

# Terminology Evaluation and Desiderata

Eneida A. Mendonca  
James J. Cimino  
Department of Biomedical Informatics  
Columbia University



# Desiderata

Go placidly amid the noise and haste, and remember what peace there may be in silence.

Speak your truth quietly and clearly; and listen to others, even the dull and ignorant; they too have their story.

If you compare yourself with others, you may become vain and bitter; for always there will be greater and lesser persons than yourself.

Exercise caution in your business affairs; for the world is full of trickery. But let this not blind you to what virtue there is; many persons strive for high ideals; and everywhere life is full of heroism.



# Why?

- Terminology is a fundamental requirement
  - Hospital and clinical systems
  - Literature databases
  - Bioinformatics databases
  - Expert systems
- Several roles for terminology
  - Representing local information
  - Mapping to standards and terminologies
  - Automated support for tools (e.g., DSS)



# Background

- Attempt to synthesize requirements expressed in the literature for the past decade
- Prepared for the IMIA Working Group 6 Jacksonville, Florida, January 1997
- Cimino JJ. Methods of Information in Medicine, 37(4/5), 1998



# Desideratum I: Content

- Must seek to be provide breadth and depth
- Structure must not limit size
- Relating atoms and molecules
  - White,conjunctiva, white conjunctiva
- Formal editorial policy and methodology



# Desideratum II: Concept Orientation

- Concepts, not terms
- Nonvagueness: at least one meaning per concept
- Nonambiguity: at most one meaning per concept
- Nonredundancy: one concept per meaning
- Explicit synonyms



# Example

- Difference between ambiguity of the meaning and ambiguity of the use
  - Myocardial infarction
    - Left ventricle MI
  - Myocardial Infarction
    - Rule out myocardial infarction
    - Family history of myocardial infarction
    - No myocardial infarction



# Desideratum III: Concept Permanence

- The meaning of a concept, once created, is inviolate
- Old concepts can't be deleted
  - Example: non-A-non-B hepatitis
- Names can be changed as long as meaning doesn't change (retronyms)
  - Example: pacemaker -> implantable pacemaker  
add transvenous pacemaker



# Desideratum IV: Nonsemantic Concept Identifiers

- Don't use the name
- Don't use a mnemonic
- Don't use a code that will run out of room
- Don't use a hierarchical code
- Meaningless integer (+/- check digit)

