

# Data and Knowledge Representation Lecture 6



# Last Time We Talked About

- Medical Coding Systems
- UMLS

# Today We Will Talk About

- Major KR schemes
  - Semantic Network
  - Frame-based Representation
  - Production Rules

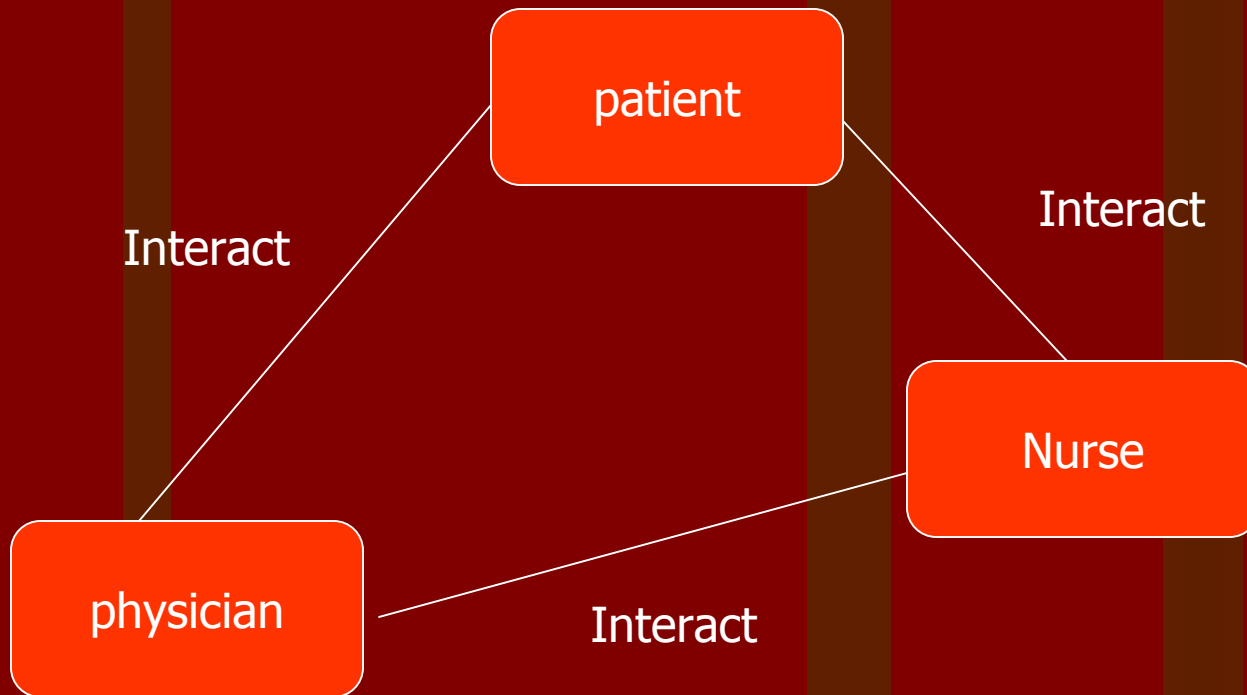
# Semantic Network

- An long existing notion: there are different pieces of knowledge of world, and they are all linked together through certain semantics.

