Logical Agents 1:

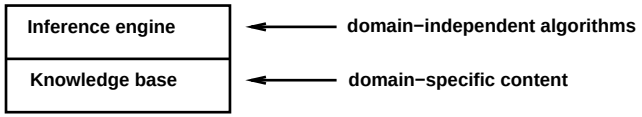
Knowledge Bases and the Wumpus World

Alan Bundy

School of informatics

(some slides courtesy of Russell and Norvig)

Knowledge bases



Knowledge base = set of sentences in a formal language

Declarative approach to building an agent (or other system):
Tell it what it needs to know

Then it can Ask itself what to do—answers should follow from the KB

Agents can be viewed at the knowledge level
i.e., what they know, regardless of how implemented

Or at the implementation level
i.e., data structures in KB and algorithms that manipulate them

A simple knowledge-based agent

function KB-Agent(*percept*) **returns** an *action*
static: *KB*, a knowledge base
t, a counter, initially 0, indicating time

Tell(*KB*, Make-Percept-Sentence(*percept*, *t*))
action ← Ask(*KB*, Make-Action-Query(*t*))
Tell(*KB*, Make-Action-Sentence(*action*, *t*))
t ← *t* + 1
return *action*

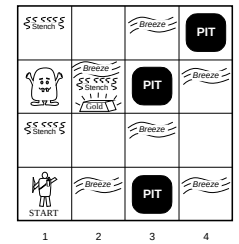
The agent must be able to:
Represent states, actions, etc.
Incorporate new percepts
Update internal representations of the world
Deduce hidden properties of the world
Deduce appropriate actions

Wumpus World PAGE description

Percepts Breeze, Glitter, Smell

Actions Left turn, Right turn,
Forward, Grab, Release, Shoot

Goals Get gold back to start
without entering pit or Wumpus square



Environment
Squares adjacent to Wumpus are smelly
Squares adjacent to pit are breezy
Glitter if and only if gold is in the same square
Shooting kills the Wumpus if you are facing it
Shooting uses up the only arrow
Grabbing picks up the gold if in the same square
Releasing drops the gold in the same square

Wumpus world characterization

Is the world fully or partially observable??

Is the world deterministic or stochastic??

Is the world episodic or sequential??

Is the world static or dynamic??

Is the world discrete or continuous??

Is the world single or multi-agent??

Wumpus world characterization

Is the world fully or partially observable?? Partially—only local perception.

Is the world deterministic or stochastic?? Deterministic—outcomes exactly specified.

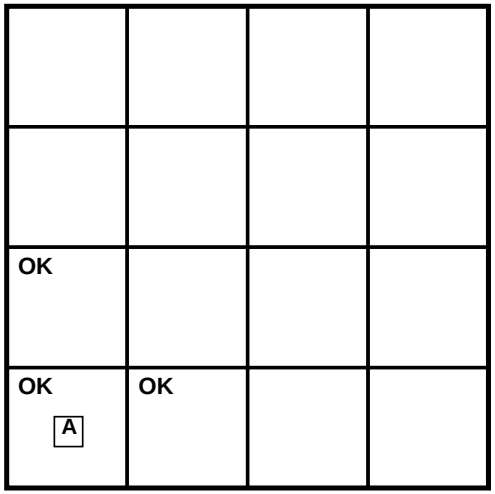
Is the world episodic or sequential?? Sequential—current moves depend on past.

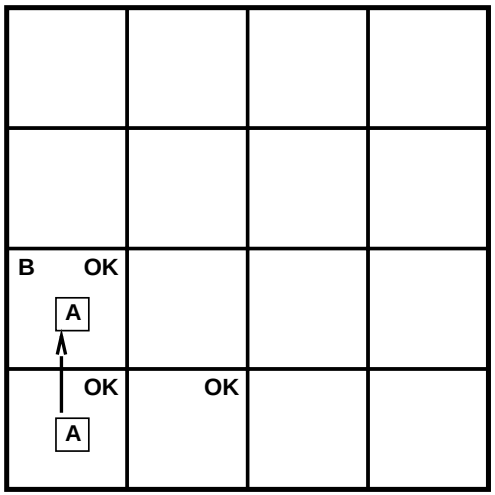
Is the world static or dynamic?? Static—Wumpus and Pits do not move.