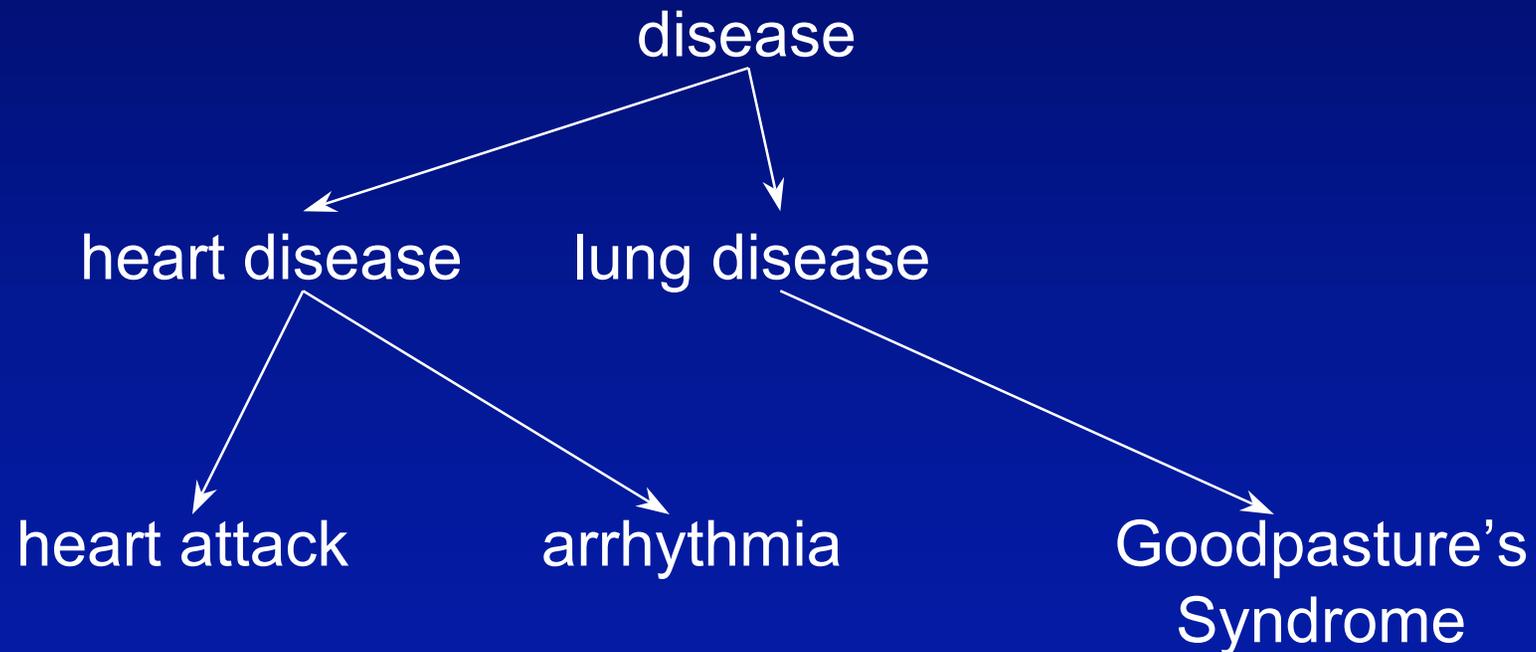


# Desideratum V: Polyhierarchy

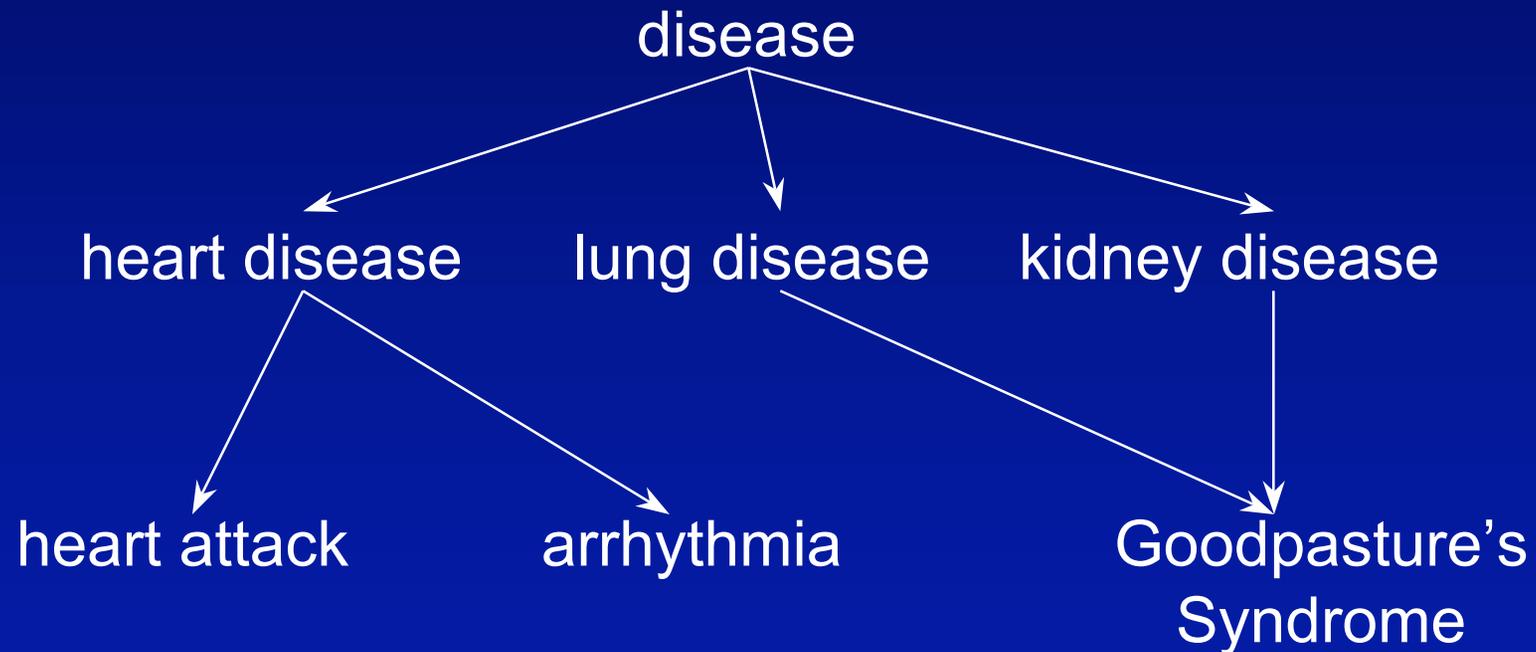
- Needed for understanding meaning
- Needed for tree walking
- Needed for inferencing
- Example: Goodpasture's Syndrome



