

# The Automation of Science

Ross D. King,  
University of Manchester, [ross.king@manchester.ac.uk](mailto:ross.king@manchester.ac.uk)



# AI Systems have Superhuman Powers to Augment Scientists

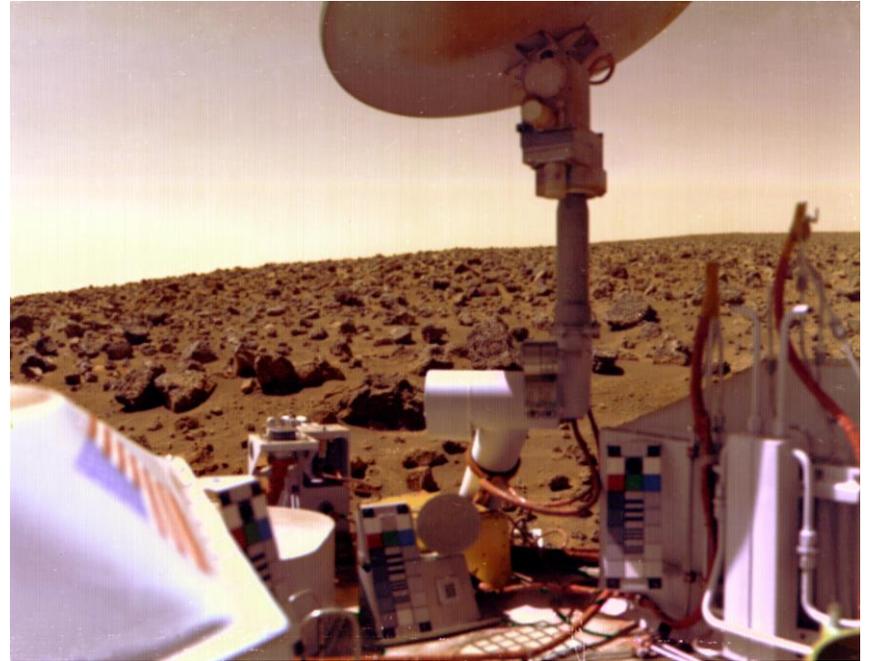
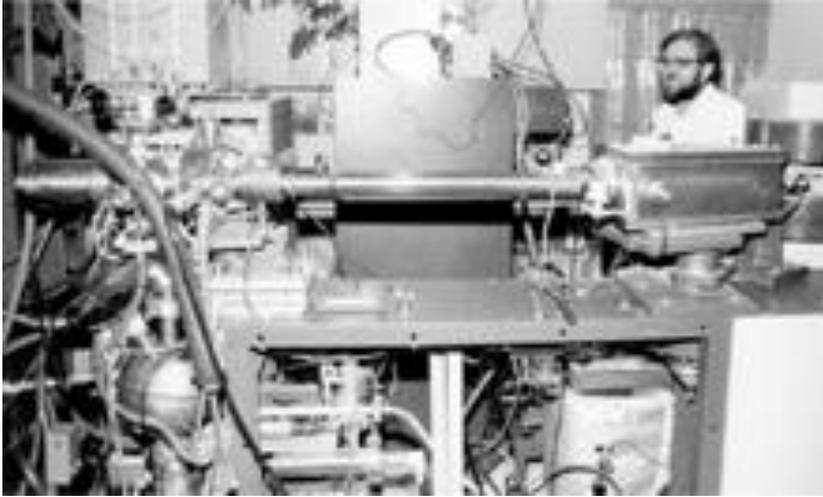
- n They flawlessly remember vast numbers of facts
- n They execute flawless logical reasoning
- n They execute near optimal probabilistic reasoning
- n They learn more rationally than humans
- n They can learn from vast amounts of data
- n They can 'read' millions of scientific papers.
- n Etc.

# Scientific Discovery

Science is a well suited application area for AI.

- n Scientific problems are abstract, but involve the real-world.
- n Scientific problems are restricted in scope – no need to know about “Cabbages and Kings”.
- n Nature is honest – no malicious agents.
- n Nature is a worthy object of our study.
- n The generation of scientific knowledge is a public good.

