Tutorial T13

AMIA Fall Symposium
Saturday, November 9, 2002

Customizing the UMLS Metathesaurus for Your Applications

Olivier Bodenreider, MD, PhD
William T. Hole, MD
Betsy L. Humphreys, MLS
Laura Roth, MLS
Suresh Srinivasan, MS

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content — O. B., L. Roth, S. Srinivasan
    - Customize with MetamorphoSys
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

UMLS Knowledge Sources

Multi-purpose tools or “intellectual middleware” for System Developers

- Metathesaurus
- Semantic Network
- SPECIALIST lexicon and lexical programs
  - T25 – Lexical Tools for UMLS Developers – Sunday, November 10, 8:30-noon.

UMLS Metathesaurus

- Concepts, terms, and attributes from many controlled “vocabularies”
  - in a common explicit database format
- New inter-source relationships, definitional information, use information
- Scope determined by combined scope of source vocabularies

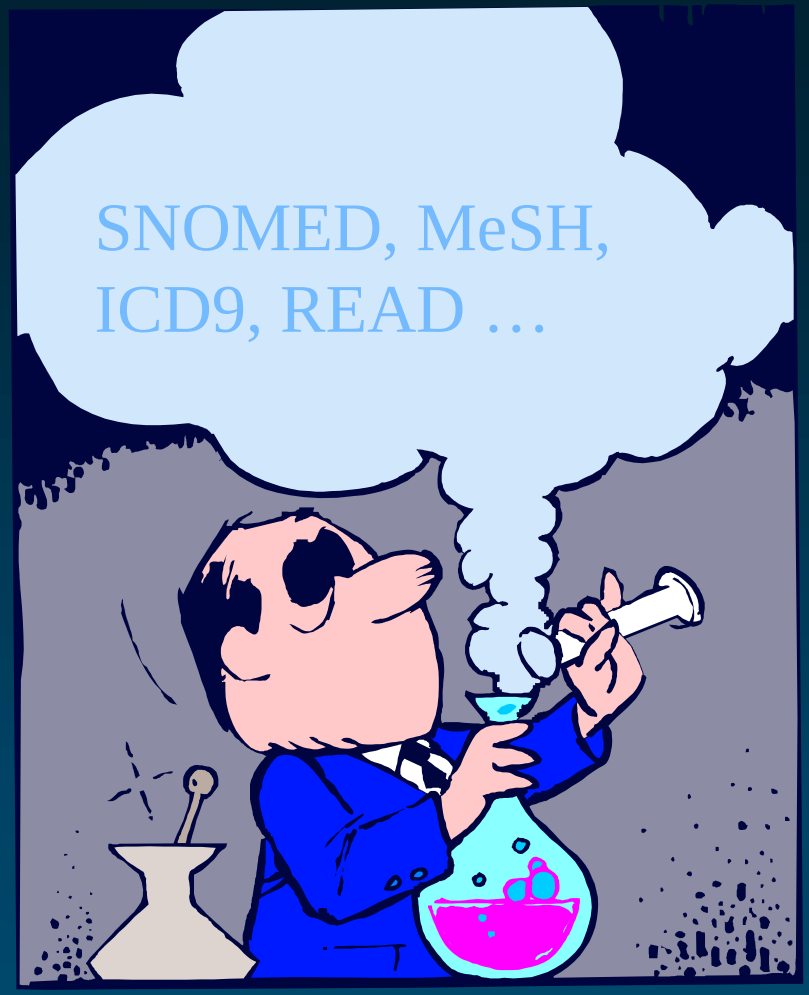
UMLS Source "Vocabularies"

- Widely varying purposes, structures, properties, but all are in essence "sets of valid values" for data elements:
  - Thesauri, e.g., MeSH
  - Statistical Classifications, e.g., ICD
  - Billing Codes, e.g., CPT
  - Clinical coding systems, e.g., SNOMED, Read , RxNorm
  - Lists of controlled terms, e.g., COSTAR, HL7 values
- All HIPAA code sets, except NDC

2002AC UMLS Metathesaurus

- ~870,000 concepts
- ~1,756,000 “terms” (Eye, Eyes, eye = 1)
- ~2,083,103 “strings”/concept names (Eye, Eyes, eye = 3)
- ~11,479,000 relationships between concepts
- >113 source vocabularies (including several “families” with multiple members)
- 15 different languages

How to combine them?



Meta Processor,
Alpha 0.001

Not really ….

- "The Metathesaurus preserves the meanings, hierarchical connections, and other relationships between terms present in its source vocabularies, while adding certain basic information about each of its concepts and establishing new relationships between concepts and terms from different source vocabularies."

Why Customize? 4 basic reasons

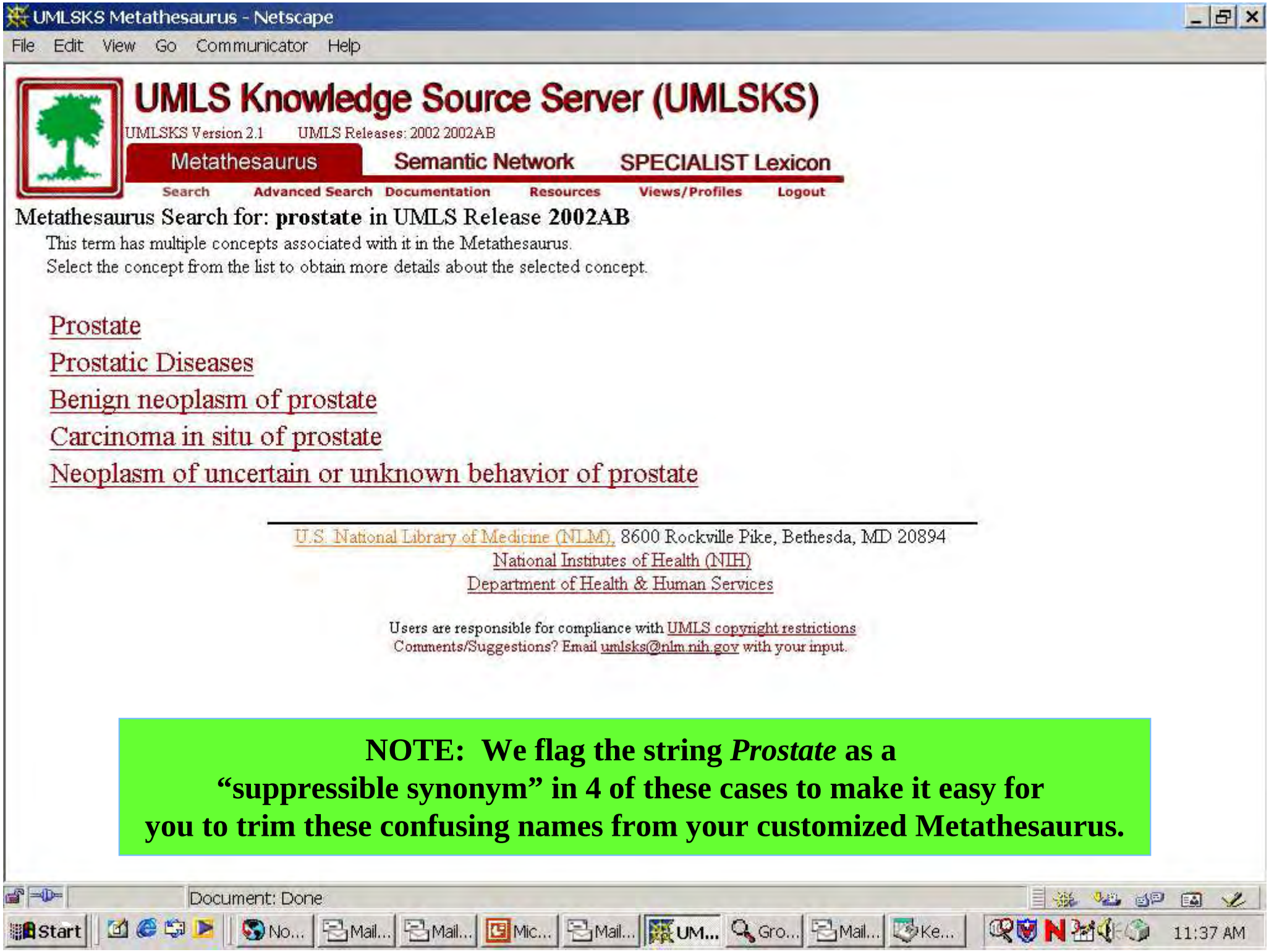
- Nobody needs or wants all of it for any specific set of purposes
  - extraneous vs. pernicious languages, concepts, strings, relationships, attributes
- You don’t have the licenses required for operational use of all source vocabularies
- The default “preferred name” is not best for your applications
- You need to add important local terminology

Possibly Extraneous, e.g.,

- Terms in languages other than English
- Redundant minor variations
- Procedure codes, when your application is focused on problems
- Vocabulary “housekeeping” attributes

Possibly Pernicious, e.g.,

- Terms that lack face validity
- Abbreviations and short forms
- Other less than beautiful “suppressible synonyms” already identified by NLM
- Relationships that reflect an alien or unhelpful “world view”


UMLSKS Metathesaurus - Netscape
File Edit View Go Communicator Help
UMLS Knowledge Source Server (UMLSKS)
UMLSKS Version 2.1 UMLS Releases: 2002 2002AB
Metathesaurus Semantic Network SPECIALIST Lexicon
Search Advanced Search Documentation Resources Views/Profiles Logout
Metathesaurus Search for: prostate in UMLS Release 2002AB
This term has multiple concepts associated with it in the Metathesaurus.
Select the concept from the list to obtain more details about the selected concept.
Prostate
Prostatic Diseases
Benign neoplasm of prostate
Carcinoma in situ of prostate
Neoplasm of uncertain or unknown behavior of prostate
U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Department of Health & Human Services
Users are responsible for compliance with UMLS copyright restrictions
Comments/Suggestions? Email umlsks@nlm.nih.gov with your input.
NOTE: We flag the string Prostate as a
"suppressible synonym" in 4 of these cases to make it easy for
you to trim these confusing names from your customized Metathesaurus.
Document: Done
Start
11:37 AM

UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 2.1 UMLS Releases: 2002 2002AB

Metathesaurus | Semantic Network | SPECIALIST Lexicon

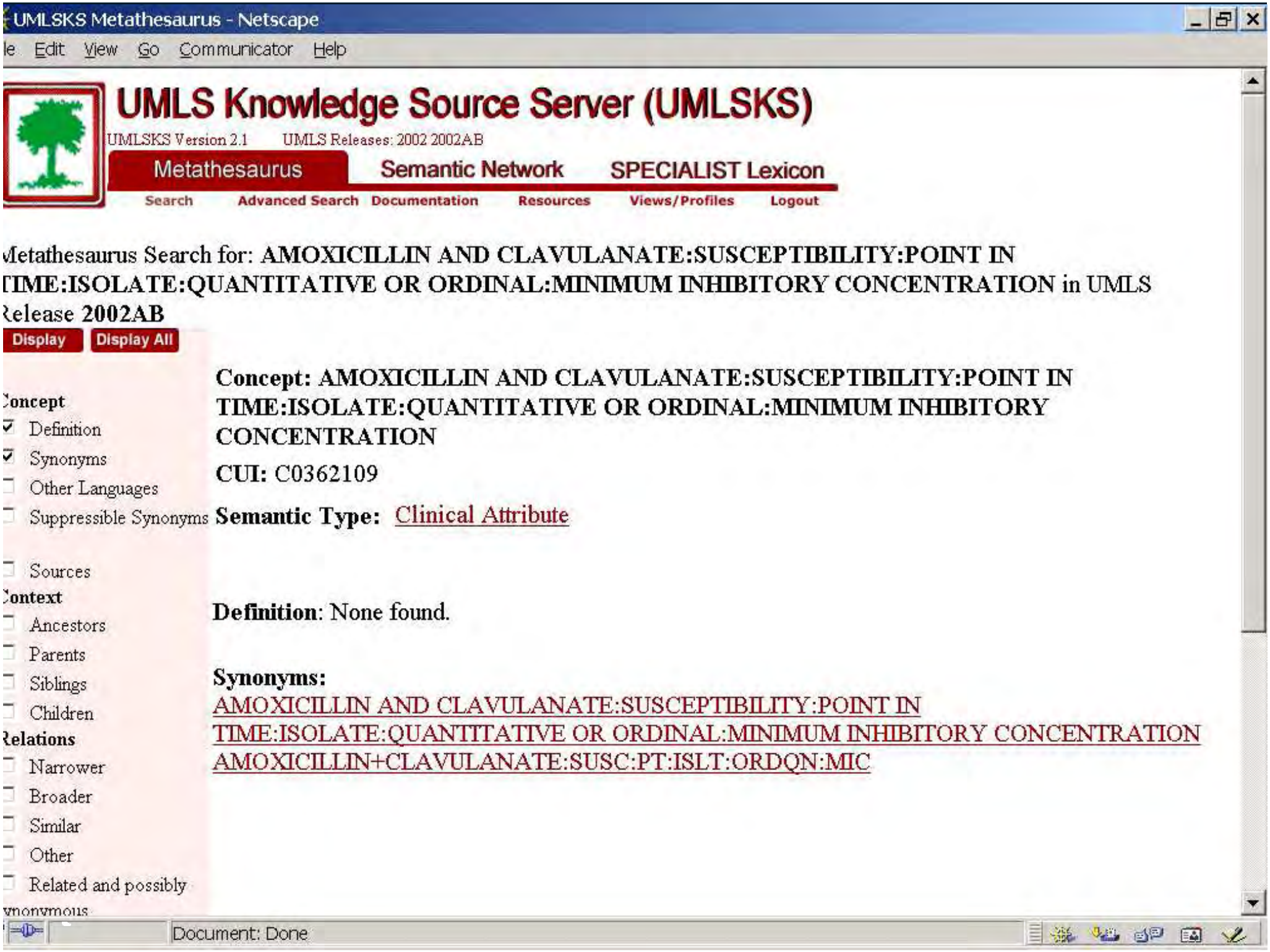
Search | Advanced Search | Documentation | Resources | Views/Profiles | Logout

Metathesaurus Search for: **ER** in UMLS Release **2002AB**

This term has multiple concepts associated with it in the Metathesaurus.
Select the concept from the list to obtain more details about the selected concept.

- Endoplasmic Reticulum
- Estrogen Receptors

U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Department of Health & Human Services


UMLSKS Metathesaurus - Netscape
le Edit View Go Communicator Help
UMLS Knowledge Source Server (UMLSKS)
UMLSKS Version 2.1 UMLS Releases: 2002 2002AB
Metathesaurus
Semantic Network
SPECIALIST Lexicon
Search
Advanced Search
Documentation
Resources
Views/Profiles
Logout
Metathesaurus Search for: AMOXICILLIN AND CLAVULANATE:SUSCEPTIBILITY:POINT IN TIME:ISOLATE:QUANTITATIVE OR ORDINAL:MINIMUM INHIBITORY CONCENTRATION in UMLS Release 2002AB
Display
Display All
Concept
Definition
Synonyms
Other Languages
Suppressible Synonyms
Sources
Context
Ancestors
Parents
Siblings
Children
Relations
Narrower
Broader
Similar
Other
Related and possibly
ynonymous
Concept: AMOXICILLIN AND CLAVULANATE:SUSCEPTIBILITY:POINT IN TIME:ISOLATE:QUANTITATIVE OR ORDINAL:MINIMUM INHIBITORY CONCENTRATION
CUI: C0362109
Semantic Type: Clinical Attribute
Definition: None found.
Synonyms:
AMOXICILLIN AND CLAVULANATE:SUSCEPTIBILITY:POINT IN TIME:ISOLATE:QUANTITATIVE OR ORDINAL:MINIMUM INHIBITORY CONCENTRATION
AMOXICILLIN+CLAVULANATE:SUSC:PT:ISLT:ORDQN:MIC
Document: Done

UMLSKS Metathesaurus - Netscape
File Edit View Go Communicator Help
Associated Expressions
Co-occurring Concepts
Co-occurring MeSH
Co-occurring AI/RHEUM
Ancestors:
MeSH
Biological Sciences (MeSH Category) [G]
Circulatory and Respiratory Physiology [G9]
Respiratory Physiology [G9.772]
Respiration [G9.772.521]
Alcohol and Other Drug Thesaurus
concepts in biomedical areas [E]
body system or organ function [EH]
respiratory system function [EM]
breathing [EM2]
Home Health Care Classification
PHYSICAL REGULATION COMPONENT [K]
Vital Signs [K33]
Respiration [K33.4]
Omaha System
DOMAIN III. PHYSIOLOGICAL [P3]
Respiration [P328]
Psychological Index Terms
Respiration []
SNOMED 1982
Function Axis []
Function and Abnormal Function of the Cardiovascular and Respiratory Systems []
Functions and Abnormal Functions of the Respiratory System []
Ventilatory Functions []
Ventilation [F-76500]
Document: Done
Start
11:59 AM

License restriction levels

- Level 0 – 61.5% of concepts
  - Basic license requirements, e.g., copyright statement and credits to NLM and producers of the vocabularies you use, no redistribution except as a part of your application
- Level 1 – 4.3% of concepts
  - Basic, plus you must negotiate with producer to translate into another language

READ the license, including the appendix

License restriction levels

- Level 2 - .0009% of concepts
  - Basic, plus you must negotiate with producer for use in the creation of health data
- Level 3 – 33.9% of concepts
  - Basic, plus you must negotiate with the producer for *any* production use. Explicit prohibition against providing access via the Internet.
- There may - or may not - be license fees associated with uses not covered by the UMLS license.

READ the license, including the appendix

Customization is critical,

but it ***requires*** a clear understanding of:

- Your functional requirements
- Characteristics of relevant UMLS source vocabularies
  - Explore these via the UMLS Knowledge Source Server
- Your license arrangements
- -- ***and*** some technical expertise
- Therefore, it is usually a team sport.

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content — O. B., L. Roth, S. Srinivasan
    - Customize with MetamorphoSys
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

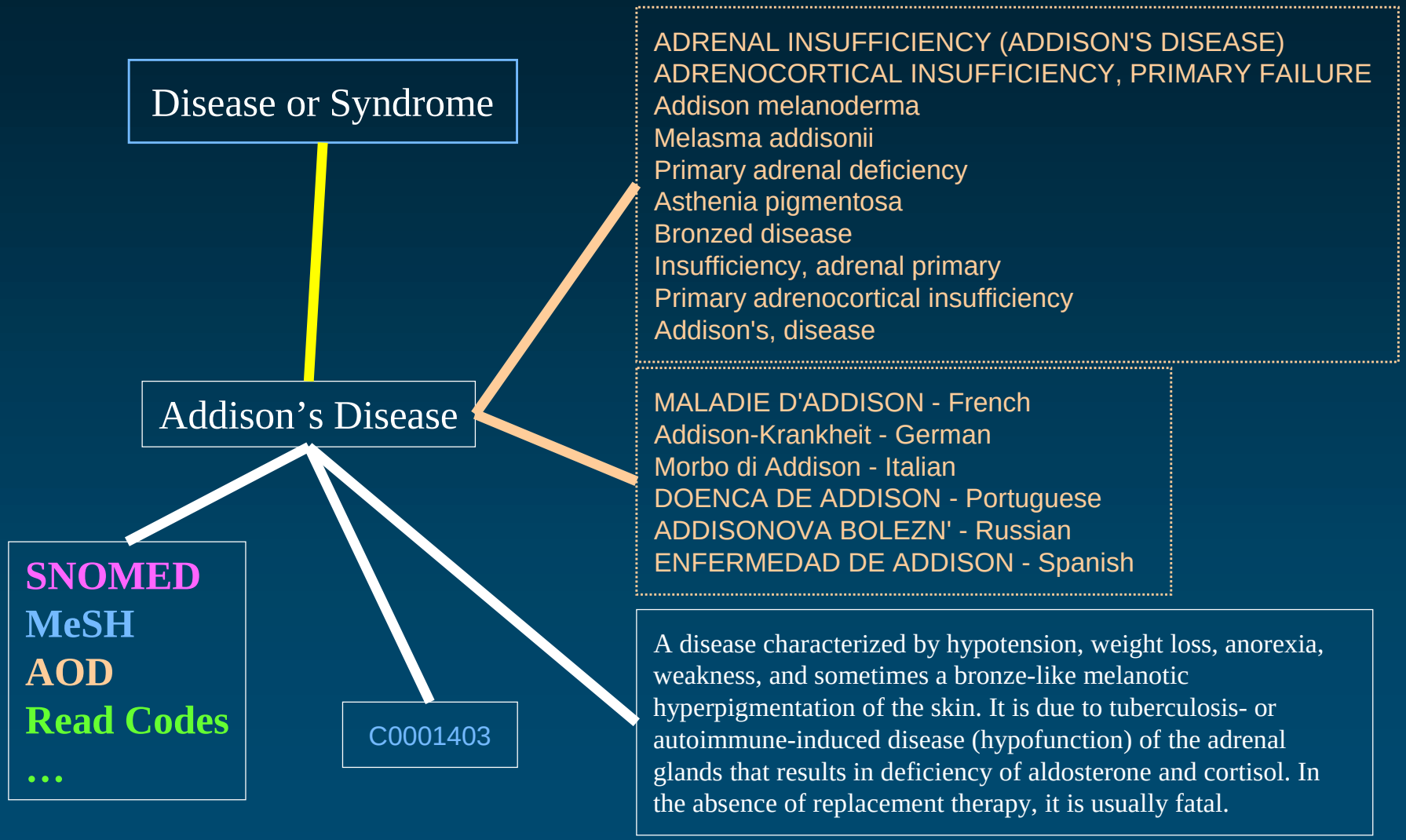
Access to UMLS data

- Local database
- Data model
  - Relational model + SQL
  - Object-oriented model + some O-O language

Metathesaurus Basic organization

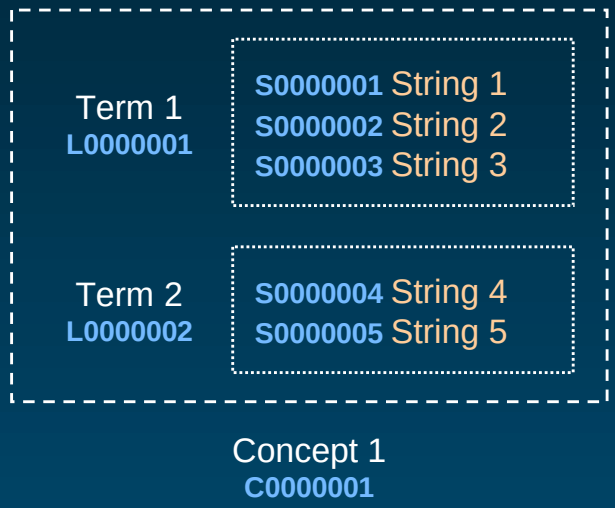
- Terms / Concepts
  - Synonymous terms are clustered into a concept
  - Properties are attached to concepts, e.g.,
    - Unique identifier
    - Definition
- Relationships
  - Concepts are related to other concepts
  - Properties are attached to relationships, e.g.,
    - Type of relationship
    - Source

Addison's Disease: Concept



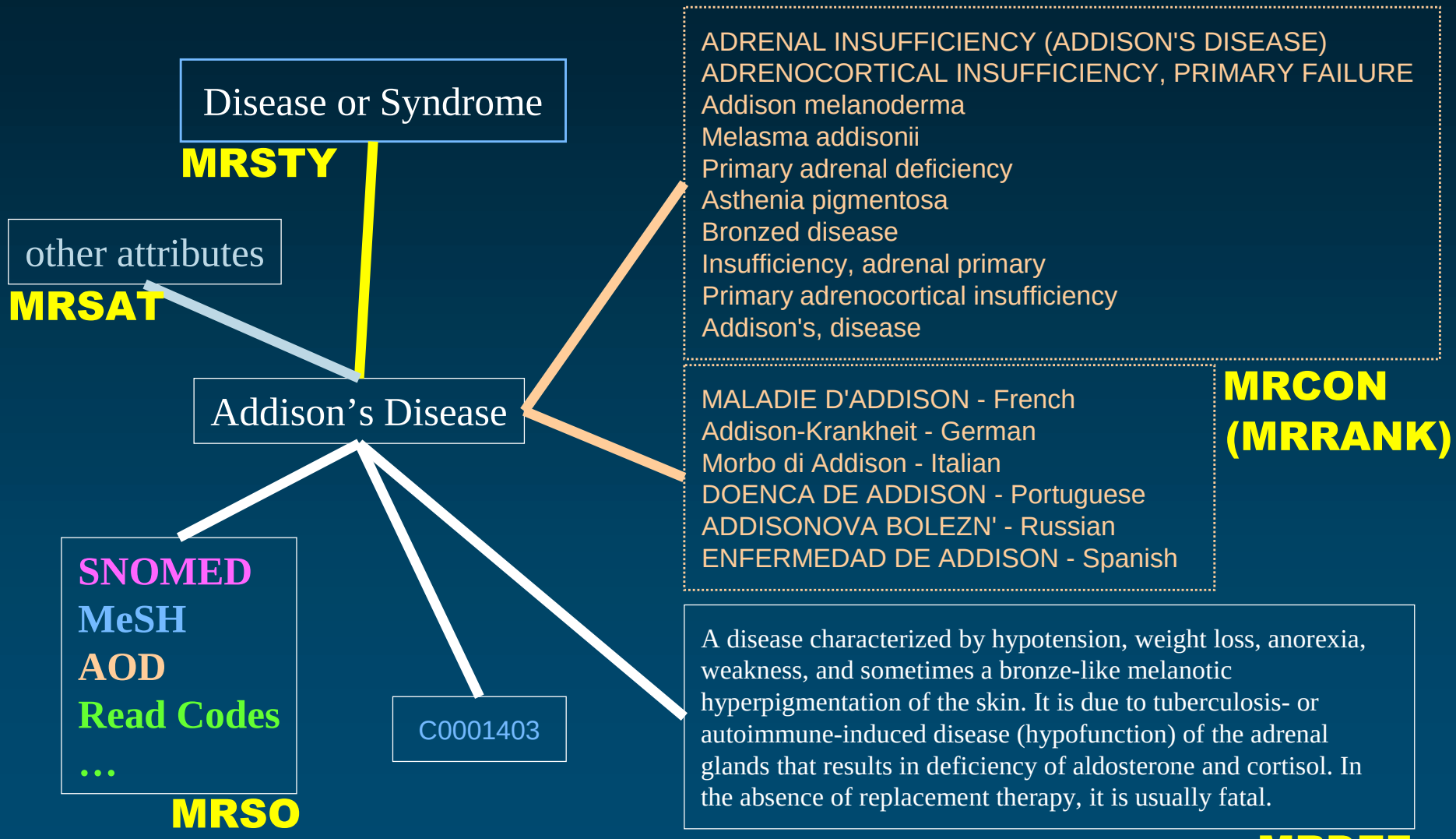
Metathesaurus Concepts

- Concept: Cluster of synonymous terms
  - ~870,000 concepts
  - identified by a CUI
- Term: Set of lexical variants
  - ~1.7 M terms
  - identified by a LUI
- String: Concept name
  - ~2 M strings
  - identified by a SUI

