The Description Logic and Concept Model of SNOMED Clinical Terms

Kent A. Spackman, MD, PhD
Oregon Health & Science University, Portland OR
Chair, SNOMED International Editorial Board

Workshop on DLs and Reasoning About Patient Data
Saarbruecken – November 2004

Basics

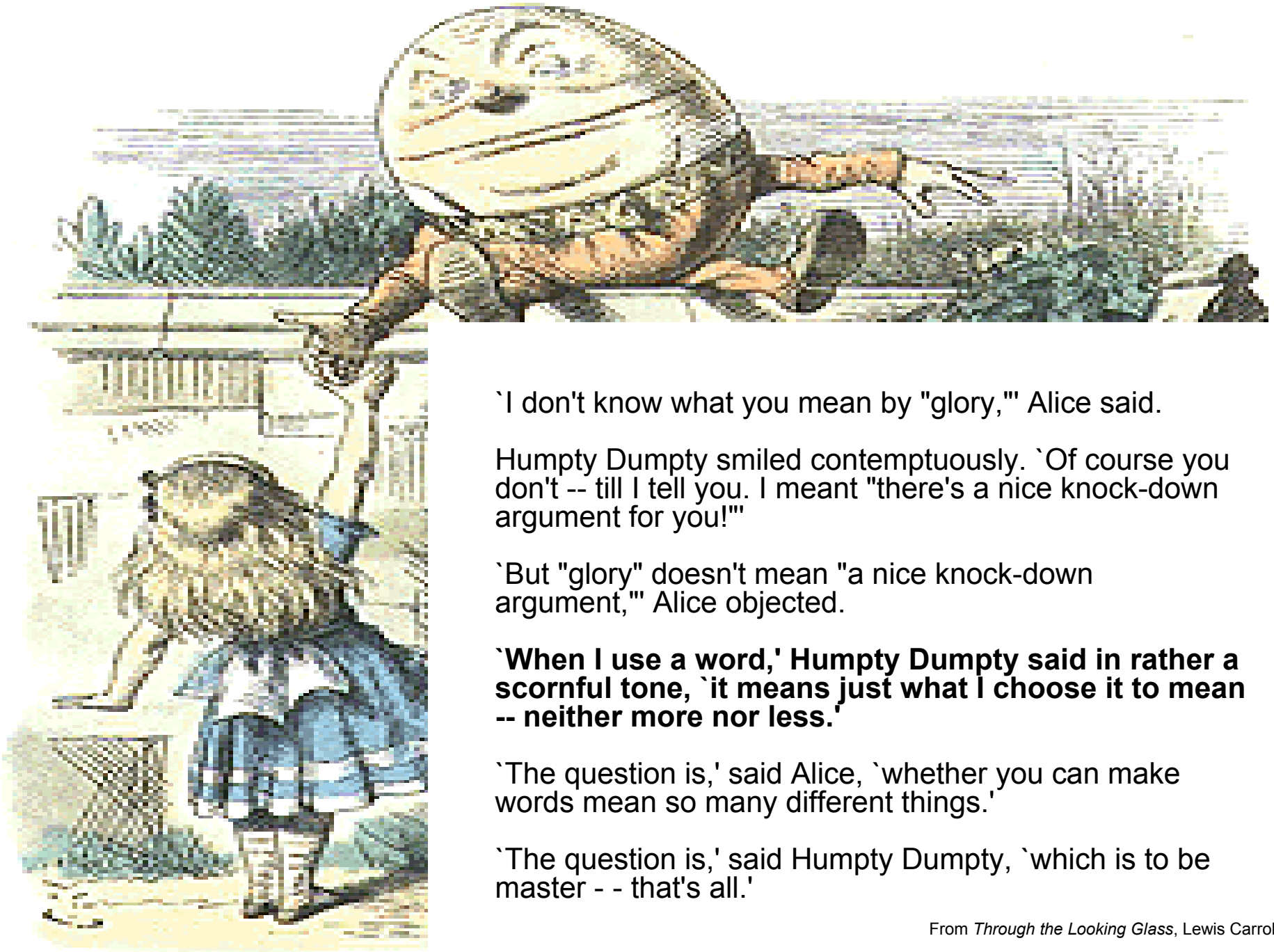
- SNOMED Clinical Terms is a terminological resource designed to support interoperability in electronic health care applications
- It has both “reference” properties and “interface” properties
  - Reference properties support compositional representation of meaning, as well as retrieval, aggregation & analysis by meaning
  - Interface properties support the user-interface linguistic, search, navigation, selection, and presentation aspects of the terminology

Reasoning About Patient Data

- Is SNOMED trying to be all things to all people? No, that is a recipe for failure

Clinicians Determine Meaning

- SNOMED is not the language police
- We are not trying to tell people which words they should or should not use (that's up to the clinical professions)
  - E.g. Should dermatologists still use the term "pyogenic granuloma" for the small blood vessel tumor that is neither pyogenic nor a granuloma?
- We also are not trying to tell clinicians how to *operationally* apply meanings to specific cases (again, that is up to the clinical professions)
  - E.g. What systolic and diastolic limits should be used for determining whether a patient has hypertension (140/90 ?)



`I don't know what you mean by "glory,"' Alice said.

Humpty Dumpty smiled contemptuously. `Of course you don't -- till I tell you. I meant "there's a nice knock-down argument for you!"'

`But "glory" doesn't mean "a nice knock-down argument,"' Alice objected.

`When I use a word,' Humpty Dumpty said in rather a scornful tone, `it means just what I choose it to mean -- neither more nor less.'

`The question is,' said Alice, `whether you can make words mean so many different things.'

`The question is,' said Humpty Dumpty, `which is to be master - - that's all.'

