Knowledge Representation and Reasoning

University "Politehnica" of Bucharest
Department of Computer Science

Fall 2012

Adina Magda Florea

The AI Debate

Language learning

Strong AI vs. Weak AI

- What is Strong AI?
- What is Weak AI?

Strong AI

- **Strong AI** is artificial intelligence that matches or exceeds human intelligence
- The intelligence of a machine can successfully perform any intellectual task that a human being can
- Advocates of "Strong AI" believe that computers are capable of true intelligence
- They argue that what intelligence is strictly algorithmic, i.e., a program running in a complex, but predictable, system of electro-chemical components (neurons).

Strong AI

- Many supporters of strong AI believe that the computer and the brain have equivalent computing power
- With sufficient technology, it will someday be possible to create machines that have the same type of capabilities as humans
- However, Strong AI's reduction of consciousness into an algorithm is difficult for many to accept

Weak AI

- The **Weak AI** thesis claims that machines, even if they *appear* intelligent, can only *simulate* intelligence [Bringsjord 1998]
- They will never actually be aware of what they are doing
- Some weak AI proponents [Penrose 1990] believe that human intelligence results from a superior computing mechanism which, while exercised in the brain, will never be present in a Turing-equivalent computer

Weak AI

- To promote the weak AI position, John R. Searle, a prominent and respected scholar in the AI community, offered the "Chinese room parable"
  - John R. Searle. "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3:417-57. 1980.
- In 1990, Roger Penrose published *The Emperor's New Mind* , a 450 page book which has been viewed by many as an attack on strong AI
- Aaron Sloman. "The emperor's real mind: review of Roger Penrose's The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics," *Artificial Intelligence*. Vol. 56. 1992. - a critique of the Penrose book.

The Singularity

- The concept of ultrasmart computers — machines with **“greater than human intelligence”** — was dubbed “The Singularity” in a 1993 paper by the computer scientist and science fiction writer **Vernor Vinge**.
- He argued that the acceleration of technological progress had led to “the edge of change comparable to the rise of human life on Earth.”

The Singularity

- The artificial-intelligence pioneer Raymond Kurzweil took the idea one step further in his 2005 book, "The Singularity Is Near: When Humans Transcend Biology."
- He sought to expand Moore's Law to encompass more than just processing power and to simultaneously predict with great precision the arrival of post-human evolution, which he said would occur in 2045.

The Singularity University

- Raymond Kurzweil is the co-founder of Singularity University, a school supported by Google that opened in June 2009 with a grand goal — to "assemble, educate and inspire a set of leaders who strive to understand and facilitate the development of exponentially advancing technologies and apply, focus and guide these tools to address humanity's grand challenges."

Open questions

- Is strong AI possible? Yes, No, Why?
- Is intelligence characterized by *deduction* (like one person thinking); *two-way interaction* (like two people talking); or *multi-way interaction*?
- Is there disembodied intelligence?
- Is language essential for intelligence? (what kind of language?)

Open questions

- Are insects intelligent?
- Can a machine feel?
- Should an AI program / machine have rights?

- Other questions?

Language evolution

- Language is dynamic
- Luc Steels – an experiment = What mechanisms for open-ended language can be grounded in situated embodied interactions by performing computer simulations and doing experiments with physical autonomous robots [Steels, 2001].
- A population of agents will **self-organise a perceptually grounded ontology and a lexicon from scratch**, without any human intervention.

Language game

- A *language game* is a routine interaction between a speaker and a listener out of a population whose members have regular interactions with each other
- Each individual agent in the population can be both speaker and hearer
- The game has a non-linguistic goal, which is some situation that speaker and hearer want to achieve cooperatively

