Artificial Intelligence
CS6380

Knowledge and Reasoning

Knowledge

- Humans know things
  - Common sense
  - Make deductions from things we already know (Knowledge Base)
    - Alice knows that if she pushes the egg off the table, it will fall and break.
    - How can we make an artificially intelligent agent do similar inferences?

- Why should an Agent do such inferences?
  - Are search methods not enough to achieve its goals?
  - Real world deployment needs to keep making such inferences on the fly
    - Too many variables
    - A fire rescue agent should inference when/how to move at each step
  - To develop a AI that can do more generic jobs
    - As supposed to doing one particular job (like the search based agents)

Knowledge

- Examples of things that search based agents cannot do:
  - In search based agent for travelling salesman:
    - Can the distance between two cities be negative?
    - Does the triangle inequality hold?
  - In search based agent for chess:
    - Can I give a check in the next three moves?
    - Am I in a better position than my opponent?
  - Deducing that some action is safe to perform before actually performing it.
- If there are some fixed set of things that the agents needs to know, it can probably be encoded in search
- But if the questions also occur on the fly then search will not be able to handle it

Knowledge

- How to empower the agents to make inferences?
- Machine Learning Approach:
  - Show the agents lot of examples and hope that the agents somehow learn them
  - For a new setting, it will answer based on what it has learnt from the examples.
- Advantages:
  - Works well in many applications
    - Especially when large amount of data is available
- Drawbacks:
  - Answers are prone to error
  - Needs lots of data to train
  - Does not do well if the question is not part of its training
- We will look at a way to make error-free inferences that does not depend on training.

Knowledge-Based Agents

- Knowledge Based Agents use a process of reasoning over an internal representation of Knowledge to decide what actions to take.

- A Knowledge Base is a set of sentences
  - Represents some assumption / observations about the world
  - Expressed in Knowledge Representation Language

- The agent can add new sentences to the knowledge base : TELL
- The agent can query if something follows from the knowledge base : ASK

- Both ASK and TELL might involve INFERENCES
  - Inference : Deriving new sentences from the old ones

Knowledge-Based Agents

- At every time step t Agent does the following:
  - TELLS the Knowledge Base what it perceives (precept)
  - ASKS the Knowledge Base what action it should perform
  - TELLS the Knowledge Base which action was performed
  - Knowledge Base will initially contain some background knowledge

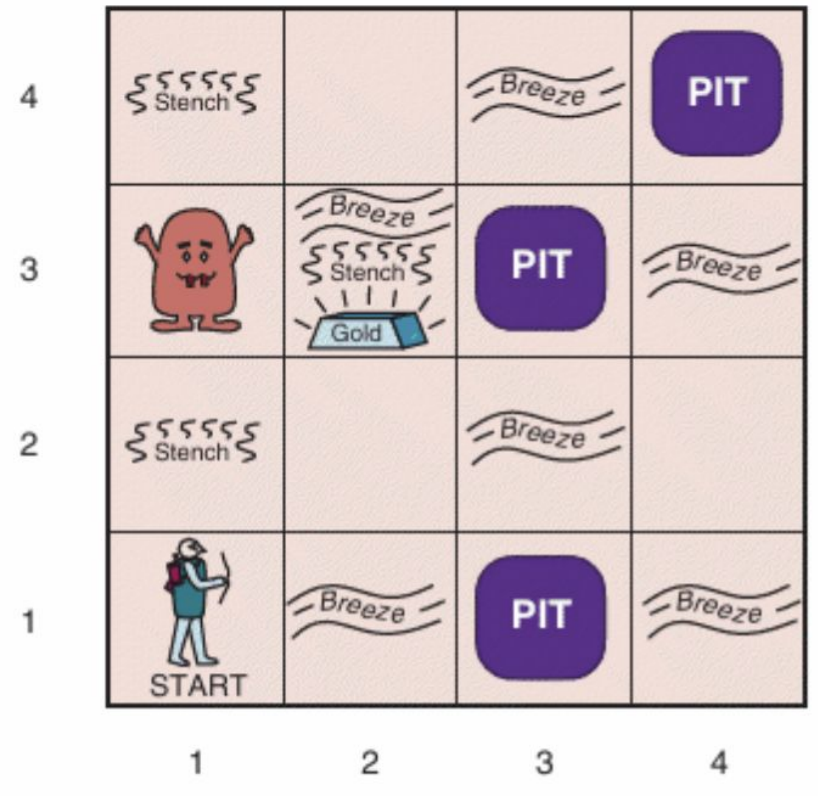
- Knowledge level abstraction does not depend on the implementation
  - Example : Automated Taxi
    - If the taxi is going from IIT Madras to Pondicherry via ECR then it knows that it should pass through Mahabalipuram
    - This is independent of how the geographic details are stored (linked list / vector / pixels ...)

