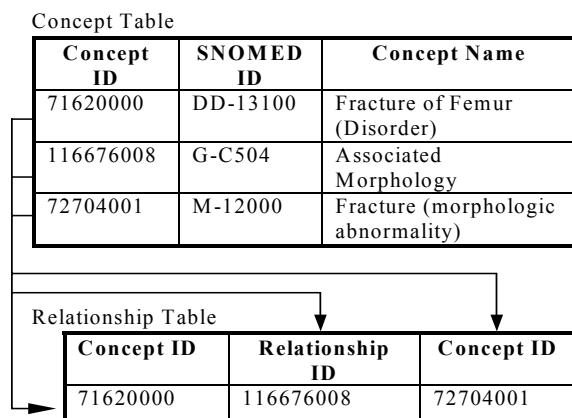


Figure 1 - The Relationships between Concept table and Relationships table

SNOMED CT concepts are linked to other concepts by specific named relationships, the full linkage of associated concepts can be staggering [5]. Concepts and relationships are stored in separate relational tables (Figure 1).

A relationship represents an association between two concepts, which contains identifiers of two logically associated concepts and the identifier of another concept that indicates the relationship type by which they are associated [5]. A relationship contains an indication of its characteristic type. This distinguishes between:

- Defining characteristics, which state things that are always true about a concept
- Qualifying characteristics, which offer options for qualifying a concept.
- Context specific characteristics, which may vary according to place or time

Defining characteristics provide the logical definitions of concepts. Concepts may be further defined by additional attributes including etiology, topography, method, etc. A qualifying characteristic is an attribute that may have one of several possible values for a particular concept. If a particular qualifier is applied to a concept, the resulting expression represents a more tightly defined subtype of that concept. A similar tightening of the definition of a concept can be achieved by allowing one or more of the defining characteristics associated with a concept to be refined. A defining characteristic is refined by an expression that applies a specified subtype of the value stated in the definition [1].

Table 1 – The refiner and qualifier attributes in SNOMED CT (released on July, 2002)

Refiners		Qualifiers
Access instrument	Has specimen	Access
Associated etiologic finding	Indirect device	Approach
Associated finding	Component	Course
Associated morphology	Interprets	Episodicity
Causative agent	Is a	Laterality
Indirect morphology	Method	Onset
Recipient category	Occurrence	Priority
Direct morphology	Part of	Severity
Direct substance	Finding site	Using
Pathological process	Procedure site	
Has definitional manifestation	Direct device	
Temporally follows	Revision status	
Subject of information	Has intent	
Has interpretation	Has focus	

The concept’s modifiers can be found through the concept’s characteristic type that refines a concept by selecting a subtype of one of its defining characteristics. For sanctioned qualifiers, a concept can be qualified by selecting one of its subtypes of the qualifier value; for unsanctioned qualifiers -

qualifiers that are not included as qualifying characteristics, may be constructed using pairs of concepts, which is a sub-type of the top-level concept “attribute” combined with an appropriate value. The refiners and qualifiers are shown in Table 1.

Generate Searching Keywords from input text strings

Effective implementation of SNOWMED CT depends on the ease and speed with which users can locate the terms and concepts that they wish to use. When typing text for a search the user is unlikely to know if their intended entry can be represented by a single concept or requires a post-coordinated expression involving additional concepts or qualifiers. In order to generate a post-coordinated SNOMED CT representation from a keyword or phrase, the application should integrate the input word and phrase to search with defining characteristic and qualifying characteristics by restricting the searches to appropriate qualifiers of a selected concept [6]. There are five tables related with keyword generation in SNOMED CT: ExcludedWords, DescWordKey, ConcWordKey, DescDualKey, and ConcDualKey. WordEquivalents table is an additional resource to assist searching and parsing of phrase. It contains equivalent, abbreviation, acronym words and word form variant. It also includes equivalent phrase and the usual role of the words, which can be considered for a post-coordinated combination of concepts that matches a phrase [6]. Based on these tables, a keyword generation algorithm can be used to text strings to generate the keywords [6]. The process for generating the single keyword consists of the following steps:

