

Foundations of Artificial Intelligence

Unit 1 - Foundations

* What is AI?

Definition can be on different grounds:

- (i) Thinking humanly - automation of activities we associate with human thinking
- (ii) Thinking rationally - the study of mental faculties using computational models
- (iii) Acting Humanly - The art of creating machines that perform things that require intelligence when performed by people.
- (iv) Acting rationally - concerned with intelligent behavior in artifacts

(*) Acting humanly - The Turing Test

- designed to provide a satisfactory operational definition of intelligence
- A computer passes the test if the human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or a computer.
- The computer would need the following capabilities:

- { (i) NLP
- (ii) knowledge representation
- (iii) automated reasoning
- (iv) machine learning

Total Turing Test Includes a video signal so that the interrogator can test the subject's perceptual abilities, and for the computer interrogator to pass physical objects.

To pass the total Turing test - the computer needs:

- (i) computer vision
- (ii) robotics

* Thinking humanity - cognitive modeling approach

Cognitive science - study of the human mind and brain, focusing on how the mind represents and manipulates knowledge, and how mental representations & processes are realized in the brain

* Thinking rationally - the laws of thought approach

Syllogism - a form of reasoning in which a conclusion is drawn from

2 given or assumed propositions

e.g. Socrates is a man - + -

Logic - a precise notation for statements about all kinds of objects in the world and the relations between them

2 issues w/ logic:

(i) not easy to take informal knowledge & state it in formal terms

(ii) even problems with just a few hundred facts can exhaust the computational resources of any computer

* Acting Rationally - the rational agent approach

Agent - An agent is something that acts. Computer agents operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, and create & pursue goals.

Rational agent one that acts so as to achieve the best outcome / best expected outcome

(OR) For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given whatever built-in knowledge the agent has.

* AI Applications

- (i) robotic vehicles
- (ii) speech recognition
- (iii) game playing
- (iv) computer vision

* Types of AI Research

→ can be theoretical or experimental

→ Two major lines are:

(i) biological - AI should study humans and imitate their psychology & physiology

(ii) phenomenal - based on studying & formalizing common sense facts.

* Branches of AI

- (i) Logical AI
- (ii) search
- (iii) NLP
- (iv) pattern recognition
- (v) epistemology - study of the kinds of knowledge required for solving problems
- (vi) ontology - study of the kind of objects that exist

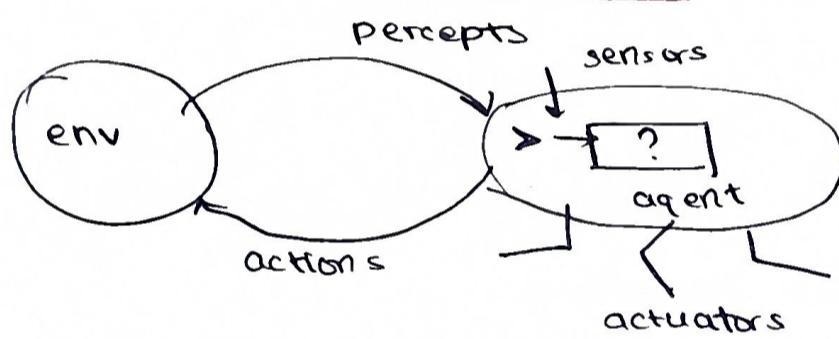
I-2 Intelligent Agents

- * **Agents** (Defn 2) - anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators. An agent is autonomous if its behavior is defined by its own experience.
- eq. in a human agent

sensors = eyes, ears, other organs

actuators, legs, mouth

Agents and Environments



Agent function: maps from percept histories to actions

$$f: P^* \rightarrow A$$

The agent program runs on the physical architecture to produce f.

agent = architecture + program

Example: In the world of a vacuum cleaner:

Percepts = location & content - eq. [A, Dirty]

Actions = left, right, suck, NoOp

- * Performance Measure - An objective criterion for the success of an agent's behavior

eq. for a vacuum cleaner - the measure could be the amount of dirt cleaned up, amount of time taken, electricity consumed etc.