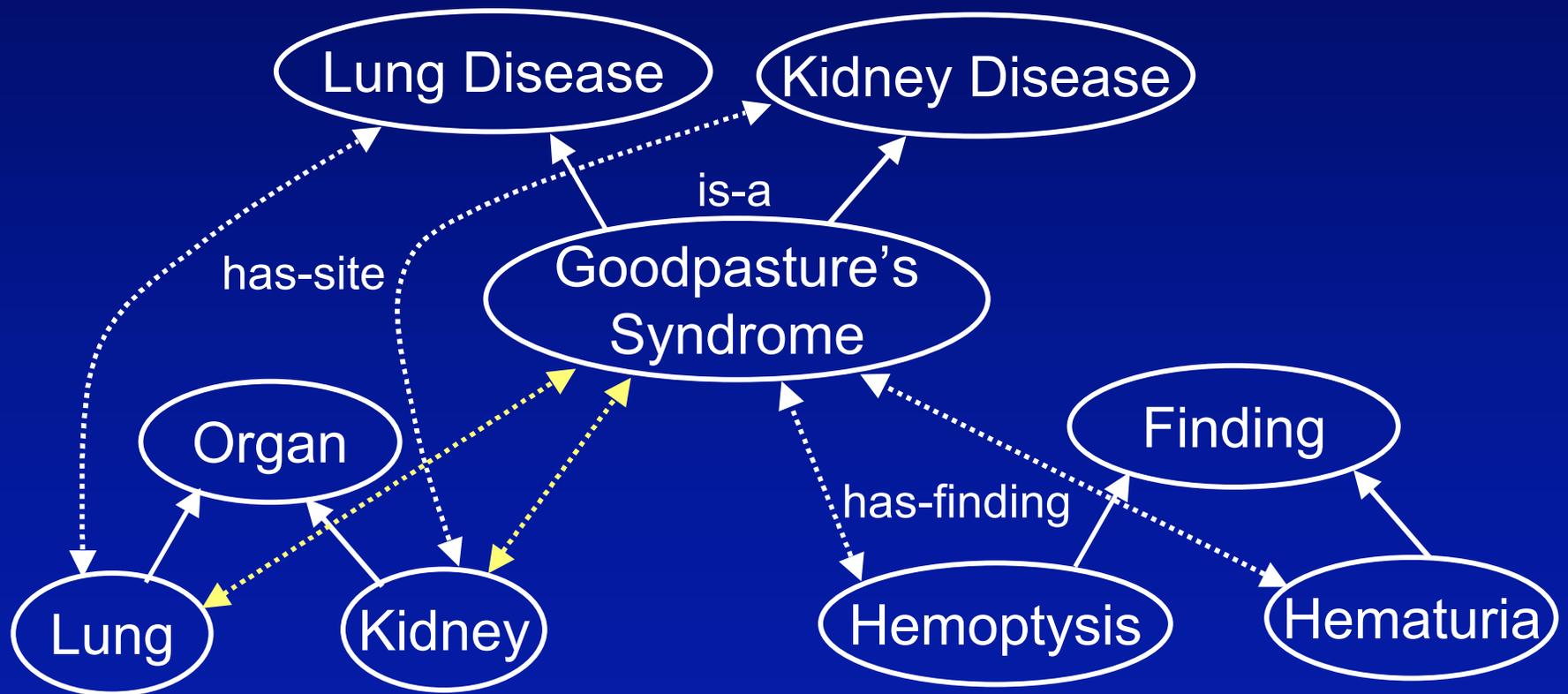
# Strict Hierarchy



# Polyhierarchy



# Semantic Network



# Desideratum VI: Formal Definitions

- Support understanding
- Support maintenance
- Structured and controlled (not narrative)
- Represented through relationships within the vocabulary
- Additional effort minimal and will pay off



# Desideratum VI: Formal Definitions

Serum Potassium Test: “a test that measures potassium ion concentration in the serum”

*Serum Potassium Test:*

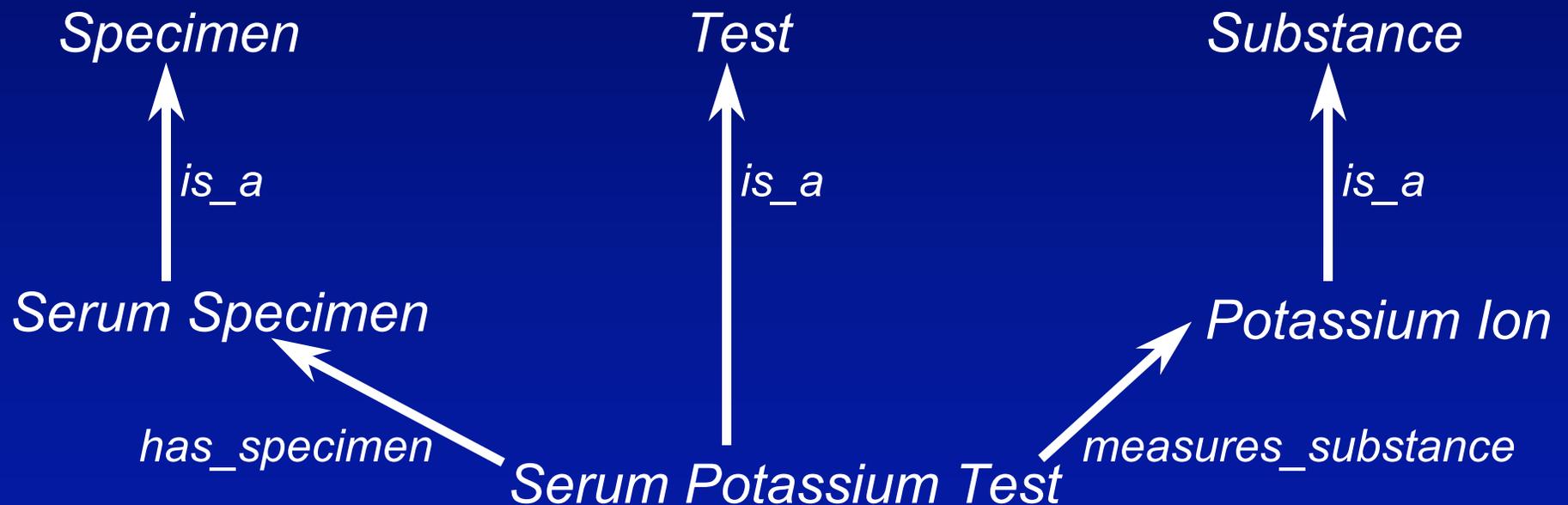
*is\_a: Test*

*has\_specimen: Serum Specimen*

*measures\_substance: Potassium Ion*



# Desideratum VI: Formal Definitions



~~NEEC~~

# Desideratum VII: Reject "Not Elsewhere Classified"

- Can never have a formal definition
  - Meaning by exclusion
- Terminology changes induce semantic drift
  - Problems interpreting historical data
- There are valid alternatives



# Desideratum VII: Reject "Not Elsewhere Classified"

- Old way:
  - “Pneumonia” coded as “Pneumonia, NOS”
  - “Informatician’s Pneumonia” coded as “Pneumonia, NEC”
- New way:
  - “Pneumonia” coded as “Pneumonia”
  - “Informatician’s Pneumonia” coded as “Pneumonia” + “Informatician”



