1. Introduction

Background

SNOMED CT is a clinical terminology with global scope covering a wide range of clinical specialties and requirements. The use of SNOMED CT expressions in Electronic Health Records (EHRs) provides a standardized way to represent clinical meanings captured by clinicians and enables the automatic interpretation of these meanings. SNOMED CT expressions are a structured combination of one or more concept identifiers used to represent a clinical idea in a logical manner. The SNOMED CT Compositional Grammar provides a lightweight syntax for the representation of SNOMED CT expressions.

In contrast, a SNOMED CT Expression Constraint is a computable rule that can be used to define a bounded set of clinical meanings represented by either precoordinated or postcoordinated expressions. Expression constraints can be used as formal constraints on the content of a particular data element in an EHR, as the intensional definition of a concept-based reference set, as a machine processable query that identifies a set of matching precoordinated or postcoordinated expressions, or as a constraint that restricts the range of an attribute defined in the SNOMED CT concept model.

History

Expression constraints have been used in projects and programs around the world for a number of years – for example HL7 TermInfo, and the NHS Logical Record Architecture.

In 2013, a draft document on "SNOMED CT Expression Constraint Syntax Specification for Terminology Binding" was developed as an assignment during the SNOMED CT Implementation Advisor (SIA) scheme.

In 2014, this work was revised and extended to support a wider range of relevant use cases to produce version 1.0 of the Expression Constraint Language specification (2015). These updates included:

- Concrete values (e.g. integers, decimals and strings) are now permitted as attribute values. This is to provide alignment with the recent extensions to SNOMED CT Compositional Grammar;
- Cardinality constraints have been introduced, and as a result the optional operator (i.e. ~) is no longer provided;
- Attributes may now be preceded by a 'descendantOf' or 'descendantOrSelfOf' operator to indicate whether attribute descendants and/or the attribute itself should be used in the matching process;
- A reverse flag has been introduced, which allows relationships to be traversed in the reverse direction;
- Exclusion has been changed from a unary operator (‘negation’) to a binary operator (‘minus’);
- A wildcard character (*) has been introduced to represent any concept in the substrate;
- A number of clarifications have been made, including the 'memberOf' operator and the default substrate upon which the expression constraints are executed.

An update to the Expression Constraint Language was then published in 2016 (version 1.1) to incorporate some additional features requested by implementers of the language. These updates include:

- Two new operators 'childOf' and 'parentOf' were added to support querying immediate children and immediate parents of a concept during user interface design;
- A new ‘dot notation’ was introduced (as an alternative to the Reverse flag) to refer to an attribute value for a concept or expression;
- The ability for a constraint operator (e.g. ‘descendantOf’) to be applied to a nested expression constraint was added;
- The ability to add comments within the text of an expression constraint was added;
- Additional optional brackets were allowed around subexpressions; and
- The non-normative syntax (previously named the ‘Full Syntax’) was renamed to the ‘Long Syntax’.

Early in 2017 version 1.2 was published, to include a new feature requested by implementers: namely, the ability for the 'memberOf' function to be applied to a set of reference set concepts defined using an expression constraint. In this version, the explanation of Operator Precedence was also moved from section 6.7 to section 5.4. Version 1.3 was then published in mid 2017 to support a range of additional features - including allowing the refinement of subexpression constraints, permitting the use of subexpression constraints to represent a set of valid attribute names and simplifying the parsing of dotted expression constraints.

For a full list of previous versions and a summary of updates, please refer to Previous Versions.

Purpose

The purpose of this document is to define and describe a formal language for representing SNOMED CT Expression Constraints. A SNOMED CT Expression Constraint is a computable rule that defines a bounded set of clinical meanings represented by either precoordinated or postcoordinated expressions. Two equivalent syntaxes are presented – a brief syntax, which is designed to be as compact as possible for interoperable communication between systems, and a long syntax, which introduces textual alternatives to the symbols from the brief syntax. This document also provides examples and guidance to assist in the implementation of this language.

Scope

This document presents the specification of an Expression Constraint Language, which can be used to represent SNOMED CT Expression Constraints. It includes a logical model of the language, two syntaxes, a set of example expression constraints and a summary of implementation considerations.
The Expression Constraint Language specified in this document is part of a consistent set of computer processable languages designed to support a variety of use cases involving the use of SNOMED CT. Other SNOMED CT computable languages, which are either complete or under development include:

- **Compositional Grammar**: designed to represent SNOMED CT expressions;
- **Query Language**: designed to express computable queries over SNOMED CT content; and
- **Template Syntax**: which allow slots to be added to expressions, expression constraints or queries that can be filled with specific values at a later time.