PEAS

Performance measure, environment, actuators, sensors

→ Those settings must be specified for intelligent agent design

Ex1 Design of an automated taxi driver.

- Performance measure - safe, fast, max profits
- environments - road, traffic, pedestrians
- actuators - steering wheel, accelerator, brake
- sensors - camera, sonar, GPS, speedometer, odometer

Environment Types

- (i) Fully Observable (vs. partially observable) - An agent's sensors give it access to the complete state of the environment at each point ~~in time~~.
- (ii) Deterministic (vs. stochastic) - The next state of the environment is completely determined by the current state and the action executed by the agent.
 If the environment is deterministic except for the actions of other agents, then the environment is strategic.
- (iii) Episodic (vs. sequential) - The agent's experience is divided into atomic episodes, and the choice of action in each episode depends only on the episode itself.
- (iv) Static (vs. dynamic) - environment is unchanged while an agent is deliberating.
 semidynamic - if environment does not change w/ time but agent's performance scale does

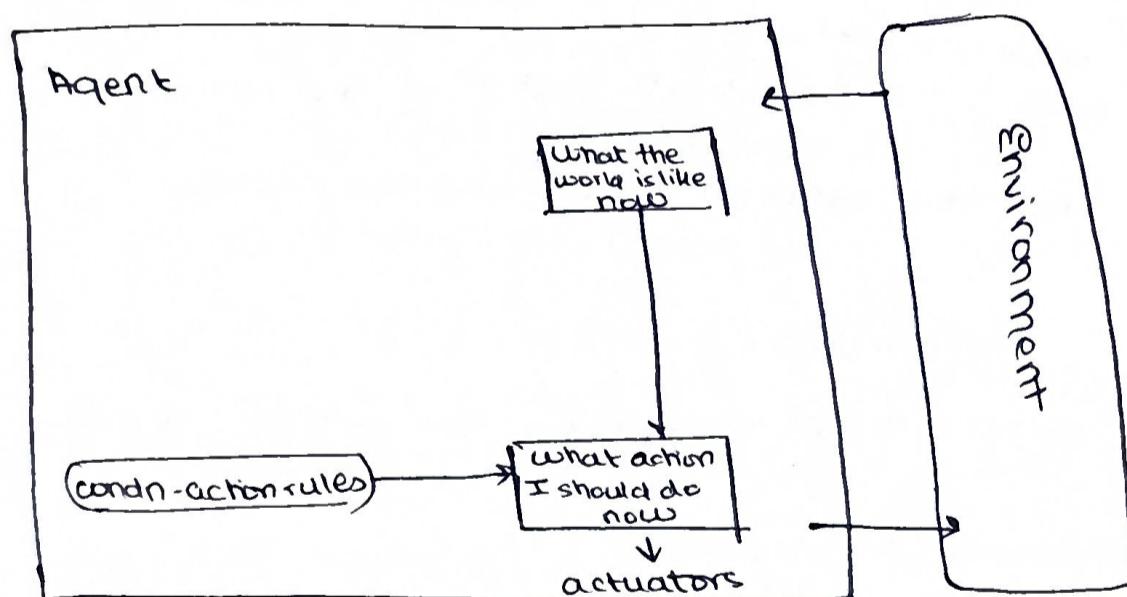
- (v) Discrete (vs. continuous) - A limited no. of distinct, clearly defined percepts and actions
- (vi) Single agent (vs. multiagent) - An agent operating by itself in an environment.

* Agent Types

Four basic types in order of increasing generality:

- (i) simple reflex agents
- (ii) model-based reflex agents
- (iii) goal-based agents
- (iv) utility-based agents.

