Necessary Normal Form for Inferred Relationships

<Use this document for standards, policies, manuals, etc. See https://confluence.ihtsdotools.org/display/FD for more information>

Date: 20180409

Version: 0.03

Status: Draft

SNOMED International
Registered in England and Wales • Company Registration Number 9915820
Reg. address: One Kingdom Street • Paddington Central • London W2 6BD • United Kingdom
Tel: +44 (0) 203 755 0974 • info@snomed.org • www.snomed.org

SNOMED International is the trading name of the
International Health Terminology Standards Development Organisation,
a private company limited by guarantee

Amendment History

Version	Date	Editor	Comments
0.01	20180226	Yongsheng Gao, Kai Kewley	First draft for comment
0.02	20180305	Yongsheng Gao, Kai Kewley	Updates following the comments from Michael Lawley and Jim Steel
0.03	20180409	Yongsheng Gao, Kai Kewley	Updates following feedback from Linda Bird and Ronald Cornet

Approvals

Version	Date	Approver	Comments
	YYYYMMDD		

Future Review Timetable

Review date	Responsible owner	Comments
YYYYMMDD	Person/group responsible	Summary of action

Table of Contents

Necessary Normal Form for Inferred Relationships

Introduction

Background

The SNOMED CT Logic Profile Enhancements have introduced new Description Logic (DL) features to support content modeling and quality improvement. The logic definitions will be represented by the OWL axiom refset that is a replacement of the RF2 stated relationship file.

As a result, the nature of the inferred relationship file in the distribution normal form (DNF) will change because the new DL features are not representable in the current relationships file. The inferred relationship file will maintain the same format and structure, but it is no longer equivalent to the stated form (containing all necessary and sufficient conditions). In fact, it is a collection of all the necessary conditions of precoordinated concepts and represents a subset of the full semantics.

Purpose

The purpose of this document is to define the Necessary Normal Form (NNF) for inferred relationships and the principles for calculating the NNF. The high level process for the calculation and a reference implementation are provided.

Scope

The scope of this document is limited to the current implementation in the International release. The representation of concrete domains and nesting are out of scope for this document.

Audience

The target audiences of this document includes:

- SNOMED CT terminology technical developers;
- SNOMED CT national release centers and content developers;
- SNOMED CT implementers and analysts, including developers of EHR and information systems.

Definition of Necessary Normal Form

The Necessary Normal Form is a replacement for the Distribution Normal Form for inferred relationships. The NNF is a precalculated distribution form for practical purposes, for example, to support the continuity of existing implementations based on relational databases and queries by the expression constraint language.

The NNF consists of the full set of necessary relationships of precoordinated concepts after removal of redundant relationships within a given concept definition.

The NNF does not include class disjointness, transitive properties, reflexive properties and sufficient conditions represented as General Concept Inclusions (GCIs) in the OWL axiom refset.

What constitutes a necessary relationship

Within the scope of a SNOMED CT terminology, necessary relationships are defined only for precoordinated concepts (aka OWL's named classes).
Let C be a precoordinated concept and D be either a precoordinated concept or a complex expression. If an axiom is in the form of SubClassOf(C D) or EquivalentClasses(C D), then all of the derivable and necessary relationships of D are necessary relationships of C.

Rules for determining redundant relationships

The following two rules should be applied to determine redundant relationships.