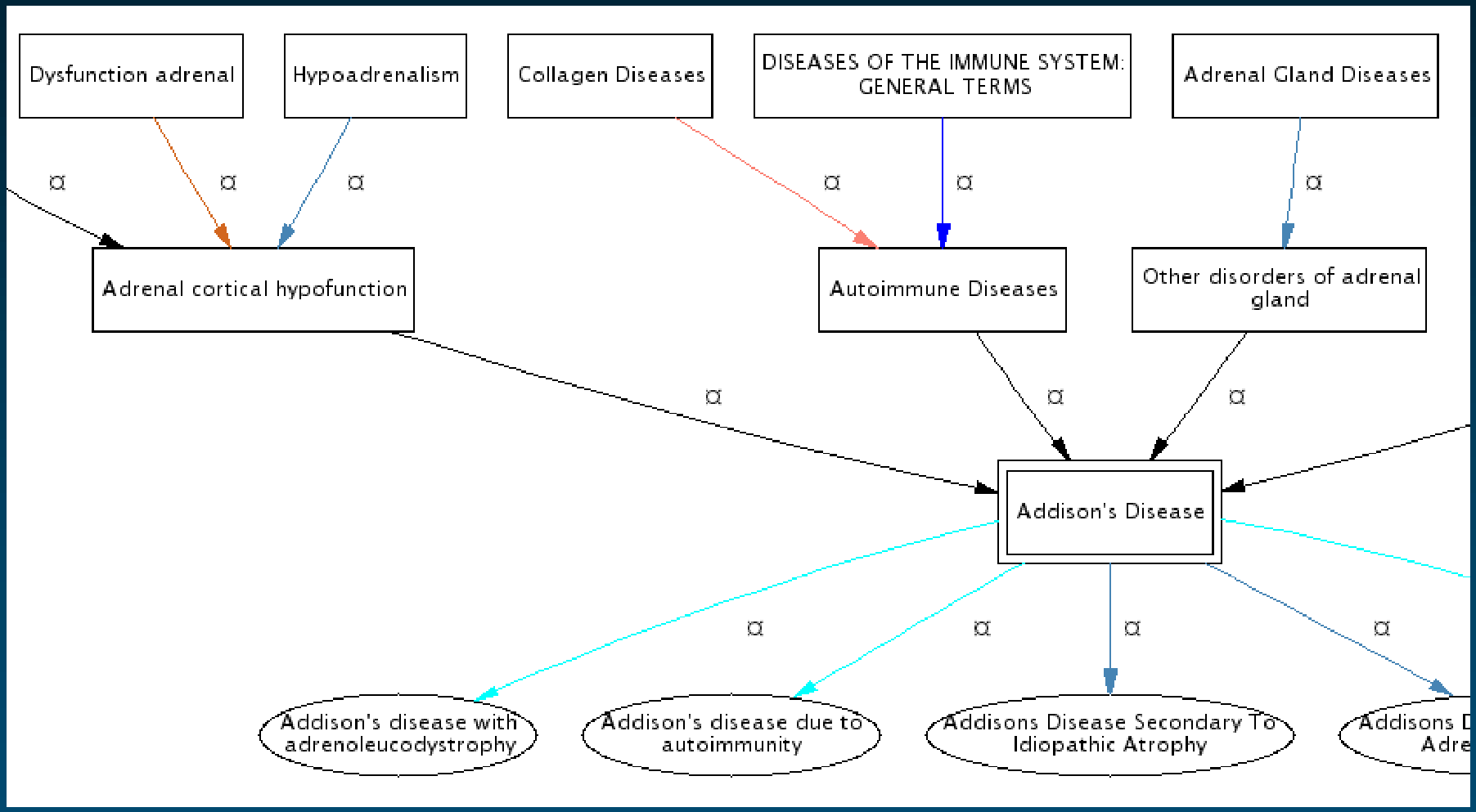


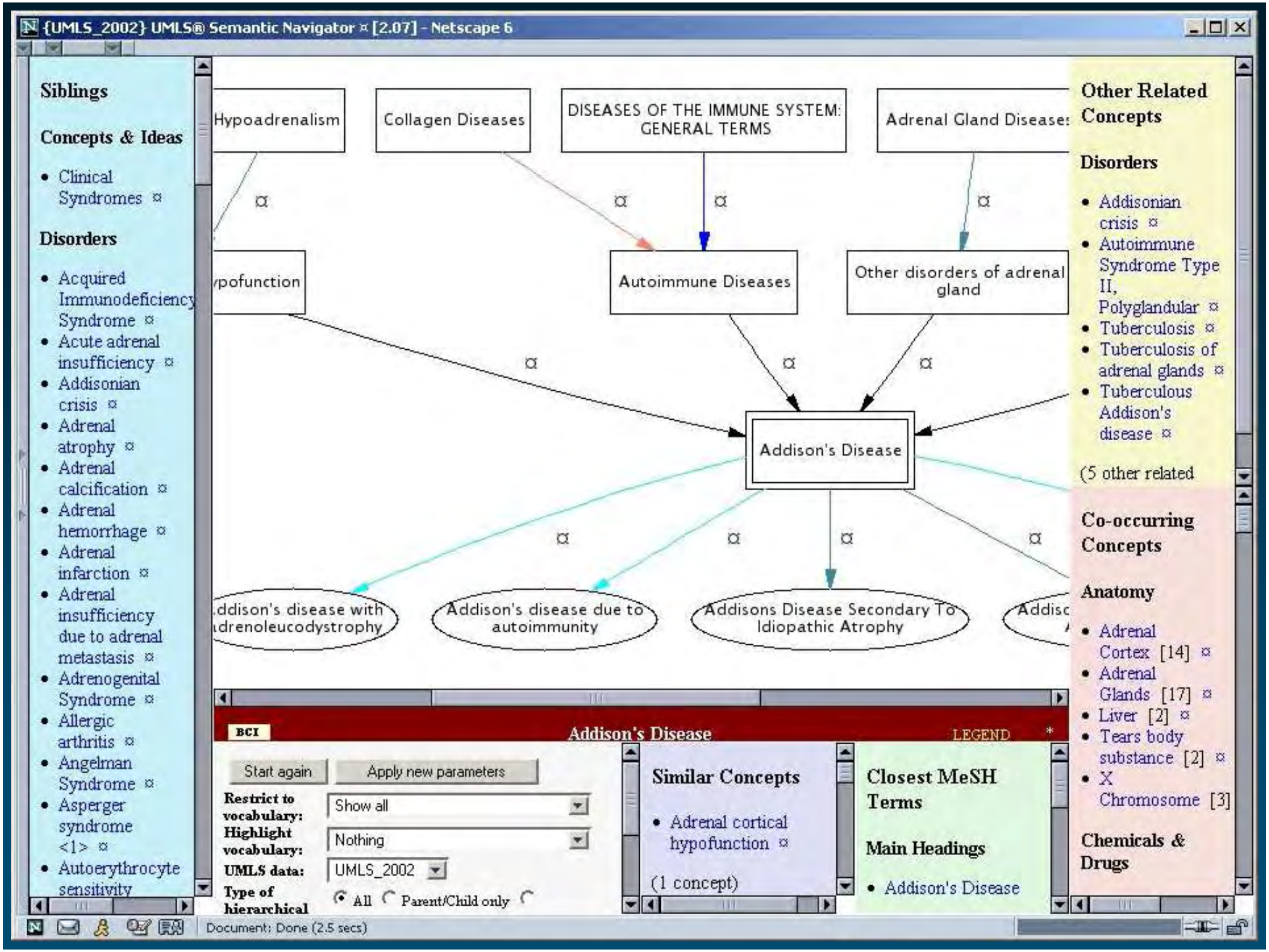
Cluster of synonymous terms

Metathesaurus files Concepts



Addison's disease Relationships

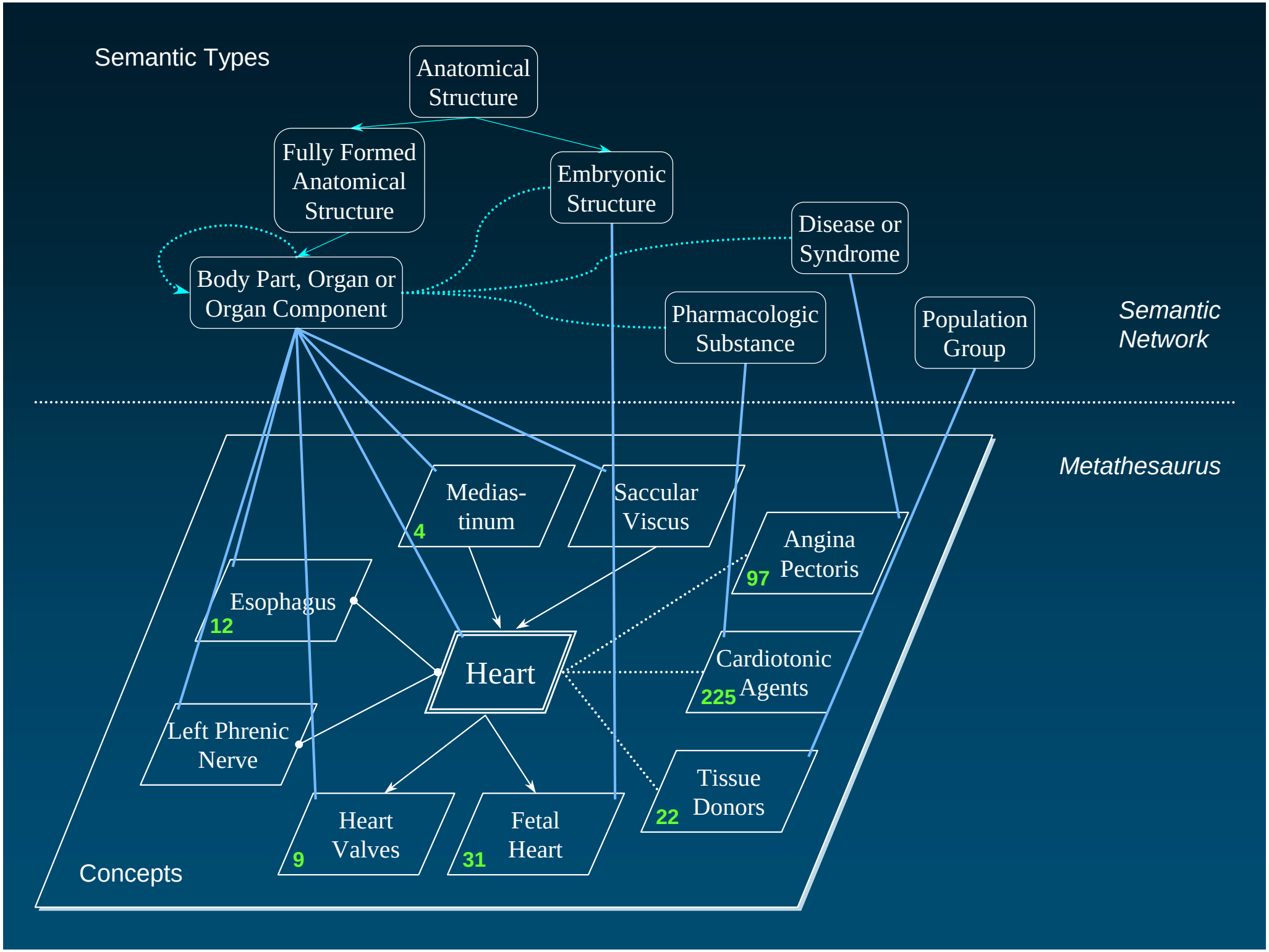



{UMLS_2002} UMLS® Semantic Navigator ¤ [2.07] - Netscape 6
Siblings
Concepts & Ideas
Clinical Syndromes ¤
Disorders
Acquired Immunodeficiency Syndrome ¤
Acute adrenal insufficiency ¤
Addisonian crisis ¤
Adrenal atrophy ¤
Adrenal calcification ¤
Adrenal hemorrhage ¤
Adrenal infarction ¤
Adrenal insufficiency due to adrenal metastasis ¤
Adrenogenital Syndrome ¤
Allergic arthritis ¤
Angelman Syndrome ¤
Asperger syndrome <1> ¤
Autoerythrocyte sensitivity
Hypoadrenalism
Collagen Diseases
DISEASES OF THE IMMUNE SYSTEM: GENERAL TERMS
Adrenal Gland Diseases
/pofunction
Autoimmune Diseases
Other disorders of adrenal gland
Addison's Disease
ddison's disease with drenoleucodystrophy
Addison's disease due to autoimmunity
Addisons Disease Secondary To Idiopathic Atrophy
Other Related Concepts
Disorders
Addisonian crisis ¤
Autoimmune Syndrome Type II, Polyglandular ¤
Tuberculosis ¤
Tuberculosis of adrenal glands ¤
Tuberculous Addison's disease ¤
(5 other related
Co-occurring Concepts
Anatomy
Adrenal Cortex [14] ¤
Adrenal Glands [17] ¤
Liver [2] ¤
Tears body substance [2] ¤
X Chromosome [3]
Chemicals & Drugs
BCI
Addison's Disease
LEGEND
Start again
Apply new parameters
Restrict to vocabulary:
Show all
Highlight vocabulary:
Nothing
UMLS data:
UMLS_2002
Type of hierarchical
All
Parent/Child only
Similar Concepts
Adrenal cortical hypofunction ¤
(1 concept)
Closest MeSH Terms
Main Headings
Addison's Disease
Document: Done (2.5 secs)

Metathesaurus Relationships

- Asserted relationships: ~5 M pairs of concepts
- Statistical relationships : ~6.5 M pairs of concepts (co-occurring concepts)

- Categorization: Relationships to semantic types from the Semantic Network


Semantic Types
Anatomical Structure
Fully Formed Anatomical Structure
Embryonic Structure
Disease or Syndrome
Body Part, Organ or Organ Component
Pharmacologic Substance
Population Group
Semantic Network
Metathesaurus
Medias-tinum
4
Saccular Viscus
Angina Pectoris
97
Esophagus
12
Heart
Cardiotonic Agents
225
Left Phrenic Nerve
Tissue Donors
22
Heart Valves
9
Fetal Heart
31
Concepts

Metathesaurus files Relationships

- Asserted relationships **MRREL**
- Statistical relationships **MRCOC**

- Categorization **MRSTY**

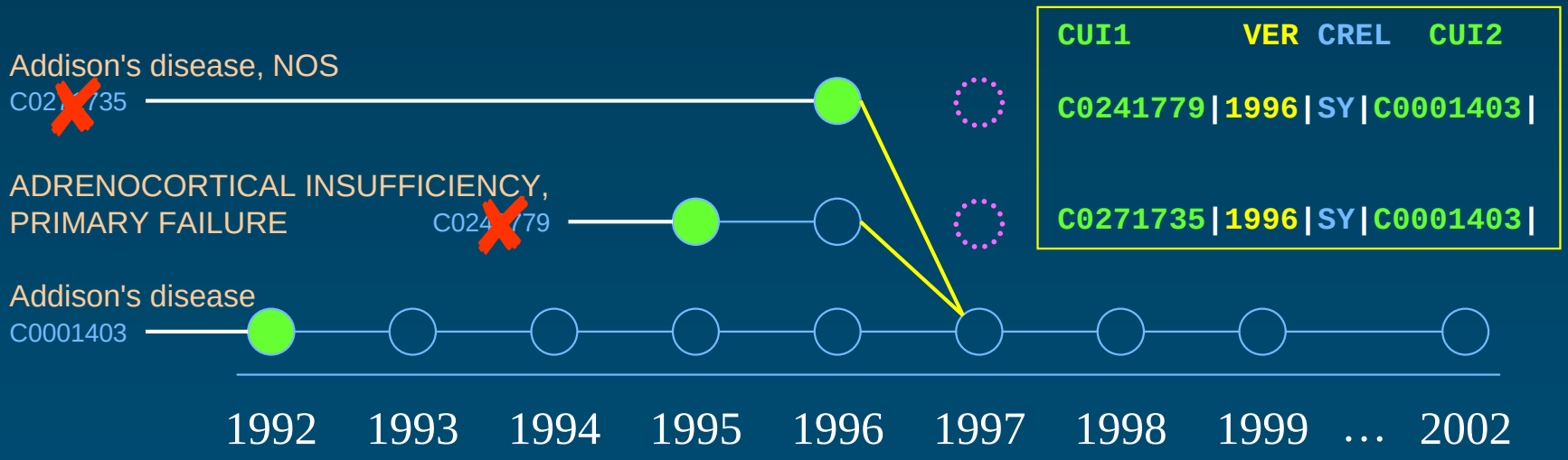
MRCXT is *not* the authoritative source of relationships

Metathesaurus Evolution over time

- Concepts never die (in principle)
  - CUIs are permanent identifiers
- What happens when they do die (in reality)?
  - Concepts can merge or split
  - Resulting in new concepts and deletions

MRCUI

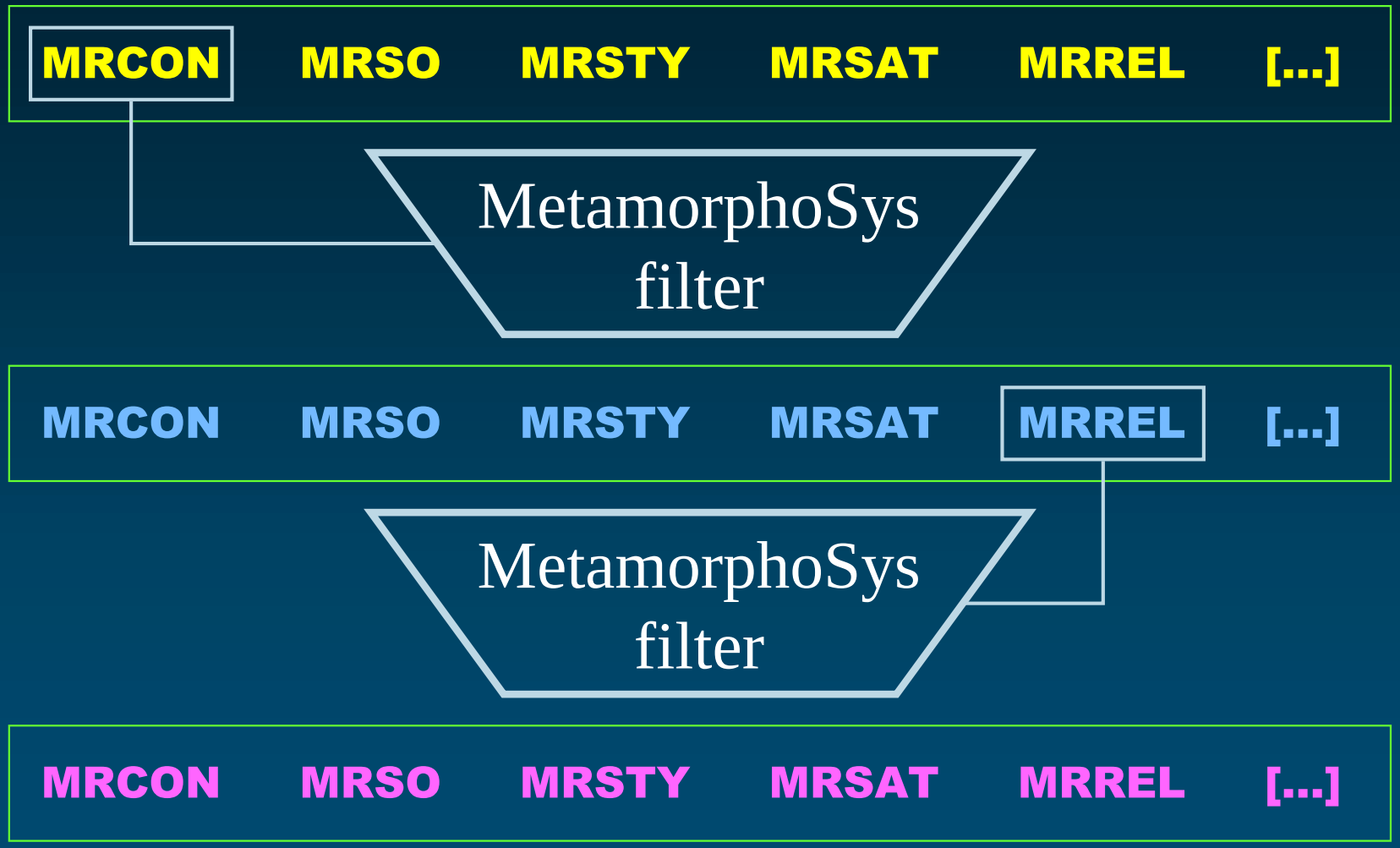
CUI1	VER	CREL	CUI2
C0241779	1996	SY	C0001403
C0271735	1996	SY	C0001403



Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content
    - Customize with MetamorphoSys — Olivier Bodenreider (1/3)
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

How does MetamorphoSys work?



Filter by language

Exclude non-English

MRCON

Filter by source

Exclude SNOMED Intl

MRSO

Concept **C0001621**

Term **L0001621**	**S0011232** *Adrenal Gland Diseases*	MeSH	
	S0011231 Adrenal Gland Disease	MeSH	
	S0000441 Disease of adrenal gland	SNOMED 2	
	~~**S0000700** Disease of adrenal gland, NOS~~	~~SNOMED Intl~~	
	S0220090 Disease, adrenal gland	MeSH	
	S0044801 Gland Disease, Adrenal	MeSH	[…]
Term **L0041793**	**S0860744** *Disorder of adrenal gland, unspecified*	ICD-10	
	S0217833 Unspecified disorder of adrenal glands	ICD-9 MedDRA	[…]
Term **L0161347**	**S0225481** *ADRENAL DISORDER*	COSTAR CCPSS	
	S0627685 DISORDER ADRENAL (NOS)	COSTAR	[…]
Term **L0181041**	**S0632950** *Disorder of adrenal gland*	CTV3	
	S0354509 Adrenal Gland Disorders	Th. Psych	[…]
Term **L0368399**	**S0586222** *Adrenal disease*	CTV3	
	S0466921 ADRENAL DISEASE, NOS	COSTAR	[…]
Term **L1279026**	**S1520972** *Nebennierenkrankheiten*	German MeSH	[…]
Term **L0162317**	**S0226798** *SURRENALE, MALADIES*	French MeSH	[…]

[…]

Filter by source

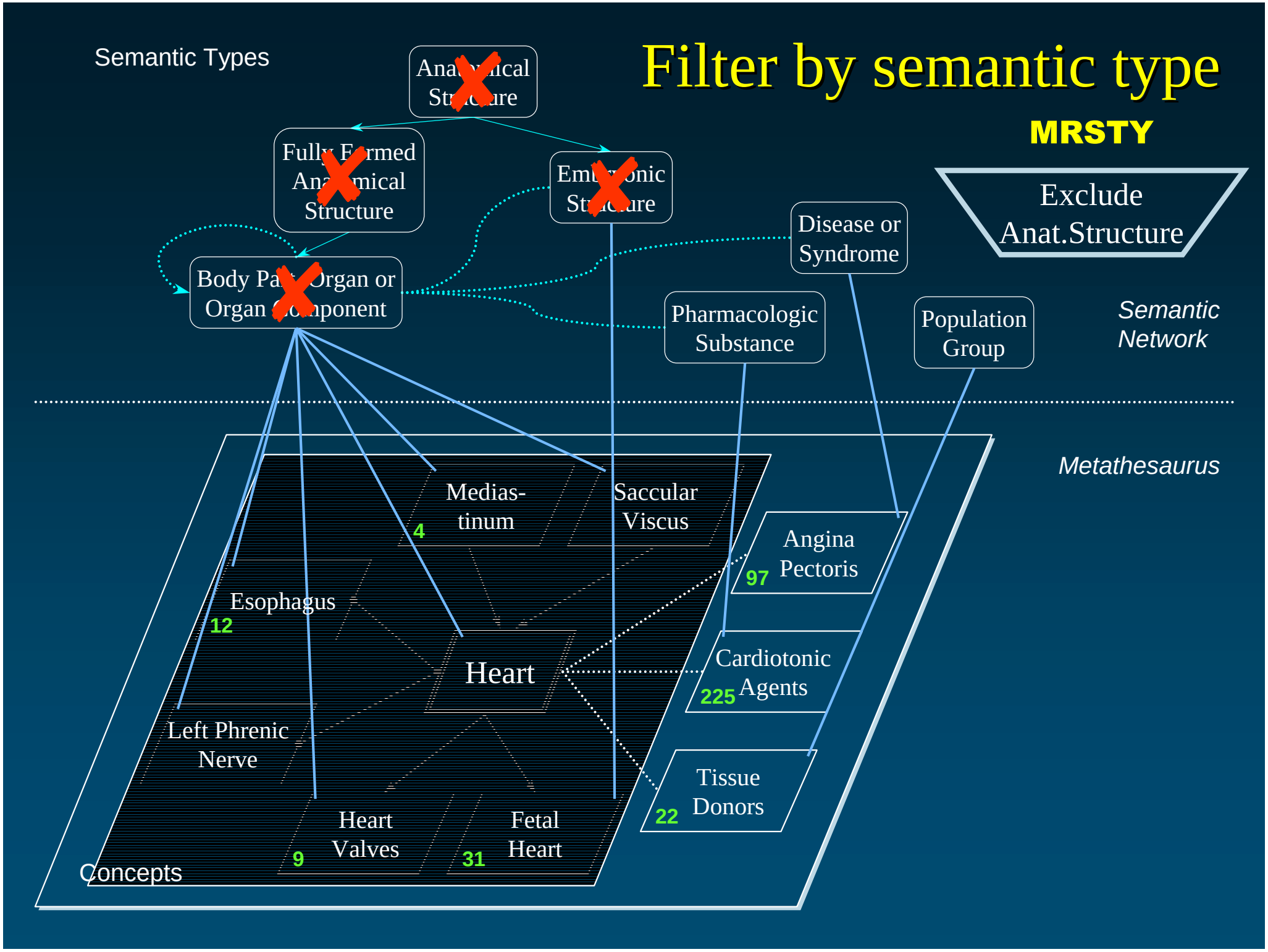
Exclude CTV3

MRSO

Concept **C0001621**

Term **L0001621**	**S0011232** *Adrenal Gland Diseases*	MeSH	
	S0011231 Adrenal Gland Disease	MeSH	
	S0000441 Disease of adrenal gland	SNOMED 2	
	S0481705 Disease of adrenal gland, NOS	SMOMED Intl	
	S0220090 Disease, adrenal gland	MeSH	
	S0044801 Gland Disease, Adrenal	MeSH	[…]
Term **L0041793**	**S0860744** *Disorder of adrenal gland, unspecified*	ICD-10	
	S0217833 Unspecified disorder of adrenal glands	ICD-9 MedDRA	[…]
Term **L0161347**	**S0225481** *ADRENAL DISORDER*	COSTAR CCPSS	
	S0627685 DISORDER ADRENAL (NOS)	COSTAR	[…]
Term **L0181041**	~~**S0630950** *Disorder of adrenal gland*~~	~~CTV3~~	
	S0354509 Adrenal Gland Disorders	Th. Psych	[…]
Term **L0368399**	~~**S0000000** *Adrenal disease*~~	~~CTV3~~	
	S0466921 ADRENAL DISEASE, NOS	COSTAR	[…]
Term **L1279026**	**S1520972** *Nebennierenkrankheiten*	German MeSH	[…]
Term **L0162317**	**S0226798** *SURRENALE, MALADIES*	French MeSH	[…]

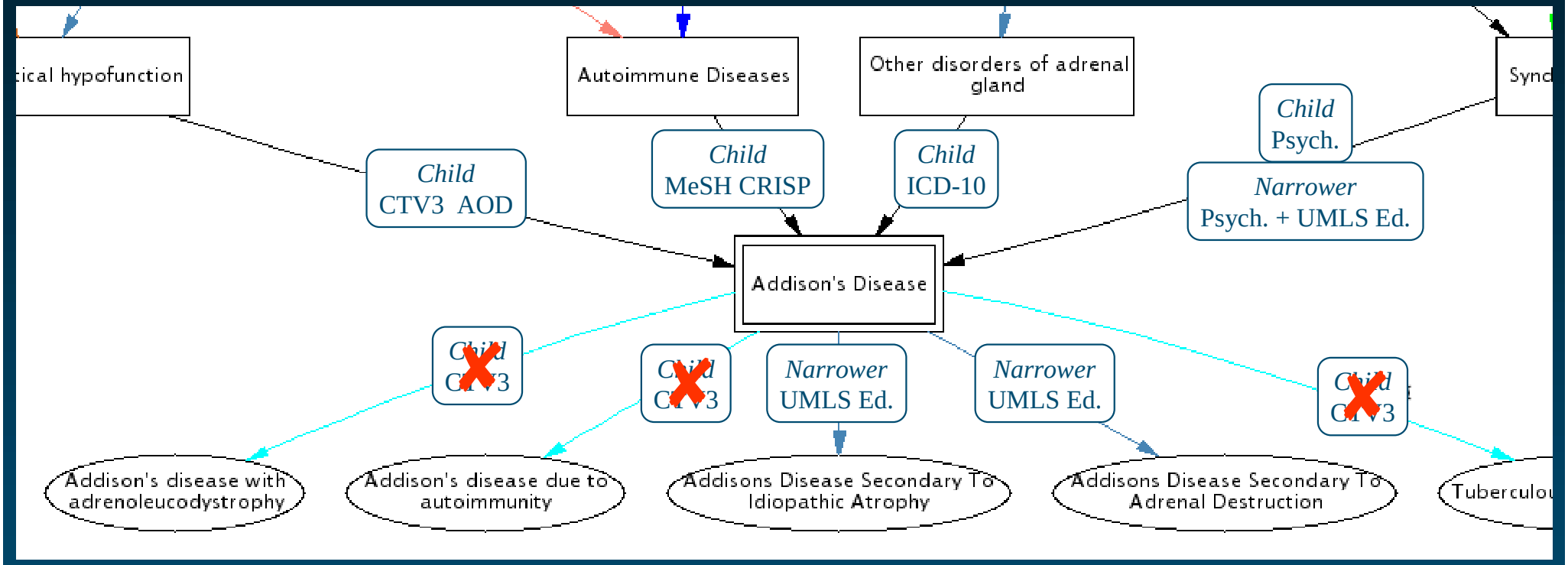
[…]


Semantic Types
Filter by semantic type
MRSTY
Exclude Anat.Structure
Fully Formed Anatomical Structure
Embryonic Structure
Body Part, Organ or Organ Component
Disease or Syndrome
Pharmacologic Substance
Population Group
Semantic Network
Metathesaurus
Medias-tinum
4
Saccular Viscus
Esophagus
12
Heart
Angina Pectoris
97
Cardiotonic Agents
225
Left Phrenic Nerve
Tissue Donors
22
Heart Valves
9
Fetal Heart
31
Concepts

Exclude relationships

Exclude
Child in CTV3

MRREL



Other MetamorphoSys actions

- Modify precedence **MRRANK**
- Exclude attribute **MRSAT**
- Exclude suppressible strings
- Write your own filter

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content
    - Customize with MetamorphoSys — Laura Roth (2/3)
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

MetamorphoSys

- A tool distributed for use with the UMLS Knowledge Sources
  - Already present in UMLS distribution in $UMLSHOME/METAMSYS directory
- Multi-platform Java software
- Creates a customized version of the Metathesaurus
- New version with added features released with 2002AD UMLS

How does MetamorphoSys work?

- What it does: removes all information from MR* files that is supplied by the excluded vocabularies
  - This includes strings, relationships, attributes, mappings, etc.
- OR removes only selected relationships or attributes but not entire concepts
- What results: A full Metathesaurus, including all the MR* files, containing information that matches what the user requested

What is new with MetamorphoSys?

- Includes 4 new filters
  - Attributes – removes only selected attributes and not entire concepts
  - Languages – removes strings from a specified language but not the whole concept (unless the concept only has strings from that particular language)
  - Relationships – removes only selected relationships
  - Semantic Types – removes concepts that contain semantic types selected for exclusion

What is new with MetamorphoSys?

- Users can create their own filters
- Undo/Redo capabilities
- Output/Input formats can vary
- Uses new versioned and versionless Source Abbreviations
- Can be run in batch mode without the GUI
- Log file contains more information

How to use MetamorphoSys

- Machine requirements
- Graphical User Interface
- Customizing with the interface

Machine requirements

- A minimum of 256 MB of physical memory, as well as 8 GB recommended free disk space
  - Full UMLS distribution needs to be present
  - MetamorphoSys needs to be in the same directory as the data
- Can run on all common Java platforms

Graphical User Interface

- Started by the MetamorphoSys program once UMLS distribution has been unpacked
  - Found in the $UMLSHOME/METAMSYS directory
  - MetamorphoSys.sh starts the program in the UNIX environment
  - MetamorphoSys.bat starts the program in the Windows environment

Graphical User Interface

- Simple to use
  - Allows users to make changes and save the changes for later use without having to edit a config file
- Composed of 4 main filters with 4 additional filters that can be selected
- Also contains advanced options for filters

Graphical User Interface components

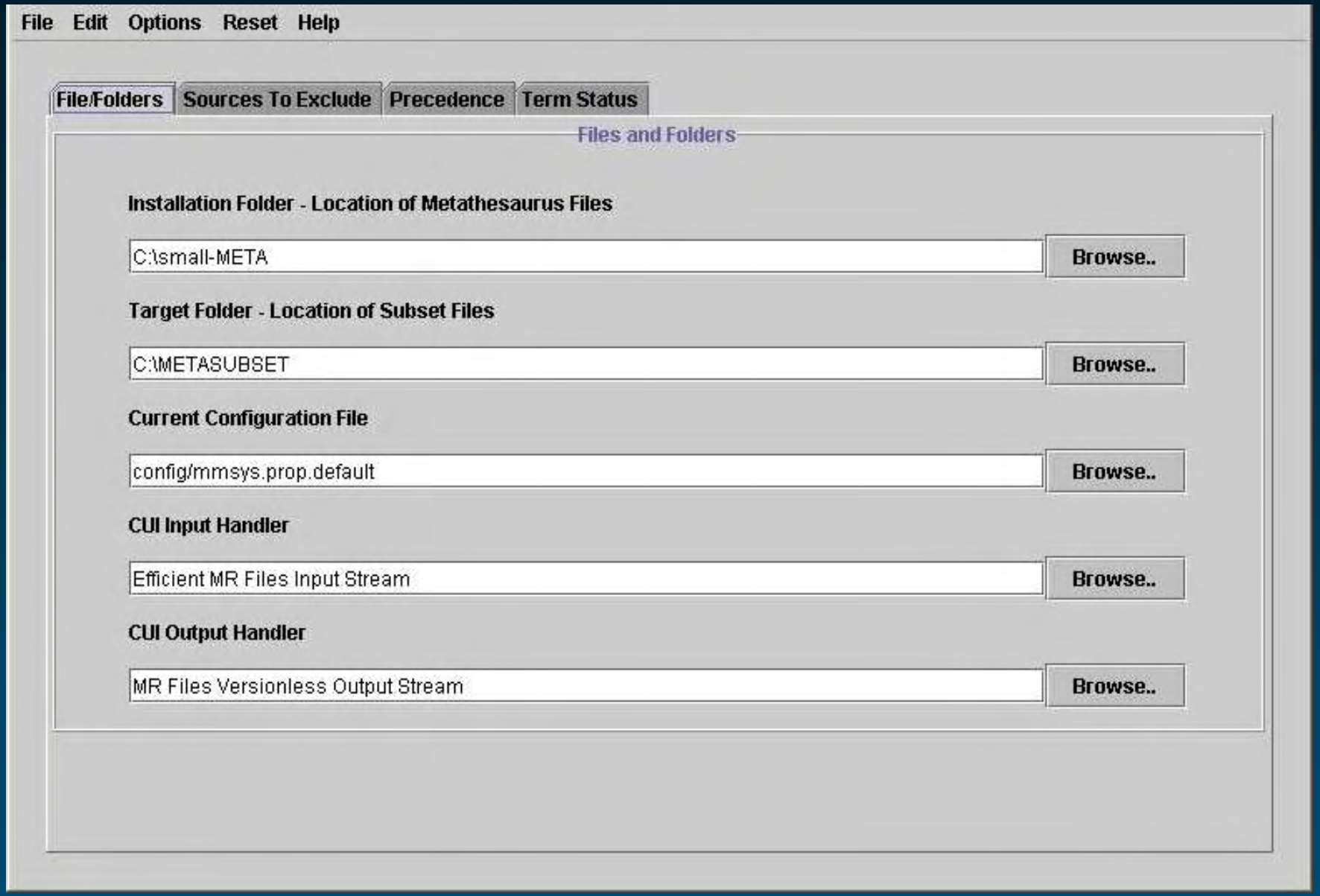
- Four main filters
  - Files/Folders
  - Sources
  - Precedence
  - Term Status

Files/ Folders

- MetamorphoSys is version aware
  - Links to Metathesaurus version it should be run against
  - On the title of the frame, the Metathesaurus version that should be used is listed
  - If a user tries to run against another version, a warning message appears

Files/ Folders

- Indicate where UMLS distribution is located
- Indicate where the customized Metathesaurus should go
- Indicate which config file should be used (default is the config file that came with MetamorphoSys but users can select their own)
- Indicate the CUI Input and Output Handler to use
- Default directories are provided but users can change if needed


File Edit Options Reset Help
File/Folders
Sources To Exclude
Precedence
Term Status
Files and Folders
Installation Folder - Location of Metathesaurus Files
C:\small-META
Browse..
Target Folder - Location of Subset Files
C:\METASUBSET
Browse..
Current Configuration File
config/mmsys.prop.default
Browse..
CUI Input Handler
Efficient MR Files Input Stream
Browse..
CUI Output Handler
MR Files Versionless Output Stream
Browse..