- Obtain source text

- Break source text into separate words
- Remove excluded words
- Process special characters
- Limit keyword length
- Remove duplicates

Store and Retrieve Post-coordinated SNOMED CT information in medical records

After find the modifiers for post-coordination, the application should allow a user to store and retrieve the post-coordinated codes. The storing method should represent the post-coordination code structure, fit the relational model, allow easy searching for concepts, and be efficient. Based on this purpose, the post-coordinated concepts can be stored using a combination of identifiers for its individual facets because the various aspects of the concept are coordinated during data entry rather than in the preparation of the terminology [1]. We need a way to store the data in a relational database that will facilitate searches and statistics. The basic structure is showed in Table 2. The Med_Code, which is the sequence order, and the ConceptID, which is the original ID from SNOMED CT, are the combined primary key. For searching key concept, its AttributeID will be “key”; for modifier, its AttributeID will be its ConceptID from SNOMED CT. The Med_Code and ConceptID will be used for storing and retrieving a set of post-coordinated codes.

Table2 - The basic structure of Post_Code table. Med_Code and ConceptID are the combined primary key

Med Code	Concept ID	SNOMED ID	Attribute ID	Attribute SNOMED ID
1	274325001 (Percutaneous liver biopsy) (Procedure)	P1-5B30B	Key	0
1	10200004 (Liver structure) (Body structure)	T-62000	363764007 (Procedure site)	G-C019
27	71620000 (Fracture of femur) (Disorder)	DD-13100	Key	0
27	72704001 (Fracture) (Morphologic abnormality)	M-12000	116676008 (Associated morphology)	G-C504

Results

System Description

SNOMED CT Post-coordinated Searching Tool is a three-tier Web architecture and developed with Active Server Page, Java Script and dynamic HTML. The Web server runs on a Microsoft Windows NT machine with Internet Information Server and is implemented using Active Server Pages. The terminology server – SNOMED CT database is built by loading SNOMED CT data files (released on July, 2002), and stored in an SQL Server Database (SQL Server 2000). The

user interaction is through a Web browser with Microsoft Explorer 5.x or greater, Netscape 6.x or greater. The database connection between application and database is through ODBC (Figure 2).

The tool main page includes an instruction about usage of this tool. The user interface is divided into two frames. The top-most frame is a searching keyword inputting part. Users can input the keyword into the text field, then click the “Search” button to view the searching results. The user also needs to select an attribute from the drop down list to add the qualifier or refiner for the post-coordinated. The second frame displays the search results. If there is a modifier for the concept, the display part will show numbers of the retrieved records, the concept name, attribute name, and modifier name. Each time

when application find a post-coordinated concept user can clicked “Code” button, the post-coordinated codes will be added into the post-coordinated record system if the codes are not in the system.

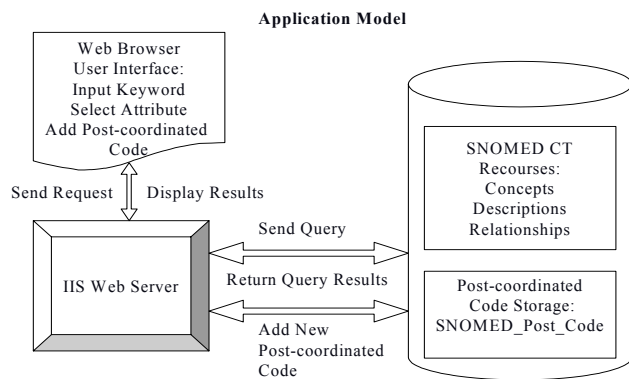


Figure 2 – The project application model

Here is an example to code SARS. The application first will separate the phrase to single words as: “severe”, “acute”, “respiratory”, and “syndrome”; then check and remove non-concept words such as “in”, “this”, “both”, “as”, ... , which are not included in the searching words in SARS; the synonyms, acronym words, abbreviations, word form variants, and word

