A Declarative Semantics for SNOMED CT Expression Constraints

January 26, 2015

Contents

1 Axiomatic Data Types

1.1 Atomic Data Types

This section identifies the atomic data types that are assumed for the rest of this specification, specifically:

- **SCTID** – a SNOMED CT identifier
- **TERM** – a fully specified name, preferred term or synonym for a SNOMED CT Concept
- **REAL** – a real number
- **STRING** – a string literal
- **GROUP** – a role group identifier
- $\mathbb{N}$ – a non-negative integer
- $\mathbb{Z}$ – an integer

We also introduce several synonyms for *SCTID*: **TODO:** change SUBJECT to SOURCE

- **SUBJECT** – a *SCTID* that appears in the *sourceId* position of a relationship.
- **ATTRIBUTE** – a *SCTID* that appears in the *typeId* position of a relationship.
- **REFSETID** – a *SCTID* that identifies a reference set

$[SCTID, TERM, REAL, STRING, GROUP]$

$SUBJECT == SCTID$
$ATTRIBUTE == SCTID$
$REFSETID == SCTID$

1.2 Composite Data Types

While we can't fully specify the behavior of the concrete data types portion of the specification at this point, it is still useful to spell out the anticipated behavior on an abstract level.

- **CONCRETEVALUE** – a string, integer or real literal
- **TARGET** – the target of a relationship that is either an SCTID or a CONCRETEVALUE

$$CONCRETEVALUE ::= string\langle\langle STRING\rangle\rangle \mid integer\langle\langle \mathbb{Z}\rangle\rangle \mid real\langle\langle REAL\rangle\rangle$$
$$TARGET ::= object\langle\langle SCTID\rangle\rangle \mid concrete\langle\langle CONCRETEVALUE\rangle\rangle$$

2 The Substrate

A substrate represents the context of an interpretation. A substrate consists of:

- **c** The set of *SCTID*s (concepts) that are considered valid in the context of the substrate. **References to any *SCTID* that is not a member of this set MUST be treated as an error.**
- **a** The set of *SCTID*s that are considered to be valid attributes in the context of the substrate. **Reference to any *ATTRIBUTE* that is not a member of this set MUST be treated as an error.**
- **r** A set of relationship quads (subject, attribute, target, group)
- **descs** The subsumption (ISA) closure from general to specific (descendants)
- **ancs** The subsumption (ISA) closure from specific to general (ancestors)
- **refset** The reference sets within the context of the substrate whose members are members are concept identifiers (i.e. are in c). While not formally spelled out in this specification, it is assumed that the typical reference set function would be returning a subset of the *refsetId*/*referencedComponentId* tuples represented in one or more RF2 Refset Distribution tables.

Issue: this is currently the simplest possible kind of Substrate; it represents the DNF form of SNOMED CT with concepts only (i.e., *Expressions* and *Expression Libraries* are not supported).

Issue: this is a *flat* representation (fixed to 1 level of nesting corresponding to role grouping) and does not properly handle Concrete Domains (which almost always involve additional nesting

Open questions: While the Core will not contain cycles, and probably NRC Extensions will not, arbitrary extensions (e.g. as a result of handling Expression libraries) may involve cycles (or at least equivalent concepts).

- **Resolution:** A reflexive, symmetric equivalence relationship (*equiv*) was added to the substrate. How it is determined is left unspecified, but a rule was added saying that an equivalent concept can not be a descendant of its equivalence.

- Is it correct to interpret the transitive closure directly against the substrate relationships (r) set or is there an implication of a reasoner being invoked somewhere, which could potentially render the transitive closure as logically derived from (r). **Answer:** If substrate includes postcoordinated expressions then subsumption would need to be calculated. **Followup:** How does one know whether it includes postcoordinated expressions? How do we express this decision formally?

- Do we want to assert that it is an error to use a non-attribute concept in a attribute position? If not, should we explicitly include a formal definition of the attribute set? **Answer:** Because the ECL doesn't declare this as an error condition, we shouldn't assert in the Z spec. **Resolution:** $a = descs\ attribute\_concept$ assertion pulled from the substrate declaration.

2.1 Substrate Components

Quad Relationships in the substrate are represented a 4 element tuples or "Quads", which consist of a subject, attribute, target and role group identifier.

```
─ Quad ──────────────────────────
│ s : SUBJECT
│ a : ATTRIBUTE
│ t : TARGET
│ g : GROUP
└────────────────────────────────
```

We will also need to recognize some "well known" identifiers: the *is_a* attribute, the *zero_group* and *attribute_concept*, the parent of all attributes

```
│ is_a : ATTRIBUTE
│ zero_group : GROUP
│ attribute_concept : SCTID
│ refset_concept : SCTID
```

2.2 Substrate

The formal definition of substrate follows, where c and r are given and the remainder are derived. The expressions below assert that:

1. All subjects and attributes and targets of type *object* in r must be members of the set of concepts, c. All attributes are also members of a.

2. The *is_a* relation is irreflexive and can only occur in the zero group.

3. *childrenOf includes* the function from a concept in c to the set of concepts that are the source of c in an *is_a* relationship in the *zero_group*. Note

that substrates that include post-coordinated expressions may include additional entries in the *childrenOf* function that are not in the relationships table and, instead, are derived through the application of a DL reasoner.

4. *parentsOf includes* the function from a concept c to the set of concepts that are the target of of c in an *is_a* relationship in the *zero_group*. Note that substrates that include post-coordinated expressions may include additional entries in the *parentsOf* function that are not in the relationships table and, instead, are derived through the application of a DL reasoner.

5. *parentsOf* and *childrenOf* are irreflexive.

6. The *equiv* relationship is symmetric and reflexive.

7. The ancestors function, *ancs* ,is the *irreflexive* transitive closure of the *parentsOf* relationship.

8. The descendants function, *descs*, is the *irreflexive* transitive closure of the inverse of the *childrenOf* relationship.

9. The descendants relationship must not contain any concepts that are declared as equivalent to the root.

10. The reference set function is a function from *subset of* c to set of SCTID's in c. A SCTID that is not in the domain of *refset* cannot appear as the target of a *memberOf* function

Substrate

$$\begin{array}{l}
c : \mathbb{P}\, SCTID \\
r : \mathbb{P}\, Quad \\
a : \mathbb{P}\, SCTID \\[1ex]
childrenOf : SCTID \nrightarrow \mathbb{P}\, SCTID \\
parentsOf : SCTID \nrightarrow \mathbb{P}\, SCTID \\[1ex]
descs : SCTID \nrightarrow \mathbb{P}\, SCTID \\
ancs : SCTID \nrightarrow \mathbb{P}\, SCTID \\
equiv : \mathbb{P}(SCTID \times SCTID) \\
refset : REFSETID \nrightarrow \mathbb{P}\, SCTID \\
\hline
a \subseteq c \\
\forall\, rel : r \bullet rel.s \in c \land rel.a \in a \land (object^{\sim} rel.t \in SCTID \Rightarrow object^{\sim} rel.t \in c) \\
\forall\, rel : r \bullet q.a = is\_a \Rightarrow (q.s \neq object^{\sim} q.t \land q.g = zero\_group) \\[1ex]
\mathrm{dom}\, childrenOf = c \land \bigcup(\mathrm{ran}\, childrenOf) \subseteq c \\
\forall\, s : c \bullet parentsOf\ s \subseteq \{q : r \mid q.s = s \land q.a = is\_a \bullet object^{\sim} q.t\} \\
\forall\, c1, c2 : c \bullet c1 \in parentsOf\ c2 \Leftrightarrow c2 \in childrenOf\ c1 \\[1ex]
\forall\, x : c \bullet (x \mapsto x) \in equiv \\
\forall\, x_1, x_2 : c \bullet x_1 \mapsto x_2 \in equiv \Rightarrow x_2 \mapsto x_1 \in equiv \\[1ex]
\forall\, s : c \bullet descs\ s = childrenOf\ s \cup \bigcup\{t : childrenOf\ s \bullet descs\ t\} \\
\forall\, t : c \bullet ancs\ t = parentsOf\ t \cup \bigcup\{s : parentsOf\ t \bullet ancs\ s\} \\
\forall\, r : equiv \bullet first\ r \in \mathrm{dom}\, descs \Rightarrow second\ r \notin (descs\ (first\ r)) \\[1ex]
\mathrm{dom}\, refset \subseteq c \land \bigcup(\mathrm{ran}\, refset) \subseteq c
\end{array}$$

strict_substrate

$$\begin{array}{l}
Substrate \\
\hline
\forall\, q : r \bullet q.a \in descs\ attribute\_concept \\
\mathrm{dom}\, refset \subseteq descs\ refset\_concept
\end{array}$$

