Standardization of Clinical Terminology for the EHR

Kent A. Spackman, MD, PhD
Professor, Dept. of Medical Informatics & Clinical Epidemiology,
Oregon Health & Science University, Portland OR
Chair, SNOMED International Editorial Board

Stanford Clinical Informatics Seminar
April 29, 2005

Standardizing terminology for the EHR

- What is it, and what do we hope to gain by standardizing?
- Some alternatives & misconceptions
- Why is it hard to do?
- An approach to standardization – the example of SNOMED Clinical Terms

What is terminology standardization?

- Creation of a terminological resource that can be implemented in software applications to represent clinically relevant information
  - In a semantically structured form that can be used by automated applications
- Codes with explicit formal definitions of meaning
- Consensus based on real clinician usage

What is terminology for?

- It is for building applications capable of:
  - Recording statements about the health and health care of individuals
    - In a way that permits retrieval according to the meaning of the statements, rather than just the words used
  - Retrieving individual cases and groups of cases
    - To enable more automated and sophisticated decision support, epidemiology, and research

What is terminology for?

- Semantic interoperability

Expected benefits of semantic interoperability

- Reduction of errors
  - Elimination of errors of omission via reminders
  - Elimination of errors of commission via alerts
- Management of costs
  - Elimination of redundant testing and investigation
- Monitoring and responding to trends & problems in the health of populations
- Expanding knowledge of diseases, treatments and outcomes

Requirements to achieve these benefits:

- Automation of systems that deal with health information requires clinical data that:
  - is recorded at the appropriate level of detail
    - not forced to be either too general or too specific
  - is consistent over time and across boundaries
  - can be transmitted without loss of meaning
  - can be aggregated at more general levels, and along multiple different perspectives
  - can be interpreted by automated systems

Misconceptions

- You can solve interoperability if you just . . .
  - Get everyone to say it the same way
  - Make a huge list of everything people have said
  - Make a common data dictionary for fields that can be filled out (like “zip code” etc)
  - Aggregate multiple small lists from each clinical specialty
  - Use special-purpose NLP or Google-style indexing and search

Misconceptions

- Achieving interoperability requires ordinary practitioners to . . .
  - Code everything
  - Use all of the terminology available
  - Document more than they really know
  - Document findings with more certainty than they actually have

Semantic interoperability is hard

- At least 3 major types of standards, and ***their interactions***, have to be taken into account
  - Terminologies / ontologies
  - Information models / architectures
  - Standards for decision support rules / guidelines

Information model – terminology model

The simplest information model:

Put all clinical data here: _____

(everything else is in the terminology used to fill in the blank)

The simplest terminology model:

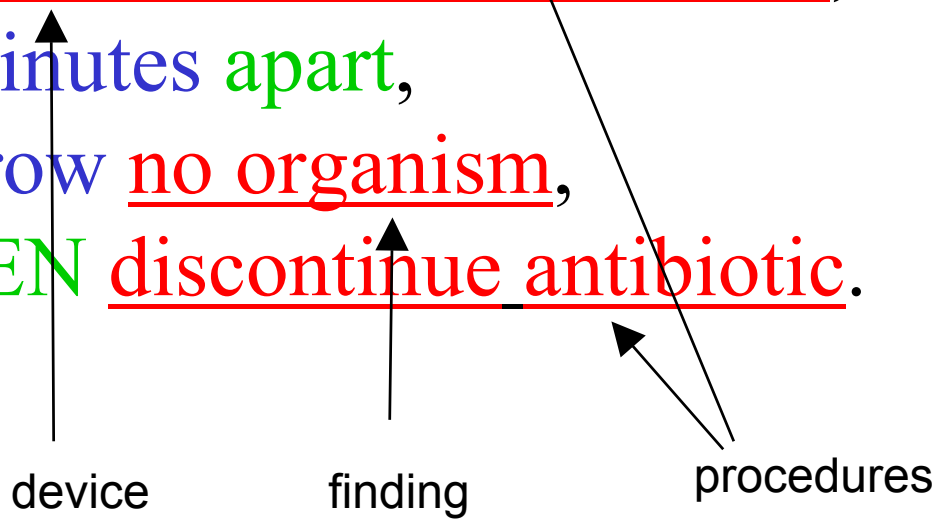
Two values: Yes & No

(everything else is a field name)