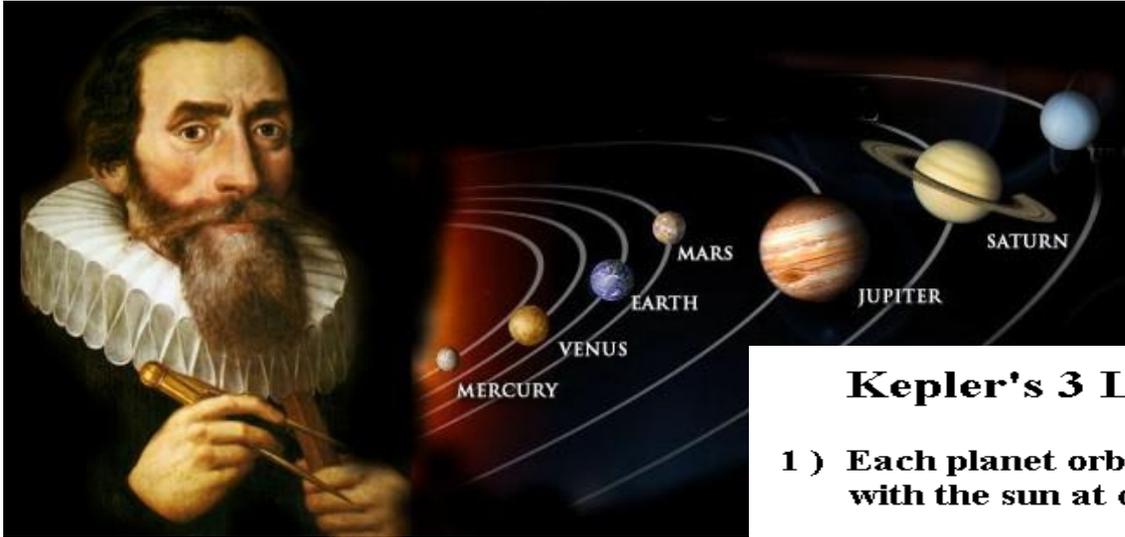
# Meta-Dendral



Analysis of mass-spectrometry data.

Joshua Lederburg, Ed. Feigenbaum, Bruce Buchanan,  
Karl Djerassi, *et al.* 1960-70s.

# Bacon



## **Kepler's 3 Laws of Planetary Motion**

- 1 ) Each planet orbits the sun in an elliptical path with the sun at one focus**
- 2 ) The radius vector ( from sun to planet ) sweeps out equal areas in equal time intervals**
- 3 ) The square of the period is proportional to the cube of the semi-major axis of the orbit**

