Analysis and Recommendations for the Persistence and Subsumption-Testing of Post-Coordinated Problem Expressions Using SNOMED-CT

Report Submitted to

Partners Healthcare System

Sujansky & Associates, LLC

March 18, 2007

Table of Contents

1. Introduction

This document is a technical white paper that describes various options for representing post-coordinated problem expressions encoded in SNOMED CT (SCT) and for supporting query processing against such expressions. Partners Healthcare System has engaged Sujansky & Associates to analyze these options in the course of designing and implementing an enterprise problem repository. Specifically, Partners is seeking information to assist in a feasibility assessment and technical design for the use of post-coordinated SCT expressions in applications that capture and use problem-list entries.

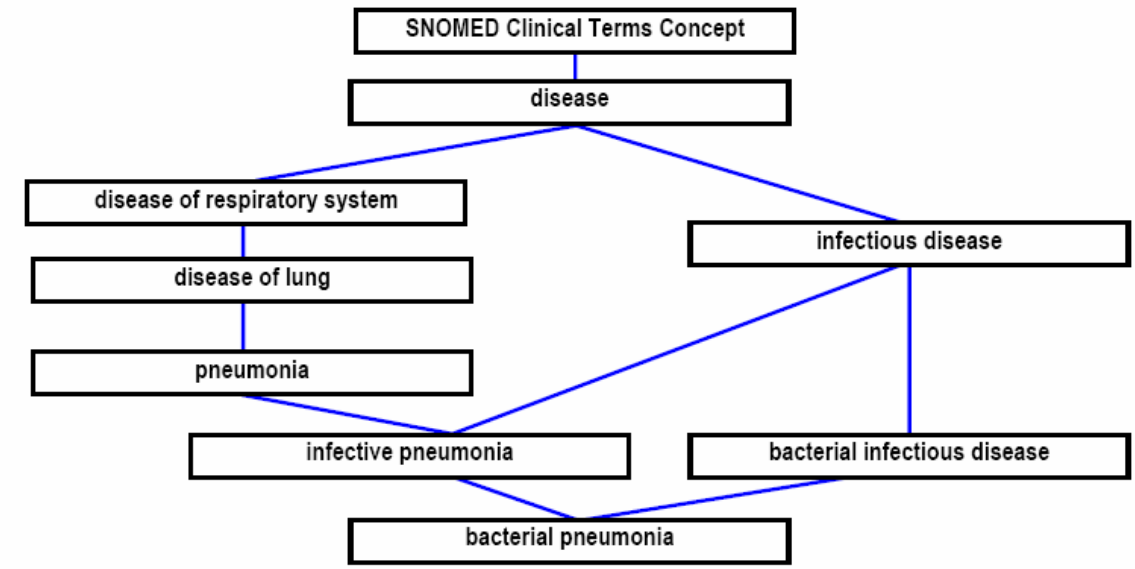
This document first provides a brief background on post-coordination in SCT and describes the scope of our analysis. The report next describes the requirements for a problem-list repository that were used to inform the analysis. The following section presents a set of relevant design dimensions and specific design options for each dimension. Lastly, the report presents a set of recommendations with respect to each design dimension, as well as an example problem instance encoded per these recommendations.

2. Background and Scope

Partners Healthcare System (PHS) is developing an enterprise-wide problem-list repository. The goal of this project is to standardize the storage of patients' problem lists in order to enable the sharing of problem-list data among clinical applications (such as inpatient and outpatient EHRs) and the centralization of services that automatically process problem-list data (such as reporting and decision support services).

The standardization of problem lists in a manner that supports automated processing requires a formal terminology model. The most important requirements of such a terminology model are that (1) it can represent any clinician-specified problem accurately and precisely and (2) it can support automated query and retrieval operations correctly and efficiently. PHS has selected SNOMED CT (SCT) for its problem-list terminology model.