A. Simple Reflex Agents



function SIMPLE-X. REFLEX-AGENT (percept)

 returns action

static: rules, a set of condition-action rules

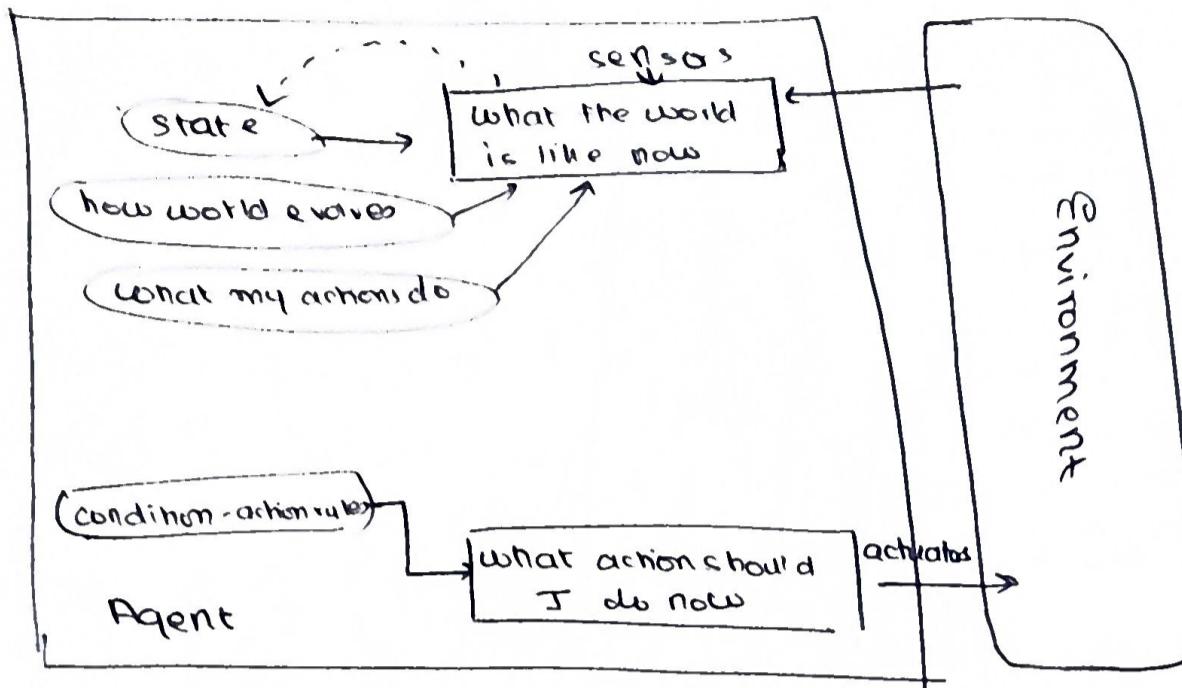
state \leftarrow INTERPRET INPUT (percept)

rule \leftarrow RULE.MATCH (state, rules)

action \leftarrow RULE.ACTION [rule]

return action

B. Model-based reflex agents



Function REFLEX, AGENT. WITH-STATE (percept)

return action

Static : state, a description of the current world state.

rule , a set of condition - action rules

state \leftarrow UPDATE - STATE (state, percept)

