Artificial Intelligence: Intelligent Autonomous Agents & Multi-Agent Systems

Mike Barley

Pat Riddle

Ian Watson

Artificial Intelligence - **I**
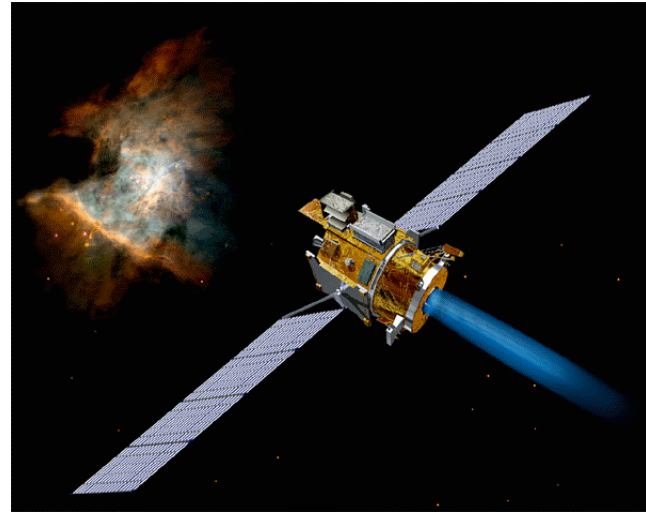
- Building "intelligent" agents
  - Autonomous

  - Rational

  - Adaptive, capable of learning

Spacecraft Control

- When a space probe makes its long flight from Earth to the outer planets, a ground crew is usually required to continually track its progress, and decide how to deal with unexpected eventualities. This is costly and, if decisions are required *quickly*, it is simply not practicable. For these reasons, organizations like NASA are seriously investigating the possibility of making probes more autonomous — giving them richer decision making capabilities and responsibilities.

- *This is not fiction:* NASA's DS1 has done it!

Deep Space 1

- http://nmp.jpl.nasa.gov/ds1/

- "Deep Space 1 launched from Cape Canaveral on October 24, 1998. During a highly successful primary mission, it tested 12 advanced, high-risk technologies in space. In an extremely successful extended mission, it encountered comet Borrelly and returned the best images and other science data ever from a comet. During its fully successful hyperextended mission, it conducted further technology tests. The spacecraft was retired on December 18, 2001." – NASA Web site

Artificial Intelligence - **II**

- Building "societies" of intelligent agents
  - Communication

  - Cooperation

  - Competition

NASA's Mars Robotic Outpost

NASA unwilling to lose a crew on Mars.

Need to have a base where humans can exist for years.

The base needs to be set up before humans arrive.

Because of communication delays, bandwidth limitations, and the construction of the base cannot be remotely controlled from Earth.

Idea is to land array of autonomous agents on Mars, where some control exploration of the surface, some control mining, some control ore refinement, some control manufacturing, and some control the construction.