3 Result

The result of applying a query against a substrate is either a (possibly empty) set of SCTID's or an *ERROR*. **TODO:** Where is unknownAttribute used?

$$\begin{array}{l}
ERROR ::= unknownOperation \mid unknownConceptReference \mid \\
\qquad\qquad unknownAttribute \mid unknownRefsetId
\end{array}$$

$$Result ::= ok\langle\!\langle \mathbb{P}\, SCTID \rangle\!\rangle \mid error\langle\!\langle ERROR \rangle\!\rangle$$

4 Interpretation of Intermediate Constructs

This section carries the interpretation of the intermediate constructs – the various forms of expressions and their combinations.

4.1 expressionConstraint

```
    expressionConstraint =
( refinedExpressionConstraint / unrefinedExpressionConstraint )
```

4.1.1 Interpretation

The interpretation of `expressionConstraint` the interpretation of the `refinedExpressionConstraint` or the `unrefinedExpressionConstraint`.

$$\begin{aligned}&\mathit{expressionConstraint} ::= \\ &\qquad \mathit{ecrec}\langle\!\langle \mathit{refinedExpressionConstraint}\rangle\!\rangle \mid \\ &\qquad \mathit{ecurec}\langle\!\langle \mathit{unrefinedExpressionConstraint}\rangle\!\rangle\end{aligned}$$

```
i_expressionConstraint :
        $Substrate \rightarrow expressionConstraint \rightarrow Result$
----------
$\forall ss : Substrate;\ ec : expressionConstraint \bullet i\_expressionConstraint\ ss\ ec =$
    if $ecrec^{\sim} ec \in refinedExpressionConstraint$
        then $i\_refinedExpressionConstraint\ ss\ (ecrec^{\sim} ec)$
    else if $ecurec^{\sim} ec \in unrefinedExpressionConstraint$
        then $i\_unrefinedExpressionConstraint\ ss\ (ecurec^{\sim} ec)$
    else $error\ unknownOperation$
```

4.1.2 unrefinedExpressionConstraint

An `unrefinedExpressionConstraint` is either a `compoundExpressionConstraint` or a `simpleExpressionConstraint`

```
unrefinedExpressionConstraint =
    compoundExpressionConstraint / simpleExpressionConstraint
```

$$\begin{aligned}&\mathit{unrefinedExpressionConstraint} ::= \\ &\qquad \mathit{ucec}\langle\!\langle \mathit{compoundExpressionConstraint}\rangle\!\rangle \mid \\ &\qquad \mathit{usec}\langle\!\langle \mathit{simpleExpressionConstraint}\rangle\!\rangle\end{aligned}$$

$$
\begin{array}{|l}
\mathit{i\_unrefinedExpressionConstraint} : \\
\qquad \mathit{Substrate} \rightarrow \mathit{unrefinedExpressionConstraint} \rightarrow \mathit{Result} \\
\hline
\forall \mathit{ss} : \mathit{Substrate};\ \mathit{uec} : \mathit{unrefinedExpressionConstraint} \bullet \\
\mathit{i\_unrefinedExpressionConstraint}\ \mathit{ss}\ \mathit{uec} = \\
\quad \textbf{if}\ \mathit{ucec}^\sim \mathit{uec} \in \mathit{compoundExpressionConstraint} \\
\qquad \textbf{then}\ \mathit{i\_compoundExpressionConstraint}\ \mathit{ss}\ (\mathit{ucec}^\sim \mathit{uec}) \\
\quad \textbf{else if}\ \mathit{usec}^\sim \mathit{uec} \in \mathit{simpleExpressionConstraint} \\
\qquad \textbf{then}\ \mathit{i\_simpleExpressionConstraint}\ \mathit{ss}\ (\mathit{usec}^\sim \mathit{uec}) \\
\quad \textbf{else}\ \mathit{error}\ \mathit{unknownOperation}
\end{array}
$$

4.1.3 refinedExpressionConstraint

```
refinedExpressionConstraint =
     unrefinedExpressionConstraint ":" refinement /
     "(" refinedExpressionConstraint ")"
```

The interpretation of `refinedExpressionConstraint` is the intersection of the interpretation of the `unrefinedExpressionConstraint` and the `refinement`, both of which return a set of SCTID's or an error.

The interpretation of the second option adds no value.

$$
\begin{array}{l}
\mathit{refinedExpressionConstraint} == \\
\qquad \mathit{unrefinedExpressionConstraint} \times \mathit{refinement}
\end{array}
$$

$$
\begin{array}{|l}
\mathit{i\_refinedExpressionConstraint} : \\
\qquad \mathit{Substrate} \nrightarrow \mathit{refinedExpressionConstraint} \nrightarrow \mathit{Result} \\
\hline
\forall \mathit{ss} : \mathit{Substrate};\ \mathit{rec} : \mathit{refinedExpressionConstraint} \bullet \\
\mathit{i\_refinedExpressionConstraint}\ \mathit{ss}\ \mathit{rec} = \\
\quad \mathit{intersect}\ (\mathit{i\_unrefinedExpressionConstraint}\ \mathit{ss}\ (\mathit{first}\ \mathit{rec}))(\mathit{i\_refinement}\ \mathit{ss}\ (\mathit{second}\ \mathit{rec}))
\end{array}
$$

4.2 simpleExpressionConstraint

The interpretation of `simpleExpressionConstraint` is the application of an optional constraint operator to the interpretation of `focusConcept`, which returns a set of SCTID's or an error. The interpretation of an error is the error.

```
simpleExpressionConstraint =
     [constraintOperator ] focusConcept
```

$$
\begin{array}{l}
\mathit{constraintOperator} ::= \mathit{descendantOrSelfOf} \mid \mathit{descendantOf} \mid \\
\qquad \mathit{ancestorOrSelfOf} \mid \mathit{ancestorOf} \\
\mathit{simpleExpressionConstraint} == \mathit{constraintOperator}[0 \mathinner{..} 1] \times \mathit{focusConcept}
\end{array}
$$

$$\begin{array}{|l}
\mathit{i\_simpleExpressionConstraint} : \\
\qquad \mathit{Substrate} \pfun \mathit{simpleExpressionConstraint} \pfun \mathit{Result} \\
\hline
\forall \mathit{ss} : \mathit{Substrate};\ \mathit{sec} : \mathit{simpleExpressionConstraint} \bullet \\
\mathit{i\_simpleExpressionConstraint}\ \mathit{ss}\ \mathit{sec} = \\
\qquad \mathit{i\_constraintOperator}\ \mathit{ss}\ (\mathit{first}\ \mathit{sec})\ (\mathit{i\_focusConcept}\ \mathit{ss}\ (\mathit{second}\ \mathit{sec}))
\end{array}$$

