

An overview of the Unified Medical Language System (UMLS)

Overview

- History of the UMLS
- Structure of the UMLS
- Content of the UMLS
- The UTS tool for the UMLS
- The NAT tool for the UMLS
- Main Tables of the UMLS

History of the UMLS

- Based on:
- [http://www.nlm.nih.gov/research/umls/pdf/
mla_nlmtheater_2005.pdf](http://www.nlm.nih.gov/research/umls/pdf/mla_nlmtheater_2005.pdf)
- Started at National Library of Medicine, 1986
- A Division of the DHHS – Department of Health and Human Services (DHHS) of the Federal Government.
- Long term R&D project -

History - Purpose

- ... the UMLS project is an effort to overcome two significant barriers to effective retrieval of machine-readable information.
- The first is the variety of ways the same concepts are expressed in different machine-readable sources and by different people.
- The second is the distribution of useful information among many disparate databases and systems.

Structure of the UMLS

- 3 Parts:
- Metathesaurus – Over 2.6 million concepts and keeps growing and growing.
- Semantic Network. 133 Semantic Types (high level concepts) organized into two trees with IS-A links. See:
- <http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=nlmumls&part=ch05>

Structure of the UMLS (cont.)

- SPECIALIST Lexicon & Lexical Tools: lexical databases and programs
- (Less of a subject for this class)
- In addition there are the UTS (UMLS Terminology Services, a kind of an online browser; SEE: uts.nlm.nih.gov)
- MetamorphoSys: An installation tool

Detailed Statistics of the UMLS 2011AB

- http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/statistics.html
- Concepts: 2,612,024
 - Number of concept names (AUIs): 10,506,764
 - Number of distinct concept names (SUIs): 8,677,735
 - Number of distinct normalized concept names (LUIs): 7,734,809
- Sources: 161
- Languages: 21

Terminologies of the Metathesaurus

- See the source vocabulary page:
- http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/source_vocabularies.html
- Some vocabularies are much more important than others.

Selection of Terminologies

- I deleted everything with fewer than 10,000 terms.
- I deleted everything older than 2000.
- I deleted everything non-English.
- Here is what is left: 39 terminologies

- AOD AOD2000 Alcohol and Other Drug Thesaurus, 2000
 - 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), 2000;ENG
 - Number of Strings: 20685
- CHV CHV2010_10 Consumer Health Vocabulary
 - University of Utah;Consumer Health Vocabulary;Salt Lake City;2010;USA;ENG
 - Number of Strings: 148383
- CPT CPT2011 Current Procedural Terminology, 2011
 - Current Procedural Terminology (CPT);American Medical Association;2011;Chicago (IL)
 - Number of Strings: 28432
- CSP CSP2006 CRISP Thesaurus, 2006
 - Computer Retrieval of Information on Scientific Projects (CRISP). Bethesda (MD): National Institutes of Health, Division of Research Grants, Research Documentation Section, 2006.
 - Number of Strings: 22775
- FMA FMA3_1 Foundational Model of Anatomy Ontology, 3_1
 - Cornelius Rosse MD DSc, Jose L.V. Mejino Jr. MD
 - Number of Strings: 139095

- GO
(updated) GO2011_06_01 Gene Ontology
The Gene Ontology;The Gene Ontology Consortium;June 1, 2011
Number of Strings: 104241
- GS GS_2011_08_02 Gold Standard Alchemy, 2011_08_02
Gold Standard Alchemy;302 Knights Run Ave, Suite 800, Tampa, FL 33602;Gold Standard
Number of Strings: 25608
- HCPCS HCPCS2011 Healthcare Common Procedure Coding System, 2011
Version of Physicians' Current Procedural Terminology (CPT) included in the Healthcare Common
Procedure Coding System (HCPCS);Baltimore, MD;Centers for Medicare & Medicaid Services;2011
Number of Strings: 11452
- HCPT HCPT2011 HCPCS Version of Current Procedural Terminology (CPT), 2011
Version of Physicians' Current Procedural Terminology (CPT) included in the Healthcare Common
Procedure Coding System (HCPCS);Baltimore, MD;Centers for Medicare & Medicaid Services;2011
Number of Strings: 19643
- HUGO
(updated) HUGO2011_05 HUGO Gene Nomenclature, 2011_05
Eyre TA, Ducluzeau F, Sneddon TP, Povey S, Bruford EA and Lush MJ;HGNC Database;The HUGO
Gene Nomenclature Database;European Bioinformatics Institute Wellcome Trust Genome Campus;United
Kingdom
Number of Strings: 136029