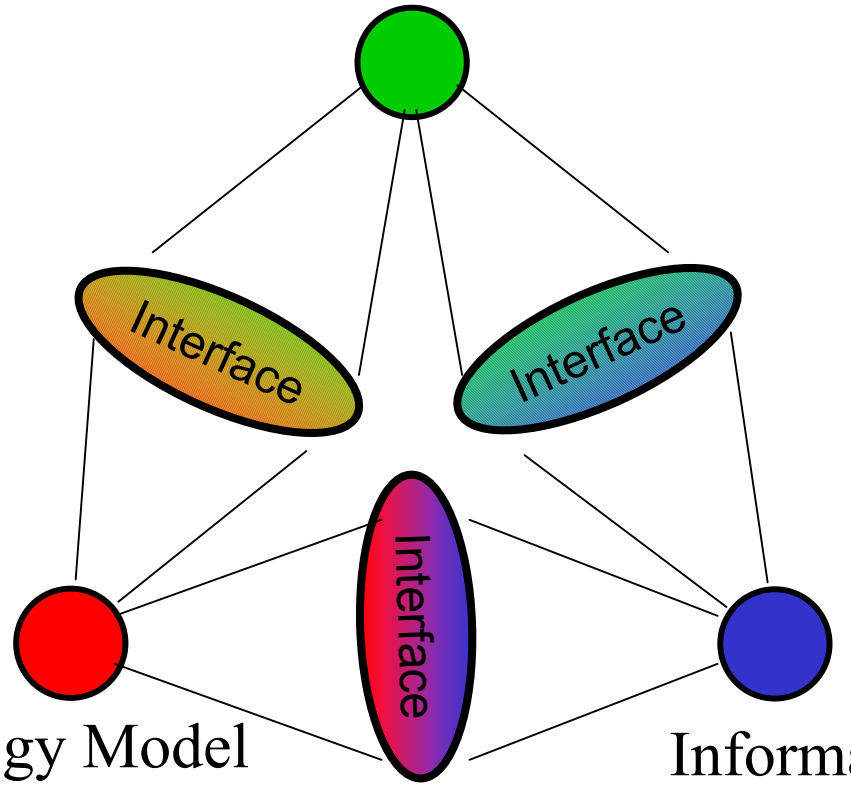
What about clinical decision support?