4.3 compoundExpressionConstraint

The interpretation of a *compoundExpressionConstraint* is the interpretation of its corresponding component.

```
compoundExpressionConstraint = conjunctionExpressionConstraint |
       disjunctionExpressionConstraint | exclusionExpressionConstraint |
        "(" compoundExpressionConstraint ")"
```

$$\begin{array}{l}
\mathit{compoundExpressionConstraint} ::= \\
\qquad \mathit{cecConj}\langle\langle \mathit{conjunctionExpressionConstraint} \rangle\rangle \mid \\
\qquad \mathit{cecDisj}\langle\langle \mathit{disjunctionExpressionConstraint} \rangle\rangle \mid \\
\qquad \mathit{cecExc}\langle\langle \mathit{exclusionExpressionConstraint} \rangle\rangle
\end{array}$$

$$\begin{array}{|l}
\mathit{i\_compoundExpressionConstraint} : \\
\qquad \mathit{Substrate} \pfun \mathit{compoundExpressionConstraint} \pfun \mathit{Result} \\
\hline
\forall \mathit{ss} : \mathit{Substrate};\ \mathit{cec} : \mathit{compoundExpressionConstraint} \bullet \\
\mathit{i\_compoundExpressionConstraint}\ \mathit{ss}\ \mathit{cec} = \\
\quad \mathbf{if}\ \mathit{cecConj}^\sim \mathit{cec} \in \mathit{conjunctionExpressionConstraint} \\
\qquad \mathbf{then}\ \mathit{i\_conjunctionExpressionConstraint}\ \mathit{ss}\ (\mathit{cecConj}^\sim \mathit{cec}) \\
\quad \mathbf{else\ if}\ \mathit{cecDisj}^\sim \mathit{cec} \in \mathit{disjunctionExpressionConstraint} \\
\qquad \mathbf{then}\ \mathit{i\_disjunctionExpressionConstraint}\ \mathit{ss}\ (\mathit{cecDisj}^\sim \mathit{cec}) \\
\quad \mathbf{else\ if}\ \mathit{cecExc}^\sim \mathit{cec} \in \mathit{exclusionExpressionConstraint} \\
\qquad \mathbf{then}\ \mathit{i\_exclusionExpressionConstraint}\ \mathit{ss}\ (\mathit{cecExc}^\sim \mathit{cec}) \\
\quad \mathbf{else}\ \mathit{error}\ \mathit{unknownOperation}
\end{array}$$

$$\begin{array}{|l}
\mathit{i\_compoundExpressionConstraint}' : \\
\qquad \mathit{Substrate} \pfun \mathit{compoundExpressionConstraint} \pfun \mathit{Result}
\end{array}$$

4.4 conjunctionExpressionConstraint

conjunctionExpressionConstraint is interpreted the conjunction (intersection) of the interpretation of two or more *subExpressionConstraints*

```
conjunctionExpressionConstraint =
        subExpressionConstraint 1*(conjunction subExpressionConstraint)
```

$$\mathit{conjunctionExpressionConstraint} == \mathit{subExpressionConstraint} \times \mathrm{seq}_1(\mathit{subExpressionConstraint})$$

$$\begin{array}{|l}
\mathit{i\_conjunctionExpressionConstraint} : \\
\qquad \mathit{Substrate} \mathrel{+\!\!\!\rightarrow} \mathit{conjunctionExpressionConstraint} \mathrel{+\!\!\!\rightarrow} \mathit{Result} \\
\hline
\forall \mathit{ss} : \mathit{Substrate};\ \mathit{cecr} : \mathit{conjunctionExpressionConstraint} \bullet \\
\mathit{i\_conjunctionExpressionConstraint}\ \mathit{ss}\ \mathit{cecr} = \\
\quad \mathbf{if}\ \mathit{tail}(\mathit{second}\ \mathit{cecr}) = \langle\rangle\ \mathbf{then} \\
\qquad \mathit{intersect}\ (\mathit{i\_subExpressionConstraint}\ \mathit{ss}\ (\mathit{first}\ \mathit{cecr}))(\mathit{i\_subExpressionConstraint}\ \mathit{ss}\ (\mathit{head}\ (\mathit{seco} \\
\quad \mathbf{else} \\
\qquad \mathit{intersect}\ (\mathit{i\_subExpressionConstraint}\ \mathit{ss}\ (\mathit{first}\ \mathit{cecr}))(\mathit{i\_conjunctionExpressionConstraint}\ \mathit{ss}\ ((\mathit{h}
\end{array}$$

4.5 disjunctionExpressionConstraint

disjunctionExpressionConstraint is interpreted the disjunction (union) of the interpretation of two or more *subExpressionConstraints*

```
disjunctionExpressionConstraint =
      subExpressionConstraint 1*(disjunction subExpressionConstraint)
```

$$\begin{array}{l}
\mathit{disjunctionExpressionConstraint} == \\
\qquad \mathit{subExpressionConstraint} \times \mathrm{seq}_1(\mathit{subExpressionConstraint})
\end{array}$$

$$\begin{array}{|l}
\mathit{i\_disjunctionExpressionConstraint} : \\
\qquad \mathit{Substrate} \mathrel{+\!\!\!\rightarrow} \mathit{disjunctionExpressionConstraint} \mathrel{+\!\!\!\rightarrow} \mathit{Result} \\
\hline
\forall \mathit{ss} : \mathit{Substrate};\ \mathit{cecr} : \mathit{disjunctionExpressionConstraint} \bullet \\
\mathit{i\_disjunctionExpressionConstraint}\ \mathit{ss}\ \mathit{cecr} = \\
\quad \mathbf{if}\ \mathit{tail}(\mathit{second}\ \mathit{cecr}) = \langle\rangle\ \mathbf{then} \\
\qquad \mathit{union}\ (\mathit{i\_subExpressionConstraint}\ \mathit{ss}\ (\mathit{first}\ \mathit{cecr}))(\mathit{i\_subExpressionConstraint}\ \mathit{ss}\ (\mathit{head}\ (\mathit{second} \\
\quad \mathbf{else} \\
\qquad \mathit{union}\ (\mathit{i\_subExpressionConstraint}\ \mathit{ss}\ (\mathit{first}\ \mathit{cecr}))(\mathit{i\_disjunctionExpressionConstraint}\ \mathit{ss}\ ((\mathit{heac}
\end{array}$$

4.6 exclusionExpressionConstraint

exclusionExpressionConstraint is interpreted the result of removing the members of the second *subExpressionConstraint* from the first

```
exclusionExpressionConstraint =
      subExpressionConstraint exclusion subExpressionConstraint
```

$$\begin{aligned}&exclusionExpressionConstraint == \\ &\qquad subExpressionConstraint \times subExpressionConstraint\end{aligned}$$

$$\begin{array}{|l}
i\_exclusionExpressionConstraint : \\
\qquad Substrate \nrightarrow exclusionExpressionConstraint \nrightarrow Result \\
\hline
\forall ss : Substrate;\ ecr : exclusionExpressionConstraint \bullet \\
i\_exclusionExpressionConstraint\ ss\ ecr = \\
\quad minus\ (i\_subExpressionConstraint\ ss\ (first\ ecr))\ (i\_subExpressionConstraint\ ss\ (second\ ecr))
\end{array}$$

4.7 subExpressionConstraint

subExpressionConstraint is interpreted as the interpretation of either a *simpleExpressionConstraint* or a *compoundExpressionConstraint*

```
subExpressionConstraint =
      simpleExpressionConstraint | "(" compoundExpressionConstraint ")
```

$$\begin{aligned}&subExpressionConstraint ::= \\ &\quad secSec\langle\langle simpleExpressionConstraint\rangle\rangle \mid \\ &\quad secCec\langle\langle compoundExpressionConstraint\rangle\rangle\end{aligned}$$

$$\begin{array}{|l}
i\_subExpressionConstraint : \\
\qquad Substrate \nrightarrow subExpressionConstraint \nrightarrow Result \\
\hline
\forall ss : Substrate;\ sec : subExpressionConstraint \bullet \\
i\_subExpressionConstraint\ ss\ sec = \\
\quad \textbf{if}\ (secSec^{\sim} sec) \in simpleExpressionConstraint \\
\qquad \textbf{then}\ i\_simpleExpressionConstraint\ ss\ (secSec^{\sim} sec) \\
\quad \textbf{else if}\ secCec^{\sim} sec \in compoundExpressionConstraint \\
\qquad \textbf{then}\ i\_compoundExpressionConstraint'\ ss\ (secCec^{\sim} sec) \\
\quad \textbf{else}\ error\ unknownOperation
\end{array}$$

