Acknowledgments

◆ Alexa T. McCray
Harvard Medical School
Boston, Massachusetts
Outline

◆ The UMLS: A two-level structure
◆ UMLS Semantic Network
  ● What is it?
    ■ Semantic types
    ■ Semantic relationships
    ■ Semantic relations
  ● How to use it?
  ● Where it get it from?
◆ Future directions
Unified Medical Language System

- **SPECIALIST Lexicon**
  - 360,000 lexical items
  - Part of speech and variant information

- **Metathesaurus**
  - 6M names from over 100 terminologies
  - 1.5M concepts
  - 8M relations

- **Semantic Network**
  - 135 high-level categories
  - 7000 relations among them
The UMLS

A two-level structure
Concept categorization

- High-level categories (semantic types)
- Assigned by the Metathesaurus editors
- Independently of the hierarchies in which these concepts are located

http://en.wikipedia.org/wiki/Glycyrrhiza
UMLS Semantic Network

What is it?
Semantic network

- Semantic types: Nodes
- Semantic relationships: Links
- Semantic relations: Triples (node₁, link, node₂)
Semantic types

- Semantic types (135)
  - tree structure
  - 2 major hierarchies
    - Entity
      - Physical Object
      - Conceptual Entity
    - Event
      - Activity
      - Phenomenon or Process
<table>
<thead>
<tr>
<th>Entity</th>
<th>Entity [continued]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Object</td>
<td>Physical Object [continued]</td>
</tr>
<tr>
<td>■ Organism</td>
<td>■ Manufactured Object</td>
</tr>
<tr>
<td>■ Plant</td>
<td>■ Medical Device</td>
</tr>
<tr>
<td>■ Alga</td>
<td>■ Research Device</td>
</tr>
<tr>
<td>■ Fungus</td>
<td>■ Clinical Drug</td>
</tr>
<tr>
<td>■ Virus</td>
<td>■ Chemical</td>
</tr>
<tr>
<td>■ Rickettsia or Chlamydia</td>
<td>■ Chemical Viewed Functionally</td>
</tr>
<tr>
<td>■ Bacterium</td>
<td>■ Pharmacologic Substance</td>
</tr>
<tr>
<td>■ Archaeon</td>
<td>■ Biologically Active Substance</td>
</tr>
<tr>
<td>■ Animal</td>
<td>■ Neuropeptide Or Biogenic Amine</td>
</tr>
<tr>
<td>■ Invertebrate</td>
<td>■ Hormone</td>
</tr>
<tr>
<td>■ Vertebrate</td>
<td>■ Enzyme</td>
</tr>
<tr>
<td>■ Amphibian</td>
<td>■ Immunologic Factor</td>
</tr>
<tr>
<td>■ Bird</td>
<td>■ Receptor</td>
</tr>
<tr>
<td>■ Fish</td>
<td>■ Indicator, Reagent, Or Diagnostic Aid</td>
</tr>
<tr>
<td>■ Mammal</td>
<td>■ Hazardous Or Poisonous Substance</td>
</tr>
<tr>
<td>■ Human</td>
<td>■ Chemical Viewed Structurally</td>
</tr>
<tr>
<td>Anatomical Structure</td>
<td>■ Organic Chemical</td>
</tr>
<tr>
<td>Embryonic Structure</td>
<td>■ Nucleic Acid, Nucleoside, Or Nucleotide</td>
</tr>
<tr>
<td>Anatomical Abnormality</td>
<td>■ Organophosphorus Compound</td>
</tr>
<tr>
<td>Congenital Abnormality</td>
<td>■ Amino Acid, Peptide, Or Protein</td>
</tr>
<tr>
<td>Acquired Abnormality</td>
<td>■ Carbohydrate</td>