Sources filter

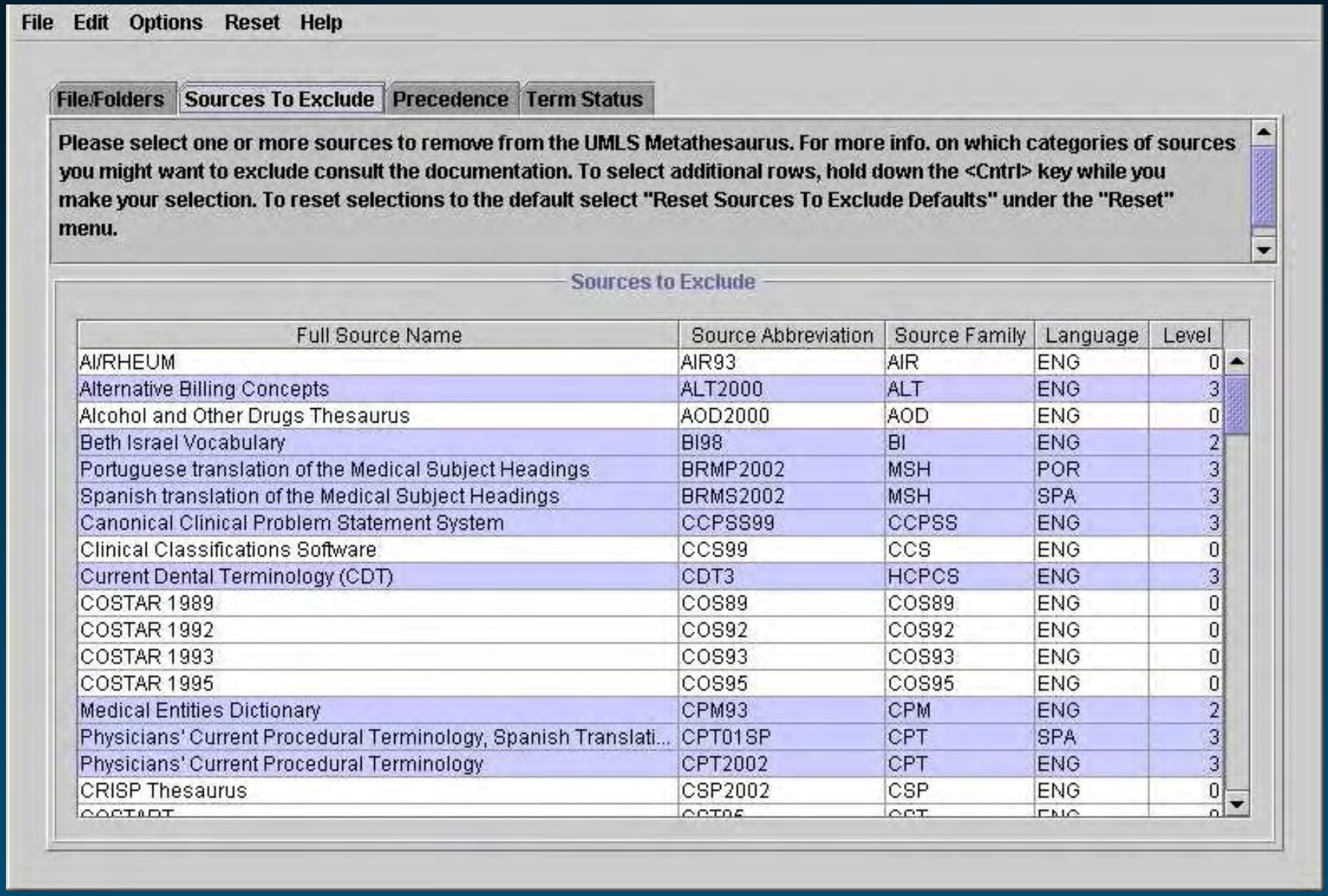
- Sources are listed alphabetically by abbreviation
  - Includes full source name, abbreviation, Source Family, language, and restriction level
  - Can be sorted on any of these fields
- Sources highlighted are the ones to be excluded
- Can change to include or exclude any vocabulary
- Options menu allows default values to be reset
- If excluding sources, want to select them before using other filters

Sources filter Source Family

- Sources are now assigned a Source Family
  - All related sources are given the same family value
  - This allows sources to be grouped together that are covered under the same licensing agreements
  - For example: WHOART and all its foreign language versions (they all have a source family value of WHO)

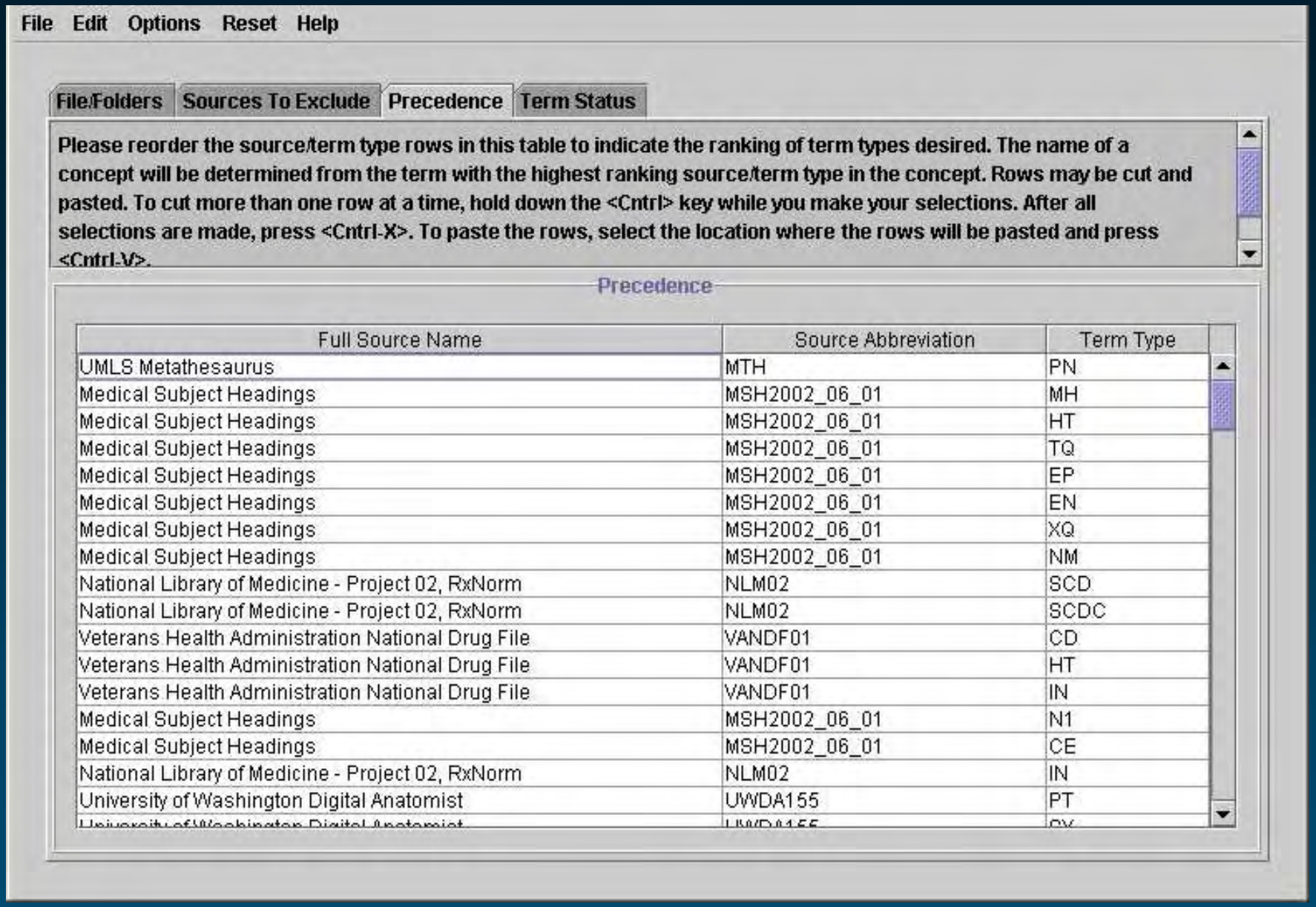
Sources filter Dependent Source

- Sources can also have a Dependent Source value
  - Sometimes sources are related in a way similar to source families but do not properly belong in the same family. These are grouped together so they can be removed together if needed
    - e.g. CPT (family=CPT) and HCPT (family=HCPCS)
  - Advanced Options allows users to create their own dependent source relationships


File Edit Options Reset Help
File/Folders
Sources To Exclude
Precedence
Term Status
Please select one or more sources to remove from the UMLS Metathesaurus. For more info. on which categories of sources you might want to exclude consult the documentation. To select additional rows, hold down the <Cntrl> key while you make your selection. To reset selections to the default select "Reset Sources To Exclude Defaults" under the "Reset" menu.
Sources to Exclude
Full Source Name | Source Abbreviation | Source Family | Language | Level
AI/RHEUM | AIR93 | AIR | ENG | 0
Alternative Billing Concepts | ALT2000 | ALT | ENG | 3
Alcohol and Other Drugs Thesaurus | AOD2000 | AOD | ENG | 0
Beth Israel Vocabulary | BI98 | BI | ENG | 2
Portuguese translation of the Medical Subject Headings | BRMP2002 | MSH | POR | 3
Spanish translation of the Medical Subject Headings | BRMS2002 | MSH | SPA | 3
Canonical Clinical Problem Statement System | CCPSS99 | CCPSS | ENG | 3
Clinical Classifications Software | CCS99 | CCS | ENG | 0
Current Dental Terminology (CDT) | CDT3 | HCPCS | ENG | 3
COSTAR 1989 | COS89 | COS89 | ENG | 0
COSTAR 1992 | COS92 | COS92 | ENG | 0
COSTAR 1993 | COS93 | COS93 | ENG | 0
COSTAR 1995 | COS95 | COS95 | ENG | 0
Medical Entities Dictionary | CPM93 | CPM | ENG | 2
Physicians' Current Procedural Terminology, Spanish Translati... | CPT01SP | CPT | SPA | 3
Physicians' Current Procedural Terminology | CPT2002 | CPT | ENG | 3
CRISP Thesaurus | CSP2002 | CSP | ENG | 0

Precedence filter

- MTH/PN source/term type is the default highest precedence source
- Sources are arranged by their rank with highest rank first
- Fields include full source name, source abbreviation, term type
  - Table can be sorted on any of these fields
- Sources can be rearranged as needed by cut/paste or drag/drop



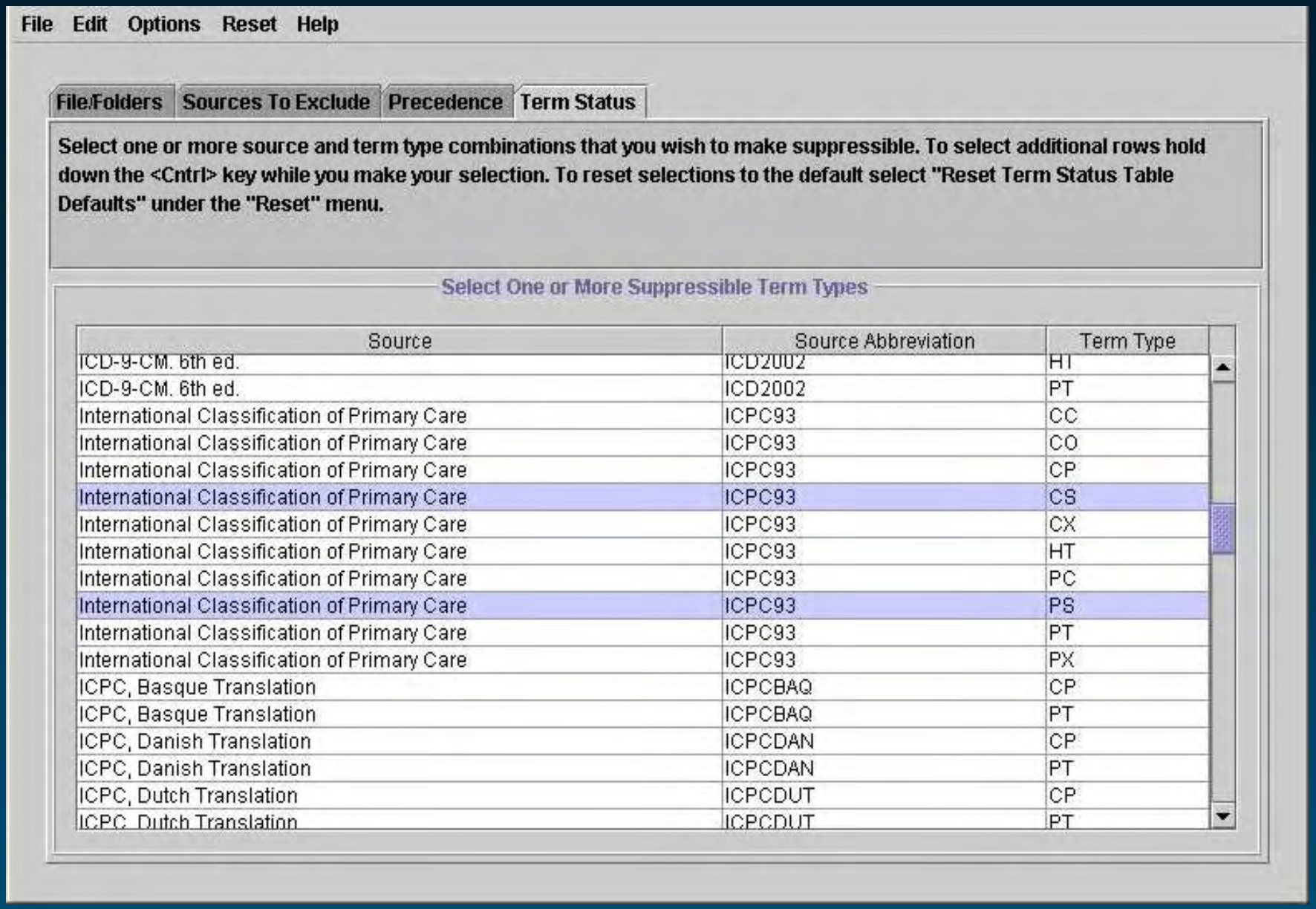
Full Source Name	Source Abbreviation	Term Type
UMLS Metathesaurus	MTH	PN
Medical Subject Headings	MSH2002_06_01	MH
Medical Subject Headings	MSH2002_06_01	HT
Medical Subject Headings	MSH2002_06_01	TQ
Medical Subject Headings	MSH2002_06_01	EP
Medical Subject Headings	MSH2002_06_01	EN
Medical Subject Headings	MSH2002_06_01	XQ
Medical Subject Headings	MSH2002_06_01	NM
National Library of Medicine - Project 02, RxNorm	NLM02	SCD
National Library of Medicine - Project 02, RxNorm	NLM02	SCDC
Veterans Health Administration National Drug File	VANDF01	CD
Veterans Health Administration National Drug File	VANDF01	HT
Veterans Health Administration National Drug File	VANDF01	IN
Medical Subject Headings	MSH2002_06_01	N1
Medical Subject Headings	MSH2002_06_01	CE
National Library of Medicine - Project 02, RxNorm	NLM02	IN
University of Washington Digital Anatomist	UWDA155	PT

Term Status filter

- Used to add or remove suppressibility
- All source-term type combinations that are suppressible are highlighted
- Can change term types that are already suppressible to non-suppressible
- New combinations can be highlighted to make suppressible

Term Status filter

- Under Advanced Options, a user can choose to remove all suppressible data from the subsetted Metathesaurus being created
- If not removed, the data is just marked as suppressible with a little "s"

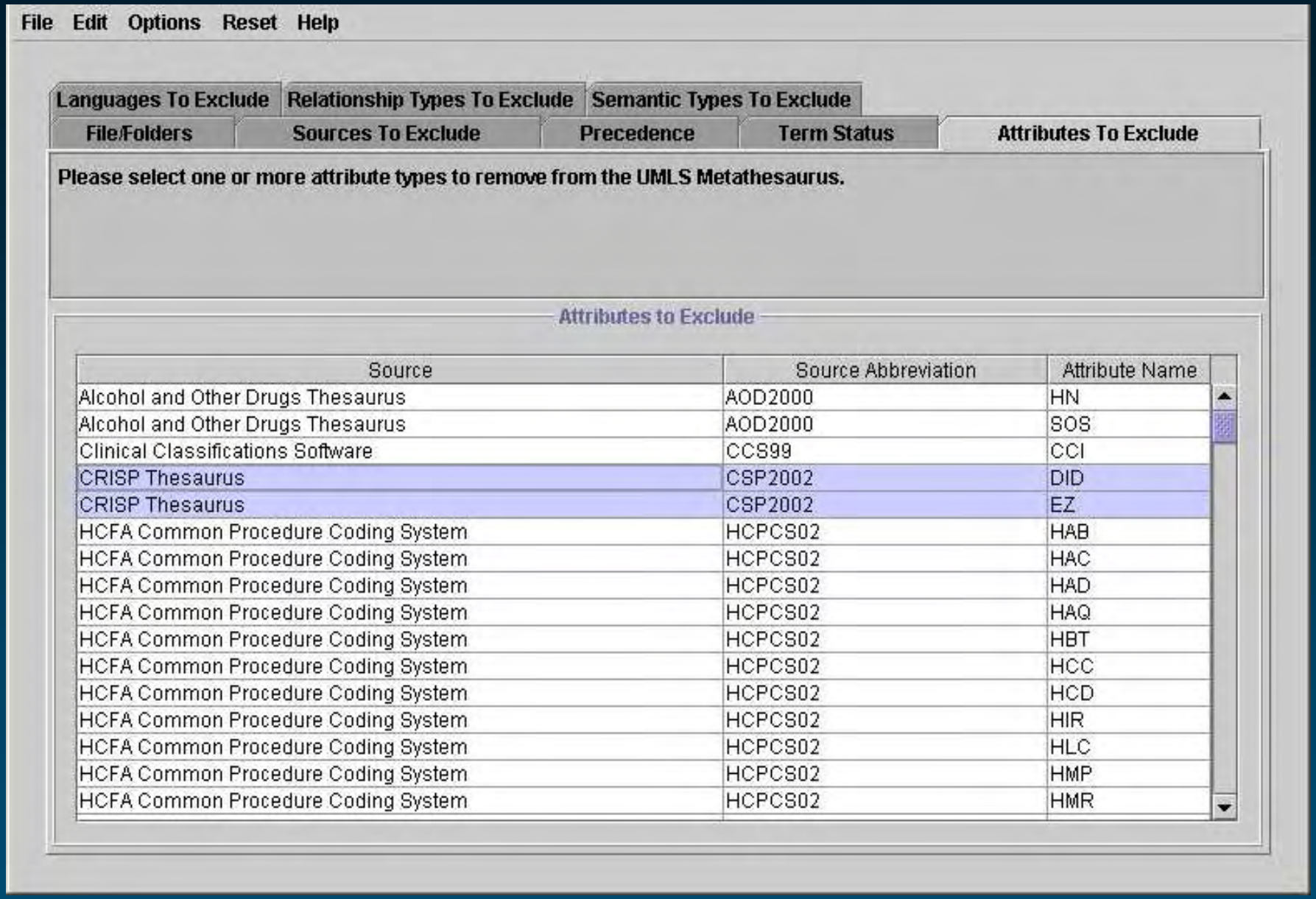

File Edit Options Reset Help
File/Folders
Sources To Exclude
Precedence
Term Status
Select one or more source and term type combinations that you wish to make suppressible. To select additional rows hold down the <Cntrl> key while you make your selection. To reset selections to the default select "Reset Term Status Table Defaults" under the "Reset" menu.
Select One or More Suppressible Term Types
Source
Source Abbreviation
Term Type
ICD-9-CM. 6th ed. ICD2002 HT
ICD-9-CM. 6th ed. ICD2002 PT
International Classification of Primary Care ICPC93 CC
International Classification of Primary Care ICPC93 CO
International Classification of Primary Care ICPC93 CP
International Classification of Primary Care ICPC93 CS
International Classification of Primary Care ICPC93 CX
International Classification of Primary Care ICPC93 HT
International Classification of Primary Care ICPC93 PC
International Classification of Primary Care ICPC93 PS
International Classification of Primary Care ICPC93 PT
International Classification of Primary Care ICPC93 PX
ICPC, Basque Translation ICPCBAQ CP
ICPC, Basque Translation ICPCBAQ PT
ICPC, Danish Translation ICPCDAN CP
ICPC, Danish Translation ICPCDAN PT
ICPC, Dutch Translation ICPCDUT CP
ICPC, Dutch Translation ICPCDUT PT

Graphical User Interface components

- 4 additional filters
  - Attributes
  - Language
  - Relationships
  - Semantic Types
- Do not automatically show up on GUI in default setting
- Can be found under File – Enable/Disable Filter

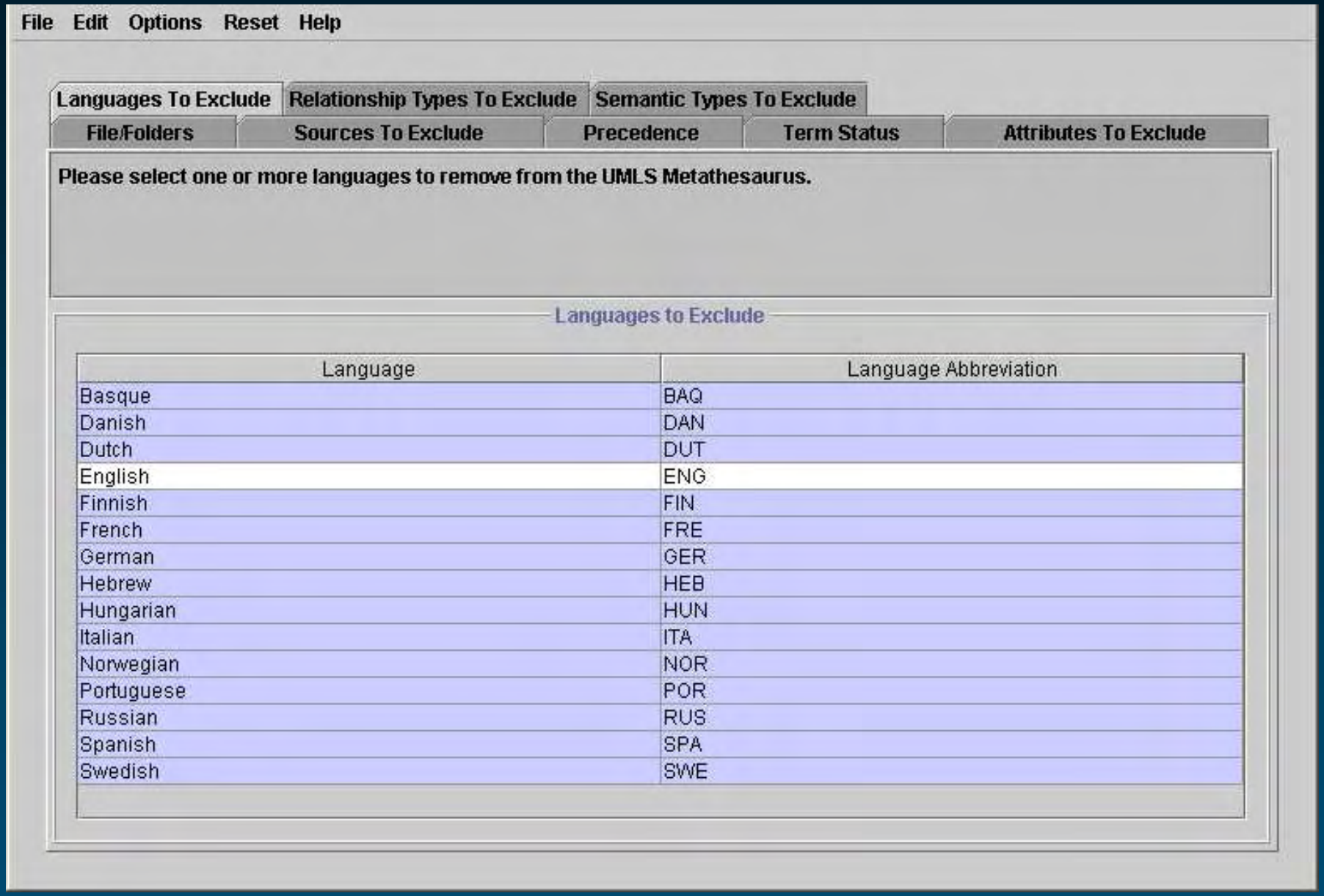
Attributes filter

- Lists source name, source abbreviation and attribute name
- If attribute is selected for exclusion, all data for this attribute is removed from MRSAT and Concept is not removed


File Edit Options Reset Help
Languages To Exclude
Relationship Types To Exclude
Semantic Types To Exclude
File/Folders
Sources To Exclude
Precedence
Term Status
Attributes To Exclude
Please select one or more attribute types to remove from the UMLS Metathesaurus.
Attributes to Exclude
Source
Source Abbreviation
Attribute Name
Alcohol and Other Drugs Thesaurus AOD2000 HN
Alcohol and Other Drugs Thesaurus AOD2000 SOS
Clinical Classifications Software CCS99 CCI
CRISP Thesaurus CSP2002 DID
CRISP Thesaurus CSP2002 EZ
HCFA Common Procedure Coding System HCPCS02 HAB
HCFA Common Procedure Coding System HCPCS02 HAC
HCFA Common Procedure Coding System HCPCS02 HAD
HCFA Common Procedure Coding System HCPCS02 HAQ
HCFA Common Procedure Coding System HCPCS02 HBT
HCFA Common Procedure Coding System HCPCS02 HCC
HCFA Common Procedure Coding System HCPCS02 HCD
HCFA Common Procedure Coding System HCPCS02 HIR
HCFA Common Procedure Coding System HCPCS02 HLC
HCFA Common Procedure Coding System HCPCS02 HMP
HCFA Common Procedure Coding System HCPCS02 HMR

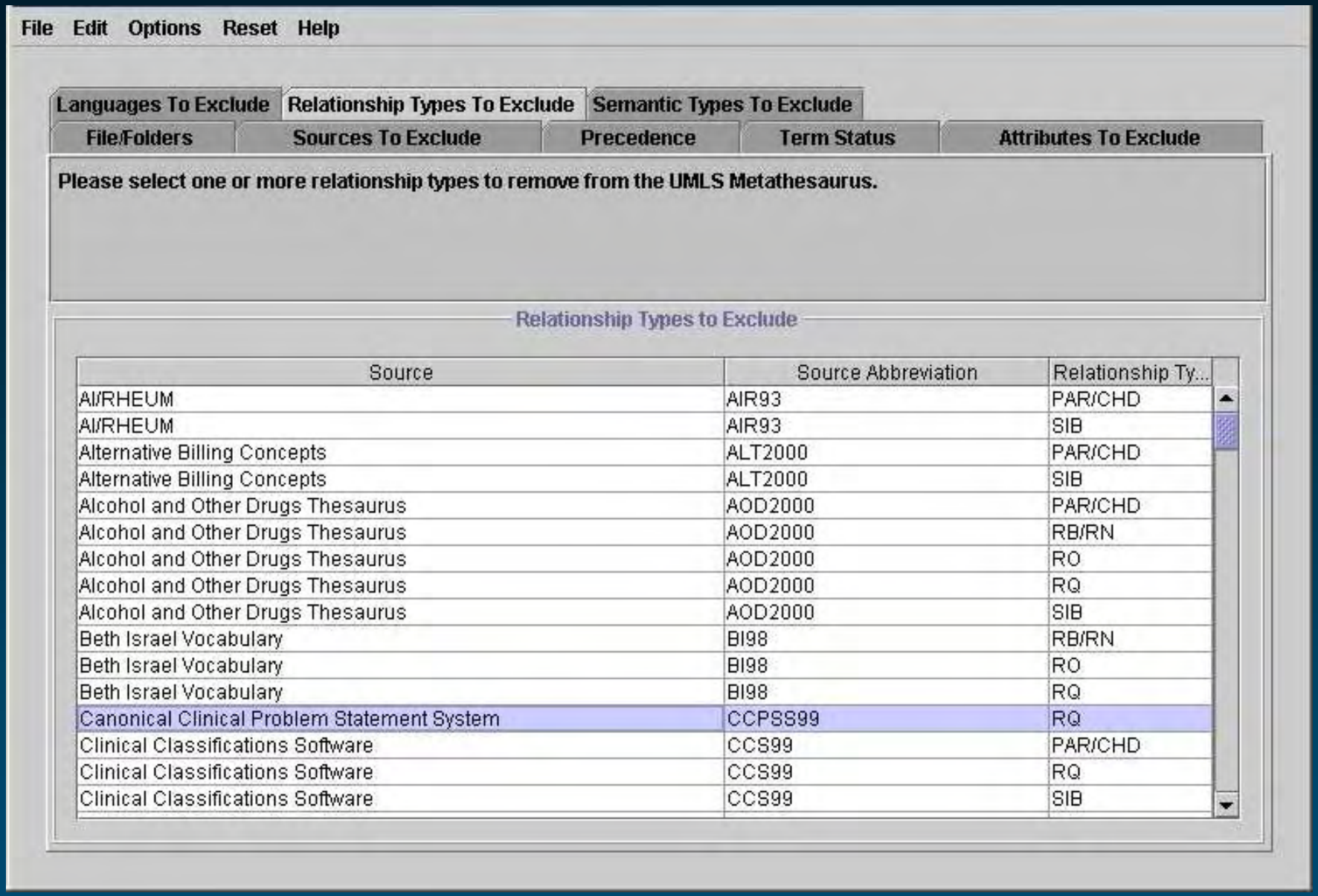
Language filter

- Lists language and language abbreviation
- Default is to exclude all non-English languages
- If language is excluded, all strings from the specified language will be removed as well as all attributes and relationships connected to those strings
- If all strings in a concept are from languages to be excluded, then the entire concept will be removed from the output subset


File Edit Options Reset Help
Languages To Exclude
Relationship Types To Exclude
Semantic Types To Exclude
File/Folders
Sources To Exclude
Precedence
Term Status
Attributes To Exclude
Please select one or more languages to remove from the UMLS Metathesaurus.
Languages to Exclude
Language
Language Abbreviation
Basque BAQ
Danish DAN
Dutch DUT
English ENG
Finnish FIN
French FRE
German GER
Hebrew HEB
Hungarian HUN
Italian ITA
Norwegian NOR
Portuguese POR
Russian RUS
Spanish SPA
Swedish SWE

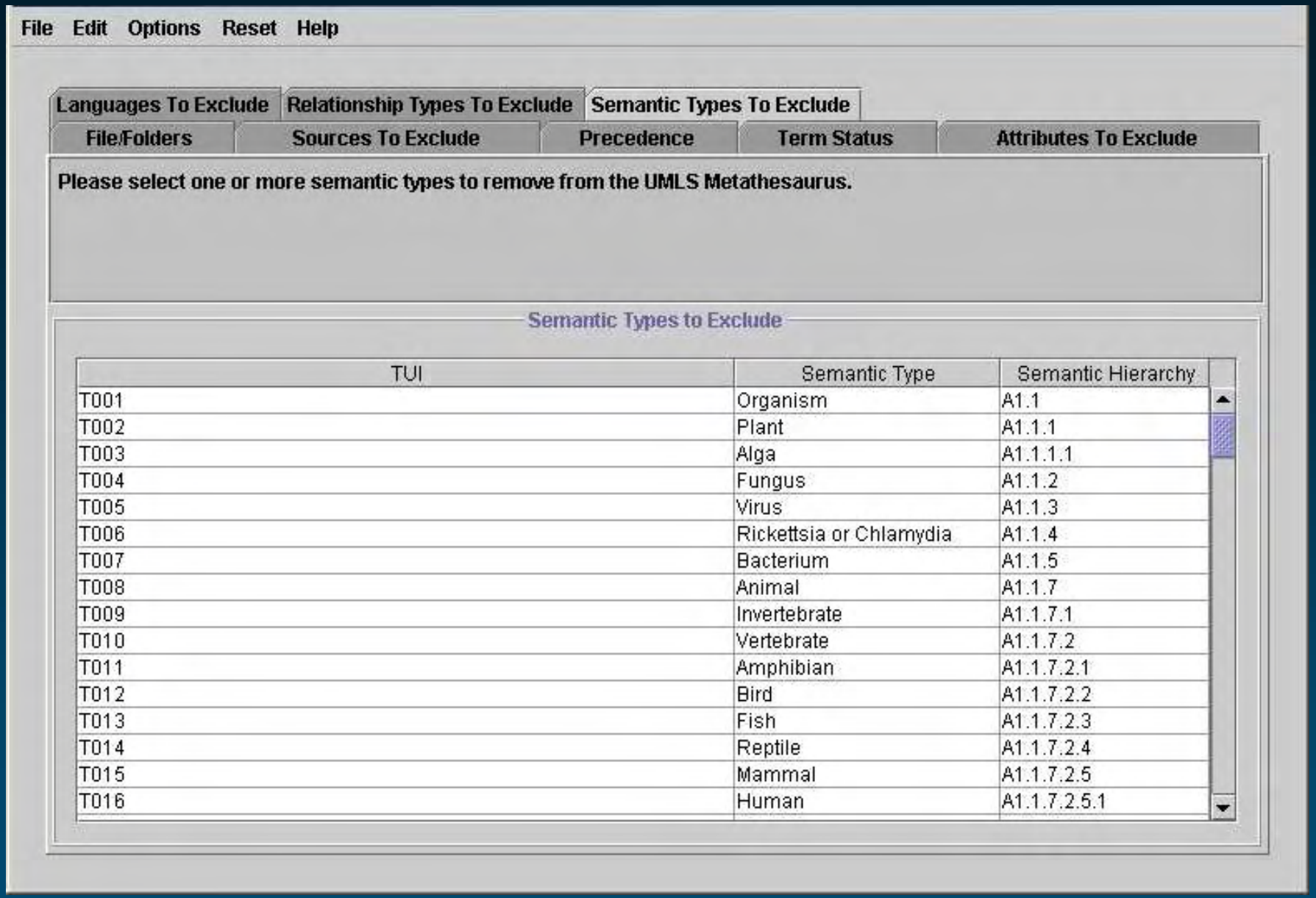
Relationships filter

- Lists source name, source abbreviation and relationship type
- This filter removes only relationship data from MRREL and not entire concepts from the output subset
- Only shows relationships from sources that will be included in the subset


File Edit Options Reset Help
Languages To Exclude
Relationship Types To Exclude
Semantic Types To Exclude
File/Folders
Sources To Exclude
Precedence
Term Status
Attributes To Exclude
Please select one or more relationship types to remove from the UMLS Metathesaurus.
Relationship Types to Exclude
Source
Source Abbreviation
Relationship Ty...
AI/RHEUM AIR93 PAR/CHD
AI/RHEUM AIR93 SIB
Alternative Billing Concepts ALT2000 PAR/CHD
Alternative Billing Concepts ALT2000 SIB
Alcohol and Other Drugs Thesaurus AOD2000 PAR/CHD
Alcohol and Other Drugs Thesaurus AOD2000 RB/RN
Alcohol and Other Drugs Thesaurus AOD2000 RO
Alcohol and Other Drugs Thesaurus AOD2000 RQ
Alcohol and Other Drugs Thesaurus AOD2000 SIB
Beth Israel Vocabulary BI98 RB/RN
Beth Israel Vocabulary BI98 RO
Beth Israel Vocabulary BI98 RQ
Canonical Clinical Problem Statement System CCPSS99 RQ
Clinical Classifications Software CCS99 PAR/CHD
Clinical Classifications Software CCS99 RQ
Clinical Classifications Software CCS99 SIB

Semantic Type filter

- Lists TUI, semantic type and hierarchy
- Removes concepts that contain at least one or all semantic types selected for exclusion


File Edit Options Reset Help
Languages To Exclude
Relationship Types To Exclude
Semantic Types To Exclude
File/Folders
Sources To Exclude
Precedence
Term Status
Attributes To Exclude
Please select one or more semantic types to remove from the UMLS Metathesaurus.
Semantic Types to Exclude
TUI
Semantic Type
Semantic Hierarchy
T001 Organism A1.1
T002 Plant A1.1.1
T003 Alga A1.1.1.1
T004 Fungus A1.1.2
T005 Virus A1.1.3
T006 Rickettsia or Chlamydia A1.1.4
T007 Bacterium A1.1.5
T008 Animal A1.1.7
T009 Invertebrate A1.1.7.1
T010 Vertebrate A1.1.7.2
T011 Amphibian A1.1.7.2.1
T012 Bird A1.1.7.2.2
T013 Fish A1.1.7.2.3
T014 Reptile A1.1.7.2.4
T015 Mammal A1.1.7.2.5
T016 Human A1.1.7.2.5.1

