

Disjoint Content Definitions for QA

Tim Williams

Objective

- Demonstrate that adding disjoint content definitions for a limited set of concepts will help prevent simple mistakes when modeling new content in SNOMED CT releases affecting both equivalence detection and content retrieval via the SNOMED CT hierarchies

Introduction

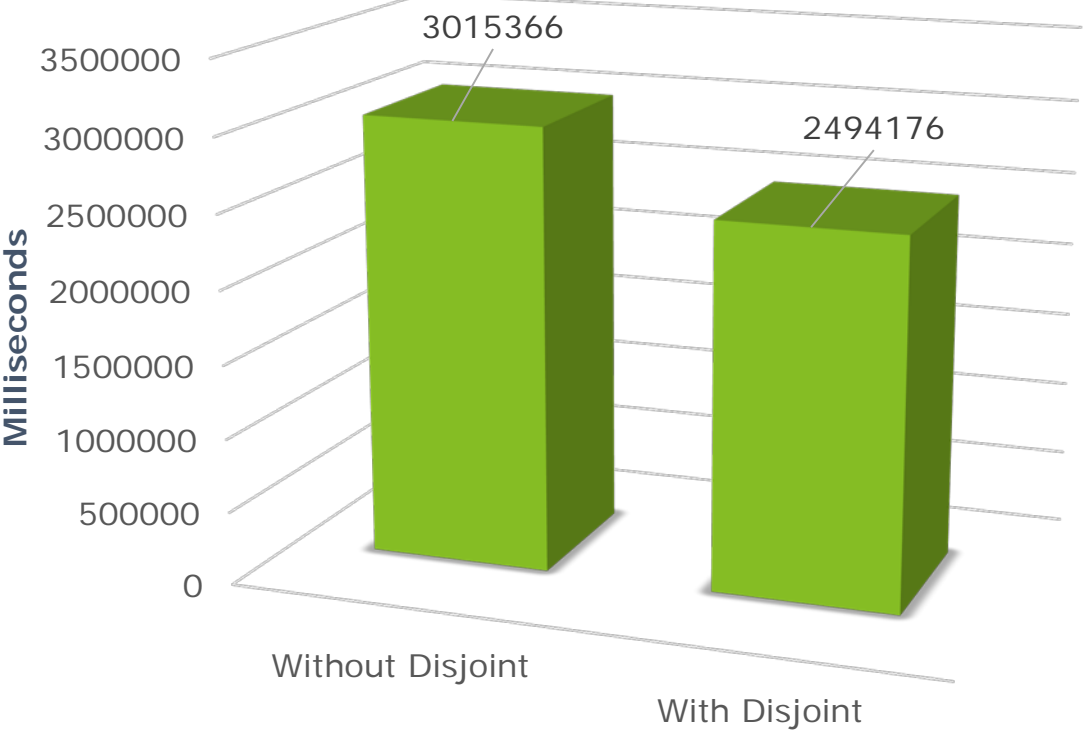
- Concepts are disjoint if they cannot have common children
- All concepts are assumed to have potential overlapping children unless they are explicitly stated to not have them
- SNOMED CT does not contain any such statements
- All top level primitives should be stated as disjoint
- **Exceptions are:**
 - 260787004 |Physical object (physical object)|
 - 373873005 |Pharmaceutical / biologic product (product)|