</tr>
<tr>
<td>Fully Formed Anatomical Structure</td>
<td>■ Lipid</td>
</tr>
<tr>
<td>Body Part, Organ, Or Organ Component</td>
<td>■ Steroid</td>
</tr>
<tr>
<td>Tissue</td>
<td>■ Eicosanoid</td>
</tr>
<tr>
<td>Cell</td>
<td>■ Inorganic Chemical</td>
</tr>
<tr>
<td>Cell Component</td>
<td>■ Element, Ion, Or Isotope</td>
</tr>
<tr>
<td>Gene or Genome</td>
<td>■ Body Substance</td>
</tr>
<tr>
<td></td>
<td>■ Food</td>
</tr>
<tr>
<td></td>
<td>■ Conceptual Entity</td>
</tr>
<tr>
<td></td>
<td>■ Idea Or Concept</td>
</tr>
<tr>
<td></td>
<td>■ Temporal Concept</td>
</tr>
<tr>
<td></td>
<td>■ Qualitative Concept</td>
</tr>
<tr>
<td></td>
<td>■ Quantitative Concept</td>
</tr>
<tr>
<td></td>
<td>■ Functional Concept</td>
</tr>
<tr>
<td></td>
<td>■ Body System</td>
</tr>
<tr>
<td></td>
<td>■ Spatial Concept</td>
</tr>
<tr>
<td></td>
<td>■ Body Space Or Junction</td>
</tr>
<tr>
<td></td>
<td>■ Body Location Or Region</td>
</tr>
<tr>
<td></td>
<td>■ Molecular Sequence</td>
</tr>
<tr>
<td></td>
<td>■ Nucleotide Sequence</td>
</tr>
<tr>
<td></td>
<td>■ Amino Acid Sequence</td>
</tr>
<tr>
<td></td>
<td>■ Carbohydrate Sequence</td>
</tr>
<tr>
<td></td>
<td>■ Geographic Area</td>
</tr>
<tr>
<td></td>
<td>■ Finding</td>
</tr>
<tr>
<td></td>
<td>■ Laboratory Or Test Result</td>
</tr>
<tr>
<td></td>
<td>■ Sign Or Symptom</td>
</tr>
<tr>
<td></td>
<td>■ Organism Attribute</td>
</tr>
<tr>
<td></td>
<td>■ Clinical Attribute</td>
</tr>
<tr>
<td></td>
<td>■ Intellectual Product</td>
</tr>
<tr>
<td></td>
<td>■ Classification</td>
</tr>
<tr>
<td></td>
<td>■ Regulation Or Law</td>
</tr>
<tr>
<td></td>
<td>■ Language</td>
</tr>
<tr>
<td></td>
<td>■ Occupation Or Discipline</td>
</tr>
<tr>
<td></td>
<td>■ Biomedical Occupation Or Discipline</td>
</tr>
<tr>
<td></td>
<td>■ Organization</td>
</tr>
<tr>
<td></td>
<td>■ Health Care Related Organization</td>
</tr>
<tr>
<td></td>
<td>■ Professional Society</td>
</tr>
<tr>
<td></td>
<td>■ Self-help Or Relief Organization</td>
</tr>
<tr>
<td></td>
<td>■ Group Attribute</td>
</tr>
<tr>
<td></td>
<td>■ Group</td>
</tr>
<tr>
<td></td>
<td>■ Professional Or Occupational Group</td>
</tr>
<tr>
<td></td>
<td>■ Population Group</td>
</tr>
<tr>
<td></td>
<td>■ Family Group</td>
</tr>
<tr>
<td></td>
<td>■ Age Group</td>
</tr>
<tr>
<td></td>
<td>■ Patient Or Disabled Group</td>
</tr>
</tbody>
</table>
Hierarchical structure

Event
- Activity
  - Behavior
    - Social Behavior
    - Individual Behavior
  - Daily or Recreational Activity
  - Occupational Activity
  - Health Care Activity
  - Laboratory Procedure
  - Diagnostic Procedure
  - Therapeutic or Preventive Procedure
  - Research Activity
  - Molecular Biology Research Technique
  - Governmental or Regulatory Activity
  - Educational Activity
  - Machine Activity