Building the Knowledge Based Agent : Declarative v/s Procedural

- Designer starts with empty knowledge base and keep TELLing what agent needs to know
  - Declarative Approach
- Encode desired behaviour directly as program code
  - Procedural Approach
- Typically a combination of these techniques are used in implementations

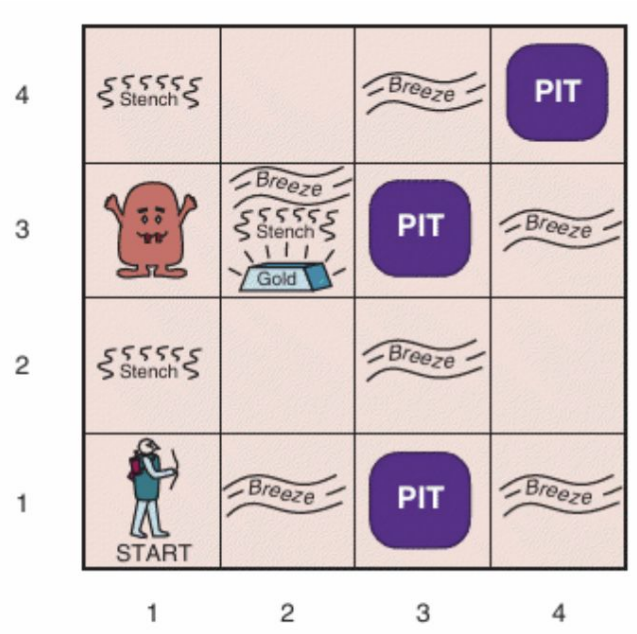
Example : The Wumpus world

- You are in a cave which is an 4 X 4 grid
- In one of the grid point there is Wumpus
  - He eats whoever enters that grid point
  - Cells adjacent to Wumpus Stinks (not diagonals)
  - Agent can shoot the Wumpus but she has only 1 arrow
- Some grid points have pits
  - Anyone who enters will fall into the pit and is stuck there forever
  - Cells adjacent to pits have Breeze (not diagonals)
- There is gold in one of these points
- Goal : Grab the gold and climb out of the cave without falling into pit or being eaten by the Wumpus



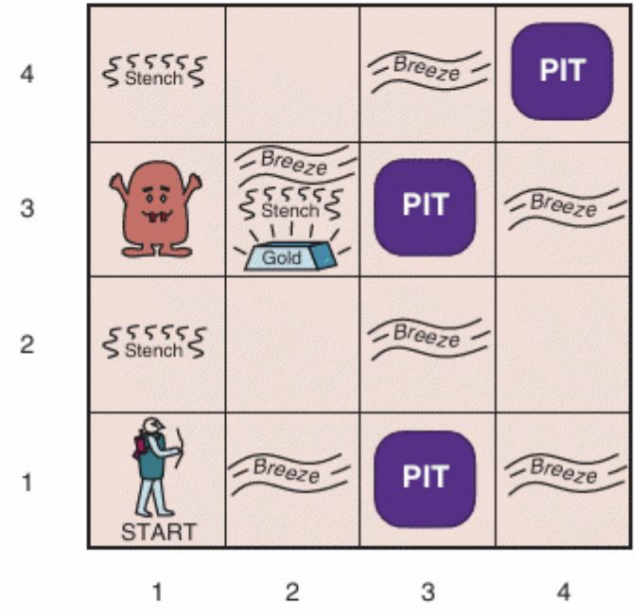
The Wumpus world : As PEAS model

- Performance Measure :
  - **+1000** for climbing out of the cave with gold
  - **+500** for climbing out of the cave without gold
  - **-1000** for falling into pit
  - **-1000** for being eaten by Wumpus
  - **-1** for every action taken
  - **-10** for using the arrow
  - Game ends when agent dies or climbs out of the cave
- Environment :
  - 4X4 grid with walls at the border
  - Agent always starts at [1,1]
  - Location of Wumpus and Gold are chosen at random (except at [1,1])
  - Every point other than the start can be a pit with probability 0.2