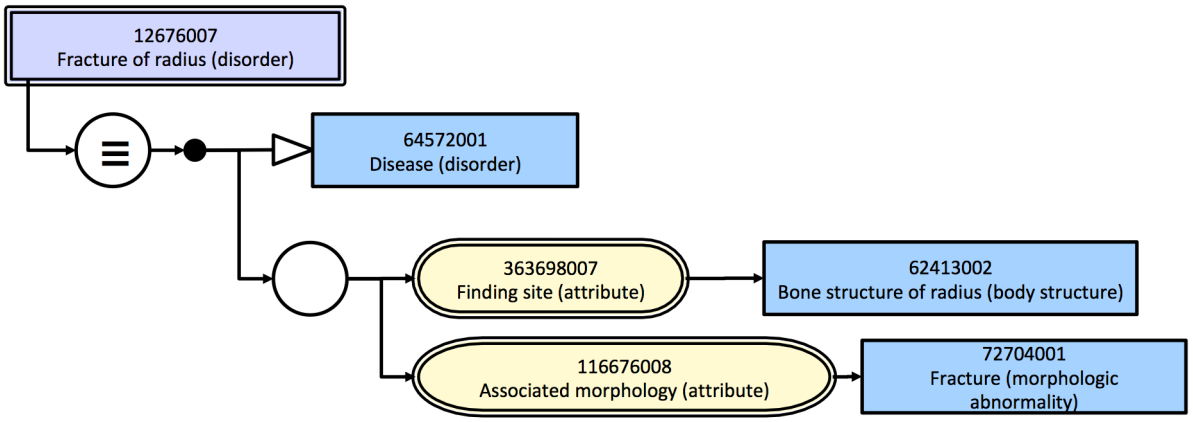
Rule 1 - Class and Role inclusions
Given two relationships, A and B, A with r = C and B with s = D, within the same role group, A is redundant if:

- r is the same as or a supertype of s, and
- C is the same as or a supertype of D

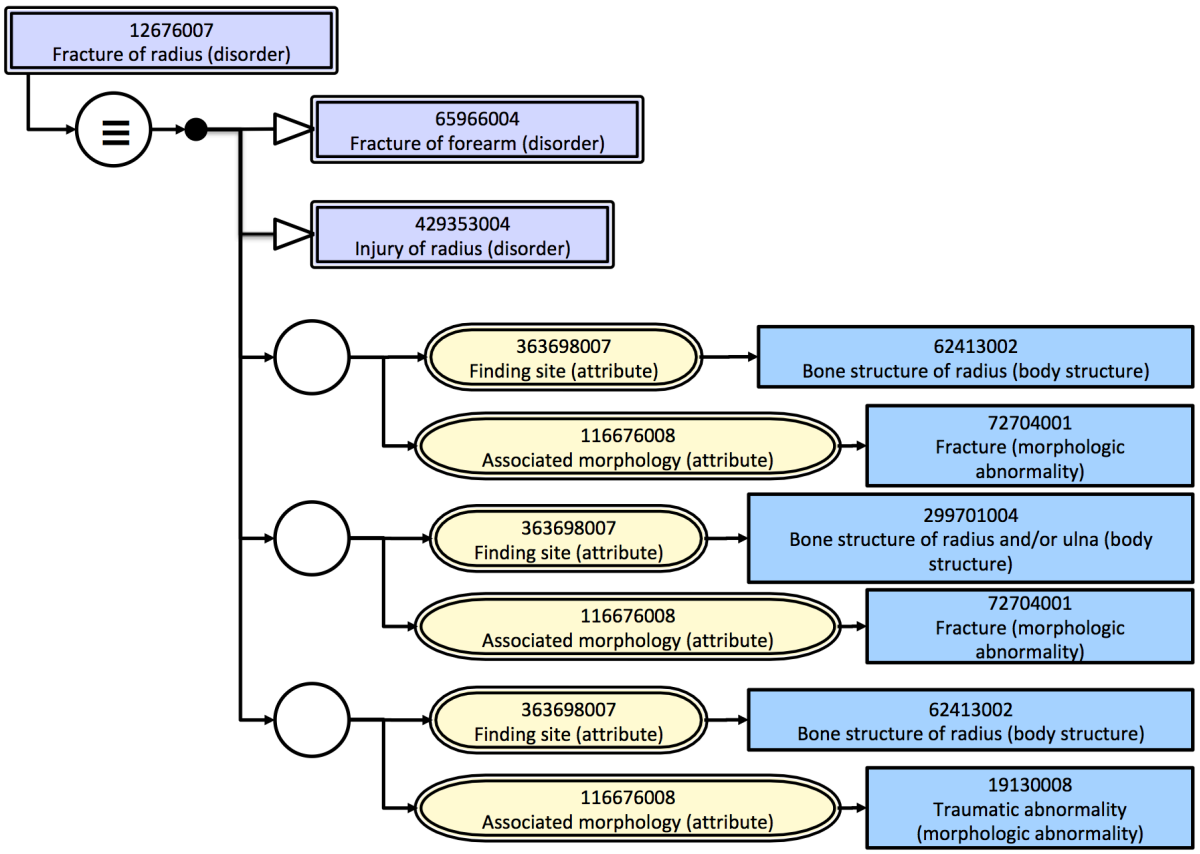
Note, "crossover relationships", where r is a supertype of s, and C is instead a subtype of D do not result in a redundant relationship.

