The UNIVERSITY of VERMONT

Benjamin Fenster
CS-251 (Wu)
15 December 2005

Notes – Introduction

I. Four Interpretations
   a. Model how human brains *think*
   b. Get computers to *act* like humans
   c. Think in a *rational* way (the goal in CS-251)
   d. Act rationally

II. Turing Test
   a. Is a machine intelligent?
   b. Prediction: By 2000, a machine could have a 30% chance of fooling a layperson for five minutes (this prediction made by Turing when "creating" the test)
   c. Components
      i. Knowledge
      ii. Reasoning: We may both know Algebra, but how fast can we each solve a problem?
      iii. Language
      iv. Understanding
      v. Machine Learning
   d. Problems
      i. Not reproducible (will give the same answers to the same questions every time)
      ii. Not constructive
      iii. Not amenable to mathematical analysis
   e. Cognitive Science (thinking humanly)
      i. Studying how humans think
      ii. Cognitive neuroscience: More about the hardware of the brain
   f. Thinking Rationally
      i. All programming languages we know are prescriptive.
      ii. We're working toward *descriptive* languages in this class (Prolog)
      iii. Problems: Not all intelligent behavior is determined by logical deliberation.
   g. Acting Rationally: Do the "right thing" (maximize goal achievement given all available information)

III. Rational Agents
   a. This is the rest of the course.
   b. An agent is a function from "percept histories" (i.e. all known information) to actions.
   c. For different environments and tasks, find the agent with the best performance, since not all information is known.
   d. But: Computational limitations make considering *all* possibilities infeasible
   e. Intelligence Consists Of
      i. Symbolic *inference*
         1. If you're too proud, you'll fail
         2. Failure is the mother of success
         3. Therefore: If you're proud, you'll succeed
      ii. Heuristics
   f. Human vs. Machine Intelligence
      i. Machines have an advantage: speed
      ii. AI Cannot yet match humans in emotion and creativity

IV. Agents and Environments
   a. Environment generates percepts
   b. Received by sensors
   c. Agent function maps percept history to actions
   d. Actuators then affect the environment
   e. Example: Vacuum Cleaner World
      i. Have two squares of carpet and want to build a vacuum agent.
      ii. Percepts: Location of the vacuum, status/contents of that location (Square B, Dirty)

iii. Actions: Move left, move right, suck, No-Op.
iv. Sequence (Precept Sequence) -> Action
    1. [A, Clean] -> Right
    2. [A, Dirty] -> Suck
    3. [A, Clean], [A, Clean] -> Right
    4. Is this the *right* function to use? Hard to say.

f. Rationality
    i. Need a performance measure to evaluate the environment sequence
    ii. "Rational" does not mean omniscient/clairvoyant/successful
    iii. Rational does mean explorative (try to find good solutions), learning, autonomous

g. PEAS: Four components of agent design
    i. Performance Measure
    ii. Environment (factors to consider)
    iii. Activators
    iv. Sensors
    v. Example: Automated Taxi
        1. P: Profit, comfort, safety
        2. E: Streets, freeways, traffic, weather
        3. A: Steering, acceleration/deceleration, horn
        4. S: GPS, video, accelerometer
    vi. Example: Internet Shopping
        1. P: Price, search time
        2. E: Terminal, Internet connection speed
        3. A: Mouse, keyboard
        4. S: Terminal, browser

h. Environment Types
    i. Observable: Can you observe (know) the environment? Can you get a clear picture of everything in the environment?
    ii. Deterministic: Given the same set of information twice, will you make the same choices twice?
    iii. Episodic (or sequential): Sequential means taking a series of steps (e.g. I'm hungry; make one decision)
    iv. Static: Information doesn't change (vs. dynamic)
    v. Discrete: Values are discrete or continuous
    vi. Single-Agent
    vii. Environment type largely determines agent design

i. Agent Types
    i. Simple reflex agents: respond to the environment only. Use a set of rules (a lookup table)
    ii. Reflex with states (same as above, but use a different set of rules depending on the agent's state)
    iii. Goal-based agents: Reach a goal through a sequence of actions
    iv. Utility-based: The best way of reaching a goal
    v. Any type of agent can also be *learning*, so it modifies its behavior over time