Artificial Intelligence and its Social/Ethical Implications

Topic Overview

- Brief History of Artificial Intelligence
- A.I. Research v. Public Interest
- Social Implications of A.I.
- What If… A.I. Surpasses Humanity
- Closing Remarks

Brief History of Artificial Intelligence

➤ Birth of Artificial Intelligence was in 1956 at the Dartmouth Conference

➤ Based on two theories:

  ➤ Principle of feedback – all intelligent behavior is the result of a feedback mechanism

  ➤ Logic Theorist – program that represented every problem as a tree; would attempt to select the branch that would most likely result in the solution

Brief History of Artificial Intelligence (cont.)

- One decade later, centers for Artificial Intelligence began to form at Carnegie Mellon and MIT

- LISP language was created – choice among Artificial Intelligence developers

- Expert Systems – Predict probability based on a certain set of conditions

- Present state of the art at MIT – Coco the humanoid robot

A.I. Research and Public Interest

- Code of Ethics - Public Interest
  - Accept full responsibility for their own work
  - Moderate interests of engineer, employer, client and users with that of the public
  - Approve software only if believed to be safe
  - Disclose any actual or potential danger to the public if believed to be associated with the software
  - Cooperate in efforts to address grave public concern caused by software

Social Implications of A.I. Systems

➢ People could easily become fully dependent on machines

➢ Possibility of increased unemployment rate

Implications of A.I. Surpassing Humanity

- What rights would machines have
- What moral obligations would Humanity have towards machines (and vice-versa; machines to Humanity)
- We could possibly end up the lower-order creatures – how would Humanity be treated?
- "Hollywoodized" example – <u>The Matrix</u>

Closing Remarks

- General questions you might have
  - Do we want to build a computer that will be like us?
  - If so, what do we need them for?
  - What will the human-computers do for humanity?
- No answers to these questions…yet research and achievement continues to progress each year - We must wait and see what the future holds