Graphical User Interface components

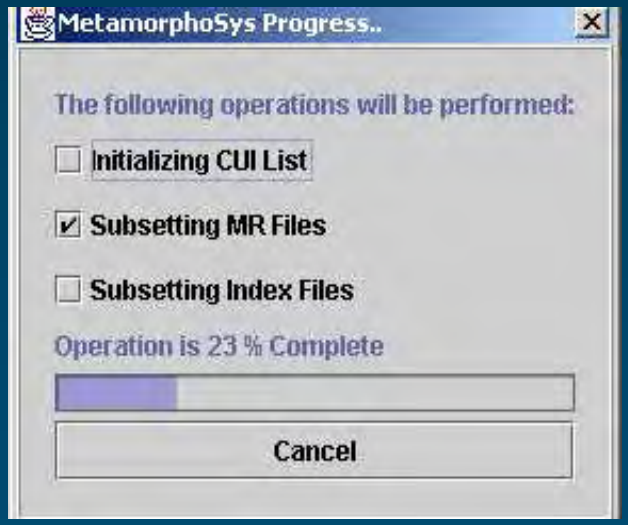
- Options Menu
  - Contains advance options for different filters
- Reset Menu
  - Resets to default values
- Help screens
  - Describes what different filters are for and what data they affect
- Undo/Redo function under Edit menu
- User created filters can be imported
  - Under File – Import Filter

Running MetamorphoSys

- Once configuration is defined, a simple file selection starts subsetting
  - Under File Menu – Begin MetamorphoSys
- Before subsetting begins, user is asked if they want the current config file (with all changes) to be saved
  - This is how a user can save changes for future runs of MetamorphoSys

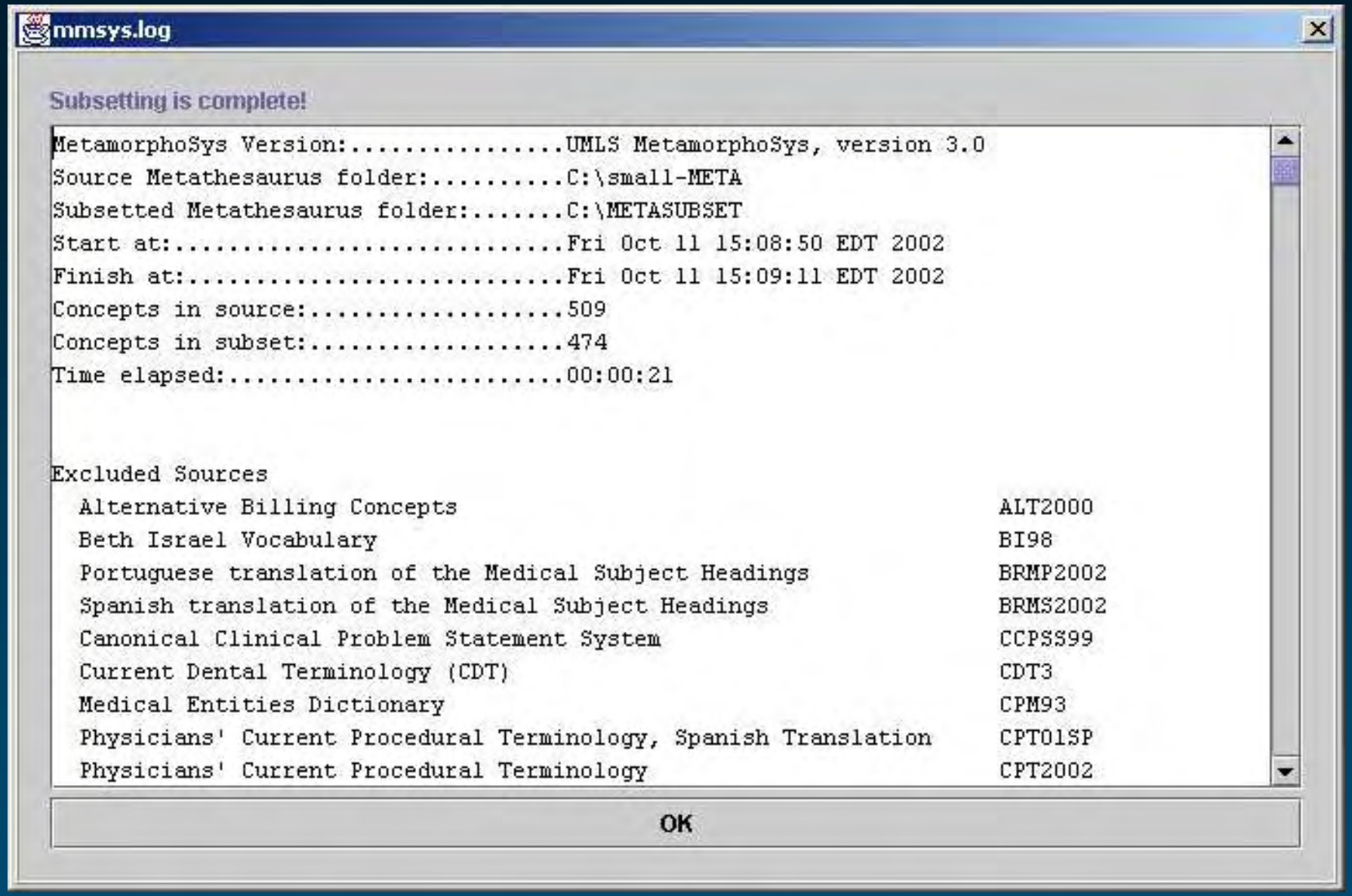
Progress Monitor

- Once subsetting begins, a progress monitor tracks process
  - Tracks progress through three major steps
  - Screen disappears only when subsetting is complete
  - “Cancel” ends the subsetting process



Log File

- After completion, a log file screen appears to indicate the process is complete and will report any errors
  - Log lists data files used, where the subsetted Metathesaurus is, name of configuration file used, number of concepts in subsetted files, time elapsed, and criteria selected to create the subset
  - Found in subset directory


mmsys.log
Subsetting is complete!
MetamorphoSys Version:.....UMLS MetamorphoSys, version 3.0
Source Metathesaurus folder:.....C:\small-META
Subsetted Metathesaurus folder:.....C:\METASUBSET
Start at:.....Fri Oct 11 15:08:50 EDT 2002
Finish at:.....Fri Oct 11 15:09:11 EDT 2002
Concepts in source:.....509
Concepts in subset:.....474
Time elapsed:.....00:00:21
Excluded Sources
Alternative Billing Concepts ALT2000
Beth Israel Vocabulary BI98
Portuguese translation of the Medical Subject Headings BRMP2002
Spanish translation of the Medical Subject Headings BRMS2002
Canonical Clinical Problem Statement System CCPSS99
Current Dental Terminology (CDT) CDT3
Medical Entities Dictionary CPM93
Physicians' Current Procedural Terminology, Spanish Translation CPT01SP
Physicians' Current Procedural Terminology CPT2002
OK

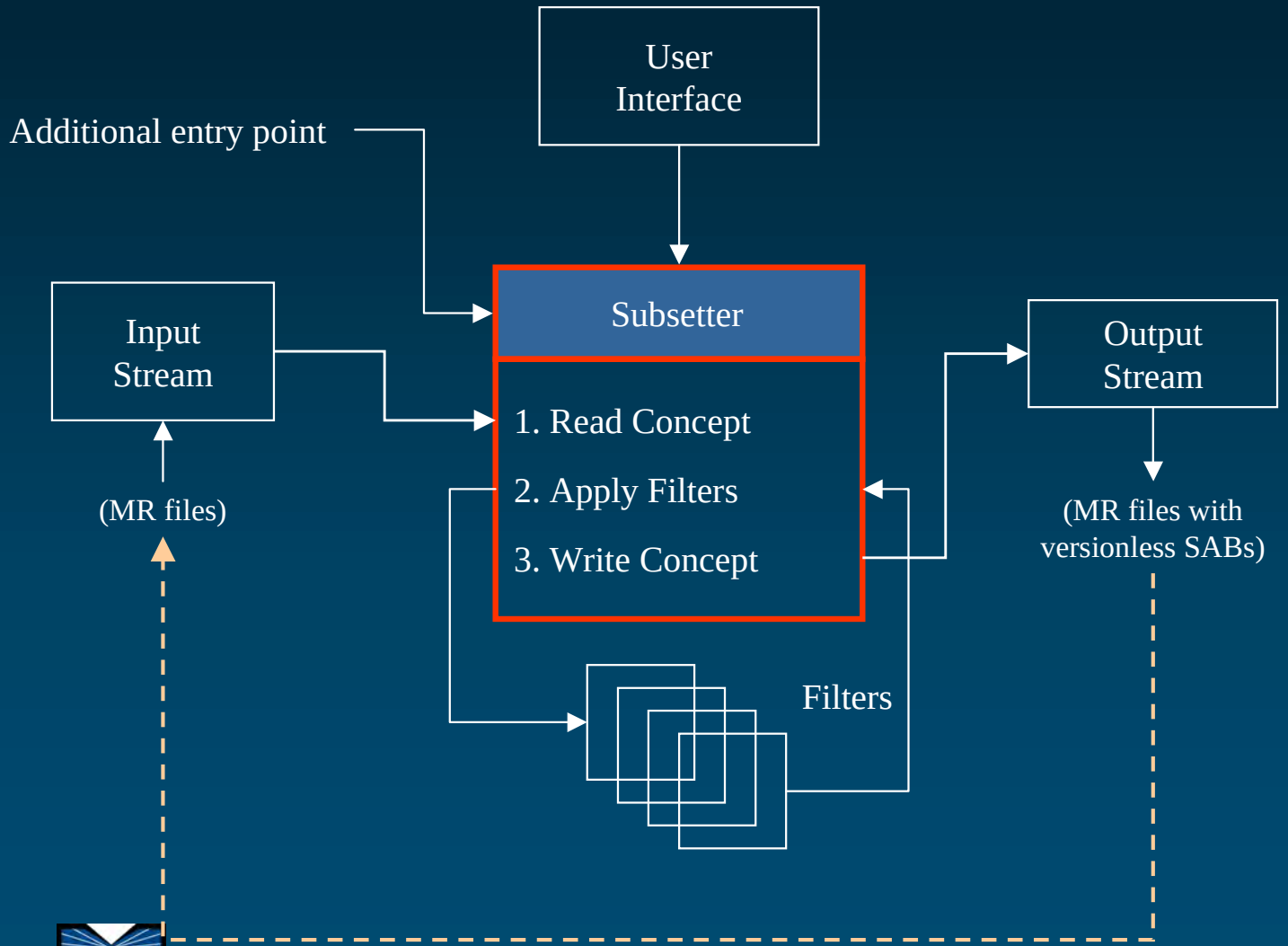
For More MetamorphoSys Information

- See README Appendix B in the tutorial handout
- Go to http://umlsinfo.nlm.nih.gov and click on the UMLS Tools section
- Read Section 2.8 of the UMLS Documentation

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content
    - Customize with MetamorphoSys — Suresh Srinivasan (3/3)
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

MetamorphoSys schematic



MetamorphoSys details

- MetamorphoSys output for:
  - Source exclusion
  - Altering precedence
  - Adding to suppressibility
- Additional Customization

Metathesaurus data for **C0001403 (“Addison’s Disease ”)**

Customize with MetamorphoSys

MRCON, MRSO data for C0001403

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70620|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70620|3|
```

MRCON, MRSO data for C0001403

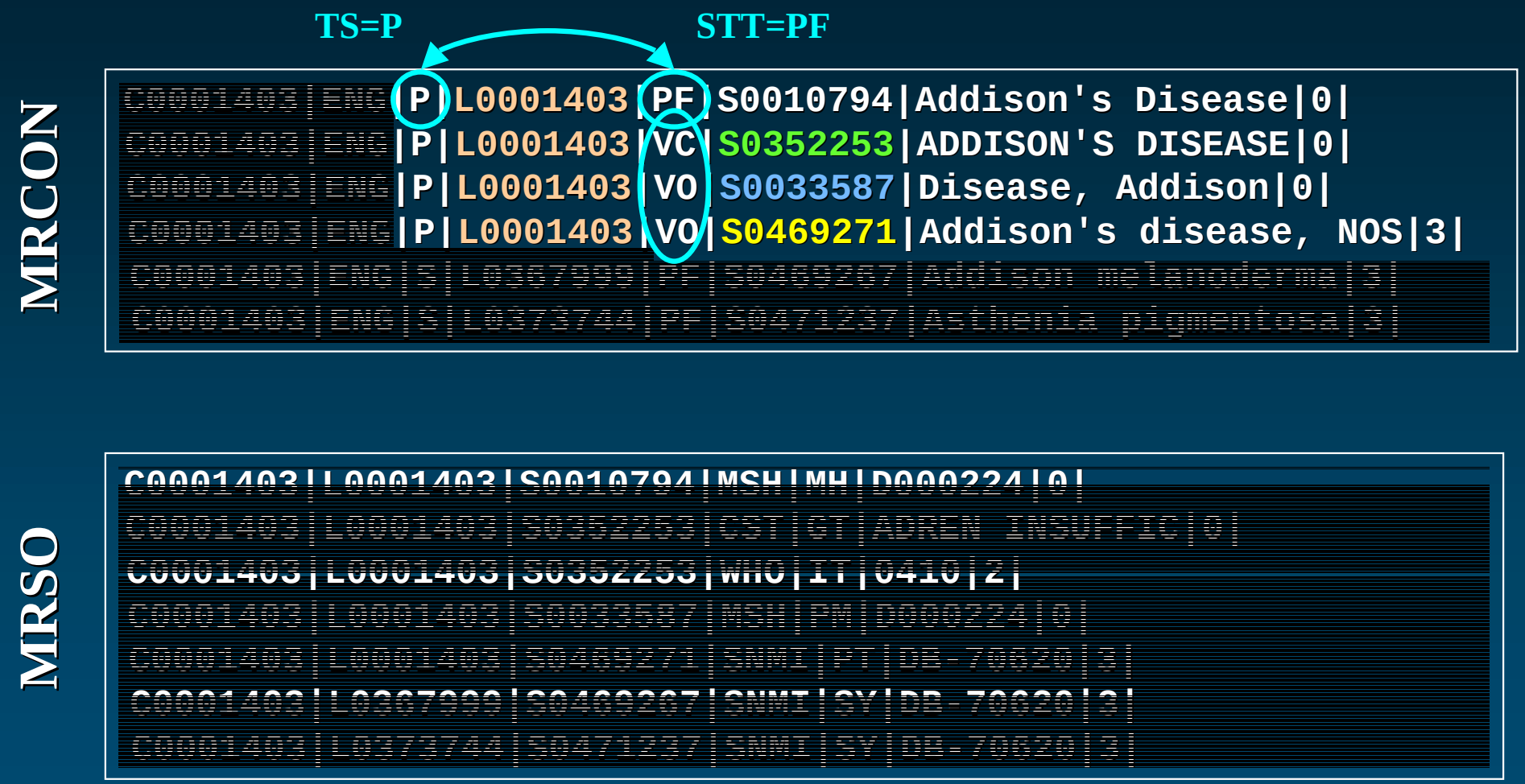
MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

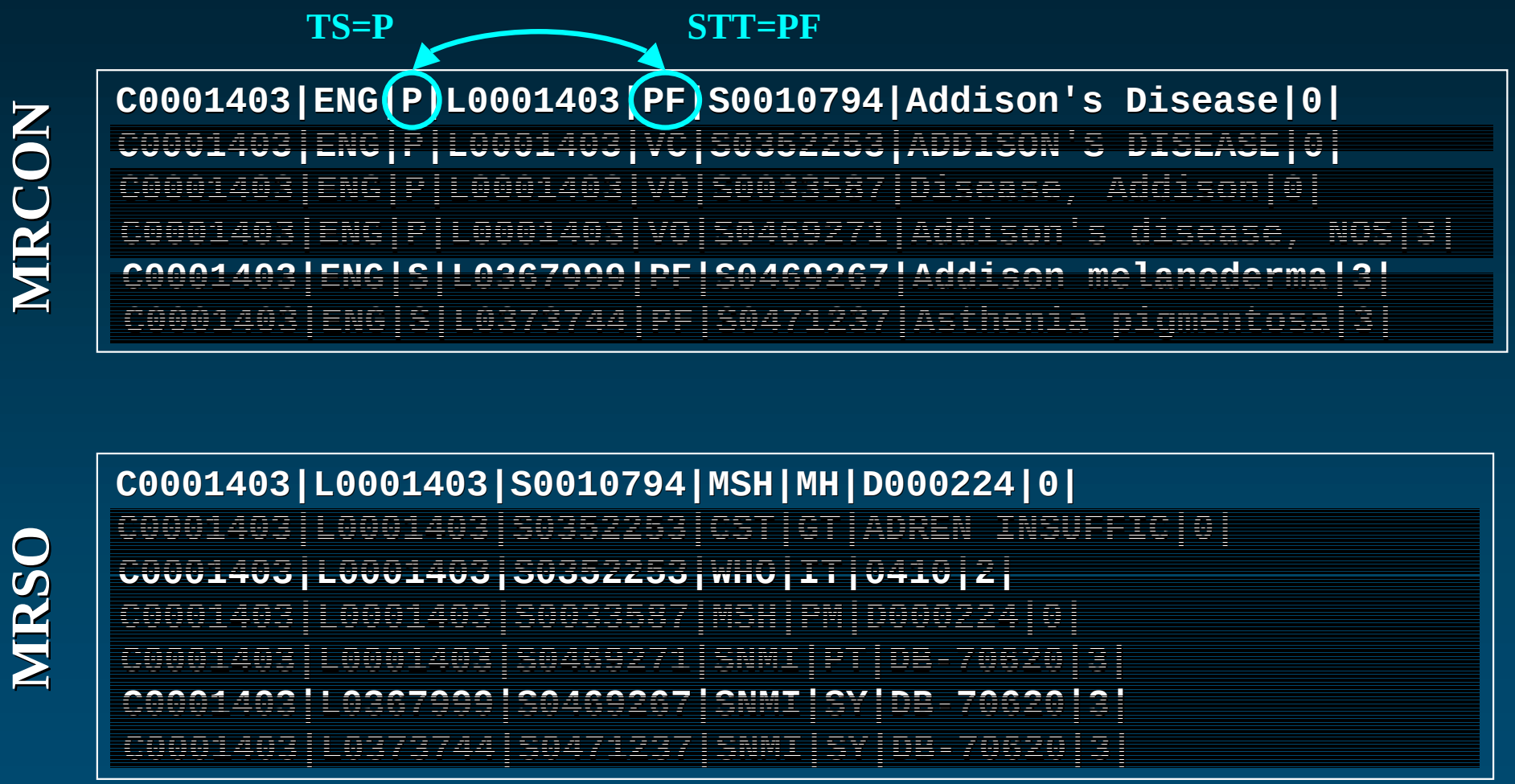
MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

MRCON, MRSO data for C0001403



MRCON, MRSO data for C0001403



MRCON, MRSO data for C0001403

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0409271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0409271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

MRCON, MRSO data for C0001403

MRCON

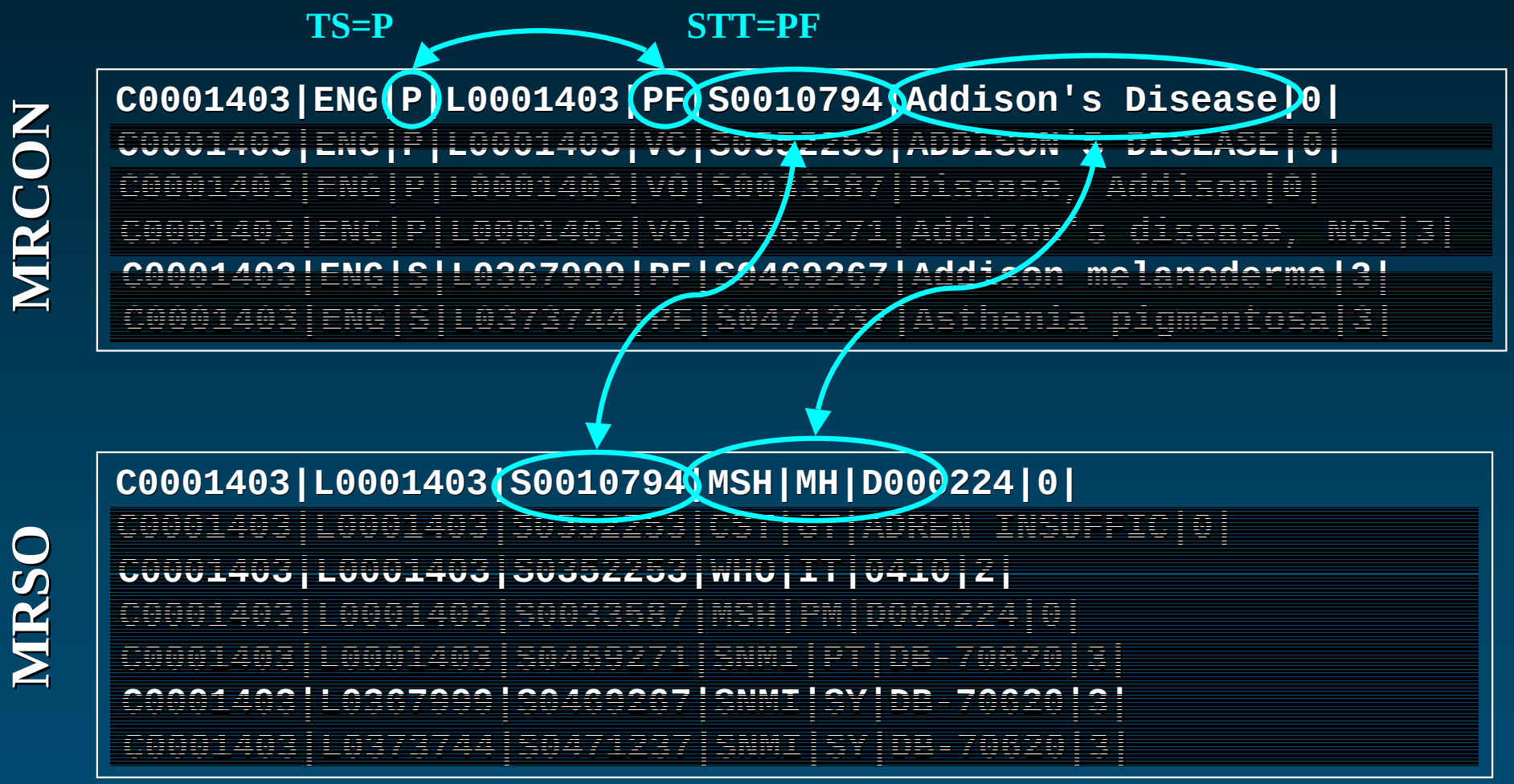
```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CSP|PT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

MRCON, MRSO data for C0001403

MRCON, MRSO data for C0001403



MRREL, MRSAT data for C0001403

MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0004364|inverse_isa|MSH|MSH||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RN|C0518010||MTH|MTH||
C0001403|RO|C0085860||MTH|MTH||
C0001403|RO|C0546992|associated_with|SNMI|SNMI||
```

Addison's Disease <has child> Tuberculous Addison's disease

MRSAT

```
C0001403|L0001403|S0010704|D000224|MN|MSH|C20.111.163|
C0001403|L0001403|S0010704|D000224|MQ|MSH|MQ000340|
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
C0001403|L0001403|S1191433|1001300|MPC|MDR|1000001300|
```

MRREL, MRSAT data for C0001403

MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0001621||PSY|PSY||
C0001403|PAR|C0004364|inverse_isa|MSH|MSH||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RN|C0518033||MTH|MTH||
C0001403|RO|C0085860||MTH|MTH||
C0001403|RO|C0546992|associated_with|SNMI|SNMI||
```

Context Relationships from Sources

MRSAT

```
C0001403|L0001403|S0010794|D000224|MN|
C0001403|L0001403|S0010794|D000224|MUE|MSH|M0000340|
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
C0001403|L0001403|S1191433|1001300|MPC|MDR|1000001300|
```

MRREL, MRSAT data for C0001403

MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0001621||PSY|PSY||
C0001403|PAR|C0004364|inverse_isa|MSH|MSH||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RN|C0518933||MTH|MTH||
C0001403|RO|C0085860||MTH|MTH||
C0001403|RO|C0546992|associated_with|SNMI|SNMI||
```

Other Relationships from Sources and MTH

MRSAT

```
C0001403|L0001403|S001079
C0001403|L0001403|S0046927
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
C0001403|L0001403|S1019433|1001300|MPC|MDR|1000001300|
```

MRREL, MRSAT data for C0001403

MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0000014O3|PAR|C0000621||PSY|PSY||
C0000014O3|PAR|C0004364|inverse_isa|MSH|MSH||
C0000014O3|RB|C0000621||MTH|MTH||
C0000014O3|RB|C0004364||CSP|CSP||
C0001403|RN|C0518033||MTH|MTH||
C0000014O3|RO|C0085860||MTH|MTH||
C0001403|RO|C0546992|associated_with|SNMI|SNMI||
```

Source Attributes

MRSAT

```
C0001403|L0001403|S0010794|D000224|MN|MSH|C20.111.163|
C0001403|L0001403|S0010794|D000224|MUI|MSH|M0000346|
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
C0001403|L0001403|S1619433|10013096|MPC|MDR|10001390|
```

Default subset using MetamorphoSys

- Removing all sources with a Source Restriction Level greater than 0
- Using default precedence ranking from MRRANK (highest precedence is MTH/PN, etc.)
- Default suppressibility and retaining suppressible rows in MRCON as TS=s
- No additional relationships and attributes removed

Default subset

Default subset: MRCON, MRSO

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70620|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70620|3|
```

Rows excluded: MRCON, MRSO

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

Restricted Sources

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70620|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70620|3|
```

Rows remaining: MRCON, MRSO

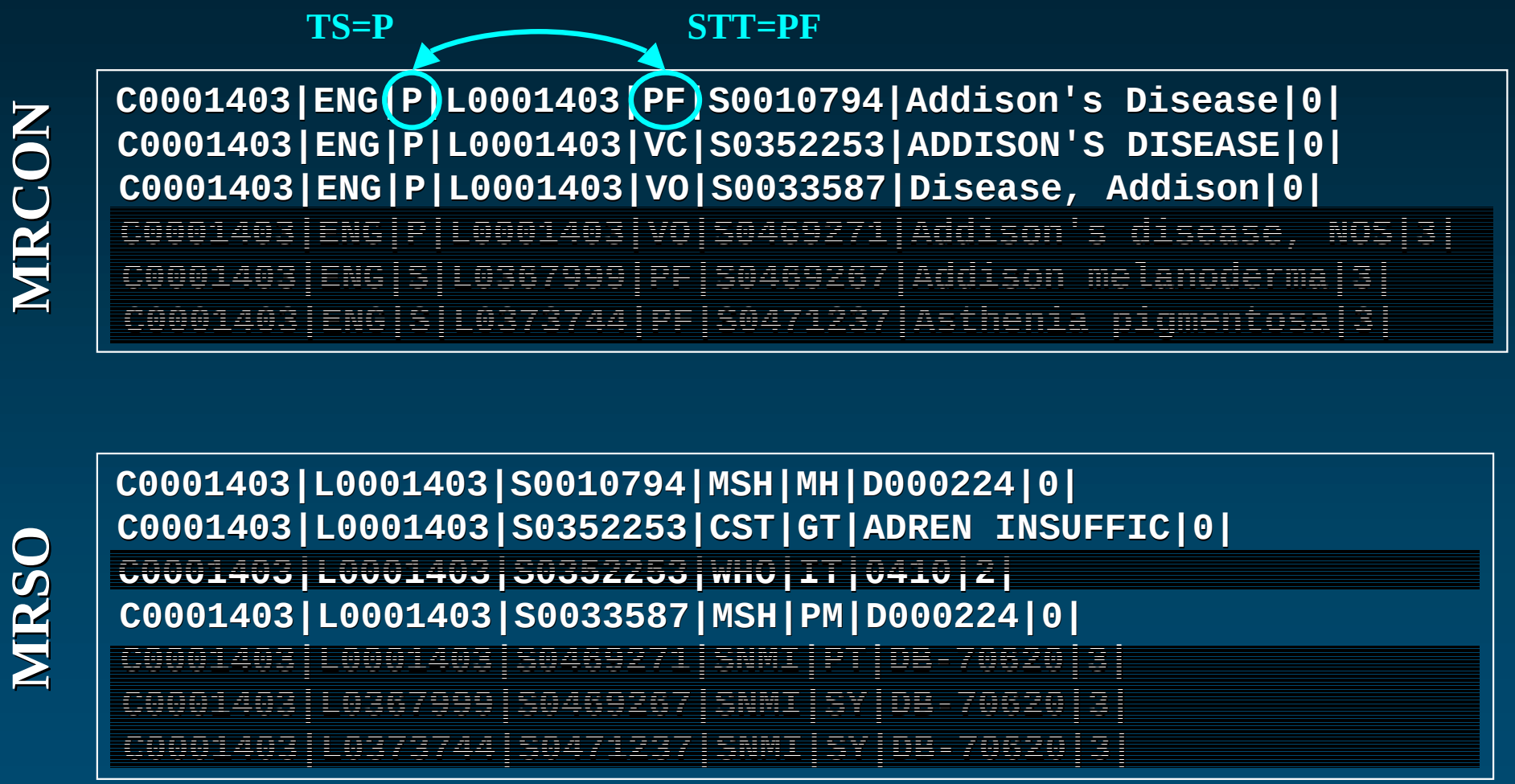
MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

Preferred name remains unchanged



S0352253 survives

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0409271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0409267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|MTH|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0409271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0409267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

Default subset: MRREL, MRSAT

MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0001621||PSY|PSY||
C0001403|PAR|C0004364|inverse_isa|MSH|MSH||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RN|C0518933||MTH|MTH||
C0001403|RO|C0085860||MTH|MTH||
C0001403|RO|C0546992|associated_with|SNMI|SNMI||
```

MRSAT

```
C0001403|L0001403|S0010794|D000224|MN|MSH|C20.111.163|
C0001403|L0001403|S0010794|D000224|MUI|MSH|M0000346|
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
C0001403|L0001403|S1619433|10013096|MPC|MDR|10001390|
```

Rows excluded: MRREL, MRSAT

MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0001621||PSY|PSY||
C0001403|PAR|C0004364|inverse_isa|MSH|MSH||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RN|C0518933||MTH|MTH||
C0001403|RO|C0085860||MTH|MTH||
C0001403|RO|C0546992|associated_with|SNMI|SNMI||
```

MRSAT

```
C0001403|L0001403|S0010794|D000224|MN|MSH|C20.111.163|
C0001403|L0001403|S0010794|D000224|MUI|MSH|M0000346|
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI|255.4|
C0001403|L0001403|S1619433|10013096|MPC|MDR|10001390|
```

Rows remaining: MRREL, MRSAT

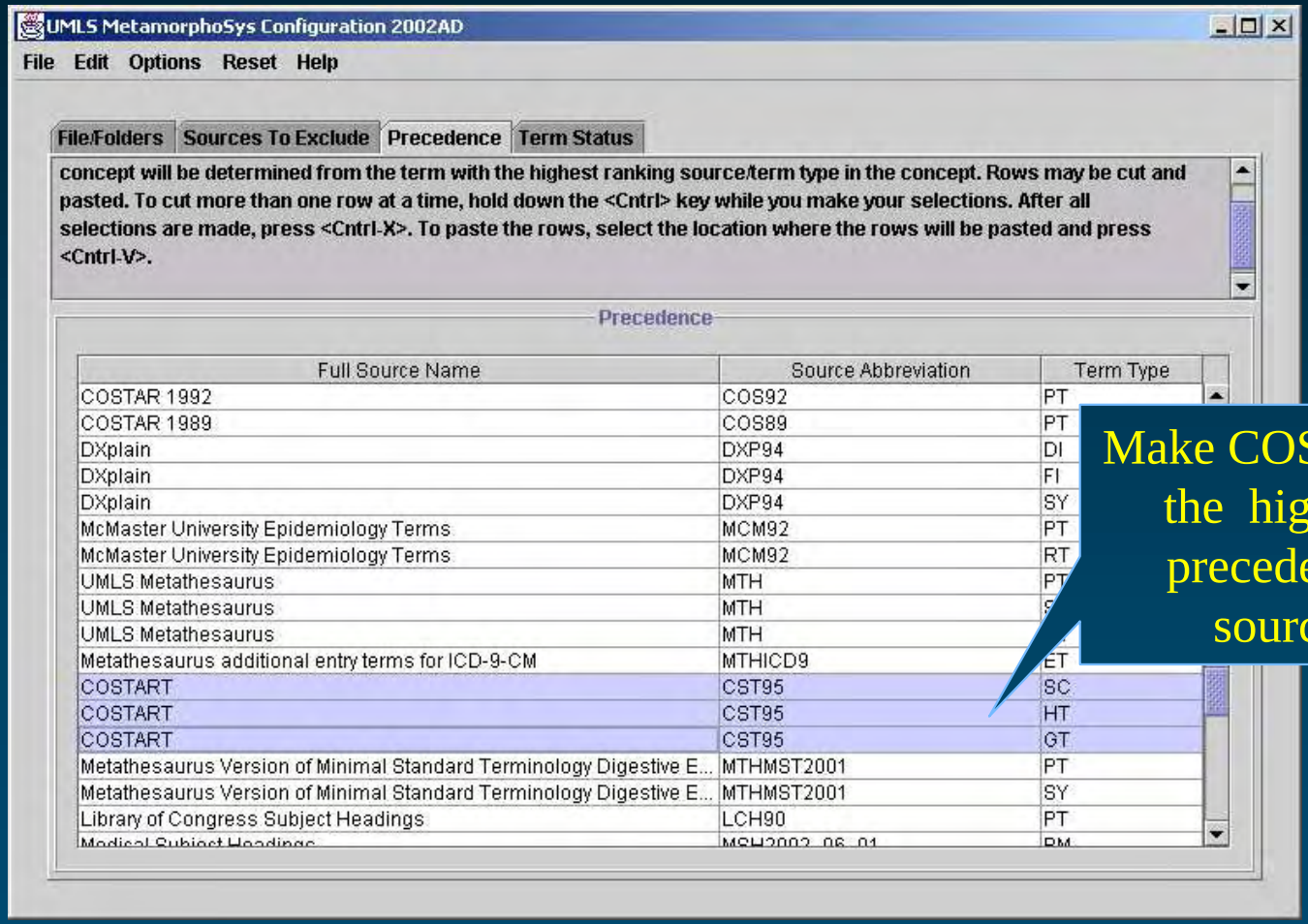
MRREL

```
C0001403|CHD|C0546992||RCD|RCD||
C0001403|PAR|C0004364|inverse_isa|MSH|MSH||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP|CSP||
C0001403|RN|C0518933||MTH|MTH||
C0001403|RO|C0085860||MTH|MTH||
```

MRSAT

```
C0001403|L0001403|S0010794|D000224|MN|MSH|C20.111.163|
C0001403|L0001403|S0010794|D000224|MUI|MSH|M0000346|
```

Changing precedence



Preferred term changes from MeSH..