From *Through the Looking Glass*, Lewis Carrol

Integration of Terminologies

- SNOMED is also not duplicating the work of established official consensus groups
- Instead, it is providing an integrated resource where the various terminologies are available for electronic interoperability applications. Examples:
  - The International Society for Blood Transfusion (ISBT) provides a set of codes and names for red cell antigens and their antibodies
  - The WHO periodically revises numerous classifications of malignant neoplasms

SNOMED interacts with (but does not directly provide):

- Realm-specific or proprietary product terminologies (e.g. proprietary drugs or devices, country-specific administrative terminology)
- EHR information model
- EHR software
- Assertional knowledge bases of medicine
- Decision support knowledge
- Decision support rule syntax

The simplest information model

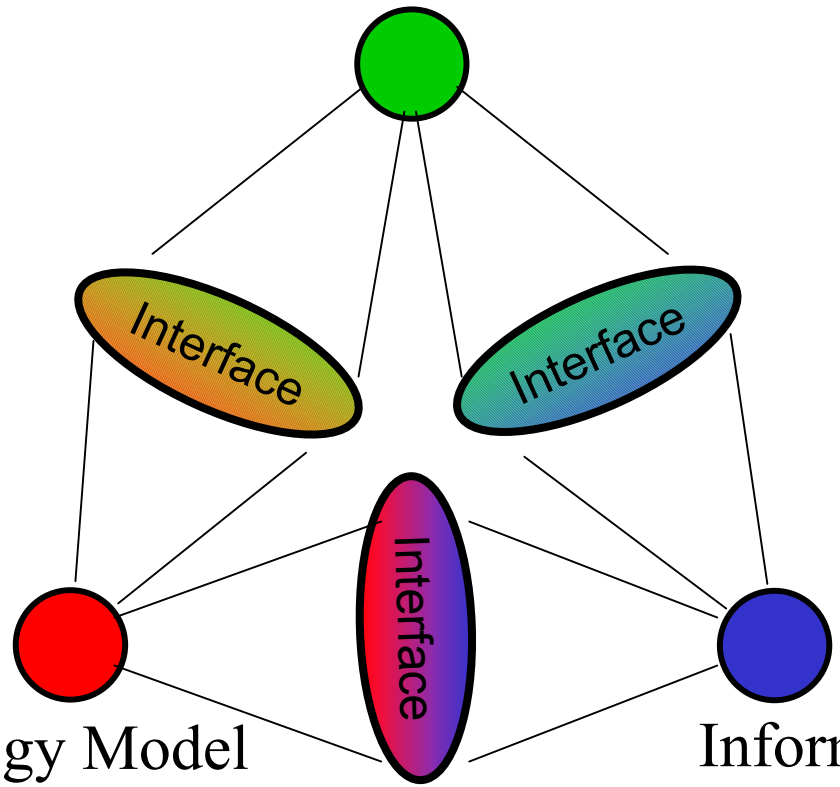
- Put all clinical data here _____

The simplest terminology model

- Two values:
  - Yes
  - No

Intermediate between these extremes there are many possible solutions!

Clinical Decision Support Model + Inference Rules



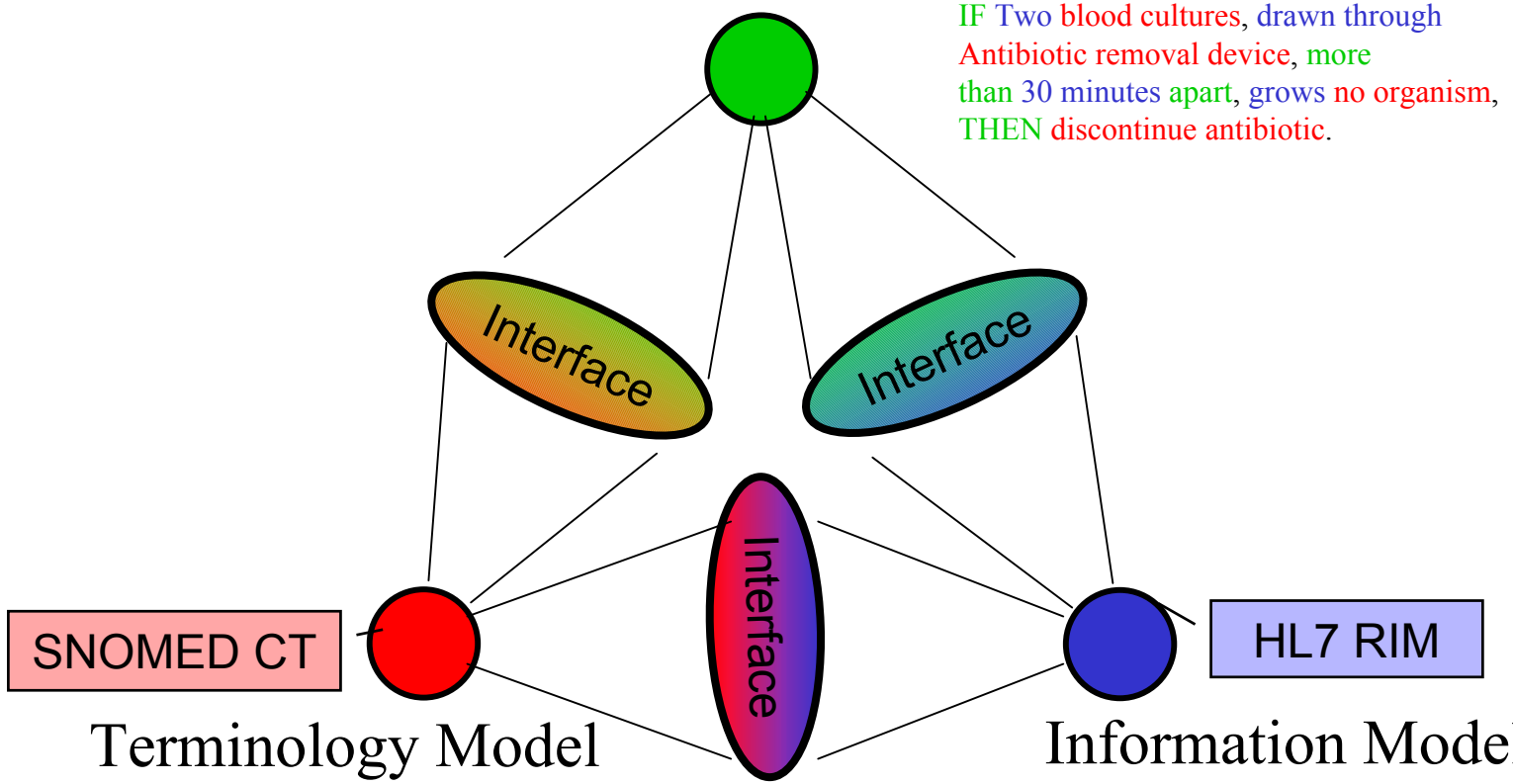
Terminology Model + Compositional Expressions