4.8 refinement

`refinement` is the interpretation of one or more `attribute` or `attributeGroup`s joined by `binaryOperator`s

```
binaryOperator = conjunction / disjunction / minus
refinement = (attribute / attributeGroup / "(" refinement ")")
         [binaryOperator  refinement]
```

$$\begin{aligned}&refTarget ::= att\langle\langle attribute\rangle\rangle \mid attg\langle\langle attributeGroup\rangle\rangle \\ &refinement == refTarget \times \mathrm{seq}(binaryOperator \times refTarget)\end{aligned}$$

```
$i\_refTarget : Substrate \rightarrow refTarget \rightarrow Result$
$i\_refinement : Substrate \rightarrow refinement \rightarrow Result$
---
$\forall\, ss : Substrate;\ rt : refTarget \bullet i\_refTarget\ ss\ rt =$
    **if** $att^{\sim} rt \in attribute$
        **then** $i\_attribute\ ss\ (att^{\sim} rt)$
    **else if** $attg^{\sim} rt \in attributeGroup$
        **then** $i\_attributeGroup\ ss\ (attg^{\sim} rt)$
    **else** $error\ unknownOperation$

$\forall\, ss : Substrate;\ ref : refinement \bullet$
    $i\_refinement\ ss\ ref =$
    $evalRefinement\ ss\ (i\_refTarget\ ss\ (first\ ref))\ (second\ ref)$
```

```
$evalRefinement :$
        $Substrate \rightarrow Result \rightarrow \mathrm{seq}(binaryOperator \times refTarget) \rightarrow Result$
---
$\forall\, ss : Substrate;\ r : Result;\ opt : \mathrm{seq}(binaryOperator \times refTarget) \bullet$
    $evalRefinement\ ss\ r\ opt =$
        **if** $opt = \langle\rangle$
            **then** $r$
        **else if** $first\ (head\ opt) = conjunction$
            **then** $evalRefinement\ ss\ (intersect\ r (i\_refTarget'\ ss\ (second\ (head\ opt))))\ (tail\ opt)$
        **else if** $first\ (head\ opt) = disjunction$
            **then** $evalRefinement\ ss\ (union\ r (i\_refTarget'\ ss\ (second\ (head\ opt))))\ (tail\ opt)$
        **else if** $first\ (head\ opt) = exclusion$
            **then** $evalRefinement\ ss\ (minus\ r (i\_refTarget'\ ss\ (second\ (head\ opt))))\ (tail\ opt)$
        **else** $error\ unknownOperation$
```

```
$i\_refTarget' : Substrate \rightarrow refTarget \rightarrow Result$
```

4.9 expressionConstraintValue

```
expressionConstraintValue = simpleExpressionConstraint /
       "(" (refinedExpressionConstraint /
       compoundExpressionConstraint) ")"
```

$expressionConstraintValue ::= ecvsec\langle\langle simpleExpressionConstraint\rangle\rangle\ |$
$\quad ecvrec\langle\langle refinedExpressionConstraint'\rangle\rangle\ |$
$\quad ecvcec\langle\langle compoundExpressionConstraint\rangle\rangle$

$[refinedExpressionConstraint']$

$$
\begin{array}{l}
i\_expressionConstraintValue : Substrate \rightarrow \\
\qquad expressionConstraintValue \rightarrow Result \\
\hline
\forall ss : Substrate;\ ecv : expressionConstraintValue \bullet \\
i\_expressionConstraintValue\ ss\ ecv = \\
\quad \textbf{if } ecvsec^{\sim} ecv \in simpleExpressionConstraint \\
\qquad \textbf{then } i\_simpleExpressionConstraint\ ss\ (ecvsec^{\sim} ecv) \\
\quad \textbf{else if } ecvrec^{\sim} ecv \in refinedExpressionConstraint' \\
\qquad \textbf{then } i\_refinedExpressionConstraint'\ ss\ (ecvrec^{\sim} ecv) \\
\quad \textbf{else if } ecvcec^{\sim} ecv \in compoundExpressionConstraint \\
\qquad \textbf{then } i\_compoundExpressionConstraint\ ss\ (ecvcec^{\sim} ecv) \\
\quad \textbf{else } error\ unknownOperation
\end{array}
$$

$$
\begin{array}{l}
i\_refinedExpressionConstraint' : \\
\qquad Substrate \nrightarrow refinedExpressionConstraint' \nrightarrow Result
\end{array}
$$

4.10 Attributes and Attribute Groups

The interpretation of an attribute. `attributeOperator` and `attributeName` determines the set of possible attributes in the substrate relationship table. `reverseFlag` and `expressionConstraintValue` determine the set of candidate targets (if `reverseFlag` is absent) or `subjects` (if `reverseFlag` is present).

`cardinality` determines the minimum and maximum matches. In all cases, only a subset of the subjects (targets if `reverseFlag` is present) in the substrate relationship table will be returned in the interpretation.

```
attribute = [cardinality] [reverseFlag] [attributeOperator] attributeName
       (comparisonOperator concreteValue / "=" expressionConstraintValue )
attributeGroup = "{" attributeSet "}"
attributeSet = attribute [binaryOperator  attributeSet] / "(" attributeSet ")"
cardinality = "[" nonNegativeIntegerValue to (nonNegativeIntegerValue / many) "]"
comparisonOperator = "=" / "!" ws "=" /
    ("n"/"N") ("o"/"O") ("t"/"T") ws "=" / "<=" / "<" / ">=" /  ">"
```

4.11 AttributeName and AttributeSubject

i_attributeName verifies that *conceptReference* is a valid concept in the substrate and interprets it as a set of (exactly one) *ATTRIBUTE*

i_attributeSubject interprets the constraintOperator, if any in the context of the attribute name and interprets it as a set of *ATTRIBUTE*s

Question: Should we add any additional validation checks?

$attributeName == conceptReference$

$attributeOperator == \{descendantOrSelfOf, descendantOf\}$
$attributeSubject == attributeOperator[0 \mathinner{\ldotp\ldotp} 1] \times attributeName$

$$\begin{array}{|l}
i\_attributeName : Substrate \pfun attributeName \pfun Result \\
i\_attributeSubject : Substrate \pfun attributeSubject \pfun Result \\
\hline
\forall ss : Substrate;\ an : attributeName \bullet i\_attributeName\ ss\ an = \\
\quad i\_conceptReference\ ss\ an \\
\forall ss : Substrate;\ as : attributeSubject \bullet i\_attributeSubject\ ss\ as = \\
\quad i\_constraintOperator\ ss\ (first\ as)\ (i\_attributeName\ ss\ (second\ as))
\end{array}$$

4.12 Attribute

`attribute` consists of an optional `cardinality` and an `attributeConstraint`.

$$unlimitedNat ::= num\langle\langle \mathbb{N} \rangle\rangle \mid many$$
$$cardinality == \mathbb{N} \times unlimitedNat$$

$$\begin{array}{|l}
\textit{attribute} \\
\hline
card : cardinality[0 \mathrel{..} 1] \\
attw : attributeConstraint
\end{array}$$

attributeConstraint is either an attribute expression constraint or a concrete value constraint.