Initial Evaluation

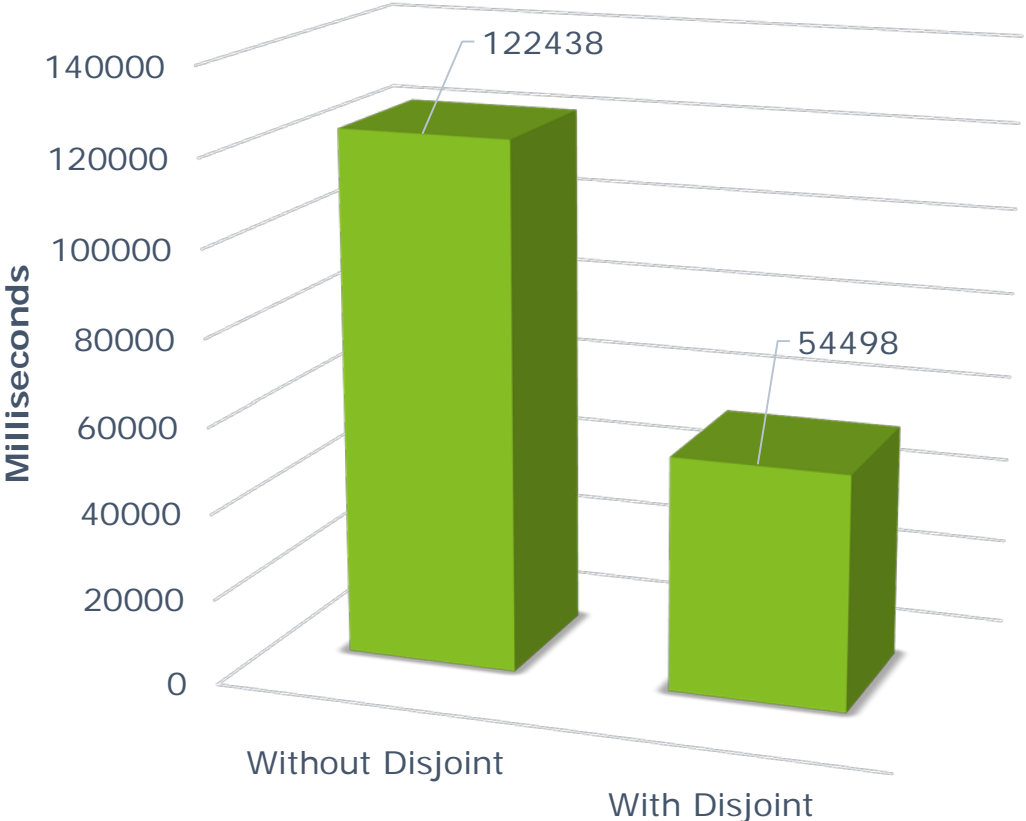
- Utilized the `tls2_StatedRelationshipsToOwlKRSS_Script_INT.pl` Perl script from the SNOMED International GitHub registry to create an OWL file from the March 2017 US Edition release
- Loaded into Protégé 5.2.0 editor and used the included Hermit reasoner using the OWL file without disjoint statements
- Added the 169 disjoint statements to the upper level primitive concepts
- Repeated the test using the SnoRocket reasoner plugin

Classifier Results

Hermit



SnoRocket



Expanded Evaluation

- Expanded work focused on the primitive hierarchies of Substance, Body structure and Situation with explicit context
- Focused on identifying all concepts that are currently disjoint from each other beginning at the top of the hierarchy and traversing downward
- Had clinicians review all potential statements to confirm that they are correct

