COMP44[illegible]nowledge Representation and Reas[illegible]g

Introduction to Knowledge Representation and Reasoning

Maurice Pagnucco
School of Computer Science and Engineering

COMP4418, Week 1

Knowledge Representation and Reasoning

Several of the lectures in the first section of this course are based on the following book:

> *Ronald Brachman & Hector Levesque Knowledge Representation and Reasoning Morgan Kaufmann, 2004. ISBN: ISBN: 978-1-55860-932-7.*

These slides will be clearly identified with the footer: B&L (2005) as in this slide. This material has been used with permission.

Up-to-date slides for this book are available from:

> *http://www.cs.toronto.edu/~hector/PublicKRSlides.pdf*

What is Knowledge?

Easier question: how do we talk about it?

We say “John knows that ...” and fill the blank with a **proposition**

- 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 bike
- affective: deep understanding

Belief: similar, but 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:

- ✚ ⟶ first aid
- 🍴 ⟶ restaurant
- “Alice” ⟶ Alice
- “John loves Mary” ⟶ the proposition that John loves Mary

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” ⟶ “1

↓ ↓ ↓

eleven two thirteen

Analogy: relationships

“John is Mary’s father” ⟶ “John is an adult male”

↓ ↓

⟶

Why Knowledge?

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

- e.g. a game-playing prog
  "because it believed its q s 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 (Daniel Dennett)

But...

Is KR just a convenient way of describing complex systems?

- sometimes anthropomorphizing is inappropriate e.g. thermostats
- can also be very misleading!
  fooling users into thinking a system knows more than it does

Why Representation

Note: intentional stance says nothing about what is / is not represented symbolically

- e.g. in game playing
  perhaps the board position is represented, but the goal of getting a knight out early is not

KR Hypothesis: (Brian Smith)

> *"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 semantic attribution, play a formal but causal and essential role in engendering the behaviour that manifests that knowledge."*

Two issues: existence of structures that

- we can interpret propositionally
- determine how the system behaves

Knowledge-based system:
one designed in this way!

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).
```


Both systems can be described intentionally

Only the 2nd has a separate collection of symbolic structures à la KR Hypothesis; its **knowledge base** (or KB)

∴ small knowledge-based system

KR and Artificial Intelligence

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”,
- ...

Some, to a certain extent

- game-playing,
- vision,
- ...

Some, to a much lesser extent

- speech,
- motor control,
- ...

Current research question:
how much of intelligent behaviour is knowledge-based?
Challenges: connectionism, others

Why Bother?

Why not “compile out” knowledge into specialized procedures?

- distribute KB to procedures that need it
  (as in Example 1)
- almost always achieves big 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 “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 [illegible]on."
- informability: debugging the KB
  "No the sky is not yellow. It's blue."
- extensibility: new relations
  "Canaries are yellow."
- new applications
  returning a list of all the white things painting pictures

Hallmark of KB'ed system:
the ability to be told facts about the world and adjust behaviour correspondingly
"Cognitive penetrability" (Zenon Pylyshyn)
actions that are conditioned by what is currently believed
e.g. do not leave the room on hearing a fire alarm if we believe that the alarm is being tested
so this action is cognitively penetrable

Why Reasoning?

Want knowledge to affect acti

not do action A if sente in KB

but do action A if world described in satisfies P

Difference:

P may not be **explicitly** represented

Need to apply what is known to particulars of given situation

Example:

"Patient x is allergic to medication m."

"Anybody allergic to medication m is also allergic to medication m'."

Is it OK to prescribe m' for x?

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

Entailment

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

- If the world is such that it satisfies the P_i, then it must also satisfy P.
- Applies to a variety of languages with truth

Inference: the process of calculating entailments

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

Sometimes want unsound / incomplete reasoning

- we won't discuss this case here

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 **propositional logic** (PL) and then move on to **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 (in particular, Horn logic)

Allen Newell’s analysis:
Knowledge level: (semantic)
deals with language, entailment
Symbol level: (computational)
deals with representation, inference

Picking a logic has issues at each level
KL: expressive adequacy, theoretical complexity, ...
SL: architectures, data structures, algorithmic complexity

Next: we begin with PL at KL