$$attributeConstraint ::= aec\langle\langle attributeExpressionConstraint \rangle\rangle \mid \\ \quad acvc\langle\langle attributeConcreteValueConstraint \rangle\rangle$$

`attributeExpressionEonstraint` is a combination of a subject constraint (target if reverse flag is true) and an expression constraint value whose interpretation yields a set of SCTID's that filter the target (subject if reverse flag).

$$[reverseFlag]$$

$$\begin{array}{|l}
\textit{attributeExpressionConstraint} \\
\hline
a : attributeSubject \\
rf : reverseFlag[0 \mathrel{..} 1] \\
cs : expressionConstraintValue
\end{array}$$

A concrete value constraint is the combination of a subject constraint and a comparison operator/concrete value whose interpretation yields a set of subject ids. The intersection of the subject and concrete value interpretation is the interpretation of `attributeConcreteValueConstraint`

$$comparisonOperator ::= eq \mid neq \mid gt \mid ge \mid lt \mid le$$

$$\begin{array}{|l}
\multicolumn{1}{l}{\textit{attributeConcreteValueConstraint}} \\
\hline
a : attributeSubject \\
op : comparisonOperator \\
v : concreteValue \\
\hline
\end{array}$$

The interpretation of an attribute is the interpretation of the cardinality applied against the underlying attribute constraint, producing a set of SCTID's or an error condition

The interpretation of an attribute constraint is the interpretation of either the attribute expression constraint or the concrete expression constraint that produces a set of Quads or an error condition.

The actual interpretations if attribute expression and concrete value are in Section 5

$$\begin{array}{|l}
i\_attribute : \\
\qquad Substrate \nrightarrow attribute \nrightarrow Result \\
i\_attributeConstraint : \\
\qquad Substrate \nrightarrow attributeConstraint \nrightarrow Quads \\
i\_attributeExpressionConstraint : \\
\qquad Substrate \nrightarrow attributeExpressionConstraint \nrightarrow Quads \\
i\_attributeConcreteValueConstraint : \\
\qquad Substrate \nrightarrow attributeConcreteValueConstraint \nrightarrow Quads \\
\hline
\forall ss : Substrate;\ a : attribute \bullet \\
\quad i\_attribute\ ss\ a = i\_cardinality\ a.card\ (i\_attributeConstraint\ ss\ a.attw) \\
\\
\forall ss : Substrate;\ aw : attributeConstraint \bullet i\_attributeConstraint\ ss\ aw = \\
\quad \mathbf{if}\ aec^\sim aw \in attributeExpressionConstraint \\
\qquad \mathbf{then}\ i\_attributeExpressionConstraint\ ss\ (aec^\sim aw) \\
\quad \mathbf{else\ if}\ acvc^\sim aw \in attributeConcreteValueConstraint \\
\qquad \mathbf{then}\ i\_attributeConcreteValueConstraint\ ss\ (acvc^\sim aw) \\
\quad \mathbf{else}\ qerror\ unknownOperation \\
\\
\forall ss : Substrate;\ aec : attributeExpressionConstraint;\ eca : expressionConstraintArgs \mid \\
\quad eca.atts = i\_attributeSubject\ ss\ aec.a \land \\
\quad eca.rf = aec.rf \land \\
\quad eca.subjOrTarg = i\_expressionConstraintValue\ ss\ aec.cs \bullet \\
\quad i\_attributeExpressionConstraint\ ss\ aec = i\_attributeExpression\ ss\ eca \\
\\
\forall ss : Substrate;\ awc : attributeConcreteValueConstraint;\ aca : concreteConstraintArgs \mid \\
\quad aca.atts = i\_attributeSubject\ ss\ awc.a \land \\
\quad aca.op = awc.op \land \\
\quad aca.t = awc.v \bullet \\
\quad i\_attributeConcreteValueConstraint\ ss\ awc = i\_concreteAttributeConstraint\ ss\ aca \\
\hline
\end{array}$$

4.13 AttributeSet and AttributeGroup

An `attributeGroup` is an `attributeSet`. An `attributeSet` consists of a sequence of one or more attribute constraints joined by `binaryOperator`s.

$$\begin{array}{l} \mathit{binaryOperator} ::= \mathit{conjunction} \mid \mathit{disjunction} \mid \mathit{exclusion} \\ \mathit{attributeGroup} == \mathit{attributeSet} \\ \mathit{attributeSet} == \mathit{attribute} \times \operatorname{seq}(\mathit{binaryOperator} \times \mathit{attribute}) \end{array}$$

$$\begin{array}{l} \mathit{i\_attributeGroup} : \mathit{Substrate} \nrightarrow \mathit{attributeGroup} \nrightarrow \mathit{Result} \\ \mathit{i\_attributeSet} : \mathit{Substrate} \nrightarrow \mathit{attributeSet} \nrightarrow \mathit{Result} \\ \hline \forall\, \mathit{ss} : \mathit{Substrate};\ \mathit{ag} : \mathit{attributeGroup} \bullet \mathit{i\_attributeGroup}\ \mathit{ss}\ \mathit{ag} = \\ \quad \mathit{i\_attributeSet}\ \mathit{ss}\ \mathit{ag} \\ \forall\, \mathit{ss} : \mathit{Substrate};\ a : \mathit{attributeSet} \bullet \mathit{i\_attributeSet}\ \mathit{ss}\ a = \\ \quad \mathit{result}\ (\mathit{evalCmpndAtt}\ \mathit{ss}\ (\mathit{i\_groupedAttribute}\ \mathit{ss}\ (\mathit{first}\ a))\ (\mathit{second}\ a)) \end{array}$$

4.14 Compound attribute evaluation

The left-to-right evaluation of `attributeSet`. The interpretation takes a *lhs* as a function from *SCTID* to *GROUP* and a *rhs* which is sequence of operator/attribute tuples and recursively interprets the *lhs* and interpretation of the head of the *rhs*.

