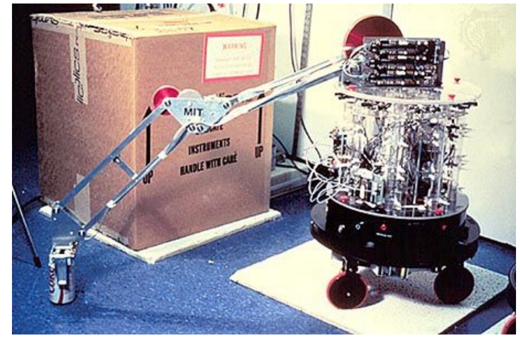
Logical Agents

Recap of last class

- Principles of intelligence
- Agent architectures
  - SMPA architecture
  - subsumption architecture

Homework



- Design the control logic for a robot that has to drive around the Trottier building and collect empty soda cans
- Robot has:

Sensors	Effectors
Laser rangefinder	three-wheel base
Compass	2-joint arm
IR beam-break (between fingers)	gripper
IR proximity sensors (around base)	
Contact sensor (on hand)	

Herbert the soda-can-collecting robot

- Laser-based table-like object finder – drives robot to table
- If robot stationary, arm control reaches out for coke can
- Laser-based coke-can object finder moves arm toward soda can
- When can breaks infrared beam between fingers, grasp reflex is activated to pickup can

Herbert

- No planning
- No representation of the environment
- No communication between modules
- Herbert could respond quickly to changed circumstances:
  - e.g., new obstacle, or object approaching on a collision course.
  - e.g., place a coke can in front of Herbert – he will pick it up. No expectations about where coke cans will be found.

Which way am I going?

- Without a map:
  - How can Herbert be sure to find its way to all parts of the environment?
  - How can Herbert be sure to bring cans back to “home?”
- video

Today's Agenda

- Using logical reasoning as the basis of a knowledge-based agent

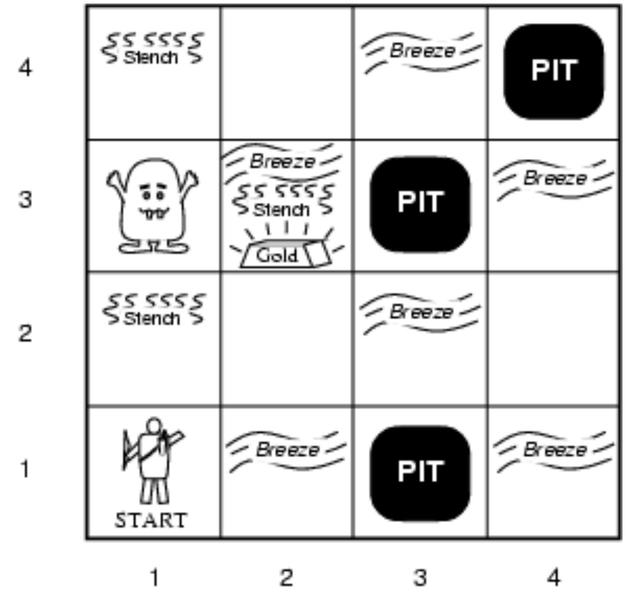
Consider...

- Reflex agents find their way from Montreal to Ottawa by dumb luck
- Chess program calculates legal moves of its king, but does not know that no piece can be on 2 different squares at the same time
- Representations we have seen for problem-solving agents is limiting

