Foundations of Artificial Intelligence

CS472/3 — Fall 1999
Lecture #16&17

Bart Selman

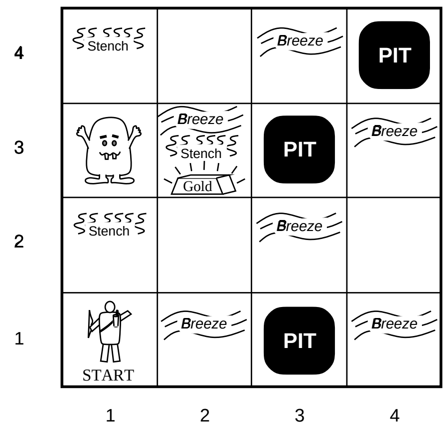
Slide CS472–1

Today's Lecture

Knowledge Representation & Reasoning, cont.
Chapter 6 & 7, R&N.

Slide CS472–2

Propositional, one more example

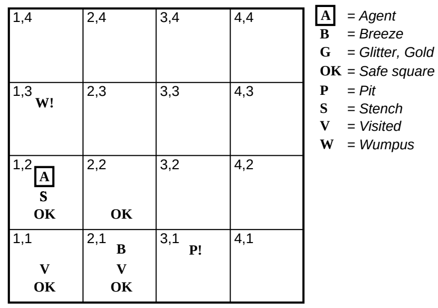


"Wumpus World" R&N, p154.

Slide CS472–3

Percept — five symbols.
E.g. [*Stench, Breeze, Glitter, None, None*]
Exercise in encoding. Important throughout AI (learning, planning etc.)

Slide CS472–4



after three moves

$\neg S_{1,1}, \neg S_{2,1}, S_{1,2}, \neg B_{1,1}, B_{2,1}, \neg B_{1,2}$

S for "stench"; B for breeze.

Slide CS472–5

Corrected slide from previous lecture.

R1: $\neg S_{1,1} \Rightarrow (\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1})$ becomes

R1a: $S_{1,1} \vee \neg W_{1,1}$ (for $\neg S_{1,1} \Rightarrow \neg W_{1,1}$)

R1b: $S_{1,1} \vee \neg W_{1,2}$ R1c: $S_{1,1} \vee \neg W_{2,1}$

R2a: $S_{2,1} \vee \neg W_{1,1}$ R2b: $S_{2,1} \vee \neg W_{2,1}$

R2c: $S_{2,1} \vee \neg W_{2,2}$ R2d: $S_{2,1} \vee \neg W_{3,1}$

R3a: $S_{1,2} \vee \neg W_{1,1}$ R3b: $S_{1,2} \vee \neg W_{1,2}$

R3c: $S_{1,2} \vee \neg W_{2,2}$ R3d: $S_{2,1} \vee \neg W_{1,3}$

R4: $\neg S_{1,2} \vee W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

R&N p. 174-175.

Slide CS472–6

Note many rule. E.g. just R3, gives up to 4 clauses **per square**.

Ex. First-order:

$\forall l_1, s, Smelly(l_1) \Rightarrow$

$(\exists l_2 At(Wumpus, l2, s) \wedge (l_2 = l_1 \vee Adjacent(l_1, l_2))$

How would you define *Adjacent* ?

How about:

$\forall i, j, , s \;\; 1 \leq i, j \leq N \; . \; Smelly(i, j) \Rightarrow$

$(\exists k, l \; . \; -1 \leq k, l \leq 1 \quad At(Wumpus, i + k, j + l, s)$

What about epressiveness? What is implicitly used? Computation?

Slide CS472–7

"Finding the Wumpus"

What holds in Wumpus KB? (12 symbols — $S_{1,1}, S_{2,1}$ etc.

$2^{12} = 4096$, truth assignments

To prove: $W_{1,3}$

Use resolution.

Compare with section 6.5 R&N for inference rule approach.

Slide CS472–8

1) $\neg S_{1,1}$ (in KB / fact)
2) $\neg W_{1,1}$ (resolution with R1a)
3) $\neg S_{2,1}$ (in KB / fact)
4) $\neg W_{1,2}$ (resolution with R1b)
5) $\neg W_{2,2}$ (res with R2c)
6) $S_{1,2}$ (in KB / fact)
7) $W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$
8) $W_{1,3} \vee W_{1,2} \vee W_{2,2}$
9) $W_{1,3} \vee W_{1,2}$
10) $W_{1,3}$
11) $\neg W_{1,3}$ (our assumption)
12) $\emptyset$ (contradiction: it follows $W_{1,3}$ **must be True.**)

Why do things this way?

Contrast with **simulation** approach.

Basic idea: axiomatic / knowledge-based / declarative approach
allows us a very large range of **queries / conclusions.**

Procedural (e.g., standard program, "simulation") would
require separate procedure for each possible question (almost).

Of course, getting the axioms / facts right can be tricky!
And, reasoning can be computationally very hard!

Slide CS472–11

So far, most success in a restricted context:
e.g. "expert systems" knowledge encoded in basic if-then rules
(similar to Horn sentences: $(p \wedge q \wedge r) \Rightarrow s$)
use forward / backward chaining (Section 9.4. R&N).

Also, more recently in **diagnosis**. (Xerox / NASA)

And, more specialized KR languages, e.g., frame-based systems
descriptin logics. (AT&T).
Tens of thousands of facts / thousands of frames ("objects").

Slide CS472–12

First-Order Logic

Richer language. Closer match to ontology / conceptual structure

objects / properties / relations / functions

Ex.:

objects: table, car, house, John...

relations: brother of, part of, has color...

properties: red, round, prime... [“unary relations”]

functions: father of, best fiend, one more than...