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

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



Clinical Decision Support Model
+ Inference Rules



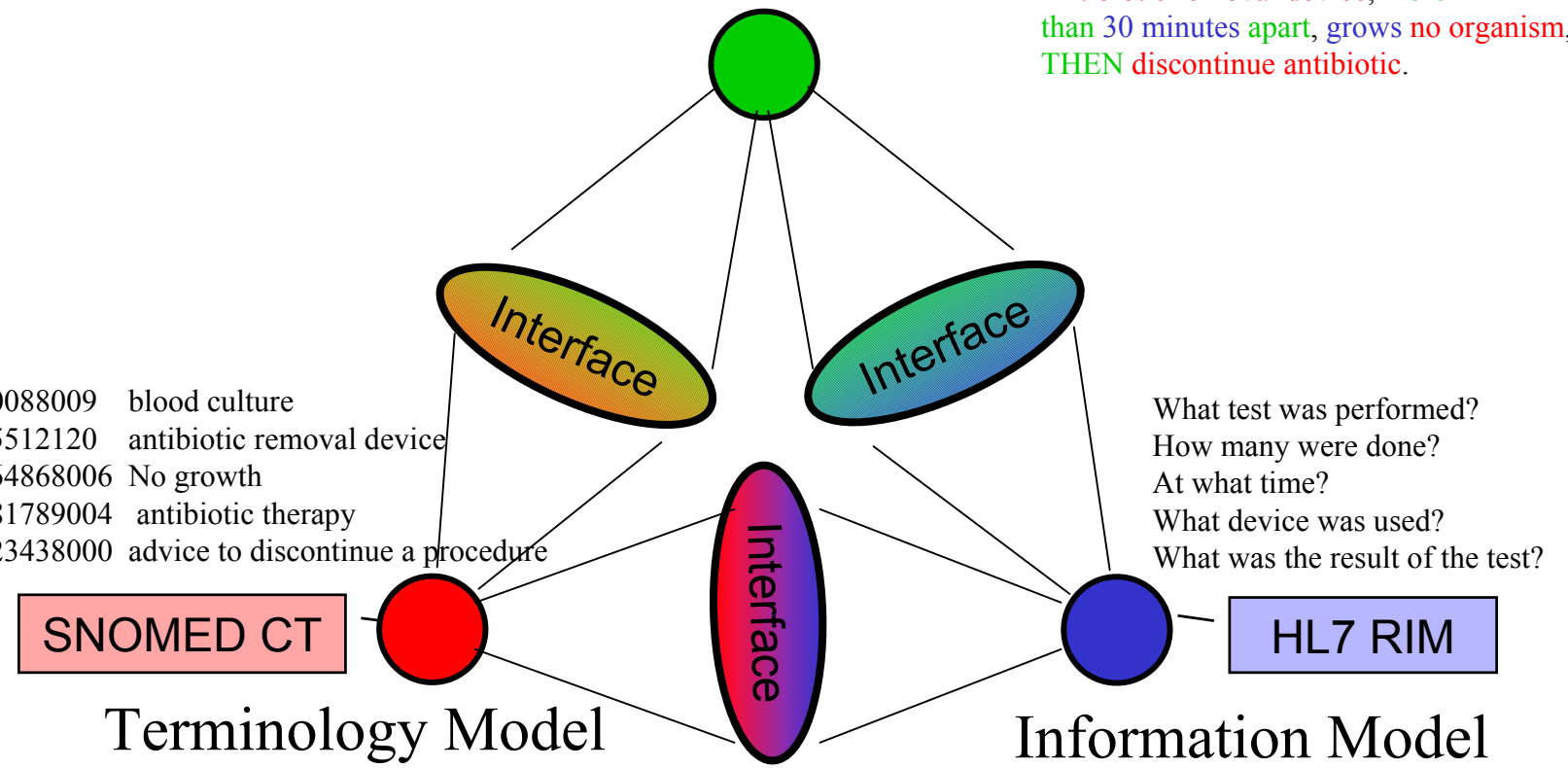
Terminology Model
+ Coded Data

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
+ Coded Data

Information Model
+ Patient Data Structures

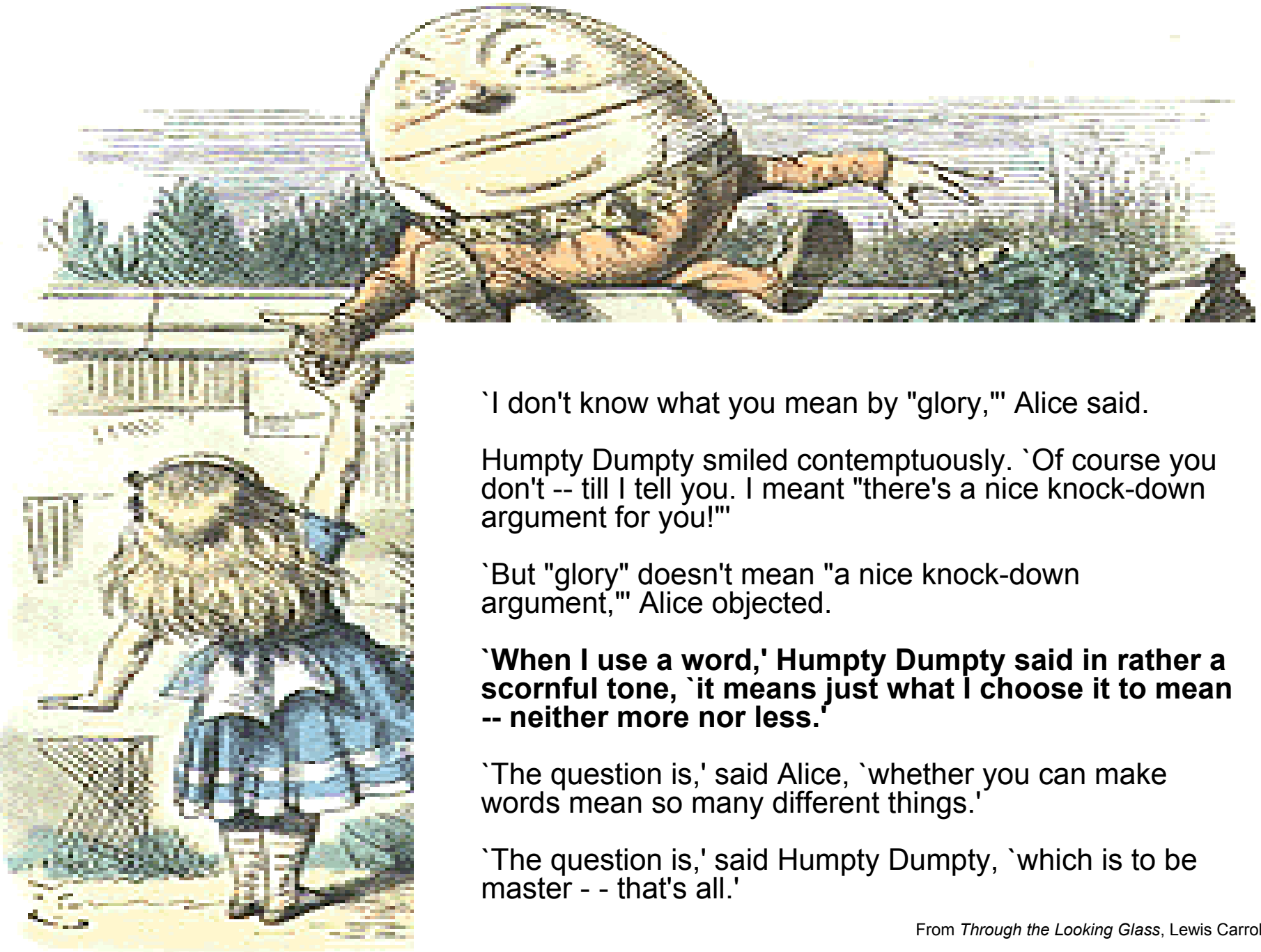
Diagram based on Figure 1 in Rector AL et al. “Interface of Inference Models with Concept and Medical Record Models” AIME 2001: 314-323

