6. Considerations

Overview

In this section, we discuss a range of topics that were considered in the design of the SNOMED CT MRCM, including:

- How the MRCM will be authored and quality checked,
- How the MRCM will be used to support the use cases described in 2. Use Cases,
- How the MRCM will be versioned when corrections or enhancements are made, and
- How the MRCM can be extended and adapted for use with SNOMED CT extensions.

Authoring and Quality Checks

The MRCM has been designed to include minimal redundancy, to make the authoring and maintenance of these rules less error-prone. For example, by using separate reference sets to represent the attribute domains and ranges, each range does not need to be repeated for each of the attribute's domains.

To support implementers who prefer to process string-based representations of the rules, three MRCM attributes have been provided that concatenate information from other fields (as per the table below). These attributes have been populated automatically to avoid inconsistencies between the structural and string-based representations of the rules. The MRCM attributes that may be generated from other attributes are:

<table>
<thead>
<tr>
<th>MRCM Reference Set</th>
<th>Attribute</th>
<th>Populated From</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRCM domain reference set</td>
<td>domainTemplateForPreCoordination</td>
<td>MRCM domain reference set .domainConstraint,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MRCM attribute domain reference set .grouped,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MRCM attribute domain reference set .attributeCardinality,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MRCM attribute domain reference set .attributeInGroupCardinality,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MRCM attribute range reference set .rangeConstraint</td>
</tr>
<tr>
<td></td>
<td>attributeRule</td>
<td>WHERE .MRCM domain reference set .referencedComponentId = .MRCM attribute domain</td>
</tr>
<tr>
<td></td>
<td></td>
<td>reference set .domainId</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AND .MRCM attribute domain reference set .referencedComponentId = .MRCM attribute</td>
</tr>
<tr>
<td></td>
<td></td>
<td>range reference set .referencedComponentId</td>
</tr>
</tbody>
</table>

In addition, the quality of the international SNOMED CT MRCM is checked using a range of mechanisms, including:

- Manual review, supported by the inclusion of human-readable terms for each concept identifier in the reference sets;
- Automated testing, including checks for the following:
  - All MRCM reference sets conform to the associated Descriptor Template,
  - All concept identifiers refer to active concepts in the relevant SNOMED CT international edition,
  - All concept identifiers refer to concepts from a value set appropriate for that field,
  - All parsable strings (e.g. expression constraints) are syntactically valid (based on the associated ABNF syntax) and refer to concepts that are active in the relevant SNOMED CT international edition,
- Feedback from direct use by SNOMED International staff in the SNOMED CT authoring tools, and
- Feedback from the SNOMED CT member, vendor and user communities.

Supporting Use Cases

Another important consideration in the design of the MRCM is how it can best be used to support the range of anticipated use cases. In this section, we describe the ways in which the MRCM design supports each of the use cases presented in 2. Use Cases.

Precoordinated Content Development

As described in 2. Use Cases, one of the key use cases for the SNOMED CT MRCM is to assist with the consistent authoring and validation of SNOMED CT concepts.

To support the authoring and validation of SNOMED CT precoordinated content, the following MRCM process can be used:

1. Identify the parents of the concept being authored:
   - Please note - Either proximal primitive parents or proximal parents may be selected, depending on the authoring approach used;
2. Determine which domains the parent concepts belong to using the MRCM domain reference set.
   - If the proximal parent modelling approach is used, the stated parents should be tested for membership in each domain using the do .mainConstraint.
If the proximal primitive parent modelling approach is used, the stated parents should be tested for membership in each domain using the `proximalPrimitiveConstraint` and all required refinements in the `proximalPrimitiveRefinement` should match at least one defining relationship stated on the concept being authored;

- Please note - if any of the proximal parents belong to a given domain, then the authored concept belongs to that domain.

- Please note - irrespective of the modelling approach used, the concept being authored can be tested for membership in a domain by checking if its inferred parents (i.e. after classification) are valid against the associated `domainConstraint`.

3. Determine the set of valid attributes for the given domains using the `MRCM attribute domain reference set`; and allow defining relationships to be added from this set.

4. For each attribute used to define the concept, ensure that the grouping and cardinality are valid according to the rules specified in the `MRCM attribute domain reference set` for the given attribute and parent domain;

5. Determine the valid range for each defining attribute using the `rangeConstraint` in the `MRCM attribute range reference set`.

Please note that when the above process is being applied to the authoring of new precoordinated content, then only rules with a `contentType` = `<< 723595009` All new precoordinated SNOMED CT content may be used. However, when the process is applied to the validation of (both new and existing) precoordinated content, only rules with a `contentType` = `<< 723594008` All precoordinated SNOMED CT content may be used. Rules with a `ruleStrength` of `Mandatory concept model rule` should be enforced when authoring and cause an error during validation, while rules with a `ruleStrength` of `Optional concept model rule` should be used as a recommendation for authoring and result in a warning during validation.