- ICD10AM ICD10AM_2000 International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification, January 2000 Release
 - International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification; 2nd Edition, published January 2000. Developed and Maintained by the National Centre for Classification in Health, University of Sydney, Faculty of Health Sciences. PO Box 170 Lidcombe, NSW, Australia 1825. Phone: +61 2 9351 9461.
<http://www.cchs.usy.edu.au/ncch/>
 - Number of Strings: 25891
- ICD10CM ICD10CM_2011_01 International Classification of Diseases, 10th Edition, Clinical Modification, 2011_01
 - U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services; ICD-10-CM [computer file]: International Classification of Diseases, Tenth Revision, Clinical Modification; October 1, 2010; Hyattsville, MD; provided for public viewing only, and not valid for use prior to October 1, 2013.
 - Number of Strings: 102764
- ICD10PCS ICD10PCS_2011 ICD-10-PCS, 2011
 - Centers for Medicare and Medicaid Services; ICD-10-PCS; Baltimore, MD; 2011
 - Number of Strings: 249863

- ICD9CM
- (updated) ICD9CM_2012 International Classification of Diseases, Ninth Revision, Clinical Modification, 2012
- U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services; ICD-9-CM [computer file]: International Classification of Diseases, Ninth Revision, Clinical Modification; October 1, 2011; Baltimore, MD
- Number of Strings: 40844
- HIPAA standard
- CHI standard
-
- ICPC2ICD10ENG ICPC2ICD10ENG_200412 ICPC2 - ICD10 Thesaurus, 200412
- International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 2002
- Number of Strings: 81799

- ICPC2ICD10ENG ICPC2ICD10ENG_200412 ICPC2 - ICD10 Thesaurus, 200412
International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 2002
Number of Strings: 81799
-
-
- ICPC2P ICPC2P_2005 ICPC-2 PLUS
ICPC-2 PLUS. Sydney: Family Medicine Research Centre, University of Sydney, December 2005.
Number of Strings: 16897
-
- LNC
(updated) LNC236 LOINC, 236
The Regenstrief Institute;Logical Observation Identifier Names and Codes (LOINC);Version 2.36.;June, 2011;Indianapolis (IN)
Number of Strings: 364173
- Context: FULL-NOSIB-MULTIPLE
- CHI standard
-
- MDDB
(updated) MDDB_2011_08_03 Master Drug Data Base, 2011_08_03
Master Drug Data Base;2011
Number of Strings: 12740
-
- MDR
(updated) MDR14_0 Medical Dictionary for Regulatory Activities Terminology
(MedDRA), 14.0
MedDRA MSSO;MedDRA [electronic resource] : Medical Dictionary for Regulatory Activities Terminology;Version 14.0;March, 2011;Reston, VA
Number of Strings: 91686

- MEDCIN
(updated) MEDCIN3_2011_07_11 MEDCIN, 3_2011_07_11
Medicomp Systems;MEDCIN;[terminology files];Chantilly VA;Medicomp Systems;July 2011;Chantilly VA Number of Strings: 730363
- MMSL
(updated) MMSL_2011_08_01 Multum MediSource Lexicon, 2011_08_01
Medisource Lexicon;Release Date: August 01, 2011;Denver, CO;Multum Information Services, Inc.
Number of Strings: 68149
- MMX
(updated) MMX_2011_08_01 Micromedex RED BOOK, 2011_08_01
Micromedex RED BOOK;August 01, 2011
Number of Strings: 50005
- MSH
(updated) MSH2012_2011_09_09 Medical Subject Headings, 2012_2011_09_09
Medical Subject Headings (MeSH);National Library of Medicine;2011;Bethesda, MD
Number of Strings: 758306
- MTH MTH UMLS Metathesaurus
UMLS Metathesaurus. Bethesda, MD: National Library of Medicine.
Number of Strings: 146787
- MTHFDA
(updated) MTHFDA_2011_06_01 Metathesaurus FDA National Drug Code
Directory, 2011_06_01
Metathesaurus Forms of FDA National Drug Code Directory;2011_06_01;Bethesda, MD;National
Library of Medicine Number of Strings: 86438