What are we standardizing?

- “It’s not what you say, it’s what you mean”

It is notoriously difficult to tell what people mean just by what they say

- From "The Economist", Charlemagne column, Sept 4, 2004
- "Decoding a Euro-diplomat takes more than a dictionary"
  - "Up to a point" means "I agree in part"?
    - Actually, it means: "No, not in the slightest"
  - "I hear what you say" means "He accepts my point of view"?
    - Actually, it means: "I disagree and do not want to discuss it any further"
  - "With the greatest respect" means "He respects my opinion"?
    - Actually, it means: "I think you are wrong, or a fool"
  - "By the way" means "This is not very important"?
    - Actually, it means: "The primary purpose of our discussion is …"
  - "I'll bear it in mind" means "I will take action based on it"?
    - Actually, it means: "I'll do nothing about it"
  - "Correct me if I'm wrong" means "I'm not sure about this"?
    - Wrong, it means: "I'm right, don't contradict me"



`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

Words alone are insufficient

1. There are national, regional and local variations in meaning of words and phrases (even within the same language)
2. Multiple meanings with the same “preferred name”
3. Combining words results in a meaning entirely different from the sum of the parts
4. Ambiguous shorthand and abbreviations are common
5. The same phrase means different things to different specialists
6. The same word or phrase means different things depending on what you are doing at the time
7. Significant differences in meaning are often obscured through use of the same word
8. Successful communication relies on making ontological distinctions that are ignored by common phrasing
9. Formal definitions are often at variance with common clinical usage
10. A general name often takes on a more specific meaning
11. A manifestation is often used to name the disorder in which it occurs

Dialect variations within a language

- What is pudding?
  - At dinner in Phoenix, Roger (from the UK) asked “Is anyone having pudding?”
  - To which I replied, “Do you mean dessert?”
  - And he said, “No, I mean pudding.”

?

One word can have multiple published definitions

- What is the scalp?
  - scalp: the skin covering the cranium (Stedman's)
  - scalp: the soft tissue envelope of the cranial vault, consists of 5 layers: the skin, connective tissue, epicranial aponeurosis + occipitofrontalis muscle, loose areolar tissue, and pericranium. (Gardner, Gray & O'Rahilly, anatomy text)
  - Epicranium (Stedman's): the muscle, aponeurosis and skin covering the cranium

We say like Humpty Dumpty "When ***I*** use the word scalp, …"

One phrase can have multiple published "authoritative" meanings

- What is juvenile rheumatoid arthritis?
  - Seropositive chronic idiopathic arthritis in child < 16 yrs ?
  - Any chronic arthritis in child < 16 yrs?
  - Is Adult-onset Still's disease included?