Event [continued]
- Phenomenon or Process
  - Human-caused Phenomenon or Process
    - Environmental Effect of Humans
    - Natural Phenomenon or Process
      - Biologic Function
        - Physiologic Function
        - Organism Function
          - Mental Process
          - Organ or Tissue Function
            - Cell Function
              - Molecular Function
              - Genetic Function
              - Pathologic Function
                - Disease or Syndrome
                - Mental or Behavioral Dysfunction
                  - Neoplastic Process
                  - Cell or Molecular Dysfunction
                  - Experimental Model of Disease
        - Pathologic Function
          - Mental or Behavioral Dysfunction
          - Neoplastic Process
          - Cell or Molecular Dysfunction
          - Experimental Model of Disease
        - Injury or Poisoning

Hierarchical structure  Event
“Biologic Function” hierarchy (isa)
## Semantic type information

<table>
<thead>
<tr>
<th>ID</th>
<th>T047</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Disease or Syndrome</td>
</tr>
<tr>
<td>Tree number</td>
<td>B2.2.1.2.1</td>
</tr>
<tr>
<td>Definition</td>
<td>A condition which alters or interferes with a normal process, state, or activity of an organism. It is usually characterized by the abnormal functioning of one or more of the host's systems, parts, or organs. Included here is a complex of symptoms descriptive of a disorder.</td>
</tr>
<tr>
<td>Examples</td>
<td>Diabetes Mellitus; Dumping Syndrome; Malabsorption Syndromes; Nephrotic Syndrome</td>
</tr>
<tr>
<td>Usage note</td>
<td>Any specific disease or syndrome that is modified by such modifiers as &quot;acute&quot;, &quot;prolonged&quot;, etc. will also be assigned to this type. If an anatomic abnormality has a pathologic manifestation, then it will be given this type as well as a type from the 'Anatomical Abnormality' hierarchy, e.g., &quot;Diabetic Cataract&quot; will be double-typed for this reason.</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>dsyn</td>
</tr>
</tbody>
</table>
Semantic relationships

- Semantic network relationships (54)
  - Hierarchical (isa = is a kind of)
    - among types
      - Animal isa Organism
      - Enzyme isa Biologically Active Substance
    - among relations
      - treats isa affects
  - Associative (non-hierarchical)
    - Sign or Symptom diagnoses Pathologic Function
    - Pharmacologic Substance treats Pathologic Function
Assiociative relationship hierarchy

- associated with
  - physically related to
    - spatially related to
      - temporally related to
        - functionally related to
          - conceptually related to
associated with

physically related to

part of
consists of
contains
connected to
interconnects
branch of
tributary of
ingredient of

spatially related to

traverses
adjacent to
surrounds
location of

temporally related to

coop-occurs with
precedes

functionally related to

conceptually related to
## Semantic relationship information

<table>
<thead>
<tr>
<th>ID</th>
<th>T202</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>ingredient_of</td>
</tr>
<tr>
<td>Tree number</td>
<td>R1.8</td>
</tr>
<tr>
<td>Definition</td>
<td><em>Is a component of, as in a constituent of a preparation.</em></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>IG</td>
</tr>
<tr>
<td>Inverse rel.</td>
<td>has_ingredient</td>
</tr>
</tbody>
</table>
Semantic relations

- 612 assertions
  - 558 (type₁, relationship, type₂)
    - Hierarchical
      - Animal isa Organism
      - Enzyme isa Biologically Active Substance
    - Associative
      - Sign or Symptom diagnoses Pathologic Function
      - Pharmacologic Substance treats Pathologic Function
  - 54 (relationship₁, isa, relationship₂)
    - treats isa affects
Semantic relations (subset)
Inheritance

Fully Formed Anatomical Structure

location of

Biologic Function

isa

Pathologic Function

isa

Disease or Syndrome

Body Part, Organ, or Organ Component

isa
Semantic relations (inherited)