- MTHICD9
 - (updated) MTHICD9_2012 International Classification of Diseases, Ninth Revision, Clinical Modification, Metathesaurus additional entry terms, 2012
 - Metathesaurus additional entry terms for ICD-9-CM [computer file]: International Classification of Diseases, Ninth Revision, Clinical Modification;National Library of Medicine;October 1, 2011;Bethesda, MD Number of Strings: 20205
- MTHSPL

 - (updated) MTHSPL_2011_08_26 Metathesaurus FDA Structured Product Labels, 2011_08_26
 - Metathesaurus Forms of the FDA Structured Product Labels;2011_08_26;Bethesda, MD;National Library of Medicine Number of Strings: 54732
- NCBI

 - (updated) NCBI2011_05_11 NCBI Taxonomy, 2011_05_11
 - U.S. Department of Health and Human Services;NCBI Taxonomy;National Institutes of Health, National Library of Medicine, National Center for Biotechnology Information;May 11, 2011;Bethesda, MD Number of Strings: 861598
- NCI

 - (updated) NCI2011_02D NCI Thesaurus, 2011_02D
 - National Cancer Institute, National Institutes of Health;NCI Thesaurus;National Cancer Institute;February 2011, Protege version;Rockville, MD Number of Strings: 238385
- HIPAA standard

- NDDF

 - (updated) NDDF_2011_08_03 National Drug Data File Plus Source Vocabulary, 2011_08_03
 - National Drug Data File Plus Source Vocabulary 2008;August 03, 2011;South San Francisco, CA;First DataBank Number of Strings: 88761

- NDFRT
(updated) NDFRT_2011_09_06 National Drug File, 2011_09_06
National Drug File - Reference Terminology;2011_09_06;Washington, DC;U.S. Department Of Veterans Affairs, Veterans Health Administration;June 2010
Number of Strings: 134828
- NIC NIC2005 Nursing Interventions Classification (NIC), 2005
Nursing Interventions Classification (NIC)
Number of Strings: 11255
- OMIM
(updated) OMIM2011_06_08 Online Mendelian Inheritance in Man,
2011_06_08
McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD);Online Mendelian Inheritance in Man, OMIM. ;McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University;ENG;<http://omim.org/>
Number of Strings: 141519
- PDQ PDQ2007_02 Physician Data Query, 2007_02
PDQ. Bethesda (MD): National Cancer Institute, February 2007.
Number of Strings: 26448
- RXNORM
(updated) RXNORM_11AA_110906F RxNorm Vocabulary,
11AA_110906F
RxNorm;META2011AA Full Update 2011_09_06;Bethesda, MD;National Library of Medicine
Number of Strings: 497893

- SNOMEDCT
(updated)
2011_07_31 SNOMEDCT_2011_07_31 SNOMED Clinical Terms,
International Health Terminology Standards Development Organization;SNOMED Clinical
Terms;July 31, 2011;Copenhagen,Denmark;International Health Terminology Standards Development
Organisation (IHTSDO)
- Number of Strings: 1181154
- CHI standard
- UMD UMD2011 UMDNS: product category thesaurus, 2011
The Universal Medical Device Nomenclature System;ECRI;Plymouth Meeting;2011
- Number of Strings: 25492
- UWDA UWDA173 University of Washington Digital Anatomist, 1.7.3
University of Washington Digital Anatomist, (UWDA). Seattle (WA): University of Washinton,
Version 1.7.3, March, 2003. Jose Mejino, M.D.; email: onard@biostr.washington.edu
- Number of Strings: 92913
- VANDF
(updated)
File, 2011_07_29 VANDF_2011_07_29 Veterans Health Administration National Drug
Veterans Health Administration National Drug File;July 29, 2011;Washington, DC;U.S. Department
of Veterans Affairs
- Number of Strings: 48910

