Data and Knowledge Representation Lecture 1

My Research

- Semantic Knowledge-based System
  - Information retrieval
  - Information integration/presentation
- Consumer Information Retrieval
- Flow Cytometry-based Proteomics
- Share Pathology Information Network

What I Intend to Cover

- Logic
- Medical Ontology

Main Textbook

- **Knowledge Representation: Logical, Philosophical, and Computational Foundations**
  by John F. Sowa

Motivation

- Representing data and knowledge for computing
  - Develop
  - Maintain
  - Share

Medical Data and Knowledge

- Large variety
- Many possible representations
- Implication of representation on computing

Example of Medical Data

This is a 51-year-old female admitted through the emergency room with syncopal episode with chest pain and also noted to have epigastric discomfort. The patient was admitted and started on Lovenox and nitroglycerin paste. The patient had serial cardiac enzymes and ruled out for myocardial infarction. The patient underwent a dual isotope stress test. There was no evidence of reversible ischemia on the Cardiolite scan. The patient has been ambulated. The patient had a Holter monitor placed but the report is not available at this time. The patient has remained hemodynamically stable. Will discharge.

http://www.angelfire.com/mt/nashrose/dsch.html

Examples of Medical Knowledge

- **Nitrates are a safe and effective treatment that can be used in patients with angina and left ventricular systolic dysfunction.**
- **On the basis of currently published evidence, amlodipine is the calcium channel antagonist that it is safest to use in patients with heart failure and left ventricular systolic dysfunction.**
- **Coronary artery bypass grafting may be indicated, in some, for relief of angina**
- **All patients with heart failure and angina should be referred for specialist assessment.**
- **Patients with angina and mild to moderately symptomatically severe heart failure that is well controlled, and who have no other contraindications to major surgery, should be considered for coronary artery bypass grafting on prognostic (as well as symptomatic) grounds.**

Challenge

- Philosophical difference
- Domain difference
- Application difference
- Developer difference
- Liability
- Cost

Formalism and Conceptualization

- Natural Language is the most expressive form of formalism and conceptualization
- Conceptualization is an abstract and simplified view of the world
- Such simplification allow computer and human alike to communicate in an unambiguous fashion (e.g. “and” vs. “&”)

Logic

- A tool for reasoning
- Provide basic concepts used in many computer science fields (AI, IR, DB, etc..)
- Used in many informatics applications

Propositional Logic

- Proposition
- Basic operators
- Language
- Truth table
- Boolean Algebra

Proposition

- A proposition is a symbolic variable whose value must be either True or False, and which stands for a natural language statement which could be either true or false
- Examples:
  - A = Smith has chest pain
  - B = Smith is depressed
  - C = It is raining

Operators

- Logic And
- Inclusive Or
- Exclusive Or
- Logic Not
- Logical Implication
- Logical Equivalence

Logical And $\wedge$

A	B	$A \wedge B$
False	False	False
False	True	False
True	False	False
True	True	True

Inclusive Logical Or (V)

A	B	A V B
False	False	False
False	True	True
True	False	True
True	True	True

Exclusive Logical Or ($\otimes$)

A	B	A $\otimes$ B
False	False	False
False	True	True
True	False	True
True	True	False

Inclusive vs. Exclusive

- Natural language “Or” can mean either
- Exclusive not often used (except in circuit design)

Medical Example

- "Heart AND Lung disease": does patients have to have both? Or either?
- "Foot AND mouth disease": what does "AND" mean in this case?
- Further reading: **Mendonca EA, Cimino JJ, Campbell KE, Spackman KA.** Evaluation of a proposed method for representing drug terminology. Proc AMIA Symp. 1999;:47-51.

Logical Not (¬)

A	¬A
False	True
True	False

Logical Implication ($\rightarrow$)

A	B	$A \rightarrow B$
False	False	True
False	True	True
True	False	False
True	True	True

Understanding "→"

- This is an operator. Although we call it "imply" or "implication", do not try to understand its semantic from the name. We could have called it "I" and still define its semantic the same way.
- A→B "means" A is sufficient, but not necessary to make B true.
  - E.g. Let A be "having cold" and B be "drink water", A → B can be interpreted as "should drink water" when "having cold". However, you can drink water even when you don't have cold. Thus A → B still is true when A is not true.

Logical Equivalence (↔)

A	B	A ↔ B
False	False	True
False	True	False
True	False	False
True	True	True

Understanding "→"

- A→B is different from A=B
  - A: a person is pregnant. B: a person is woman.
  - In this case, A→B is true, A=B is not.
- Use formal logic to represent knowledge of the real world, not the other way around.

Well-Formed Formulas

- Formula
  - A term (string) in prepositional logic
- Well-formed formula (WFF)
  - A term that is constructed correctly according to propositional logic syntax rules

WFF

- Constants: *False, True*
- Variables: *P, Q, R*
- If *a* is WFF, ¬*a* is WFF
- If *a* and *b* are WFF, *a*∧*b* are WFF
- If *a* and *b* are WFF, *a*∨*b* are WFF
- If *a* and *b* are WFF, *a*→*b* are WFF
- If *a* and *b* are WFF, *a*↔*b* are WFF
- Any formula that cannot be constructed using these rules are not WFF

Precedence of Logical Operators

- $\neg$
- $\wedge$
- $\vee$
- $\rightarrow$
- $\leftrightarrow$

Let Try An Example

- Order Test A for all male over 70, smokers with family history of cancer, and women with chronic cough and family history of cancer. Otherwise, do not order it.
  - Male: a person being male
  - Old: a person being over 70
  - Smoker: a person being a smoker
  - Cough: a person having chronic cough
  - FHC: a person having family history of cancer
  - OrderA: Order Test A

$(Male \wedge Old) \vee (Smoker \wedge FHC) \vee (\neg Male \wedge Cough \wedge FHC) \leftrightarrow OrderA$

Examples

- Smokers are those who are currently smoking or had quit smoking for less than 6 months
- A document is completed only after signed by both the chief resident and the attending physician.
- Smith is depressed whenever it rains

A Few Comments

- Use parentheses if precedence not clear
- Very similar to programming language operators' precedence
- Precedence in natural language depend more on context
  - E.g. "no heart and lung disease"
  - E.g. "no family history and healthy life style".

Truth Table

- An easy way to evaluate propositions

A	B	A ∨ B	¬B	(A ∨ B) ∧ ¬B
0	0	0	1	0
0	1	1	0	0
1	0	1	1	1
1	1	1	0	0

Let Try An Example

- Order Test A for all male over 70, smokers with family history of cancer, and women with chronic cough and family history of cancer. Other wise, do not order it.

(Male ∧¬Young) V (Smoker ∧ FHC) V (¬Male ∧ Cough ∧ FHC) ↔ OrderA

Male	Young(<=70)	Smoker	FHC	Cough	Order Test A
T	T	T	T	T	T
T	T	T	T	F	T
T	T	T	F	T	F
T	T	T	F	F	F
T	T	F	T	T	F
……					

Tautology and Contradiction

- Male V ¬Male
- Tautology: proposition that is always true
- Healthy Λ ¬Healthy
- Contradiction: proposition that is always false

Extra Reading

- Aho's book chapter 12
- Sowa's book p1-39

Homework