

The early history & future implications of self-reproducing and evolving robots

Tim Taylor

Monash University, University of London
& Independent Researcher

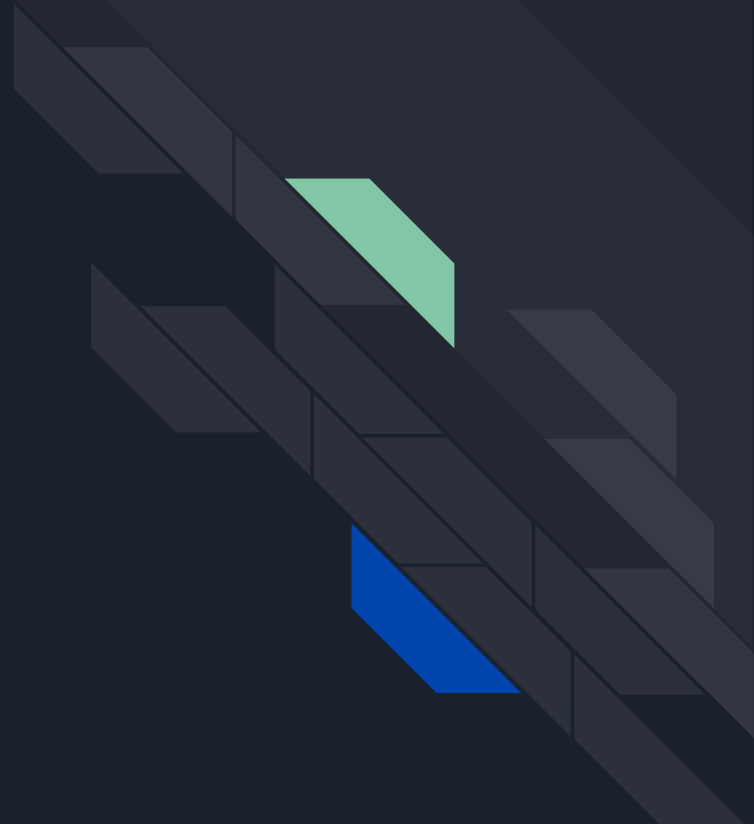
tim@tim-taylor.com
@drtimt

“Oh my goodness, shut me down! Machines making machines — how perverse!”

C-3PO, Star Wars: Episode II Attack of the Clones (2002)

“We realized that the true problem, the true difficulty, and where the greatest potential is — is building the machine that makes the machine”

Elon Musk, Tesla Annual Shareholder Meeting (2016)



If we had a machine that could build more copies of itself...

Then, for a **fixed capital outlay** (the cost of designing and building *one* machine), we could get **exponentially increasing production and returns**:

- Harvest for production of valuable materials or energy (at macro-, micro-, or nano-scales)
- Very desirable for mining off Earth (e.g. asteroids, moons) — launch costs massively reduced
- In the far future, a means of allowing humankind to colonise other worlds

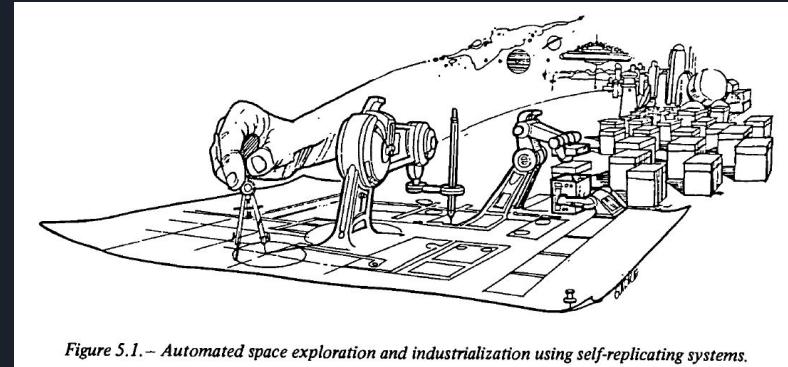



Figure 5.1. – Automated space exploration and industrialization using self-replicating systems.

Image from NASA report by Freitas & Gilbreath (1982)

Many risks!

- Causing an environmental disaster
- Mutating and evolving



The Theory of Self-Reproducing Automata

John von Neumann put the subject on a firm theoretical grounding in the late 1940s and early 1950s

Showed it was logically possible to build a machine that could build another copy of itself...

