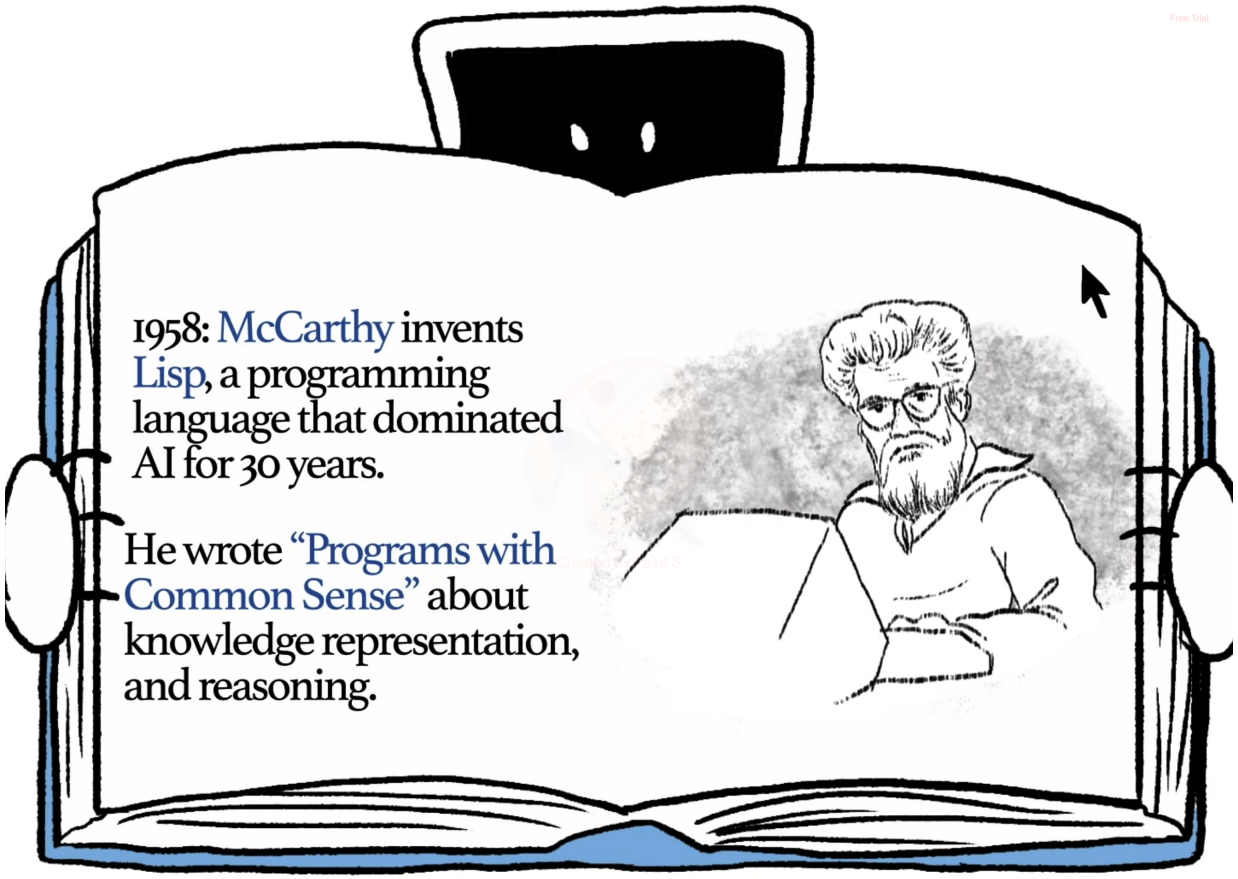
Logical Agents

aiphabet

Your AI Journey Starts Here

Knowledge-Based Agents



Knowledge-Based Agents

"The idea is that an agent can represent knowledge of its world, its goals and the current situation by sentences in logic and decide what to do by inferring that a certain action or course of action is appropriate to achieve its goals."