2001 Mars Odyssey

2003 Mars Exploration Rovers

2005 Mars Reconnaissance Orbiter

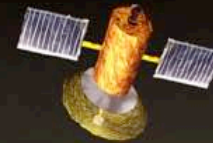**2007** ASI Telecom / CNES Aerocapture / Aerial Scouts / Netlanders / Smart Lander & Rover
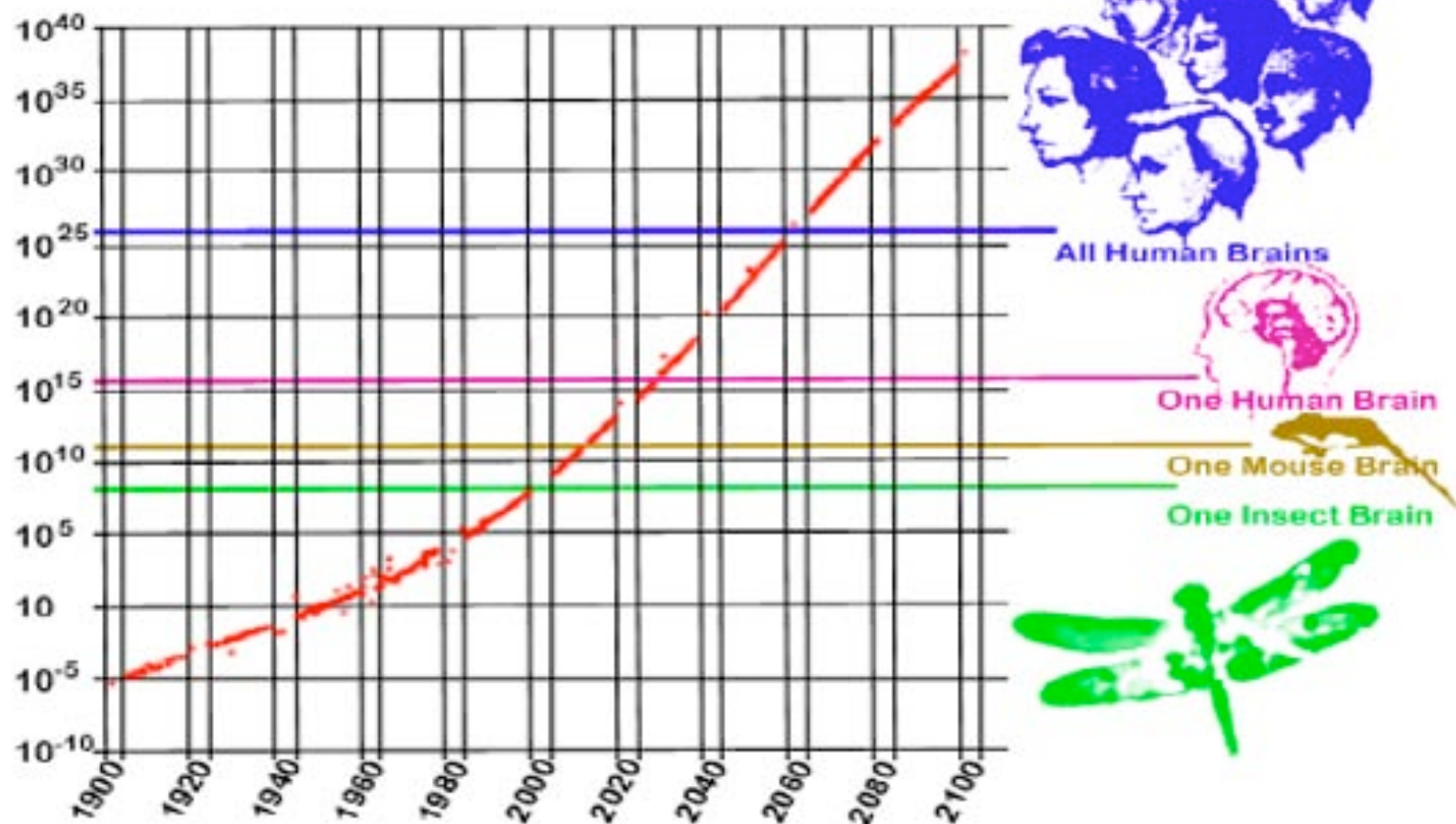
2009 ASI/U.S. SAR

2011 CNES Return / Mars Sample Return (with Smart Lander & Rover)

$1,000 of Computing Buys

All Human Brains

One Human Brain

One Mouse Brain

One Insect Brain

Kurzweil's predictions - 2009

- $1K computer does 1 trillion calculations / second
- Computers embedded in clothes and jewelry.
- Most routine business transactions involve a human and a virtual personality
- Translating telephones are common
- Human musicians routinely jam with cybernetic musicians.

Kurzweil's predictions - 2019

- $1K computer has approximately the computational capacity of a human brain.
- Computers are invisible and embedded everywhere.
- 3D virtual reality displays in glasses are primary way of communicating with people and accessing computer-generated information.
- All-encompassing virtual environments allow virtually any kind of interaction over any distance.
- Virtual personalities (e.g., teachers, lovers) are common.

Kurzweil's predictions - 2029

- $1K computer has approximately the computational capacity of 1000 human brains.
- Neural interfaces to computers have been perfected, and computational augmentation of human brains is available.
- Computers have read all available human literature.
- There is growing discussion about the legal rights of computers.

Kurzweil's predictions - 2099

- The number of software-based humans vastly exceeds the number of those using carbon-based neurons.

- Humans who do not use the enormous computational augmentations available cannot meaningfully participate in dialog with those who do.

- Life expectancy is no longer a viable term in relation to intelligent beings.

A Brave New World of Autonomous Agents

- They will extend our abilities by giving us surrogates who can work in cyberspace on our behalf.

- They will prepare the way for us in outer space by creating an infrastructure there to support us when we arrive.

- *What's the problem?*

#24

DETECTIVE

SPOONER

PRECINCT 28

The Problem!

- Currently, we have no reason to trust them!

- Asimov's 3 Laws of Robotics doesn't exist!

- **We** need to make them trustworthy!

Current Research Going On Here

- Safe integration of planning and execution.

- Safe integration of learning and planning.

- Looking at techniques for proving that an intelligent agent architecture possesses safety properties.

Summary

- AI is one of today's most challenging and revolutionizing fields.
- There are tremendous benefits that might be had from this field.
- There are also frightening perils that might arise from this field.
- There is work going on here to investigate both aspects.