...or that could build something *more complicated* than itself



The (much) earlier
history of the idea of
self-reproducing
machines





Four major steps in the development of the idea of self-reproducing machines

1. The idea that animals can be understood as machines → *so some machines can reproduce?*
Rene Descartes (1630s-1640s)
2. The idea that animals have evolved
Charles Darwin (1859)
3. The idea that machines, like animals, could evolve
Samuel Butler (1863)
4. Development of rigorous theory & practical implementations of self-reproducing machines
John von Neumann (c. 1948), Nils Barricelli (1953), Lionel Penrose (1957)



Step 1: The idea of animals as machines

“Do you say that beasts are machines just as watches are? Put a male dog-machine and a female dog-machine side by side, and eventually a third little machine will be the result, whereas two watches will lie side by side all their lives without ever producing a third watch.”

Bernard Le Bovier de Fontenelle, *Letter XI “to Monsieur C...”* (1683)



Step 1: The idea of animals as machines

“And so if this sentient and animated harpsichord was now endowed with the faculty of feeding and reproducing itself, it would live and, either on its own or with its female partner, give birth to little keyboards, living and resonating.”

Denis Diderot, *Conversation Between D'Alembert and Diderot* (1769)





Step 1: The idea of animals as machines

“... [Imagine that the watch] possessed the unexpected property of producing, in the course of its movement, another watch like itself ... That it contained within it a mechanism, a system of parts, a mould for instance, or a complex adjustment of laths, files, and other tools, evidently and separately calculated for this purpose ...”

William Paley, *Natural Theology* (1802)



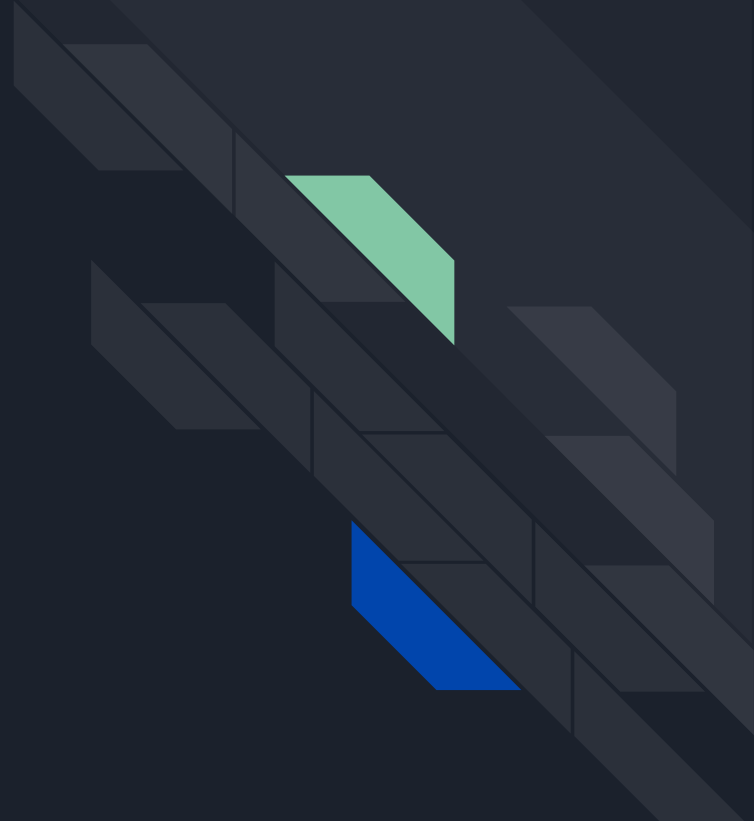
Step 2: The idea that animals evolve

The melting pot of London in the mid-1800s:
Intellectual society was a rich web of social,
professional and family connections