- Three different published terminology standards, all incompatible
  - JRA (juvenile rheumatoid arthritis) – US
  - JCA (juvenile chronic arthritis) – UK
  - JIA (juvenile idiopathic arthritis) – International

Word combinations ≠ the sum of their parts

- What is a "pyogenic granuloma" ?
  - Pyogenic = pus forming
  - Granuloma = a collection of inflammatory cells of a particular type

- Pyogenic granuloma = a benign tumor of small blood vessels of the skin
- It is neither *pyogenic* nor a *granuloma*.

Ambiguous shorthand and abbreviations

- What is “general paresis”?
  - General = affecting all skeletal muscles
  - Paresis = weakness
- GPI = a form of tertiary neurosyphilis characterized by generalized weakness

Same phrase means different things to different specialists

- What is “acute inflammation”?
  - To the GP, it is inflammation with an acute onset, characterised by redness, heat, swelling and pain.
  - To the pathologist, it is inflammation in which polymorphonuclear leukocytes predominate, as opposed to chronic inflammation, in which “mononuclear cells” (lymphocytes, plasma cells, monocytes, histiocytes) predominate.

What you are doing at the time changes the meaning of words

- What is the “fundus”?
  - When caring for a pregnant patient –
  - When examining the eyes –
  - When doing a gastroscopy –
  - When doing a cholecystectomy –

Subtle distinctions are often implicit

- What is a laceration?
  - Torn or jagged wound

    vs
  - Accidental cut wound

- Perineal **laceration** during O-P delivery

  vs
- **Laceration** of thumb while using kitchen knife

Formal ontologists insist on a difference between matter/stuff and individual entities

- What is aspirin?
  - Some **aspirin** – the chemical ASA
  - An **aspirin** – a tablet containing ASA

Is there an error in this hierarchy?

- Radiographic procedures
  - Angiography procedures
    - Magnetic resonance angiography procedures

Is there an error in this hierarchy?

Radiographic procedures

Angiography procedures

Magnetic resonance angiography procedures

It is common for a general name to acquire a more specific meaning

Is there an error in this hierarchy?

psoriasis

- psoriasis with arthropathy
  - juvenile psoriatic arthritis
    - juvenile psoriatic arthritis without psoriasis

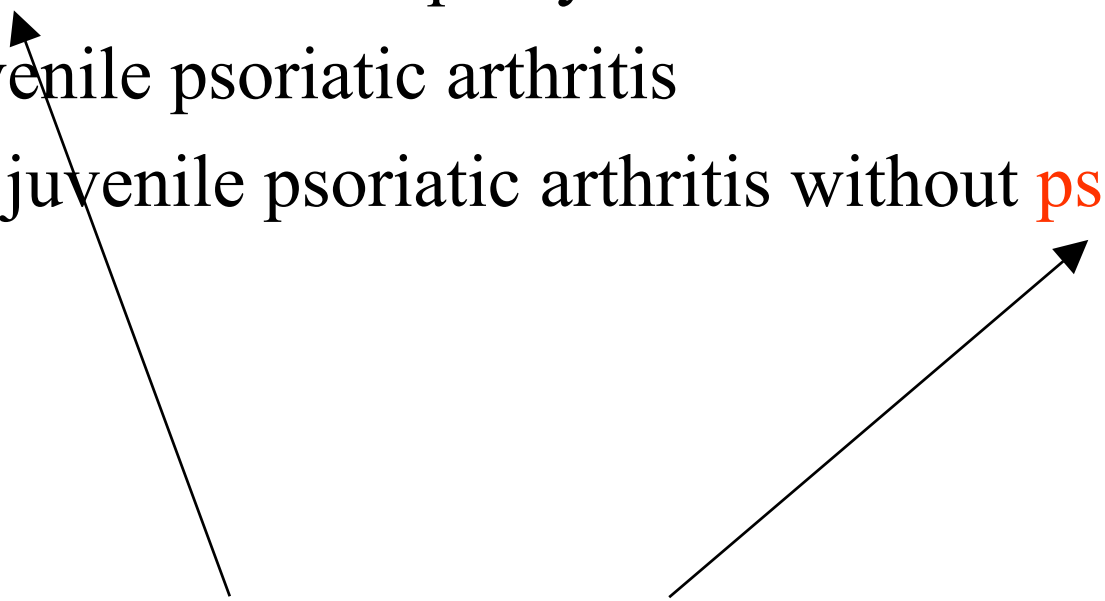
Is there an error in this hierarchy?