Information Model + Patient Data Structures

Diagram based on Figure 1 in Rector AL et al. “Interface of Inference Models with Concept and Medical Record Models” AIME 2001: 314-323

Clinical Decision Support Model + Inference Rules

IF Two blood cultures, drawn through Antibiotic removal device, more than 30 minutes apart, grows no organism, THEN discontinue antibiotic.



Terminology Model + Expressions

30088009 blood culture
55512120 antibiotic removal device
264868006 No growth
281789004 antibiotic therapy
223438000 advice to discontinue a procedure

Information Model + Patient Data Structures

What test was performed?
How many were done?
At what time?
What device was used?
What was the result of the test?

Diagram based on Figure 1 in Rector AL et al. “Interface of Inference Models with Concept and Medical Record Models” AIME 2001: 314-323

SNOMED's approach

- URU (understandable, reproducible, usable) criteria
- Evolutionary design

URU Criteria

1. Definitions should be Understandable by average clinicians, given brief explanations
2. We assess understandability by examining Reproducibility
3. We can ignore distinctions for which we see no Use in health care

Evolutionary Design

- Evolution without pre-ordained design
- Accumulation of desirable features
- Heterogeneity of perspectives

- Dealing with Scale
  - Participatory consensus-based approach
    - Involve the experts
  - Semantics-based concurrency control
    - Description logic underpinnings
  - Configuration management tools
    - Keith Campbell's "Galapagos" tool set

Emergence as a Standard: Recent Events

- Government Actions – US and UK
  - US National License (2003)
  - ANSI – Terminology Distribution Structure Standard (2003)
  - US NCVHS – HIPAA recommendation (2003)
  - US Government CHI Initiative recommendation (2004)
  - UMLS release (2004)
  - UK NPfIT adoption

Description Logic Foundation In a Nutshell:

- SNOMED's reference properties are based on a description logic
  - Conjunction
  - Existential restrictions
  - Role hierarchies
- Plus "role groups"
- Plus role composition
  - Currently:
    - direct-substance ◦ has-active-ingredient

DL Basics

- Concepts are given formal definitions
  - e.g. “Red car” is a kind of “Car”, and has color “Red”
  - e.g. “Lung disorder” is a kind of disorder, and has site “Lung”
- Definitions are expressed in description logic
  - conjunction (logical “and” ⊓)
  - existentially quantified role restrictions (∃R.C)
    - red_car = car ⊓ ∃ color.red
    - lung_disorder = disorder ⊓ ∃ site.lung

Concept & role forming operators & terminological axioms

Name of construct	Notation	Semantics
Primitive concept	A	$A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
Primitive role	R	$R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
Top	$\top$	$\Delta^{\mathcal{I}}$
Bottom	$\bot$	$\emptyset$
Conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$
Exists restriction	$\exists R.C$	$\{x \mid \exists y.R^{\mathcal{I}}(x,y) \wedge C^{\mathcal{I}}(y)\}$
Disjunction	$C \sqcup D$	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$
Negation	$\neg C$	$\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$
Value restriction	$\forall R.C$	$\{x \mid \forall y.R^{\mathcal{I}}(x,y) \rightarrow C^{\mathcal{I}}(y)\}$
Role composition	$R_1 \circ \ldots \circ R_n$	$R^{\mathcal{I}}_1 \circ \ldots \circ R^{\mathcal{I}}_n$
Restricted role value maps a.k.a. "right identity"	$R \circ S \sqsubseteq R$	$xRy \wedge ySz \rightarrow xRz$
Concept definition	$A \equiv C$	$A^{\mathcal{I}} = C^{\mathcal{I}}$
Primitive concept introduction	$A \sqsubseteq C$	$A^{\mathcal{I}} \subseteq C^{\mathcal{I}}$
Primitive role introduction (role hierarchy)	$R \sqsubseteq S$	$R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$

"Right identity" row added since Saarbrucken conference

Evolution of SNOMED's DL

SNOMED version	Concept & Role-forming Operators	Role axioms	Language name	Role grouping?
Initial development	$(\sqcap, \exists R.C)(\ )$		$\mathcal{EL}$	No
SNOMED RT	$(\sqcap, \exists R.C)(\circ)$		$\mathcal{EL}^{\circ}$	No
SNOMED CT (Jan02-Jan04)	$(\sqcap, \exists R.C)(\ )$		$\mathcal{EL}$	Yes
SNOMED CT (Jul04)	$(\sqcap, \exists R.C)(\circ)$	$R \sqsubseteq S$	$\mathcal{ELH}^{\circ}$	Yes

Notation follows Donini in Ch.3 Description Logic Handbook

All but last row have been corrected since Saarbrucken conference: RT used right identity; initially CT did not.

