

- The agent function maps from percept histories to actions
- For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.
- Agents can perform actions in order to modify future percepts so as to obtain useful information (information gathering, exploration)
- An agent is autonomous if its behavior is determined by its own experience (with ability to learn and adapt)
- PEAS: Performance measure, Environment, Actuators, Sensors

Task Environments & their properties

- Fully vs. partially observable vs. unobservable
- Single vs. multi-agent
- Competitive vs. cooperative & partially cooperative (partially competitive)
- Deterministic vs. non-deterministic vs. stochastic
- Episodic vs. Sequential
- Static vs. Dynamic
- Discrete vs. continuous
- Known vs. unknown

Different types of Agents

- Simple reflex agent: These agents select actions on the basis of the current percept, ignoring the rest of the percept history. These agents work only when the environment is fully observable.
- Model-based reflex agent: The knowledge about “how the world works” — whether implemented in simple Boolean circuits or in complete scientific theories — is called a model of the world. This agent uses a state which is a function of current state, action, percept and rules. It uses this state to decide the next action.
- Goal-based agent: It keeps track of the world state as well as a set of goals it is trying to achieve, and chooses an action that will (eventually) lead to the achievement of its goals.
- Utility-based agent: It uses a model of the world, along with a utility function that measures its preferences among states of the world. Then it chooses the action that leads to the best expected utility, where expected utility is computed by averaging over all possible outcome states, weighted by the probability of the outcome.
- General Learning Agent: A learning agent can be divided into four conceptual components
 - **learning element**, which is responsible for making improvements
 - **performance element**, which is responsible for selecting external actions. The performance element is what we have previously considered to be the entire agent: it takes in percepts and decides on actions.
 - The learning element uses feedback from the **critic** on how the agent is doing and determines how the performance element should be modified to do better in the future.
 - **Problem generator**, It is responsible for suggesting actions that will lead to new and informative experiences.

Ways to represent states and transition between them

- **Atomic representation**: a state (such as B or C) is a black box with no internal structure.
- **Factored representation**: a state consists of a vector of attribute values; values can be Boolean, realvalued, or one of a fixed set of symbols.
- **Structured representation**: a state includes objects, each of which may have attributes of its own as well as relationships to other objects.

Logic

- Well formed formula (wff): An atomic proposition is a wff
- Logical inference is used to create new sentences that logically follow from a given set of propositional logic sentences (Knowledge Base, KB).
- sound - inference rule does not create any contradictions
- complete - inference rule is able to produce every expression that logically follows from (is entailed by) the KB.

Interpretation for propositional logic is a mapping assigning a truth value to each of the simple sentences of the language.

- Satisfiability: We say that an interpretation i satisfies a sentence \iff it is true under that interpretation
- Validity: A sentence is valid \iff it is satisfied by every interpretation.
- Unsatisfiability: \iff it is not satisfied by any interpretation

We would like to know that given some sentences, whether other sentences are or are not logical conclusions. This relative property is known as logical entailment. Formally, a set of sentences Δ logically entails a sentence φ (written $\Delta \models \varphi$) if and only if every interpretation that satisfies Δ also satisfies φ . For example, the set of sentences $\{p, p \vee q\}$ logically entails the sentence (q) .

CNF Form

- Eliminate \iff , replacing $\alpha \iff \beta$ with $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$.
- Eliminate \rightarrow , replacing $\alpha \rightarrow \beta$ with $\neg\alpha \vee \beta$
- Requires \neg to appear only in literals, so we “move \neg inwards” i.e. distribute it in a bracket.
- Distribute \vee over \wedge , so final looks like $(\alpha \vee \beta) \wedge (\alpha \vee \gamma) \iff (\alpha \vee (\beta \wedge \gamma))$

To show that $KB \models \alpha$, we show that $(KB \wedge \neg\alpha)$ is unsatisfiable

- CNFSentence $\rightarrow Clause_1 \wedge \dots \wedge Clause_n$
- Clause $\rightarrow Literal_1 \wedge \dots \wedge Literal_m$
- Literal $\rightarrow Symbol | \neg Symbol$
- Symbol $\rightarrow \$P | Q | R | \dots \$$
- DefiniteClauseForm $\rightarrow (Symbol_1 \wedge \dots \wedge Symbol_l) \rightarrow Symbol$
- GoalClauseForm $\rightarrow (Symbol_1 \wedge \dots \wedge Symbol_l) \rightarrow False$
- HornClauseForm $\rightarrow DefiniteClauseForm | GoalClauseForm$

Read forward chaining and backward chaining