—John McCarthy in Concepts of Logical AI, 2000

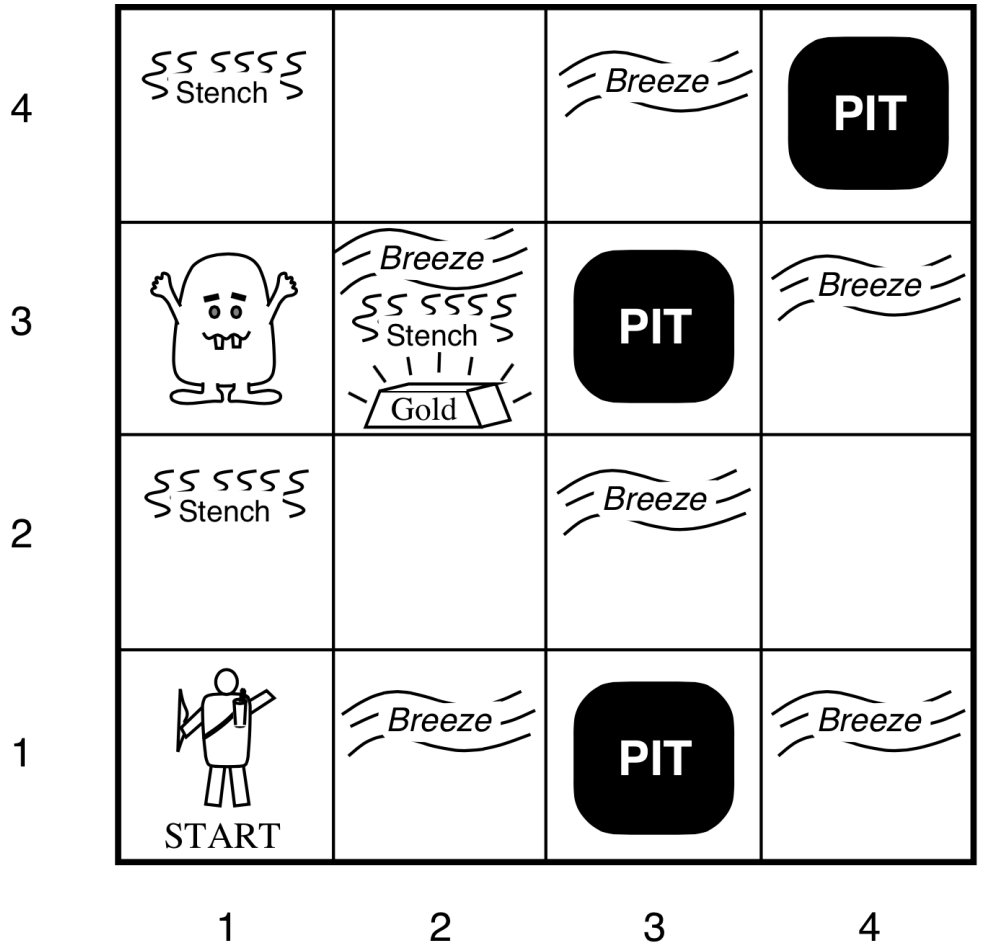
http://www-formal.stanford.edu/jmc/concepts-ai/concepts-ai.html

Knowledge-Based Agents

- Intelligent agents need **knowledge** about the world to choose good actions/decisions.
- Knowledge is a set of **sentences** in a knowledge representation language (formal language).
- A sentence is an assertion about the world.
- A knowledge-based agent is composed of:
  1. Knowledge base —**domain-specific** content
  2. Inference mechanism —**domain-independent** algorithms

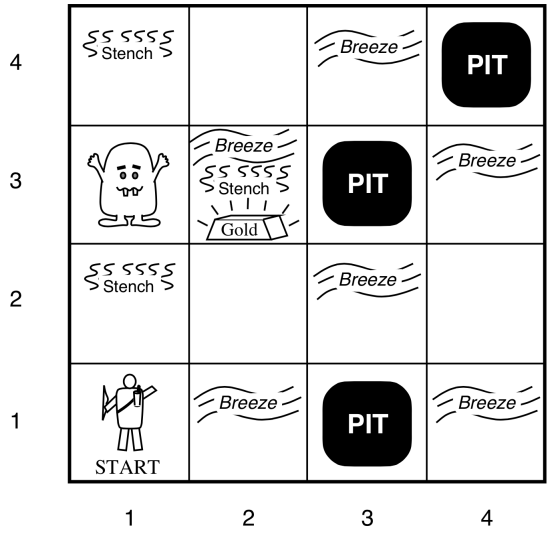
The Wumpus World

Gregory Yob (1975)