Class inclusion example for rule 1:

Stated relationships

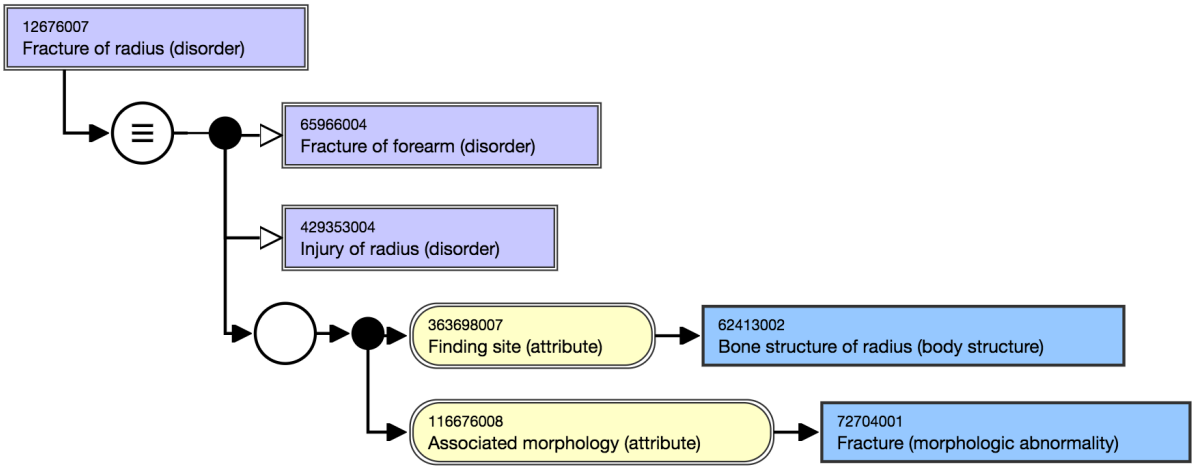


Inferred relationships before the removal of redundant relationship



For |Fracture of radius|, the relationship |Finding site| = |Bone structure of radius and/or ulna| is inherited from |Fracture of forearm|, which is a redundant relationship because |Radius bone structure| is a subtype of |Bone structure of forearm|. The relationship |Associated morphology| = |Traumatic abnormality| is inherited from |Injury of radius|, which is a redundant relationship because |Fracture| is a subtype of |Traumatic abnormality|.

Inferred relationships after the reduction



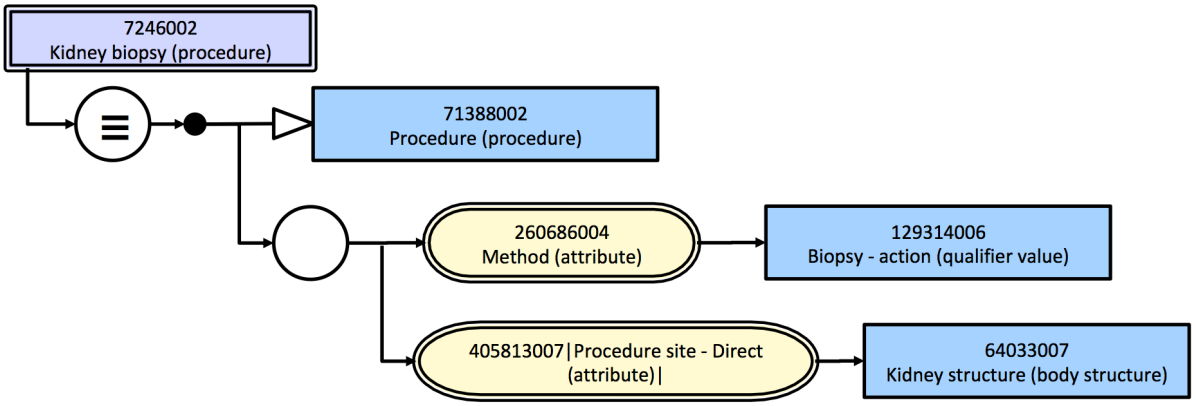
Example in OWL axiom refset and RF2 relationship file (NNF)

