Wumpus World - The story

"Lurking somewhere in the cave is the terrible wumpus, a beast that eats anyone who enters its room. The wumpus can be shot by an agent, but the agent has only one arrow. Some rooms contain bottomless pits that will trap anyone who wanders into these rooms (except for the wumpus, which is too big to fall in). The only mitigating feature of this bleak environment is the possibility of finding a heap of gold"

Wumpus World PEAS description

Performance measure

gold +1000, death -1000

-1 per step, -10 for using the arrow

The game ends either when the agent dies or when the agent climbs out of the cave.

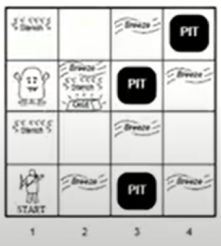
Environment

A 4 × 4 grid of rooms.

The agent always starts in t[1,1], facing to the right.

The locations of the gold and the wumpus are chosen randomly, with a uniform distribution, from the squares other than the start square.

Each square other than the start can be a pit, with probability 0.2.



Actuators Left turn, Right turn,

Press Esc to exit full screen

Forward, Grab, Release, Shoot, Climb

Sensors 5 sensors, each gives a single bit of information

In the square containing the wumpus and in the directly (not diagonally) adjacent squares, the agent will perceive a Stench.

In the squares directly adjacent to a pit, the agent will perceive a Breeze.

In the square where the gold is, the agent will perceive a Glitter.

When an agent walks into a wall, it will perceive a Bump.

When the wumpus is killed, it emits a woeful Scream that can be perceived anywhere in the cave.

[Stench, Breeze, None, None, None]

Wumpus world characterization

Observable?? No—only local perception

Deterministic?? Yes—outcomes exactly specified

Episodic?? No—sequential at the level of actions

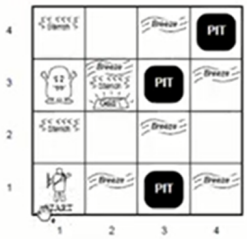
Static?? Yes—Wumpus and Pits do not move

