

1. (21 points) For each of the following assertions, say whether it is true or false.

(a) **T** One of our goals in AI is to design an agent program that implements the agent function—the mapping from percepts to actions.

(b) **T** A rational agent must behave as if it possesses a utility function whose expected value it tries to maximize.

*Explanation.* (See Section 2.4.5 of the text.)

(c) **F** An agent that senses only partial information about the state cannot be perfectly rational.

*Explanation.* Rationality does not require the task environment be fully observable. Partial observability and nondeterminism are ubiquitous in the real world, and so, therefore, is decision making under uncertainty. A rational utility-based agent chooses the action that maximizes the expected utility of the action outcomes—that is, the utility the agent expects to derive, on average, given the probabilities and utilities of each outcome. (See Section 2.4.5 of the text.)

(d) **F** Every agent is rational in an unobservable environment.

*Explanation.* It is easy to come up with irrational agents in any environment. In general, once we specify a performance measure and utility function (e.g., minimize fuel consumption), we can then create an agent with poor performance/utility (e.g., recklessly waste fuel by revving the engine, spinning wheels, and burning rubber while doing nothing useful).

(e) **T** There exist pure reflex agents that do not behave rationally in some task environments.

*Explanation.* Pure reflex agents ignore consider only the present percept values, ignoring all past historical percept values. Consider a simple reflex vacuum cleaner agent with only a dirt sensor and no location sensor. If it is on a clean square, what should it do? If it moves left in response to a clean tile, then when it's on tile A it is stuck forever. If it moves right on a clean tile, then when it's on tile B it is stuck forever. (See Section 2.4.1 for more details and examples.)

(f) **T** There exists a task environment in which every agent is rational.

*Explanation.* Consider, for example, an environment in which all actions have the same reward.

(g) **F** A perfectly rational poker-playing agent never loses.

*Explanation.* In this case, the answer is clear because of the stochastic nature of the environment, which makes it possible for an unskilled novice to occasionally get lucky and win against a strong, e.g., perfectly rational opponent.

(h) **T** In general, an agent's choice of action at any given instant can depend on its built-in knowledge and on the entire percept sequence observed to date.

*Explanation.* (See Section 2.1 of the text.)

(i) **T** The agent function is an abstract mathematical description; the agent program is a concrete implementation, running within some physical system.

*Explanation.* (See Section 2.1 of the text.)

- (j) **F** Every agent function is implementable by some program/machine combination.

*Explanation.* An agent function may, for some input, enter an infinite loop or cause the computer to run out of memory or disk space before completing its task.

- (k) **T** Suppose an agent selects its action uniformly at random from the set of possible actions. There exists a deterministic task environment in which this agent is rational.

*Explanation.* Consider (as in Q 1.5) an environment in which all actions have the same reward.

- (l) **T** It is possible for a given agent to be perfectly rational in two distinct task environments.

*Explanation.* Consider the vacuum cleaner agent. One environment goal could be to keep each square clean; the other environment goal could be to eliminate dirt. Assuming that, in both environments, the only thing making the squares unclean is dirt, then, for both environments the rational thing to do would be to clean when the agent detects dirt.

- (m) **F** The “transition model” of the world depends on how the world evolves independently of the agent. It does not consider the affects of the agent’s action on the world.

*Explanation.* The “transition model” is based on information about how the world changes over time, which can be divided roughly into two parts: how the world evolves independently of the agent *and* the effects of the agents’ actions on the world. (See Section 2.4.3 of the text.)

2. (3 points) Consider a simple thermostat that turns on a furnace when the temperature is at least 3 degrees below the setting, and turns off a furnace when the temperature is at least 3 degrees above the setting. Is a thermostat an instance of a simple reflex agent, a model-based reflex agent, or a goal-based agent?

- ☒ **simple reflex agent**
- ☐ model-based reflex agent
- ☐ goal-based agent

3. (4 points) How would you characterize the task environment of a crossword puzzle?

- ☒ **fully observable**
- ☐ partially observable
- ☐ unobservable

4. (4 points) An behavior (action) of an agent is described by which of the following?

- ☐ Perception sequence
- ☒ **Agent function**
- ☐ Sensors and Actuators
- ☐ Environment in which agent is performing