Knowledge-Based Agents

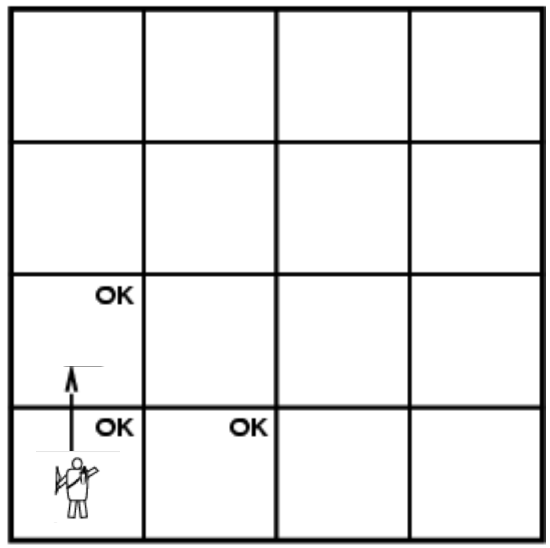
- combine general knowledge with current percepts to infer hidden aspects of current state
- Knowledge base (KB): set of sentences represented in a knowledge representation language; represents assertions about the world
- Inference rule: way to derive new sentences from existing ones
- we add new sentences to KB and query what is known by TELL and ASK operations

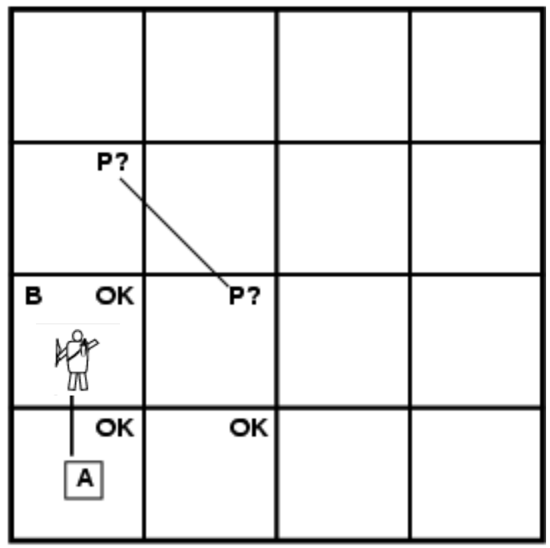
Wumpus World

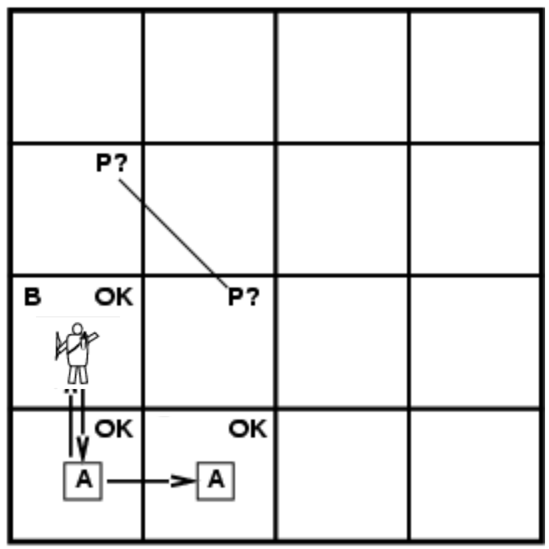


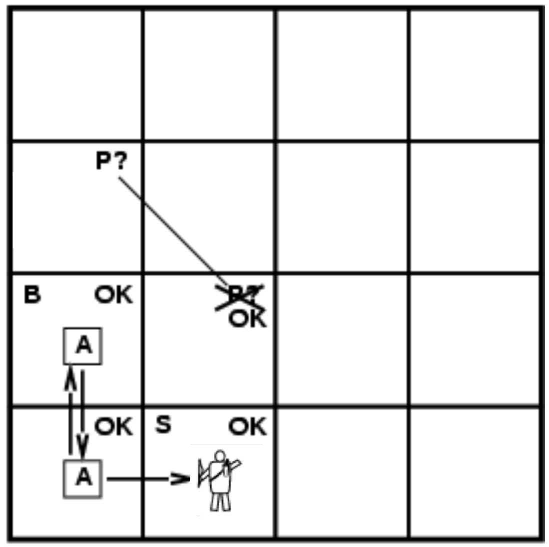
- Environment
  - 4x4 grid of rooms
  - Agent starts in [1,1]
  - Gold and wumpus locations chosen randomly
  - Each square other than [1,1] can be a pit with P(0.2)
- Actuators
  - *Left turn, Right turn, Forward*
  - Agent dies if it enters a square containing a pit or live wumpus
  - Can climb out of the cave from square [1,1]
- Sensors:
  - Stench (S): in cells directly adjacent to wumpus (W)
  - Breeze (B): in cells directly adjacent to pit (P)