# Desideratum VIII: Multiple Granularities

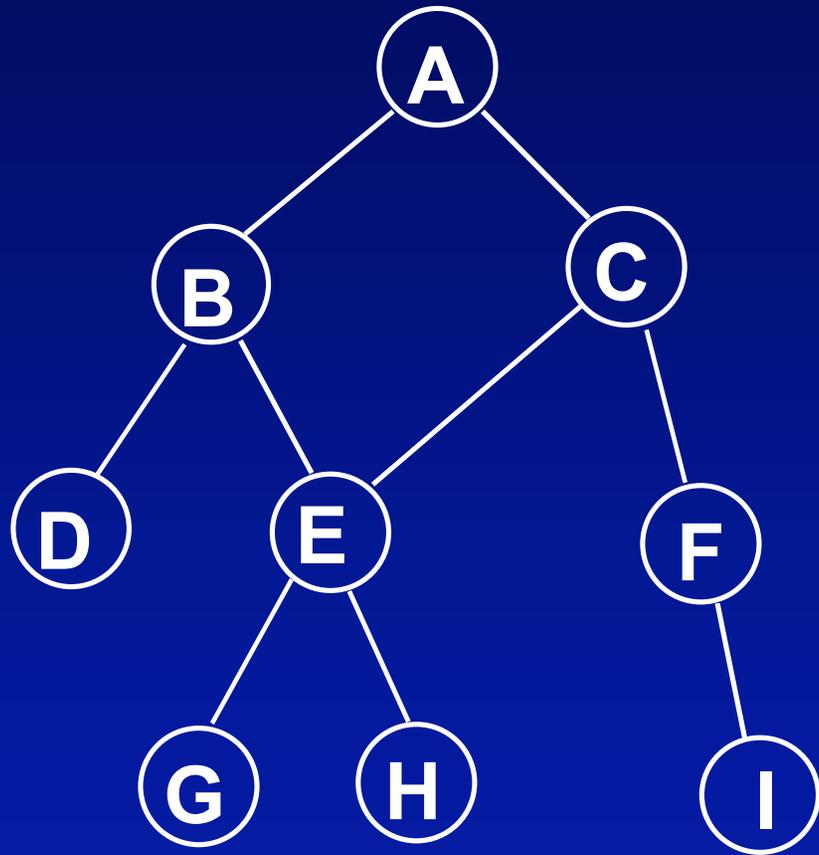
- Different levels for different purposes
- Uncertainty is allowed, imprecision is not (we must be precise about our uncertainty)

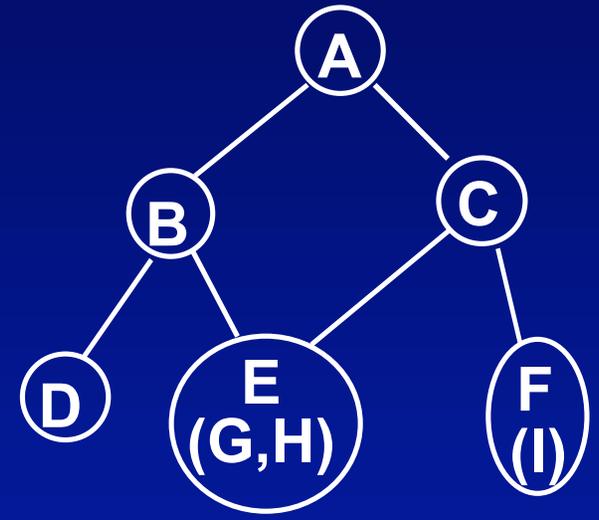
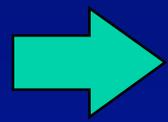
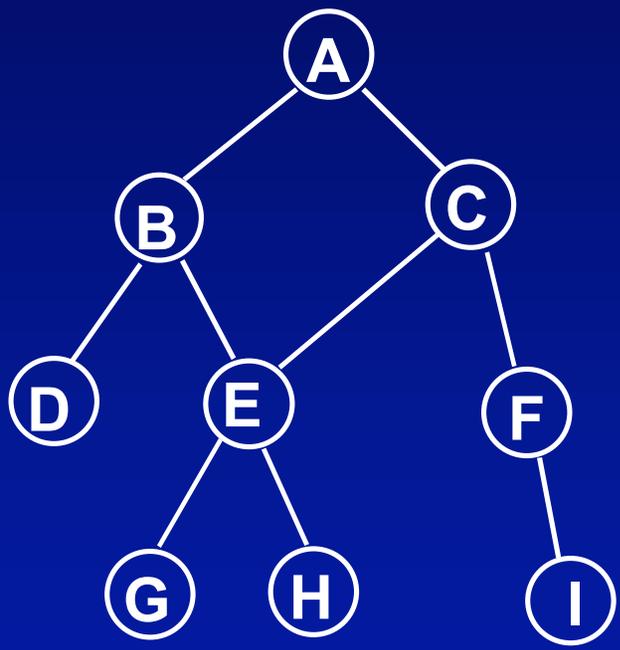