Q. Contrast relation with function.

Syntax and Semantics

Study section 7.1 of R&N carefully.

Note: Semantics can be defined more formally.

See e.g. “A course in mathematical logic” Bell & Machover.

R&N provide main ideas behind semantics.

Semantics give by **interpretations** (propositional
analogue: truth assignment)

Sentence ("formula") evaluates to True or False under a given interpretation. If True, interpretation is called a **model** of the sentence.

We hope that models of KB are close to actual state of affairs (world). But often, we also have unexpected (non-standard) models.

Relation between mathematical notion of interpretation / model and actual physical world interesting philosophical issue. We'll ignore it.

Each interpretation is defined over a given **domain** U (set of individuals / objects).

Slide CS472–15

Constant symbols: *A, B, C, John, chair-1, house-10...*

In interpretation these symbols correspond to elements of U.
(two constants can define the same element in U).
(morning-star / evening-star)

Predicate symbols: *Round, Brother, Part-of,...*

Each predicate symbol correspond to a relation on U.
E.g., a binary predicate, corresponds to a binary relation.
If U equals { car, tires, steering wheel, house }.
[tires, car], and [steering wheel, car].
could be the intended interpretation of "part-of"

Slide CS472–16

But there may be another interprtation with "part-of" given by

[car, tires], [steering wheel, car].

How do we prevent this?

Slide CS472–17

Language	Ontological Commitment (What exists in the world)	Epistemological Commitment (What an agent believes about facts)
Propositional logic	facts	true/false/unknown
First-order logic	facts, objects, relations	true/false/unknown
Temporal logic	facts, objects, relations, times	true/false/unknown
Probability theory	facts	degree of belief 0…1
Fuzzy logic	degree of truth	degree of belief 0…1

Slide CS472–18

Function symbols: *Cosine, FatherOf, LeftLegOf* ...

Correspond to functions defined on U.

Relation to predicate symbols?

Slide CS472–19

Again, "what's meant by embodying **knowledge** about the world?"

Example:

1) $On(A, Fl) \Rightarrow Clear(B)$
2) $(Clear(B) \wedge Clear(C)) \Rightarrow On(A, Fl)$
3) $Clear(B) \vee Clear(A)$
4) $Clear(B)$
5) $Clear(C)$

Slide CS472–20

One interpretation:

U is the set { `A, B, C, Floor` }.

1) mapping constant symbols to elements of U.

e.g., A to `A`, B to `B`, C to `C`

and Fl to `Floor`

Could we have mapped Fl to `A`??

2) mapping of relation symbol On to relation on U.

e.g., On = { [`B, A`], [`A, Floor`], [`C, Floor`]}.

3) mapping of relation (property) $Clear$ to a unary rel. on U.

e.g., $Clear$ = { [`B`], [`C`] }.

Slide CS472–21

Yet others . . .

```
B                                 C
A   C        A   B   C            A    B
------       -------            -----
floor          floor              floor
```

Including completely different interpretations!

E.g., use integers for domain. (Lowenheim 1915)

Slide CS472–22

Try to add sufficient axioms (facts) to rule out unwanted models. E.g., add $clear(A)$.

Slide CS472–23

Terms —- a logical expressions that refers to an object. Constant symbols are terms. Functions applied to constant symbols. $FatherOf(John)$. Also, **variables** are terms (later) and functions applied to variables or other terms.

The interpretation is given by whatever the Constant or Function maps to in U (vars later).

If no vars, called **atomic terms**.

Slide CS472–24

Atomic sentences — A predicate symbol applied to atomic terms.

E.g. $Married(FatherOf(Richard), MotherOf(John))$

Evaluated to **true** if predicate symbol holds between the objects referred to by the arguments.

Complex sentences — add **logical connectives**.

E.g. $Older(John, 30) \Rightarrow Older(Jane, 29)$

Slide CS472–25

Quantifiers

Universal Quantification $\forall$ —

E.g., $\forall x \;\; Cat(x) \Rightarrow Mammal(x)$

Think of as:

$(Cat(Spot) \Rightarrow Mammal(Spot)) \quad \wedge$

$(Cat(Felix) \Rightarrow Mammal(Felix)) \quad \wedge$

$(Cat(John) \Rightarrow Mammal(John)) \quad \wedge$

$\ldots$

Intuition: Expand over all object symbols.

Slide CS472–26

Existential Quantification $\exists$ —

E.g., $\exists x \;\; Sister(x, Spot) \wedge Cat(x)$

Think of as:

$(Sister(Spot, Spot) \wedge Cat(Spot)) \quad \vee$

$(Sister(Rebecca, Spot) \wedge Cat(Rebecca)) \quad \vee$

$(Sister(Felix, Spot) \wedge Cat(Felix)) \quad \vee$

...

Intuition: Expand over all object symbols.

Slide CS472–27

Equality $=$ —

E.g. $father(John) = Henry$

True iff refer to same object of U in interpretation.

(identity relation)

Slide CS472–28

See Chapter 7 of R&N for more discussion and fine details.

E.g. can't switch quantifiers around.

Compare $\forall x \exists y Loves(x, y)$ vs.

Compare $\exists x \forall y Loves(x, y)$

Reflex Agent

Directly connects percepts to actions:

$\forall s, b, u, c, t \quad Percept([s, b, Glitter, u, c], t) \Rightarrow Action(Grab, t)$

Or, more indirectly:

$\forall s, b, u, c, t \quad Percept([s, b, Glitter, u, c], t) \Rightarrow AtGold(t)$

$\forall t \quad AtGold(t) \Rightarrow Action(Grab, t)$

Why more flexible? Limitations of reflex approach?

Next time: Representing Change —
some surprising difficulties!