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

..to COSTART (CST95)

MRCON

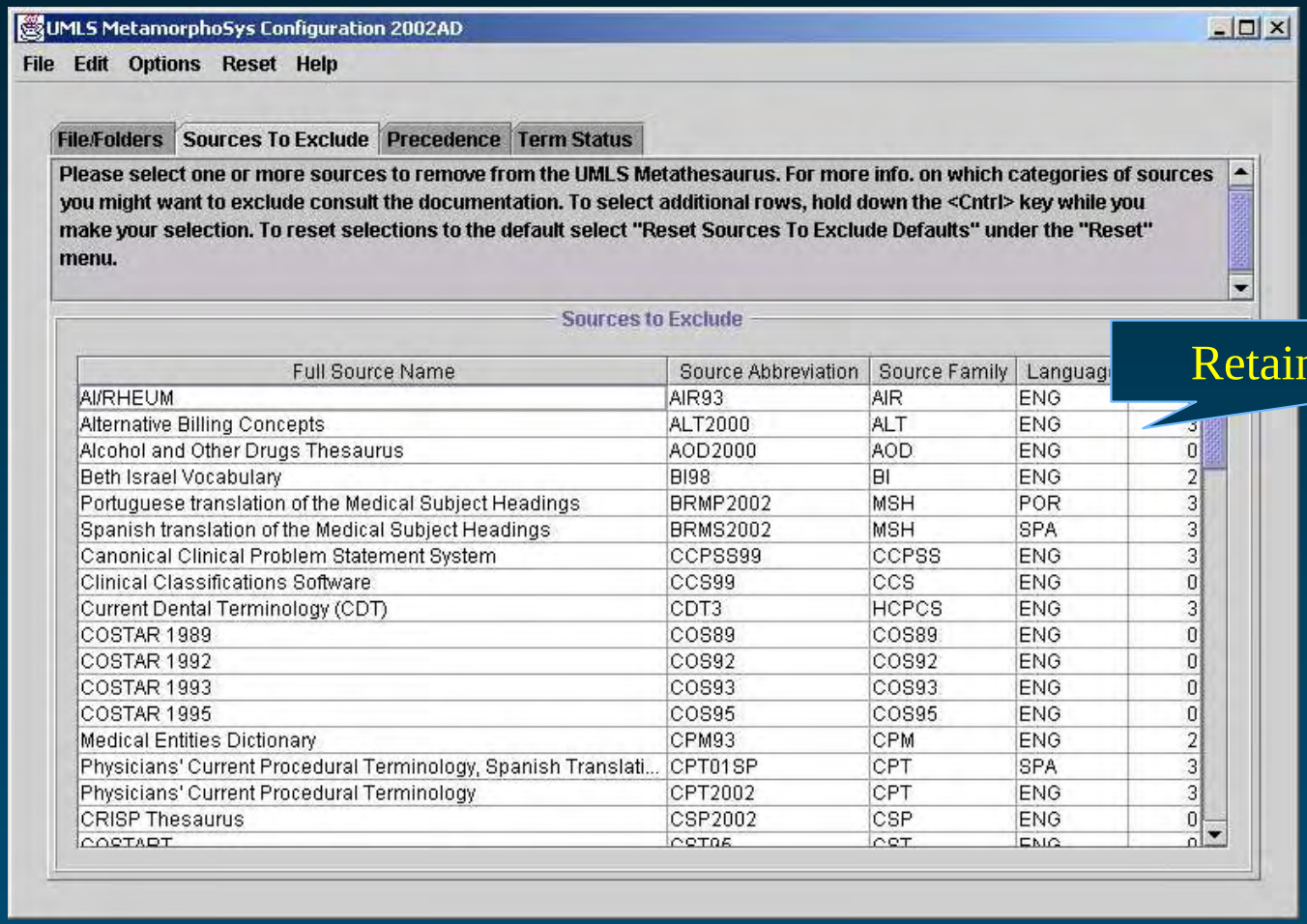
```
C0001403|ENG|P|L0001403|PF|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VC|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

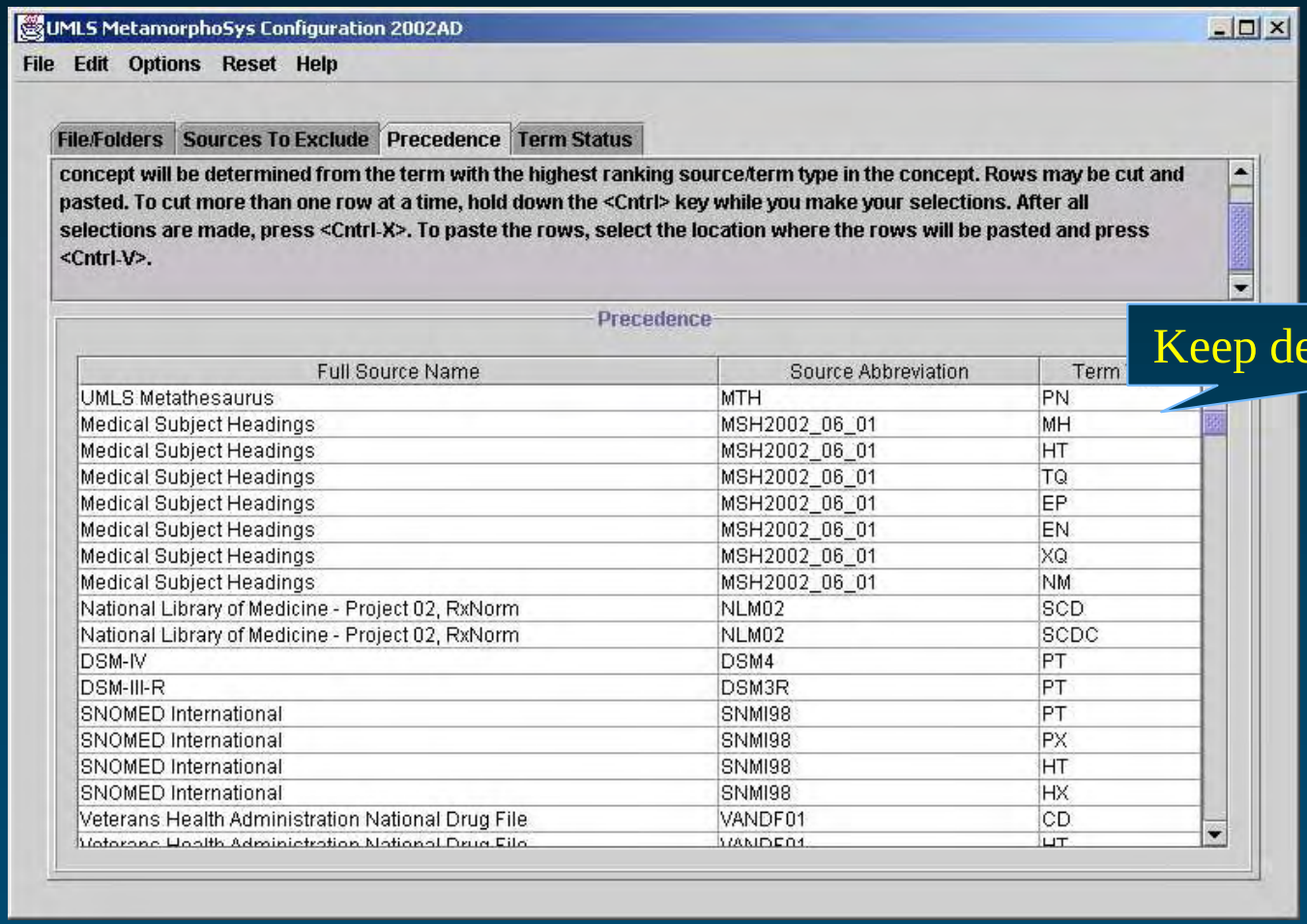
```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70020|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70020|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70020|3|
```

TS, STT and LRL get recomputed

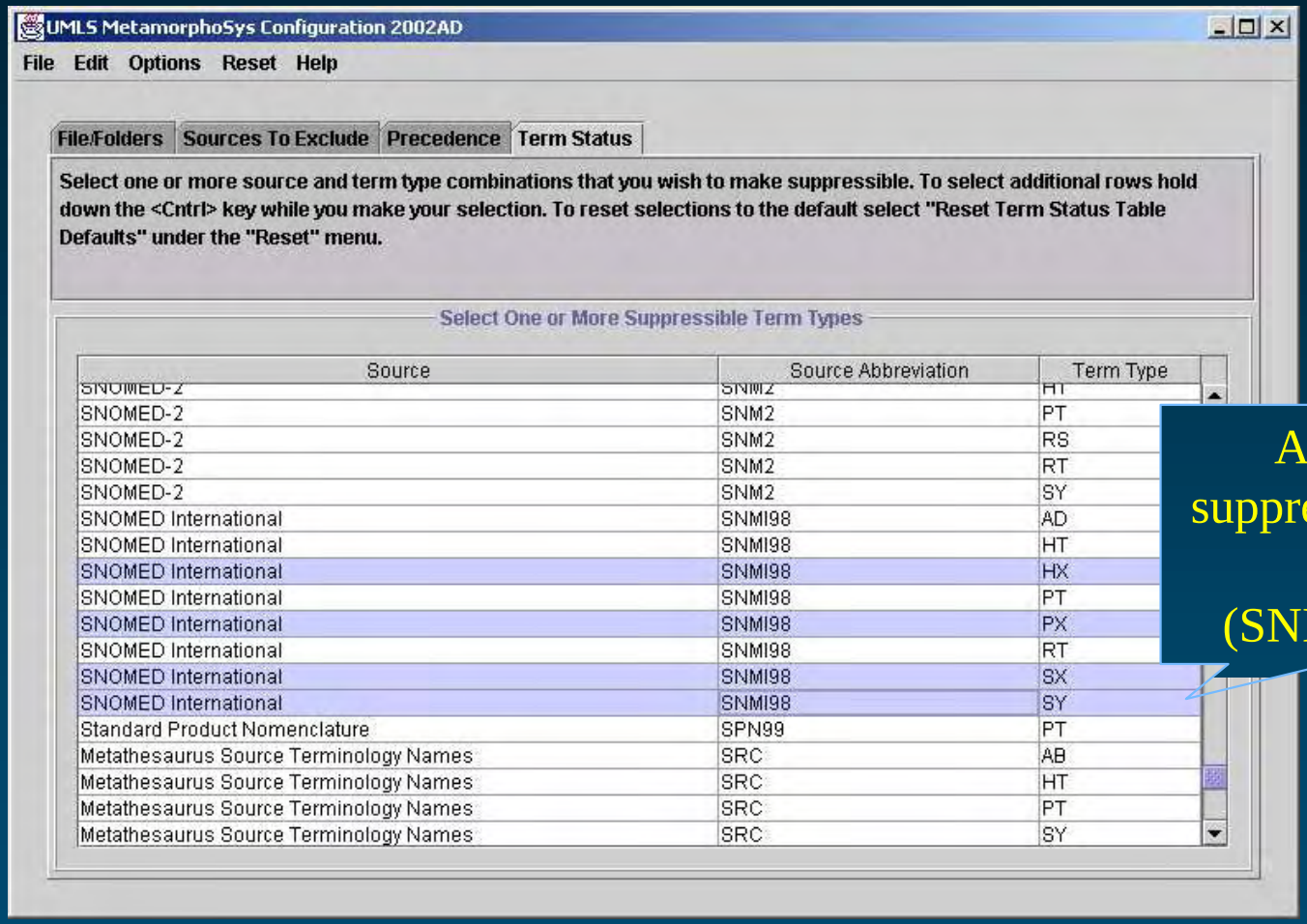
Adding to default suppressibility



Adding to default suppressibility



Adding to default suppressibility



Add new suppressible term type (SNMI98/SY)

Adding to default suppressibility

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|S|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70620|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70620|3|
```

TS goes from “S” to “s”

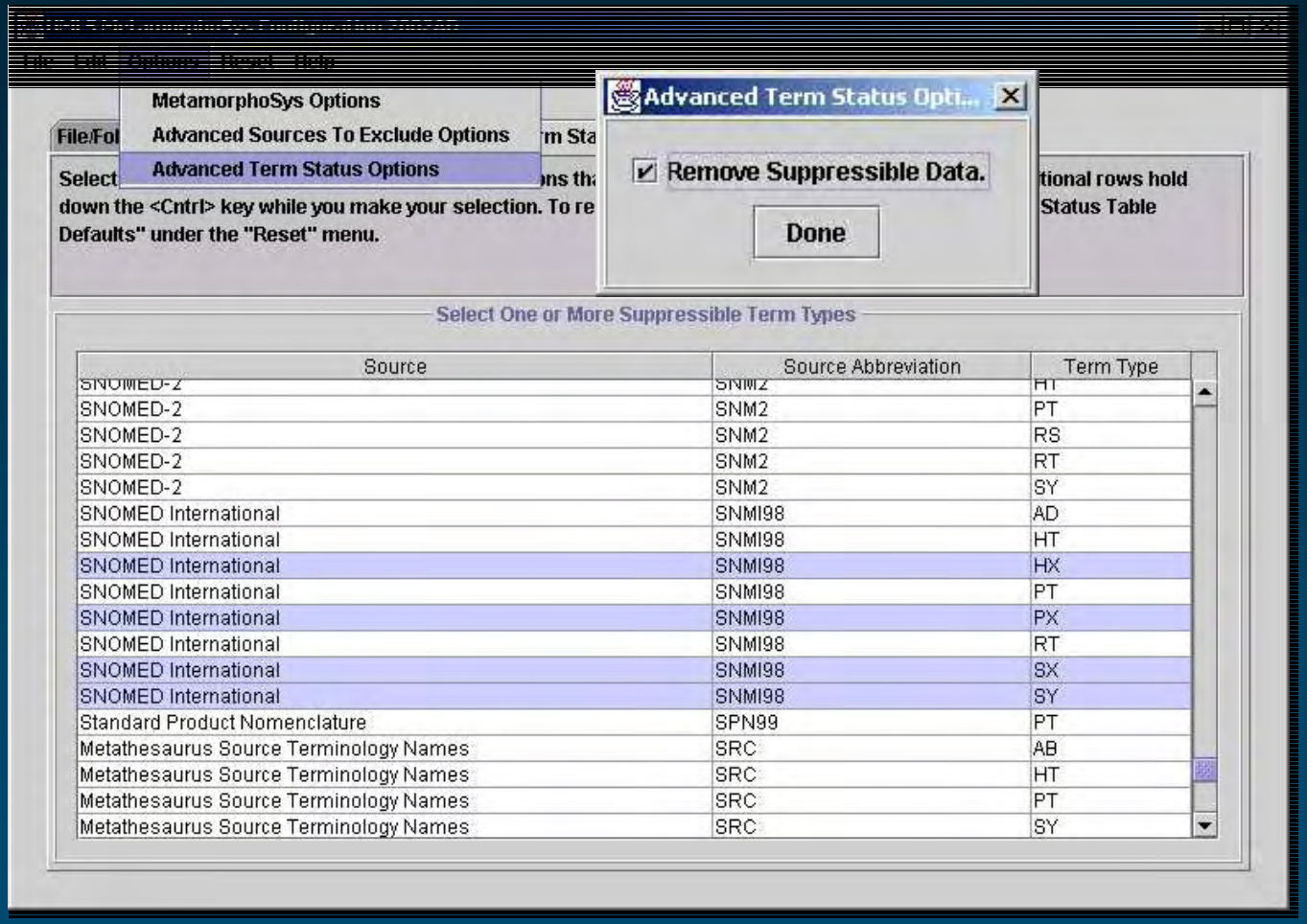
MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|s|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|s|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70620|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70620|3|
```

Removing suppressible data



Then, associated data are removed

MRCON

```
C0001403|ENG|P|L0001403|PF|S0010794|Addison's Disease|0|
C0001403|ENG|P|L0001403|VC|S0352253|ADDISON'S DISEASE|0|
C0001403|ENG|P|L0001403|VO|S0033587|Disease, Addison|0|
C0001403|ENG|P|L0001403|VO|S0469271|Addison's disease, NOS|3|
C0001403|ENG|P|L0367999|PF|S0469267|Addison melanoderma|3|
C0001403|ENG|S|L0373744|PF|S0471237|Asthenia pigmentosa|3|
```

MRSO