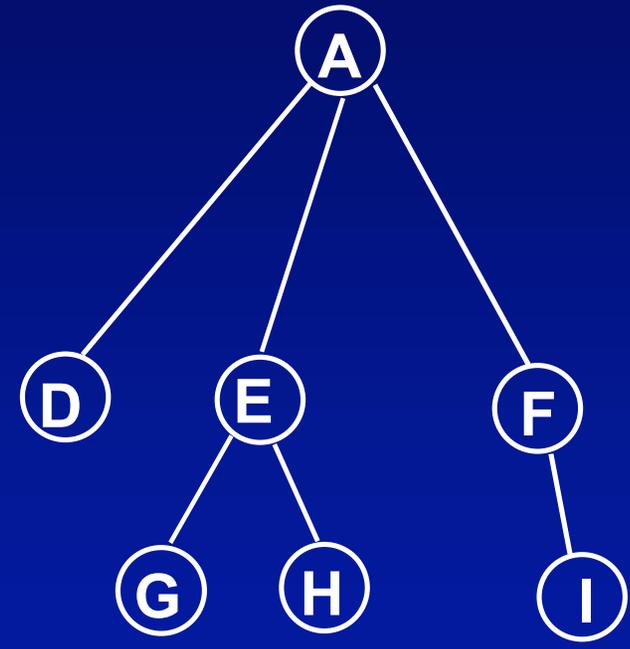
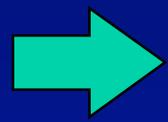
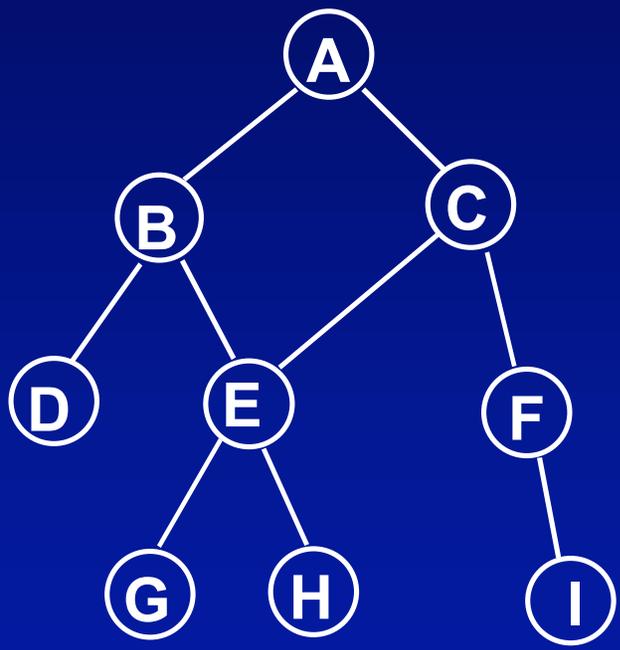
# Desideratum IX: Multiple Consistent Views

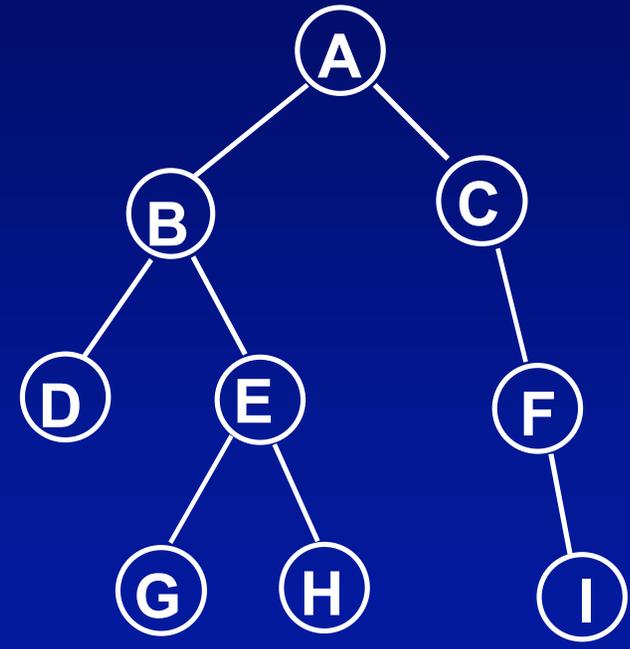
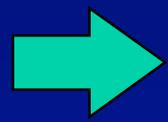
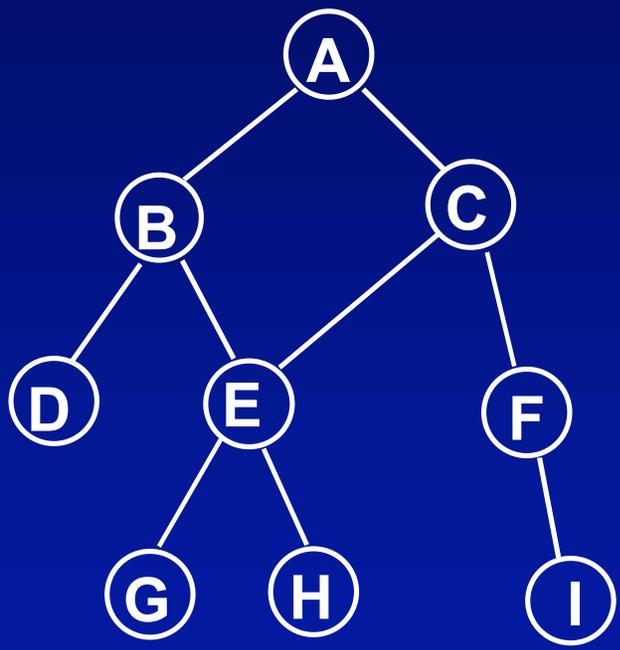
- Multiple views for multiple purposes
  - More or less granular? Strict hierarchy?
- Must not lead to inconsistency

