Knowledge Representation and Reasoning

CS227

Spring 2011

Outline

- Three Example Systems
- Goals / Design of the course
- Some Basic Definitions

Example Systems

- We will take a look at three implemented systems
  - Cognitive Assistant (SIRI)
  - Smart Textbook (Inquire)
  - Computational Knowledge Engine (Wolfram Alpha)
- For each system, we will look at
  - What knowledge must it represent?
  - What reasoning must it do?
  - What would it take to extend it?
  - Where does it fail?
  - How is it different from (current) Google?

Cognitive Assistant SIRI

- See Demo at: http://www.youtube.com/watch?v=MpjpVAB06O4&feature=player_embedded
- What knowledge must it represent?
  - Restaurants, movies, events, reviews, …
  - Location, tasks, web sources, …
- What reasoning must it do?
  - Nearest location, date for tomorrow, AM vs PM, etc
- What would it take to extend it?
  - More sources, different sources,
- Where does it fail?
  - Completely different environment, completely different task
- Differences from Google
  - Dialog driven, task-oriented, location aware, …

Smart Textbook Inquire

- What knowledge must it represent? (Demo in the class)
  - Concepts, definitions, relationships, descriptions
- What reasoning must it do?
  - Follow relationships, answer questions
- What would it take to extend it?
  - Must be customized to a new domain, must have methods for handling each kind of question
- Where does it fail?
  - Does not capture all the content in the book, limited forms of reasoning
- How is it different from Google?
  - Very specific domain targeted at a specific class of user situated in an educational context

Wolfram Alpha

- Try out examples at: http://www.wolframalpha.com/examples/
  - We will focus on the nutrition example
- What knowledge must it represent?
  - Different kinds of foods, their nutrition composition, caloric values
- What reasoning must it do?
  - Mathematical computations based on portions
- What would it take to extend it?
  - Add more data on foods and nutrition composition
- Where does it fail?
  - Does not know about recipes, how to combine foods, …
- How is it different from Google?
  - Data driven as opposed to document driven, mathematical reasoning

Goals of the Course

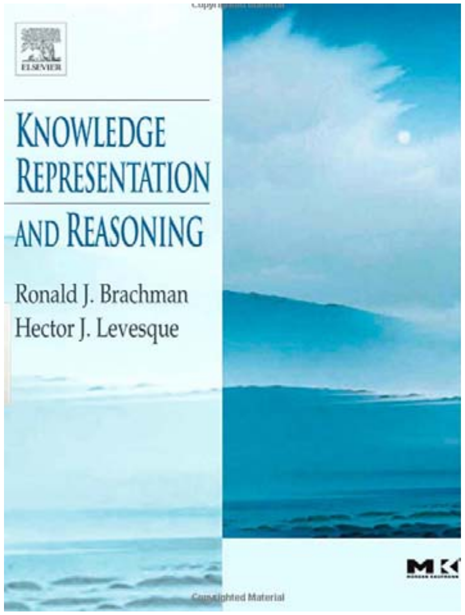
- Introduction to techniques used to represent symbolic knowledge
- Associated methods of automated reasoning

- The three systems that we saw
  - use symbolic knowledge representation and reasoning
  - But, they also use non-symbolic methods
    - Non-symbolic methods are covered in other courses (CS228, CS229, …)

- This course would be better labeled as a course on Symbolic Representation and Reasoning
  - The non-symbolic representations are also knowledge representations but are not covered in this course

Design of the Course

- Textbook:
  - Knowledge Representation & Reasoning by Brachman & Levesque (available online)
- Lectures
  - Tuesday and Thursday, 12:50-2:05, 300-300
- Grades
  - Four Assignments (40%), Mid-term (25%), Final (35%)
- Prerequisites
  - First order logic and Resolution (at the level of CS157)
    - There will be two tutorial sections to cover this material
    - The textbook chapters 2-4 provide adequate background
  - Discrete mathematics (data structures and algorithms)
  - A course in AI (knowledge of Lisp or Prolog)



Design of the Course

- Course website
  - http://cs227.stanford.edu
- Topics:
  - Object-oriented representation, description logics, ontologies, logic programming, constraint programming, action representation and reasoning, abstraction/reformulation/approximation
- Tests
  - Mid-term, Date: TBA
    - Will be held in the evening
    - Please let us know about any conflicts ASAP
  - Final, Date: TBA
- Staff mailing list
  - cs227-spr1011-staff@lists.stanford.edu
- Projects
  - Only with the approval of the instructor

Design of this Course

- Will this course require programming?
  - We will work with several off-the-shelf representation and reasoning tools
  - We will not be writing any new tools from scratch
  - The focus will be on applying representation techniques to real world knowledge and using existing tools to reason with that knowledge
  - Minor programming may be needed for some assignments