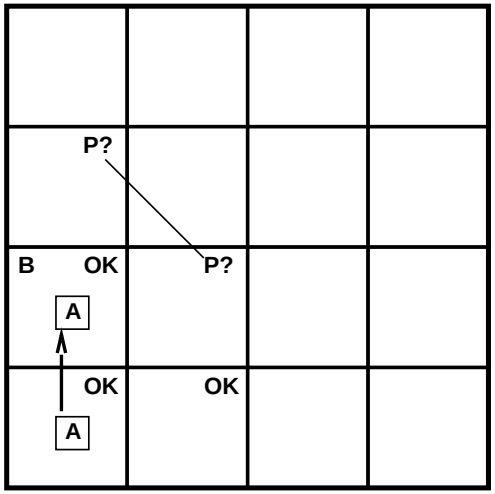
Is the world discrete or continuous?? Discrete—changes associated with moves.

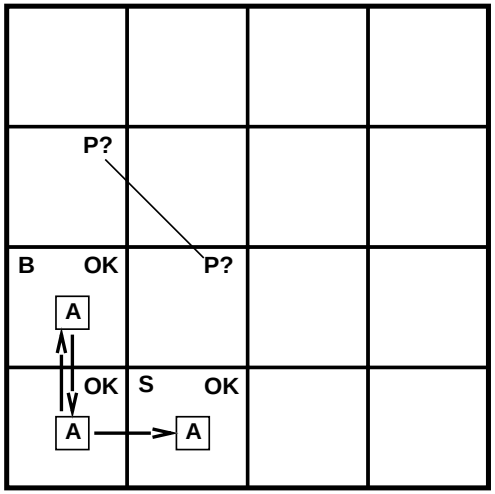
Is the world single or multi-agent?? Single—just the player.