of AUIs by Language

Language	Name Count (AUIs)	% of Metathesaurus
ENG	7651535	72.82%
SPA	1258098	11.97%
JPN	249845	2.38%
DUT	220625	2.1%
FRE	198051	1.88%
GER	185805	1.77%
POR	157675	1.5%
ITA	136654	1.3%
CZE	131918	1.26%

RUS	106860	1.02%
HUN	91125	0.87%
POL	42127	0.4%
SWE	26311	0.25%
FIN	25489	0.24%
KOR	11771	0.11%
SCR	8844	0.08%
LAV	1406	0.01%
DAN	723	0.01%
NOR	722	0.01%
BAQ	695	0.01%
HEB	485	0%

Languages used in the UMLS (2011AB)

- http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/statistics.html
- Statistics by terminologies

Source Counts by Language

Language	Name count	% of Metathesaurus
ENG	105	65.22%
DUT	6	3.73%
GER	6	3.73%
SPA	6	3.73%
FRE	5	3.11%
*	5	3.11%
ITA	4	2.48%
POR	4	2.48%
CZE	2	1.24%

FIN	2	1.24%
HUN	2	1.24%
JPN	2	1.24%
KOR	2	1.24%
SWE	2	1.24%
BAQ	1	0.62%
DAN	1	0.62%
HEB	1	0.62%
LAV	1	0.62%
NOR	1	0.62%
POL	1	0.62%
RUS	1	0.62%
SCR	1	0.62%

What was that *?

- *Note: 5 sources contribute relationships which have no associated language values

Structure of the Metathesaurus

- Concepts, Terms, Strings, Atoms
- Each one has its own system of ID numbers.
- CUI = Concept Unique Identifier
- LUI = Lexical Unique Identifier (Really term identifier, but the T was used for Types)
- SUI = String Unique Identifier
- AUI = Atom Unique Identifier

Atoms are Easy

- An atom is a pair consisting of a string and the terminology it is coming from.
- The same (identical) string might exist in the UMLS coming from several different terminologies.
- But an atom is unique in the whole UMLS.

Atom Unique Identifier

- Note **Headache** twice!
- The first time it is followed by (SNOMED).
- The second time it is followed by (MeSH).
- That means the string Headache exists twice in the UMLS, once coming from SNOMED, once from MeSH.
- Each occurrence has its own ID number, the AUI. A2882187 versus A0066000.

A1412439 headaches (BI)

S1459113 headaches

A2882187 Headache (SNOMED)

A0066000 Headache (MeSH)

S0046854 Headache

L0018681 headache

A1641293 Cranial Pain (MeSH)

S1680378 Cranial Pain

L1406212 cranial pain

A0418053

HEAD PAIN CEPHALGIA (DxP)

S0375902

HEAD PAIN CEPHALGIA

L0290366 cephalgia head pain

C0018681 Headache

String Unique Identifier

- Now compare **headaches** and **Headache**
- A HUMAN can tell that this is the same.
- A computer can tell that this is the same by using simple algorithms: Capitalize all letters. Drop plural “s” at the end.
- The UMLS assign different string IDs to strings that look different to a computer.
- S1459114 versus S0046854

A1412439 headaches (BI)

S1459113 headaches

A2882187 Headache (SNOMED)

A0066000 Headache (MeSH)

S0046854 Headache

L0018681 headache

A1641293 Cranial Pain (MeSH)

S1680378 Cranial Pain

L1406212 cranial pain

A0418053

HEAD PAIN CEPHALGIA (DxP)

S0375902

HEAD PAIN CEPHALGIA

L0290366 cephalgia head pain

C0018681 Headache

Lexical Unique Identifier

- Now compare the two terms (= lexical units) **headache** and **cranial pain**.
- There is no algorithmic way a computer would know this is the same thing.
- Even a human would not know this is the same thing, unless he went to medical school.
- The UMLS assigns a different LUI to every term. L0018681 versus L1406212.