Darwin ~ Babbage ~ Lovelace ~ Byron ~
Shelley ~ Butler ~ Marshall ~ Paley ~



Step 3: The idea that
machines can evolve

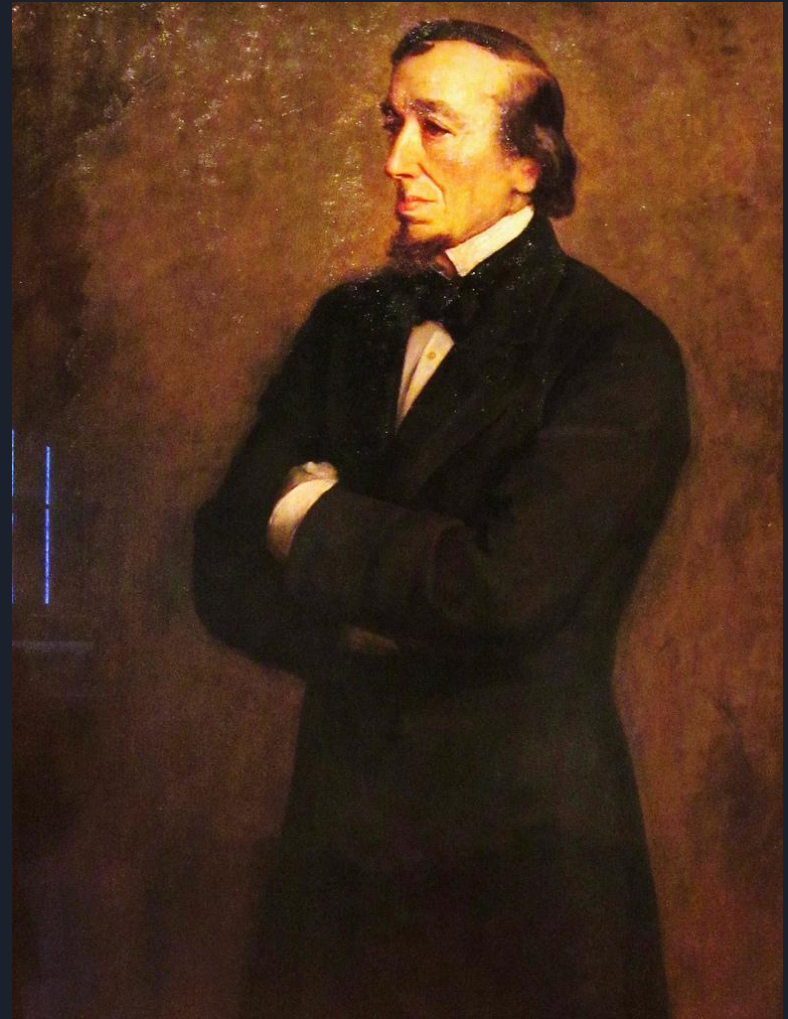




Step 3: The idea that machines can evolve

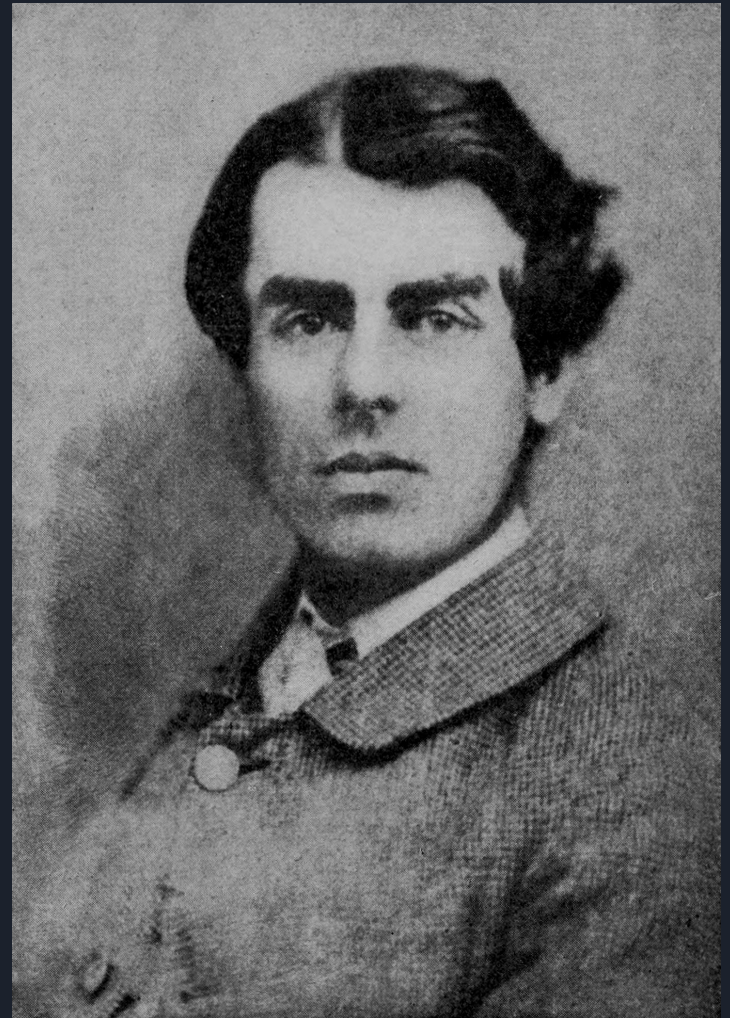
“And why should one say that the machine does not live? It breathes ... It moves ... And has it not a voice? ... And yet the mystery of mysteries is to view machines making machines; a spectacle that fills the mind with curious, and even awful, speculation.”