SCT consists of a large set of pre-defined medical concepts (currently > 350,000 concepts) that are hierarchically organized and inter-related. The size of SCT helps the terminology meet the first requirement noted above, i.e. adequate coverage of the problems that clinicians need to document. The hierarchical and other relationships within SCT help it meet the second requirement, i.e., support for relevant query and retrieval operations. To illustrate the contents and structure of the SCT terminology, the following graphic shows a subset of the pre-defined concepts that might be used to populate a problem list:



Using SCT, a clinician can document that a patient has bacterial pneumonia by specifying on the patient's problem list the unique SCT identifier for that concept, i.e., the clinician would add the following entry:

```
Bacterial Pneumonia (ConceptID = 53084003)
```

Additionally, a reporting program would use the SCT hierarchy to automatically retrieve the same patient's record in response to the query "retrieve all patients with any infectious disease (Infectious Disease : ConceptID = 40733004) on their problem list". The program could automatically determine whether bacterial pneumonia (Bacterial Pneumonia : ConceptID = 53084003) is an infectious disease (Infectious Disease : ConceptID = 40733004) by using a process called *subsumption testing*.

Importantly, SCT also supports the ability to express new medical concepts by combining pre-existing ones. This process, called *post-coordination*, enables clinicians who use SCT to express problems that do not appear as pre-defined concepts in the terminology, thereby vastly increasing SCT's expressive power. For example, a clinician could document that a patient has "bacterial pneumonia caused by methicillin-resistant Staph. Aureus" by combining the pre-existing concept "bacterial pneumonia" with the pre-existing concept "Methicillin Resistant Staph. Aureus" and specifying that the latter is the "causative agent" of the former. The patient's problem list would then contain an entry consisting of the following expression:

```
Bacterial Pneumonia (ConceptID = 53084003) :
       Causative Agent (ConceptID=246075003) =
              Methicillin Resistant Staph. Aureus (ConceptID=115329001)
```

If specified correctly, post-coordinated expressions also support subsumption testing. Hence, the patient whose problem list contains the expression above would also be identified by the query "find all patients with any infectious disease (Infectious Disease : ConceptID = 40733004) on their problem list."

For additional background information on the SNOMED terms and concepts used in this report, please see the glossary in Appendix A.

Although very useful, post-coordination creates a number of practical challenges for information systems that support this capability. The foremost challenge, and the one that has been most studied, is the design of user interfaces that enable clinicians to create post-coordinated expressions efficiently, intuitively, and in a manner that is consistent with the SCT terminology model. Although somewhat easier in the specific domain of problem list maintenance, these user-interface issues remains a significant challenge. However, another important set of challenges pertain to the management of post-coordinated expressions *after* they have been specified by clinicians. These management tasks include the appropriate persistence of post-coordinated problems in a patient database and efficient subsumption testing against problem lists that include post-coordinated problems. This report addresses those data-management challenges, which include:

- Determining the *degree of transformation and normalization* to apply to post-coordinated expressions when they are persisted in a database. What transformations and normalizations appropriately balance the needs of storage efficiency, retrieval performance, terminology evolution, and medicolegal requirements?
- Determining the specific *structure and syntax* for representing post-coordinated expressions when they are persisted in a database. What structure and syntax appropriately balance the needs of storage efficiency, retrieval performance, interoperability, and software evolution?
- Determining the appropriate way to represent the *contextual modifiers* for problems within post-coordinated expressions. For example, representing modifiers that indicate whether a problem is a current diagnosis for the patient, a past medical problem of the patient, or a disorder in the patient's family history.
- Determining appropriate strategies for *optimizing the performance of subsumption testing* against post-coordinated concepts (a critical but inherently costly operation). Potential

strategies include maintaining a transitive closure of the SCT hierarchy and incorporating post-coordinated concepts into the SCT terminology model ("just-in-time pre-coordination").

The report first summarizes requirements of the PHS problem-list repository and assumptions about the PHS environment that pertain to the persistence and processing of post-coordinated SCT expressions. The report then describes various options for addressing the challenges listed above, including advantages and disadvantages, and concludes with specific technical recommendations.