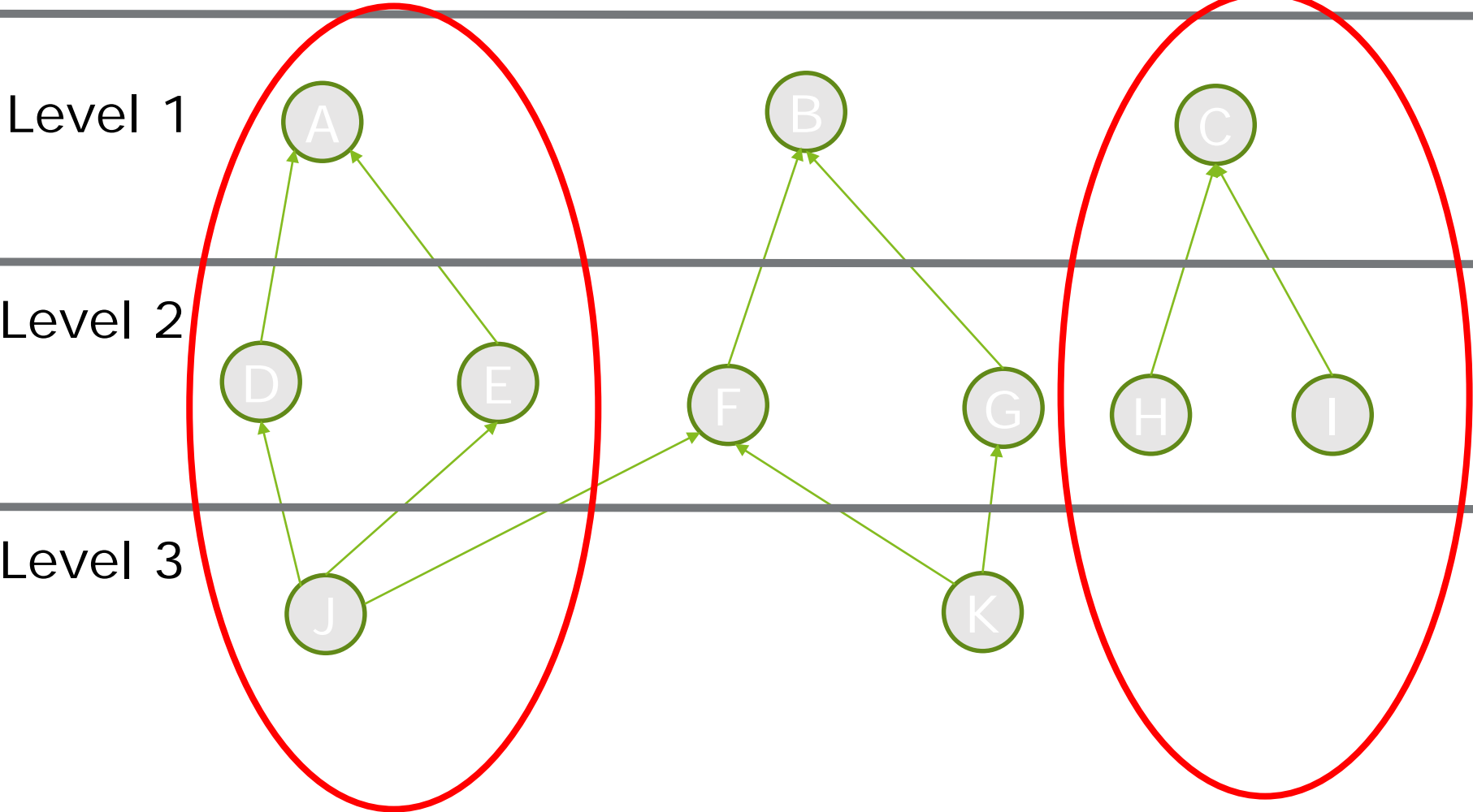
Expanded Approach

To efficiently process and identify the most beneficial content, the following steps were taken:

- Traversed the SNOMED CT hierarchies to identify the appropriate disjoint definitions for each concept; this review paid particular attention to primitive hierarchies like body structure and substances where disjoint content statements provided the most benefit