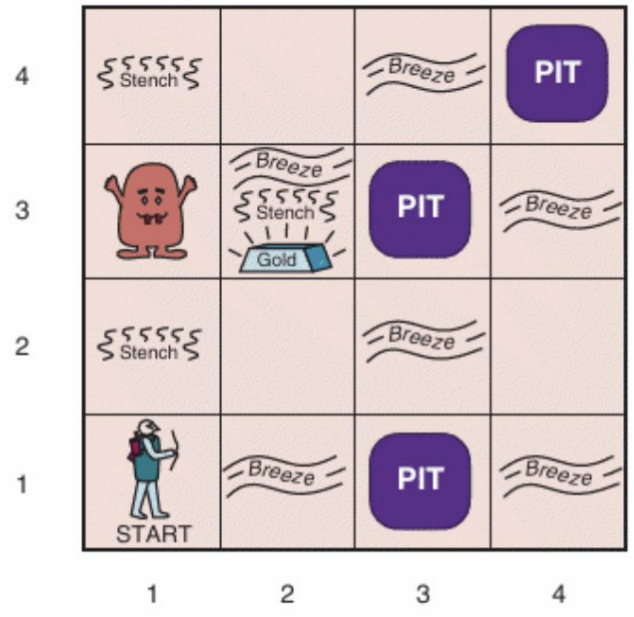
The Wumpus world : As PEAS model

- Actuators:
  - Agent can move FORWARD, TURNLEFT, TURNRIGHT
    - If an agent tries to move forward and bumps into a wall then the agent does not move
    - Agent dies if she enters a square containing a pit or Live Wumpus (entering a square with dead Wumpus is ok)
  - GRAB can be used to pick gold if the agent and gold are in the same square
  - SHOOT can be used to fire an arrow in a straight line in the direction that the agent is facing
    - The Arrow continues till it hits (and kills) Wumpus or hits the wall
    - Agent has only one arrow, so only the first shoot action has an effect
  - CLIMB can be used to climb out of the cave but only from [1,1]



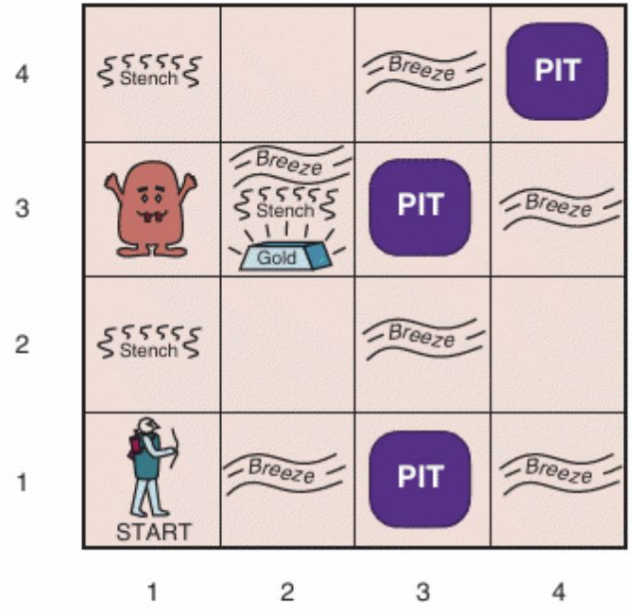
The Wumpus world : As PEAS model

- Sensors: Agent has 5 sensors
  - In the squares directly adjacent to Wumpus, Agent perceives STENCH
  - In the squares directly adjacent to Pit, Agent perceives BREEZE
  - In the square that contains gold, Agent perceives GLITTER
  - If the agent walks into wall, she will perceive BUMP
  - When Wumpus dies, it emits a SCREAM that can be perceived anywhere in the cave
- Percept is given to the agent in the form of a list at every step:
  - Example : Stench, Breeze, Glitter, No Bump, No Scream would be [STENCH, BREEZE, GLITTER, None, None]