A1412439 headaches (BI)

S1459113 headaches

A2882187 Headache (SNOMED)

A0066000 Headache (MeSH)

S0046854 Headache

L0018681 headache

A1641293 Cranial Pain (MeSH)

S1680378 Cranial Pain

L1406212 cranial pain

A0418053

HEAD PAIN CEPHALGIA (DxP)

S0375902

HEAD PAIN CEPHALGIA

L0290366 cephalgia head pain

C0018681 Headache

Concept Unique Identifier

- Cranial Pain
- Headache
- Cephalgia Head Pain
- all have the same CUI: C0018681

A1412439 headaches (BI)

S1459113 headaches

A2882187 Headache (SNOMED)

A0066000 Headache (MeSH)

S0046854 Headache

L0018681 headache

A1641293 Cranial Pain (MeSH)

S1680378 Cranial Pain

L1406212 cranial pain

A0418053

HEAD PAIN CEPHALGIA (DxP)

S0375902

HEAD PAIN CEPHALGIA

L0290366 cephalgia head pain

C0018681 Headache

So how does this help the Computer?

- By assigning the same CUI to headache and cranial pain the computer KNOWS that these two words mean the same thing.
- By assigning the same LUI to headaches and Headache the computer knows that these two words are basically the same, except for differences in English syntax (singular vs. plural.)

- By assigning the same SUI but different AUIs to headache (SNOMED) and headache (MeSH) the computer knows that the term headache comes from two different source terminologies into the UMLS.
- Do we absolutely need the SUI? I am not sure. But the basic approach is, everything gets its own ID number. We saw this for the SNOMED.

A1412439 headaches (BI)

S1459113 headaches

A2882187 Headache (SNOMED)

A0066000 Headache (MeSH)

S0046854 Headache

L0018681 headache

A1641293 Cranial Pain (MeSH)

S1680378 Cranial Pain

L1406212 cranial pain

A0418053

HEAD PAIN CEPHALGIA (DxP)

S0375902

HEAD PAIN CEPHALGIA

L0290366 cephalgia head pain

C0018681 Headache

Let's look at the figure one more time

- Everything (more or less) is color-coded.
- Note the green box (LUI) around headache.
- Inside of this box are two yellow boxes (SUI).
- This says that these two SUIs belong to this LUI.
- Similarly note the red box containing the three green boxes.
- Three LUIs belong to one CUI.

Test Question

- Where is the color coding wrong?
- The CUI and the word Headache should be in RED, just as the box they belong to.
- The terminologies (SNOMED, MeSH etc. should be in BLUE).

Another Example

- This has a slightly different color coding.
- It is also missing the Atoms.

Concept
C0001621

Term L0001621	<p>S0011232 Adrenal Gland Diseases S0011231 Adrenal Gland Disease S0000441 Disease of adrenal gland S0481705 Disease of adrenal gland, NOS S0220090 Disease, adrenal gland S0044801 Gland Disease, Adrenal</p> <p>[...]</p>
Term L0041793	<p>S0860744 Disorder of adrenal gland, unspecified S0217833 Unspecified disorder of adrenal glands</p>
Term L0161347	<p>S0225481 ADRENAL DISORDER S0627685 DISORDER ADRENAL (NOS)</p> <p>[...]</p>
Term L0181041	<p>S0632950 Disorder of adrenal gland S0354509 Adrenal Gland Disorders</p> <p>[...]</p>
Term L0368399	<p>S0586222 Adrenal disease S0466921 ADRENAL DISEASE, NOS</p> <p>[...]</p>
Term L1279026	<p>S1520972 Nebennierenkrankheiten</p> <p>GER</p>
Term L0162317	<p>S0226798 SURRENALE, MALADIES</p> <p>FRE</p> <p>[...]</p>



Mapping Rules

- The next page contains mapping rules between LUIs, CUIs, SUIs, AUIs.
- You don't have to know those.

