Data and Knowledge Representation Lecture 5

Last Time We Talked About

- Medical vocabulary
- Survey of coding systems

Today We Will Talk About

- Continue survey of medical coding systems
- UMLS

SNOMED

- Systematized Nomenclature of Human and Veterinarian Medicine
- Developed by the College of American Pathologists (1971)
- SNDO, SNOP, SNOMEDSNOMED
- SNOMED International
- SNOMED-RT (Reference Terminology)
- SNOMED-CT (merger with Read)

Slides borrowed from Dr. Cimino J of Columbia Univ.

SNOMED International

Chemicals, Drugs, and Biological Products	14,846	
Diseases/Diagnoses		35,834
Function		19,221
General Linkage/Modifiers		1,569
Living Organisms		24,614
Manufacturers of Pharmaceuticals Human/Veterinary		363
Morphology		5,875
Occupations		1,949
Physical Agents, Forces and Activities	1,600	
Procedures		30,723
Social context	1,013	
Topography		12,936
		150,343

Slides borrowed from Dr. Cimino J of Columbia Univ.

SNOMED III - Coding Examples

"D3-15000" "01" "Myocardial infarction, NOS" "(T-32020) (M-54700)"

T-32020 = Myocardium, NOS

M-54700 = Infarction, NOS

"D3-15010" "01" "Microinfarct of heart" "(T-32000) (M-54701)"

T-32000 = Heart, NOS

M-54701 = Focal Infarct

Mother died of myocardial infarct

S-10120, S-13030, D3-15000

S-10120, F-A7860, T-32020, M-54700

SNOMED RT

Fully Specified Name: Myocardial infarction (disorder)

Concept ID: 22298006

Definition:

Is a (attribute) Myocardial disease (disorder)

Is a (attribute) Structural disorder of heart (disorder)

Associated morphology (attribute) Infarct (morphologic abnormality)

Finding site (attribute) Myocardium structure (body structure)

SNOMED RT

Fully Specified Name: Heart disease in mother complicating pregnancy, childbirth AND/OR puerperium (disorder)

Concept ID: 78381004

Definition:

Is a (attribute)Cardiac complication (disorder)

Is a (attribute)Complication related to pregnancy (disorder)

Finding site (attribute) Heart structure (body structure)

Qualifiers:

Onset (attribute)Onsets (qualifier value)

Severity (attribute)Severities (qualifier value)

Episodicity (attribute)Episodicities (qualifier value)

Course (attribute)Courses (qualifier value)

SNOMED RT

Fully Specified Name: Needle biopsy (procedure)

Concept ID: 129249002

Definition:

Is a (attribute) Biopsy (procedure)

Method (attribute) Biopsy - action (qualifier value)

Using (attribute) Biopsy needle, device (physical object)

Qualifiers:

Priority (attribute) Priorities (qualifier value)

Appropriate values for the Priority (attribute) relationship type

Deferred (qualifier value)
Denied (qualifier value)
Elective (qualifier value)
Emergency (qualifier value)
Immediate (qualifier value)
Reclassified (qualifier value)
Reclassified and rescheduled (qualifier value)
Repeat elective (qualifier value)
Repeat emergency (qualifier value)
Rescheduled (qualifier value)
Routine (qualifier value)
Scheduled (qualifier value)
Urgency (qualifier value)

Read Clinical Codes

- Developed by James Read in the 80s
- Adopted by UK NHS in 1990
- Allows post-coordination
- Merged with SNOMED

READ

182..A Y7CmDC P Chest pain
Xa0wWK Y7CmFC P Pleurodynia
182Z.A Y7CmGC P Chest pain NOS
Xa0wWK Y7CmIC S Painful breathing -pleurodynia
1826.A Y7CmJC P Parasternal pain
1823.A Y7CmLC P Precordial pain
1821.A Y7CmNC P Chest pain not present
X75rWC Y7CmYC P Pain in heart
1829.A Y7CmZC P Retrosternal pain

Gabrieli Medical Nomenclature

- Single large hierarchy
- More complex terms as you move down
- Being adopted by ASTM as a standard

Nursing terminologies

- Many initiatives worldwide
- North American Nursing Diagnosis Association (NANDA) codes
- Nursing Outcomes Classification (NOC)
- Georgetown Home Health Care Classification (HHCC)
- Omaha System
  - Problems, interventions, outcomes