Need for Role Groups

- When a single concept may have more than one value for a particular attribute
  - for example, “bone fusion with tendon transfer”
    - method = fusion, site = bone, and
    - method = transfer, site = tendon
- And, one attribute-value pair needs to be associated with another.
  - How can we specify that the fusion is done to the bone and not to the tendon? and that the transfer is done to the tendon and not to the bone?

Role Groups as a Solution

- Informally:
  - don't nest or create subprocedures
  - simply "group" the attribute-value pairs
- Using curly braces as a syntactic marker:

  { site=bone, method=fusion }, {site=tendon, method=transfer}
- Or, in tabular form, use a "group" column:

attr	value	group
site	bone	1
method	fusion	1
site	tendon	2
method	transfer	2

Role Grouping Logical Form: A Nested Existential Restriction

- $\exists R_G.(\exists R_1.C_1 \sqcap \exists R_2.C_2) \sqcap \exists R_G.(\exists R_3.C_3)$

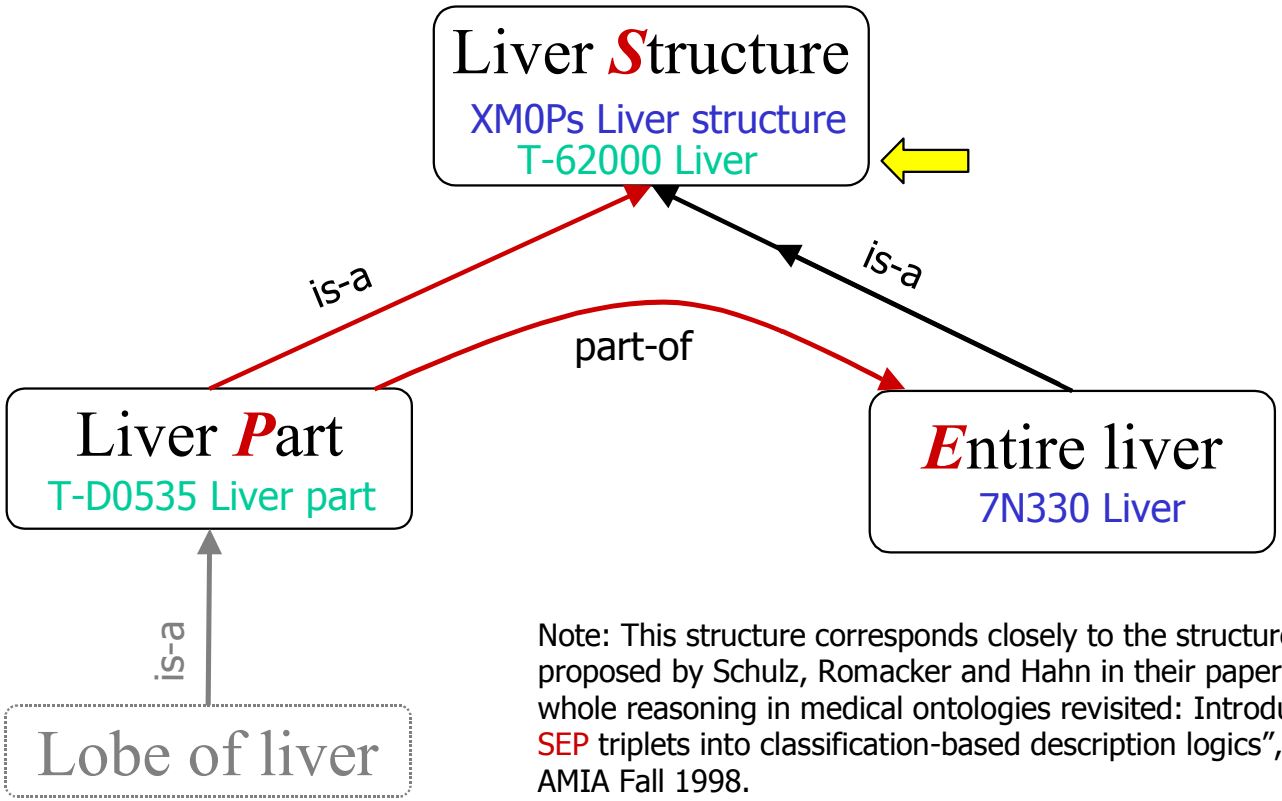
- Distributed as rows in relationships table:
  - $C\ R_3\ C_3$ 0
  - $C\ R_1\ C_1$ 1
  - $C\ R_2\ C_2$ 1

Role Composition

- $R \circ S \sqsubseteq R$
- $xRy \wedge ySz \rightarrow xRz$

- femurFracture site femur
- headOfFemurFracture site headOfFemur
- headOfFemur part-of Femur
- Allows the automated inference that:
  - headOfFemurFracture is-a FemurFracture

Avoiding Role Composition by Using SEP Triplets



Note: This structure corresponds closely to the structure proposed by Schulz, Romacker and Hahn in their paper "Part-whole reasoning in medical ontologies revisited: Introducing SEP triplets into classification-based description logics", Proc. AMIA Fall 1998.