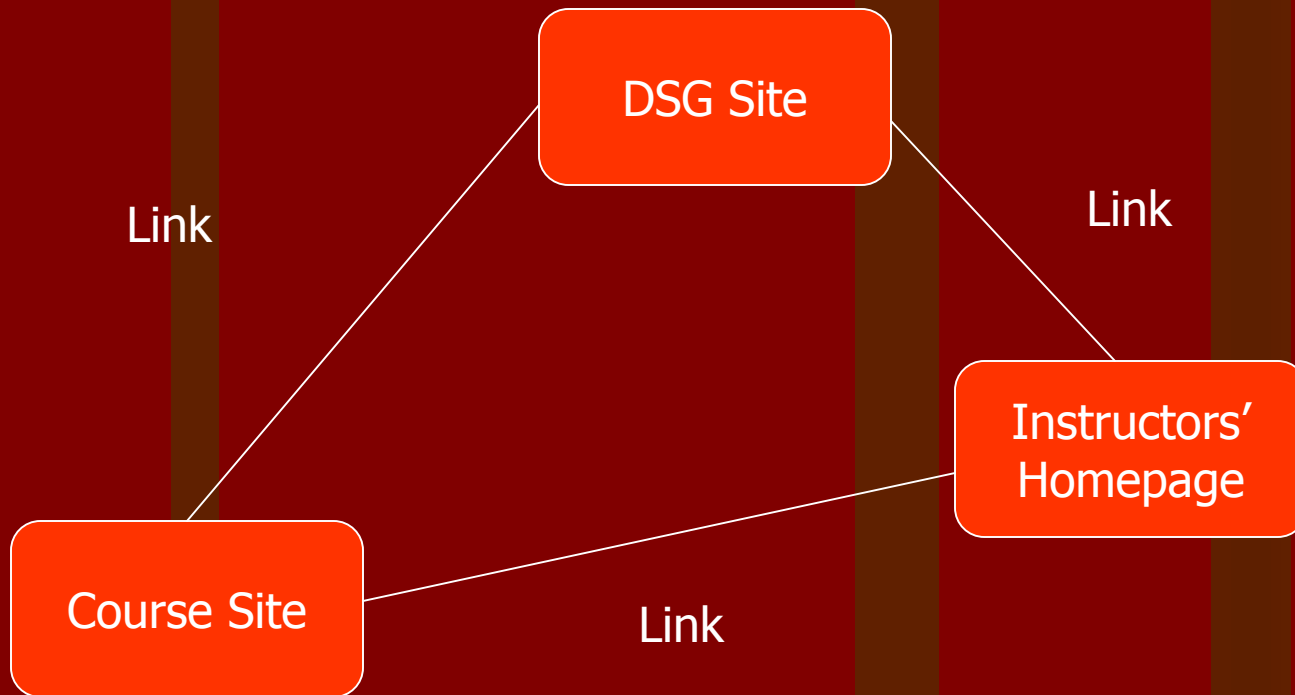
# Basic Components

- Nodes
  - Represent concepts
- Arcs
  - Represent relations
- Labels for nodes and arcs

# Little Constraint




# Little Constraint



Web

# Relation

- Directed or non-directed
  - Multiple relations between two concepts
  - Can have different properties
    - Reflexive (e.g. concurrence)
    - Transitive (e.g. causal)
    - Symmetric (e.g. sibling)
    - .....
- 



# Some Often Used Relations in Biomedical Domain

- **IS A**
- IS PART OF
- CAUSE OF
- MEASURES
- CO-OCCURS
- .....

