Knowledge Representation and Reasoning

Administrivia

- This class will (almost always) meet twice a week. You are expected to attend all class meetings.
  - I will be taking attendance in class, but only to get an idea about who is regularly attending.
- You are expected to work individually on all assignments and submit them on time.
- Participation in class discussion is encouraged and rewarded.

Course Assessment

- Your grade in this course will be based on class participation, one exam, almost-weekly assignments, and a term paper or final project.
- Weights:
  - Participation: 5%
  - Assignments: 15%
  - Final Exam: 40 %
  - Term Paper: 40 %

KRR: An Introduction

Lecture 1

January 30, 2018

What is KRR?

- Common wisdom has it that KRR is a sub-field of AI.
- It is concerned with the study of different ways of representing (purported) information in computers and drawing conclusions from it.
- Here are some of the issues involved.
  - The design of formal languages to represent information.
  - The study of properties of such languages.
  - The design of reasoning methods for drawing conclusions.
  - The study of properties of such methods.
  - The design of theories that provide information about particular domains.
    - Typically, theories about mundane domains such as time, space, action, and causality are the most challenging.

Why KRR?

- The girl with a flower.
- The moving of blocks.
- The busy street.
- The straw in the water. (Thanks to Nasr Kasrin.)
- The checkmate.

For each of the above, knowledge needs to be

1. stored (and, hence, represented) and
2. used to create/explicate further knowledge.

KRR Systems

- A KRR system consists of two *main* components:
  1. A **knowledge base** (KB).
  2. An **inference engine.**
- The KB is a set of well-formed formulas (**WFFs**) of an appropriate formal language.
- The inference engine is a set of algorithms used to derive new WFFs from the WFFs in the KB.
  - Typically, those conclusions are added to the KB.
- Two functions, generically called **Ask** and **Tell**, are used to query the KB and to add new WFFs to it, respectively.
- A KRR system may also have a **truth maintenance system** as an optional component.

KRR vs. Database

- The field of KRR strongly overlaps that of *deductive* databases.
- The difference is three-fold:
  1. **Application:** A KRR system is typically a component of a larger AI system. (A robotics system, a natural language understanding system, etc.)
  2. **Focus:** A KRR system is more concerned with knowledge of mundane domains and with commonsense reasoning patterns.
     - Mundane domains are not as well-defined as typical database domains.
  3. **Approach:** KRR theories have deep roots in philosophy and linguistics (and sometimes psychology).

The Different Senses of "know"

- Need to distinguish different senses of "know":[a]
  - I know that whales are mammals.
  - I know Stuart Shapiro.
  - I know the author of "Animal Farm".
  - I know this place!
  - I know how to ride a bicycle. **(This is an example of procedural knowledge.)**
  - I know whether whales are mammals.
- KRR focuses on **propositional knowledge**:

$$x \text{ knows that } p$$

[a] Doing KRR, we often get involved in semantics. KRR and semantics are, at some level, indistinguishable.

What is (Propositional) "Knowledge"?

- Common philosophical reply: *Knowledge is justified true belief.*
- That is, "x knows that p" is true iff
  1. p is true.
  2. x believes that p.
  3. x is justified in believing that p.
- Give examples to show that each of the above conditions is necessary for knowledge.
- Are they jointly sufficient? See the link to Gettier's paper.

But . . .

- This definition is too restrictive.
- In practice, what we do is closer to "belief representation".
  - But more on this later.
- In particular, recall that an important aspect of KRR is coming up with *theories* of mundane domains.
  - Such theories are typically philosophical, and cannot be proven *true* in any strict sense.

What is "Representation"?

- A representation is a 1-1 correspondence between two sets (with some qualifications).
- The domain constitutes the set of entities to be represented.
- The range is a set of entities interpreted as symbols, each **denoting** a particular element of the domain.
- Different KRR theories differ with respect to what they take themselves to be representing.
  - They may directly represent (a subset of) "the world".
  - They may represent the "mind" of a cognitive agent (and, hence, the conceptualization of the world therein).
- The difference is philosophical but practically significant.
  - For example, the beliefs of an agent could be contradictory; the world cannot.

What is "Reasoning"?

- Reasoning is the activity of arriving at correct conclusions.
- It involves at least two things:
  1. Inference.
  2. Explicating implicit knowledge needed for plausible conclusions to follow.
- **Example**: From "All students took the exam", it is natural to conclude that "Someone took the exam". But why?

Types of Reasoning

- Reasoning comes in different types:
  - Deductive.
  - Abductive.
  - Inductive.
  - Analogical.
- In this course, we shall focus on deductive reasoning (since most KRR systems are deductive).

A Taste of KRR

- Read the following passage carefully.

  Adam felt very hungry. He walked into the kitchen, and saw a glass jar on the table. The jar was full of cookies that Adam liked. He grabbed the lid, but since it was loose, the jar fell on the floor. A small piece of glass caused his shin to bleed.

A Taste of KRR (Cont'd)

- Now, answer the following questions.
  - Why did Adam walk into the kitchen?
  - Was there a jar in the kitchen?
  - Was there a table in the kitchen?
  - Was the jar made of glass?
  - Did Adam like to eat cookies?
  - Why did the jar fall on the floor?
  - Did the jar break?
  - Did Adam break the jar?
  - Was Adam injured?

On the Necessity of KRR to AI

- Two views:
  1. **Necessary:** Brian Cantwell Smith's **knowledge representation hypothesis.**

     Any mechanically embodied intelligent process will be comprised of structural ingredients that a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits, and b) independent of such external semantical attribution, play a formal but causal and essential role in engendering the behavior that manifests that knowledge.

  2. **Not necessary:** Rodney Brooks's anti-hypothesis.

     Representation is the wrong unit of abstraction in building the bulkiest parts of intelligent systems.

Next Time . . .

- Propositional logic.