Role Hierarchies

- Equipment
  - Direct-device
  - Indirect-device
  - Using
  - Access-instrument

DL features we might want to consider

- Negation
- Value restrictions
- Number restrictions
- Disjunction
- Cyclic (recursive) definitions
- Transitive roles
- General concept inclusion axioms

Negation $\neg C$

- Head injury without loss of consciousness
  - ? (headInjury) $\sqcap$ $\neg$ (lossOfConsciousness)
  - Problems with this approach:
    - Assertional statement (instance?) appears to be mixed with Terminological entity (type, class, universal)
- We can get along without negation

Value restriction $\forall R.C$

- Not an intuitive construct
  - person $\sqcap \forall hasCar.Jaguar$
  - Includes people who have no car, but if they had one it would have to be a Jaguar . . . . Do we encounter this kind of concept in common-sense thinking?
- Creates pernicious interactions with disjunction and negation that tend to make structural subsumption algorithms incomplete
- But it was included in $\mathcal{ALC}$ and $\mathcal{FL}^{-}$, so it seems languages including it have been studied extensively.
- We can definitely get along without value restriction

Disjunction $C \sqcup D$

- Some high-level aggregators are naturally disjunctive
- We can address this need partially by using navigation hierarchies

- There is no urgent need for adding disjunction

Cyclic definitions, number restrictions

- ? No significant need for these at present

Transitive roles

- $xRy \wedge yRz \rightarrow xRz$
- Useful for causal/associational chains
- Interaction with role hierarchy is interesting & useful

- Example:
  - Varicella (chicken pox)
  - Herpes zoster
  - Post-herpetic neuralgia

General concept inclusion axioms

- Extremely useful feature
- Demonstrated conclusively by GALEN project
- Probably compatible with a polynomial-time structural subsumption algorithm

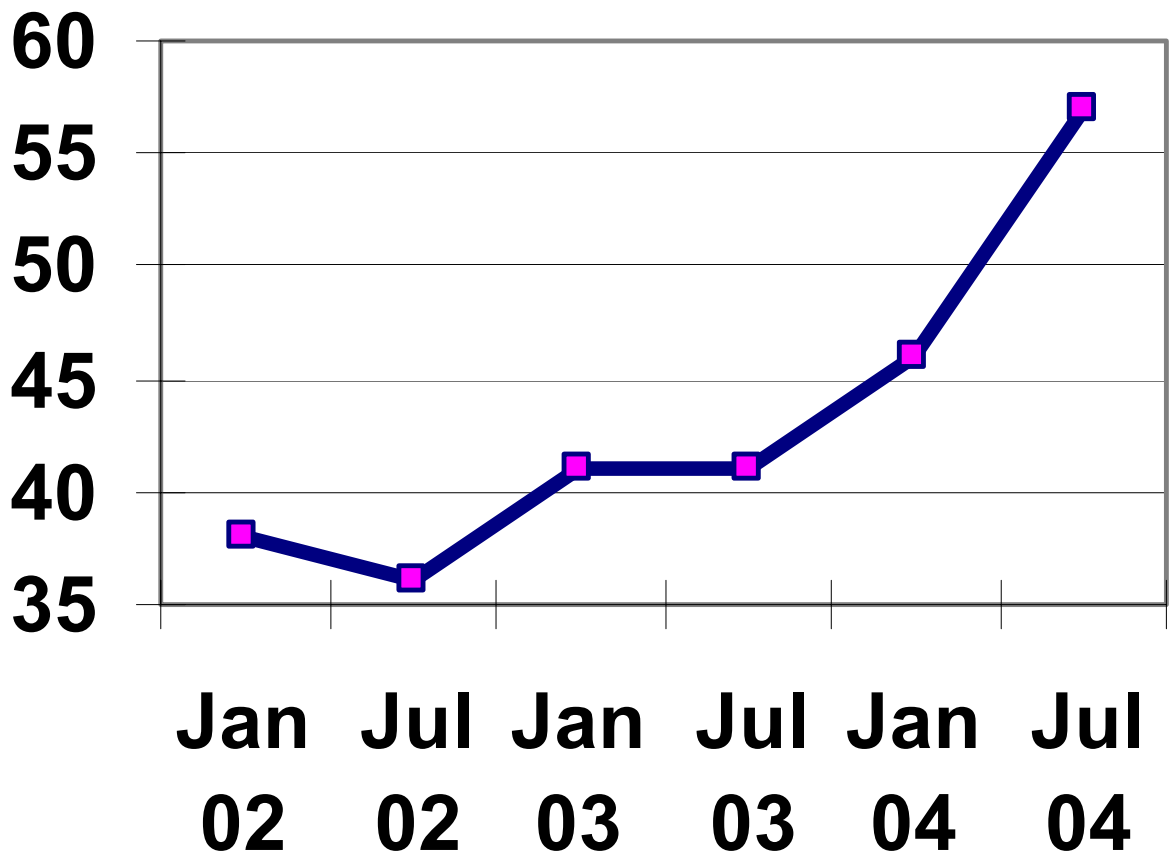
What classifier has SNOMED used?

- 1996: Kaiser Convergent Medical Terminology Project agreement with College of American Pathologists (CAP)
  - Kaiser Colorado was using K-REP description logic classifier from IBM
- 1999: UK National Health Service agreement with CAP
  - 3 year development project
  - K-REP developers had moved on to Apelon Corp., developed Ontylog classifier

SNOMED's Concept Model

- Top-level categories
- Attributes

Number of attributes (relationship types) in the SNOMED concept model



Concept Model

- Currently has 55 attributes (down from 57 in July 04)
  - Eliminated “subject of information” which overlapped with “subject relationship context”.
  - Eliminated “part-of” in anatomy. These are true but not in the “sufficient” set for fully defined lateralized concepts, because of the SEP model.

High level categories

- 18 “roots” but some of these aggregate groups.