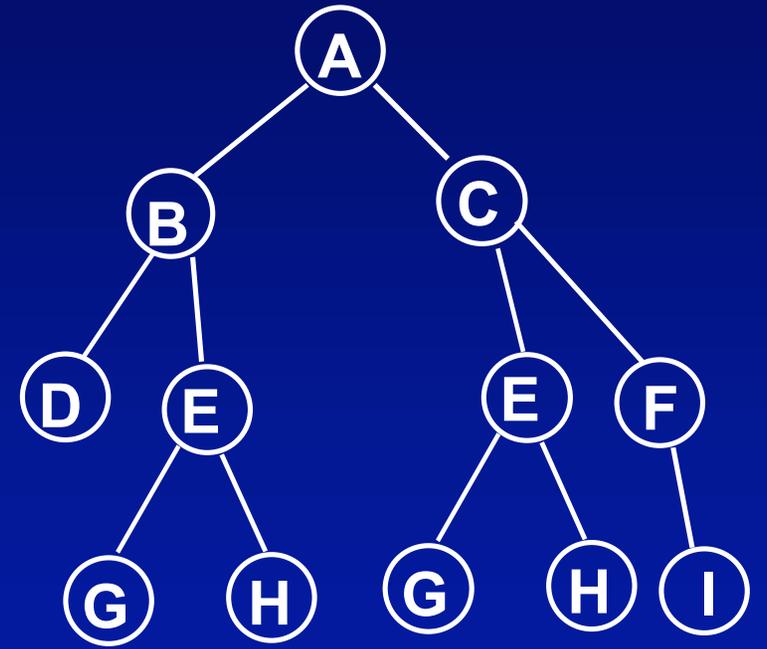
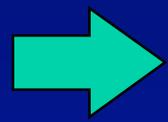
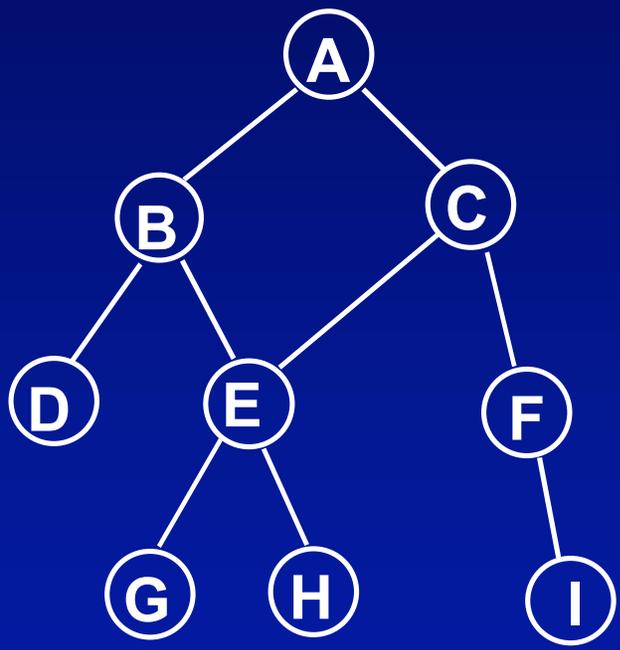


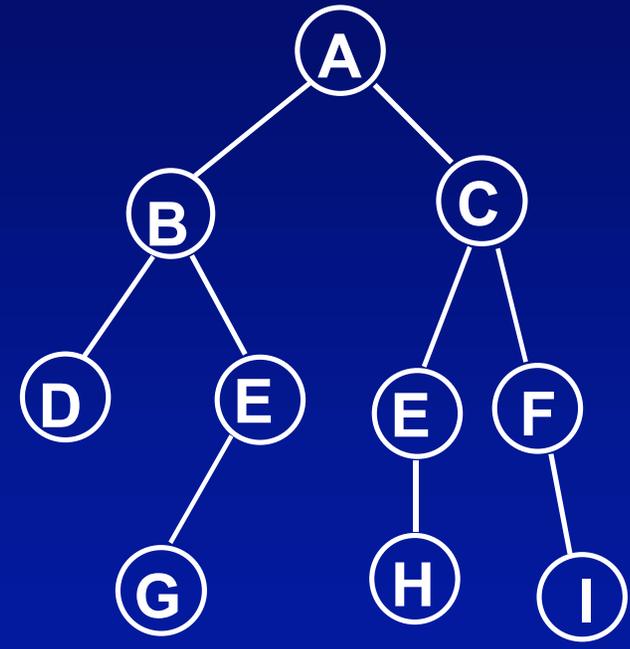
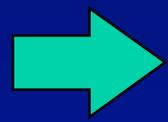
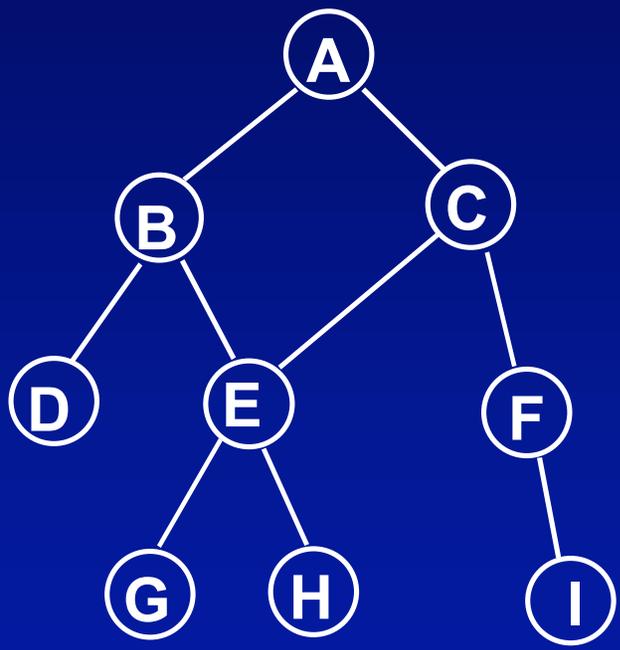