Inference at Play in the wumpus world

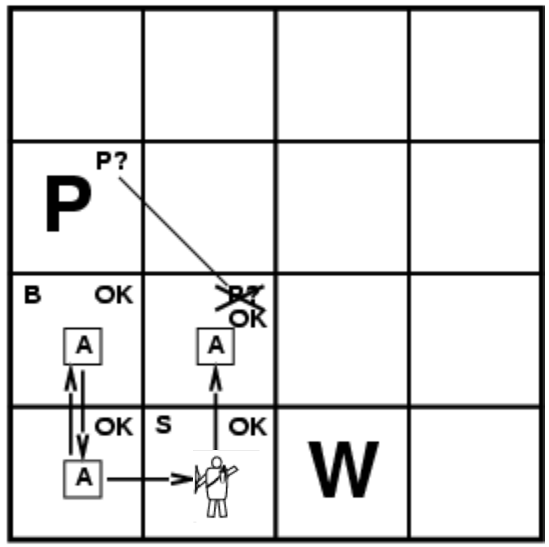



P?
B
OK
P?
OK
OK
A

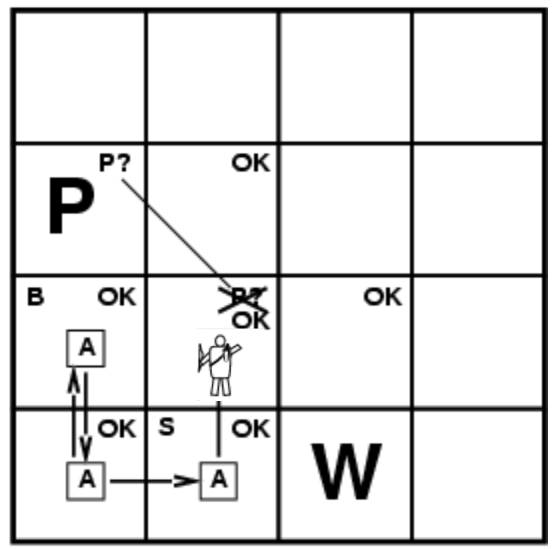

P?
B
OK
P?
OK
OK
A
A



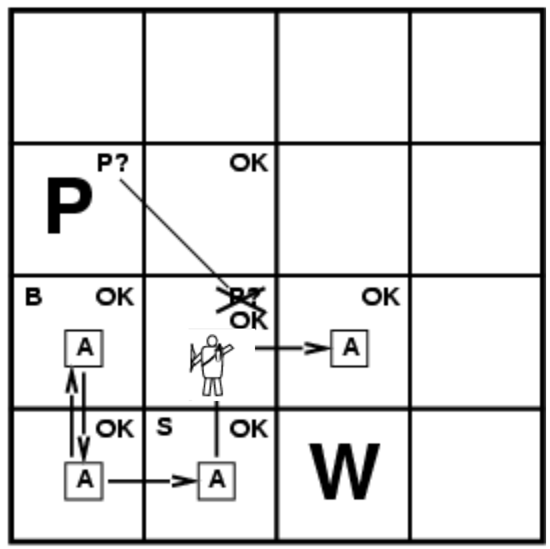
P?
B
OK
P?
OK
A
OK
S
OK
A



P?
P
B
OK
P?
OK
A
A
OK
S
OK
A
W



P?
OK
P
B
OK
P?
OK
OK
A
OK
S
OK
A
A
W



P?
OK
P
B
OK
P?
OK
OK
A
A
OK
S
OK
A
A
W

That was easy

- Now how do we develop a logical agent to apply the same reasoning?