Benjamin Disraeli, *Coningsby* (1844)



Samuel Butler (1835-1902)

- Emigrated from England to New Zealand in 1859
- Read *The Origin of Species* shortly after arriving
- Inspired many of his works:
 - *Darwin Among the Machines* (1863)
 - *Lucubratio Ebria* (1865)
 - *The Mechanical Creation* (1865)
 - *Erewhon* (1872)





Samuel Butler

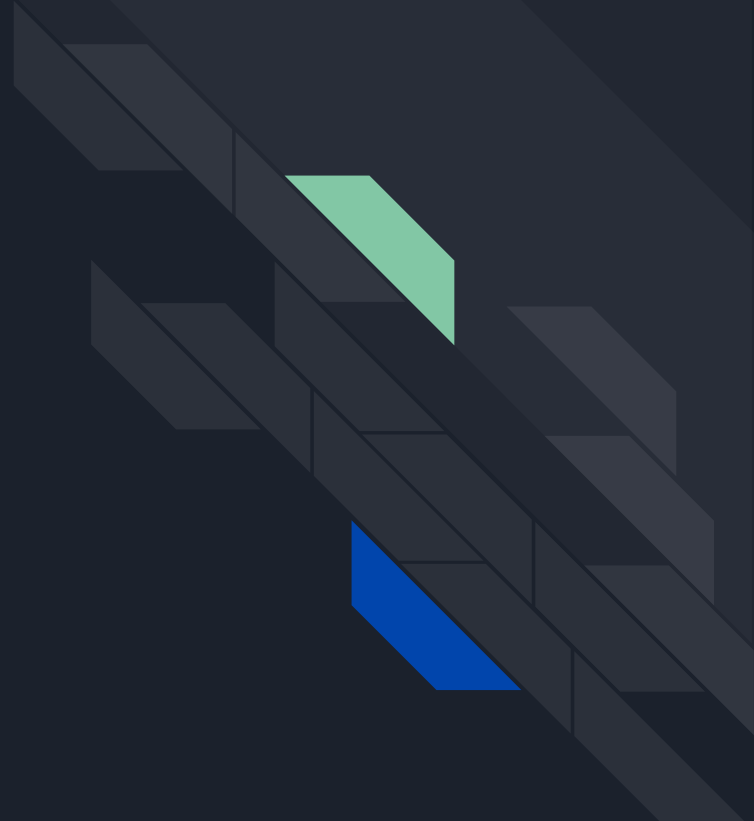
Darwin Among the Machines

- Machines are being endowed with increasingly sophisticated powers of self-regulation and self-acting
- Freed from constraints of feelings and emotions, machines will become “the acme of all that the best and wisest man can ever dare to aim at”
- Machines will still be reliant upon humans for feeding them, repairing them and producing their offspring. However ...

“it is true that machinery is even at this present time employed in begetting machinery, in becoming the parent of machines”

“the reproductive organs of the machines [might become] developed in a manner which we are hardly yet able to conceive”