$$\text{i.e. } T^2 = k a^3 \quad \text{for some constant } k$$

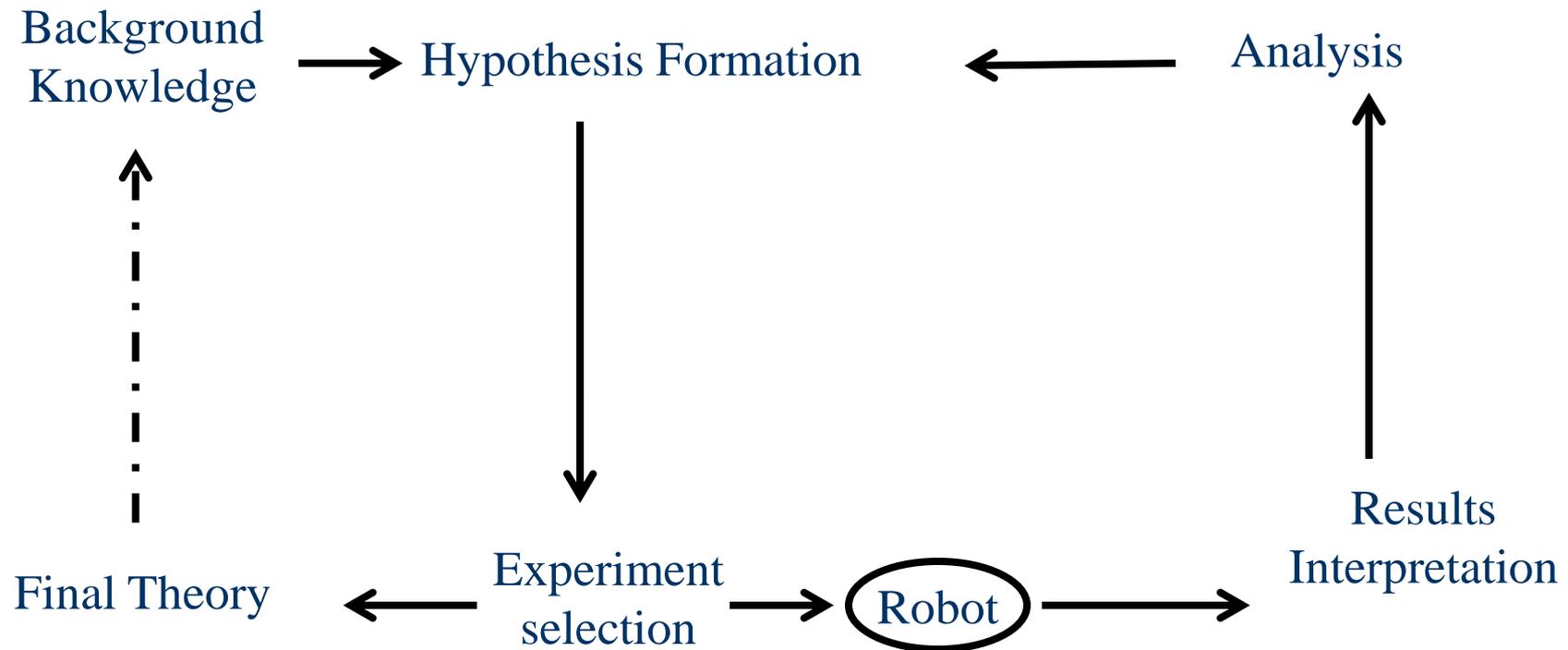
**Figure 11.1**

Rediscovering physics and chemistry: Langley, Bradshaw, Simon (1979).

# Robot Scientists

# The Concept of a Robot Scientist

Computer systems capable of originating their own experiments, physically executing them, interpreting the results, and then repeating the cycle.



# Potential Advantages of Robot Scientists

- n Faster scientific discovery
- n Cheaper experimentation
- n They are more easily reproduced and trained than human scientists.
- n They are harder working than human scientists.
- n Improved knowledge/data sharing, and reproducibility.

# Robot Scientist Timeline

- n 1999-2004 Initial Robot Scientist Project
  - Limited Hardware: Collaboration with Douglas Kell (Aber Biology), Steve Oliver (Manchester), Stephen Muggleton (Imperial)

**King et al. (2004) *Nature*, 427, 247-252**
  
- n 2004-2011 Adam – Yeast Functional Genomics
  - Sophisticated Laboratory Automation: Collaboration with Steve Oliver (Cambridge).

**King et al. (2009) *Science*, 324, 85-89**
  
- n 2008-2015 Eve – Drug Design for Tropical Diseases
  - Sophisticated Laboratory Automation: Collaboration with Steve Oliver (Cambridge)