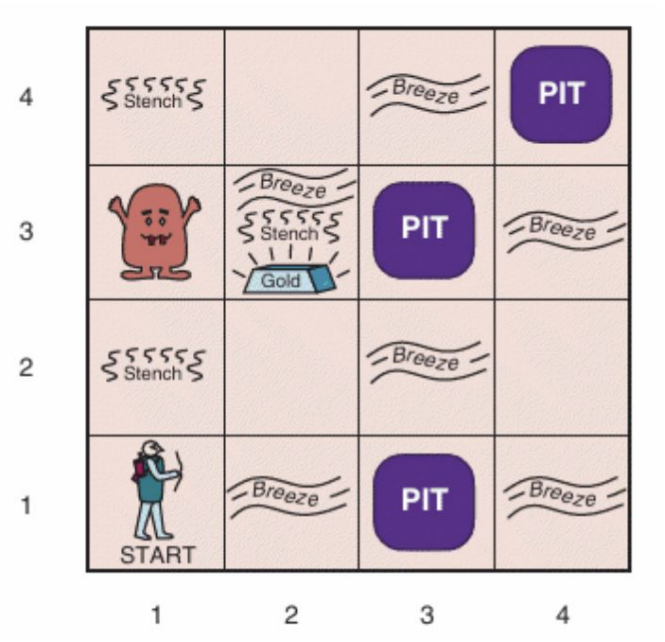


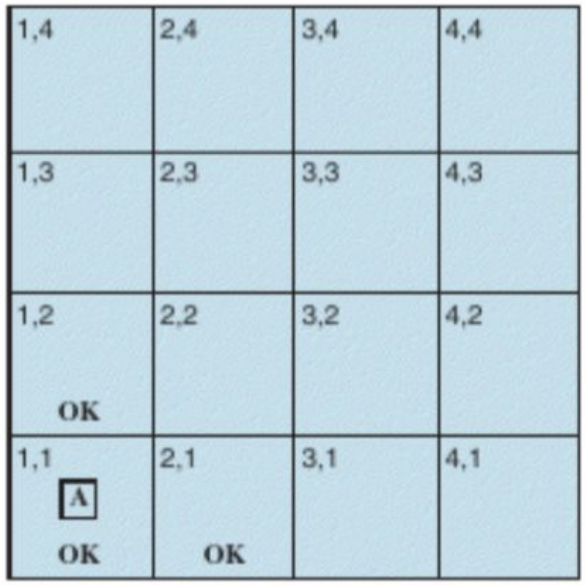
The Wumpus world

- Deterministic, Discrete, Static, Single-Agent
- Sequential
  - Reward may come only after many actions are taken
- Agent does not know the full configuration of the system (location of Wumpus, gold, pits..)
- Overcoming the ignorance seems to require logical reasoning