The Wumpus World

- 4 X 4 grid of rooms
- Rooms adjacent to Wumpus are smelly, and squares adjacent to pit(s) are breezy
- Glitter if and only if gold is in the same square
- Shooting kills Wumpus if you are facing it; uses up the only arrow
- Wumpus emits a horrible scream when it is killed that can be heard anywhere
- Grabbing picks up gold if in same square

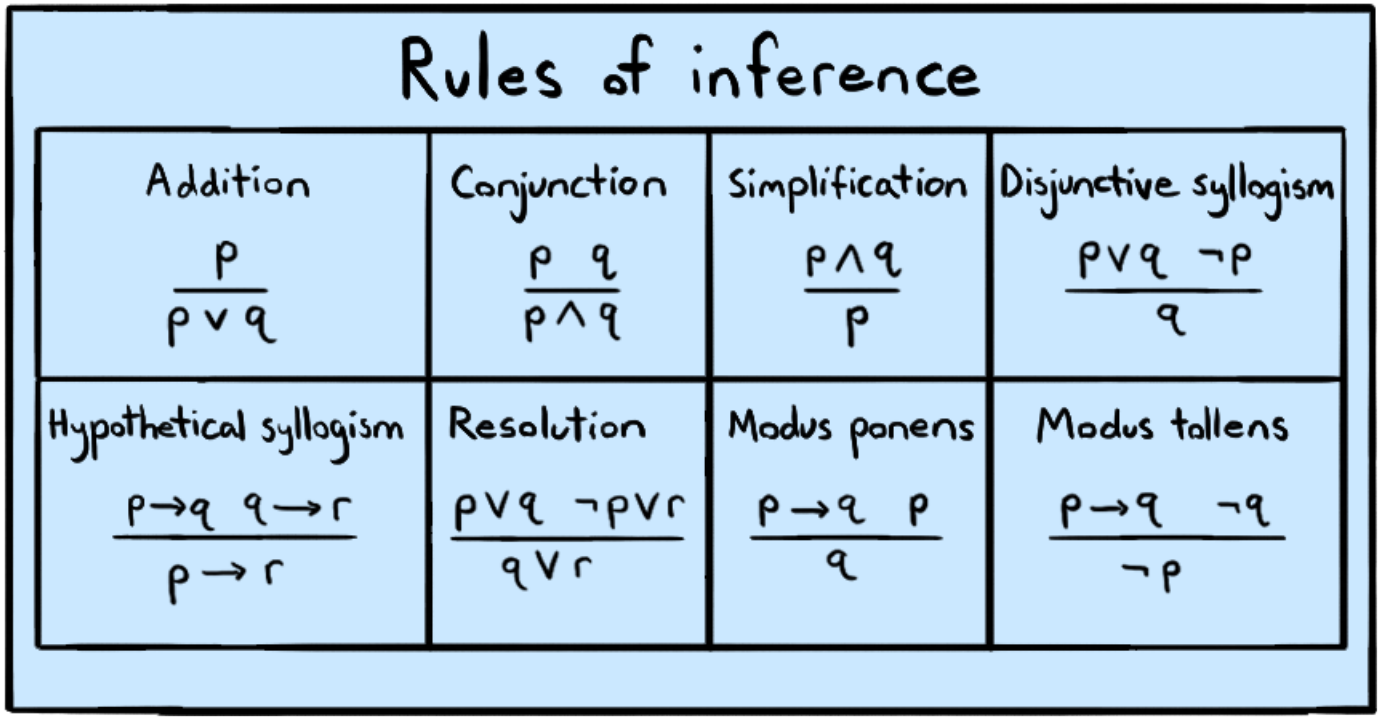


Gregory Yob (1975)

