Artificial morality can be implemented using several different approaches

- Top-down approaches are based on rules
- Bottom-up approaches are learned by the system based on feedback
- Hybrid approaches incorporate aspects of both top-down and bottom-up approaches

Top-down approaches are based on rules

- The rules could be based on human ethical systems, such as the Golden Rule or the Ten Commandments
- Rules could be derived from either effects ("consequentialist" – e.g. utilitarianism) or intentions ("deontological" – e.g. Kant)
- In practice, computation based on such rules could be challenging

Bottom-up approaches are learned by the system based on feedback

- The system should gradually adjust its behavior based on rewards or selective pressure
- Examples include:
  - Turing's 1950 discussion of educating a computer like a child
  - Wilson's sociobiology investigations demonstrating some kind of moral behavior as a result of evolution
- The concept of "moral grammar" could be important
- Because machine learning systems are imperfect, unexpected (or even dangerous) behavior could occur with only a small change in inputs

Hybrid approaches incorporate aspects of both top-down and bottom-up approaches

- Possible implementations of hybrid approaches include: “Von Neumann Machines and neural networks, genetic and learning algorithms, rule and natural language parsers, virtual machines and embodied robots”
- No details of how to implement hybrid approaches are given