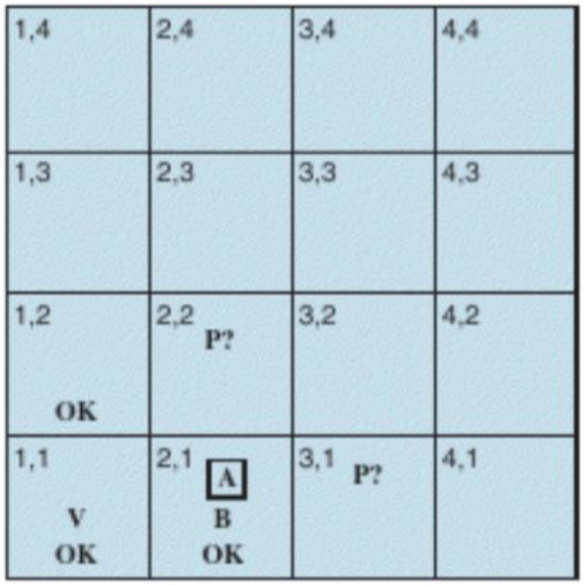


The Wumpus world



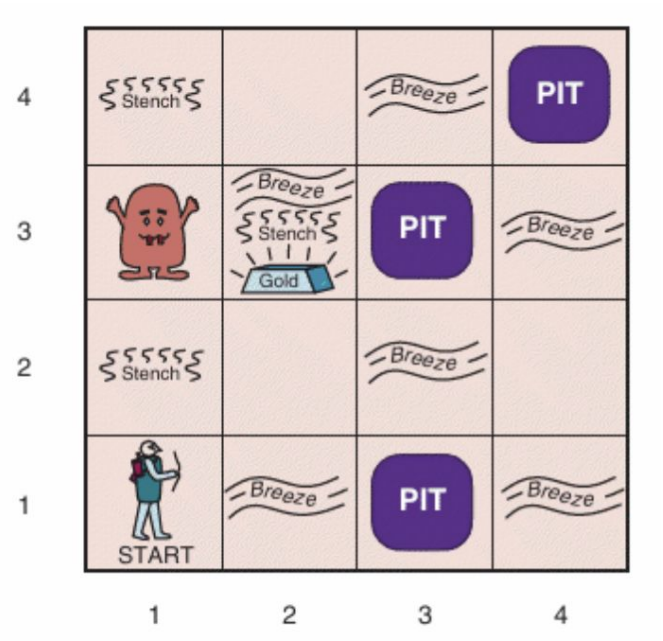


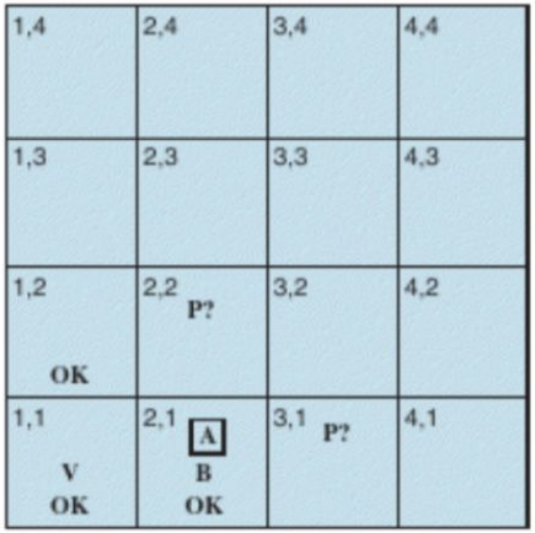
A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

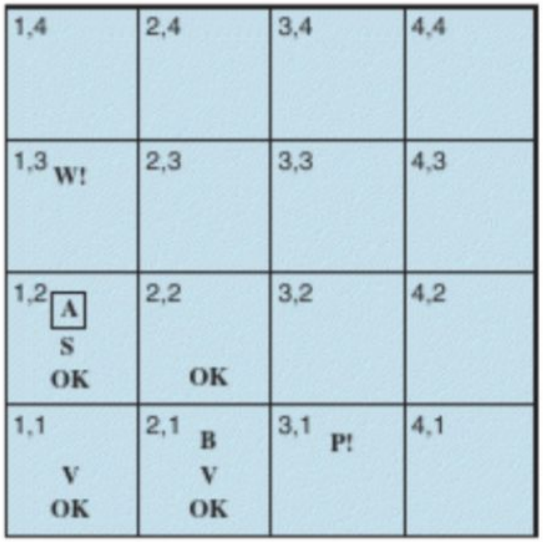


- At [1,1] agent receives [None, None, None, None, None]
- From this Agent concludes that [1,2] and [2,1] are OK ([1,1] is also OK)
- Cautious agent will only move to OK squares
- Say agent moves to [2,1] where agent perceives [None, Breeze, None, None, None]
- So there must be a pit in [2,2] or [3,1] or both (but not [1,1])