$$\begin{array}{l}
\mathit{evalCmpndAtt} : \\
\qquad \mathit{Substrate} \to \mathit{IDGroups} \to \mathrm{seq}(\mathit{binaryOperator} \times \mathit{attribute}) \to \mathit{IDGroups} \\
\mathit{gintersect}, \mathit{gunion}, \mathit{gminus}, \mathit{gfirstError} : \\
\qquad \mathit{IDGroups} \to \mathit{IDGroups} \to \mathit{IDGroups} \\
\hline
\forall\, ss : \mathit{Substrate};\ lhs : \mathit{IDGroups};\ rhs : \mathrm{seq}(\mathit{binaryOperator} \times \mathit{attribute}) \bullet \\
\mathit{evalCmpndAtt}\ ss\ lhs\ rhs = \\
\mathbf{if}\ rhs = \langle\rangle \\
\quad \mathbf{then}\ lhs \\
\mathbf{else\ if}\ \mathit{first}\ (\mathit{head}\ rhs) = \mathit{conjunction} \\
\mathbf{then}\ \mathit{evalCmpndAtt}\ ss\ (\mathit{gintersect}\ lhs (\mathit{i\_groupedAttribute}\ ss\ (\mathit{second}\ (\mathit{head}\ rhs))))\ (\mathit{tail}\ rhs) \\
\mathbf{else\ if}\ \mathit{first}\ (\mathit{head}\ rhs) = \mathit{disjunction} \\
\mathbf{then}\ \mathit{evalCmpndAtt}\ ss\ (\mathit{gunion}\ lhs (\mathit{i\_groupedAttribute}\ ss\ (\mathit{second}\ (\mathit{head}\ rhs))))\ (\mathit{tail}\ rhs) \\
\mathbf{else\ if}\ \mathit{first}\ (\mathit{head}\ rhs) = \mathit{exclusion} \\
\mathbf{then}\ \mathit{evalCmpndAtt}\ ss\ (\mathit{gminus}\ lhs (\mathit{i\_groupedAttribute}\ ss\ (\mathit{second}\ (\mathit{head}\ rhs))))\ (\mathit{tail}\ rhs) \\
\mathbf{else}\ \mathit{gerror}\ \mathit{unknownOperation} \\
\forall\, a, b, r : \mathit{IDGroups} \mid \\
\quad r = \mathbf{if}\ \mathit{gerror}^\sim a \in \mathit{ERROR} \lor \mathit{gerror}^\sim b \in \mathit{ERROR}\ \mathbf{then}\ \mathit{gfirstError}\ a\ b \\
\quad \mathbf{else}\ \mathit{gv}\ (\mathit{gv}^\sim a \cap \mathit{gv}^\sim b) \bullet \\
\quad \mathit{gintersect}\ a\ b = r \\
\forall\, a, b, r : \mathit{IDGroups} \mid \\
\quad r = \mathbf{if}\ \mathit{gerror}^\sim a \in \mathit{ERROR} \lor \mathit{gerror}^\sim b \in \mathit{ERROR}\ \mathbf{then}\ \mathit{gfirstError}\ a\ b \\
\quad \mathbf{else}\ \mathit{gv}\ (\mathit{gv}^\sim a \cup \mathit{gv}^\sim b) \bullet \\
\quad \mathit{gunion}\ a\ b = r \\
\forall\, a, b, r : \mathit{IDGroups} \mid \\
\quad r = \mathbf{if}\ \mathit{gerror}^\sim a \in \mathit{ERROR} \lor \mathit{gerror}^\sim b \in \mathit{ERROR}\ \mathbf{then}\ \mathit{gfirstError}\ a\ b \\
\quad \mathbf{else}\ \mathit{gv}\ (\mathit{gv}^\sim a \setminus \mathit{gv}^\sim b) \bullet \\
\quad \mathit{gminus}\ a\ b = r
\end{array}$$

$$\begin{array}{l}
\mathit{i\_groupedAttribute} : \\
\qquad \mathit{Substrate} \to \mathit{attribute} \to \mathit{IDGroups} \\
\hline
\forall\, ss : \mathit{Substrate};\ a : \mathit{attribute} \bullet \\
\quad \mathit{i\_groupedAttribute}\ ss\ a = \mathit{i\_groupCardinality}\ (\mathit{i\_attributeConstraint}\ ss\ a.\mathit{attw})\ a.\mathit{card}
\end{array}$$

4.15 Group Cardinality

The interpretation of cardinality within a group impose additional constraints:

- $[0 \mathinner{..} n]$ – the set of all substrate concept codes that have at least one group (entry) in the substrate relationships and, at most n matching entries in the same group

- $[0\,..\,0]$ – the set of all substrate concept codes that have at least one group (entry) in the substrate relationships and *no* matching entries
- [1..∗] – (default) at least one matching entry in the substrate relationships
- $[m_1\,..\,n_1]op[m_2\,..\,n_2]...$ – set of substrate concept codes where there exists at least one group where all conditions are simultaneously true

The interpretation of a grouped cardinality is a function from a set of SCTID's to the groups in which they were qualified.

The algorithm below partitions the input set of Quads by group and validates the cardinality on a per-group basis. Groups that pass are returned

TODO: This assumes that q.t is always type object. It doesn't say what to do if it is concrete **TODO:** the *gresult* function seems to express what is described below more simply

$$\begin{array}{l}
\mathit{i\_groupCardinality} : \\
\qquad \mathit{Quads} \rightarrow \mathit{cardinality}[0\,..\,1] \rightarrow \mathit{IDGroups} \\
\hline
\forall\, \mathit{quads} : \mathit{Quads};\ \mathit{oc} : \mathit{cardinality}[0\,..\,1];\ \mathit{uniqueGroups} : \mathbb{P}\,\mathit{GROUP}; \\
\qquad \mathit{quadsByGroup} : \mathit{GROUP} \mathrel{+\!\!\!\rightarrow} \mathbb{P}\,\mathit{Quad} \mid \\
\quad \mathit{uniqueGroups} = \{ g : \mathit{qids}\ \mathit{quads} \bullet q.g \} \land \\
\quad \mathit{quadsByGroup} = \{ g : \mathit{uniqueGroups};\ q : \mathbb{P}\,\mathit{Quad} \mid \\
\qquad q = \{ e : \mathit{qids}\ \mathit{quads} \mid e.g = g \} \bullet g \mapsto (\mathit{evalCardinality}\ \mathit{oc}\ q) \} \bullet \\
\mathit{i\_groupCardinality}\ \mathit{quads}\ \mathit{oc} = \\
\quad \mathit{gv}\, \{ \mathit{sctid} : \mathit{SCTID};\ \mathit{groups} : \mathbb{P}\,\mathit{GROUP} \mid \mathit{sctid} \in \{ q : \bigcup(\mathrm{ran}\ \mathit{quadsByGroup}) \bullet \\
\qquad \mathbf{if}\ \mathit{qidd}\ \mathit{quads} = \mathit{subjects}\ \mathbf{then}\ q.s\ \mathbf{else}\ \mathit{object}^{\sim}\, q.t \} \land \\
\quad \mathit{groups} = \{ g : \mathrm{dom}\ \mathit{quadsByGroup} \mid (\exists\, q : \mathit{quadsByGroup}\ g \bullet \\
\qquad \mathit{sctid} = \mathbf{if}\ \mathit{qidd}\ \mathit{quads} = \mathit{subjects}\ \mathbf{then}\ q.s\ \mathbf{else}\ \mathit{object}^{\sim}\, q.t) \} \bullet \\
\qquad \mathit{sctid} \mapsto \mathit{groups} \}
\end{array}$$

4.16 Cardinality

Interpretation: *cardinality* is tested against a set of quads with the following rules:

1. Errors are propagated
2. No cardinality or passing cardinality returns the subjects / targets of the set of quads
3. Otherwise return an empty set

$$\begin{array}{l}
\mathit{i\_cardinality} : \\
\qquad \mathit{cardinality}[0\,..\,1] \rightarrow \mathit{Quads} \rightarrow \mathit{Result} \\
\hline
\forall\, \mathit{oc} : \mathit{cardinality}[0\,..\,1];\ \mathit{qr} : \mathit{Quads} \bullet \\
\mathit{i\_cardinality}\ \mathit{oc}\ \mathit{qr} = \\
\quad \mathbf{if}\ \mathit{qerror}^{\sim}\, \mathit{qr} \in \mathit{ERROR}\ \mathbf{then}\ \mathit{result}\,(\mathit{gresult}\ \mathit{qr}) \\
\quad \mathbf{else}\ \mathit{result}\,(\mathit{gresult}\,(\mathit{qv}\,(\mathit{evalCardinality}\ \mathit{oc}\,(\mathit{qids}\ \mathit{qr}), \mathit{qidd}\ \mathit{qr})))
\end{array}$$

evalCardinality Evaluate the cardinality of a set, returning the set if it meets the constraints, otherwise return the empty set.

$$evalCardinality : cardinality[0\,..\,1] \to \mathbb{P}\,Quad \to \mathbb{P}\,Quad$$

```
$\forall\, oc : cardinality[0\,..\,1];\ s : \mathbb{P}\,Quad \bullet evalCardinality\ oc\ s =$
    $\textbf{if}\ \#oc = 0\ \textbf{then}\ s$
    $\textbf{else if}\ (\#s \geq first\,(head\ oc)) \wedge$
        $(second\,(head\ oc) = many \vee num^{\sim}(second\,(head\ oc)) \leq \#s)$
    $\textbf{then}\ s$
    $\textbf{else}\ \emptyset$
```

5 Substrate Interpretations

This section defines the interpretations that are realized against the substrate.