High level categories

1. **clinical findings** (no URU criteria for a distinction between finding and disorder)
2. **procedures** (any "act" or collection of acts)
3. observable entities (a "question", or the object of a "what is the x" question. Too many things mixed up here from an ontological perspective).
4. body structures (anatomical, morphological)
5. organisms
6. pharmaceutical/biologic products (drugs, etc)
7. substances ("matter", not individuals)
8. physical objects (devices, etc)
9. physical forces
10. events
11. specimens
12. environments & geographic locations
13. social contexts
14. context-dependent categories
15. qualifier values
16. staging and scales
17. attributes

special concepts

FSN Tags

- Finding
- Disorder
- Procedure
- Organism
- Substance
- Product
- Physical force
- Physical object
- Attribute
- Person
- Social concept
- Religion/philosophy
- Lifestyle

- Body structure
- Morphological abnormality
- Context-dependent category
- Environment
- Geographic location
- Qualifier value
- Event
- Observable entity (many types here)
- Navigational concept
- Namespace concept
- Inactive concept
- Tumor staging
- Staging scale
- Assessment scale

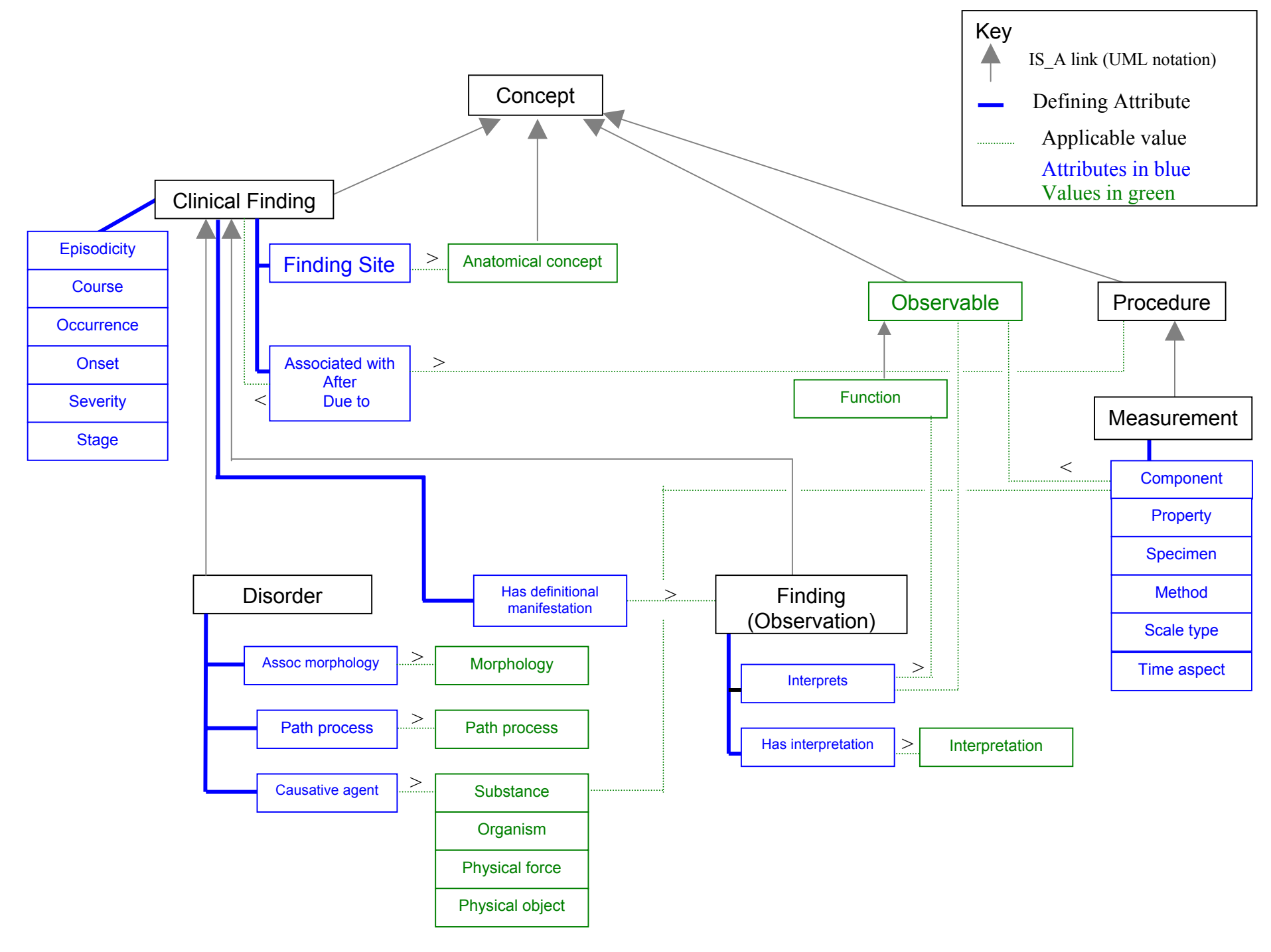
Clinical Findings Attributes

Defining Attributes	Permissible Values
*Finding site	*Anatomical concept* 257728006
*Associated morphology	*Morphologically abnormal structure* 49755003
Associated with	*Clinical Finding* 404684003, *Procedure* 71388002, *Event* 272379006, *Organism* 410607006, *Substance* 105590001, *Pharmaceutical/biological product* 373873005, *Physical object* 260787004, or *Physical force* 78621006.
--Causative agent	*Organism* 410607006, *Substance* 105590001, *Pharmaceutical/biological product* 373873005, *Physical object* 260787004, or *Physical force* 78621006.
--Due to	*Clinical Finding* 404684003, *Procedure* 71388002, *Event* 272379006
--After	*Clinical Finding* 404684003, *Procedure* 71388002, *Event* 272379006
Severity	*Severities* 272141005
Onset	*sudden* 255363002, *gradual* 255343009
Course	*Courses* 288524001
Episodicity	*Episodicities* 288526004
§Interprets	*Observable entity* 363787002, *Laboratory procedure* 108252007.
§Has interpretation	*Findings values* 260245000, *Result comments* 281296001
Pathological process	*Pathological process* 308489006
Has definitional manifestation	*Clinical finding* 404684003
Occurrence	*Periods of life* 282032007
Stage	*General clinical stage for disease AND/OR neoplasm* 106240007

* FINDING SITE and ASSOCIATED MORPHOLOGY should be grouped

§ INTERPRETS and HAS INTERPRETATION should be grouped.