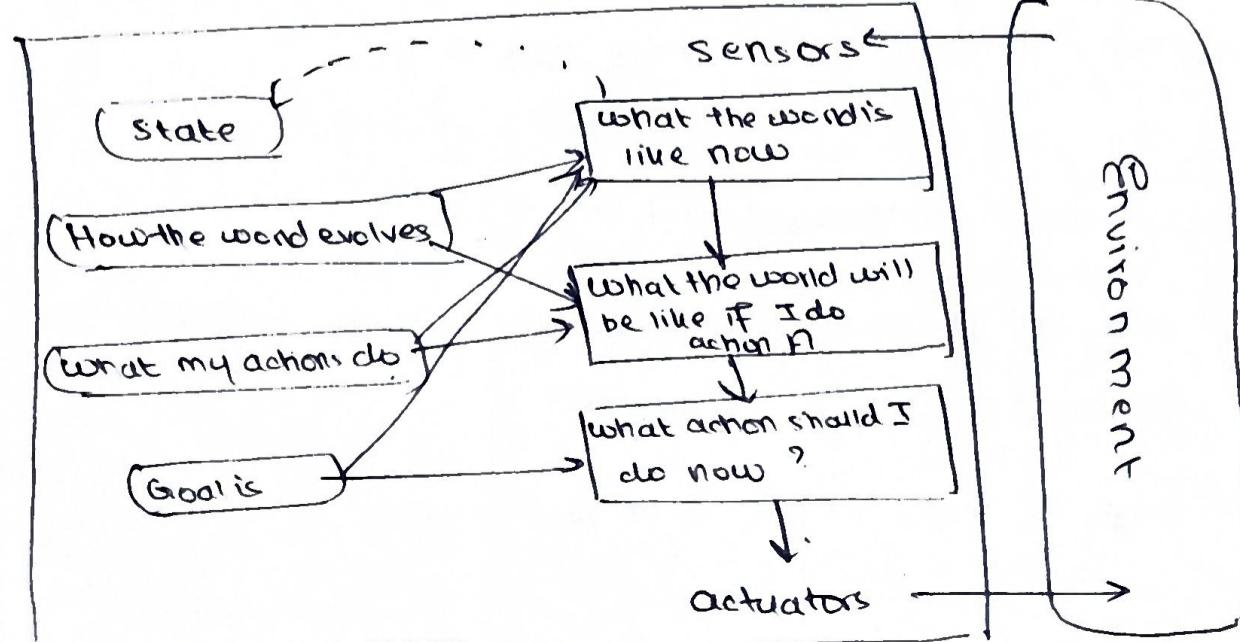
rule \leftarrow RULE - MATCH (state, rules)

action \leftarrow RULE - ACTION [rule]

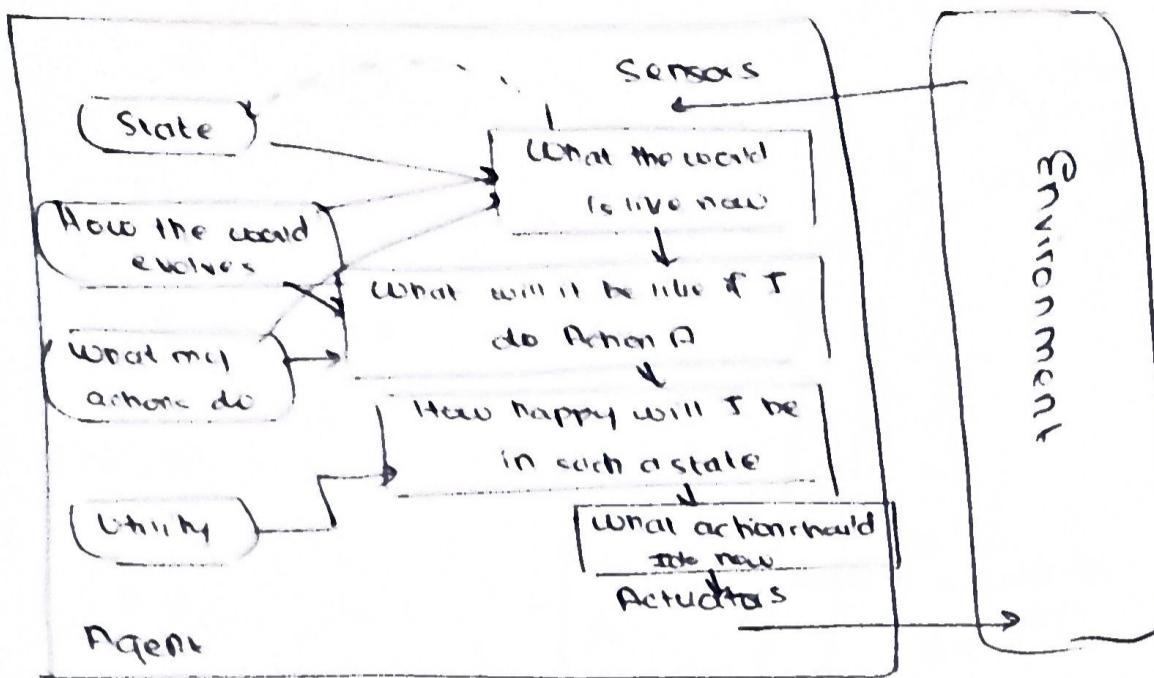
state \leftarrow UPDATE - STATE (state, action)