- Clinical Terminologists independently reviewed the assignment results and any disagreements were resolved with the assistance of a Senior Clinical Terminologist
- Stored all disjoint definitions in a Reference Set in SNOMED CT RF2 Format

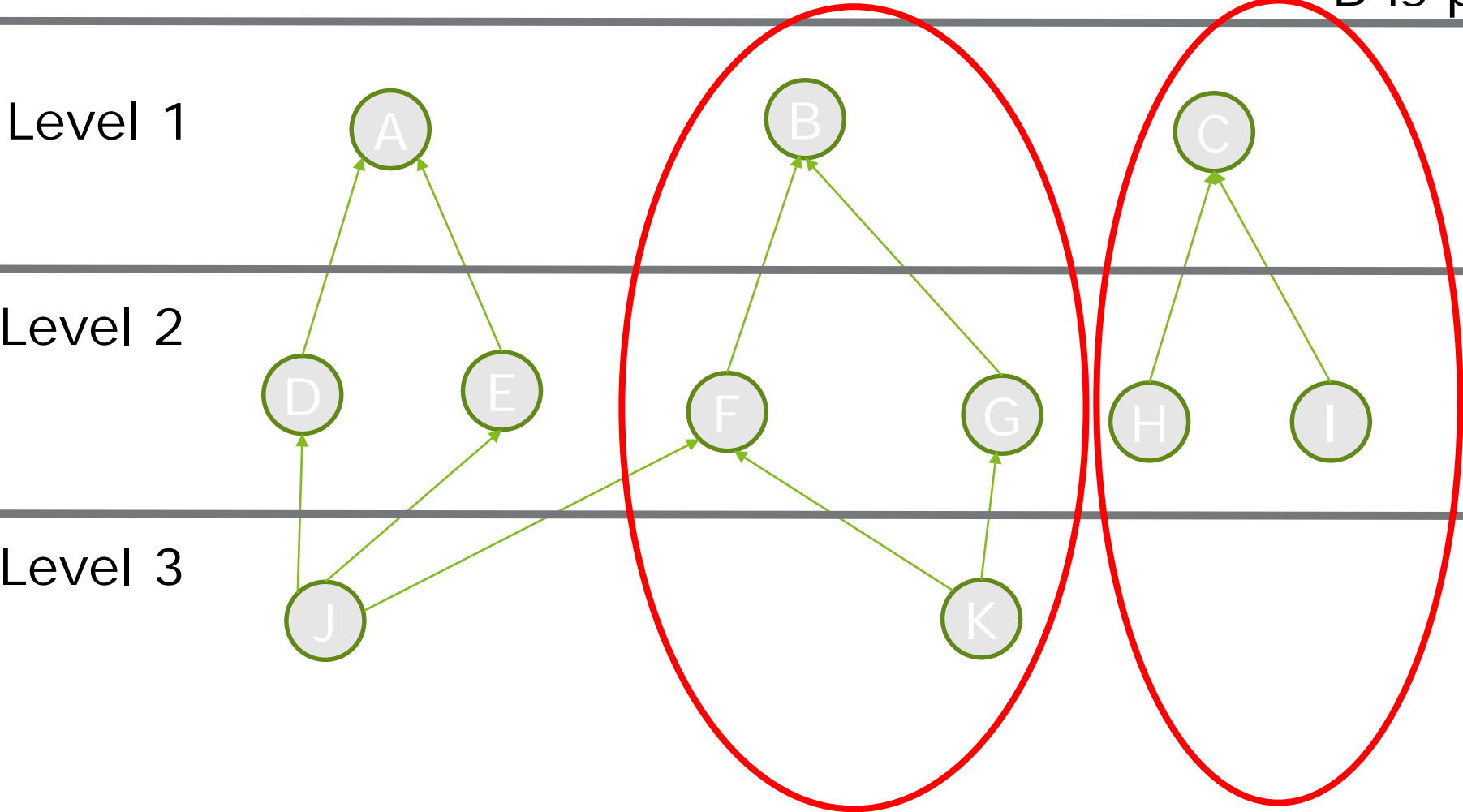
Query Strategy to Identify Potential Disjoint Content