The sections below use the following terms, as defined by the SCT terminology model. For readers not familiar with these terms, a glossary appears in Appendix A.

concept
Concept
Attribute
Relationship
Concept Definition
Pre-coordinated concept
Expression
Post-coordinated Expression
Refinement
Focus Concept
Subsumption Testing
Equivalence Testing
Predicate Expression
Candidate Expression

1. Requirements and Assumptions

[Omitted in public version]

2. Options for Persistence and Management of Post-Coordinated Problems

[Omitted in public version]

3. Recommendations

[Omitted in public version]

6. Appendix A: Glossary

Term	Definition	Example
concept	A clinical idea	The concept of a broken femur bone
Concept	A clinical idea to which a unique SNOMED ConceptId has been assigned	Fracture of Femur (SNOMED ConceptID = 71620000)
Attribute	A type of association between two Concepts	Finding Site
Relationship	An association between two Concepts	Fracture of Femur : Finding Site = Bone Structure of Femur
Concept Definition	A collection of Relationships that logically defines the meaning of a Concept in SNOMED	Fracture of Femur : IS-A = Injury of Thigh, Finding Site = Bone Structure of Femur, Morphology = Fracture
Pre-coordinated concept	A concept that is pre-defined as a Concept in SNOMED	Fracture of Femur (SNOMED ConceptID = 71620000) : IS-A = Injury of Thigh, Finding Site = Bone Structure of Femur, Morphology = Fracture
Expression	A collection of references to one or more Concepts used to express an instance of a clinical idea (i.e., in a particular patient) An expression may consist of a single ConceptID or a large collection of related Concepts	Fracture of Femur OR Fracture of Femur : Finding Site = Structure of Head of Femur : Laterality = Left, Morphology = Spiral Fracture
Post-coordinated Expression	An expression created to represent an instance of a clinical idea that does not exist as a pre-defined Concept in SNOMED	Fracture of Femur : Finding Site = Structure of Head of Femur : Laterality = Left, Morphology = Spiral Fracture (clinical idea = spiral fracture of the left femur)

Refinement	The further specification or addition of Relationships to a predefined Concept to express a more specific concept	<u>Pre-defined Concept</u> Fracture of Femur : IS-A = Injury of Thigh, Finding Site = Bone Structure of Femur, Morphology = Fracture <u>Refinement</u> Fracture of Femur : Finding Site = Structure of Head of Femur : Laterality = Left, Morphology = Spiral Fracture, Severity = Severe
Focus Concept	The core concept that is refined in a post-coordinated expression	Fracture of Femur (SNOMED ConceptID = 71620000) IN Fracture of Femur : Finding Site = Structure of Head of Femur : Laterality = Left, Morphology = Spiral Fracture, Severity = Severe
Subsumption Testing	The logical determination of whether a concept (as represented by an Expression) is more specific than another concept (also represented by an Expression). If so, the more specific concept *is subsumed by* the more general concept, and the more general concept *subsumes* the more specific concept.	Fracture of Femur (see Concept Definition above) SUBSUMES Fracture : Finding Site = Structure of Head of Femur : Laterality = Left, Morphology = Spiral Fracture (the head of the femur is a part of the femur, and a spiral fracture is a kind of fracture) Fracture of Femur (see Concept Definition above) DOES NOT SUBSUME Fracture : Finding Site = Bone Structure of Shaft of Fibula Morphology = Transverse Fracture (the shaft of the fibula is not a part of the femur)

Equivalence Testing	The logical determination of whether a concept (as represented by an Expression) is exactly the same as another concept (also represented by an Expression). If so, the two concepts are *equivalent.*	Fracture of Femur (see Concept Definition above) IS EQUIVALENT TO Traumatic Injury : Finding Site = Bone Structure of Femur Morphology = Fracture Fracture of Femur (see Concept Definition above) IS NOT EQUIVALENT TO Traumatic Injury : Finding Site = Structure of Head of Femur : Laterality = Left, Morphology = Spiral Fracture (the second concept is more specific than "Fracture of Femur"; although it is subsumed by it, it is not equivalent to it)
Predicate Expression	The Expression that is being tested as the *more general* concept in a subsumption test. This is typically the expression that appears in a query.	Fracture of Femur in the subsumption tests above
Candidate Expression	The Expression that is being tested as the *more specific* concept in a subsumption test. This is typically the expression that appears in the patient record.	Post-coordinated Expressions in the subsumption tests above