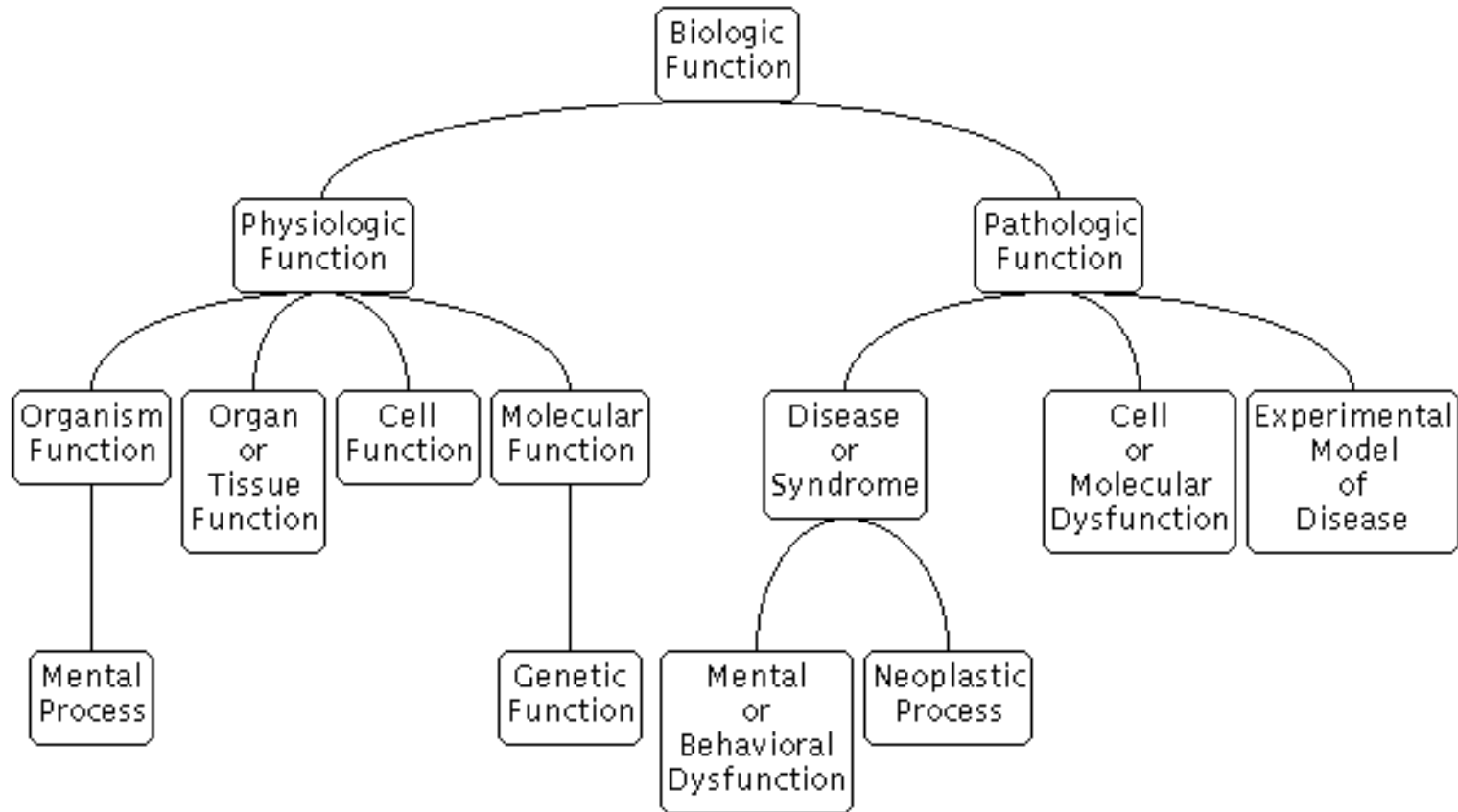
# Major Limitation

- Lack of Semantics
  - No formal semantic of the relations
    - E.g. Does "ISA" mean subclass, member, etc?
  - Possible multiple interpretations
  - Restricted expressiveness
    - E.g. can not distinguish between instance and class