The Wumpus world

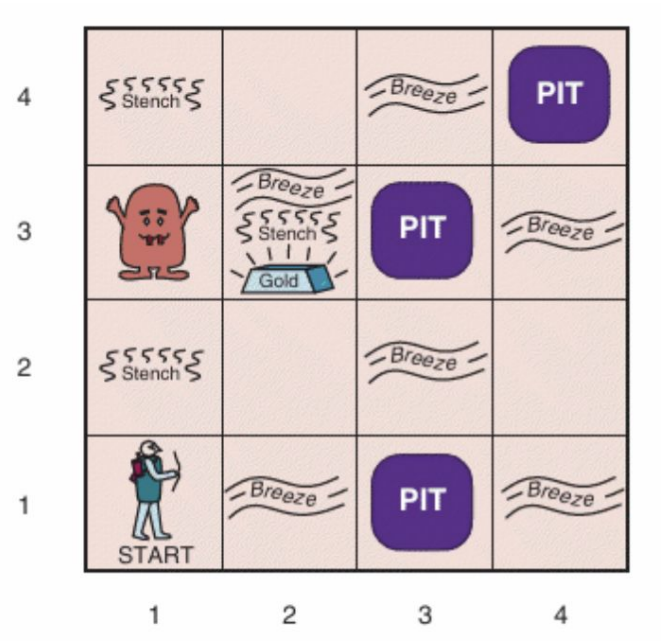






- Cautious agent at this point will turn back and visit [1,2]
- At [1,2] agent receives [Stench, None, None, None, None]
  - This means Wumpus is there in one of [1,1] or [2,2] or [1,3]
- Wumpus cannot be in [1,1]
- Can Wumpus be in [2,2] ?
  - (No)
- So Agent can deduce that Wumpus is in [1,3]
- Also, [2,2] is OK (otherwise there would be breeze), so [3,1] has a pit
- This inference needs to combine knowledge gained in different times and different places

The Wumpus world



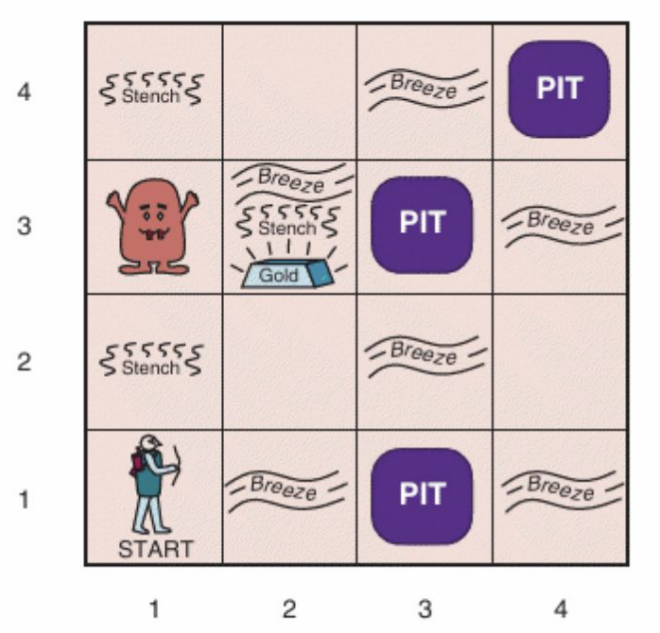
1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 A S OK	2,2 OK	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

1,4	2,4 P?	3,4	4,4
1,3 W!	2,3 A S G B	3,3 P?	4,3
1,2 S V OK	2,2 V OK	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

- Agent now moves to the next unexplored OK square [2,2] where she receives [None, None, None, None, None]
- So [2,3] and [3,2] are OK
- Agent say moves to [2,3]
- Agent detects GLITTER, so agent grabs gold, returns to [1,1] and climbs out of cave
  - Agent can climb out only from [1,1]
  - So agent should find a path from current position to [1,1]

The Wumpus world

- Agent makes conclusions at every step from the available information
- At each step the conclusions are guaranteed to be correct if the available information is correct
- How can the agent represent these information and draw conclusions?



Knowledge Base

- What to store in the Knowledge Base?
- English sentences?
  - “If there is breeze in a square then there is a pit in at least one of the adjacent squares”
  - Not machine friendly / ambiguous
- Mathematical Logic gives unambiguous way to represent information
  - Meaning is clear from the statement
  - Amenable to make deductions and inferences
- Most commonly used Logical Formalisms:
  - Propositional Logic
  - First Order Logic / Predicate Logic
  - Modal Logic : Epistemic / Temporal ...

Logic

- Syntax of the Logic : Specifies what are all the well formed formulas
  - Example : In arithmetic, x+y = 4 is well-formed but x += is not well-formed
  - In propositional logic : p Λ q is well-formed but pq Λ is not well-formed

- Semantics defines the truth of a well-formed formula in each possible world.
  - Example :
    x+y=4 is TRUE in a world where x = 3 and y =1 (and x=y=2 etc.), but false in x = 1 and y=1
  - In propositional Logic : p Λ q is true in a world where both p and q are TRUE but FALSE in all the worlds where at least one of the two is FALSE

- A Model is Mathematical Abstraction of the real world
  - A well formed formula is evaluated to TRUE/FALSE in a given model
  - If a formula α and a Model m, if α evaluates to TRU
    m satisfies α or m is a model of α
  - $M(\alpha) = \{ m \mid m \text{ satisfies } \alpha \}$ is the set of all models in