Relationship to Other Courses

- This course is a good follow up to
  - CS157: Computational Logic
  - CS221: Introduction to Artificial Intelligence
  - CS270: Modeling Bio-Medical Systems
- This course is complementary to:
  - CS228: Probabilistic Graphical Models
- This course can be followed by:
  - CS223: Rational Agency and Intelligent Interaction
  - CS224: Multi-agent systems
  - CS227B: General Game Playing
  - Application of techniques in your respective projects
  - Research opportunities in symbolic representation and reasoning

Mini Project

- Represent a chapter from a Biology textbook and answer the questions at the back of the book
  - It is high school level knowledge and each of us should know it
    - Develop confidence in approaching any domain with the formal tools you will learn in this course
  - Primary focus on representation and reasoning
  - Provides natural progression:
    - one question, multiple questions, novel questions
    - Structured representations, inference rules, special purpose reasoners

Outline

- ✓ Three Example Systems
- ✓ Goals / Design of the course
- Some Basic Definitions

What is knowledge?

Easier question: how do we talk about it?

We say "John knows that ..." and fill the blank with a <u>proposition</u>

– can be true / false, right / wrong

Contrast: "John fears that ..."

– same content, different attitude

Other forms of knowledge:

- know how, who, what, when, ...
- sensorimotor: typing, riding a bicycle
- affective: deep understanding

Belief: not necessarily true and/or held for appropriate reasons

and weaker yet: "John suspects that ..."

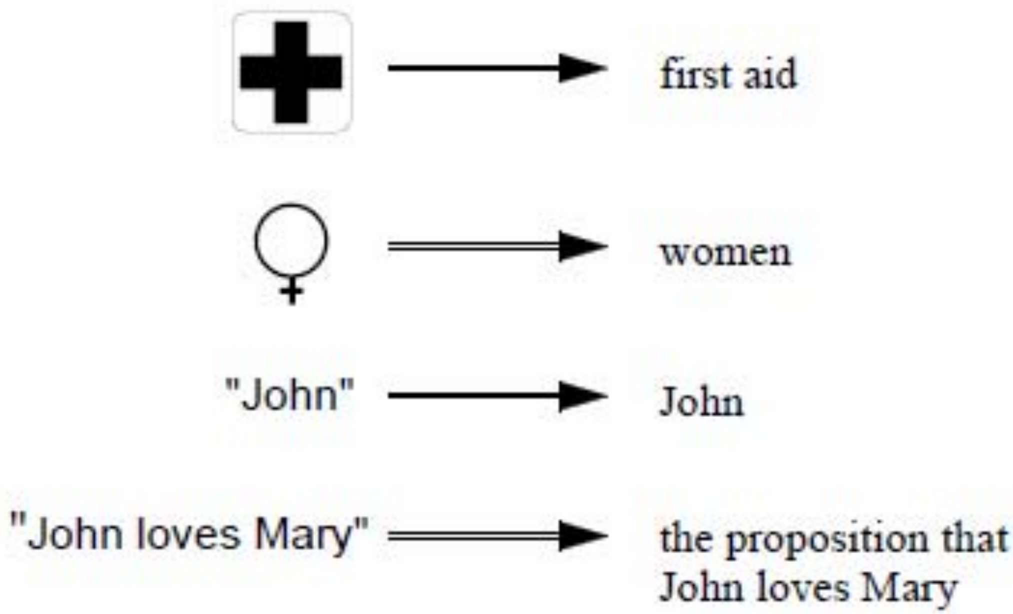
Here: no distinction

the main idea

taking the world to be one way and not another

What is representation?

Symbols standing for things in the world



Knowledge representation:

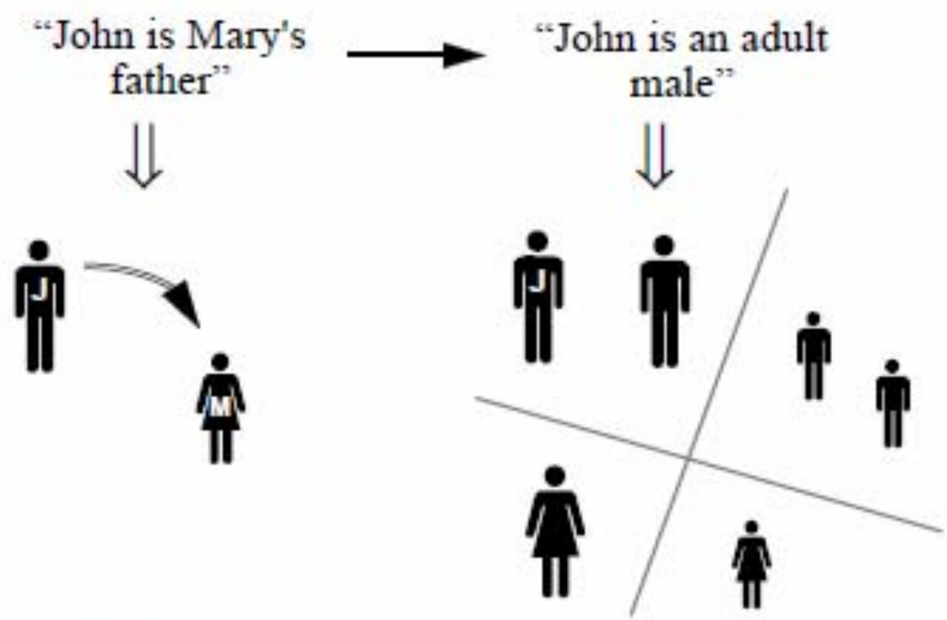
symbolic encoding of propositions believed
(by some agent)

What is reasoning?