GALEN

- European initiative
- Reference model for medical concepts
- Formalism called Structured Meta Knowledge
  - Similar to description logic

LOINC

- Logical Observations, Identifiers, Names, Codes (LOINC)
- Consortium led by Clem McDonald and Stan Huff
- Originally lab results
- Now extended to include clinical observations
- Recently, merged into SNOMED

National Drug Codes

- Developed by FDA
- Widely used in US
- Codes based on drug manufacturer
- Codes have little class hierarchy
- Codes are reused at manufacturer's discretion

MeSH

- Medical Subject Headings
- Developed by NLM
- Indexes medical literature
  - Medline
- Terms are in hierarchies and appear in multiple places in hierarchies

Unified Medical Language System

- A long term NLM project
- Designed to facilitate the retrieval and integration of biomedical information from various sources
- Components
  - Metathesaurus
  - Semantic Network
  - SPECIALIST Lexicon and Lexical Programs
  - (Information Source)

Metathesaurus

- Metathesaurus Concept Names
  - MRCON
- Relationships between Different Concept Names
  - MRREL, MRCOC, MRATX
- Attributes
  - MRSAT, MRDEF, MRSTY, MRLO, MRRANK
- Source Information and contexts
  - MRSO, MRCXT
- Indexes
  - MRXW.BAQ, MRXW.DAN, MRXW.DUT, MRXW.ENG, MRXW.FIN, MRXW.FRE, MRXW.GER, MRXW.HEB, MRXW.HUN, MRXW.ITA, MRXW.NOR, MRXW.POR, MRXW.RUS, MRXW.SPA, MRXW.SWE, MRXNW.ENG, MRXNS.ENG

MRCON: Concept Names

Col.	Description
CUI	Unique identifier for concept
LAT	Language of Term
TS	Term status
LUI	Unique identifier for term
STT	String type
SUI	Unique identifier for string
STR	String
LRL	Least Restriction Level

MRCON: Concept Names

C0002871|ENG|P|L0002871|PF|S0013742|Anemia|0|
C0002871|ENG|P|L0002871|VP|S0013787|Anemias|0|
C0002871|ENG|P|L0002871|VC|S0352787|ANEMIA|0|
C0002871|ENG|P|L0002871|VC|S0414880|anemia|0|
C0002871|ENG|P|L0002871|VO|S0470197|Anemia, NOS|3|
C0002871|ENG|S|L0280031|PF|S0803242|Anaemia|3|

TS

- P: Preferred Name
- S: Synonym
- s: Suppressible synonym (possibly problematic for some applications, e.g. abbreviations)

STT

- PF: Preferred form of term
- V: Followed by one or more of the following types of variation, in this order:
  - C: Varies from the preferred term only in upper-lower case
  - W: Contains same words as the preferred form, disregarding order and punctuation
  - S: Singular of the preferred form
  - P: Plural of the preferred form
  - O: Other variant of the preferred form

MRREL: Related Concepts

Col.	Description
CUI1	Unique identifier of first concept
REL	Relationship of second to first concept
CUI2	Unique identifier of second concept
RELA	Relationship attribute
SAB	Abbreviation of the source of relationship
SL	Source of relationship labels
MG	Machine-generated and unverified indicator (optional)

MRREL: Related Concepts

C0002871|CHD|C0002891|isa|MSH2001|MSH2001||

C0002871|RB|C0221016||MTH|MTH||

C0002871|RL|C0002886|mapped_to|SNMI98|SNMI98||

C0002871|RO|C0002886|clinically_associated_with|CCPSS99|CCPSS99||

Megaloblastic anemia due to folate deficiency, NOS (C0151482) has clinically_associated_with relationship to Anemia (C0002871)

REL

- RB: has a broader relationship
- RN: has a narrower relationship
- RO: has relationship other than synonymous, narrower, or broader
- RL: the relationship is similar or "alike". PAR: has parent relationship in a Metathesaurus source vocabulary
- CHD: has child relationship in a Metathesaurus source vocabulary
- SIB: has sibling relationship in a Metathesaurus source vocabulary.
- AQ: is an allowed qualifier for the first concept in a Metathesaurus source vocabulary.