- 1 CUI -> 1 or more AUIs
- 1 CUI -> 1 or more SUIs
- 1 CUI -> 1 or more LUIs
- 1 AUI -> 1 SUI (exactly one!)
- 1 AUI -> 1 LUI (exactly one!)
- 1 AUI -> 1 CUI (exactly one!)
- 1 SUI -> 1 or (typically) many AUIs
- 1 SUI -> 1 LUI
- 1 SUI -> 1 or more CUIs (but typically 1 CUI)
- 1 LUI -> 1 or (typically) many AUIs
- 1 LUI -> 1 or (typically) many SUIs
- 1 LUI -> 1 or more CUIs (but typically 1 CUI)

The UTS tool for the UMLS

- You need a license.
- Please apply for one.
- <https://uts.nlm.nih.gov>

The NAT tool for the UMLS

- <http://nat.njit.edu/>

Main Tables of the UMLS

- Just like the SNOMED, the UMLS is released as relational tables.
- MRCONSO (Concept information)
- MRSTY (Semantic types – next lecture)
- MRREL (Relationships between concepts)
- MRHIER (Relationship paths)
- MRDOC (Documentation)

Release Formats

- We only look at the RRF (Rich Release Format)
- Theoretically the UMLS promises “source transparency” which means that every source can be pulled out of the UMLS just as it looked when it was put into the UMLS.
- In practice there are problems with source transparency.

Columns

- Information about columns can be found here:
- http://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/release/columns_data_elements.html
- Important columns include AUI, LUI, SUI, CUI, of course. Also if a table contains two CUIs then they are coded CUI1, CUI2.

More Important Columns

- LAT ... Language of Term (ENG, SPA,...)
- STR ... String
- REL ... Relationship Label (PAR, CHD, ...)
- RELA ... Relationship Attribute. Some terminologies have additional information about what kind of relationship this is.

Other Important Columns

- SAB ... Source Abbreviation, e.g.
SNOMEDCT_2011_01_31,
MEDCIN3_2010_12_14
- RSAB... Root Source Abbreviation, e.g.,
- SNOMEDCT,
MEDCIN
- VSAB ... A superset of SABs. Contains older
versions also.