**Williams et al. (2015) *Royal Society Interface*, DOI 10.1098/rsif.2014.1289**
  
- n 2015-2018 Eve – Human cells - Cancer, Yeast - Aging
  - DARPA, CHIST-ERA

Adam

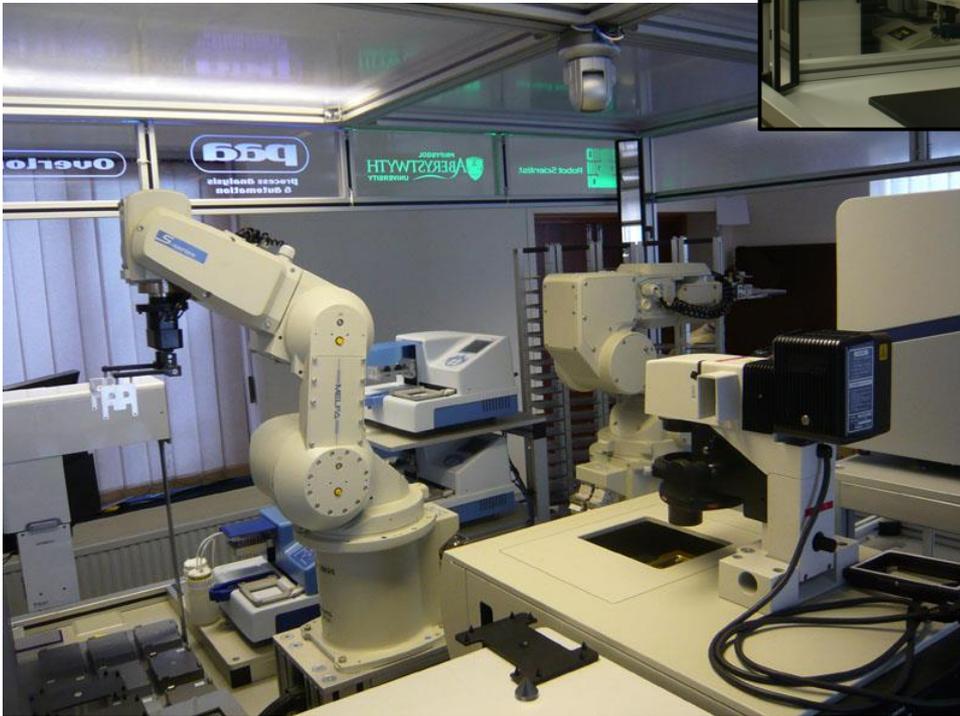
# Adam

- n Adam was the first AI system to both explicitly form hypotheses and experiments, and physically do the experiments.
- n Adam generated and confirmed novel functional-genomics hypotheses concerning the identify of genes encoding enzymes catalysing orphan reactions in the metabolic network of the yeast *S. cerevisiae*.
- n Adam's conclusions have been manually verified using bioinformatic and biochemical evidence.

# Adam Movie

Eve

# Eve



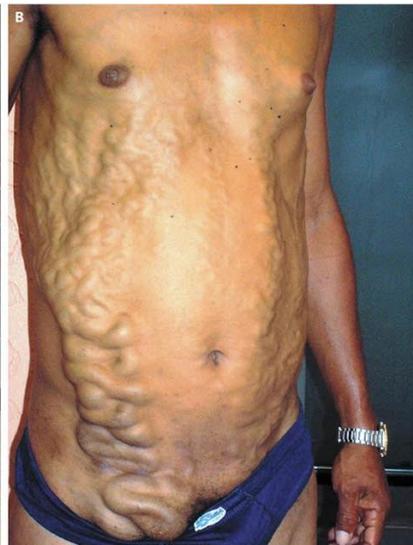
# Parasitic Diseases targeted



Malaria



Shistosomiasis



Leishmania

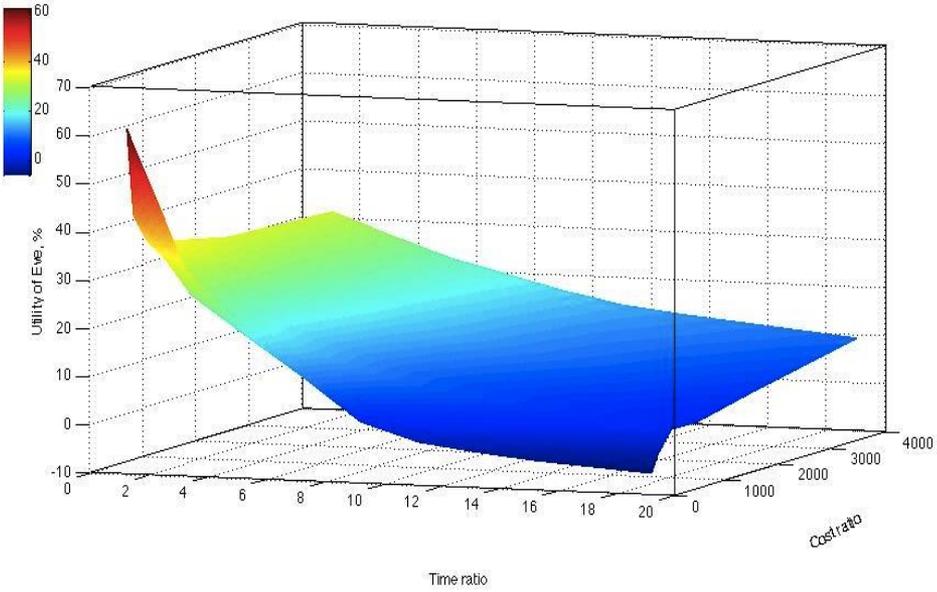
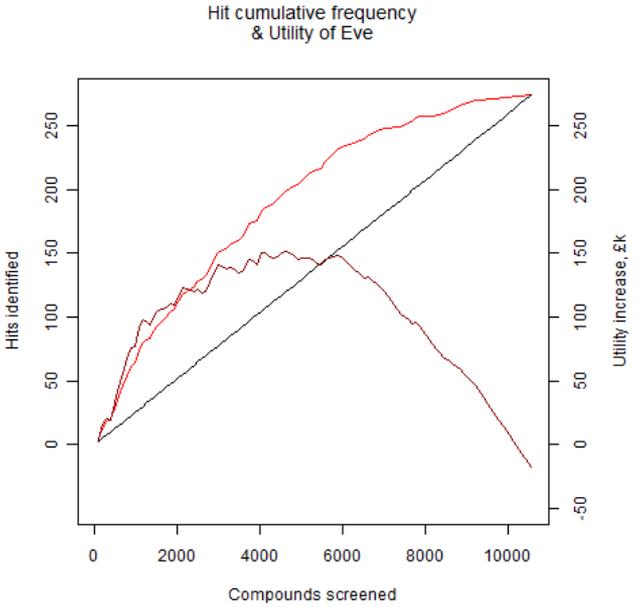
Chagas