equivalences also will be checked through WordsEquivalents table, which are also not included in this example. Secondly, the application will identify the modifier attribute values: “severe”, “acute”; then generate the keywords: “respiratory” and “syndrome”. The application will search the first eight characters of the keywords in ConcWordKey table. If the keywords existed in the SNOMED CT terminology severe, the ConceptID will link to concept and relationship tables. Then the application will find the concept that is fully defined, with defining/qualify characteristic and can be refined by selecting subtypes. If there is no any sanctioned qualifier matching with the root concept, the application will look for unsanctioned qualifier, which has been selected by user. In this example, the application found some predefined concepts with sanctioned qualifier (“acute”), but there is no any match with qualifier “severe”. Then the application will search the unsanctioned qualifier “severe” with the root concepts, such as “acute respiratory disease” or “acute respiratory infections”. The application finally get the root concepts with the qualifier attribute for “severity” valuing as “severe” (Figure3). Thus the application facilitates a search for a rational post-coordinated combination as:

“acute respiratory disease” (severity = “severe”)

Then user can click “Code” button to add all ConceptID and SNOMEDID for above concepts into Post_Code to construct the post-coordinated codes.

There are 5 distinct records for “severe acute respiratory syndrome” and “Severity (attribute)” in this database.

You can select a concept from the list to insert into our SNOMED code table

Searching Concept / (SNOMED ID)	Attribute / (SNOMED ID)	Value / (SNOMED ID)	Post-coordination
Acute respiratory disease (disorder) / (D2-00001)	Severity / (G-C197)	Severe (severity modifier) (qualifier value) / (G-A003)	<input type="button" value="Code"/>
Acute respiratory infections (disorder) / (D2-0004E)	Severity / (G-C197)	Severe (severity modifier) (qualifier value) / (G-A003)	<input type="button" value="Code"/>
Acute respiratory alkalosis (disorder) / (D6-24220)	Severity / (G-C197)	Severe (severity modifier) (qualifier value) / (G-A003)	<input type="button" value="Code"/>
Acute respiratory syncytial virus bronchitis (disorder) / (DE-33A16)	Severity / (G-C197)	Severe (severity modifier) (qualifier value) / (G-A003)	<input type="button" value="Code"/>
Acute respiratory acidosis (disorder) / (D6-23620)	Severity / (G-C197)	Severe (severity modifier) (qualifier value) / (G-A003)	<input type="button" value="Code"/>

Figure 3. Application interface: searching records display

User Evaluation of the Search Tool

Preliminary user testing by Pathologists showed that this web-base SNOMED post-coordinated code searching tool is easy to use and meets their requirements. Further evaluation will be continued.

Discussion

As a complementary coding method of pre-coordination, post-coordination improves the functionality and usability of SNOMED CT. Through a semantic network linkage, post-

coordinated codes can be constructed by root-concepts with a set of qualifiers and definers in SNOMED CT terminology [7]. However, search for the post-coordinated expressions is not a simple work. In theory, all multiple word searches may could resolvable through post-coordination of several concepts [6]. If each word is checked independently the total number of matches for a phrase containing several words is the union of the sets of matches for each word. Thus unless some rationalization is applied, the more words there are in a search string the more possible matches and post-coordinated combinations need to be considered [6]. A good user interface should rationalize the records to provider the user with the minimum number of options that they need to access a particular concept. Different search strategies may be appropriate to the nature or the restriction and the number of terms within the restricted scope. In this application, we create an attribute table to store all attribute values. When user inputs the search words or phrase, the application will first to identify the attribute, so that the keywords can be limited to those root-concepts. Thus correctly inputting the attributes values, such as “routine”, “urgent”, “emergency”, is a critical issue for successful identifying keywords.