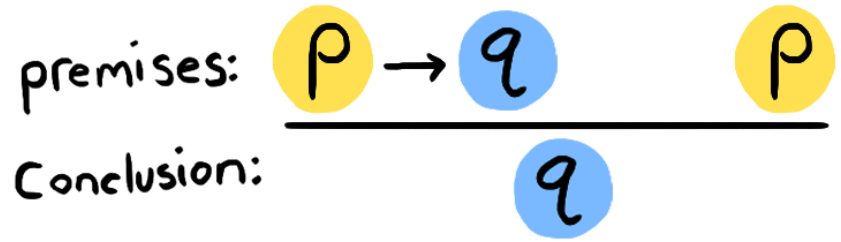
Logic: Review

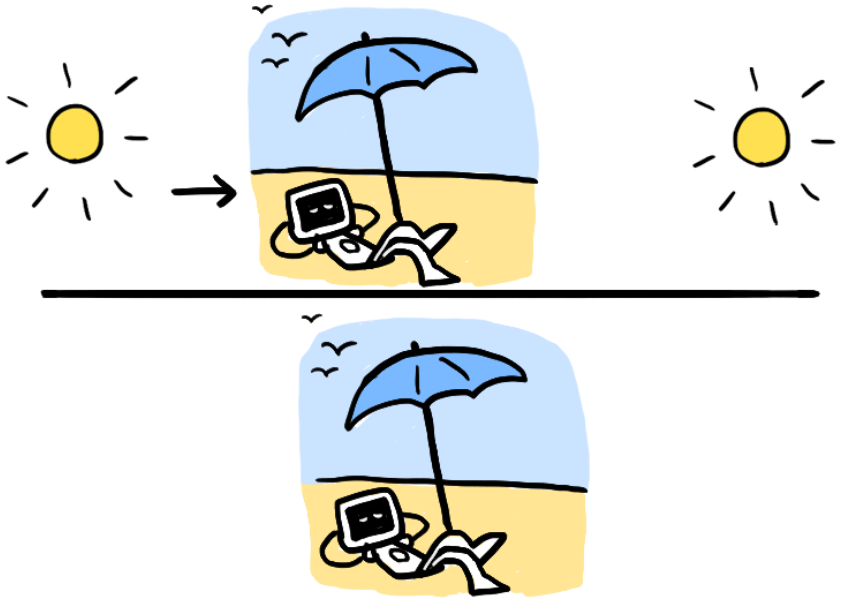
- **Knowledge base:** a set of sentences in a formal representation, logic
- **Logic:** formal language for representing knowledge
  - **Syntax:** defines well-formed sentences in the language
  - **Semantic:** defines the truth or meaning of sentences in a world
- **Inference:** a procedure to derive a new sentence from other ones
- **Logical entailment:** a relationship between sentences; it means that a sentence follows logically from other sentences