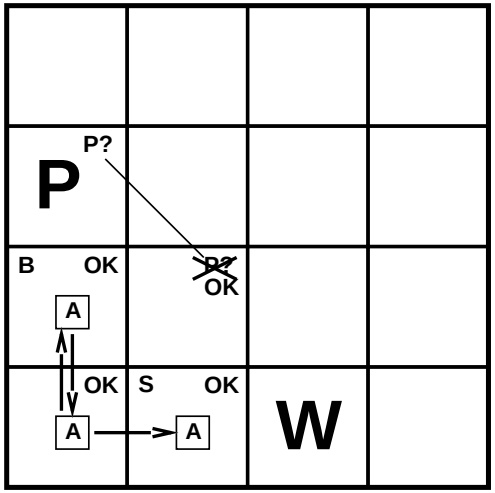
Exploring a Wumpus World

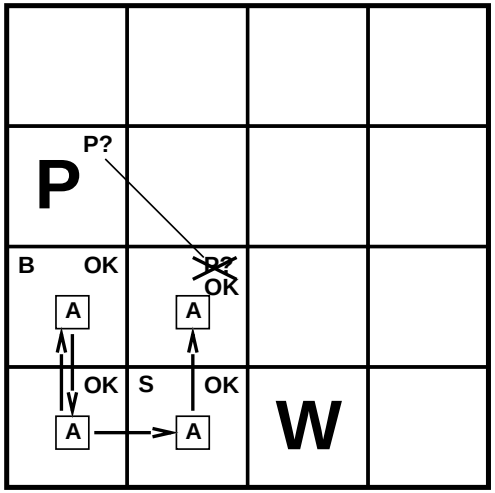


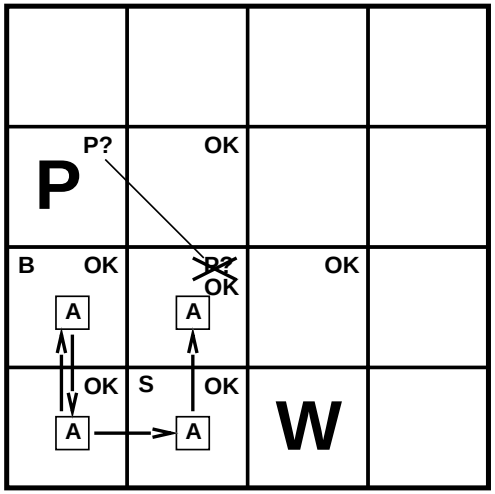


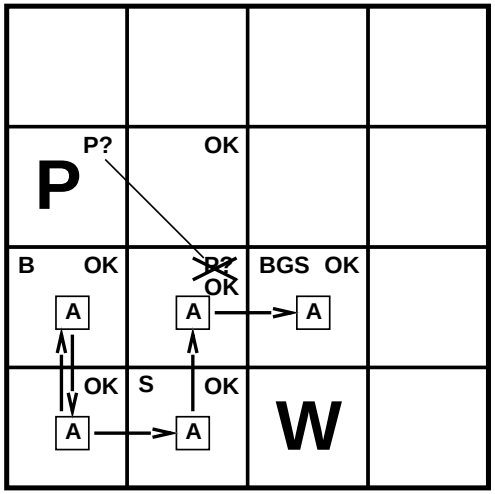

P?
B
OK
P?
A
OK
OK
A


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


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

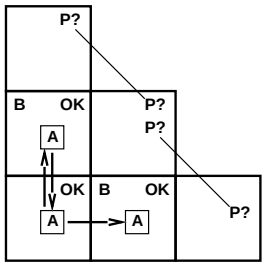

P
P?
B
OK
P?
OK
A
A
OK
S
OK
W
A
A

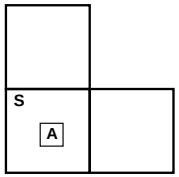




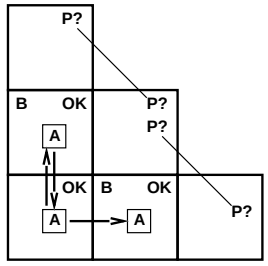
Mid Lecture Exercise

What should we do in the following situations?



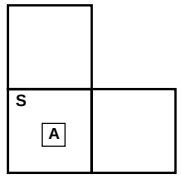


Discussion of Exercise



Breeze in (1,2) and (2,1)
$\Rightarrow$ no safe actions

Assuming pits uniformly distributed,
(2,2) is most likely to have a pit



Smell in (1,1)
$\Rightarrow$ cannot move
Can use a strategy of coercion:
shoot straight ahead
Wumpus was there $\Rightarrow$ dead $\Rightarrow$ safe
Wumpus wasn't there $\Rightarrow$ safe

Representation: Propositional Variables

♢ $W_{i,j}$ means there is a Wumpus in square (i, j).

♢ $S_{i,j}$ means there is a stench in square (i, j).

♢ $P_{i,j}$ means there is a pit in square (i, j).

♢ $B_{i,j}$ means there is a breeze in square (i, j).

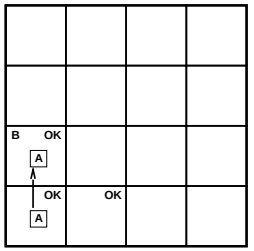
♢ $G_{i,j}$ means there is gold (and a glitter) in square (i, j).

♢ $V_{i,j}$ means that square (i, j) has been visited.

♢ $OK_{i,j}$ means that square (i, j) is safe.

For a 4×4 board there are $7 \times 4 \times 4 = 112$ propositional variables.