-- Denotes sub-role of ASSOCIATED WITH


Key
IS_A link (UML notation)
Defining Attribute
Applicable value
Attributes in blue
Values in green
Concept
Clinical Finding
Episodicity
Course
Occurrence
Onset
Severity
Stage
Finding Site
Anatomical concept
Associated with
After
Due to
Observable
Function
Procedure
Measurement
Component
Property
Specimen
Method
Scale type
Time aspect
Disorder
Has definitional manifestation
Finding (Observation)
Assoc morphology
Morphology
Path process
Path process
Causative agent
Substance
Organism
Physical force
Physical object
Interprets
Has interpretation
Interpretation

Procedure Attributes

Defining Attributes	Permissible Values
Method	*Action* 129264002
*Procedure site *Procedure site direct *Procedure site indirect	*Anatomical concept* 257728006
*Procedure device *Direct device *Indirect device *Using *Access instrument	*Device* 49062001, *Physical force* 78621006 *Device* 49062001 *Device* 49062001 *Device* 49062001, *Physical force* 78621006 *Endoscope* 37270008
*Procedure morphology *Direct morphology *Indirect morphology	*Morphologically abnormal structure* 49755003
*Direct substance	*Substance* 105590001, *Pharmaceutical/biologic product* 1212316016
*Access	*Surgical access values* 309795001
*Approach	*Procedural approach* 103379005
Has intent	*Intents* 363675004
Has focus	*Clinical finding* 404684003, *Procedure* 71388002
Priority	*Priorities* 272125009
Revision status	*Primary operation* 261424001, *Revision – value* 255231005, or *Part of multistage procedure* 257958009.
Recipient category	*Person* 125676002, *family* 125676002, *community* 133928008, *Donor* 261008006

* These attributes should be grouped with the *method* attribute to which they apply.

Measurement Attributes

Additional Attributes for Measurement Procedures Only (6)	
Defining Attributes	**Permissible Values**
Component	*Substance* 105590001, *Observable entity* 363787002, *Cell structure* 4421005, *Organism* 410607006
Measurement method	*Laboratory procedure categorized by method* 127789004
Scale type	*Quantitative* 30766002, *Ordinal value* 117363000, *Ordinal or quantitative value* 117365007, *Nominal value* 117362005, *Narrative value* 117364006, *Text value* 117444000
Time aspect	*Time frame* 7389001
Property	*Measurement property* 118598001
Specimen	*Specimen* 123038009

Other Attributes

Attributes for Modeling Context-dependent Concepts (6)

Defining Attributes	**Permissible Values**
Temporal context	*Temporal context value (qualifier value)* 410510008
Subject relationship context	*Person* 125676002
Associated finding	*Clinical finding* 404684003
Finding context	*Finding context value (qualifier value)* 410514004
Associated procedure	*Procedure* 71388002
Procedure context	*Context values for actions (qualifier value)* 288532009

Attribute for Modeling Body Structures (1)

Defining Attributes	**Permissible Values**
Laterality	*left* 7771000, *right* 24028007, and *bilateral* 51440002

Attribute for Modeling Pharmaceutical/Biologic Products (1)

Defining Attributes	**Permissible Values**
Has active ingredient	*Substance* 105590001

SNOMED CT concepts in context

- Context refers to the effects of embedding a concept code in a clinical statement
  - A code is embedded in a clinical statement when it is used in a clinical record
- Embedding a code in a clinical statement
  - Adds information
    - Date of finding or action
    - Author, performer, etc.
  - May also elaborate its meaning in one of several ways
    - Subtype qualification
    - Axis modification
    - Affirmation or Negation
    - Combination

Context: Representation

- Context may be represented in various ways
- Pre-coordination of SNOMED CT concepts
  - Example
    - 160303001
    - Family History of diabetes mellitus
- Post-coordination of SNOMED CT concepts that together express a more specific concept
  - Example
    - 57177007+(246090004=73211009)
    - Family history + (Assoc. finding = Diabetes mellitus)
- Structures and attributes specific to a proprietary data model or a standard reference model
  - Examples from HL7 RIM
    - ActRelationship class
    - moodCode attribute

Subtype qualifiers & Axis modifiers

- Subtype Qualification
  - An elaboration of a concept that results in a concept that is a subtype of the original unelaborated concept
    - Examples
      - Asthma + severe
      - Fracture reduction + open approach
- Axis Modification
  - An elaboration of a concept that results in a concept that is not a subtype of the original unelaborated concept
    - Examples
      - Asthma + Family History
      - Hip replacement + Planned

Context: Soft defaults

- Many SNOMED concepts have a “soft default” set of context attribute values
  - The defaults apply when a concept is placed in the context of a record without any explicit expression of context
  - If alternative values are applied these may either
    - Refine the context
      - Family history of asthma → Family history - father has asthma
    - Axis modify
      - asthma” → “family history of asthma”
  - An attribute in some cases may behave as an axis modifier and in other cases as a subtype qualifier
- The values of some context attributes are more rigidly fixed
  - In these cases the context can be refined but cannot be shifted out of its axis

Context Model: Clinical findings & disorders

- ***Finding context***
  - Affirmed (known present) (default)
  - Negated (known absent)
  - Uncertain
  - Expectation
    - Prognosis / Goal / At risk, etc
- ***Temporal context***
  - Current or specified time (default)
  - Past
- ***Subject relationship context***
  - Subject of record (default)
  - Family member
    - Specific family members
  - Others
    - Donor / Fetus / Contact, etc.