An alternative approach to authoring and validating precoordinated concepts is to use the `domainTemplateForPrecoordination` from the `MRCM domain reference set` to ensure compliance with the full set of attribute rules. The `domainTemplateForPrecoordination` can also be specialized into a concept authoring template that meets the needs of a subclass of concepts, while still conforming to the overall rules of the domain.

Other features provided by the `MRCM` that are helpful for precoordinated content development include:

- The `MRCM domain reference set` and the `MRCM attribute domain reference set` use domains defined in terms of a single top-level hierarchy. This enables concepts being authored to be tested for subsumption with each domain concept, using techniques that are relatively simple to implement (e.g. using a transitive closure table).

- The `MRCM attribute range reference set` includes a set of attribute rules that combine the domain, grouping, cardinality and range constraints into a single expression constraint, which can facilitate the efficient implementation of content validation.

- The `ruleStrengthId` attribute in the `MRCM attribute domain reference set` and the `MRCM attribute range reference set` allows both mandatory rules (resulting in errors) and optional rules (resulting in warnings) to be specified.

- The `contentType` attribute in the `MRCM attribute domain reference set` and the `MRCM attribute range reference set` allows the scope of each rule to be restricted to a particular type of SNOMED CT content, including:
  
  - All SNOMED CT content – for rules which apply to both precoordinated and postcoordinated content;
  - All precoordinated SNOMED CT content – for rules which apply to all precoordinated content (but not to postcoordinated content);
  - All new precoordinated SNOMED CT content – for rules which apply to newly authored (i.e. stated) precoordinated content, which may not necessarily be valid for all existing content; and
  - All postcoordinated SNOMED CT content – for rules which apply to all postcoordinated content (but not to precoordinated content).

Expressions, Constraints and Queries

2. Use Cases describes using the SNOMED CT `MRCM` to assist the authoring and validation of SNOMED CT expressions, SNOMED CT expression constraints, and SNOMED CT queries. These languages allow one or more focus concepts to be defined, which when tested against a domain in the `MRCM` enables possible attribute refinements, cardinalities and valid ranges to be found.

To support the authoring and validation of SNOMED CT postcoordinated expressions the following `MRCM` process can be used:

1. Determine which domains the expression belongs to using the `MRCM domain reference set`;
   a. An expression belongs to a given domain if any focus concept is either:
      a. Valid when tested against the associated `domainConstraint`, or
      b. Valid when tested against the associated `proximalPrimitiveConstraint` and all required refinements in the `proximalPrimitiveRefinement` either match a defining relationship on the given focus concept, or match a refinement condition added to the expression being authored;
   b. Determine the set of valid attributes for the given domains using the `MRCM attribute domain reference set`; and allow refinements to be added using attribute concepts from this set;
   c. For each attribute used to define the concept, ensure that the grouping and cardinality are valid according to the rules specified in the `MRCM attribute domain reference set` for the given attribute and parent domain;
   d. Determine the valid range for each attribute using the `rangeConstraint` in the `MRCM attribute range reference set`;

Please note that in the above process only rules with a `contentType` = `<< 723595009` All postcoordinated SNOMED CT content should be used. Rules with a `ruleStrength` of `Mandatory concept model rule` should be enforced when authoring and cause an error during validation, while rules with a `ruleStrength` of `Optional concept model rule` should be used as a recommendation for authoring and result in a warning during validation.

An alternative approach to authoring and validating postcoordinated expressions is to use the `domainTemplateForPostcoordination` from the `MRCM domain reference set` to ensure compliance with the full set of attribute rules. The `domainTemplateForPostcoordination` can also be specialized into an expression authoring template that meets the needs of a particular use case, while still conforming to the overall rules of the domain.

To support the authoring and validation of SNOMED CT expression constraints and queries the following `MRCM` process can be used:

1. Determine which domains the constraint or query belongs to using the `MRCM domain reference set`;
   a. An expression constraint or query belongs to a given domain if either:
      a. Any focus concept is valid when tested against the associated `proximalPrimitiveConstraint`;
      b. The constraint or query belongs to a parent domain of the given domain.
   b. Please note: The concept model requirements for constraints and queries are more relaxed than those for precoordinated expressions, because their role is to define the membership of a set of concepts or expressions, rather than restrict the manner in which a single concept or expression may be defined.
2. Determine the set of valid attributes for the given domains using the MRCM attribute domain reference set, and allow refinements to be added using attribute concepts from this set;

3. For each attribute used to define the concept, ensure that the grouping and cardinality are consistent with the rules specified in the MRCM attribute domain reference set for the given attribute and parent domain. In particular:
   - An attribute should only be grouped in an expression constraint or query if it specified as grouped = 1 for that domain in the MRCM attribute reference set. Any attribute can appear as ungrouped in an expression constraints.
   - A cardinality constraint used in an expression constraint or query should be the same, or stricter than the cardinality specified for the given attribute and domain in the MRCM attribute domain reference set.