# Why Tropical Diseases?

- n Millions of people die of these diseases, and hundreds of millions of people suffer infection.
- n It is clear how to cure these diseases – kill the parasites.
- n They are “neglected”, so avoid competition from the Pharmaceutical industry.

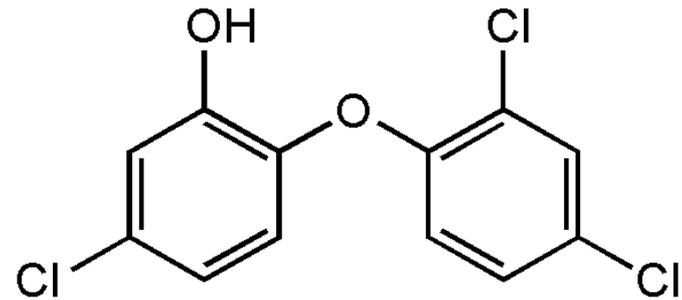
# The Economics of Intelligent Screening



$$\Delta \text{Utility of Eve} = \sum_1^{Nm} (Tm + Cm) + \sum_1^{Nx} (Tc + Cc - Uh) + \sum_1^{Ne} (Tm - Tc + Cm - Cc)$$

- Nm - Number of compounds not assayed by Eve
- Tm - Cost of the time to screen a compound using the mass screening assay
- Cm - Cost of the loss of a compound in the mass screening assay
- Nx - Number of hits missed by Eve
- Tc - Cost of the time to screen a compound using a cherry-picking (confirmation or intelligent) assay
- Cc - Cost of the loss of a compound in a cherry-picking assay
- Uh - Utility of a hit
- Ne - Number of compounds assayed by Eve

# Triclosan Repositioned for Malaria



- n Simple compound
- n Known to be safe – used in toothpaste.
- n Targets both DHFR and FAS-II – well established targets.
- n Demonstrated activity using multiple wet experimental techniques.
- n Works against wild-type and drug-resistant *Plasmodium falciparum*, and *Plasmodium vivax*.

# Future Prospects

# The Future?

- n In chess there is a continuum of ability from novices up to Grandmasters.
- n I argue that this is also true in science, from the simple research of Adam and Eve, through what most human scientists can achieve, up to the ability of a Newton or Einstein.
- n If you accept this, then just as in chess, it is likely that advances in computer hardware and software will drive the development of ever smarter Robot Scientists.
- n In favour of this argument are the impressive on going development of AI and laboratory robotics.

# Vision

- n The collaboration between Human and Robot Scientists will produce better science than either can alone – human/computer teams still play better chess than either alone.
- n Scientific knowledge will be primarily expressed in logic with associated probabilities and published using the Semantic Web.
- n The improved productivity of science leads to societal benefits: better food security, better medicines, etc.
- n The Physics Nobel Frank Wilczek is on record as saying that in 100 years' time the best physicist will be a machine.

# Conclusions

- n Science is a well suited application area for AI.
- n Automation is becoming increasingly important in scientific research e.g. DNA sequencing, drug design.
- n The Robot Scientist concept is the logical next step in scientific automation.
- n The Robot Scientist Adam was the first machine to have discovered novel scientific knowledge.
- n The collaboration between Human and Robot Scientists will produce better science than either can alone.

# Acknowledgments

- n The Robot Scientist team: Manchester, Aberystwyth, Cambridge, Brunel, Leuven, Thailand. (BBSRC)
- n Chicago Big Mechanism consortium. (DARPA)
- n AdaLab consortium. (CHIST-ERA)