The compositional grammar is designed to provide a common foundation for the additional functionality added by the other languages.

This document does not include a full description of how to implement an expression constraint parser, classifier or interpreter. It does not describe how to transform an expression constraint into other languages, such as OWL, SPARQL or SQL; or how to determine whether two expression constraints are equivalent. It also does not describe how to implement an EHR which uses expression constraints to constrain or query its content, or a terminology server which uses expression constraints to query its content. Instead, it provides a specification, examples and general guidance to assist in the implementation of expression constraints in any of these applications.

**Audience**

The target audiences of this document include:

- SNOMED National Release Centres;
- SNOMED CT designers and developers, including designers and developers of EHR systems, information models, data entry interfaces, storage systems, decision support systems, retrieval and analysis systems, communication standards and terminology services;
- SNOMED CT terminology developers, including concept model designers, content authors, map developers, subset and constraint developers and release process managers.

It should be noted that this document contains both technical and non-technical content. In particular, the detailed logical model and formal syntax is specifically focussed at more technical readers. Less technical readers are encouraged to read the introductory material (including the use cases and requirements) and the extensive set of examples that is presented. It should also be noted that even though complex expression constraints are possible, most expression constraints are likely to be very simple, such as those described in Simple Expression Constraints.

**Document Overview**

This document defines the SNOMED CT Expression Constraint Language and describes how and where it may be implemented. Chapter 2 begins by describing the use cases in which it is anticipated that SNOMED CT Expression Constraint Language will be used. Chapter 3 then describes the requirements used to guide the definition of this language. In Chapter 4, the logical model of the Expression Constraint Language is presented, while in Chapter 5 two syntaxes are defined using an ABNF serialisation of the logical model. Chapter 6 then presents some examples of expression constraints that conform to the SNOMED CT Expression Constraint syntaxes, and Chapter 7 discusses some implementation considerations. Appendix A – Examples Of Valid Expressions provides some examples of precoordinated and postcoordinated expressions that satisfy each of the expression constraints presented earlier in the document. Appendix B – Examples Of Invalid Expressions then provides some examples that do not satisfy these expression constraints.

**Glossary**

The following table contains the definition of terms used within this document. Please refer to the SNOMED International Glossary for additional SNOMED CT definitions.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Backus-Naur Form (ABNF)</td>
<td>A language used to define the formal syntax of another language (as defined by Internet Standard 68, RFC 5234).</td>
</tr>
<tr>
<td>Compositional Grammar</td>
<td>The set of rules that govern the way in which SNOMED CT expressions are represented as a plain text string.</td>
</tr>
<tr>
<td>Concept Model</td>
<td>A set of rules that determines the permitted sets of relationships between particular types of concepts.</td>
</tr>
<tr>
<td>Expression</td>
<td>A structured combination of one or more concept identifiers used to express a clinical idea.</td>
</tr>
<tr>
<td>Expression Constraint</td>
<td>A computable rule that can be used to define a bounded set of clinical meanings.</td>
</tr>
<tr>
<td>Extensonal Reference Set</td>
<td>A reference set whose membership is defined by enumeration.</td>
</tr>
<tr>
<td>Intensional Reference Set</td>
<td>A reference set whose membership is defined using a serialised query.</td>
</tr>
<tr>
<td>Machine Readable Concept Model</td>
<td>A representation of the rules that comprise the SNOMED CT Concept Model in a form that can be processed by computer software and applied to validate content.</td>
</tr>
<tr>
<td>Postcoordinated Expression</td>
<td>Representation of a clinical meaning using a combination of two or more concept identifiers is referred to as a postcoordinated expression.</td>
</tr>
<tr>
<td>Precoordinated Expression</td>
<td>Representation of a clinical meaning using a single concept identifier is referred to as a precoordinated expression.</td>
</tr>
<tr>
<td>Reference Set</td>
<td>A SNOMED CT file structure consisting of a set of references to SNOMED CT components.</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Substrate</td>
<td>The SNOMED CT content over which an expression constraint is evaluated or a query is executed.</td>
</tr>
</tbody>
</table>