Learning Agents

MSE 2400 EaLiCaRA
Dr. Tom Way

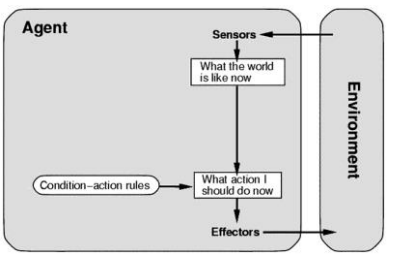
Learning Agent

- An Agent that observes its performance and adapts its decision-making to improve its performance in the future.

Agent

- Something that does something
- Computational Agent – a computer that does something

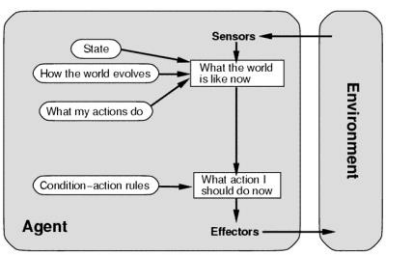
Simple Reflex Agent



Simple Reflex Agent

- The action to be selected only depends on the most recent percept, not a percept sequence
- As a result, these agents are stateless devices which do not have memory of past world states

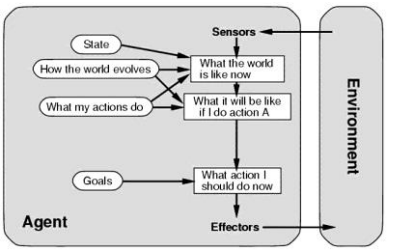
Model-based Reflex Agent



Model-based Reflex Agent

- Have internal state which is used to keep track of past states of the world (i.e., percept sequences may determine action)
- Can assist an agent deal with at least some of the observed aspects of the current state

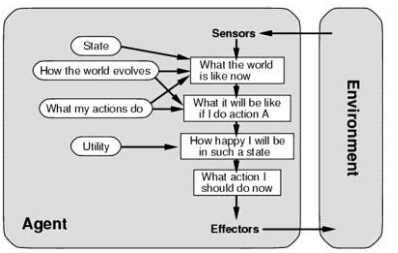
Goal-based Agent



Goal-based Agent

- Agent can act differently depending on what the final state should look like
- Example: automated taxi driver will act differently depending on where the passenger wants to go

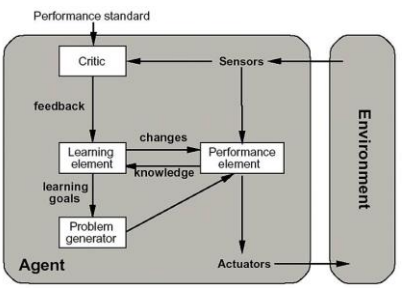
Utility-based Agent



Utility-based Agent

- An agent's utility function is an internalization of the performance measure (which is external)
- Performance and utility may differ if the environment is not completely observable or deterministic

Learning Agent (in general)



Learning Agent Parts (1)

- Environment – world around the agent
- Sensors – data input, senses
- Critic – evaluates the input from sensors
- Feedback – refined input, extracted info
- Learning element – stores knowledge
- Learning goals – tells what to learn

Learning Agent Parts (2)

- Problem generator – test what is known
- Performance element – considers all that is known so far, refines what is known
- Changes – new information
- Knowledge – improved ideas & concepts
- Actuators – probes environment, triggers gathering of input in new ways

Intelligent Agents should…

- accommodate new problem solving rules incrementally
- adapt online and in real time
- be able to analyze itself in terms of behavior, error and success.
- learn and improve through interaction with the environment (embodiment)
- learn quickly from large amounts of data
- have memory-based exemplar storage and retrieval capacities
- have parameters to represent short and long term memory, age, forgetting, etc.

Classes of Intelligent Agents (1)

- Decision Agents – for decision making
- Input Agents - that process and make sense of sensor inputs (neural networks)
- Processing Agents - solve a problem like speech recognition
- Spatial Agents - relate to physical world

Classes of Intelligent Agents (2)

- World Agents - incorporate a combination of all the other classes of agents to allow autonomous behaviors
- Believable agents - exhibits a personality via the use of an artificial character for the interaction

Classes of Intelligent Agents (3)

- Physical Agents - entity which percepts through sensors and acts through actuators.
- Temporal Agents - uses time based stored information to offer instructions to a computer program or human being and uses feedback to adjust its next behaviors.

How Learning Agents Acquire Knowledge

- Supervised Learning
  - Agent told by teacher what is best action for a given situation, then generalizes concept F(x)
- Inductive Learning
  - Given some outputs of F(x), agent builds h(x) that approximates F on all examples seen so far is SUPPOSED to be a good approximation for as yet unseen examples

How Learning Agents Acquire Concepts (1)

- Incremental Learning: update hypothesis model only when new examples are encountered
- Feedback Learning: agent gets feedback on quality of actions it chooses given the h(x) it learned so far.

How Learning Agents Acquire Concepts (2)

- Reinforcement Learning: rewards / punishments prod agent into learning
- Credit Assignment Problem: agent doesn't always know what the best (as opposed to just good) actions are, nor which rewards are due to which actions.

Examples

- Eliza - http://www.simonebaldassarri.com/eliza/eliza.html
- Mike - http://www.rong-chang.com/tutor_mike.htm
- iEinstein - http://www.pandorabots.com/pandora/talk?botid=ea77c0200e365cfb
- More Cleverbots - https://www.existor.com/en/
- Chatbots - http://www.chatbots.org/