Discrete?? Yes

Single-agent?? Yes—Wumpus is essentially a natural feature

13:07 / 24:43

Exploring a wumpus world

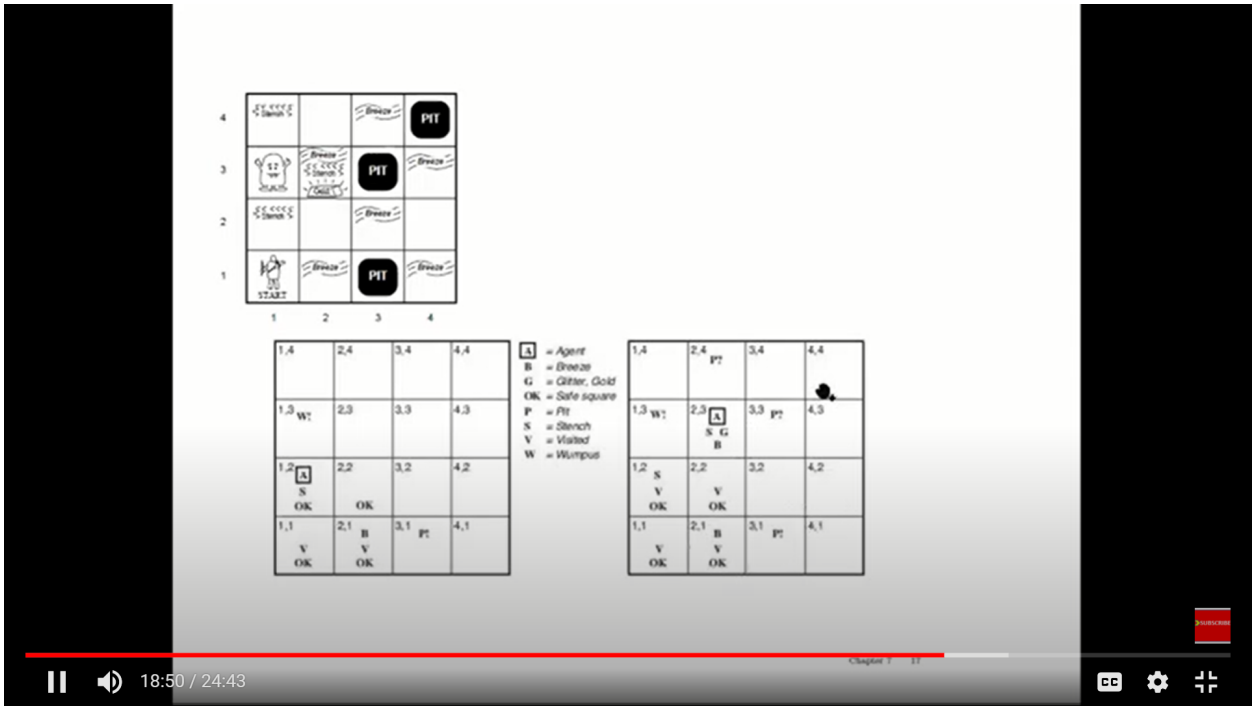


Exploring a wumpus world

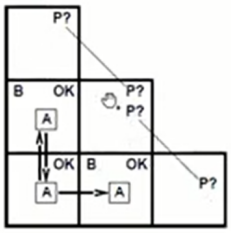
1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1

A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2 P?	3,2	4,2
1,1 V OK	2,1 A B OK	3,1 P?	4,1