referenced ComponentId	owlExpression [1] (stated relationships)	Inferred Relationships in Necessary Normal Form			
		sourceId	destinationId	relationshipGroup	typeId
125605004	EquivalentClasses(:125605004 ObjectIntersectionOf(:64572001 ObjectSomeValuesFrom(:609096000 ObjectIntersectionOf(ObjectSomeValuesFrom(:116676008 :72704001) ObjectSomeValuesFrom(:363698007 :272673000)))))	125605004	284003005	0	116680003
		125605004	72704001	1	116676008
		125605004	**272673000**	**1**	**363698007**
12676007	EquivalentClasses(:12676007 ObjectIntersectionOf(:64572001 ObjectSomeValuesFrom(:609096000 ObjectIntersectionOf(ObjectSomeValuesFrom(:116676008 :72704001) ObjectSomeValuesFrom(:363698007 :62413002)))))	12676007	65966004	0	116680003
		12676007	429353004	0	116680003
		12676007	72704001	1	116676008
		12676007	**62413002**	**1**	**363698007**
62413002	SubClassOf(:62413002 :299701004)	**62413002**	**299701004**	**0**	**116680003**

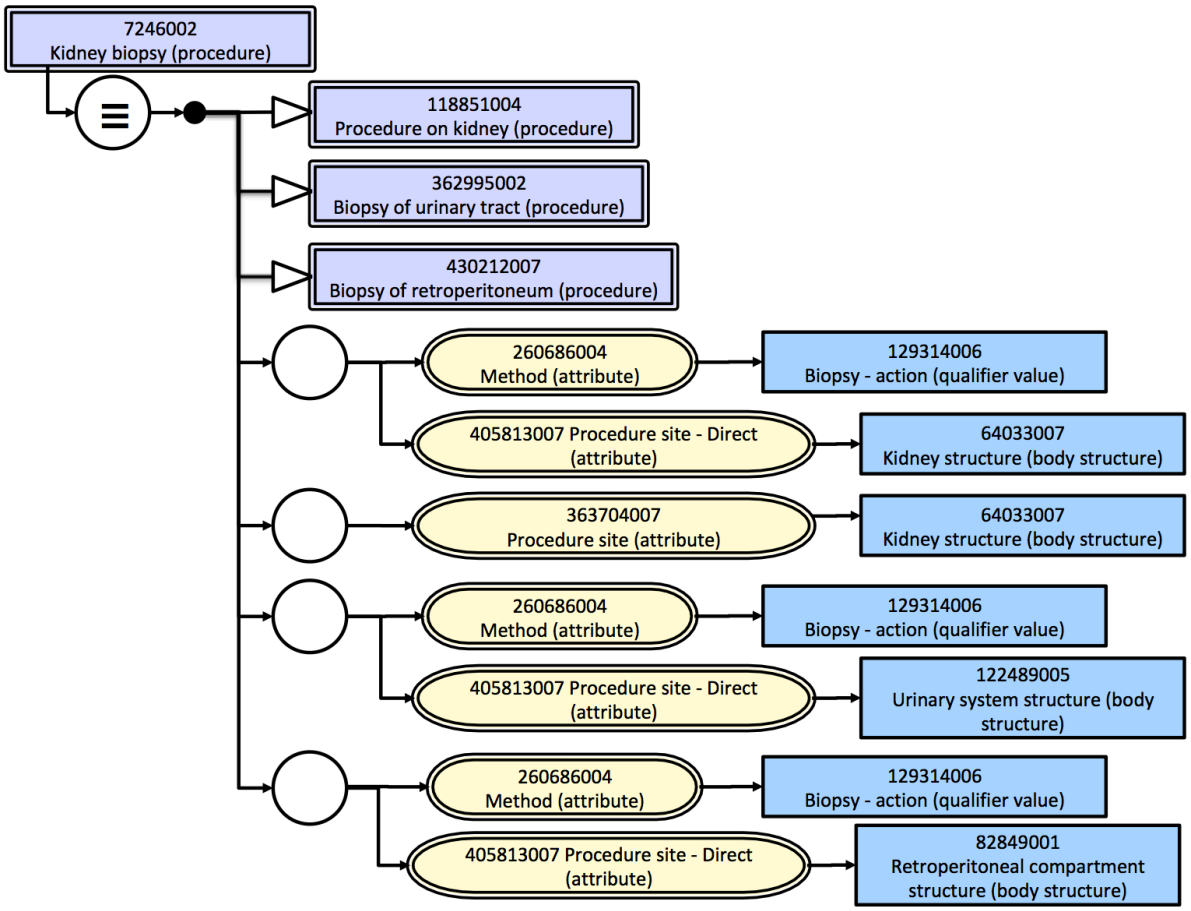
[1] 609096000 |Role group (attribute)|

Role inclusion example for rule 1

Stated relationships



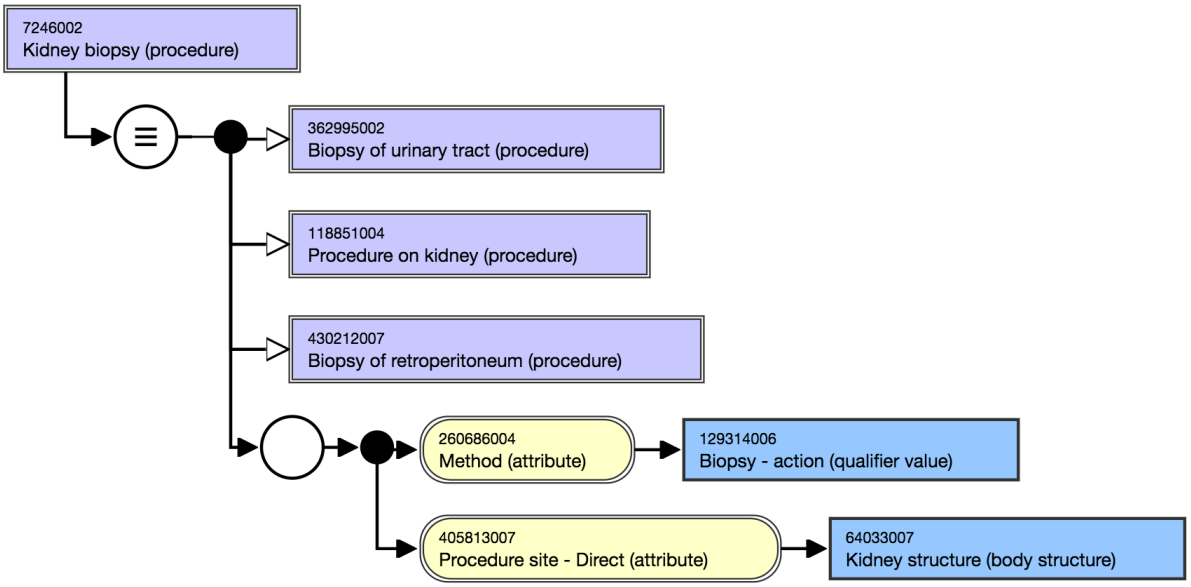
Inferred relationships before the removal of redundant relationship



For concept |Kidney biopsy|, the relationship 363704007 |Procedure site| = 64033007 |Kidney structure| is inherited from 118851004 |Procedure on kidney|, which is a redundant relationship to 405813007 |Procedure site - Direct| = |Kidney structure| because |Procedure site - Direct| is a subtype of |Procedure site|. Because |Kidney structure| is a subtype of

|Urinary system structure| and |Retroperitoneal compartment structure|, the inherited relationships for |Procedure site - Direct| are also redundant.