```
C0001403|L0001403|S0010794|MSH|MH|D000224|0|
C0001403|L0001403|S0352253|CST|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO|IT|0410|2|
C0001403|L0001403|S0033587|MSH|PM|D000224|0|
C0001403|L0001403|S0469271|SNMI|PT|DB-70620|3|
C0001403|L0367999|S0469267|SNMI|SY|DB-70620|3|
C0001403|L0373744|S0471237|SNMI|SY|DB-70620|3|
```

MetamorphoSys and MRCUI

- MRCUI has a row for every ‘dead’ CUI
- Provides a map or pointer to a ‘live’ CUI
- Map can be SY or a close relationship

CUI1	VER	CREL	CUI2	MAPIN
C0079158	1997AA	SY	C0009081	Y
C0079138	2001AA	RO	C0037440	Y

- Mapping work is ongoing

MetamorphoSys and MRCUI (contd.)

- MetamorphoSys preserves all MRCUI rows
- If CUI2 is not in subset
  - Changes MAPIN to ‘N’
  - Adds a row for CUI2 with CREL=SUBX

CUI1	VER	CREL	CUI2	MAPIN
C0079158	1997AA	SY	C0009081	Y
C0079138	2001AA	RO	C0037440	N
C0037440	2002AD	SUBX		N

MetamorphoSys configuration

- Program maintains the configuration as Java properties file
- Do not edit this file directly!
- Can be saved for future runs
  - Default (*mmsys.prop.default*) should not be deleted
- Configuration is generic
  - Can be ported across versions of UMLS
  - Uses versionless SAB
- Settings for all filters can be saved

General comments on MetamorphoSys

- Configured to run with a specific release from its install directory – its use with other releases will cause unpredictable results
- Propagates string-level suppressibility created and maintained by editors
- Writes a log file (mmsys.log) in the subset directory that contains information about how that subset was generated
- Can be run iteratively – order matters

Custom filters

- Coded in the Java language
- Implement Filter and extend AbstractFilter
- Have access to concept data and config data
- Additional data externally provided, if needed
- Have “undo” functionality
- Test filters come with MetamorphoSys
  - See $MMSHOME/ext folder

AbstractFilter Class

- GUI-related abstract behavior
- Provides default behavior for events when filter configuration changes
- Subclasses only have to call the fireDataChanged() method when configuration changes

Filter Interface

- Specifies how custom filter presents itself (GUI)
- Logic of the MetamorphoSys subsetting function
- Some methods:

getPanel()	Return GUI panel
getFilterProperties()	Properties for filter
hasDataChanged()	Filter data changed?
applyFilter(Cui cui)	Applies logic to concept

How to install a custom filter

- Develop, debug and test filter (Java)
- Compile with $MMSHOME/classes/mms.jar
  - Package name for core classes: gov.nih.nlm.mms
- Create a JAR file with filter and helper classes
- Copy your jar file to $MMSHOME/ext
- New filter should be available on next run
- Use File->Import to access the new filter

Examples of custom filters

- Test filters come with MetamorphoSys
  - See $MMSHOME/ext folder
- Used internally at NLM for license compliance and for other applications
- Check umlsinfo.nlm.nih.gov for more

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content — O. B., L. Roth, S. Srinivasan
    - Customize with MetamorphoSys
    - Advanced techniques — Olivier Bodenreider
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

Advanced customization techniques

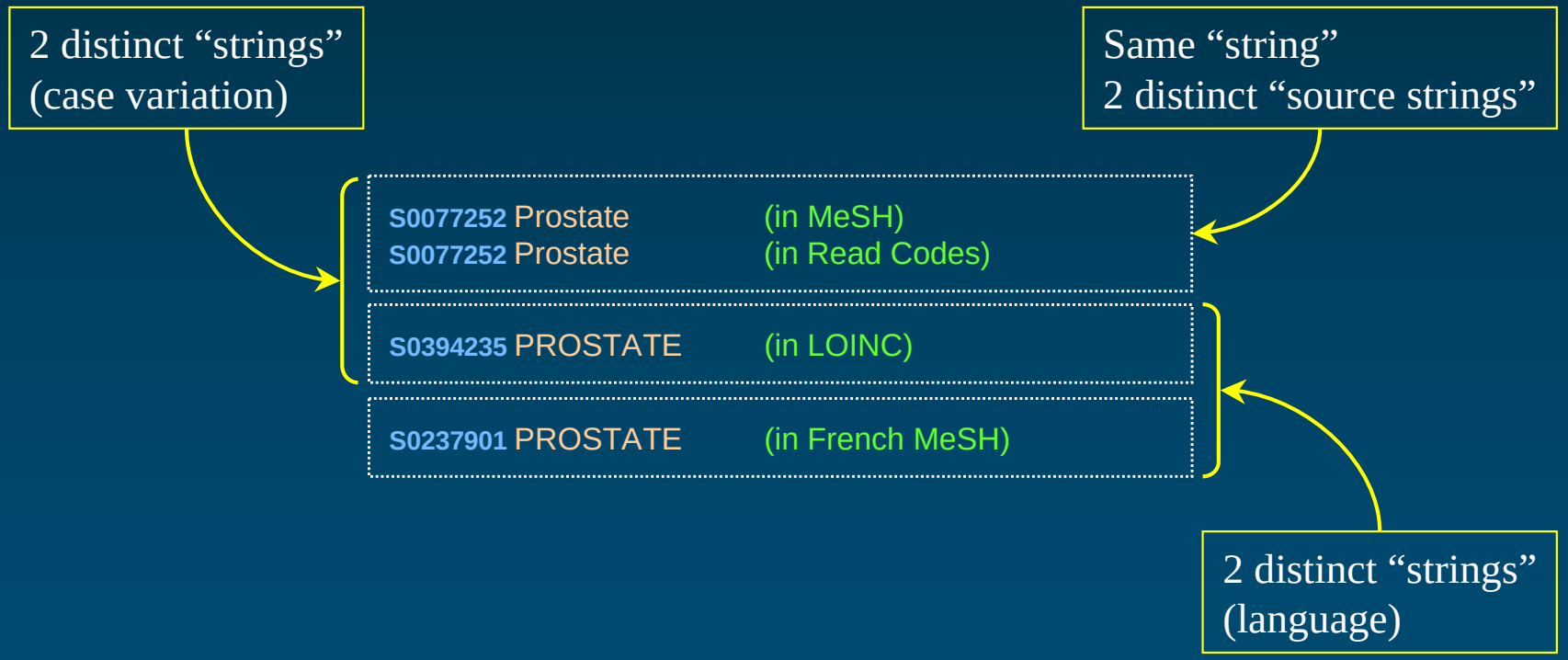
- Customize strings
- Customize synonyms
- Customize relationships
  - Semantic approach
  - Structural approach
  - Statistical approach

Advanced Techniques

Customize Strings

Background Strings

- Located in MRCON
- ~2.3 million “source strings”



Background String attributes

- Language
- Preferred name in a source
- Lexical variants (case, inflection, word order, …)
- Other variants
  - Underspecification marker (Other, NOS)
  - Classification-specific marker (NEC)

Background More string attributes

- Source
- Term type (= type of string in a given source)
- Code in a given source
- Source-specific attributes
  - MN: Position in the hierarchy (MeSH)
  - SIC: ICD-9-CM code mapped to (SNOMED)
  - LFR: French name for a LOINC term
  - ICN: ICD-9-CM coding information
  - […]

Background Implicit string attributes

- Number of (families of) source vocabularies providing the string
- Presence in a target corpus

Motivation

- Reduce volume
- Select useful strings for natural language processing
- Select target-specific strings
- Filter out
  - Source-specific strings (e.g., truncated strings)
  - Purpose-specific strings (e.g., classification-specific strings, inverted terms)

Methods

- Identify string properties
- Combine the properties in order to create filters

Methods Identify string properties (1)

- Properties based on morphology (identified through regular expressions)
  - /, / for inverted terms 238,000
  - /[0-9]/ for strings containing digits 376,000
  - /^other|not elsewhere classified|NEC|without mention/ for classification feature 28,000
  - […]
  - Number of words in the string

Methods Identify string properties (2)

- Properties based on UMLS features

  - Redundancy: Number of (families of) source vocabularies providing this string 95,000
  - Term type (MRSO/TTY)
    - Chemical names 318,000
    - Branded drug names or supplies 62,000
    - Abbreviations and truncated strings 126,000

    […]
- Properties based on a corpus
  - e.g., strings found in MEDLINE 144,000

Methods Combine properties

- Using logical operators (AND, OR, NOT)
- 2 approaches
  - *A priori* model of the strings in a given context
  - Classification techniques against a target
- Traditional sensitivity/specificity balance

- e.g.: select English strings
  - Excluding chemical names
  - Excluding inverted terms
  - Found in more than one source vocabulary

Example of use

- Select UMLS strings useful for natural language processing

McCray A.T, Bodenreider O., Malley, J.D., Browne A.C.
Evaluating UMLS strings for natural language processing.
Proc AMIA Fall Symp. 2001:448-452

STR	NB WORDS	ALLCAPS ALWAYS	ALL CLSP	ALL UNSP	ANY PARENTHETICAL	CT COMMA SPACE	CT NON ALPHANUM	CT NUMBERS	CT PUNCTUATION	CT SYMBOLS	MI AND OR	NB SOURCES	SUPPRESSIBLE ALWAYS	TTY CHEMICAL	TTY LOINC	TTY METADATA	TTY PHRASE	TTY PRESCRIPTION	TTY SHORT FORM
ADDISON DISEASE ✔	2											3							
Addison melanoderma	2											1							
Addisons Disease	2											2							
Addison's disease ✔	2											8							
Addison's disease NOS	3			x								1							
Addison's disease, NOS	3			x		x	x					1							
ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)	4	x			x		x					1							
ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE	4	x				x	x					1							
Asthenia pigmentosa	2											1							
Bronzed disease	2											1							
DISEASE ADDISON'S	2	x										1							
Disease, Addison ✔	2					x	x					1							
Disease, Addisons	2					x	x					1							
Disease, Addison's ✔	2					x	x					1							
Disease;Addisons	2						x		x			1							
Melasma addisonii	2											1							
Primary adrenal deficiency	3											1							
Primary adrenocortical insuff	3											1	x						x
Primary adrenocortical insufficiency ✔	3											2							

Discussion

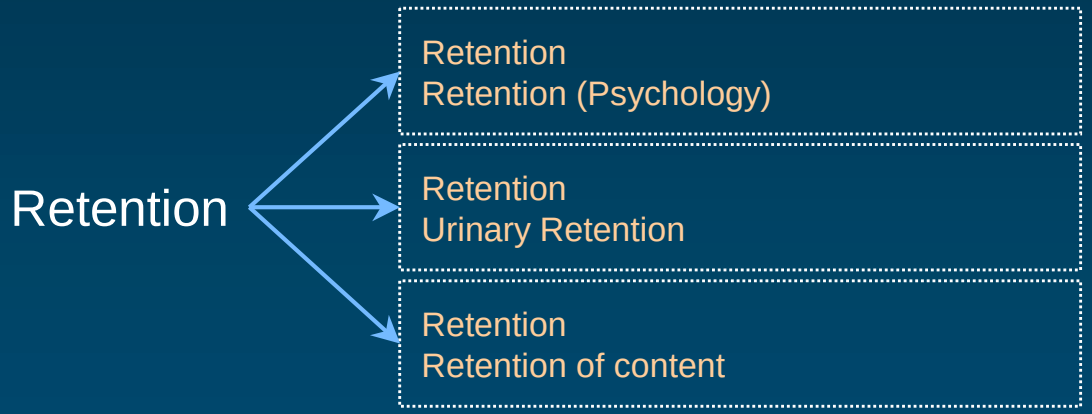
- Restricting to a given language is easier done through sources
- Filtering out strings may result in removing concepts
- Term status is relative to the preferred name, but does not identify the canonical form

Advanced Techniques

Customize Synonyms

Background

- Metathesaurus concepts are clusters of synonymous terms
- Polysemous terms may appear in more than one concept



Background

- Metathesaurus synonymy is not necessarily linguistic synonymy
  - Not fully specified terms
  - Granularity issues
  - Generic / prototypical

Prostate ✔ (in MeSH)
Prostatic gland

prostate ✘ (in COSTAR)
Prostatic Diseases

Prostate ✘ (in ICD-10)
Benign neoplasm of prostate

Posttransfusion hepatitis
Posttransfusion viral hepatitis

Asplenia
Congenital asplenia

Background

Myocardial Infarction

- Additionally, Metathesaurus synonyms include
  - Translated terms

    Infarctus du myocarde (French)
    Myocardinfarkt (German)
  - Lexical variants

    Myocardial Infarctions (plural)
    Infarction, Myocardial (permutation)
    Infarctions (Myocardial) (parentheses)
  - Acronyms

    MI
    MI - Myocardial infarction
  - Various kinds of terms (truncated, obsolete, …) as provided by source vocabularies

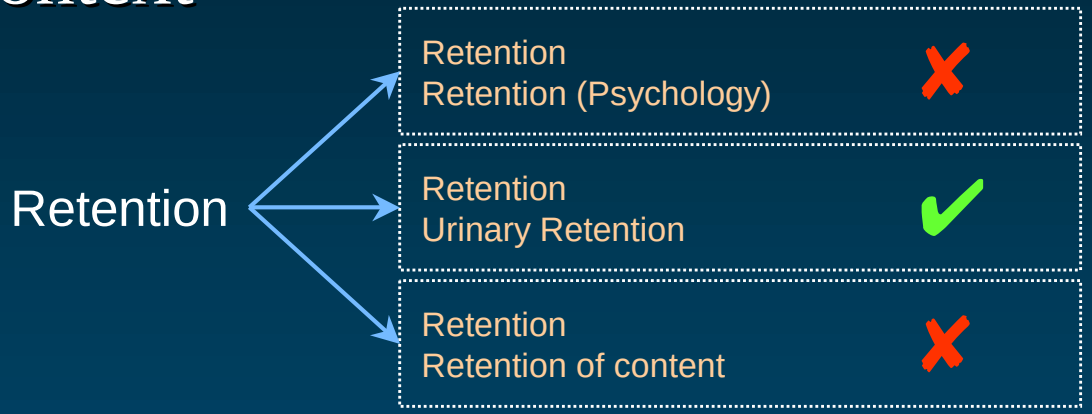
Background

- Some vocabularies implement their own notion of "synonymy"

depression and suicide	(preferred term)
suicide and depression	(synonym)
depression	(synonym)
suicide	(synonym)
cancer patients and suicide and depression	(synonym)
cancer patients and depression and suicide	(synonym)

Motivation

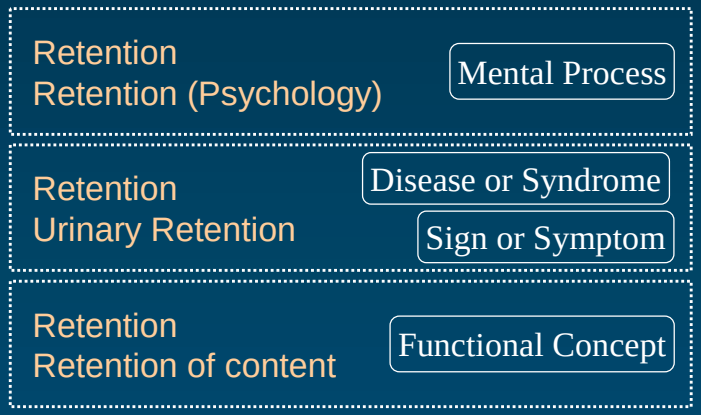
- Associate the right meaning with a string in a given context



- From the several strings associated with a meaning, select the most appropriate ones in a given context

Methods Associate the right meaning

- Use the “suppressible synonym” flag
  - Identifies not fully specified names
  - A fully specified name usually exists among the synonyms (sometimes created by NLM)
- Restrict the domain
  - In order to limit polysemy
  - Implies
    - A priori knowledge
    - Interaction with users



- Word sense disambiguation research area

Methods Most appropriate strings

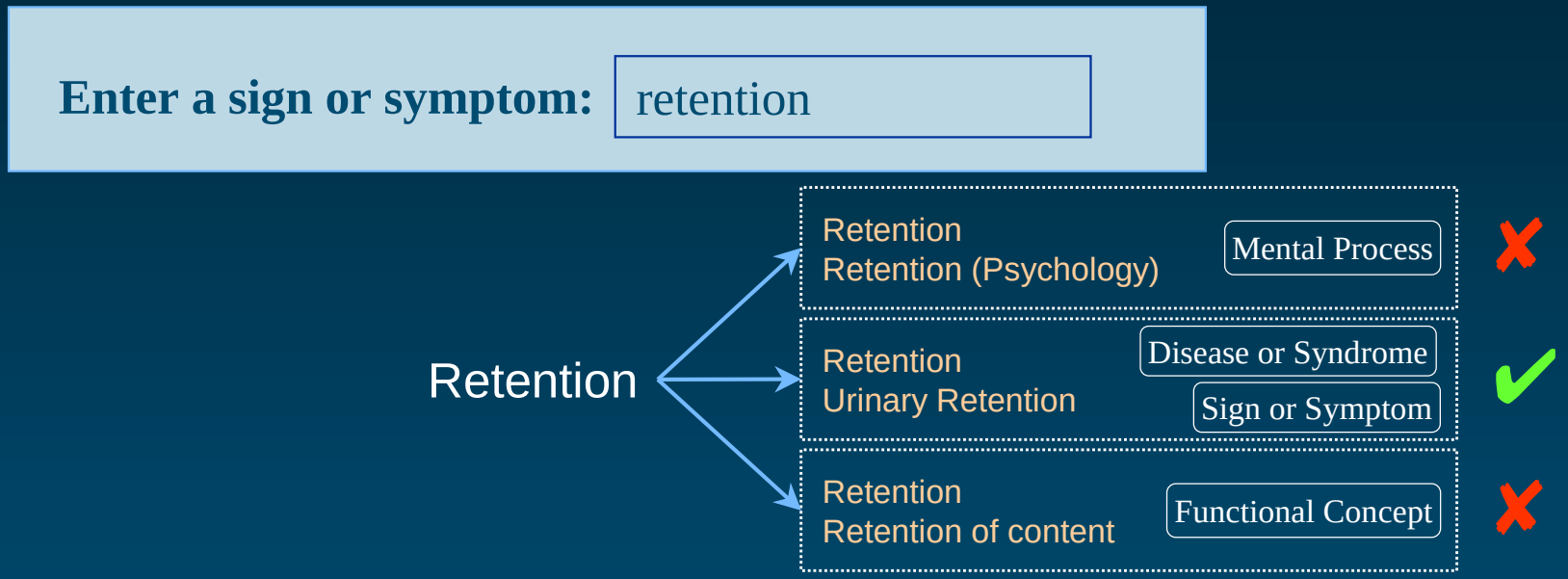
- Recognize and filter out lexical variants
  - Canonical form
  - Normalization
- Filter against a corpus
  - To find the most common form in your target

MEDLINE 1999

Fallen arch	
Fallen arches	
Flat foot NOS	
Flat Feet	✔
Flatfeet	✔
Flatfoot	✔
Foot, Flat	
Low medial arch of foot	
Pes Planus	✔
Pes planovalgus	✔
Pes valgus	✔

Example of use

- Disambiguate according to the context



- Filter redundant lexical variants from a list of terms in a Metathesaurus concept

Discussion

- Word sense disambiguation
  - Never trivial
  - Still open research area (linguistics)
  - Often involves statistical analysis of the context
- The Metathesaurus partially addresses the issue of not fully specified terms

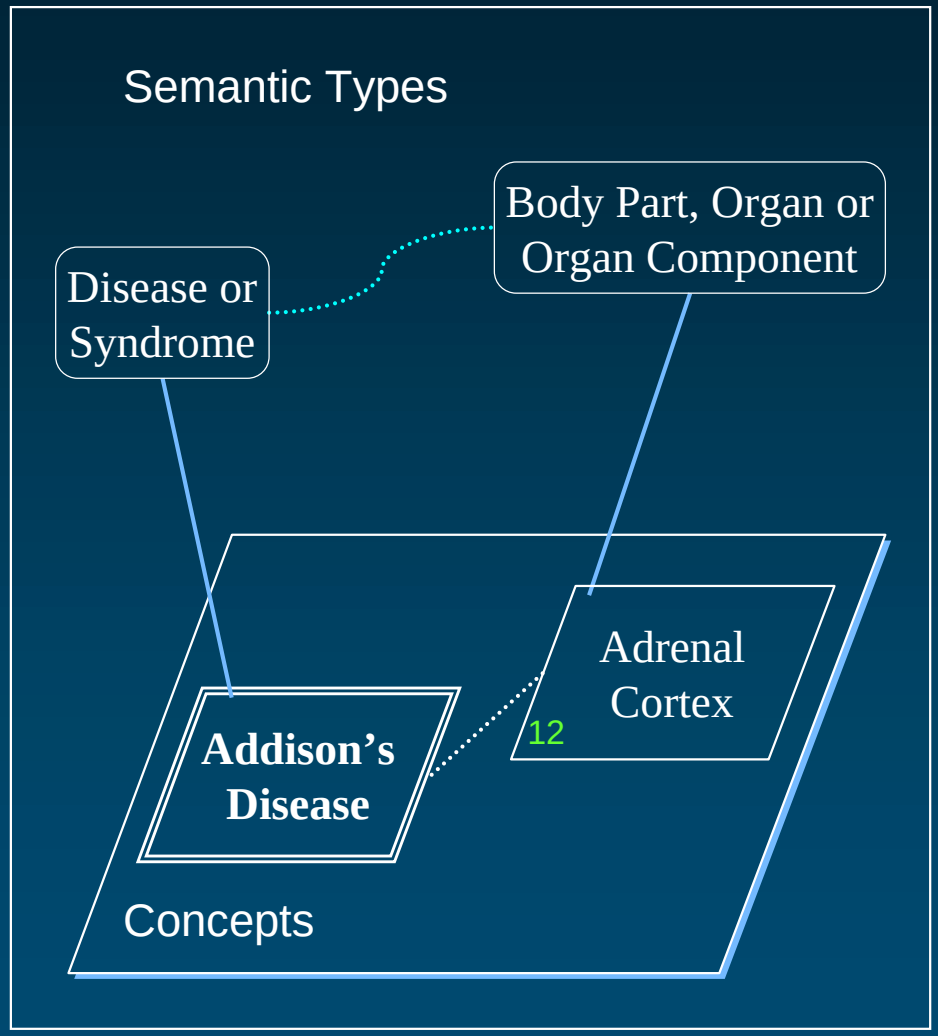
Advanced Techniques

Customize Relationships

1. Semantic Approach

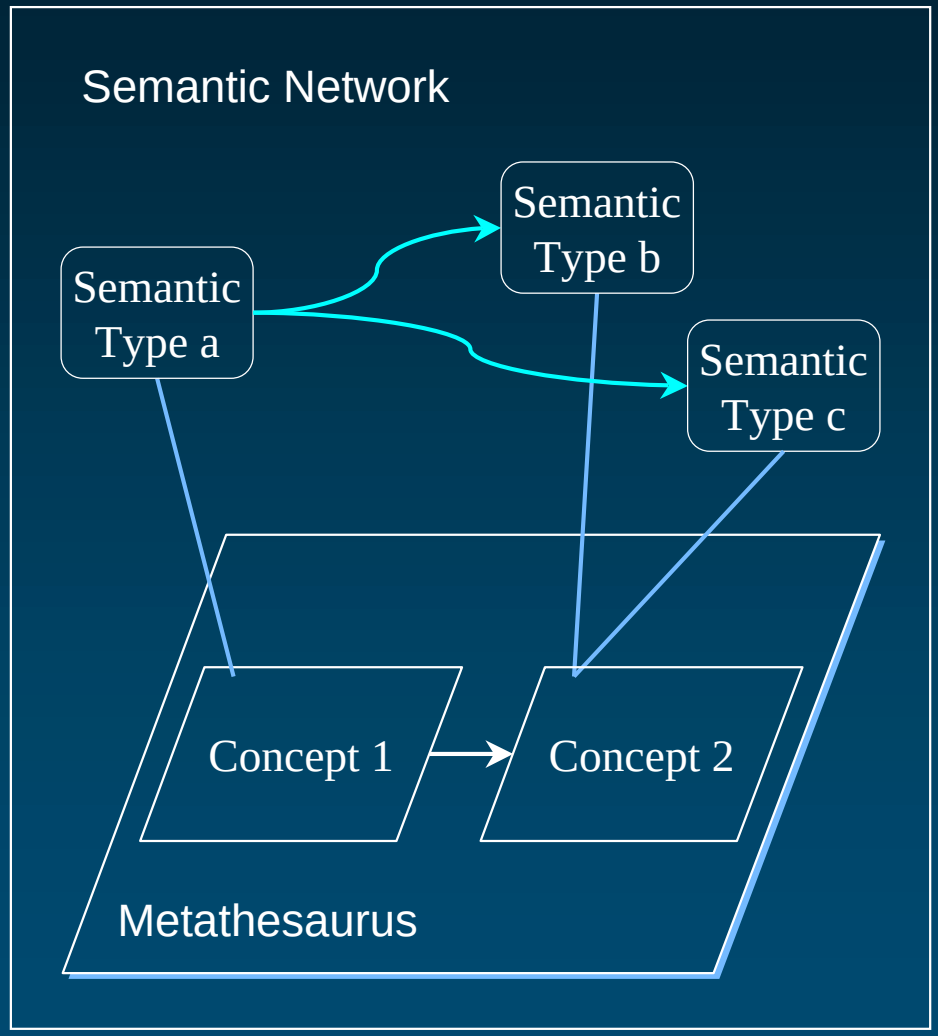
Background UMLS structure (nodes)

- Two-level structure
  - Semantic Network (135 semantic types)
  - Metathesaurus (870,000 concepts)



Background UMLS structure (links)

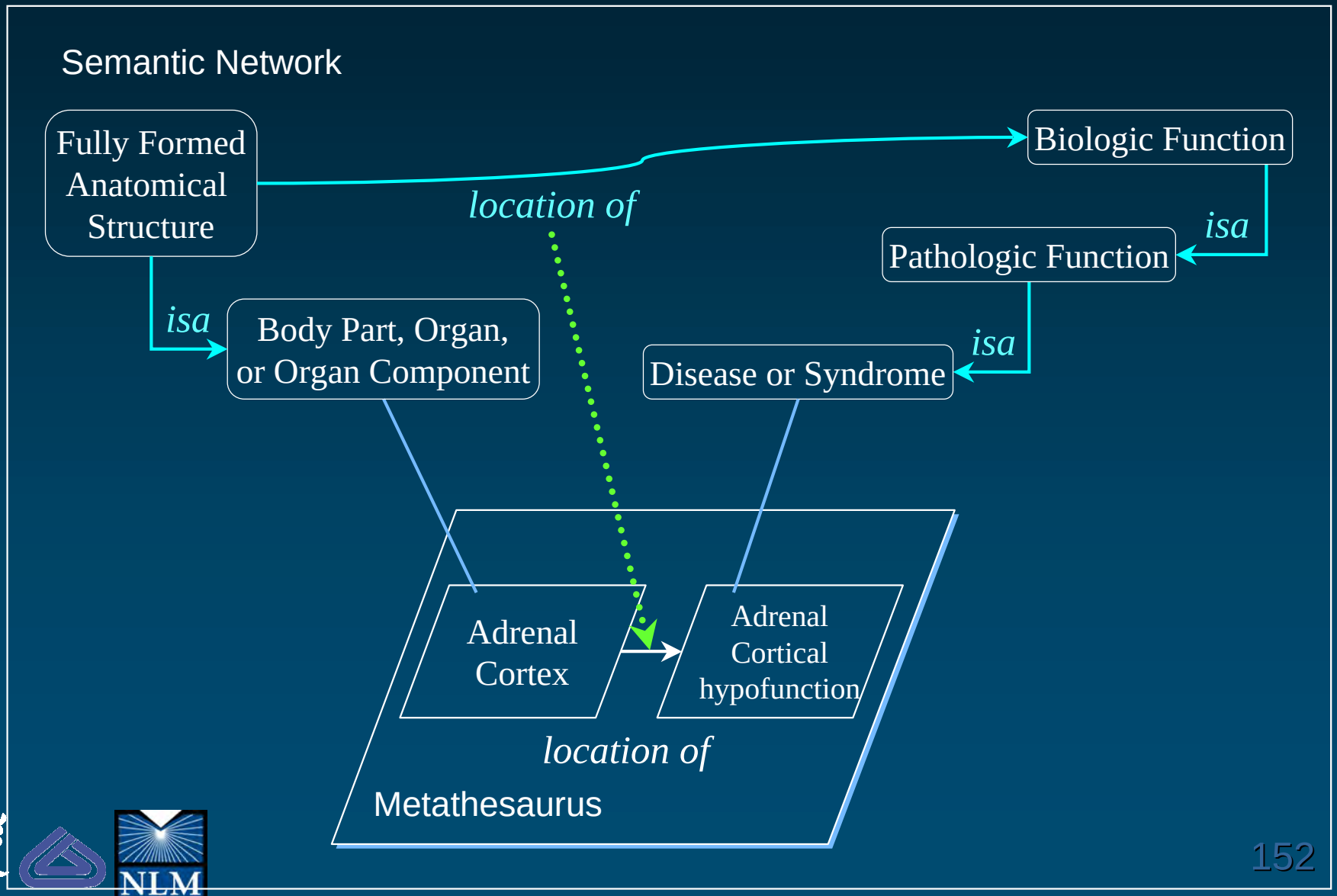
- Semantic network relationships
- Categorization
- Interconcept relationships



Background UMLS structure (links)

- Semantic network relationships
  - Hierarchical or associative
  - General (definitional) knowledge
  - May or may not hold at the concept level
- Categorization
  - Links each concept to (at least) one broad category
  - Either *isa* or *is an instance of* relationships
- Interconcept relationships
  - Hierarchical, associative or statistical
  - Factual knowledge

Relationships can inherit semantics



Motivation

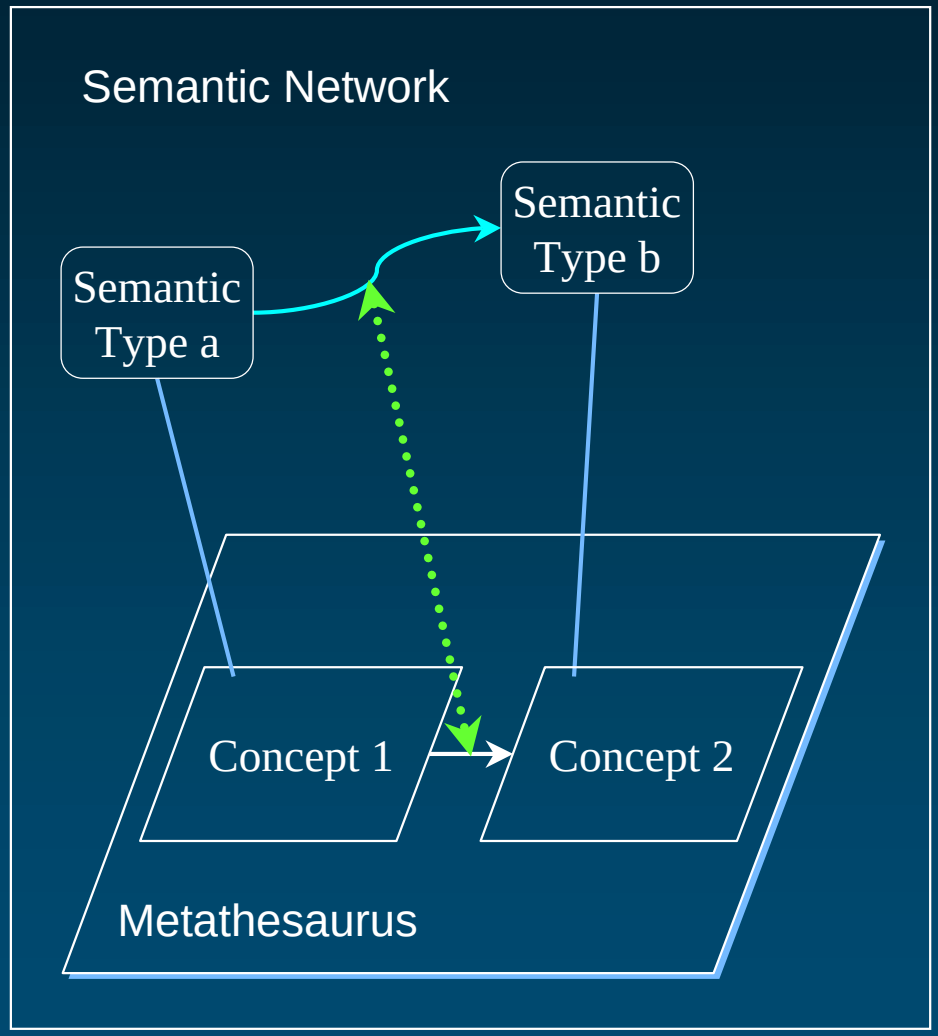
- Check the consistency of the two levels
  - Semantic network
  - Metathesaurus
- Check the consistency between
  - Semantic network relationships
  - Interconcept relationships
- Discrepancies may indicate
  - Inaccurate relationship
  - Inaccurate categorization

Motivation

- More generally
  - The Semantic Network represents some kind of upper-level ontology of the biomedical domain
  - The organization of Metathesaurus concepts
    - is *expected* to be compatible with the upper level
    - is *required* to be compatible with the upper level if reasoning is to be supported

Methods

- For each pair of related concepts
  - Get their semantic types
  - Get all the “expanded” semantic network relationships between the two semantic types (transitive closure)
  - Compare
    - Interconcept relationship
    - Sem. Net. relationships



Methods

- Possible outcome
  - ICR = SNR → validate
  - ICR descendant of SNR → validate
  - ICR and SNR not compatible → reject
  - Unspecified ICR (no RELA) → infer/reject
  - ICR not in the Semantic Network

ICR: Inter-concept relationship
SNR: Semantic Network relationship

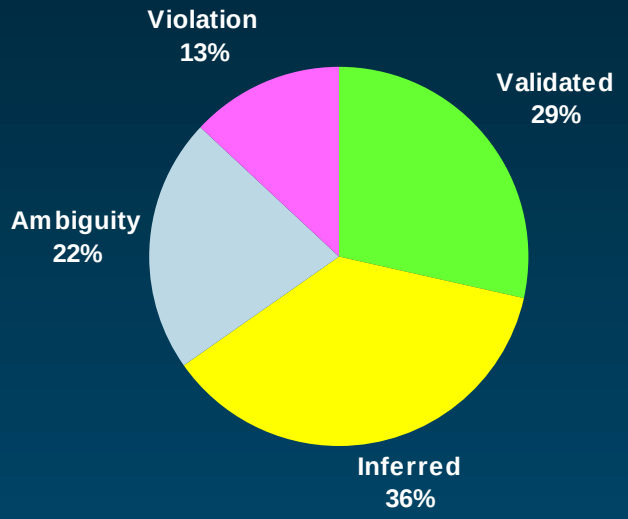
Example of use

- Validate, infer or reject interconcept relationships by comparison to the relationships defined between the semantic types assigned to the concepts

McCray A.T, Bodenreider O.
A conceptual framework for the biomedical domain.
In: Green R, Bean CA, Myaeng SH, editors. *The semantics of relationships: an interdisciplinary perspective*.
Boston: Kluwer Academic Publishers; 2002. p. 181-198.

Example of use Results

- 6894 interconcept relationships
  - among the 3764 concepts in the semantic neighborhood of “Heart”



Discussion

- Interconcept relationships recorded in the Metathesaurus are not censored
- The Semantic Network
  - Provides semantic constraints
  - Can be used to select Metathesaurus relationships that are “semantically sound”
- Limitations
  - Ambiguous SN relationships
  - Unspecified Metathesaurus relationships
  - Need for some manual review

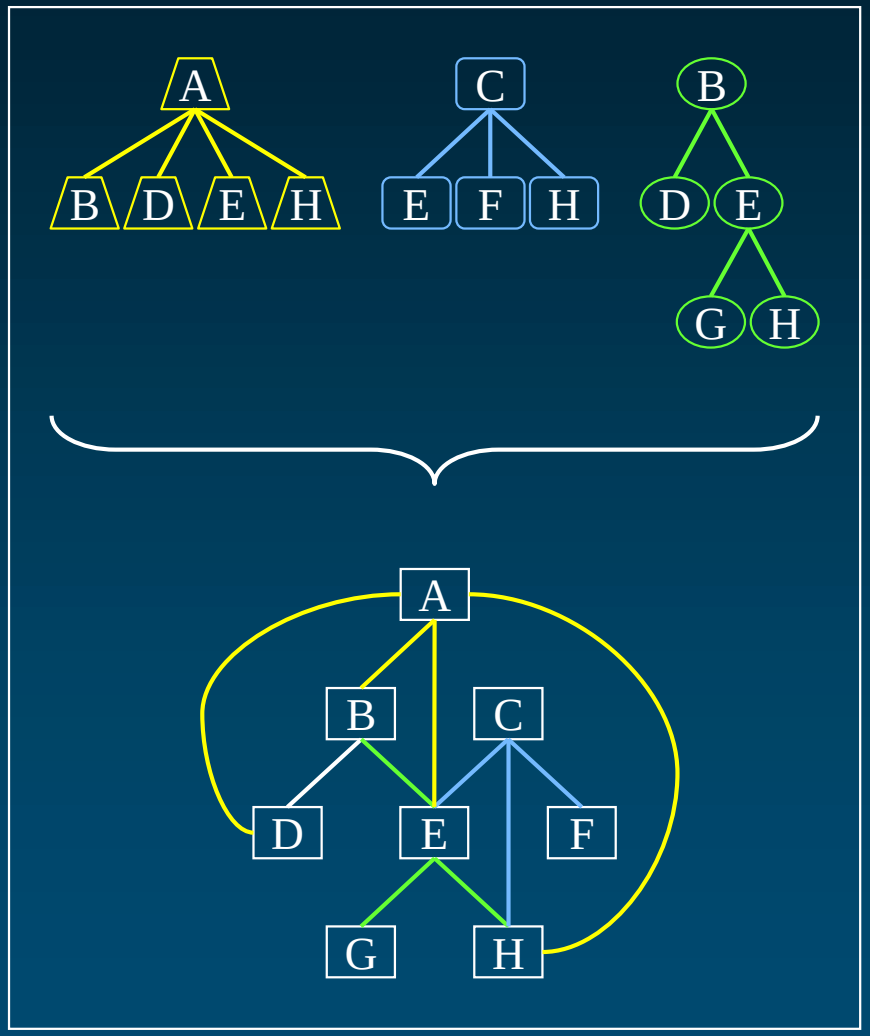
Advanced Techniques

Customize Relationships

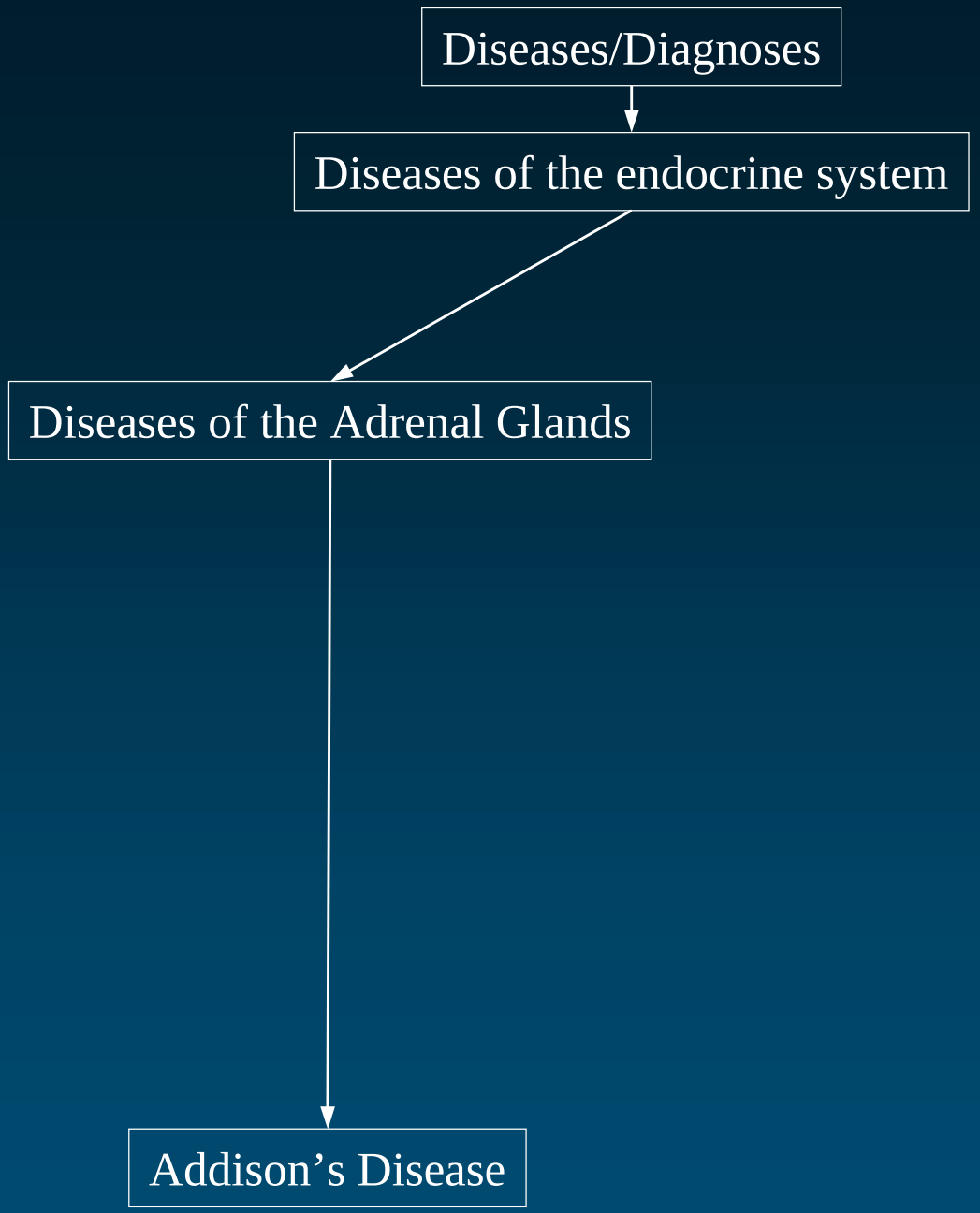
❷ Structural Approach

Background

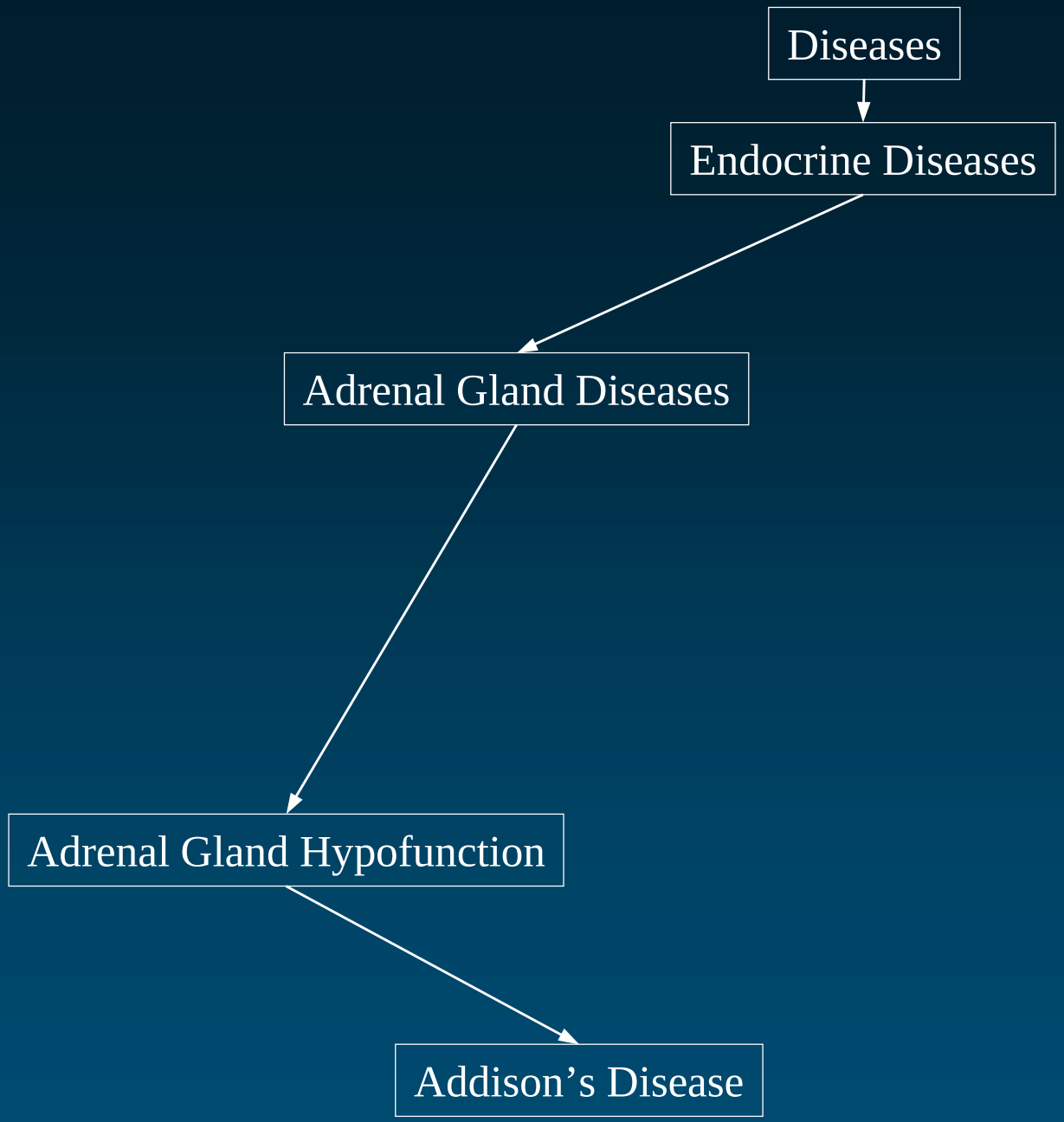
- The Metathesaurus is often seen as a bunch of trees
- Trees can be combined into a (directed) graph
- Hierarchies (esp. taxonomies) are based on partial ordering relationship
- Hierarchical relationships in the Metathesaurus are expected to result in a Directed Acyclic Graph (DAG)



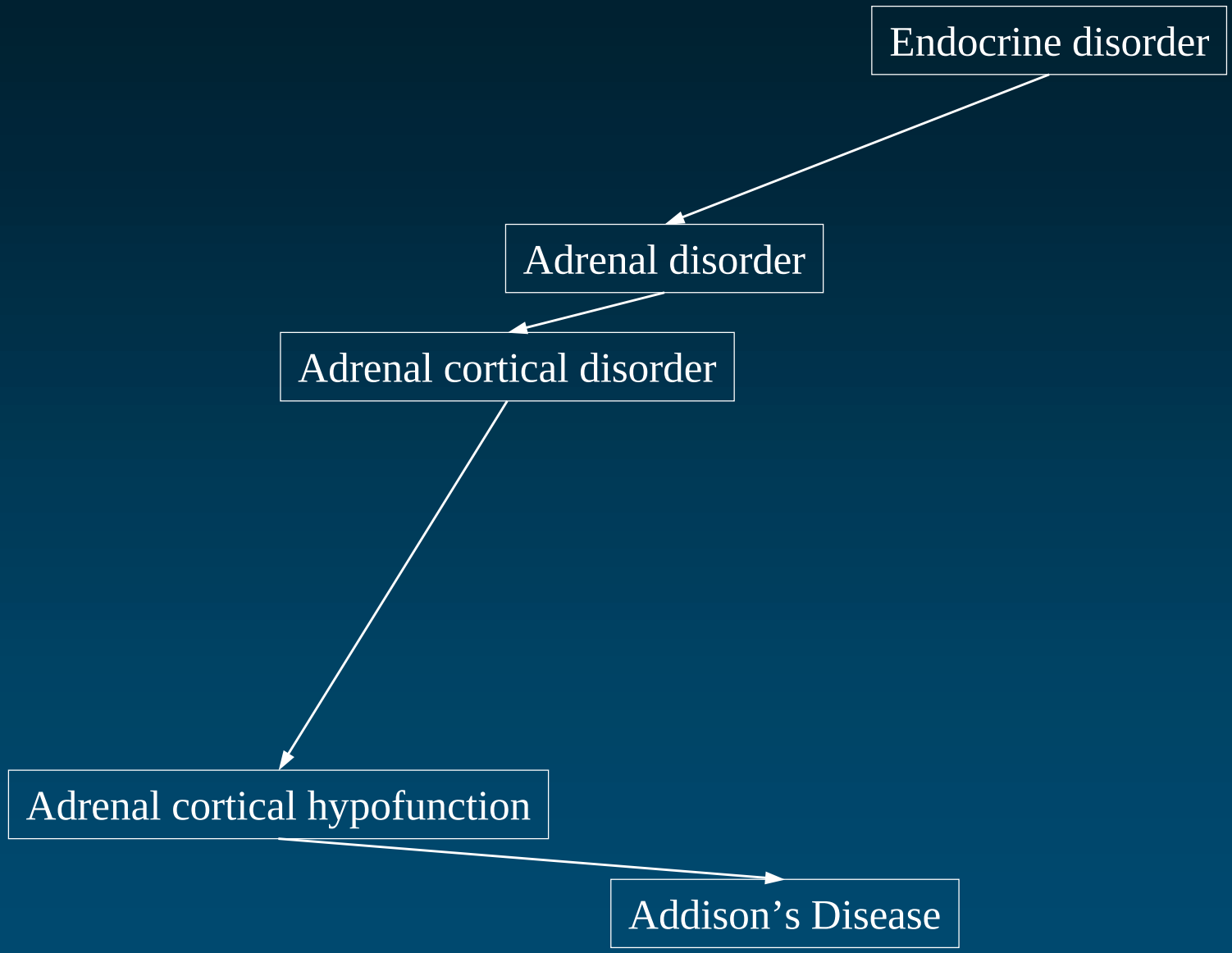
SNOMED International *tree*



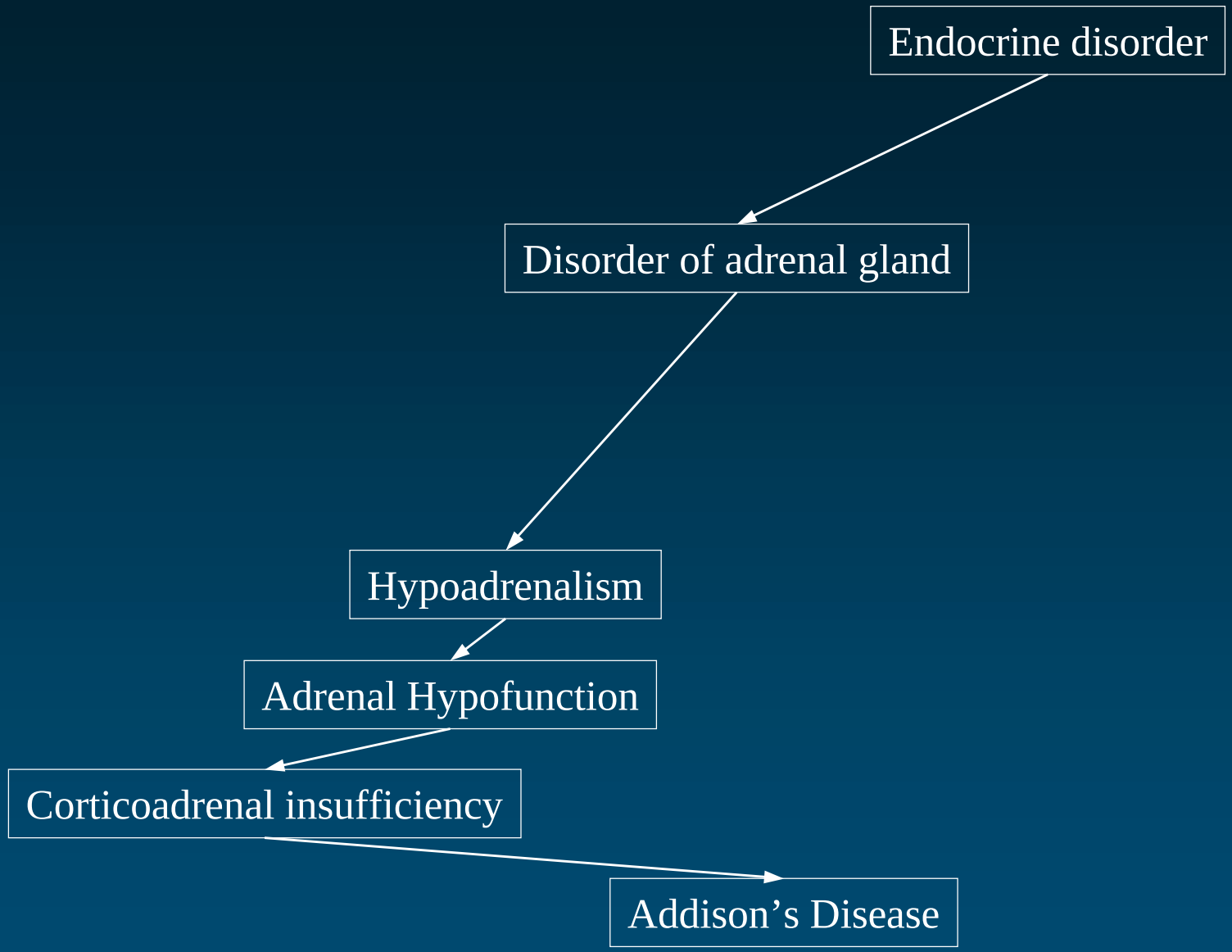
MeSH *tree*


Diseases
Endocrine Diseases
Adrenal Gland Diseases
Adrenal Gland Hypofunction
Addison’s Disease

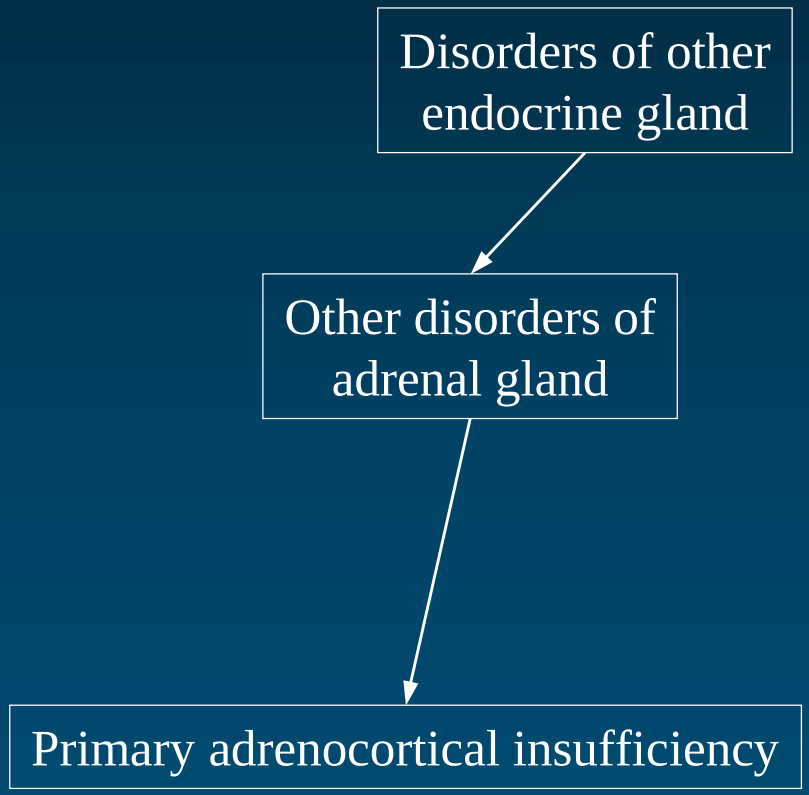
AOD *tree*


Endocrine disorder
Adrenal disorder
Adrenal cortical disorder
Adrenal cortical hypofunction
Addison’s Disease

Read Codes *tree*



ICD-10 *tree*


Disorders of other endocrine gland
Other disorders of adrenal gland
Primary adrenocortical insufficiency

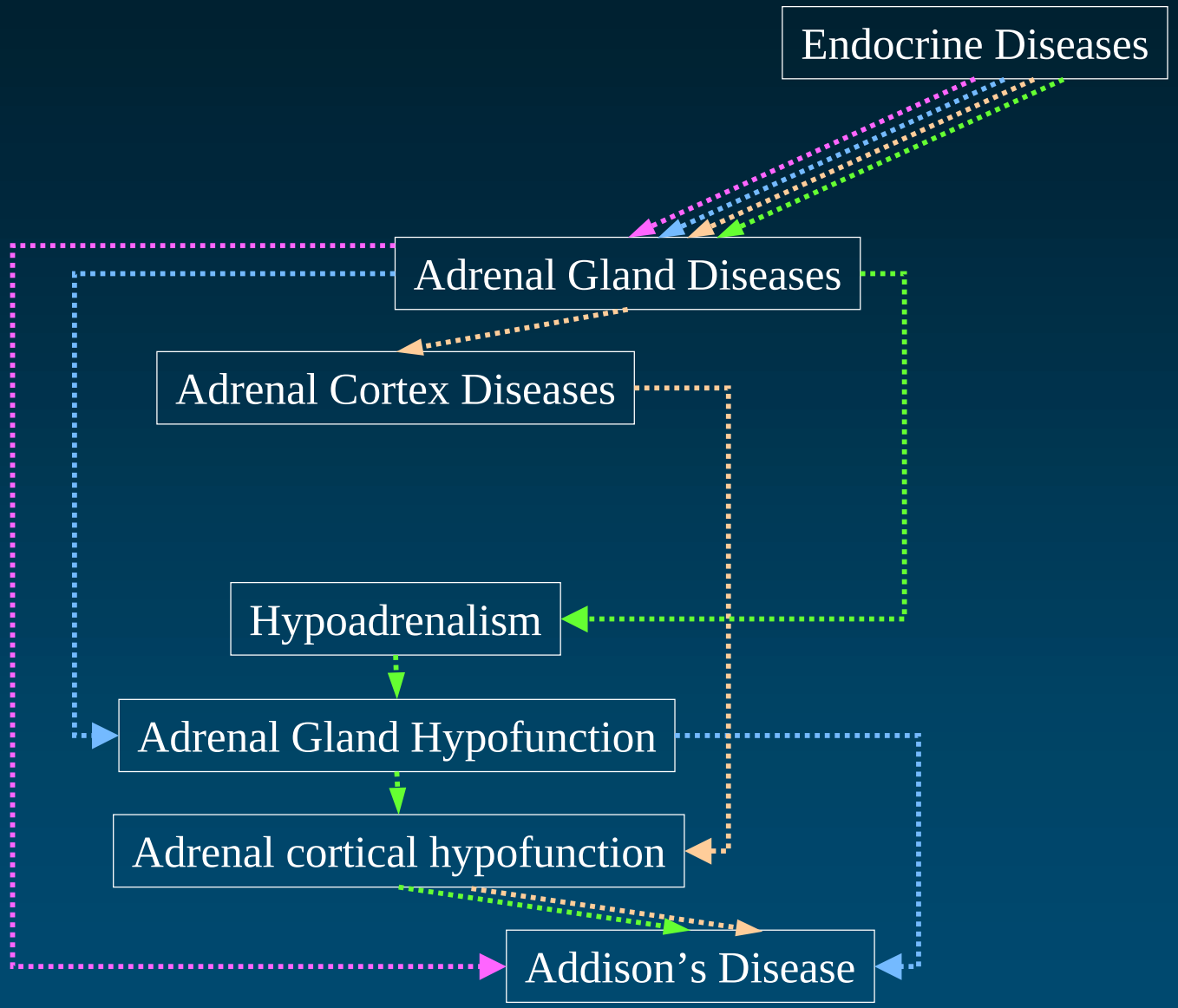
Metathesaurus *graph*

SNOMED

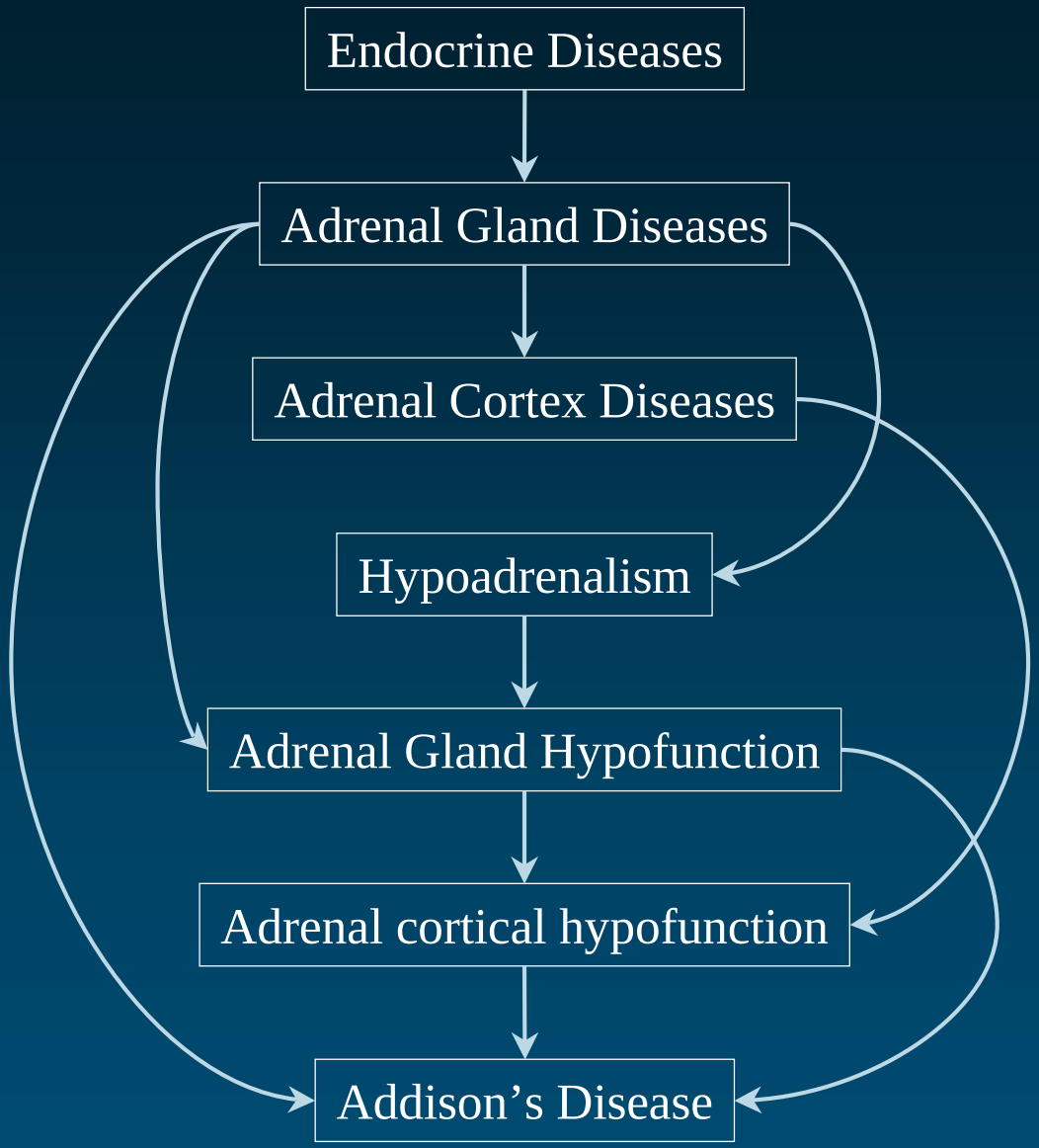
MeSH

AOD

Read Codes

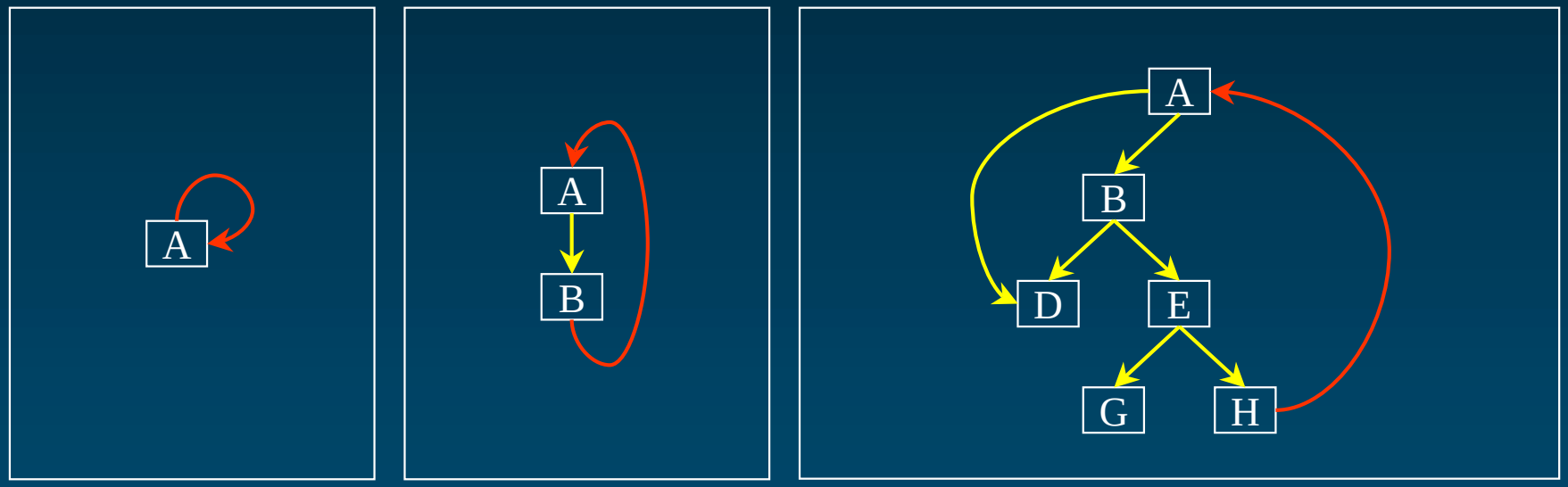


Metathesaurus ***graph***



Circular hierarchical relationships

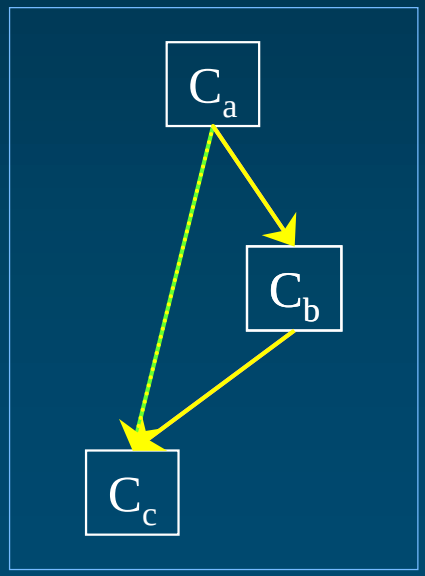
"back edge" from a child concept to a parent concept



Reflexive Direct Indirect

Motivation

- Circular hierarchical relationships are indicative of potential semantic issues
  - Wrong relationships
  - Non-hierarchical “hierarchical” relationships
- Some graph operations cannot be performed unless graph is acyclic
  - Transitive reduction



Methods

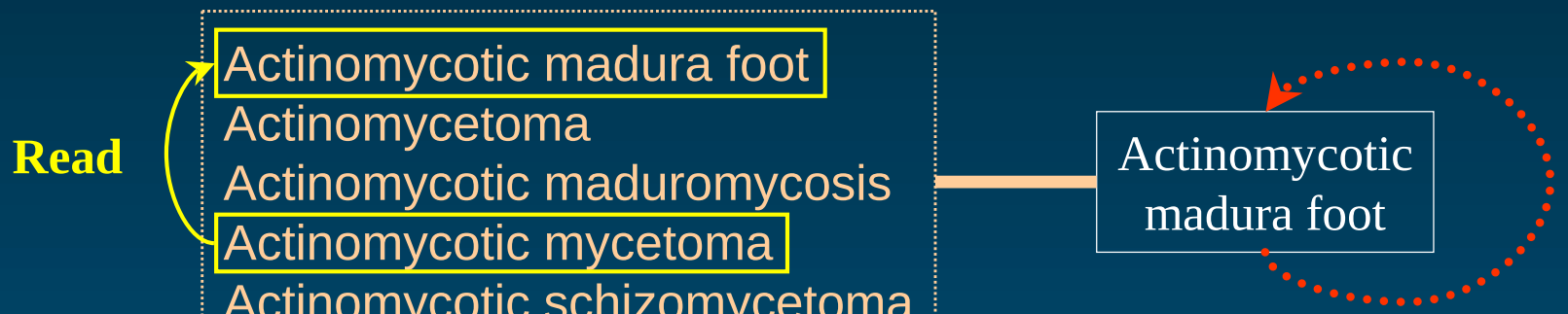
- Identify cycles
  - Reflexive: CUI1 = CUI2
  - Direct: CUI1|PAR/RB|CUI2 and CUI1|CHD/RN|CUI2
  - Indirect: graph analysis (depth-first search)
- Break cycles
  - Reflexive: remove all (or ignore)
  - Direct: remove (at least) one of the two links
    - Contexts (original trees), redundancy
  - Indirect: remove (at least) one link
    - Manual review

Example of use

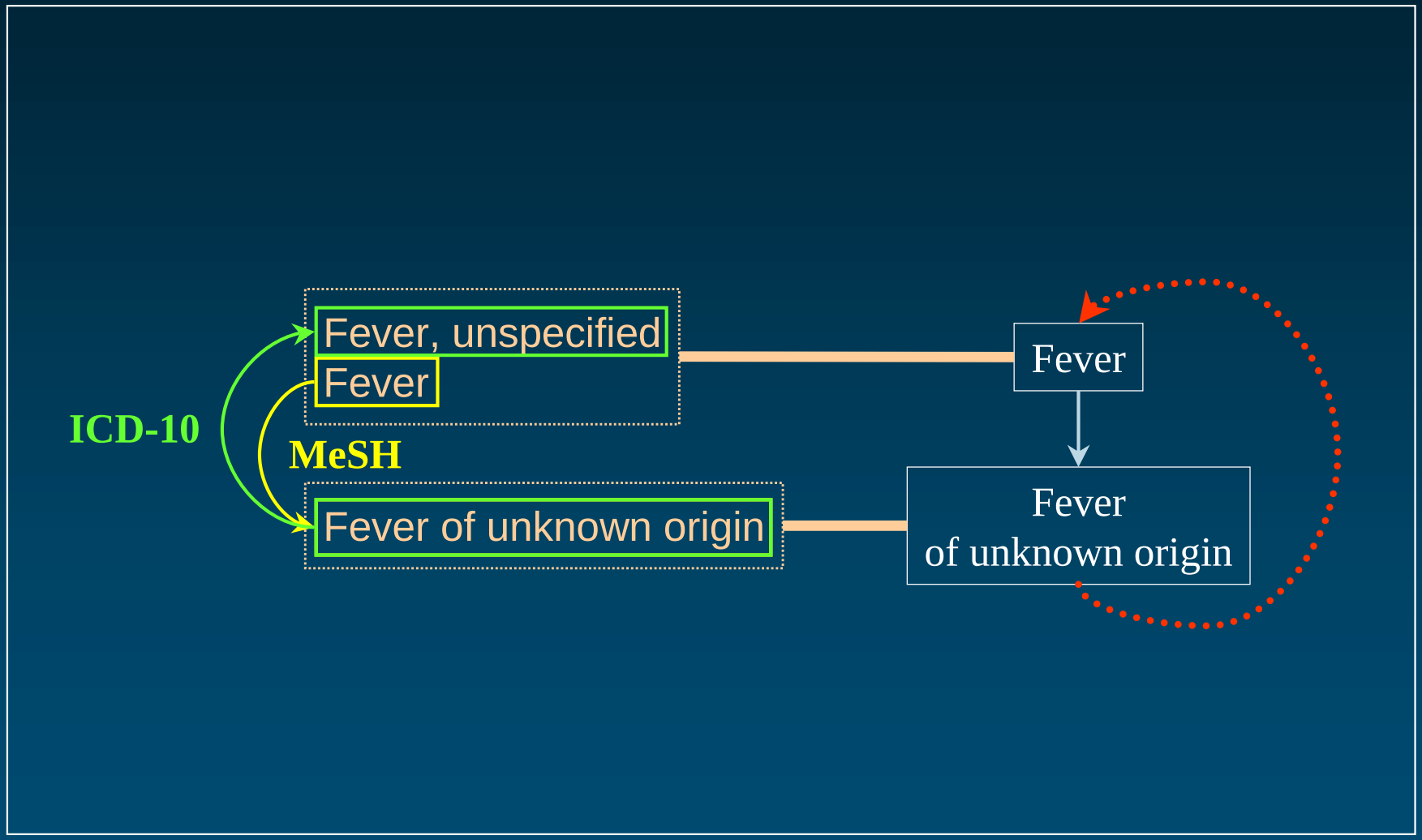
- Create an acyclic Metathesaurus
- Removed
  - 13,000 reflexive relationships
  - 1800 direct relationships
  - 120 indirect relationships

Bodenreider O.
Circular Hierarchical Relationships in the UMLS: Etiology, Diagnosis, Treatment, Complications and Prevention.
Proc AMIA Fall Symp. 2001:57-61

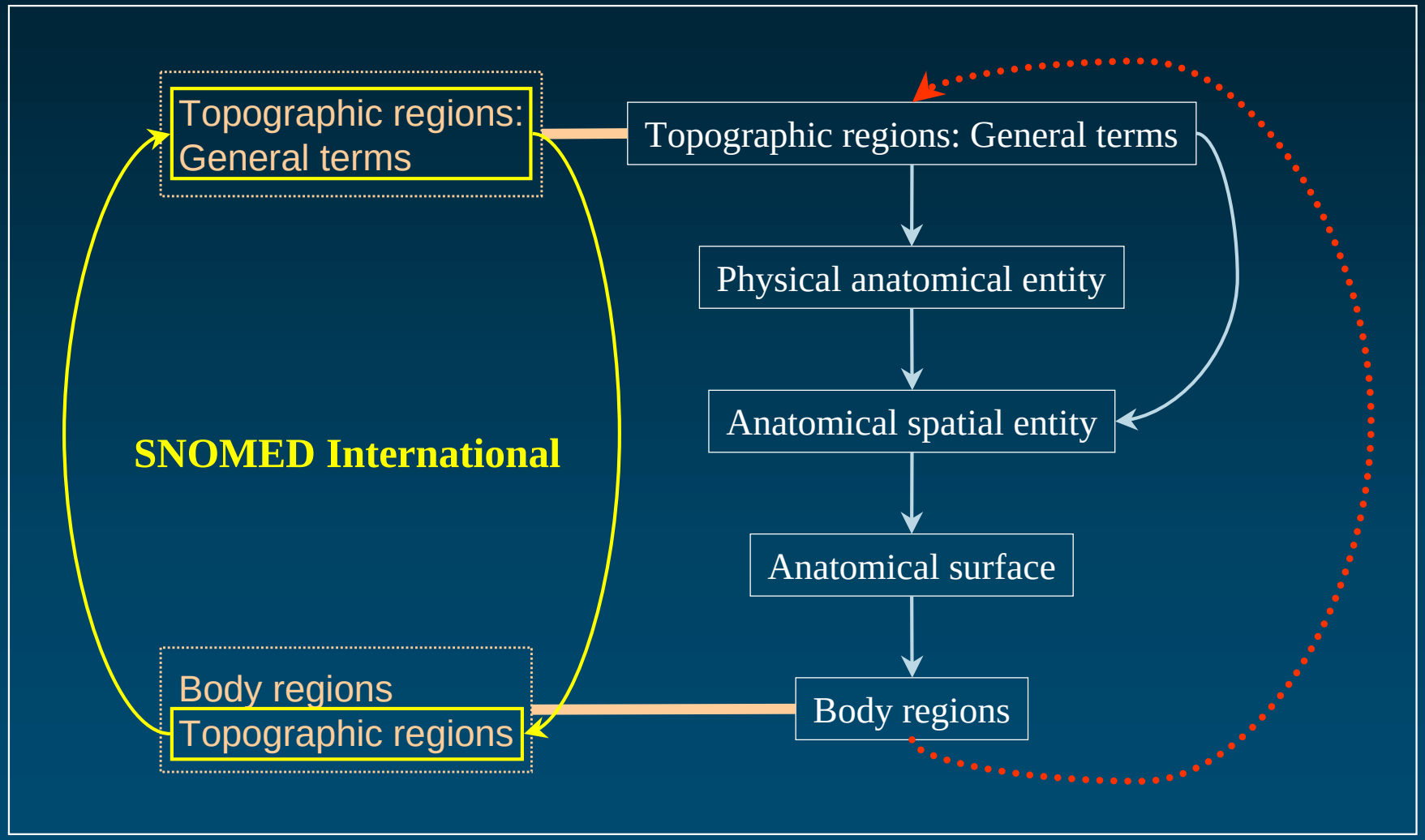
Example Reflexive relationship



Example Direct relationship



Example Indirect relationship



Discussion

- Small number of cycles, but large number of concepts having at least one cycle among the graph of their ancestors / descendants
- Methods based on redundancy
  - are no substitute for a careful review
  - But represent a trade-off between cost and efficacy
- Controls based on structure could be performed at the level of data entry

Advanced Techniques

Customize Relationships

❸ Statistical Approach

Background Statistical Knowledge

- Several kinds of knowledge in the Metathesaurus recorded as interconcept relationships
  - Symbolic: based on the meaning (MRREL)
    - “Addison’s disease” isa “disease”
    - “Addison’s disease” associated with “Addisonian crisis”
  - Statistical: based on the co-occurrence of MeSH descriptors in MEDLINE citations (MRCOC)
    - “Addison’s disease” coc “adrenal glands” [19/808]
    - “Addison’s disease” coc “prostatic neoplasms” [2/808]
    - “Addison’s disease” coc “quality of life” [2/808]

An example from MEDLINE

Cugini P, Letizia C, Cerci S, Di Palma L, Battisti P, Coppola A, Scavo D.
A chronobiological approach to circulating levels of renin, angiotensin-converting enzyme, aldosterone, ACTH, and cortisol in Addison's disease.
***Chronobiol Int* 1993 Apr;10(2):119-22**

This study deals with a chronobiological approach to the circadian rhythm of the renin-angiotensin-aldosterone system (RAAS) and the ACTH-cortisol axis (ACA) in patients with Addison's disease (PAD). The aim is to explore the mechanism(s) for which the circadian rhythmicity of the RAAS and ACA takes place. The study has shown that both the RAAS and ACA are devoid of a circadian rhythm in PAD. The lack of rhythmicity for renin and ACTH provides indirect evidence that their rhythmic secretion is in some way related to the circadian oscillation of aldosterone and cortisol. This implies a new concept: a positive feedback may be included among the mechanisms which chronoregulate the RAAS and ACA.

PMID: 8388783, UI: 93272348

- Addison's Disease/physiopathology
- Addison's Disease/blood*
- Adolescence
- Adult
- Aldosterone/blood*
- Circadian Rhythm*
- Corticotropin/blood*
- Female
- Human
- Hydrocortisone/blood*
- Male
- Middle Age
- Peptidyl-Dipeptidase A/blood*
- Renin/blood*

Background Co-occurences

- Relationships
  - Pair of concept identifiers
  - Frequency of co-occurrence
  - Source of co-occurrence
- Semantics of the relationship: undefined
  - Some redundancy with symbolic relationships
  - “Addison’s disease” coc “prostatic neoplasms” [2/808]

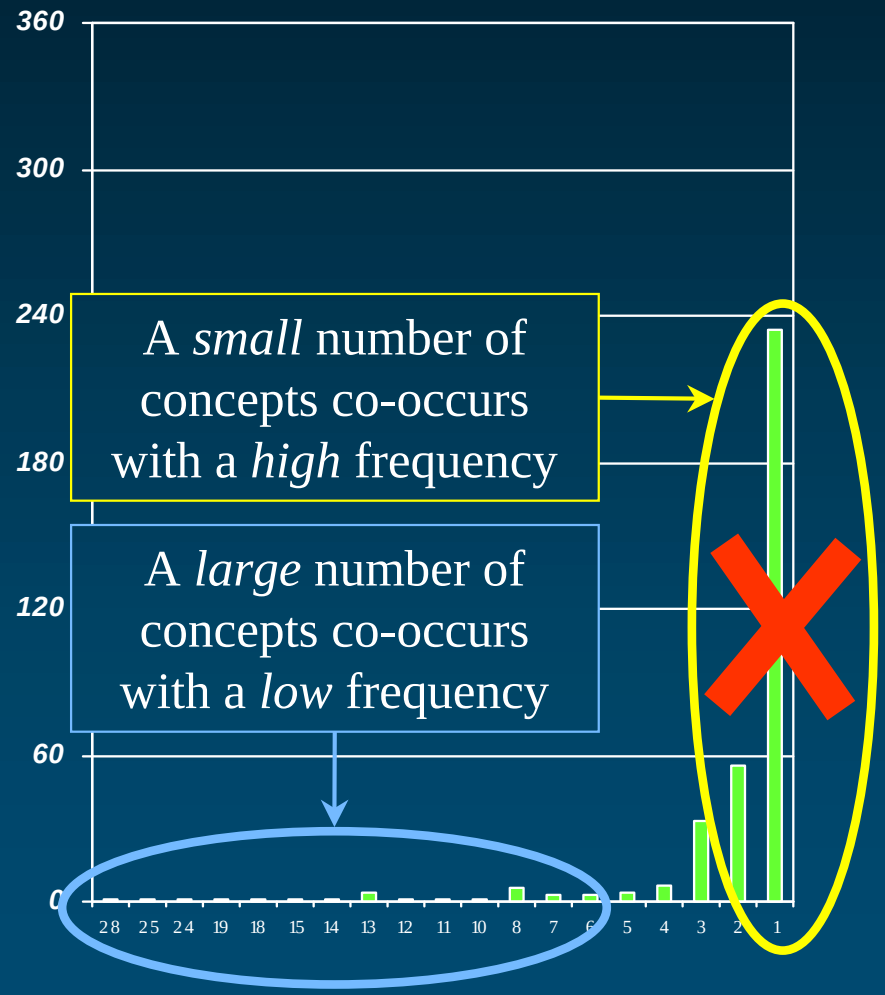
> - *Addison's disease secondary to prostatic carcinoma. A case report.*
> - *Retropubic radical prostatectomy in a patient with chronic adrenal insufficiency*

Background Co-occurences

- Only co-occurrence between “starred” descriptors is recorded in the Metathesaurus
- Relative frequency of co-occurrence
  - Freq(A and B) / Freq(A)
  - Freq(A and B) / Freq(B)
  - Surrogate for the strength of the link
- Frequency distribution may help select the most significant co-occurrences

Addison's Disease: Co-occurring concepts

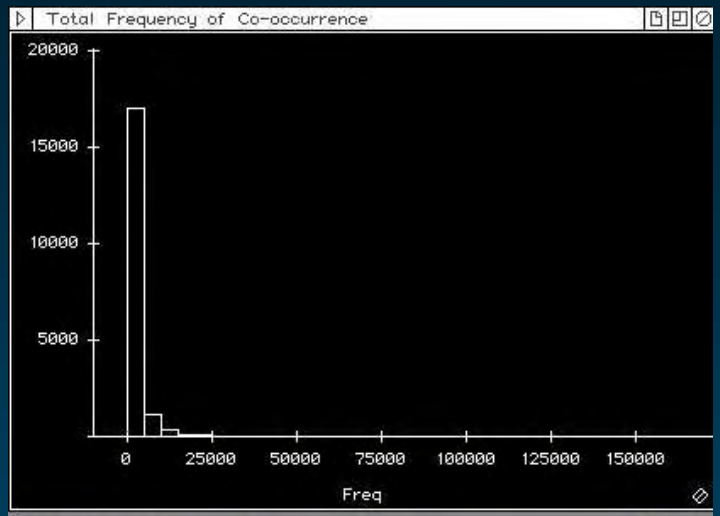
```
28  Autoimmune Diseases
25  Autoantibodies
24  Hydrocortisone
19  Adrenal Glands
18  Steroid 21 Monooxygenase
15  Corticotropin
14  Adrenal Gland Neoplasms
13  Adrenal Cortex
13  Adrenal Gland Diseases
13  Glucocorticoids
13  Polyendocrinopathies, Autoimmune
12  Diabetes Mellitus, Insulin Dependent
11  Tuberculosis, Endocrine
10  Adrenoleukodystrophy
 8  Adrenal gland hypofunction
 8  Autoantigens
 8  Cushing Syndrome
 8  Hypothyroidism
 8  Tuberculosis
 8  Chronic lymphocytic thyroiditis
    […]
 1  Circadian Rhythm
    […]
```

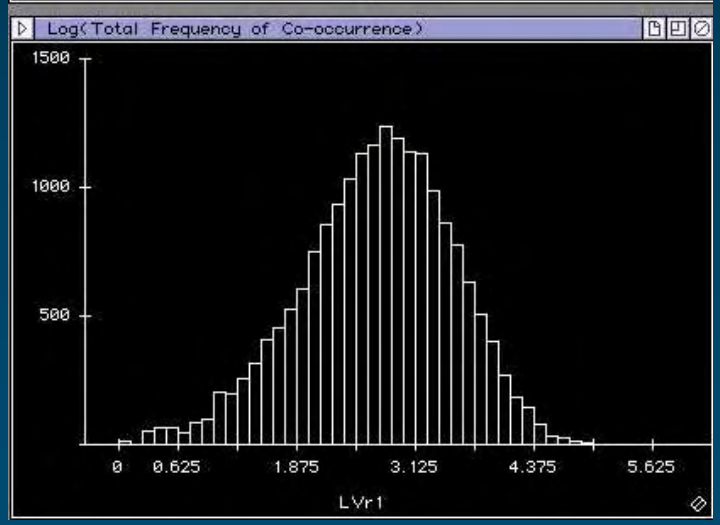


Total frequency of co-occurrence

- Number of co-occurring concepts
  - Min: 1
  - Max: 164,762
  - Median: 585

164762	Brain
137102	Liver
126009	Neurons
105382	Calcium
102109	Postoperative Complications
101955	DNA Binding Proteins
93425	Breast Neoplasms
86878	RNA, Messenger
83578	Transcription Factors
82987	Escherichia coli
82840	T Lymphocytes
82629	Aging
81442	Hypertension





Motivation

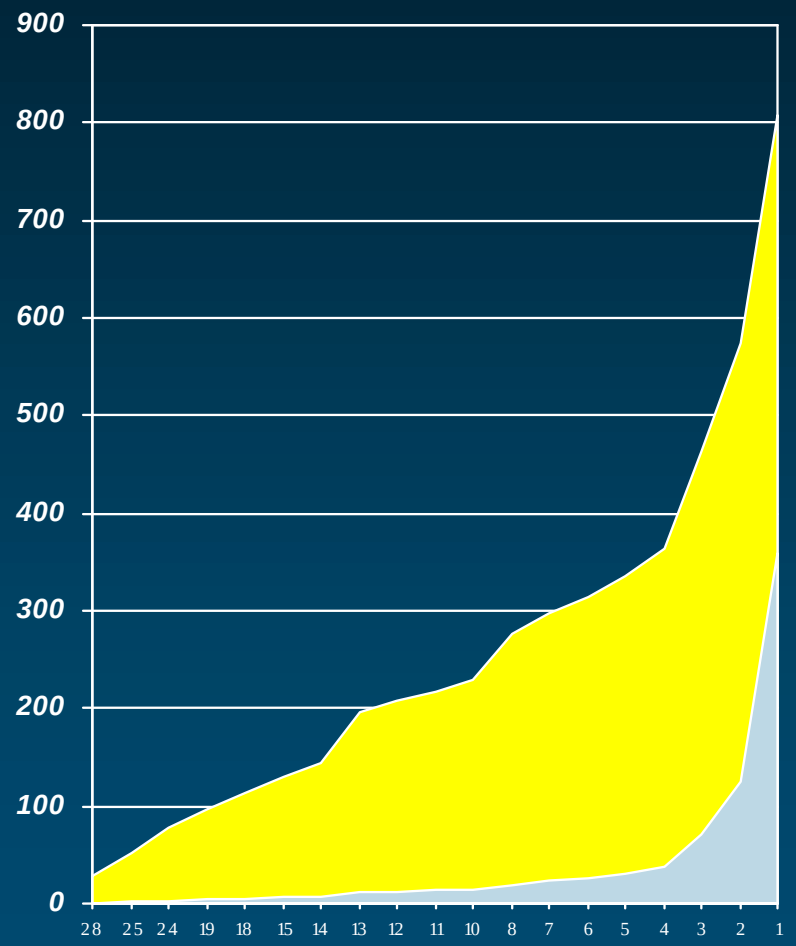
- Reduce the volume
- Select significant associations
  - For display purposes
  - Discover unexpected associations
  - Select candidate associative relationships for UMLS editors to review

Methods

- Threshold on relative frequency of co-occurrence
  - Fixed threshold
    - Absolute (e.g., at least 2)
    - Relative (e.g., at least 1%)
  - Percentile
    - e.g., 90th percentile
    - Problem with long distribution tails
  - Dynamic approach
    - Smallest number of pairs representing the largest fraction of the total frequency

Methods

- 19 classes (concepts with the same frequency)
- Total frequency: 808
- Add classes until the benefit of adding the next class becomes insignificant



Example of use Visualization

- Display only a reasonable number of co-occurring concepts
- Addison's disease
  - Co-occurring concepts: 360
  - *Displayed:* *126 (35%)*
  - Total frequency of co-occurrence: 808
  - *Represented:* *574 (71%)*

Discussion

- Only 6 percent of the relationships between co-occurring concepts are redundant with symbolic relationships in the Metathesaurus
- A more sophisticated statistical analysis is necessary to refine the filter
- Additional filters may be applied
  - E.g., minimum value for the total frequency of co-occurrence

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content — O. B., L. Roth, S. Srinivasan
    - Customize with MetamorphoSys
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

Two key questions

- Are the *meanings* already in the Metathesaurus?
- How will you maintain your system as your vocabulary and the Metathesaurus change?

Create Unique Identifiers for *your* terminology

- For your concepts, use:

  'CA000001 …' as CUIs instead

  of UMLS 'C0000001 ….' for CUIs

- Similarly, use 'LA000001 ...' for LUIs

  and 'SA000001 …' for SUIs, as needed

- Create a table which can map your UIs to UMLS UIs

e.g., `Your CUI|UMLS CUI|`

Is the meaning in the Metathesaurus?

- Use the ‘norm’ program to normalize your terms
- look for matches to the Normalized String Index (MRXNS).
- Use other sensible approaches to searching:
  - normalized word searches;
  - explore alternate naming styles and conventions

Hole, W.T., Srinivasan, S.
Discovering Missed Synonymy in a Large Concept-Oriented Metathesaurus.
Proc AMIA Fall Symp. 2000:354-8

Map your terms to Unique Identifiers

- Use Meta CUIs when synonyms are found
- Use *your* CUIs where no synonyms are found
- Store the map for future use

- You will probably want to assign Semantic Types for your new concepts

Bonus Add relationships, attributes

- As you look for synonyms, add relationships to the Metathesaurus when you add a new concept
  - Assign a REL and RELA to label the particular kinds of relationships you need and will use,
    e.g. to map or aggregate
- Add attributes (e.g. version ID, categories)

Updating to a New Meta Release

- Repeat MetamorphoSys and processing scripts used for the previous release
- Re-use previously found UIs for your terms to add your synonyms, etc. to the new Meta
- Look for new Meta Concepts which are synonyms of your concepts not previously found in Meta
- Check for any deleted or changed CUIs in MRCUI