◆ 6864 assertions
  ● 6752 (type₁, relationship, type₂)
    ■ Hierarchical
      – Mammal isa Organism (from: Animal isa Organism)
      – Enzyme isa Chemical
        (from: Enzyme isa Biologically Active Substance)
    ■ Associative
      – Body Part, Organ, or Organ Component
        location_of Disease or Syndrome
        (from: Fully Formed Anatomical Structure
         location_of Biologic Function)
  ● 112 (relationship₁, isa, relationship₂)
    – treats isa functionally_related_to (from: treats isa affects)
## Semantic relation information

<table>
<thead>
<tr>
<th>Sem. Type&lt;sub&gt;1&lt;/sub&gt;</th>
<th>Relationship</th>
<th>Sem. Type&lt;sub&gt;2&lt;/sub&gt;</th>
<th>Link status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Fully Formed Anatomical Structure</td>
<td></td>
<td>• D</td>
</tr>
<tr>
<td></td>
<td>• location_of</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Biologic Function</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
UMLs Semantic Network

How to use it?
Why a semantic network?

- Semantic Types serve as high level categories assigned to Metathesaurus concepts, *independently of their position in a hierarchy*

- A relation between 2 Semantic Types (ST) indicates a possible link between 2 concepts that have been assigned to those STs
  - The relationship may or may not hold at the concept level
  - Other relationships may apply at the concept level
Identify concepts for a given category

- Semantic type
  - List of all concepts having this semantic type

- Concept
  - List of all descendants
Semantics of Metathesaurus relations

Semantic Network

- Fully Formed Anatomical Structure
- Biologic Function
- Pathologic Function
- Disease or Syndrome
- Body Part, Organ, or Organ Component

Arrows:
- isa
- location of

Nodes:
- Adrenal Cortex
- Adrenal Cortical hypofunction

Metathesaurus
UMLS Semantic Network

Where to get it from?
UMLS distribution

- Part of the UMLS distribution
  - Installed by MetamorphoSys (option)
UMLS Semantic Network website

- No intellectual property restrictions on the Semantic Network
- Available outside UMLS license agreement
- UMLS Semantic Network website
  http://semanticnetwork.nlm.nih.gov/
The UMLS Semantic Network

The UMLS Semantic Network is one of three UMLS Knowledge Sources developed as part of the Unified Medical Language System project. The network provides a consistent categorization of all concepts represented in the UMLS Metathesaurus.

The UMLS Semantic Network is available in unit record and relational format and is distributed subject to these terms and conditions.

- Download the UMLS Semantic Network files

The UMLS semantic network reduces the complexity of the Metathesaurus by grouping concepts according to the semantic types that have been assigned to them. For certain purposes, however, an even smaller and coarser-grained set of semantic type groupings may be desirable. The UMLS semantic groups have been established for this purpose.

Further information about the UMLS Semantic Network can be found at:

- UMLS Semantic Network Fact Sheet
- UMLS Semantic Network Documentation


Updated: February 1, 2006

http://semanticnetwork.nlm.nih.gov/
Semantic Network files

The UMLS Semantic Network Files

The UMLS Semantic Network files are distributed subject to these terms and conditions.

Download the current Semantic Network files:

- The SRFL file (File Description)
- The SRFLD file (Field Description)
- The SRDEF file (Basic information about the Semantic Types and Relations)
- The SRSTR file (Structure of the Network)
- The SRSTRE1 file (Fully inherited set of Relations (UIs))
- The SRSTRE2 file (Fully inherited set of Relations (Names))
- The SU file (Unit Record)
Section 3

SEMANTIC NETWORK

3.0 Introduction

The Semantic Network consists of (1) a set of broad subject categories, or Semantic Types, that provide a consistent categorization of all concepts represented in the UMLS Metathesaurus, and (2) a set of useful and important relationships, or Semantic Relations, that exist between Semantic Types. This section of the documentation provides an overview of the Semantic Network, and describes the files of the Semantic Network. Sample records illustrate structure and content of these files.