RELA

- Any of the relationships defined in the UMLS Semantic Network
- A more specific relationship provided by the source vocabulary identified

SAB

- AIR93
  - AI/RHEUM. Bethesda (MD): National Library of Medicine, Lister Hill Center, 1993.
- ALT2000
  - Alternative Billing Concepts (AltLink). Version 983. Las Cruces (NM): Alternative Link LLC, 2000. Contact: Alternative Link LLC; 1065 S. Main St.; Bldg. C; Las Cruces, NM 88005; phone: (505) 527-0636; fax: (505) 523-4152; http://www.alternativelink.com; mail@alternativelink.com.
- AOD99
  - Alcohol and Other Drug Thesaurus: A Guide to Concepts and Terminology in Substance Abuse and Addiction. 3rd ed. [4 volumes]. Bethesda (MD): National Institute on Alcohol Abuse and Alcoholism (NIAAA) and Center for Substance Abuse Prevention (CSAP), 1999.
- BI98
  - Beth Israel OMR Clinical Problem List Vocabulary. Version 1.0. Boston (MA): Beth Israel Deaconess Medical Center, 1999. Contact: Howard Goldberg, MD.; hgoldber@bidmc.harvard.edu.
- BRMP2001
  - Descritores em Ciencias da Saude [Portuguese translation of MeSH]. Sao Paulo (Brazil): Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/WHO, 2001.
- …….

Semantic Network

- SRDEF: Basic information about the Semantic Types and Relations
- SRSTR: Structure of the Network.
- SRSTRE1: Fully inherited set of Relations (UI's).
- SRSTRE2: Fully inherited set of Relations (names).

STSTRE2

- Acquired Abnormality|isa|Anatomical Abnormality|
- Acquired Abnormality|isa|Anatomical Structure|
- Acquired Abnormality|isa|Physical Object|
- Acquired Abnormality|isa|Entity|
- Acquired Abnormality|affects|Alga|
- Acquired Abnormality|affects|Amphibian|
- Acquired Abnormality|affects|Animal|
- Acquired Abnormality|affects|Bacterium|
- Acquired Abnormality|affects|Bird|

THE SPECIALIST LEXICON AND LEXICAL PROGRAMS

- The SPECIALIST lexicon has been developed to provide the lexical information needed for the SPECIALIST Natural Language Processing System (NLP).
- A general English lexicon that includes many biomedical terms
- “Lexical Methods for Managing Variation in Biomedical Terminologies", A.T. McCray, S. Srinivasan, A.C. Browne, in the Proceedings of the 18th Annual Symposium on Computer Applications in Medical Care, 1994, 235-239.

SPECIALIST LEXICON

{base=abdominal delivery
entry=E0006453
cat=noun
variants=uncount
variants=reg }

Size

- Metathesaurus: About 800,000 concepts, 2 million terms
- Semantic Netowork:134 semantic types and 54 relationships
- Lexicon: About 30,000 words
- Include 56 families of vocabularies
- 13 Languages

Use of UMLS

- Natural Language Processing
  - Liu H, Johnson SB, Friedman C. Automatic Resolution of Ambiguous Terms Based on Machine Learning and Conceptual Relations in the UMLS. J Am Med Inform Assoc. 2002 Nov-Dec;9(6):621-36.
- Information Retrieval
  - Hersh W, Mailhot M, Arnott-Smith C, Lowe H. Selective automated indexing of findings and diagnoses in radiology reports. J Biomed Inform. 2001 Aug;34(4):262-73.

Use of UMLS

- Knowledge Discovery
  - Weeber M, Klein H, Aronson AR, Mork JG, de Jong-van den Berg LT, Vos R. Text-based discovery in biomedicine: the architecture of the DAD-system. Proc AMIA Symp. 2000;:903-7.
- Knowledge Acquisition
  - Weeber M, Klein H, Aronson AR, Mork JG, de Jong-van den Berg LT, Vos R. Text-based discovery in biomedicine: the architecture of the DAD-system. Proc AMIA Symp. 2000;:903-7.

Use of UMLS

- Mediation
  - Aymard S, Joubert M, Fieschi D, Fieschi M. Mediation services with health information sources. Proc AMIA Symp. 2000;:37-41.
- Decision Support
  - Geissbuhler A, Miller RA. Clinical application of the UMLS in a computerized order entry and decision-support system. Proc AMIA Symp. 1998;:320-4.

Reading

- http://www.nlm.nih.gov/research/umls/UMLSDOC.HTML
- http://umlsinfo.nlm.nih.gov/education.html