Manipulation of symbols encoding propositions to produce representations of new propositions

Analogy: arithmetic

"1011" +	"10"	→ "1101"
⇓	⇓	⇓
eleven	two	thirteen



Why KR&R?

- KR Hypothesis (Brian Smith)
  - Any mechanically embodied intelligent process will be comprised of structural ingredients that
    - We as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits
    - Independent of such external semantic attribution, play a formal but causal and essential role in engendering the behavior that manifests that knowledge
- Two issues: existence of structures that
  - We can interpret
  - Determine how the system behaves

Two examples

Example 1

```
printColour(snow) :- !, write("It's white.").
printColour(grass) :- !, write("It's green.").
printColour(sky) :- !, write("It's yellow.").
printColour(X) :- write("Beats me.").
```

Example 2

```
printColour(X) :- colour(X,Y), !,
        write("It's "), write(Y), write(".").
printColour(X) :- write("Beats me.").

colour(snow,white).
colour(sky,yellow).
colour(X,Y) :- madeof(X,Z), colour(Z,Y).
madeof(grass,vegetation).
colour(vegetation,green).
```

Only the 2nd has a separate collection of symbolic structures à la KR Hypothesis

its knowledge base (or KB)

∴ a small knowledge-based system

Benefits of Explicit Representation

- We can add new tasks and easily make them depend on previous knowledge
  - Enumerating objects vs painting objects
- Extend the existing behavior by adding new beliefs
  - Assert that canaries are yellow
- Debug faulty behavior by locating the erroneous beliefs
  - By changing the color of sky we change any routine that uses that information
- Explain and Justify the behavior of the system
  - The program did X because Y

Benefits of Reasoning

- Given
  - Patient X allergic to medication M
  - Anyone allergic to medication M is also allergic to medication M’
- Reasoning helps us derive
  - Patient X is allergic to medication M’

Entailment

Sentences $P_1, P_2, \ldots, P_n$ entail sentence P iff the truth of P is implicit in the truth of $P_1, P_2, \ldots, P_n$.

If the world is such that it satisfies the P_i then it must also satisfy P.

Applies to a variety of languages (languages with truth theories)

Inference: the process of calculating entailments

- sound: get only entailments
- complete: get all entailments

Sometimes want unsound / incomplete reasoning

for reasons to be discussed later

Logic: study of entailment relations

- languages
- truth conditions
- rules of inference

Using logic

No universal language / semantics

- Why not English?
- Different tasks / worlds
- Different ways to carve up the world

No universal reasoning scheme

- Geared to language
- Sometimes want "extralogical" reasoning

Start with first-order predicate calculus (FOL)

- invented by philosopher Frege for the formalization of mathematics
- but will consider subsets / supersets and very different looking representation languages

KR&R and AI

- KR&R started as a field in the context of AI research
  - Need explicitly represented knowledge to achieve intelligent behavior
    - Expert systems, language understanding, …
- Many of the AI problems today heavily rely on statistical representation and reasoning
  - Speech understanding, vision, machine learning, natural language processing
    - For example, the recent Watson system relies on statistical methods but also uses some symbolic representation and reasoning
- Some AI problems require symbolic representation and reasoning
  - Explanation, story generation
  - Planning, diagnosis
  - Abstraction, reformulation, approximation
  - Analogical reasoning
- KR&R today has many applications outside AI
  - Bio-medicine, Engineering, Business and commerce, Databases, Software engineering, Education

Some Long-Term Problems that need Knowledge Representation

- Read a chapter in a textbook and answer questions at the end of the chapter
- Einstein in a box: The quality of reasoning that distinguishes an ordinary human from a top scientist
  - Answer the same questions as a national academy of science member
- Learn how to repair a mobile robot and successfully demonstrate the capability by repairing one on Mars
- Encyclopedia on Demand
  - Produce a 5000 word or less encyclopedia style article on a given subject by summarizing from the relevant information available on the web in less than 24 hours

Suggested Readings

- Required Reading
  - Chapter 1 of Brachman & Levesque textbook
  - Chapters 2-4 if you do not have prior background in FOL
- Optional Readings
  - Three Open Problems in AI. Raj Reddy. In the Journal of ACM, Vol 50, No. 1, 2003.
  - Some Challenges and Grand Challenges for Computational Intelligence. Edward A. Feigenbaum. In the Journal of ACM, Vol 50, No. 1, 2003.
  - Systems that Know What they're Doing. Ron Brachman. Intelligent Systems, Vol 17, no. 6, pp 67-71.