3.1 Overview

The purpose of the Semantic Network is to provide a consistent categorization of all concepts represented in the UMLS Metathesaurus and to provide a set of useful relationships between these concepts. All information about specific concepts is found in the Metathesaurus. The Network provides information about the set of basic semantic types, or categories, which may be assigned to these concepts, and it defines the set of relationships that may hold between the semantic types. The Semantic Network contains 135 semantic types and 54 relationships. The Semantic Network serves as an authority for the semantic types that are assigned to concepts in the Metathesaurus. The Network defines these types, both with textual descriptions and by means of the information inherent in its hierarchies.

The semantic types are the nodes in the Network, and the relationships between them are the links. There are major groupings of semantic types for organisms, anatomical structures, biologic function, chemicals, events, physical objects, and concepts or ideas. The current scope of the UMLS semantic types is quite broad, allowing for the semantic categorization of a wide range of terminology in multiple domains.
UMLS Semantic Network

Future directions
Future directions

- Harmonize with other top-level ontologies of biomedicine
  - Semantic types (e.g., with BioTop)
  - Semantic relationships (e.g., with the OBO Relation Ontology, with relationships in GALEN)
**Overview**

The increasing need for advanced ontology-based knowledge management in the life sciences and the description of its basic ontological entities is generally being acknowledged but, up until now, the development of biological ontologies lacks adherence to foundational principles of ontology design.

In light of this we introduce **BioTop**, a description-logic-based upper ontology of the most important entities in the domain of molecular biology, founded upon the formal principles mentioned above.

**BioTop** was originally based on GENIA, an ontology applied for corpus annotation in text mining contexts. In contrast to this, a major desideratum for **BioTop** is to describe as many classes as possible in terms of necessary and sufficient conditions.

To further ensure ontological validity we employ BFO as the top layer for the actual **BioTop** classes.

We use OWL-DL as a formal knowledge representation language and can thus use a terminological reasoner for classification in order to maintain consistency during the ontology engineering and also to automatically infer its full multihierarchical structure.

**Users and Examples**

- BOOTStrep: Bootstrapping Of Ontologies and Terminologies STategic REsearch Project
- UMLS-Mapping: A mapping of the Universal Medical Language System (UMLS) Semantic Network (SN) to **BioTop**
- taxdemo: A small and demonstrative ontology of how biological taxa can be based on **BioTop** and described in the (submitted) ISMB paper (You can download the implementation or get either a HTML view or a graphical view of it.)
# OBO relations

Summary Table

<table>
<thead>
<tr>
<th>name</th>
<th>transitive</th>
<th>symmetric</th>
<th>reflexive</th>
<th>anti-symmetric</th>
<th>documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>is_a</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>part_of</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>integral_part_of</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>proper_part_of</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>located_in</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>contained_in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>adjacent_to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>transformation_of</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>derives_from</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>preceded_by</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>has_participant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>has_agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
<tr>
<td>instance_of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>View detailed summary</td>
</tr>
</tbody>
</table>

http://www.obofoundry.org/ro/
GALEN relations  *isIngredientOf*
Future directions

- Harmonize with other top-level ontologies of biomedicine
  - Semantic types (e.g., with BioTop)
  - Semantic relationships (e.g., with the OBO Relation Ontology, with relationships in GALEN)
- Align relationships between Metathesaurus and Semantic Network
affects
has direct substance
physically related to
spatially related to
metabolic site of
metabolizes
has physiologic effect
has direct morphology
functionally related to
has indirect morphology
temporally related to
physiologic effect
occurs in
has direct substance
uses
manifestation of
indicates
result of
conceptually related to
definitional manifestation of
interprets
has specimen procedure
occurs after
process of pathological process of
access instrument of
direct device of
indirect device of
procedure device of
may treat
may prevent
encodes gene product
causative agent of
cause of
exhibits
References

- McCray AT, Nelson SJ.  
  *The representation of meaning in the UMLS.*  

- McCray AT. 
  *An upper level ontology for the biomedical domain.* 

- McCray AT, Bodenreider O. 
  *A conceptual framework for the biomedical domain.* 
  In: Green R, Bean CA, Myaeng SH, editors. The semantics of relationships: an interdisciplinary perspective. 

http://semanticnetwork.nlm.nih.gov/