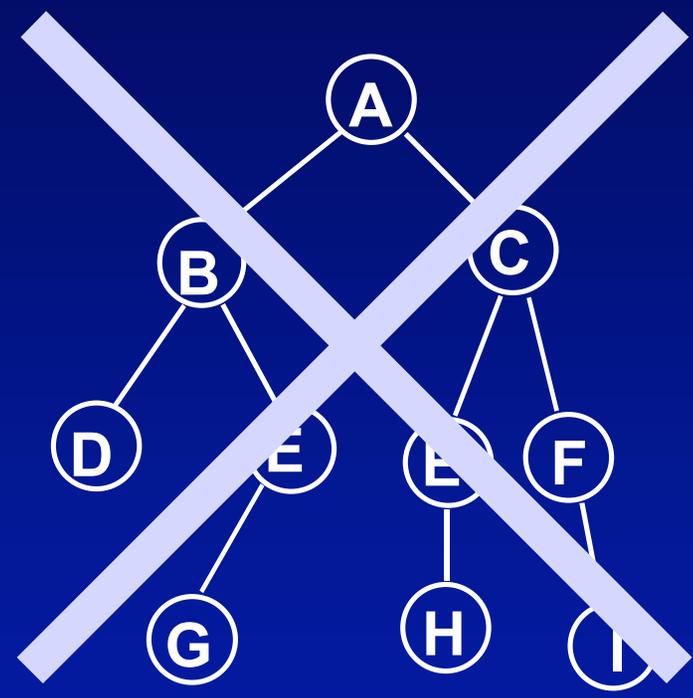
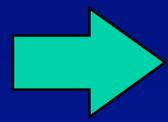
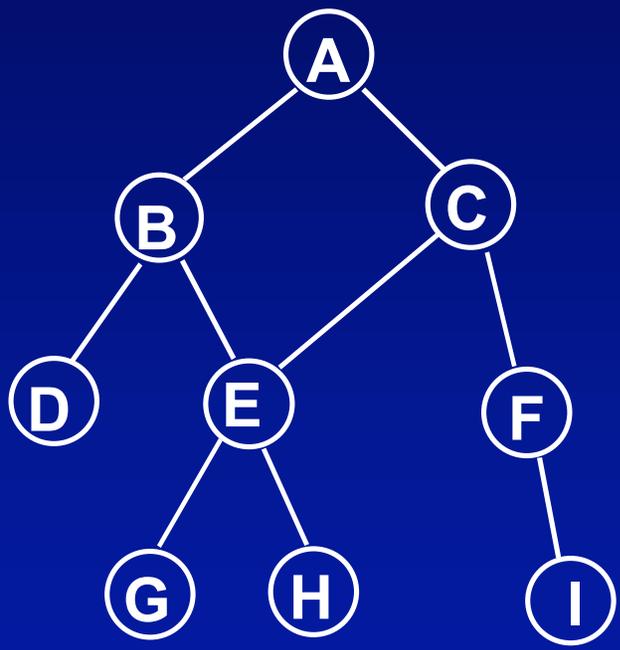












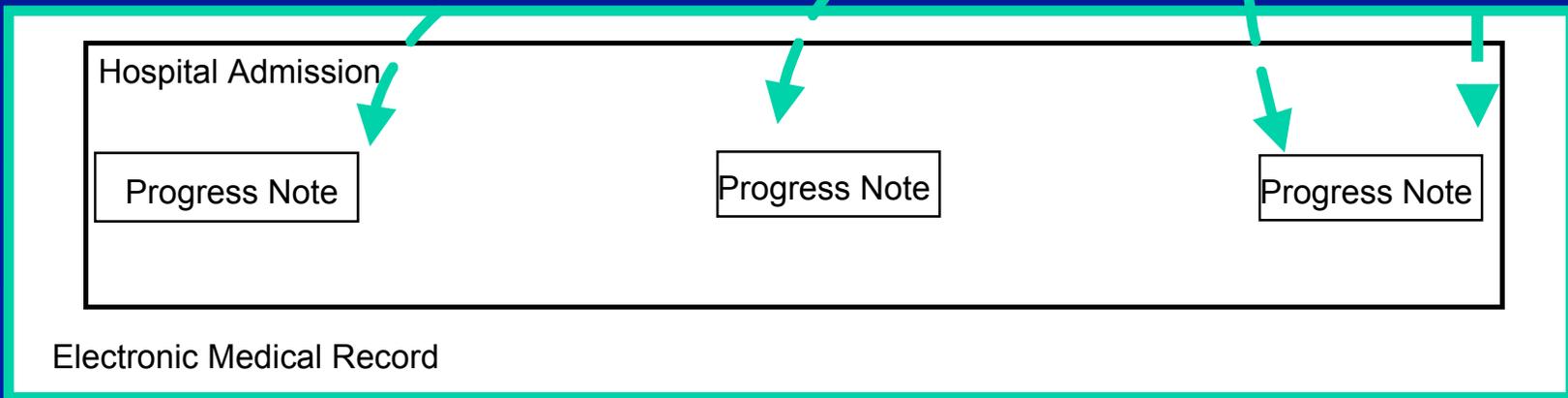
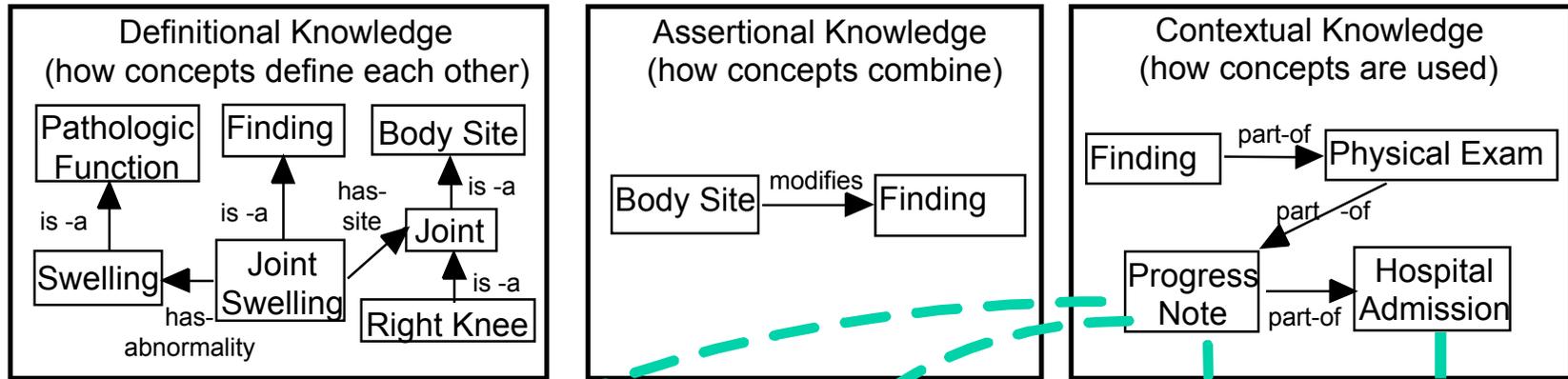
# Desideratum X: Representing Context

