Introduction: Knowledge Representation and Reasoning

Yuliya Lierler

University of Nebraska at Omaha

January 14, 2013

Some slides are due to Marc Denecker – course on Complex system modeling
Knowledge representation and Reasoning

- KRR: the formal science of knowledge and its use for solving problems
- a subdiscipline of AI
Some examples of knowledge

Examples of critical pieces of knowledge, in the context of specific tasks.

▶ In a tool to support a student to choose his study-program for the Master Informatics:
  ▶ A student chooses 1 or 2 from the 7 possible specialisations.
  ▶ The compulsory courses of a chosen specialisation must be selected.
  ▶ The total number studypoints of selected courses is between 55 and 65.

▶ In a tool to warn the nurse in intensive care that supervision of a patient is necessary:
  ▶ At latest 1 hour after a glycemia measurement of less than 40mgr, a new measurement must be carried out or an alarm is generated.
KR and AI

Much of AI involves building systems that are knowledge-based ability derives in part from reasoning over explicitly represented knowledge
– language understanding,
– planning,
– diagnosis,
– “expert systems”, etc.

Some, to a certain extent game-playing, vision, etc.

Some, to a much lesser extent speech, motor control, etc.
Why knowledge?

For sufficiently complex systems, it is sometimes useful to describe systems in terms of beliefs, goals, fears, intentions.

  e.g. in a game-playing program

  “because it believed its queen was in danger, but wanted to still control the center of the board.”

  more useful than description about actual techniques used for deciding how to move

  “because evaluation procedure P using minimax returned a value of +7 for this position

= taking an intentional stance (Dan Dennett)
Wikipedia:
The intentional stance (Daniel Dennett): the level of abstraction in which we view the behavior of a thing in terms of mental properties.

"Here is how it works: first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in most instances yield a decision about what the agent ought to do; that is what you predict the agent will do." (Daniel Dennett, The Intentional Stance, p. 17)
Why bother?

Why not “compile out” knowledge into specialized procedures?
  • distribute KB to procedures that need it
  • almost always achieves better performance

No need to think. *Just do it!*
  – riding a bike
  – driving a car
  – playing chess?
  – doing math?
  – staying alive??

Skills (Hubert Dreyfus)
  • novices think; experts *react*
  • compare to current “expert systems”:
    knowledge-based!
Advantage

Knowledge-based system most suitable for open-ended tasks

can structurally isolate reasons for particular behaviour

Good for

• explanation and justification
  – “Because grass is a form of vegetation.”

• informability: debugging the KB
  – “No the sky is not yellow. It's blue.”

• extensibility: new relations
  – “Canaries are yellow.”

• extensibility: new applications
  – returning a list of all the white things
  – painting pictures

Yuliya Lierler

Introduction: Knowledge Representation and Reasoning
Cognitive penetrability

Hallmark of knowledge-based system:

the ability to be *told* facts about the world and adjust our behaviour correspondingly

for example: read a book about canaries or rare coins

Cognitive penetrability (Zenon Pylyshyn)

actions that are conditioned by what is currently believed

an example:

we normally leave the room if we hear a fire alarm
we do not leave the room on hearing a fire alarm
if we believe that the alarm is being tested / tampered
    can come to this belief in very many ways
so this action is cognitively penetrable

a non-example:

blinking reflex
Why reasoning?

Want knowledge to affect action

\[ \text{not} \quad \text{do action } A \text{ if sentence } P \text{ is in KB} \]

\[ \text{but} \quad \text{do action } A \text{ if world believed in satisfies } P \]

Difference:

\[ P \text{ may not be } \text{explicitly} \text{ represented} \]
Need to apply what is known in general
to the particulars of a given situation

Example:

“Patient \( x \) is allergic to medication \( m \).”

“Anybody allergic to medication \( m \) is also
allergic to \( m' \).”

Is it OK to prescribe \( m' \) for \( x \)?

Usually need more than just DB-style retrieval of facts in the KB
Knowledge exists!

- Human experts have it and use it to accomplish a task or solve a problem.
- Human programmers that write a program for a task or problem, use it and “compile” it into their algorithm and program.

We want to understand this.