Propositional logic

negation: ¬S is true iff S is false
conjunction: $S_1 \wedge S_2$ is true iff S_1 is true and S_2 is true
disjunction: $S_1 \vee S_2$ is true iff S_1 is true or S_2 is true
implication: $S_1 \Rightarrow S_2$ is true iff S_1 is false or S_2 is true
biconditional: $S_1 \Leftrightarrow S_2$ is true iff $S_1 \Rightarrow S_2$ is true and $S_2 \Rightarrow S_1$ is true

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
false	*false*	*true*	*false*	*false*	*true*	*true*
false	*true*	*true*	*false*	*true*	*true*	*false*
true	*false*	*false*	*false*	*true*	*false*	*false*
true	*true*	*false*	*true*	*true*	*true*	*true*

Standard Logical Equivalences

$$
\begin{aligned}
(\alpha \wedge \beta) &\equiv (\beta \wedge \alpha) \quad \text{commutativity of } \wedge \\
(\alpha \vee \beta) &\equiv (\beta \vee \alpha) \quad \text{commutativity of } \vee \\
((\alpha \wedge \beta) \wedge \gamma) &\equiv (\alpha \wedge (\beta \wedge \gamma)) \quad \text{associativity of } \wedge \\
((\alpha \vee \beta) \vee \gamma) &\equiv (\alpha \vee (\beta \vee \gamma)) \quad \text{associativity of } \vee \\
\neg(\neg\alpha) &\equiv \alpha \quad \text{double-negation elimination} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\beta \Rightarrow \neg\alpha) \quad \text{contraposition} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\alpha \vee \beta) \quad \text{implication elimination} \\
(\alpha \Leftrightarrow \beta) &\equiv ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)) \quad \text{biconditional elimination} \\
\neg(\alpha \wedge \beta) &\equiv (\neg\alpha \vee \neg\beta) \quad \text{de Morgan} \\
\neg(\alpha \vee \beta) &\equiv (\neg\alpha \wedge \neg\beta) \quad \text{de Morgan} \\
(\alpha \wedge (\beta \vee \gamma)) &\equiv ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma)) \quad \text{distributivity of } \wedge \text{ over } \vee \\
(\alpha \vee (\beta \wedge \gamma)) &\equiv ((\alpha \vee \beta) \wedge (\alpha \vee \gamma)) \quad \text{distributivity of } \vee \text{ over } \wedge
\end{aligned}
$$

Inference in Propositional Logic: Modus Ponens

$$\frac{\alpha \rightarrow \beta \,, \quad \alpha}{\beta}$$

means that whenever $\alpha \rightarrow \beta$ and α are given, we can infer β

Wumpus world sentences

Let $P_{i,j}$ be true if there is a pit in [i, j]
Let $B_{i,j}$ be true if there is a breeze in [i, j]

"There is no pit in [1, 1]":

$R_{1:}$ $\neg P_{1,1}$

"a square is breezy if and only if there is an adjacent pit":

$R_{2:}$ $B_{1,1} \leftrightarrow (P_{1,2} \lor P_{2,1})$

$R_{3:}$ $B_{1,2} \leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{1,3})$

percepts for first two squares visited:

$R_{4:}$ $\neg B_{1,1}$

$R_{5:}$ $B_{1,2}$

Reasoning in Propositional Logic

$R_{2:}$ $B_{1,1} \leftrightarrow (P_{1,2} \vee P_{2,1})$

$R_{6:}$ $(B_{1,1} \rightarrow (P_{1,2} \vee P_{2,1})) \wedge ((P_{1,2} \vee P_{2,1}) \rightarrow B_{1,1})$ — bicond. elimin.

$(P_{1,2} \vee P_{2,1}) \rightarrow B_{1,1}$ — and-elimination

$\neg B_{1,1} \rightarrow \neg(P_{1,2} \vee P_{2,1})$ — contrapositive equivalence

$\neg(P_{1,2} \vee P_{2,1})$ — Modus Ponens with R_4

$\neg P_{1,2} \wedge \neg P_{2,1}$ — de Morgan's rule

$R_{7:}$ $\neg B_{2,1}$ — percept

What can you infer from R_7?