- Needed: a grammar to show usage
  - Avoid implicit assumptions based on the context
  - Descriptive vocabulary + rules for its use (model)
- "What is sensible to say"
- Consider modeling "events"



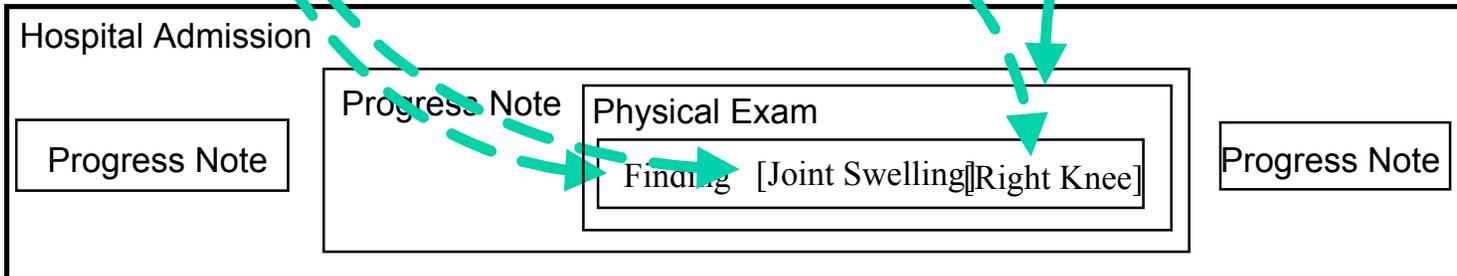
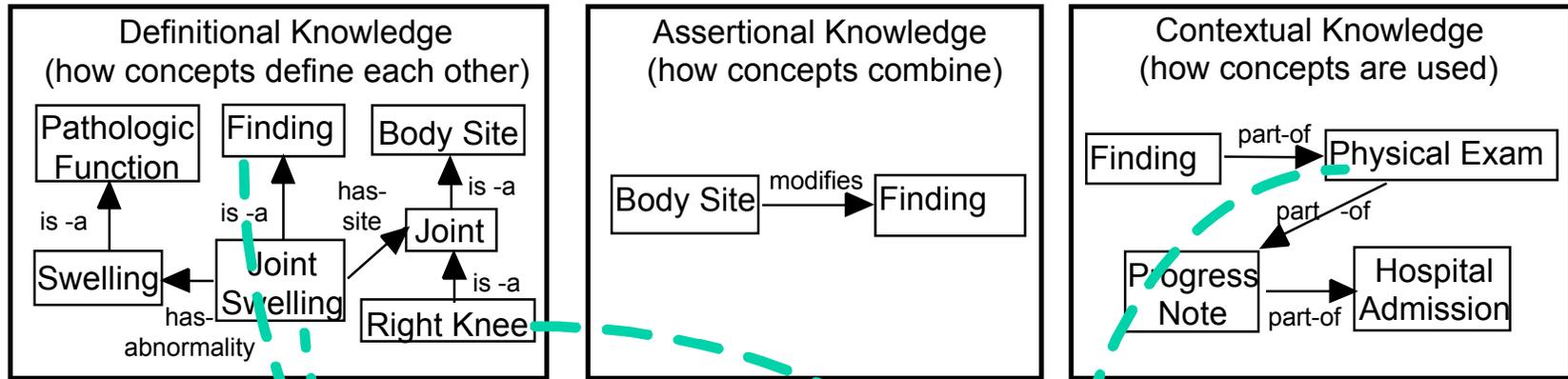
# Representing Context

## Controlled Medical Terminology



# Representing Context

## Controlled Medical Terminology



## Electronic Medical Record

# Desideratum XI: Graceful Evolution

- Will always need to fix mistakes
- Medical knowledge will grow
- Bad reasons
- Good reasons



# Desideratum XI: Graceful Evolution

- Bad reasons:
  - Redundancy
  - Major name changes
  - Code reuse
  - Changed codes



# Desideratum XI: Graceful Evolution

- Good reasons:
  - Simple addition
  - Refinement
  - Minor name changes
  - Precoordination
  - Disambiguation
  - Obsolescence
  - Discovered redundancy



# Desideratum XII: Recognize Redundancy

- Synonyms are good
- Redundant concepts are bad
- Redundant expressions are inevitable
- Example:
  - Year 1: "Pneumonia", "Left Lower Lobe"
  - Year 2: "Left Lower Lobe Pneumonia"

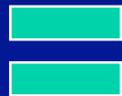


# Desideratum XII: Recognize Redundancy

Finding

Pneumonia

Left Lower Lobe



Left Lower Lobe Pneumonia  
is-a: Pneumonia  
has-site: Left Lower Lobe  
participates-in: Finding



# Status Report

I. Content

II. Concept Orientation

III. Concept Permanence

IV. Nonsemantic Concept Identifiers

V. Polyhierarchy

VI. Formal Definitions



# Status Report

- VII. Reject "Not Elsewhere Classified"
- VIII. Multiple Granularities
- IX. Multiple Consistent Views
- X. Representing Context
- XI. Graceful Evolution
- XII. Recognize Redundancy



# Practical Considerations

- Think “Concept Orientation”
- Have an editorial policy
- Consider explicit definitions
- Nonsemantic identifiers
- Include “is-a” attribute
- Pay attention to maintenance issues
- Include data dictionary in terminology
- Please: reject NEC!