Language game

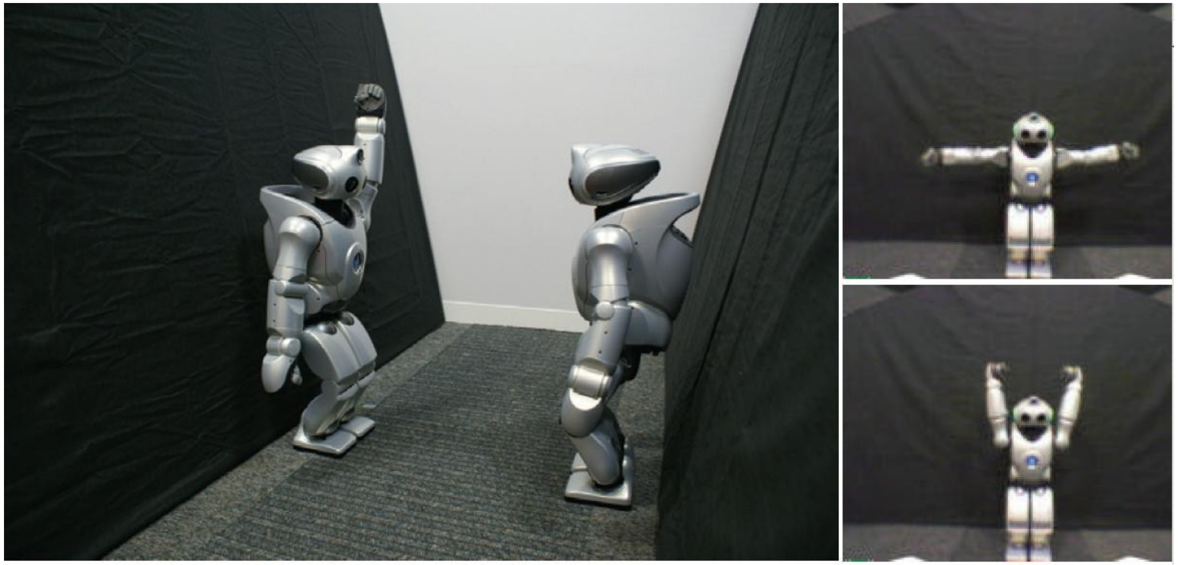
- Speaker and hearer can use **bits of language** but they can also use **pointing gestures and non-verbal interaction**, so that not everything needs to be said explicitly.
- A typical example of a language game is the *Color Naming Game* - a game where the speaker uses a color to draw the attention of the hearer to an object in the world [Steels and Belpaeme, 2005].

Action game

- **Action Game** (Luc Steels) – one robot asks another robot to take on a body posture (such as stand or sit)
- Evolve an ontology and lexicon for body postures and the visual image schemata they generate
- Words such as “stand”, “sit” and “lie”
- Two humanoid robots face each other and play the Action Game

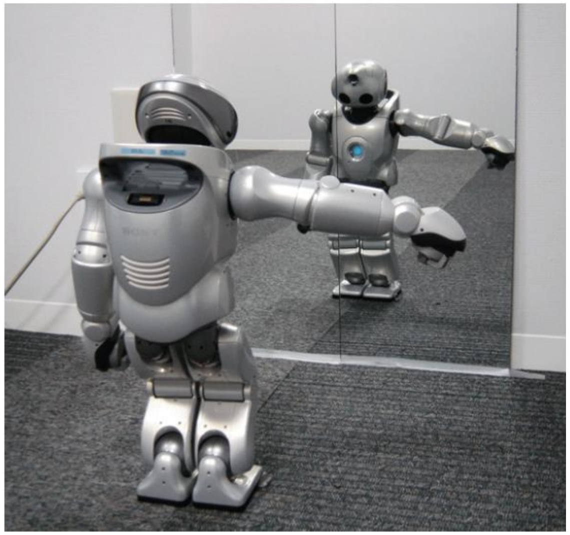
Action game

- One robot (the speaker) asks another robot (the hearer) to perform an action.
- The speaker then observes the body posture achieved by the action and declares the game a success if the body posture is the desired one.
- Otherwise the speaker provides feedback by doing the action himself.



First experiment

- Kinesthetic teaching - the robot can acquire by itself the right motor commands to achieve a particular body posture
- Kinesthetic teaching means that the experimenter moves the robot's body from a given position to a target body posture.



First experiment

- A population of 10 agents
- Each individual has coordinated motor behavior and visual bodyimage through the mirror for 10 postures.
- 100 % success is reached after about 2000 games.
- After 1000 games, which means 200 games per agent, there is already more than 90 % success.

Second experiment

- The robots no longer use a mirror to learn about the relation between a visual body image schema and their own bodily action
- They coordinate this relationship through language
- Language will enforce coordination
- If a speaking robot *R*1 asks *R*2 to achieve a posture P using a word W, the game will only be successful if for *R*2, W is associated with an action A so that P is indeed achieved

Second experiment

- A population of 5 agents and 5 postures
- 100 % success is reached after about 4000 games (1600 games per agent) and stays stable

- The speed of convergence can be improved significantly if the hearer uses his own selfbody model (a stick-figure simulation of the impact of motor commands on body parts) in order to guess which actions have to be performed in order to reach the body posture that is shown by the speaker [Steels and Spranger, 2008b]