```
C0435517|1999|SY|C0435516|
C0361163|1998|DEL||
C0785652|2000|SY|C0775088|
```

Outline of Tutorial

- Why customize? — Betsy Humphreys
- Metathesaurus basics — Olivier Bodenreider
- How to customize?
  - Removing content — O. B., L. Roth, S. Srinivasan
    - Customize with MetamorphoSys
    - Advanced techniques
  - Adding “local” content — Bill Hole
- Preview - Coming attractions — Bill Hole

What's to come in November (2002AD)

- Simple Update Model
  - Only rows with actual changes are in the update
  - Changes are rows to delete, rows to add
- Versionless Source Abbreviations
  - MR files go "Versionless"
    - e.g., the SAB 'MSH2002_06_01' becomes 'MSH'
  - You can always look up current version a new file, MRSAB:

VCUI	RCUI	VSAB	RSAB	Source Official Name	..
<CUI1>	<CUI2>	MSH2002_06_01	MSH	Medical Subject Headings	..

  - Will allow simple updates in 2003

MRSAB - Source abbreviations

- Information about all source vocabularies, e.g., Names, contacts, versions, dates, ...
  - Details in documentation
- Both Versioned and Versionless source abbreviations (SABs)
- “CURVER” field flags versions in the release
- MetamorphoSys will make MR files with either type of SAB, as you wish

What's to come in 2002AD, continued

- New Semantic Type, "Drug Delivery Device"
  - Used in RxNorm Clinical Drug Vocabulary
  - For more RxNorm info, see:

    http://umlsinfo.nlm.nih.gov/RxNorm.html

Recent vocabulary changes

- RxNorm Clinical Drug Terminology, see:

  http://umlsinfo.nlm.nih.gov/Rxnorm

- NCBI Taxonomy
- Quarterly MeSH updates
  - 2003 MeSH in November Release
    (will be used in MEDLINE from January)
- Medical Device updates (UMDNS, SPN)

Coming in 2003

- Many vocabulary updates
- Simple update files
- Gene Ontology (GO), see:

  http://www.geneontology.org

- …

Goals for 2003 and beyond

- Views
  - e.g., Natural Language Processing subset
  - Identified by an attribute added to each MR file
- Rich Data Formats, e.g. XML
  - e.g., atomic format representing all source information explicitly, more navigable hierarchies, sharable views
  - Smart update model
  - UMLS Objects and Tools
- Complete Source Transparency

Goals for 2003 and beyond

- MetamorphoSys will become the “install” program for the UMLS Metathesaurus
- Variety of output formats will be possible (Relational, XML, Atomic)
- MetamorphoSys will be able to act as an update client for the Metathesaurus

We need
User Community input!

Resources

WWW: http://www.nlm.nih.gov/research/umls/
http://umlsks.nlm.nih.gov
http://umlsinfo.nlm.nih.gov

E-mail: custserv@nlm.nih.gov

umls-users listserv:

To subscribe to the listserv, send a message to

listserv@nlm.nih.gov

which includes the following line:

subscribe umls-users

To post a message to the umls-users listserv, AFTER subscribing, send email to:

umls-users@nlm.nih.gov

Tutorial T25

AMIA Fall Symposium
Sunday, November 10, 2002
8:30 am - 12:00 noon

Lexical Tools for UMLS Developers

Allen C. Browne

Guy Divita

Chris J. Lu

Appendix

MRCON Concepts

CUI	LAT	TS	LUI	STT	SUI	STR	LRL
C0001403	ENG	P	L0001403	PF	S0010794	Addison's Disease	0
C0001403	ENG	P	L0001403	VC	S0352253	ADDISON'S DISEASE	0
C0001403	ENG	P	L0001403	VO	S0010792	Addison Disease	0
C0001403	ENG	P	L0001403	VO	S0033587	Disease, Addison	0
C0001403	ENG	P	L0001403	VO	S0469271	Addison's disease, NOS	3
C0001403	ENG	S	L0278071	PF	S0352321	ADRENAL INSUFFICIENCY (ADDISON'S DISEASE)	0
C0001403	ENG	S	L0278422	PF	S0352329	ADRENOCORTICAL INSUFFICIENCY, PRIMARY FAILURE	0
C0001403	ENG	S	L0367999	PF	S0469267	Addison melanoderma	3
C0001403	ENG	S	L0368000	PF	S0496840	Melasma addisonii	3
C0001403	ENG	S	L0368398	PF	S0506528	Primary adrenal deficiency	3
C0001403	ENG	S	L0373744	PF	S0471237	Asthenia pigmentosa	3
C0001403	ENG	S	L0377831	PF	S0473611	Bronzed disease	3
C0001403	ENG	S	L0494940	PF	S0718028	Primary adrenocortical insufficiency	3
C0001403	ENG	s	L0494937	PF	S0718027	Primary adrenocortical insuff	3
C0001403	FIN	P	L1510041	PF	S1805950	Addisonin tauti	3
C0001403	FRE	S	L1272481	PF	S1514427	MALADIE D'ADDISON	2
C0001403	GER	P	L1229627	PF	S1471573	Addison Krankheit	3
C0001403	GER	S	L1288823	PF	S1530769	Primaere Nebennierenrindeninsuffizienz	1
C0001403	ITA	P	L1276837	PF	S1518783	Morbo di Addison	3
C0001403	POR	P	L0324623	PF	S0432928	DOENCA DE ADDISON	2
C0001403	RUS	P	L0889403	PF	S1093220	ADDISONOVA BOLEZN'	3
C0001403	SPA	P	L0342625	PF	S0450930	ENFERMEDAD DE ADDISON	3

[…]

MRSO Sources