In healthcare there are many words with equivalent meanings. Some words may be used as synonym, abbreviation or acronym, or variant form. For example, “MI” as “myocardial infarction” or “mitral incompetence”; “inflamed”, “inflammatory”, “inflammation”; “pulmonary”, “lung”, “respiratory”. A simple keyword index with phrase and pattern matching may fail to find the terms that a user requires. This problem can be addressed by enabling searches using word equivalent table in SNOMED CT terminology [6].

In this application, we did not include post-coordination by combination. This type of post-coordination is to combine two or more primitive concepts representing a combined diagnosis or procedure. However, combinations like this should only be used to represent concepts that can be regarded as discrete reusable clinical statements. They should not be used to construct arbitrarily complex representations of multiple statements to a particular record [1]. Thus we did not add this feature into the application in case it would make the relationship between combined concepts uncertain.

Post-coordination may cause same concept coded in many different ways. Unconstrained post-coordination can lead to an unwieldy number of possible expressions. The allowable expressions can be constrained by limiting the roles available. One possible approach to optimization of retrieval is to transform the original stored information into an equivalent representation with the minimum number of post-coordinated components. . When new pre-coordinated codes have been issued, the equivalent post-coordinated codes should be transfer to pre-coordinated codes.

Web technology provides an easy-to-use, inexpensive, cross-platform solution that allows even a casual user to successfully navigate the complex terminology. This tool may significantly affect biomedicine and medical informatics [8][9]. This application is easy to use and can be accessed from anyplace where one can connect to the Internet. The users do not need any knowledge about post-coordination, they only needs

to input the searching keywords or phrase to identify matched concepts in SNOMED CT terminology severe.

Conclusion

Post-coordination is a complementary coding method for those codes that do not exist in a vocabulary. It provides more specific and detailed coding information to refine or qualify the pre-coordinated codes. The web-based SNOMED post-coordination code-searching tool can help the coder code post-coordinated codes using SNOMED CT. The critical issue of successful search is the keyword generation. The keyword generation algorithm that has been applied in this application is relevant to search phrase to enable effective searches. When new pre-coordinated codes have been issued, the equivalent post-coordinated codes should be transfer to pre-coordinated codes.

References

- [1] SNOMED Clinical Terms Guide: Technical Implementation Guide, College of American Pathology. July 2002.
- [2] Clue-5 for SNOMED® Clinical Terms, Clue-5 home page: <http://www.clininfo.co.uk/clue5/clue.htm>
- [3] Rallins MC, Spachman KA, Wilcke J. Representing non-defining attributes in a clinical terminology to support practical applications: Examples from SNOMED clinical terms. Proc. AMIA Annual Fall Symp 2001, 1006.
- [4] Keizer NF, Abu-Hanna A. Understanding terminological systems II: Experience with conceptual and formal representation of structure. *Methods Inform Med* 2000; 39:22-29.
- [5] SNOMED Clinical Terms Guide: Core Structure, College of American Pathology. July, 2002
- [6] SNOMED Clinical Terms Technical Specification, Word Search Tables. College of American Pathology. September, 2001.
- [7] Dolin RH, Spaceman KA, Markwell D. Selective retrieval of pro- and post-coordinated SNOMED concepts. Proc. AMIA Annual Fall Symp, 2002, 210-214
- [8] Lowe HJ, Lomax EC, Polonkey SE. The World Wide Web: a review of an emerging internet-based technology for the distribution of biomedical information. *J. Am Med Inform Assoc.* 1996;3(1):1-14.
- [9] Cimino JJ. Beyond the superhighway: exploiting the Internet with medical informatics. *J. Am Med Inform Assoc.* 1997; 4 (4): 279-284.

Address for correspondence

Cai Wu, Department of Family and Community, Baylor College of Medicine. 3701 Kirby, Suite 600. Houston, TX 77098.

E-mail: cwu@bcm.tmc.edu