- Languages and methodologies to represent knowledge and describe the “world”
- Developing Inference techniques to solve tasks and problems using such representations.
KRR and End User

In addition to
- Building modeling languages and methodologies
- Building inference systems

KRR is concerned with
- **Modeling** domains of interest

Modeling is also of major concern to end users of KR – *knowledge engineers*
Modelling in science

- Modelling is done in all formal empirical sciences
  - empirical: a “real”, non-mathematical domain is being studied.
  - formal: a mathematical modelling of this domain
    - by stating mathematical equations and properties (axioms)
    - by defining mathematical objects — operations, relations, ..
  - Evaluation of correctness of the modelling by experimental validation

- Physics, Probability, Economy, Climate, Geology, Chemistry, Biology, Natural Language understanding, . . . ,
An example – the two body problem

- Newtons gravity theory restricted to two bodies:

  One body’s acceleration is proportional to the mass of the other object and inverse proportional to the square of the distance.

\[
\frac{d^2 F_1(t)}{d^2 t} = G \times \frac{m_2}{||F_2(t) - F_1(t)||^2}
\]

\[
\frac{d^2 F_2(t)}{d^2 t} = G \times \frac{m_1}{||F_2(t) - F_1(t)||^2}
\]

where

- \( F_1(t), F_2(t) \): spatial positions of objects 1, 2 at time \( t \).
- \( m_1, m_2 \): masses of objects 1 and 2
- \( G \): the gravitational constant.
How does scientific modelling work?

- Scientific theories characterise mathematical images of the domain: structures.

A structure is a mathematical object consisting of:

- A universe: a set of objects.
- A set of distinguished objects, functions, relationships in this universe.

In the 2-body theory:

- The universe consists of the real numbers \( \mathbb{R} \)
- Two functions \( F_1, F_2 : \mathbb{R}^+ \rightarrow \mathbb{R}^3 \)

- Structures that satisfy the theory represent possible states of affairs of the domain.
  - Structures that satisfy the 2-body theory are solutions of the 2-body problem and represent the ways two isolated bodies could behave.
How does scientific modelling work?

- Structures: The universal tool of formal empirical sciences to make mathematical images of the empirical reality.
- A structure is, in a sense, a very inaccurate image of the world:
  - Idealisations of the world.
  - An enormous amount of information is abstracted away.
  - E.g., Newton’s apple and the moon look the same in his theory, as two moving volume-less points.
- Modelling (almost) necessarily involves making approximations, i.e., errors
  - Newton’s theory is only accurate at low relative speeds.
- Nevertheless, the success of mathematical modelling has been phenomenal.
- A major difficulty in modelling is to find the “right” abstractions.
Problem solving in the 2-body problem

This mathematical modelling can be used to solve a range of problems:

- E.g., compute trajectories of planets.
- E.g., if I drop an object from the Empire State building, when will it hit the ground?
- E.g., compute the velocity of object 2 at time 10, given initial positions and speeds.
- E.g., given that at time 0, both objects have zero speed and occur within the sphere with radius 1, is it possible that they ever leave this sphere, and if so, when?
- ...
Modelling in this course

Modelling in this course

= Making a formal scientific theory of some domain or system expressed in a formal language

Our formal theories will characterise a class of structures, mathematical images of possible states of affairs in the domain.

Domains – could be anything

▶ A domain of family relations, university lecture schedules, . . .
▶ Dynamic systems
  ▶ Specification of a dynamic system
  ▶ Specifying desired properties of dynamic systems to be verified
  ▶ Specifying a planning domain,
▶ Programs and desired properties of programs
Modelling in this course

Modelling in this course

Making a formal scientific theory of some domain or system expressed in a **formal language**

Our formal theories will characterise a class of structures, mathematical images of possible states of affairs in the domain.

Domains – could be anything

- A domain of family relations, university lecture schedules, . . .
- Dynamic systems
  - Specification of a dynamic system
  - Specifying desired properties of dynamic systems to be verified
  - Specifying a planning domain,
- Programs and desired properties of programs
In this course primarily KR paradigm considered:

- Answer Set Programming
- Its language
- Modeling in ASP (assignments)
- Inference technology of ASP

Wake up Question: What is KRR concerned with as a scientific discipline?