```
CUI      LUI      SUI      SAB     TTY  SCD   SRL
C0001403|L0001403|S0010792|MSH2000|EN|D000224|0|
C0001403|L0001403|S0010794|MSH2000|MH|D000224|0|
C0001403|L0001403|S0010796|MSH2000|PM|D000224|0|
C0001403|L0001403|S0010796|PSY94|PT|00810|3|
C0001403|L0001403|S0219379|ICD91|IT|255.4|0|
C0001403|L0001403|S0220088|ICD91|IT|255.4|0|
C0001403|L0001403|S0220088|MSH2000|PM|D000224|0|
C0001403|L0001403|S0352252|CCPSS99|PT|0022753|3|
C0001403|L0001403|S0352252|DXP94|SY|NOCODE|0|
C0001403|L0001403|S0352253|CST95|GT|ADREN INSUFFIC|0|
C0001403|L0001403|S0352253|WHO97|IT|0410|2|
C0001403|L0001403|S0354372|AOD95|DE|0000005430|0|
C0001403|L0001403|S0354372|CSP98|PT|0060-3321|0|
C0001403|L0001403|S0354372|LCH90|PT|U000061|0|
C0001403|L0001403|S0354372|RCD99|PT|C1541|3|
C0001403|L0001403|S0354372|SNM2|SY|D-2332|3|
C0001403|L0001403|S0469271|SNMI98|PT|DB-70620|3|
C0001403|L0278071|S0352321|COS93|PT|U000087|0|
C0001403|L0278422|S0352329|DXP94|SY|NOCODE|0|
C0001403|L0367999|S0469267|SNMI98|SY|DB-70620|3|
C0001403|L0494937|S0718027|RCD99|AB|C1541|3|
C0001403|L0494940|S0718028|ICD10|PT|E27.1|3|
C0001403|L0494940|S0718028|RCD99|SY|C1541|3|
[…]
```

MRDEF Definitions

```
CUI      SAB     DEF
C0001403|MSH2000|A disease characterized by hypotension, weight loss, anorexia,
weakness, and sometimes a bronze like melanotic hyperpigmentation of the skin. I
t is due to tuberculosis  or autoimmune induced disease (hypofunction) of the ad
renal glands that results in deficiency of aldosterone and cortisol. In the abse
nce of replacement therapy, it is usually fatal.|
```

MRSTY Semantic Types

```
CUI      TUI  STY
C0001400|T040|Organism Function|
C0001403|T047|Disease or Syndrome|
C0001406|T083|Geographic Area|
C0001407|T114|Nucleic Acid, Nucleoside, or Nucleotide|
C0001407|T123|Biologically Active Substance|
```

MRATX Associated Expressions

```
CUI      SAB    REL     ATX
 Closed fracture of malar and maxillary bones, NOS
C0009045|MSH2000|B|<Zygomatic Fractures> OR <Maxillary Fractures>|
 Unilateral congenital dislocation of hip
C0009702|MSH2000|B|<Hip Dislocation, Congenital> AND <Femur Head>/<abnormalities>|
 Suture of bladder
C0010700|MSH2000|B|<Bladder>/<surgery>|
```

MRCXT Contexts

```
CUI      SUI      SAB    SCD    CXN CXL RNK   CXS              CUI2      HCD REL XC
C0001403|S0469271|SNMI98|DB-70620|1|ANC|1|SNOMED International|C0220967||||
C0001403|S0469271|SNMI98|DB-70620|1|ANC|2|DISEASES/DIAGNOSES|C0338067||||
C0001403|S0469271|SNMI98|DB-70620|1|ANC|3|DISEASES OF THE END. SYSTEM|C0014130||||
C0001403|S0469271|SNMI98|DB-70620|1|ANC|4|DISEASES OF THE ADRENAL GLANDS|C0001621||||
C0001403|S0469271|SNMI98|DB-70620|1|CCP||Addison's disease, NOS|C0001403|DB-70620|||

C0001403|S0718028|ICD10|E27.1|1|ANC|1|ICD, Tenth Revision (ICD 10)|C0391804||||
C0001403|S0718028|ICD10|E27.1|1|ANC|2|End., nutr. and metabolic diseases|C0694452||||
C0001403|S0718028|ICD10|E27.1|1|ANC|3|Disorders of other endocrine glands|C0178257||||
C0001403|S0718028|ICD10|E27.1|1|ANC|4|Other disorders of adrenal gland|C0494313||||
C0001403|S0718028|ICD10|E27.1|1|CCP||Primary adrenocortical insuff.|C0001403|E27.1|||

(*   C0001403|S0010794|MSH2000)
*|D000224|1|ANC|1|MeSH|C0220876||||
*|D000224|1|ANC|2|Diseases (MeSH Category)|C0012674|C|||
*|D000224|1|ANC|3|Endocrine Diseases|C0014130|C19|||
*|D000224|1|ANC|4|Adrenal Gland Diseases|C0001621|C19.53|isa||
*|D000224|1|ANC|5|Adrenal Gland Hypofunction|C0001623|C19.53.264|manifestation_of||
*|D000224|1|CCP||Addison's Disease|C0001403|C19.53.264.263|has_manifestation||
*|D000224|1|SIB||Adrenoleukodystrophy|C0001661|C19.53.264.270|has_manifestation||
*|D000224|1|SIB||Hypoaldosteronism|C0020595|C19.53.264.480|has_manifestation||
```

MRSAT Simple concept attributes

```
CUI      LUI      SUI      SCD    ATN SAB     ATV
C0001403|L0001403|S0010792|D000224|EV|MSH2000|ADDISON DIS|
C0001403|L0001403|S0010794|D000224|AN|MSH2000|an autoimmune dis with adrenal hypofunction|
C0001403|L0001403|S0010794|D000224|DC|MSH2000|1|
C0001403|L0001403|S0010794|D000224|DE|MSH2000|ADDISONS DIS|
[…]
C0001403|L0001403|S0010794|D000224|M93|MSH2000|*120|
C0001403|L0001403|S0010794|D000224|M93|MSH2000|162|
C0001403|L0001403|S0010794|D000224|MED|MSH2000|*116|
C0001403|L0001403|S0010794|D000224|MED|MSH2000|167|
C0001403|L0001403|S0010794|D000224|MMR|MSH2000|19940628|
C0001403|L0001403|S0010794|D000224|MN|MSH2000|C19.53.264.263|
C0001403|L0001403|S0010794|D000224|MN|MSH2000|C20.111.163|
C0001403|L0001403|S0010794|D000224|TH|MSH2000|NLM (1966)|
C0001403|L0001403|S0352252|0022753|CCF|CCPSS99|44|
C0001403|L0001403|S0354372|C1541|RID|RCD99|Y41X1|
C0001403|L0001403|S0469271|DB-70620|SIC|SNMI98|255.4|
C0001403|L0367999|S0469267|DB-70620|SIC|SNMI98|255.4|
[…]
C0001403|L0494937|S0718027|C1541|RID|RCD99|Y41X2|
C0001403|L0494940|S0718028|C1541|RID|RCD99|Y41X2|
C0001403||||DA|MTH|19900930|
C0001403||||MR|MTH|20000101|
C0001403||||ST|MTH|R|
```

MRLO Locators

```
CUI      ISN                 FR  UN         SUI    SNA SOUI
C0001403|MEDLINE(1990-1995)|228|*CITATIONS|S0010794|||
C0001403|MEDLINE(1996-Fall 1999)|116|*CITATIONS|S0010794|||
C0001403|DXPLAIN|||S0352252|||
C0001403|DXPLAIN|||S0352329|||
```

MRRANK Name Ranking

```
RANK SAB TTY SUPRES
0324|MTH|PN|N|
0323|MTH|MM|N|
0322|MSH2000|MH|N|
0321|MSH2000|HT|N|
0320|MSH2000|TQ|N|
0319|MSH2000|GQ|N|
0318|MSH2000|LQ|N|
0317|MSH2000|EP|N|
0316|MSH2000|EN|N|
0315|MSH2000|XQ|N|
0314|MSH2000|NM|N|
0313|DSM4|PT|N|
0312|DSM3R|PT|N|
0311|SNMI98|PT|N|
0310|SNMI98|PX|Y|
0309|SNMI98|HT|N|
0308|SNMI98|HX|Y|
0307|NDDF99|CD|N|
0306|NDDF99|IN|N|
0305|MDDB99|CD|N|
0304|MMX99|CD|N|
0303|MMX99|IN|N|
0302|RCDSA|PT|N|
[…]
```

MRREL Inter-concept Relationships

```
CUI1    REL CUI2   RELA  SAB  SL      MG
C0001403|AQ|C0205470||MSH2000|MSH2000||
C0001403|AQ|C0348026||MSH2000|MSH2000||
C0001403|CHD|C0271737||RCD99|RCD99||
C0001403|CHD|C0342477||RCD99|RCD99||
C0001403|PAR|C0001623|manifestation_of|MSH2000|MSH2000||
C0001403|PAR|C0004364|inverse_isa|MSH2000|MSH2000||
C0001403|PAR|C0405580||AOD95|AOD95||
C0001403|PAR|C0405580||RCD99|RCD99||
C0001403|PAR|C0494313||ICD10|ICD10||
C0001403|RB|C0001621||MTH|MTH||
C0001403|RB|C0004364||CSP98|MTH||
C0001403|RL|C0405580|mapped_from|SNMI98|SNMI98||
C0001403|RN|C0518933||MTH|MTH||
C0001403|RN|C0518934||MTH|MTH||
C0001403|RO|C0020615|clinically_associated_with|CCPSS99|CCPSS99||
C0001403|RO|C0041296||MTH|MTH||
C0001403|RO|C0085860|mapped_to|CSP98|CSP98||
C0001403|RO|C0151467|clinically_similar|RAM99|RAM99||
C0001403|RO|C0152889|associated_with|SNMI98|SNMI98||
C0001403|RO|C0405580|mapped_from|CST95|CST95||
C0001403|SIB|C0001661||MSH2000|MSH2000||
C0001403|SIB|C0002880||CSP98|CSP98||
[…]
```

MRCOC Co-occurrences

```
CUI1     CUI2    SOC COT COF COA
C0001403|C0000737|MBD|L|1|CO 1,DI 1|
C0001403|C0000833|MBD|L|1|DT 1|
C0001403|C0000833|MED|L|1|DT 1,MI 1,RA 1|
C0001403|C0001175|MBD|L|1|CO 1|
C0001403|C0001180|MBD|L|1|CO 1|
C0001403|C0001418|MBD|L|2|ET 2|
C0001403|C0001430|MED|L|1|BL 1,CO 1|
C0001403|C0001613|MBD|L|5|PP 2,CN 1,DI 1,HI 1,IM 1,SU 1|
C0001403|C0001613|MED|L|7|IM 4,ET 2,PP 2,BL 1,CL 1,PA 1|
C0001403|C0001614|MED|L|1|BL 1,CI 1|
C0001403|C0001617|MBD|L|1|BL 1|
C0001403|C0001618|MBD|L|1|IM 1|
C0001403|C0001618|MED|L|3|BL 2,CO 2,ET 1,PA 1|
C0001403|C0001621|MBD|L|10|ET 7,DI 3,PA 3,BL 1,CO 1,DT 1,PP 1|
C0001403|C0001621|MED|L|3|ET 3,DI 2|
C0001403|C0001623|MBD|L|7|DI 3,ET 2,PP 2,<> 1,CN 1,DT 1,IM 1,PA 1,TH 1|
C0001403|C0001623|MED|L|1|DI 1,ET 1|
C0001403|C0001624|MBD|L|10|ET 9,DI 2,DT 1,PA 1|
C0001403|C0001624|MED|L|3|DI 2,ET 2|
C0001403|C0001625|MBD|L|12|ET 4,CO 3,RA 3,SU 3,IM 2,BL 1,DT 1,EN 1,MI 1,PA 1,PP 1|
C0001403|C0001625|MED|L|7|IM 3,DI 2,PP 2,RA 2,BL 1,CO 1,ET 1,HI 1,PA 1,TH 1|
C0001403|C0001627|MBD|L|1|DT 1|
[…]
```

MRCON Suppressible synonyms

CUI	LAT	TS	LUI	STT	SUI	STR	LRL
C0154009	ENG	P	L0180842	PF	S0245368	Benign neoplasm of prostate	0
C0154009	ENG	P	L0180842	VO	S1650872	PROSTATE NEOPLASM BENIGN	3
C0154009	ENG	P	L0180842	VO	S1912324	Neoplasm benign;prostate	3
C0154009	ENG	P	L0180842	VO	S1933166	Neoplasm benign, prostate	3
C0154009	ENG	S	L0524756	PF	S0599238	Benign tumor of prostate	3
C0154009	ENG	S	L0524757	PF	S0599632	Benign tumour of prostate	3
C0154009	ENG	S	L0524758	PF	S0598914	Benign prostatic tumor	3
C0154009	ENG	S	L0524759	PF	S0598915	Benign prostatic tumour	3
C0154009	ENG	s	L0033572	PF	S0999020	Prostate <3>	0
C0154009	ENG	s	L0033572	VO	S0077252	Prostate	3
C0154009	GER	P	L1258213	PF	S1500159	Gutartige Neubildung: Prostata	1

MRCUI Concept history

```
CUI1      VER CREL  CUI2
C0241779|1996|SY|C0001403|
C0271735|1996|SY|C0001403|
```

SRDEF Basic information

```
RT  TUI  STY/RL  STN/RTN  DEF     EX     UN     NH     ABR     RIN
STY|T001|Organism|A1.1|Generally, a living individual, including all plants and
animals.|Homozygote; Radiation Chimera; Sporocyst|||||
STY|T002|Plant|A1.1.1|An organism having cellulose cell walls, growing by
synthesis of inorganic substances, generally distinguished by the presence of
chlorophyll, and lacking the power of locomotion.  Plant parts are included here
as well.|Pollen; Potatoes; Vegetables|||||
STY|T003|Alga|A1.1.1.1|A chiefly aquatic plant that contains chlorophyll, but does
not form embryos during development and lacks vascular tissue.|Chlorella;
Laminaria; Seaweed|||||
STY|T004|Fungus|A1.1.2|A eukaryotic organism characterized by the absence of
chlorophyll and the presence of a rigid cell wall.  Included here are both slime
molds and true fungi such as yeasts, molds, mildews, and mushrooms.|Aspergillus
clavatus; Blastomyces; Helminthosporium; Neurospora|||||
[…]
RL|T132|physically_related_to|R1|Related by virtue of some physical attribute or
characteristic.||||PR|physically_related_to|
RL|T133|part_of|R1.1|Composes, with one or more other physical units, some larger
whole.  This includes component of, division of, portion of, fragment of, section
of, and layer of.||||PT|has_part|
[…]
RL|T186|isa|H|The basic hierarchical link in the Network.  If one item "isa"
another item then the first item is more specific in meaning than the second
item.||||IS|inverse_isa|
[…]
```

SRSTR Structure

```
STY/RL              RL      STY/RL                                  LS
Biologic Function|affects|Organism|D|
Biologic Function|isa|Natural Phenomenon or Process|D|
Biologic Function|process_of|Organism|D|
Biologic Function|produces|Biologically Active Substance|D|
Biologic Function|produces|Body Substance|D|
[…]
Disease or Syndrome|conceptually_related_to|Experimental Model of Disease|DNI|
Disease or Syndrome|isa|Pathologic Function|D|
Disease or Syndrome|produces|Tissue|D|
[…]
Medical Device|isa|Manufactured Object|D|
Medical Device|prevents|Injury or Poisoning|D|
Medical Device|prevents|Pathologic Function|D|
Medical Device|treats|Anatomical Abnormality|D|
Medical Device|treats|Injury or Poisoning|D|
Medical Device|treats|Pathologic Function|D|
Medical Device|treats|Sign or Symptom|D|
[…]
Mental Process|process_of|Plant|B|  blocks  Biologic Function|process_of|Organism|D|
[…]
part_of|isa|physically_related_to|D|
[…]
```

SRSTRE2 Structure (expanded)

```
STY                    RL          STY
Disease or Syndrome|isa|Pathologic Function|
Disease or Syndrome|isa|Biologic Function|
Disease or Syndrome|isa|Natural Phen. or Pr.|
Disease or Syndrome|isa|Phenomenon or Process|
Disease or Syndrome|isa|Event|
Disease or Syndrome|affects|Alga|
Disease or Syndrome|affects|Amphibian|
Disease or Syndrome|affects|Animal|
Disease or Syndrome|affects|Archaeon|
Disease or Syndrome|affects|Bacterium|
Disease or Syndrome|affects|Biologic Function|
Disease or Syndrome|affects|Bird|
Disease or Syndrome|affects|Cell Function|
Disease or Syndrome|affects|Cell or Molecular Dysfunction|
[…]
```

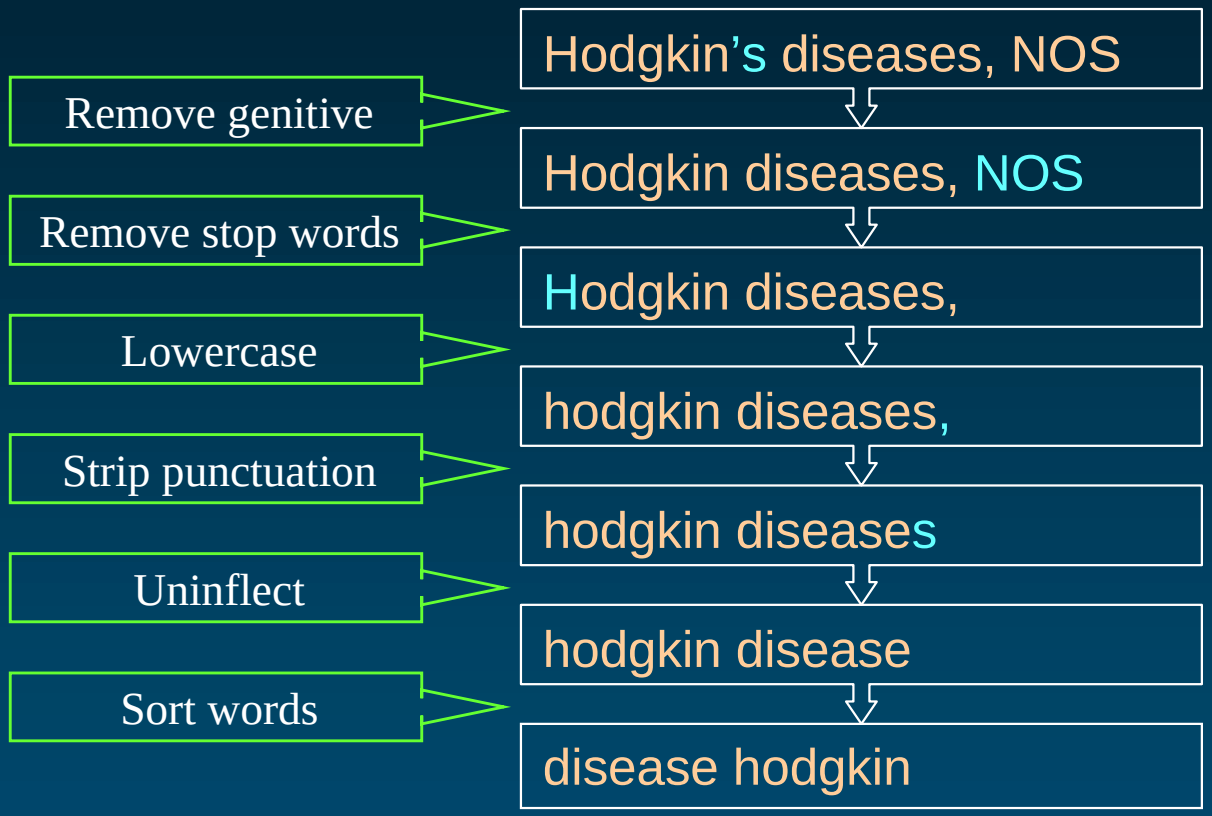
```
Pathologic Function |isa|Biologic Function|
Biologic Function|isa|Natural Phen. or Process|
Natural Phen. or Process|isa|Phen. or Process|
Phenomenon or Process|isa|Event|
```

from `Biologic Function|affects|Organism|D|`

Normalization Example

Hodgkin Disease
HODGKINS DISEASE
Hodgkin's Disease
Disease, Hodgkin's
Hodgkin's, disease
HODGKIN'S DISEASE
Hodgkin's disease
Hodgkins Disease
Hodgkin's disease NOS
Hodgkin's disease, NOS
Disease, Hodgkins
Diseases, Hodgkins
Hodgkins Diseases
Hodgkins disease
hodgkin's disease
Disease, Hodgkin

Normalization



Addison's Disease: Co-occurring concepts

```
25  Autoimmune Diseases
21  Autoantibodies
20  Hydrocortisone
19  Adrenal Glands
16  Steroid 21 Monooxygenase
13  Adrenal Gland Diseases
13  Adrenal Gland Neoplasms
12  Polyendocrinopathies, Autoimmune
12  Adrenal Cortex
11  Tuberculosis, Endocrine
10  Corticotropin
10  Glucocorticoids
 9  Diabetes Mellitus, Insulin Dependent
 8  Thyroiditis, Autoimmune
 8  Tuberculosis
 8  Hypothyroidism
 8  Adrenal gland hypofunction
 8  Autoantigens
 8  Adrenoleukodystrophy
[…]
 1  Circadian Rhythm
[…]
```