A is potential disjoint of C



Query Strategy to Identify Potential Disjoint Content

A is potential disjoint of C
B is potential disjoint of C

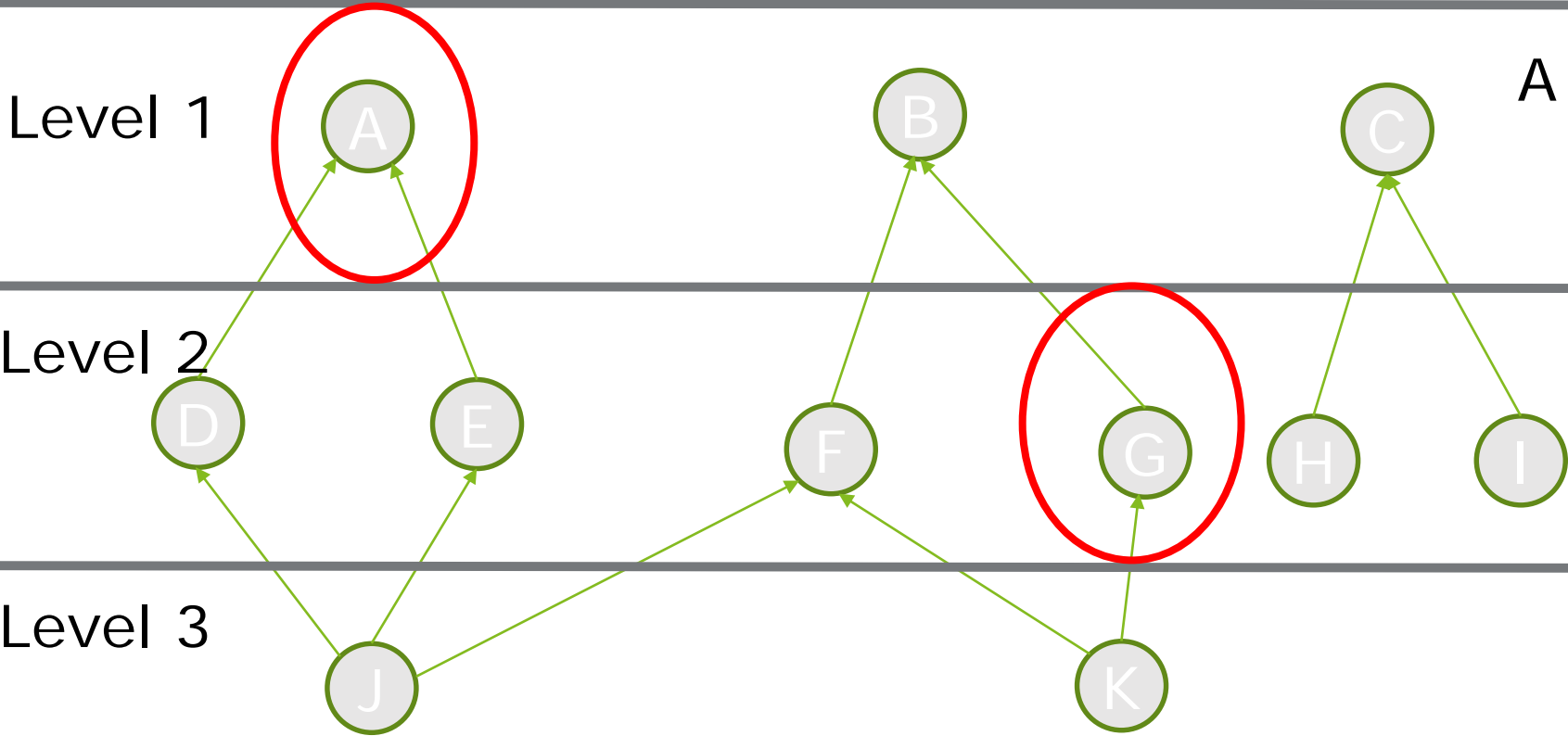


Query Strategy to Identify Potential Disjoint Content

A is potential disjoint of C

B is potential disjoint of C

A is potential disjoint of G



Query Strategy to Identify Potential Disjoint Content

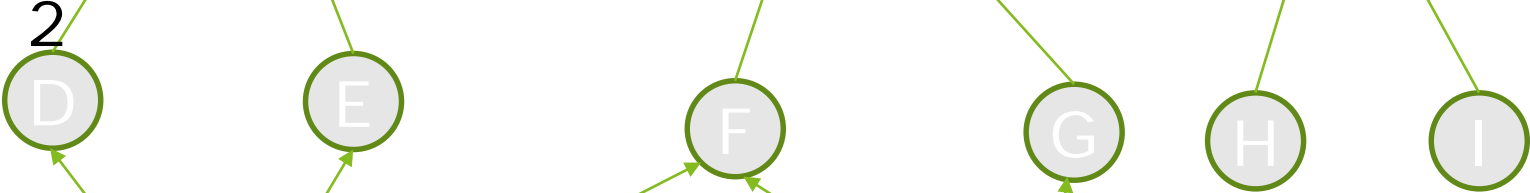
A is potential disjoint of C
B is potential disjoint of C

Level 1



A is potential disjoint of G

Level 2



J is potential disjoint of K

Level 3



Expanded Results

- Statements by the Numbers:
 - 13 Substance Statements
 - 1193 Body Structure Statements
 - 12 Situation with Explicit Content Statements

These disjoint statements only cover the immediate children for the all the hierarchies listed above, except for Body structure where we evaluated concepts three levels below the upper level primitive to identify potential disjoint content.

Conclusion

- The Body structure and Substance hierarchies will have limited use cases for extension and post-coordination once the work to redesign the concept model for these hierarchies is complete
- The Situation with explicit context hierarchy is one where heavy post-coordination and/or extension will take place; however most of this work will involve assigning a single parent that is a direct subtype of the upper level primitive
- Focusing on addressing any modeling issues in these hierarchies and introducing a mechanism for blocking the editing of the stated view or changes in the inferred view of stable concepts without editorial approval would be a more productive use of resources
- Focus only on specific concepts that could benefit from disjoint statements to fix issues with the inferred view
- Without these statements to detect disjoint content, there is a potential for modeling incorrect parents for SNOMED CT concepts, which affects both equivalence detection and content retrieval via the SNOMED CT hierarchies
- Adding disjoint content statements to the SNOMED CT definitions will assist both SNOMED CT International and extension content creators by providing explicitly stated, built in Quality Assurance to prevent errors in assigning parents
- The creation of disjoint statements should focus on the upper level primitive hierarchies and their direct descendants
- Adding disjoint statements to the below the upper level primitive concepts will provide limited benefit for error checking
- Assigning further statements may become more useful once the redesign of the concept model for the various hierarchies is complete

Acknowledgments

Acknowledgement to the Veterans Health Administration (VHA) and Dr. Keith E. Campbell for their support in developing and testing the disjoint content.