Context Model: Procedures

- ***Procedure context***
  - Pre-start
    - To be done
  - Post-start
    - In-progress
    - Ended
      - Done (default)
- ***Temporal context***
  - Current or specified time (default)
  - Past
- ***Subject relationship context***
  - Subject of record (default)
  - Family member
  - Other

Context: Affirmation & Negation

- A concept may be affirmed or negated in a clinical situation
  - e.g. No headache, No FH of heart disease
- Negation has profound effects
  - Inverts the meaning of a concept
  - Changes to direction of subsumption
    - Affirmation also affirms super-types
      - Fracture of femur
      - Is there a “fracture of a bone”? – Yes
      - Is there a “fractured neck of femur”? – Not known
    - Negation also negates sub-types
      - No fracture of femur
      - Is there a “fracture of a bone”? – Not known
      - Is there a “fractured neck of femur”? – No

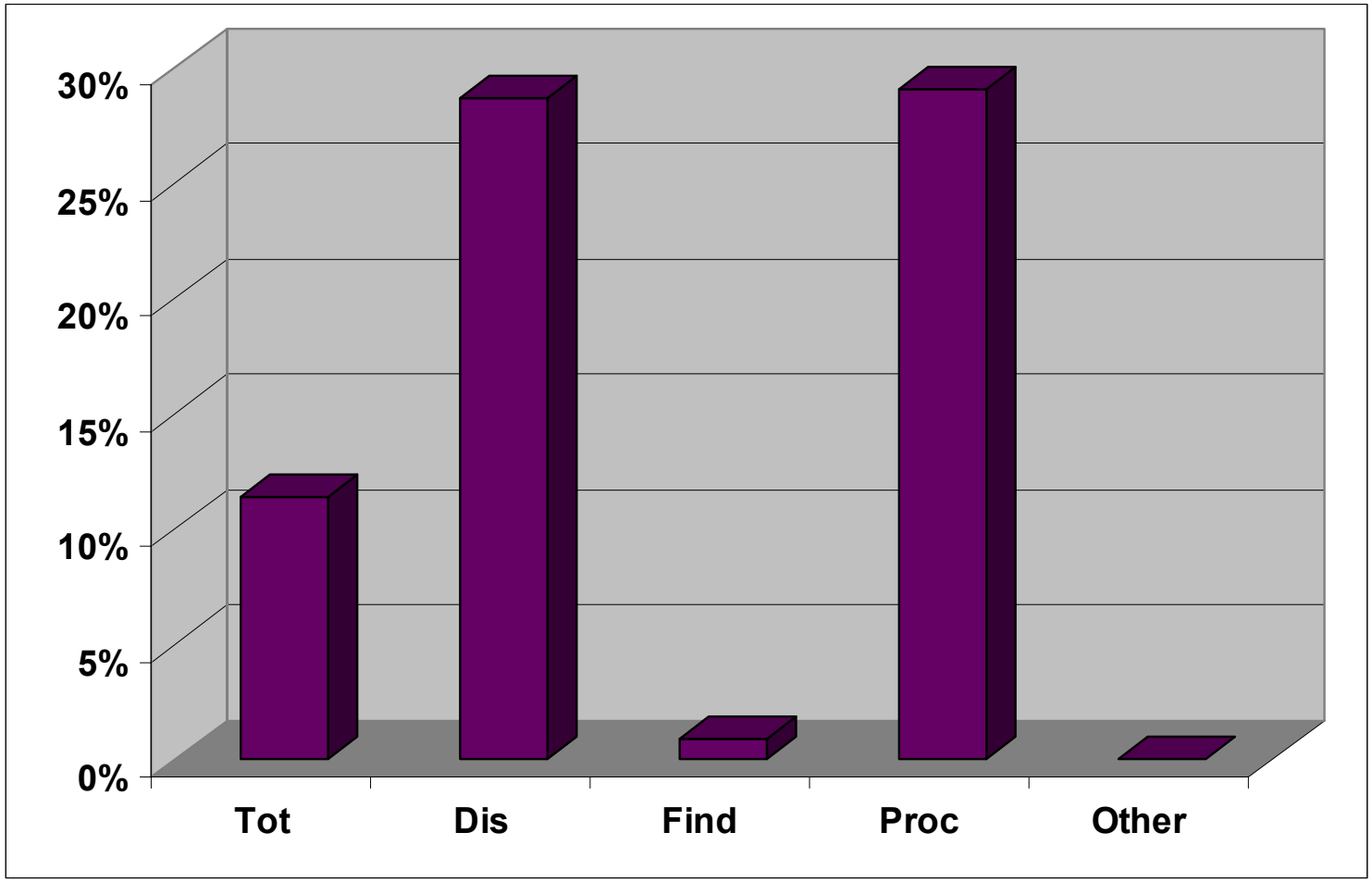
Context: Combination & Linkage

- Combination
  - Two or more assertions may be embedded in a clinical situation in a way that links them together
- Examples of combinations that invoke context (affirmation, negation, etc):
  - Simple combination of concepts
    - “Head injury” with “loss of consciousness”
  - One concept present and another absent
    - “Head injury” without “loss of consciousness”
- Examples of combinations that do not invoke context:
  - An explicit typed relationship between concepts (this is not axis modifying, so is a subtype qualifier not an axis modifier)
    - “Pain in hand” due to “rheumatoid arthritis”
  - Temporal relationships between statements
    - “Pain in knee” following “exercise”

X with Y: not simple

- Head injury with LOC
  - Most complete definition of this (simple) phrase is:
    - Two things are affirmed present (invoking context):
      - Head injury
        - » and
      - LOC due to head injury
  - Need to clearly distinguish from
      - LOC
        - » And
      - Head injury following LOC

Percentage of SNOMED CT concept codes that are “fully defined”



Eventually should reach ~70% or more of disorders, findings & procedures

A Final Note About Process

- Open working group meetings + on-line discussion forums
- Active working groups:
  - Concept model working group
  - Mapping working group
  - Content-area focused working groups
    - Primary care
    - Nursing
    - Genomics
    - Anesthesiology, pathology, dermatology, ophthalmology, …
- Upcoming in-person meeting dates:
  - Feb 2, 2005, S. California
  - June 14-15, 2005, Chicago
  - Oct 5, 2005, London

- Internationalization agenda