Inferred relationships after reduction



Example in OWL axiom refset and RF2 relationship file (NNF)

referenced ComponentId	owlExpression[2] (stated relationships)	Inferred Relationships in Necessary Normal Form			
		sourceId	destinationId	relationship Group	typeId
118851004	EquivalentClasses(:118851004 ObjectIntersectionOf(:71388002 ObjectSomeValuesFrom(:609096000 ObjectSomeValuesFrom(:363704007 :64033007))))	118851004	71388002	0	116680003
		118851004	**64033007**	**1**	**363704007**
7246002	EquivalentClasses(:7246002 ObjectIntersectionOf(:71388002 ObjectSomeValuesFrom(:609096000 ObjectIntersectionOf(ObjectSomeValuesFrom(:260686004 :129314006) ObjectSomeValuesFrom(:405813007 :64033007)))))	7246002	118851004	0	116680003
		7246002	362995002	0	116680003
		7246002	430212007	0	116680003
		7246002	129314006	1	260686004
		7246002	**64033007**	**1**	**405813007**
405813007	SubObjectPropertyOf(:405813007 :363704007)	**405813007**	**363704007**	**0**	**116680003**

[2] 609096000 |Role group (attribute)|

Rule 2 - Property chains including transitive properties
Given attribute r, s and t with a property chain SubObjectPropertyOf(ObjectPropertyChain(t s) r), and two relationships A and B, A with r = C and B with u = D, within the same role group, A is redundant if:

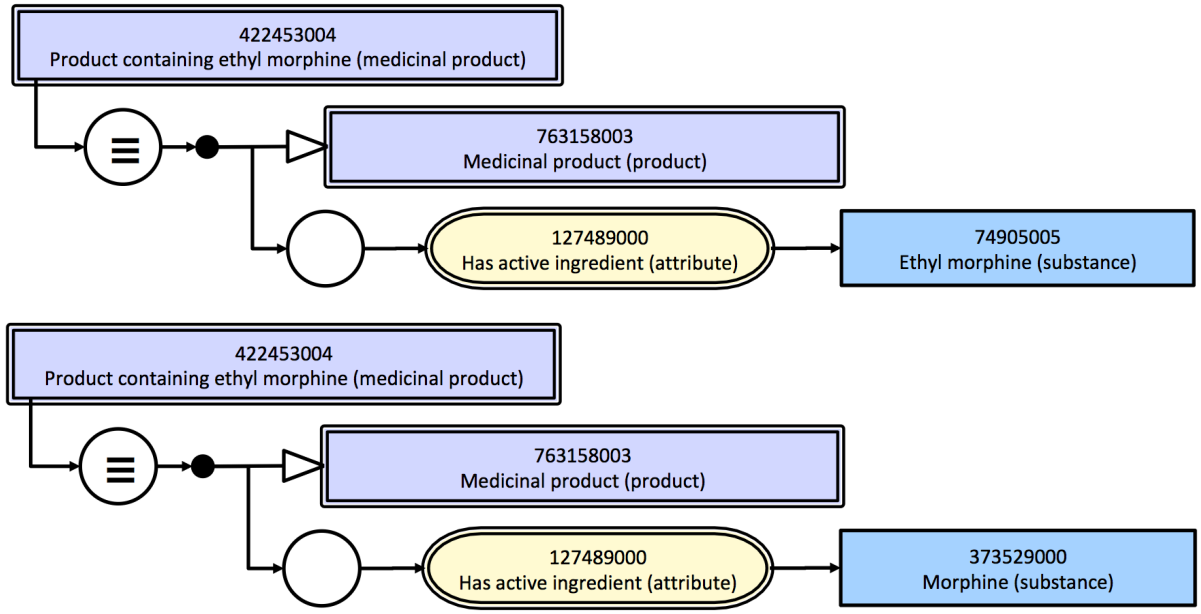
- Attribute u is the same as or a subtype of t, and
- D has relationship to C via attribute s

Note the following:

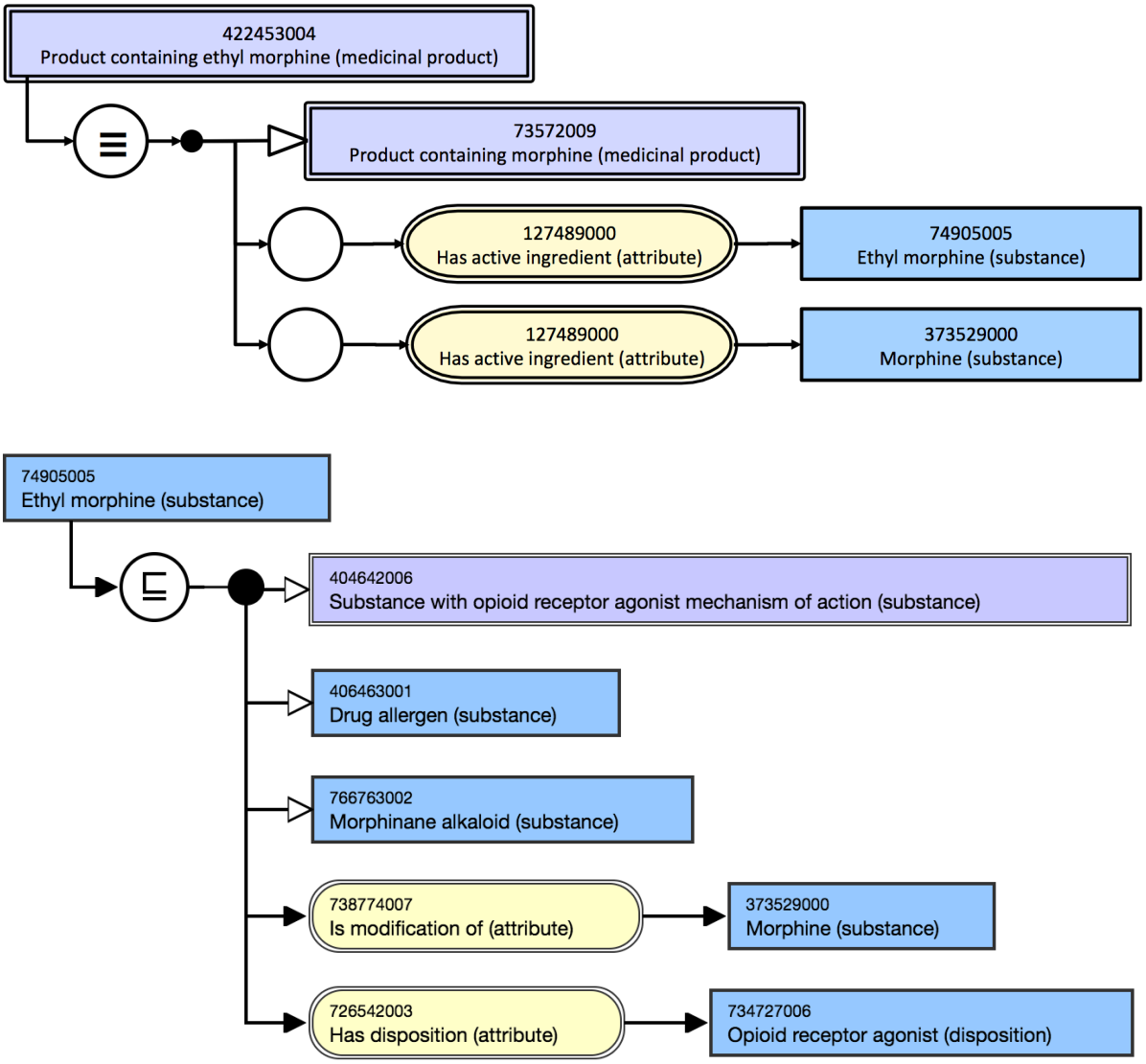
- C does not need to subsume D
- Attribute t does not need to be the same as or a subtype of r
- Transitive properties are defined by a property chain in the form of SubObjectPropertyOf(ObjectPropertyChain(r r) r) and thus it is a special case of the above.