psoriasis

psoriasis with arthropathy

juvenile psoriatic arthritis

juvenile psoriatic arthritis without psoriasis



It is common for the disorder and its manifestation to have the same name

Some formal definitions are in conflict with ordinary usage

- What is the leg?
  - 1) same as "lower limb"
  - 2) just the part from the knee to the ankle
    - Stedman's "the segment of the inferior limb between the knee and the ankle"
    - Dorland's "that section of the lower limb between the knee and ankle"

SNOMED
Systematized Nomenclature of Medicine

- A clinical terminology standardization effort
- An organization supporting that effort
- A database containing the codes and terms and knowledge representation resulting from the effort

Rationale for the Development of SNOMED Clinical Terms

- SNOMED CT was developed in response to the needs of the users
  - In particular, a need for better reference properties of the terminology
- User needs remain the driving motivation behind decisions about what SNOMED will or will not undertake

Users Told Us That Existing Terminologies Were Inadequate

- Lack of content coverage at the desired level of generality / specificity
  - ICD-9-CM and CPT codes usually too general
- Lack of computability
  - e.g. single hierarchy in ICD
- Lack of single consistent meaning for concepts and their relationships
  - e.g. UMLS “anemia”, “hypertension”, “aspirin”
- Perceived deficiencies in change management and updates
  - re-use of codes

Users Told Us That Existing Terminologies Were Inadequate

- Need formal concept representation principles
  - To help solve the dilemma of pre-coordination vs post-coordination, also called enumeration vs composition
  - To permit complete and consistent retrieval and aggregation of data

Large organizations implementing systems using SNOMED CT as a terminology standard

- Kaiser Permanente
- United Kingdom National Health Service
  - National Programme for Information Technology (NPfIT)
- US Veterans Administration

Settings Where SNOMED CT is in Current Use

- Surgical Pathology Departments
- Clinical Laboratories
- General Practices
- Internal Medicine Clinics
- Public Health Agencies

Comprehensive

- Can a single terminology be truly comprehensive?
  - It depends on how you frame the question.
  - Comprehensive across what scope?

Carving Up Clinical Terminology: What Doesn't Work

Vertical domains overlap too much.

The "separate interlocking" idea ignores high cost of integration.

Examples of Overlap

- Measurement of phenytoin level (lab, drugs)
- Diabetic foot exam (nursing, allied health, GP, specialists)
- Diabetes mellitus type 2
- Headache
- …

Carving Up Clinical Terminology: What Might Work

A Global Clinical Terminology Need Not Have:

- "Released from restriction order under Section 42 of the Mental Health Act 1983 (England and Wales)"
  - This has a Read code: Xa9Dh
- "Multnomah County (Oregon) Jail Cell Number"
  - This concept was added to the MedicaLogic / GE Medical Systems data dictionary for a Portland, Oregon implementation site

Concepts in the Global Terminology Are Not Specific to a Particular:

- Government realm (nation, region, state/province, county, etc.)
- Corporate Entity (HMO, private practice, etc.)
- Implementation site (hospital, outpatient clinic, lab)
- Software package (EHR, EMR, etc.)

- This means that concepts that *are* uniquely specific to these entities will have to be represented separately

Towards Implementation: Global Scope ≠ 'Big Bang' Implementation

- The terminology itself must be global
- Implementations must be local
  - Local: limited in scope
  - Local: specialized for the site of use

Integration: Catch-22 ?

- Terminology must be **integrated** across professional groups and specialties
- Terminology development is usually done by professional groups
  - To make their work possible, they limit it according to the concepts and terms they commonly use
- No over-arching professional health care organization existed to carry out the integration

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 ?)