5.1 AttributeExpressionConstraint

expressionConstraintArgs carries the arguments necessary to evaluate an attribute expression

- **rf** – If present, return subjects matching attribute/target subjects
- **subjOrTarg** – Set of subject or target SCTIDS (or error)
- **atts** – Set of attributes to evaluate

expressionConstraintArgs

$rf : reverseFlag[0\,..\,1]$
$subjOrTarg : Result$
$atts : Result$

Evaluate an attribute expression, returning the set of quads in the substrate relationship table with the supplied subject / attribute if *rf* is absent and attribute / target if *rf* is present. The interpretation returns an error there is an error in either the subject/targets or attributes. **TODO:** We don't say what to do if the target (t) is concrete

```
$i\_attributeExpression :$
        $Substrate \rightarrow expressionConstraintArgs \rightarrow Quads$
-----
$\forall ss : Substrate;\ args : expressionConstraintArgs;$
    $sti : \mathbb{P}\ TARGET;\ atts : \mathbb{P}\ ATTRIBUTE \mid$
$sti = ok^{\sim} args.subjOrTarg \land atts = ok^{\sim} args.atts \bullet$
$i\_attributeExpression\ ss\ args =$
    **if** $error^{\sim}(bigunion\{args.subjOrTarg, args.atts\}) \in ERROR$
    **then**
        $qfirstError\{args.subjOrTarg, args.atts\}$
    **else if** $\neg\ (sti \cup atts) \subseteq ss.c$
    **then**
        $qerror\ unknownConceptReference$
    **else if** $\#args.rf = 0$
    **then**
        $qv(\{s : SUBJECT;\ t : sti;\ a : atts;\ g : GROUP;\ re : ss.r \mid$
            $re.t = t \land re.a = a \bullet re\}, subjects)$
    **else**
        $qv(\{s : sti;\ a : atts;\ t : TARGET;\ g : GROUP;\ re : ss.r \mid$
            $re.s = object^{\sim} t \land re.a = a \bullet re\}, targets)$
```

5.2 ConcreteAttributeConstraint

```
concreteValue =  QM stringValue QM / "#" numericValue
stringValue = 1*(anyNonEscapedChar / escapedChar)
numericValue = decimalValue / integerValue
integerValue = ( ["-"/"+"] digitNonZero *digit ) / zero
```

The mechanism for interpreting *concreteValue* is not fully specified at this point. Much of the rest of the machinery is focused around interpreting constraints in the context of quads - subject, attribute, target and group, so the function is currently defined as returning a *Quads*, but another type or interpretation may be in order.

$$concreteValue ::= stringValue \mid integerValue \mid realValue$$

concreteConstraintArgs The arguments to the concrete value function.

- **atts** – list of attributes to test
- **op** – operator
- **t** – value to test against

Note: *reverseFlag* has no meaning for concrete constraints.

```
$concreteConstraintArgs$
$atts : Result$
$op : comparisonOperator$
$t : concreteValue$
```

$$\begin{array}{l} \mathit{i\_concreteAttributeConstraint} : \\ \qquad \mathit{Substrate} \rightarrow \mathit{concreteConstraintArgs} \rightarrow \mathit{Quads} \end{array}$$

5.3 ConstraintOperator

```
constraintOperator =
        descendantOrSelfOf / descendantOf /  ancestorOrSelfOf / ancestorOf
```

Interpretation: Apply the substrate descendants (*descs*) or ancestors (*ancs*) function to a set of SCTID's in the supplied *Result*. Error conditions are propagated.

$$\begin{array}{l} \mathit{i\_constraintOperator} : \\ \qquad \mathit{Substrate} \nrightarrow \mathit{constraintOperator}[0 \mathrel{..} 1] \nrightarrow \mathit{Result} \nrightarrow \mathit{Result} \\ \mathit{completeFun} : (\mathit{SCTID} \nrightarrow \mathbb{P}\,\mathit{SCTID}) \rightarrow \mathit{SCTID} \rightarrow \mathbb{P}\,\mathit{SCTID} \end{array}$$

```
$\forall ss : Substrate;\ oco : constraintOperator[0 \mathrel{..} 1];\ refset : Result \bullet$
$i\_constraintOperator\ ss\ oco\ refset =$
    if $error^{\sim} refset \in ERROR \lor \#oco = 0$
        then $refset$
    else if $head\ oco = descendantOrSelfOf$
        then $ok(\bigcup\{id : ids\ refset \bullet completeFun\ ss.descs\ id \cup \{id\}\})$
    else if $head\ oco = descendantOf$
        then $ok(\bigcup\{id : ids\ refset \bullet completeFun\ ss.descs\ id \setminus \{id\}\})$
    else if $head\ oco = ancestorOrSelfOf$
        then $ok(\bigcup\{id : ids\ refset \bullet completeFun\ ss.ancs\ id \cup \{id\}\})$
    else if $head\ oco = ancestorOf$
        then $ok(\bigcup\{id : ids\ refset \bullet completeFun\ ss.ancs\ id \setminus \{id\}\})$
    else $error\ unknownOperation$

$\forall f : (SCTID \nrightarrow \mathbb{P}\,SCTID);\ id : SCTID \bullet completeFun\ f\ id =$
    if $id \in \operatorname{dom} f$ then $f\ id$ else $\emptyset$
```

5.4 FocusConcept

```
focusConcept = [memberOf] conceptReference
```

Interpretation: focusConcept is interpreted as the interpretation of conceptReference itself or, if memberOf is specified, the interpretation of the memberOf function applied to the conceptReference

$$\mathit{focusConcept} ::= \mathit{mo}\langle\!\langle \mathit{conceptReference} \rangle\!\rangle \mid \mathit{cr}\langle\!\langle \mathit{conceptReference} \rangle\!\rangle$$

```
$i\_focusConcept : Substrate \rightarrow focusConcept \rightarrow Result$
---
$\forall ss : Substrate;\ fc : focusConcept \bullet$
$i\_focusConcept\ ss\ fc =$
    $\textbf{if}\ cr^{\sim}fc \in conceptReference$
        $\textbf{then}\ i\_conceptReference\ ss\ (cr^{\sim}fc)$
    $\textbf{else if}\ mo^{\sim}fc \in conceptReference$
        $\textbf{then}\ i\_memberOf\ ss\ (mo^{\sim}fc)$
    $\textbf{else}\ error\ unknownOperation$
```

5.4.1 memberOf

Interpretation: memberOf is interpreted as application *refset* function to the conceptReference SCTID or an error if the SCTID isn't in the domain of the *refset* function.

```
$i\_memberOf : Substrate \rightarrow conceptReference \rightarrow Result$
---
$\forall ss : Substrate;\ cr : conceptReference \bullet$
$i\_memberOf\ ss\ cr =$
    $(\textbf{let}\ sctid == toSCTID\ cr \bullet$
        $\textbf{if}\ sctid \notin \mathrm{dom}\ ss.refset\ \textbf{then}\ error\ unknownRefsetId$
    $\textbf{else}\ ok\ (ss.refset\ sctid))$
```

5.5 ConceptReference

```
conceptReference = conceptId [ "|" Term "|"]
conceptId = sctId
```

Interpretation: conceptReference is interpreted as *SCTID* if it is a member the list of concepts, *c* in the substrate, otherwise as an *error*.

$[conceptReference]$

```
$toSCTID : conceptReference \rightarrow SCTID$
$i\_conceptReference : Substrate \rightarrow conceptReference \rightarrow Result$
---
$\forall ss : Substrate;\ c : conceptReference \bullet i\_conceptReference\ ss\ c =$
    $\textbf{if}\ (toSCTID\ c) \in ss.c\ \textbf{then}\ ok\{(toSCTID\ c)\}$
    $\textbf{else}\ error\ unknownConceptReference$
```

6 Glue and Helper Functions

This section carries various type transformations and error checking functions

6.1 Types

- **direction** – an indicator whether a collection of quads was determined as subject to target (*subjects*) or target to subject (*targets*)
- **Quads** – a collection of *Quad*s or an error condition. If it is a collection *Quad*s, it also carries a direction indicator that determines whether it represents a set of subjects or targets.
- **IDGroups** – a map from *SCTID*s to the *GROUP* they were in when they passed if successful, otherwise an error indication.

