Why Logic-Based AI was Cool But Fell Out of Favor

Search agents, even CSP, cannot develop new rules or insights or knowledge, even after the solution is found.

e.g. Search can solve a jigsaw puzzle ("is this right?" w/ all pieces)
CSP can solve it faster w/ constraints on connections } 5

Neither generates knowledge:
→ flat-edge pieces are puzzle border
→ try connecting similar colored pieces

Knowledge-Based Agents were huge in 1970s-1980s.
→ keep "KB" of facts
→ some way to add new facts
→ method to "reason" about facts

Need an interesting example to make it concrete.

Hunt the Wumpus. Written in BASIC for GECOS TSS.
Alongside Colossal Cave Adventure (FORTRAN) one of the first computer games. Paper rolls. Acoustic couplers.

Wumpus World

→ maze of caves, no map, don't know rooms/connections

→ partially observable - no search/CSP!

→ Some general knowledge:

- wumpus - horrible, stinky, ravenous, suction cup feet, doesn't move
- pits - bottomless, fatal (dehydration)
- gold - somewhere? (my precious)
- bow & one arrow - ought to kill wumpus (?)
- really dark - vision range 0 (current location)

→ Win: escape w/ gold

→ Lose: die (pit, wumpus)

→ Sensors:

- wumpus stinks (range 1, manhattan)
- pits make breezes (" ")
- gold glitters (range 0)
- bumping into walls hurts
- wumpus screams horribly on death (range ∞)

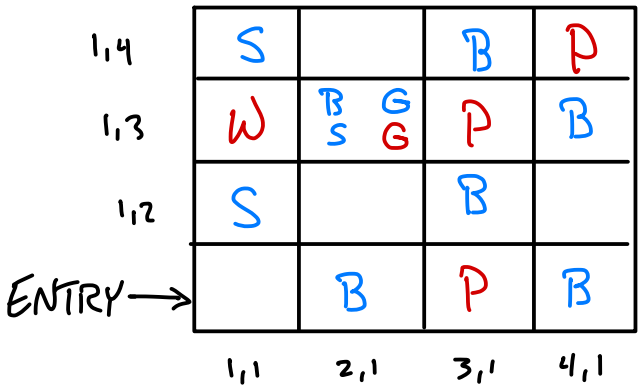
→ Actions: deterministic

- move U, D, L, R
- shoot U, D, L, R (once)
- grab gold

Static, discrete, sequential, single agent, known rules but not "map"

→ Perfect for a KB agent: specialize in inferring unknown in P/D env!

Omniscient map example



Percepts: Glitter Breeze Stench

Hidden State: Wumpus Pit Gold

Sadly, percepts have no strength/direction.

Propositional Logic Agent

→ Possible world: collection of boolean propositions

- Atoms: real world meaning, P_{12} means "pit in 1,2"
- Operations: $\wedge \vee \neg \rightarrow \underset{\text{iff}}{\leftrightarrow}$ Constants: T F Grouping: ( )

→ KB stores atoms & formulas:

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

English rule: squares next to a pit are breezy

→ Senses atoms, uses KB to infer hidden information, updates KB

e.g. sense $\neg B_{22}$, learn $\neg P_{12}$ $\neg P_{21}$ $\neg P_{32}$ $\neg P_{23}$

→ Agent effectively asks itself questions & formulates proofs:

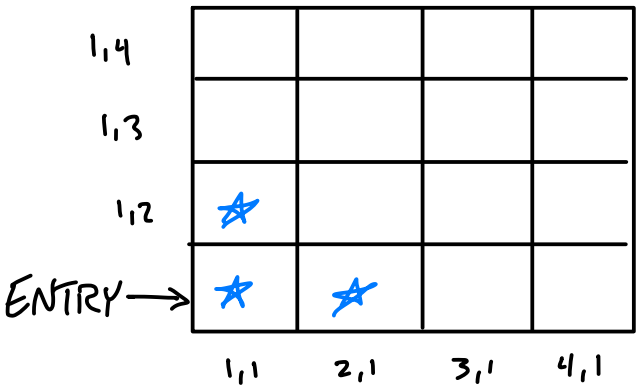
"Am I in a world where 1,3 is safe to enter?"