Example for property chain for rule 2:

Stated relationships

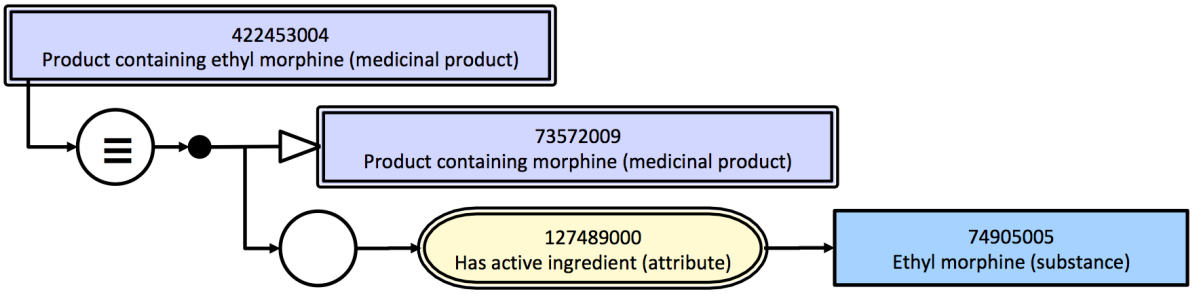


Inferred relationships before the removal of redundant relationship



For |Product containing ethyl morphine|, the relationship |has active ingredient| = |Morphine| is inherited from |Product containing morphine|, which is a redundant relationship to |has active ingredient| = |Ethyl morphine| because |Ethyl morphine| |Is modification of| |Morphine| and |has active ingredient| has a property chain with |Is modification of|.

Inferred relationships after reduction



Example in OWL axiom refset and RF2 relationship file (NNF)

referenced ComponentId	owlExpression[3] (stated relationships)	Inferred Relationships in Necessary Normal Form			
		sourceId	destinationId	relationshipGroup	typeId
73572009	EquivalentClasses(:73572009 ObjectIntersectionOf(:763158003 ObjectSomeValuesFrom(:609096000 ObjectSomeValuesFrom(:127489000 :373529000))))	73572009	764887005	0	116680003
		73572009	360204007	0	116680003
		73572009	**373529000**	**1**	**127489000**
422453004	EquivalentClasses(:422453004 ObjectIntersectionOf(:763158003 ObjectSomeValuesFrom(:609096000 ObjectSomeValuesFrom(:127489000 :74905005))))	422453004	73572009	0	116680003
		422453004	**74905005**	**1**	**127489000**
127489000	**SubObjectPropertyOf(ObjectPropertyChain(:127489000 :738774007) :127489000))**	N/A	N/A		N/A
74905005	SubClassOf(:74905005 ObjectIntersectionOf(:440327007 ObjectSomeValuesFrom(:738774007 :373529000)))	74905005	440327007	0	116680003
		74905005	**373529000**	**0**	**738774007**

Technical implementation for calculating the NNF

This fairly complex process uses the stated form and the output of the reasoner to calculate the necessary normal form which is represented in the relationship RF2 file.

The most straightforward way to produce the necessary normal form would be to use the Snomed OWL Toolkit or the Classification Service REST API which is language agnostic.

[3] 609096000 |Role group (attribute)|

High Level Process

Classification

1. Read the Stated Form from RF2 files.
    a. The following files are required: Concept, Stated Relationship, OWL Ontology Reference Set, OWL Axiom Reference Set and MRCM Attribute Domain Reference Set.
2. Use the OWL API to infer the class hierarchy
    a. Build the Ontology object using:
        i. Axioms from the OWL Axiom Reference Set, making a note of any Transitive property axioms.
        ii. Axioms created by converting Stated Relationships to OWL Axioms using the MRCM Attribute Domain Reference Set for list of attributes which should not be grouped in the given domain.
    b. Use a reasoner to pre-compute the class hierarchy.

Necessary Normal Form Calculation

Calculating the necessary normal form happens in two passes of the hierarchy.

1. Walk the class hierarchy in a top-down, breadth first, order.
    a. For each class visited gather the stated attributes of this class and each inferred parent.
    b. Compare the attributes and remove those which are found to be redundant because they are less specific in terms of depth in the hierarchy.
    c. During this first pass build a hierarchy for property chains and transitive properties.
2. Walk the class hierarchy again in the same order reducing the attributes of each class further.
    a. Compare the attributes and remove those which are found to be redundant because they are less specific in terms of depth in one of the alternate hierarchies.

For fine level detail the best source of information is the Java class org.snomed.otf.owltoolkit.normalform.RelationshipNormalFormGenerator which performs the Necessary Normal Form calculation.