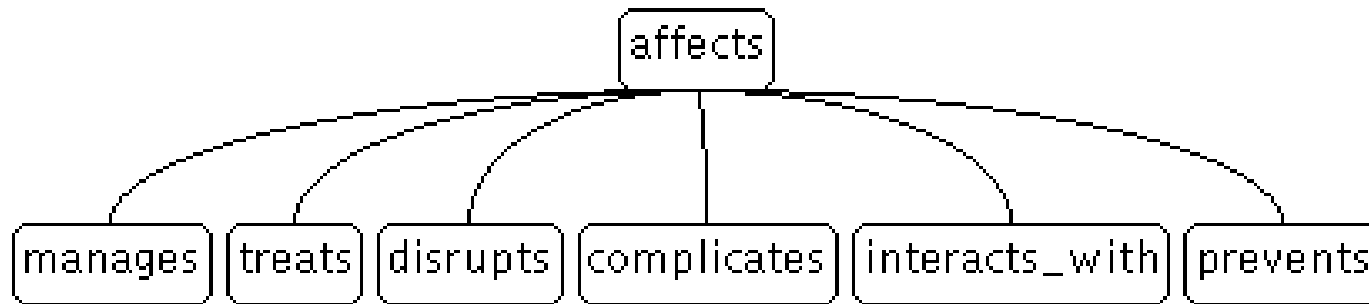
# Extension

- Extending expressivity (distinguish different types of concepts and relations")
  - Distinguish between "some" and "all"
  - Distinguish between "existence" and "intension"
  - Distinguish between "definition" and "assertion"
- Add semantic rigor
  - Map to logic (Sowa – CG)

# UMLS Semantic Network – Concept Hierarchy




# UMLS Semantic Network – Relation Hierarchy






# Frame-based Network

- Distinguish instance vs. class
  - Hierarchical structure (superclass and subclass)
  - Multiple hierarchy
  - Slots
    - Member slot
    - Own slot
- 

# Slot

- Frame identifying information
  - Relationship between frames
  - Descriptors of requirements for frame match
  - Procedural information
  - Default information
  - Restrictions and constraints
  - New instance information
- 



# Strength

- Help organize knowledge hierarchically
- Procedure information
- Support multiple inheritance

# Weakness

- Expressiveness (e.g. quantifier)
- Inheritance
  - Sub classing (override slot value)
  - Multiple inheritance
- Large complex knowledge system

# Example: MED

<http://www.cpmc.columbia.edu/homepages/dab7001/topics/index.html>



# Example: Protégé

<http://protege.stanford.edu/>

# Production Rules

- Also called IF-THEN rules
- Many forms:
  - IF condition THEN action
  - IF premise THEN conclusion
  - IF proposition  $p_1$  and proposition  $p_2$  are true THEN proposition  $p_3$  is true

# Components

- Rule base
- Inference engine
- Working memory

# Inference

- Modus ponens
  - Forward chaining
- Modus tollens
  - Background chaining

# Example: MYCIN


IF the identity of the germ is not known  
with certainty

AND the germ is gram-positive

AND the morphology of the organism is  
"rod"

AND the germ is aerobic

THEN there is a strong probability (0.8) that  
the germ is of type enterobacteriaceae





# Pro and Con

- Pro
  - Modular
  - Natural
- Con
  - Not efficient
  - Not expressive

# Exercise

*"Mrs Z.N. : 35 yr old farmer labourer from the Kwazulu-Natal province. She presented with a one week history of right sided pleuritic chest pain. The onset was sudden. This was associated with a warm flushing feeling and dizziness. There was also a pain in the right upper quadrant of her abdomen. There was no history of shortness of breath, cough or wheezing. There were no cardiac symptoms, no symptoms of malaise or loss of weight. No history of fevers. She was not a smoker and drank no alcohol. Of note is that she admitted to eating the entrails of the goats that she keeps." -*

<http://www.wits.ac.za/fac/med/pulmonology/case1.htm>

# Exercise

- Represent the information in semantic network
- Represent the information in frame-based network

# Exercise

- The thyroid gland is located at the base of your neck in front of your trachea (or windpipe). It has two sides and is shaped like a butterfly.
- The thyroid gland makes, stores, and releases two hormones - T4 (thyroxine) and T3 (triiodothyronine). Thyroid hormones control the rate at which every part of your body works. This is called your metabolism. Your metabolism controls whether you feel hot or cold or tired or rested. When your thyroid gland is working the way it should, your metabolism stays at a steady pace -not too fast or too slow.
- If no cancer cells are found, your doctor may prescribe a thyroid hormone to decrease the size of your nodule. Or, your doctor may suggest surgery to remove it. If cancer cells are found, further treatment will be needed. Thyroid cancer usually can be treated with success.

# Excise

- Which representation scheme to choose?



# Knowledge Representation Process

- Identify Needs
  - Conceptualization
  - Formalization
  - Implementation
  - Evaluation
- 

# Reading

- Sowa Chap. 4