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

Convergence

- SNOMED brings many terminologies together into one integrated whole

Approach to convergence

- Leverage existing work
  - acknowledge and reach agreement on intellectual property issues
  - expand and refine a common base
- Properly handle differences of meaning:
  - a process for discovering the differences
  - a formalism in which to represent them
- Provide an organization that can responsively & authoritatively maintain the terminology

Some of the Specialty and Professional Groups Working with SNOMED:

- Primary care
- Surgery, pediatrics, internal medicine, mental health, ob/gyn
- Pharmacy/drugs
- Pathology
- Anesthesiology
- Radiology
- Dermatology
- Ophthalmology
- Nursing
- Veterinary medicine
- Dentistry

Primary Care

NHS Information Authority

- Most GPs (general practitioners) in the UK currently use the Read Codes (version 2)
- Regular updates to the Read Codes are also added to SNOMED CT

Medicine, Surgery, Peds, OB/GYN, Mental Health

KAISER PERMANENTE®

- Kaiser's CMT (convergent medical terminology) project
  - a key component of its electronic health record projects
  - has been actively working with SNOMED since 1996

Pathology

- CAP committees continue to contribute to the maintenance of SNOMED
- Cancer checklists are the main area of active expansion
- ICD-O collaboration continues for morphology (M-codes) and topography (T-codes)

Anesthesiology

- IOTA: International Organization for Terminology in Anesthesia
  - Originated with the Anesthesia Patient Safety Foundation (APSF) Data Dictionary Task Force
  - Working with the British SCATA (Society for Computing and Technology in Anaesthesia)

Radiology

- The Digital Imaging Communications (DICOM) standard uses SNOMED codes for its structured reporting standard
- Regular additions to SNOMED are received from DICOM for image and image-related documentation messaging

Dermatology

- The British Association of Dermatology has produced a Diagnostic Index with over 5,300 diagnostic concepts in dermatology
- They are working with SNOMED to ensure the appropriate representation of these concepts in SNOMED CT, and to produce a subset for dermatologist to “navigate” these concepts

Ophthalmology

- The American Academy of Ophthalmology has supported several ophthalmologists to work with SNOMED on eye-related terminology

Nursing

- SNOMED incorporates integrated concepts from, and maps to, the following nursing terminologies:
  - NANDA (North American Nursing Diagnosis Association)
  - NIC (Nursing Interventions Classification)
  - NOC (Nursing Outcomes Classification)
  - PNDS (Perioperative Nursing Data Set)
  - Omaha
  - Georgetown
  - Home Health Care Classification

Dentistry

- The American Dental Association has produced a subset of SNOMED for dentists, called “SNODENT”
- There is also a mapping to CDT2 (Current Dental Terminology), which is used for dental billing
- There is an ADA liaison on the SNOMED Int’l Editorial Board

Veterinary Medicine

- The American Veterinary Medicine Association (AVMA) has collaborated with SNOMED continuously since the 1980's
- One of the four primary editors of SNOMED 3 was a veterinarian
- A substantial number of veterinarians use SNOMED, and the AVMA provides support for a web site and moderated discussion list
- Provides unique contributions in the area of comparative medicine, teratology, and living organisms, as well as substantial general clinical quality improvements

Drugs

- UK-specific manufacturing and proprietary brand names are maintained in a UK-specific extension
- This extension will be used for prescribing and dispensing (replacing the Read codes)
- Generic clinical drugs (name, strength, dose form) are maintained in the SNOMED "core" for US and UK

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

SNOMED's approach?

- Careful representation of meaning
  - Consensus process
    - URU criteria: understandable, reproducible, useful
  - Evolutionary design
  - Formal description logic foundation

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

Description Logic Foundation

- SNOMED is based on the description logic known as $\mathcal{ELH}$
  - Conjunction
  - Existential restrictions
  - Role hierarchies
- Plus "role groups" (Ref: Spackman et al. 2002 AMIA Proceedings)
- Plus role composition
  - So far, only one:
    - direct-substance o 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)	$R \sqsubseteq S$	$R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$

DL features we might want to consider

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

Building terminology standards is time-consuming and expensive, despite extensive volunteer effort

- Read Codes
  - 34 million pounds through 1999
- SNOMED RT
  - 17 million dollars over 5 years 1997-2002
- Merger of the two -> SNOMED CT
  - Additional 9 million dollars over 3 years 1999-2002
- Not counting multiple millions of dollars per year in in-kind contributions from Kaiser Permanente and clinical organizations and volunteers

SNOMED Emergence as a Standard

- 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

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, anesthesiology, pathology, dermatology, ophthalmology, …
- Upcoming in-person meeting dates:
  - June 14-15, 2005, Chicago
  - Oct 5, 2005, London
- Internationalization agenda