Representation: Knowledge Base



$\neg W_{1,1}$	$\neg S_{1,1}$	$\neg P_{1,1}$	$\neg B_{1,1}$	$\neg G_{1,1}$	$V_{1,1}$	$OK_{1,1}$
$\neg W_{2,1}$	–	$\neg P_{2,1}$	–	–	$\neg V_{2,1}$	$OK_{2,1}$
$\neg W_{1,2}$	$\neg S_{1,2}$	$\neg P_{1,2}$	$B_{1,2}$	$\neg G_{1,2}$	$V_{1,2}$	$OK_{1,2}$

Facts become known either via sensors as a result of agent actions or via inference using facts and rules.

Representation: Propositional Rule Schemas

♢ A square is safe iff it contains no Wumpus and no pit.

$$OK_{i,j} \Leftrightarrow (\neg W_{i,j} \wedge \neg P_{i,j})$$

♢ A stench iff a Wumpus in an adjacent square.

$$S_{i,j} \Leftrightarrow (W_{i-1,j} \vee W_{i+1,j} \vee W_{i,j-1} \vee W_{i,j+1})$$

♢ A breeze iff a pit in an adjacent square.

$$B_{i,j} \Leftrightarrow (P_{i-1,j} \vee P_{i+1,j} \vee P_{i,j-1} \vee P_{i,j+1})$$

NB – drop disjunct if that square does not exist, *e.g.* $P_{0,1}$.

Representation: Inference 1

♢ Rule: $S_{1,2} \Leftrightarrow (W_{2,2} \vee W_{1,1} \vee W_{1,3})$

♢ KB: $\neg S_{1,2}, \neg W_{1,1}$

♢ Inference: $\neg W_{2,2}$, $\neg W_{1,3}$

via truth table.

$S_{1,2}$	$W_{2,2}$	$W_{1,1}$	$W_{1,3}$	$W_{2,2} \vee W_{1,1} \vee W_{1,3}$	$S_{1,2} \Leftrightarrow W_{2,2} \vee W_{1,1} \vee W_{1,3}$
f	–	f	–	–	t
f	–	f	–	–	t
f	–	f	–	–	t
f	–	f	–	–	t

Representation: Inference 2

$S_{1,2}$	$W_{2,2}$	$W_{1,1}$	$W_{1,3}$	$W_{2,2} \vee W_{1,1} \vee W_{1,3}$	$S_{1,2} \Leftrightarrow W_{2,2} \vee W_{1,1} \vee W_{1,3}$
f	t	f	t	$-$	t
f	f	f	t	$-$	t
f	t	f	f	$-$	t
f	f	f	f	$-$	t

$S_{1,2}$	$W_{2,2}$	$W_{1,1}$	$W_{1,3}$	$W_{2,2} \vee W_{1,1} \vee W_{1,3}$	$S_{1,2} \Leftrightarrow W_{2,2} \vee W_{1,1} \vee W_{1,3}$
f	t	f	t	$\otimes$	t
f	f	f	t	$\otimes$	t
f	t	f	f	$\otimes$	t
f	f	f	f	f	t

Only last line of table avoids contradiction, so $W_{2,2}$ and $W_{1,3}$ both false.

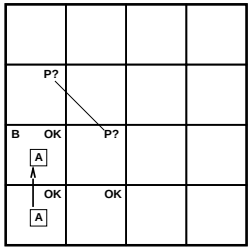
Representation: Inference 3

♢ Rule: $B_{1,2} \Leftrightarrow (P_{2,2} \vee P_{1,1} \vee P_{1,3})$

♢ KB: $B_{1,2}, \neg P_{1,1}$

♢ Inference: $P_{2,2} \vee P_{1,3}$

Must maintain disjunct until uncertainty is resolved by further exploration.



Conclusion

♢ Agents can be built using logic.

♢ Knowledge base of logical formulae describes agent's memory.
Can use propositional variables.

♢ New knowledge from sensors or ...

♢ ... Inference engine draws conclusions from knowledge.
Can use truth tables.

♢ Applied to simple Wumpus game.