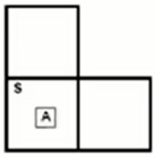


Other tight spots

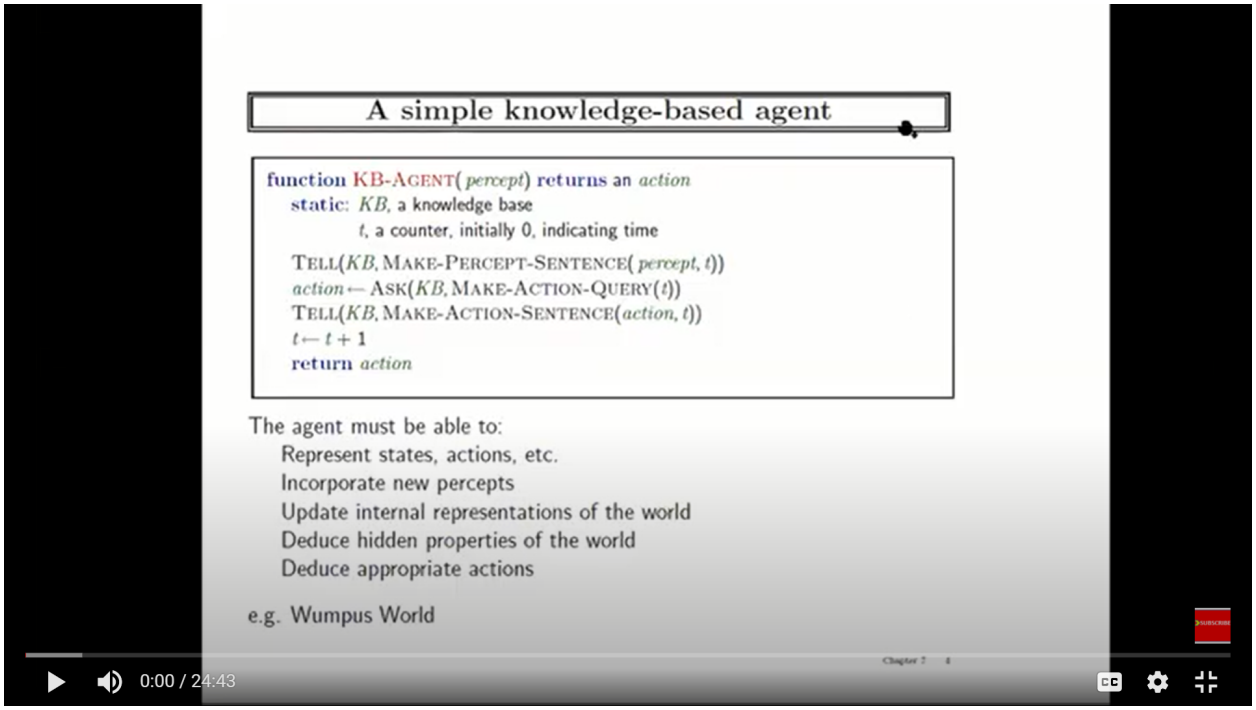


Breeze in (1,2) and (2,1)
⇒ no safe actions

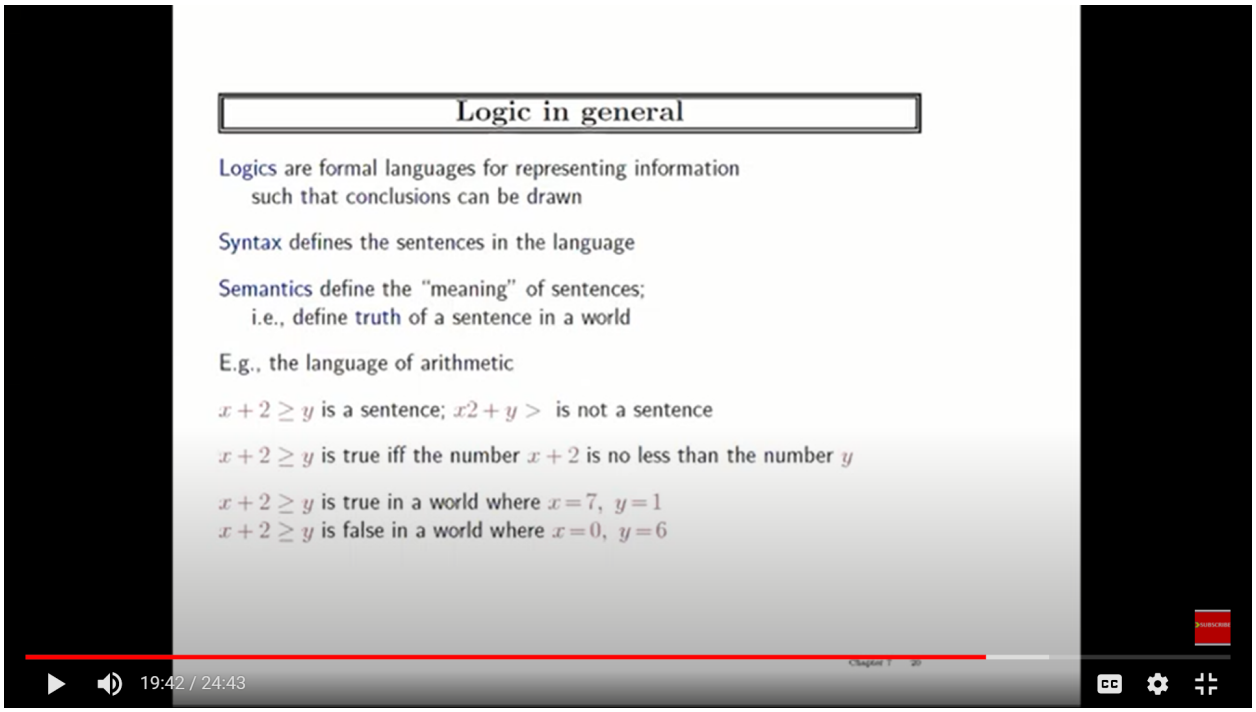
Assuming pits uniformly distributed,
(2,2) has pit with prob ?????



Smell in (1,1)
⇒ cannot move


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
e.g. Wumpus World
0:00 / 24:43

==========


Logic in general
Logics are formal languages for representing information
such that conclusions can be drawn
Syntax defines the sentences in the language
Semantics define the "meaning" of sentences;
i.e., define truth of a sentence in a world
E.g., the language of arithmetic
$x+2 \geq y$ is a sentence; $x2+y >$ is not a sentence
$x+2 \geq y$ is true iff the number $x+2$ is no less than the number y
$x+2 \geq y$ is true in a world where $x=7,\ y=1$
$x+2 \geq y$ is false in a world where $x=0,\ y=6$
19:42 / 24:43

Entailment

Entailment means that one thing follows from another:

$KB \models \alpha$

Knowledge base KB entails sentence α
if and only if
α is true in all worlds where KB is true

E.g., the KB containing "the Giants won" and "the Reds won"
entails "Either the Giants won or the Reds won"

E.g., $x + y = 4$ entails $4 = x + y$

Entailment is a relationship between sentences (i.e., syntax)
that is based on semantics

Models

Logicians typically think in terms of models, which are formally
structured worlds with respect to which truth can be evaluated

We say m is a model of a sentence α if α is true in m

$M(\alpha)$ is the set of all models of α

Then $KB \models \alpha$ if and only if $M(KB) \subseteq M(\alpha)$

E.g. KB = Giants won and Reds won
α = Giants won