- "I know... and have observed... therefore..."
- Possible answers: yes, no, don't know

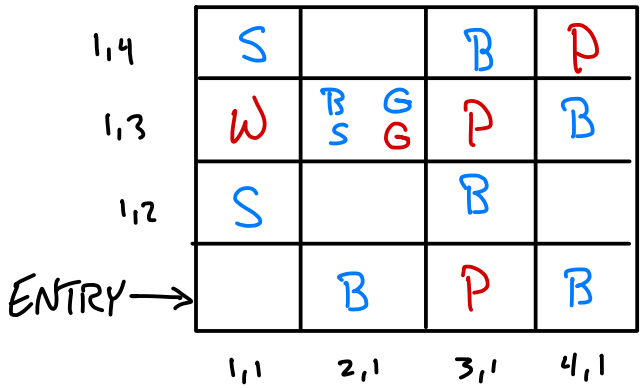
→ Heuristic was human-provided knowledge "estimate"

→ KB agent starts w/ some human-provided true knowledge, discovers more as it goes!

Hunt the Propositional Wumpus



Reference



Action

KB

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

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

$K_{ij} \leftrightarrow (\neg P_{ij} \wedge \neg W_{ij})$ "ij is 'kay to enter (safe)"

$\neg B_{11}$ $\neg S_{11}$ ← percepts

$\neg P_{11}$ $\neg P_{12}$ $\neg P_{21}$ $\neg W_{11}$ $\neg W_{12}$ $\neg W_{21}$ K_{11} K_{12} K_{21} ← inferences

---

both adjacent rooms safe

move U

$\neg B_{12}$ S_{12} → implies $(W_{12} \vee W_{02} \vee W_{22} \vee W_{11} \vee W_{13})$ ← percepts

$\neg P_{13}$ $\neg P_{22}$ $(W_{22} \vee W_{13})$ no new K! ← infer.

replace

---

only one new safe room

move D, R

B_{21} $\neg S_{21}$ → implies $(P_{21} \vee P_{11} \vee P_{31} \vee P_{20} \vee P_{22})$ ← per.

→ implies $(\neg W_{21} \wedge \neg W_{11} \wedge \neg W_{31} \wedge \neg W_{20} \wedge \neg W_{22})$

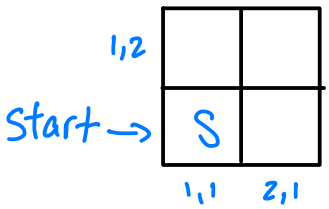
P_{31} $\neg W_{31}$ $\neg W_{22} \rightarrow W_{13}$ $\neg K_{31}$ K_{22} $\neg K_{13}$ ← infer.

---

can now safely explore 2,2...

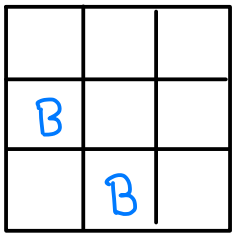
This kind of deduction is what makes KB agents special, & good for certain P|O environments!

Limitations of "Hard Logic"



No safe rooms to explore. What to do?
Coerce the environment. Clever use of actions guarantees useful new knowledge.

Shoot up generates K_{12}!
(either: scream → was W_{12} and now dead, or no scream → $\neg W_{12}$)



No safe rooms to explore & no way to coerce. Time to go home (or guess and risk dying)!

K_{31}? K_{22}? K_{13}? all answer "don't know"

Limits:
- common to get "don't know" to all relevant questions
- real sensors are noisy: wind speed 1.5 ± 0.7 mph
  → if "breezy" ≥ 1.0 mph, probably breezy here?
- real actions can fail (or have some uncertainty)
  → advance 10m (slip/stick wheel, pebble... accidental turn)

Solution: allow for probabilistic knowledge!
- maybe 2,1 is sometimes breezy, sometimes not
- over time agent could estimate $P(B_{21}) = 0.85$
- therefore $P(P_{22}) = 0.85$

Maybe I even know how accurate my sensors/actions are:
$P(\text{gold} \mid \text{glitter}) = 0.9$ $P(\text{left} \mid \text{"go left"}) = 0.99$ (engine stall)

Why virtually all AI/ML today is really probabilistic inference!