Rules of Inference

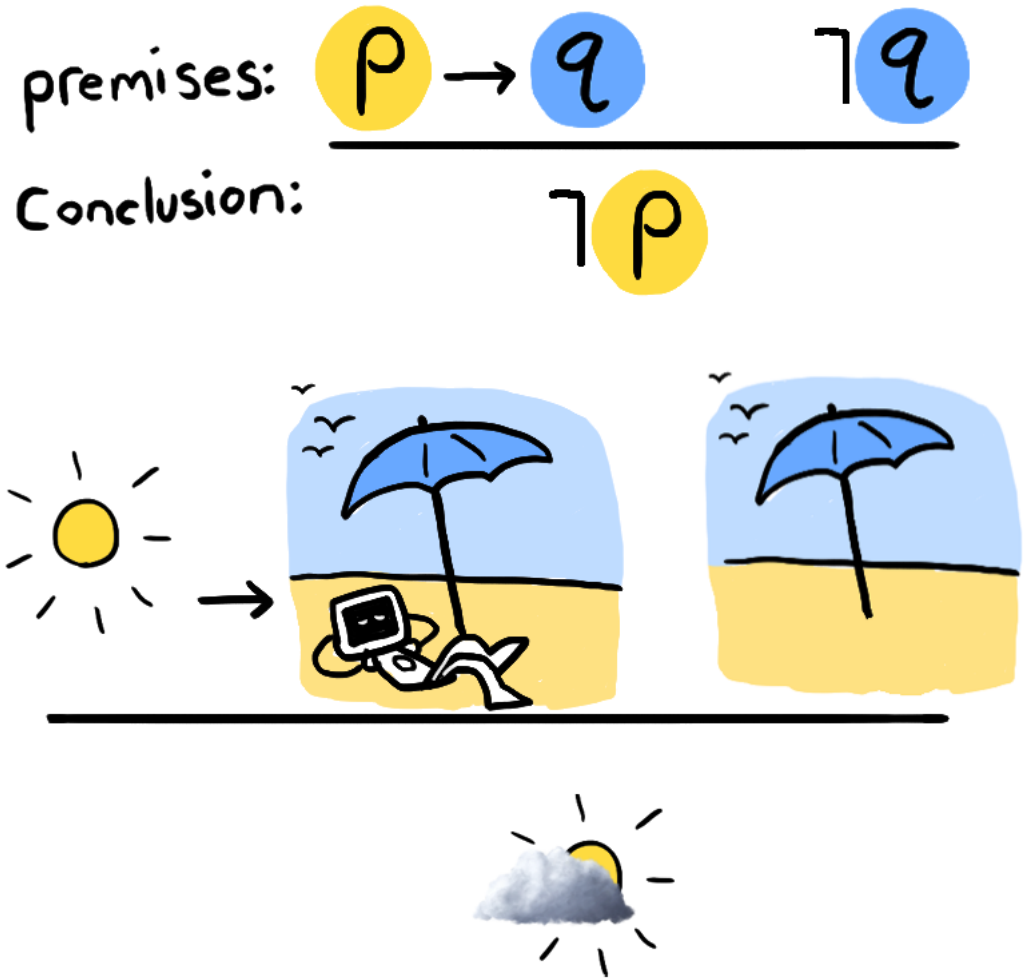


Inference (Modus Ponens)





Inference (Modus Tollens)

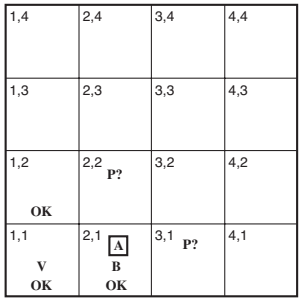


Soundness & Completeness

- We want an inference algorithm that is:
  1. **Sound:** does not infer false formulas, that is, derives only entailed sentences
  2. **Complete:** derives **all** entailed sentences

Wumpus World Inference

Let's build the KB for the reduced Wumpus world.



- 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 $[i, j]$.

R_1: $\neg P_{1,1}$

- "A square is breezy if and only if there is an adjacent pit".

R_2: $B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

R_3: $B_{2,1} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$

R_4: $\neg B_{1,1}$

R_5: $B_{2,1}$

Wumpus World Inference

Questions: $KB \models P_{1,2}?$ $KB \models P_{2,2}?$

R_1: $\neg P_{1,1}$

R_2: $B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

R_3: $B_{2,1} \Leftrightarrow P_{1,1} \vee P_{2,2}$

R_4: $\neg B_{1,1}$

R_5: $B_{2,1}$

Model Checking

- Truth table for inference
- Model: assignment of true or false to every propositional symbol
- Check that α is true in every model in which KB is true

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
false	*false*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*false*	*false*
false	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*false*	*true*	*false*	*false*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
false	*true*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*false*	*true*	*true*	*false*
false	*true*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*false*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*true*	*false*	*false*	*true*	*false*	*false*	*true*	*true*	*false*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
true	*true*	*true*	*true*	*true*	*true*	*true*	*false*	*true*	*true*	*false*	*true*	*false*

Summary

- Building logical agents was a main research trend in AI before the mid-1990s
- Logic is used in AI to represent the environment of the agent and reason about that environment
- PL is not expressive enough to describe all the world around us.
- PL is not compact. It can't express a fact for a set of objects without enumerating all of them.
- - Do not handle uncertainty, probability does
  - Rule-based and do not use data, machine learning does
  - It is hard to model every aspect of the world
  + Intelligibility of models: models are encoded explicitly

APPENDIX

Wumpus world: PEAS and Environment

Wumpus World PEAS

- **Performance measure**
  - Gold $+1000$, death (eaten or falling in a pit) -1000, -1 per action taken, -10 for using the arrow
  - The games ends either when the agent dies or comes out of the cave
- **Environment**
  - 4 X 4 grid of rooms
  - Agent starts in square [1,1] facing to the right
  - Locations of the gold and Wumpus are chosen randomly with a uniform distribution from all squares except [1,1]
  - Each square other than the start can be a pit with probability of 0.2

Wumpus World PEAS

- **Actuators**
  - Left turn, right turn, forward, grab, release, shoot
- **Sensors**
  - Stench, breeze, glitter, bump, scream
  - Represented as a five-element list
  - Example: stench, breeze, none, none, none]

Wumpus World Environment

- Partially observable
- Static
- Discrete
- Single-agent
- Deterministic
- Sequential