return action

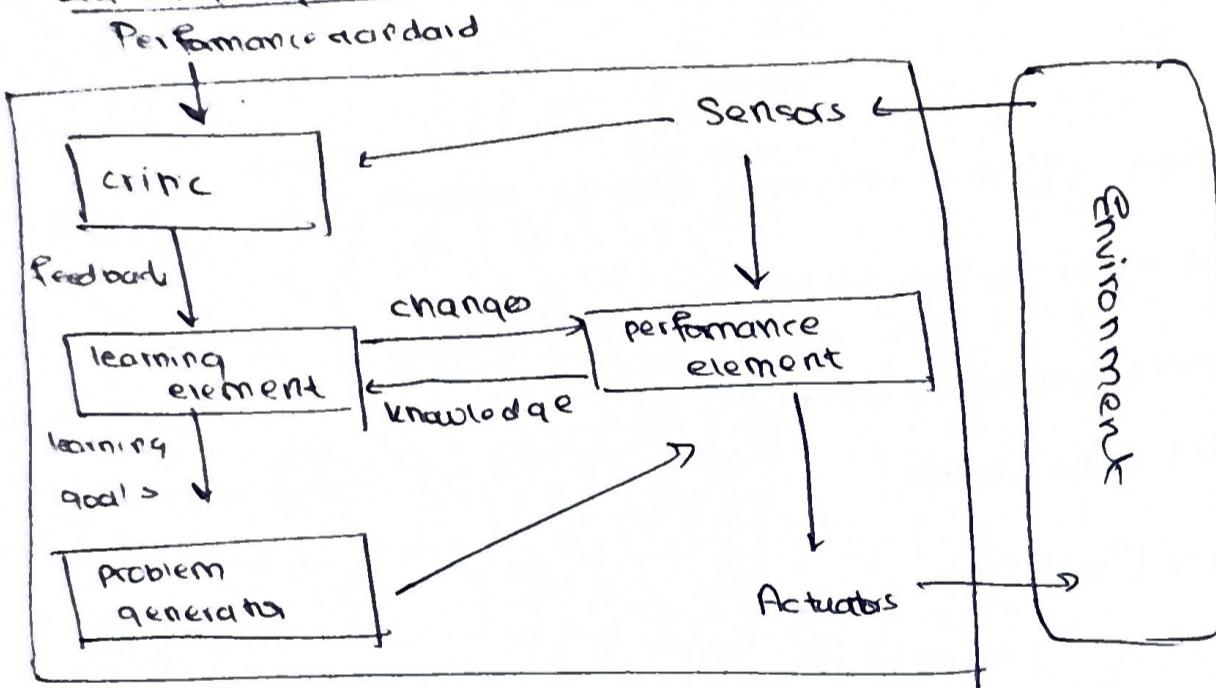
C. Goal-based Agents



D. Utility-based agents



E. Learning Agents



WeakAI - designated to perform a specific set of tasks

weak because it is limited in its capabilities

e.g. speech recognition, image processing, NLP

StrongAI - called artificial general intelligence (AGI)

meant to mimic the cognitive abilities of a human being

can perform a wide range of tasks

→ can learn & adapt to new env. & info

Environment Types - Examples

- ① Fully Observable vs. Partially Observable

- ② Deterministic vs. Stochastic

- ③ Episodic vs. Sequential

↓ ↓

place 2 checkers / chess

pick robot

- ④ Static vs. Dynamic

↓ ↓

entering an roller coaster

empty house

- ⑤ Discrete vs. Continuous
↓ ↓
chess self-driving car actions

- ## (6) Single Agent vs. multiagent → swarm

→ others
self-driving cars
competitive, collaborative

* known vs. unknown