Resolution

- If you know:
  - it's raining or it's snowing (a v b)
- and you also know:
  - it's not raining (¬a)
- then you can conclude:
  - it's snowing (b)

Resolution

unit resolution inference:

if y_i and m are complementary literals, i.e., $y_i \wedge m = 0$

$$\frac{y_1 \vee \ldots \vee y_{i-1} \vee y_i \vee y_{i+1} \vee \ldots y_k, \qquad m}{y_1 \vee \ldots \vee y_{i-1} \vee \qquad y_{i+1} \vee \ldots y_k}$$

Generalizes to full resolution rule:

$$\frac{y_1 \vee y_2, \neg y_1 \vee y_3}{y_2 \vee y_3}$$

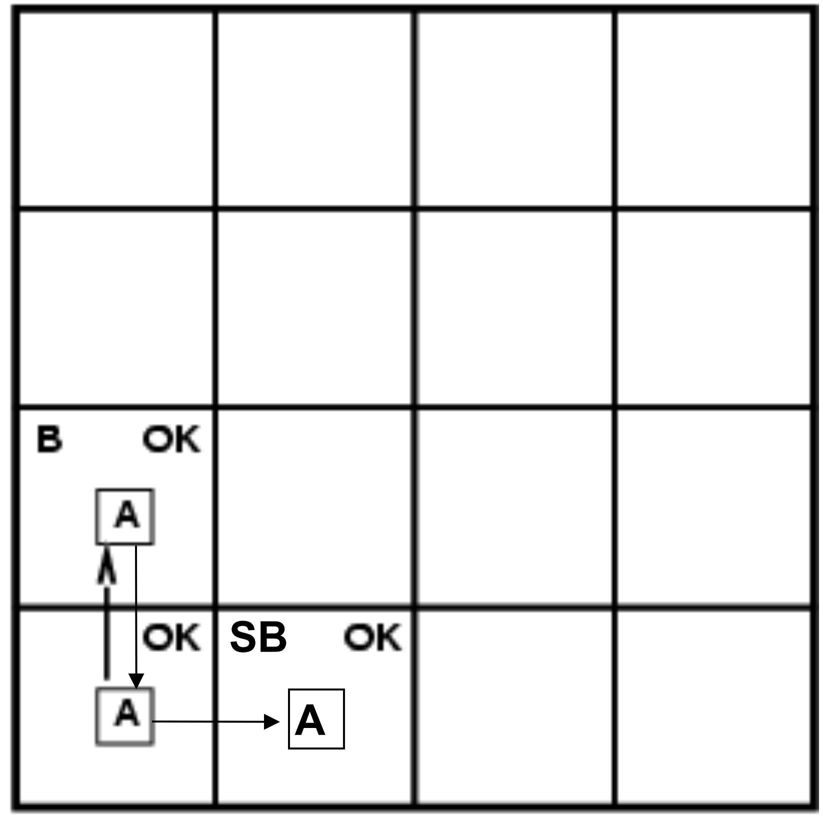
So continuing…

$R_{11:}$ $P_{1,1} \vee P_{2,2} \vee P_{1,3}$ bicon. elim. of R_3 & MP with R_5

$P_{1,3}$ resolution with R_1 and R_9

Exercise

- New rule: The Wumpus cannot be in the same square as a pit
- Can you determine where the Wumpus is?
- What about a pit?



First-Order Logic (FOL)

- Propositional logic: propositions (sentences)
- FOL adds quantification (∀ , ∃) and predicates

Predicates

- assume that Spot and Fido are dogs
- then the predicate, Dog(x)
  - returns TRUE if x is Spot or Fido

Unification

- Let *p* and *q* be sentences in FOL
- Let *U* be a **unifier**, i.e., some set of substitutions of values for variables
- *subst(U,x)* is the result of applying the substitutions of *U* to sentence *x*
- If *subst(U,p)* = *subst(U,q)* then *UNIFY(p,q)* = U
  - The **unification** of p and q is the result of applying U to both of them.

Example

Dog(Spot)

Man(John)

UNIFY((Bites(Spot,y) ∧ Dog(x) ∧ Man(y)),
(Bites(z,John) ∧ Dog(z) ∧ Man(John))) =
{x/Spot; y/John; z/Spot}