$direction ::= subjects \mid targets$
$Quads ::= qv\langle\langle \mathbb{P}\, Quad \times direction \rangle\rangle \mid qerror\langle\langle ERROR \rangle\rangle$

$IDGroups ::= gv\langle\langle SCTID \nrightarrow \mathbb{P}\, GROUP \rangle\rangle \mid gerror\langle\langle ERROR \rangle\rangle$

6.2 Result transformations

- **ids** – return the set of *SCTID*s in a result or the empty set if there is an error
- **qids** – return the set of *Quads* in a quads result or an empty set if there is an error
- **qidd** – return the direction of a *Quads* result. Undefined if error
- **gids** – return the *SCTID* to group map in an id group or an empty map if there is error
- **gresult** – convert a set of quads int a set of id groups
- **result** – convert a set of id groups into a simple result.

$$
\begin{array}{l}
ids : Result \rightarrow \mathbb{P}\, SCTID \\
qids : Quads \rightarrow \mathbb{P}\, Quad \\
qidd : Quads \nrightarrow direction \\
gids : IDGroups \rightarrow SCTID \rightarrow \mathbb{P}\, GROUP \\
gresult : Quads \rightarrow IDGroups \\
result : IDGroups \rightarrow Result \\
\hline
\forall\, r : Result \bullet ids\ r = \\
\quad \textbf{if } error^{\sim} r \in ERROR \textbf{ then } \emptyset \\
\quad \textbf{else } ok^{\sim} r \\[1ex]
\forall\, q : Quads \bullet qids\ q = \\
\quad \textbf{if } qerror^{\sim} q \in ERROR \textbf{ then } \emptyset \\
\quad \textbf{else } first\,(qv^{\sim} q) \\[1ex]
\forall\, q : Quads \bullet qidd\ q = \\
\quad second\,(qv^{\sim} q) \\[1ex]
\forall\, g : IDGroups \bullet gids\ g = \\
\quad \textbf{if } gerror^{\sim} g \in ERROR \textbf{ then } \emptyset \\
\quad \textbf{else } gv^{\sim} g \\[1ex]
\forall\, q : Quads \bullet gresult\ q = \\
\quad \textbf{if } qerror^{\sim} q \in ERROR \textbf{ then } gerror(qerror^{\sim} q) \\
\quad \textbf{else if } qidd\ q = subjects \\
\qquad \textbf{then } gv\,\{s : SCTID \mid (\exists\, qr : qids\ q \bullet s = qr.s) \bullet \\
\qquad s \mapsto \{qr : qids\ q \bullet qr.g\}\} \\
\quad \textbf{else} \\
\qquad gv\,\{t : SCTID \mid (\exists\, qr : qids\ q \bullet t = object^{\sim} qr.t) \bullet \\
\qquad t \mapsto \{qr : qids\ q \bullet qr.g\}\} \\[1ex]
\forall\, g : IDGroups \bullet result\ g = \\
\quad \textbf{if } gerror^{\sim} g \in ERROR \textbf{ then } error(gerror^{\sim} g) \\
\quad \textbf{else } ok(\mathrm{dom}(gv^{\sim} g))
\end{array}
$$

Definition of the various functions that are performed on the result type.

- **firstError** – take a set of *Result*s with one or more errors and return an aggregation / summary as a *Result*. (Not fully defined)
- **qfirstError** – take a set of *Result*s with one or more errors and return an aggregation / summary as a *Quads* instance. (not fully defined)

- **union** – return the union of two *Result SCTID*s or *ERROR* if either of them have an error.
- **intersect** – return the intersection of two *Result SCTID*s or *ERROR* if either of them have an error.

- **minus** – return the set of *SCTID*s in the first set of results that are not in the second or *ERROR* if either of them have an error.
- **bigunion** – return the union of a set *Result SCTID*s or *ERROR* if any of the results in the set have an error.
- **bigintersect** – return the intersection of a set *Result SCTID*s or *ERROR* if any of the results in the set have an error.

firstError is not fully defined – it takes a set of results with one or more errors and returns an aggregation of all of the errors.

$$\begin{array}{|l}
\hline
\mathit{union}, \mathit{intersect}, \mathit{minus} : \mathit{Result} \rightarrow \mathit{Result} \rightarrow \mathit{Result} \\
\mathit{bigunion}, \mathit{bigintersect} : \mathbb{P}\,\mathit{Result} \rightarrow \mathit{Result} \\
\mathit{firstError} : \mathbb{P}\,\mathit{Result} \nrightarrow \mathit{Result} \\
\mathit{qfirstError} : \mathbb{P}\,\mathit{Result} \nrightarrow \mathit{Quads} \\
\hline
\forall\, x, y : \mathit{Result} \bullet \mathit{union}\; x\; y = \\
\quad \textbf{if } \mathit{error}^\sim x \in \mathit{ERROR} \textbf{ then } x \\
\quad \textbf{else if } \mathit{error}^\sim y \in \mathit{ERROR} \textbf{ then } y \\
\quad \textbf{else } \mathit{ok}\,((\mathit{ok}^\sim x) \cup (\mathit{ok}^\sim y)) \\
\forall\, x, y : \mathit{Result} \bullet \mathit{intersect}\; x\; y = \\
\quad \textbf{if } \mathit{error}^\sim x \in \mathit{ERROR} \textbf{ then } x \\
\quad \textbf{else if } \mathit{error}^\sim y \in \mathit{ERROR} \textbf{ then } y \\
\quad \textbf{else } \mathit{ok}\,((\mathit{ok}^\sim x) \cap (\mathit{ok}^\sim y)) \\
\forall\, x, y : \mathit{Result} \bullet \mathit{minus}\; x\; y = \\
\quad \textbf{if } \mathit{error}^\sim x \in \mathit{ERROR} \textbf{ then } x \\
\quad \textbf{else if } \mathit{error}^\sim y \in \mathit{ERROR} \textbf{ then } y \\
\quad \textbf{else } \mathit{ok}\,((\mathit{ok}^\sim x) \setminus (\mathit{ok}^\sim y)) \\
\forall\, \mathit{rs} : \mathbb{P}\,\mathit{Result} \bullet \mathit{bigunion}\; \mathit{rs} = \\
\quad \textbf{if } \exists\, r : \mathit{rs} \bullet \mathit{error}^\sim r \in \mathit{ERROR} \textbf{ then } \mathit{firstError}\; \mathit{rs} \\
\quad \textbf{else } \mathit{ok}\,(\bigcup\{r : \mathit{rs} \bullet \mathit{ids}\; r\}) \\
\forall\, \mathit{rs} : \mathbb{P}\,\mathit{Result} \bullet \mathit{bigintersect}\; \mathit{rs} = \\
\quad \textbf{if } \exists\, r : \mathit{rs} \bullet \mathit{error}^\sim r \in \mathit{ERROR} \textbf{ then } \mathit{firstError}\; \mathit{rs} \\
\quad \textbf{else } \mathit{ok}\,(\bigcap\{r : \mathit{rs} \bullet \mathit{ids}\; r\}) \\
\hline
\end{array}$$

7 Appendix 1

Representing optional elements of type T. Representing it as a sequence allows us to determine absence by $\# T = 0$ and the value by *head* T.

$$T[0 \mathbin{..} 1] == \{s : \mathrm{seq}\, T \mid \# s \leq 1\}$$