Other Important Columns

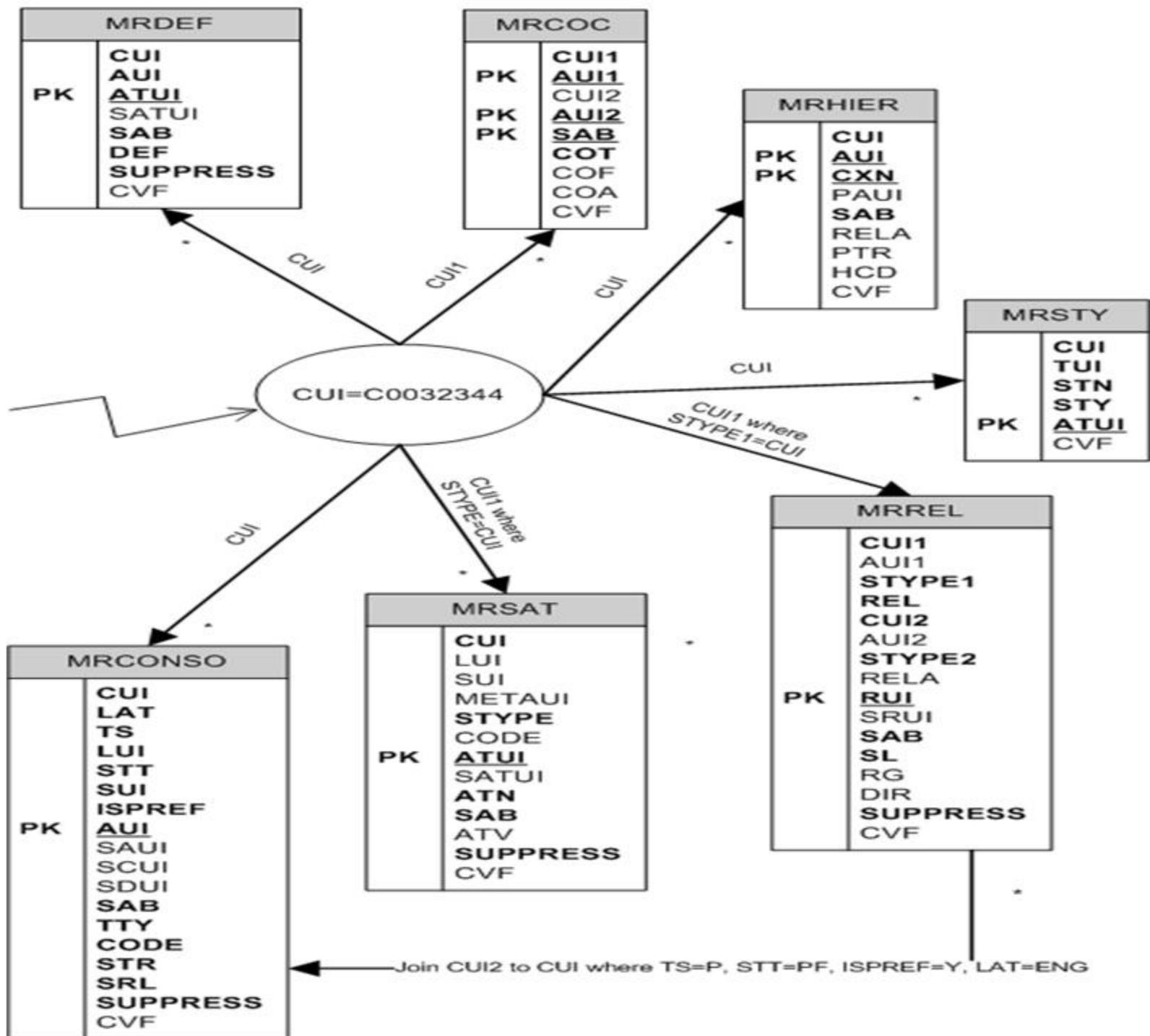
- STY ... Semantic Type (next lecture)
- TUI ... Tree number of Semantic Type
(essentially a path of semantic types)

A Good Place to Review Basics

- http://www.nlm.nih.gov/research/umls/new_users/online_learning/index.htm

Query Diagrams for the UMLS

- http://www.nlm.nih.gov/research/umls/implementation_resources/query_diagrams/index.html
- There is a LOT of stuff there.
- We'll look at one example, with no details.



Example Queries

- 1. Find all atoms of a UMLS concept.
- SELECT * FROM mrconso
WHERE cui = 'C0032344';
- 3. Find all source definitions associated with a UMLS concept.
- SELECT * FROM mrdef
WHERE cui = 'C0032344';

Example Queries - 2

- 7.a. Find all relationships for a UMLS concept.
- SELECT * FROM mrrel
WHERE cui1 = 'C0032344'
- 7.b. Find all inverse relationships for a UMLS concept.
- SELECT * FROM mrrel
WHERE cui2 = 'C0032344';
AND stype2 = 'CUI';

What is STYPE?

- STYPE: Relationships can be defined between concepts, but they don't have to be.
- STYPE1 and STYPE2 tells you for a specific relationship instance (a line in the table MRREL) between what it is defined.
- E.g. it can be defined between concepts or atoms.
- Possible values: SDUI, CODE, CUI, SCUI, AUI

What are those?

- SDUI: Source asserted descriptor identifier.
Basically means an ID number directly from a source terminology for a term.
- CUI: We did that one. Concepts
- SCUI: Source asserted concept identifier.
Basically means an ID number directly from a source terminology for a concept.
- AUI: We did that one. Atoms

What are those? - 2

- CODE:CODE is a synonym for one of these identifiers that is the “Most useful source asserted identifier (if the source vocabulary has more than one identifier), or a Metathesaurus-generated source entry identifier (if the source vocabulary has none).”