4. Determine the valid range for each attribute using the MRCM attribute range reference set.

Please note that if the expression constraint or query is being applied to a substrate that includes only precoordinated content, then only the MRCM rules with a contentType = << 723594008. All precoordinated SNOMED CT content should be used. However, if the expression constraint or query is being applied to a substrate that may include both precoordinated and postcoordinated content, then the less restrictive rules (e.g. with broader ranges) that apply where contentType = << 7235950009. All postcoordinated SNOMED CT content may be used. Rules with a ruleStrength of mandatory concept model rule should be enforced when authoring and cause an error during validation, while rules with a ruleStrength of optional concept model rule should be used as a recommendation for authoring and result in a warning during validation.

**Natural Language Processing**

The MRCM can also be used to support Natural Language Processing (NLP), as described in 2. Use Cases. When terms used in free text are associated with a particular SNOMED CT concept, this concept can be tested for membership in specific MRCM domains using the process above for authoring postcoordinated expressions. As per this process, if the concept is found to belong to one of the domains, then the MRCM attribute domain reference set can be used to determine its possible attributes, and the MRCM attribute range reference set used to determine the valid values of these attributes. This process can thus be used to indicate possible ways in which the discovered concept may be linked (via suitable SNOMED CT attributes) to concepts found in the surrounding text.

Optimizations to the MRCM rules may also be adopted by NLP implementations to simplify the testing of membership in a potential range. For example, each range constraint could be split into its separate subhierarchies to allow simple techniques, such as a transitive closure table, to be used.

For more information on using SNOMED CT in Natural Language Processing, please refer to Data Analytics with SNOMED CT.

**Terminology Binding to Information Models**

2. Use Cases describes using the MRCM to suggest possible SNOMED CT terminology bindings for an information model, to ensure that the bindings are consistent (at least at a high level) with the SNOMED CT concept model, and to support the design of expression templates.

The processes above for authoring SNOMED CT postcoordinated expressions, expression constraints and queries can similarly be applied to support terminology binding. Alternatively, new use case specific MRCM reference sets could be developed, which restrict the available attributes and ranges that may be used in a specific implementation scenario.

**Versioning**

As mentioned in 3. Requirements (D.5), the concept model must be able to be changed between releases to fix identified issues and enhance future releases.

The RF2 reference set format used in the design of the MRCM provides a standard way of versioning the MRCM and retaining a history of changes, to enable the concept model to evolve over time. Using this standard approach, the effectiveTime field is used to reflect the date on which each rule version comes into effect, while the active field is used to indicate whether a rule is active or inactive. A new version of an MRCM rule may be created to update the mutable fields in the reference set, while changing immutable fields requires the rule to be inactivated, and a new rule to be created.

When a new or updated rule is introduced to the MRCM, it may take a period of time before all existing content conforms to the new rule. In these situations, a content type of 723593002 (All new pre-coordinated SNOMED CT content) can be used, to ensure that all new content obeys the rule. Over time, as the old content is updated to reflect the new rule, the content type may be changed to either 723594008 (All pre-coordinated SNOMED CT content) or 723596005 (All SNOMED CT content). This approach can be used to ensure that all new content added to SNOMED CT conforms to high quality concept model rules, even though reviewing and correcting existing content may take additional time.

**Extension**

As described in 3. Requirements (D.6), the international MRCM rules defined by SNOMED International must be able to be extended and adapted by organizations developing SNOMED CT extensions, to support the concept model requirements of their extension content.

To support this requirement, SNOMED CT extension developers will be able to copy the international MRCM rules defined by SNOMED International into one or more extension MRCM reference sets and edit these reference sets to add new rules and restrict or extend the existing rules as required. New MRCM reference sets created for a SNOMED CT extension must be designed to ensure consistency and data integrity between editions. Please note that if no changes to the SNOMED CT concept model are required to support an extension, no additional MRCM reference sets need to be created, as the international MRCM can be reused by the extension.

The 723563008 MRCM Module Scope Reference Set is designed to allow extension developers to specify which MRCM reference sets should be applied to the content in each module. The 723563008 MRCM Module Scope Reference Set should therefore be referred to, to determine whether an extension is reusing the international MRCM, or using a customized extension MRCM to support its content.