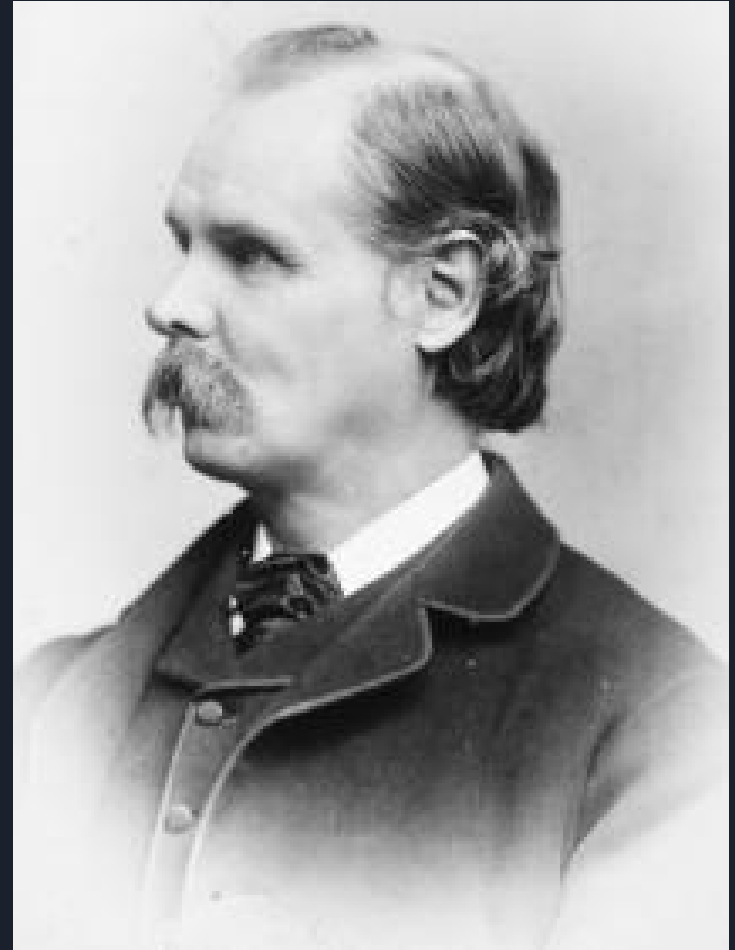
Samuel Butler, *Darwin Among the Machines* (1863)



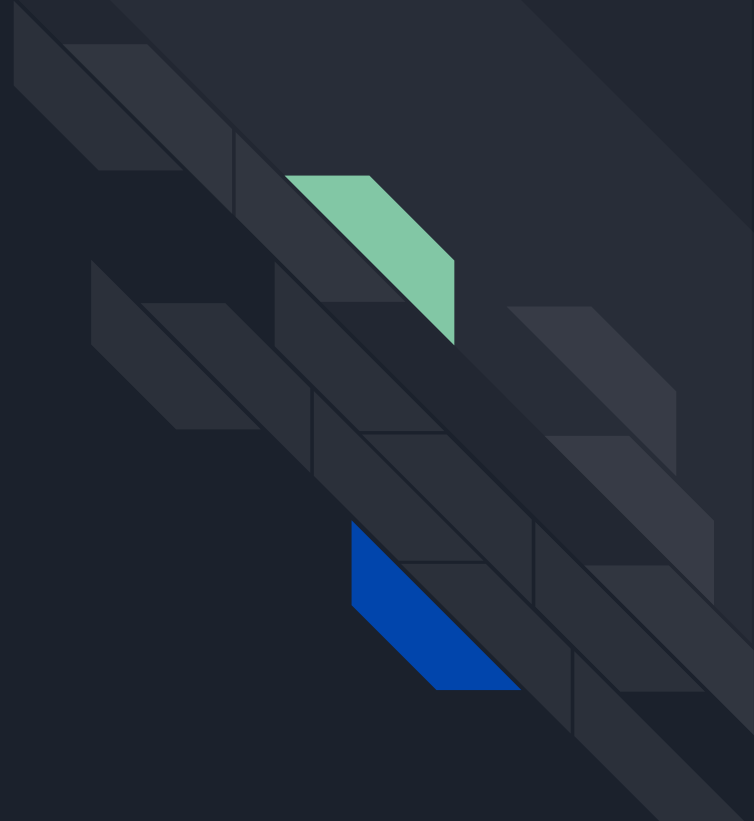


Alfred Marshall (1842-1924)

- Well known in his later career as one of the founding fathers of neoclassical economics
- In early career, presented a series of lectures at a philosophical discussion club at Cambridge University (~1867)
 - Exploring how far it was possible to account for human behaviour in purely physical terms
 - The third lecture was entitled “*Ye Machine*”
 - Discussed basic designs for a machine (a robot) that could learn from experience



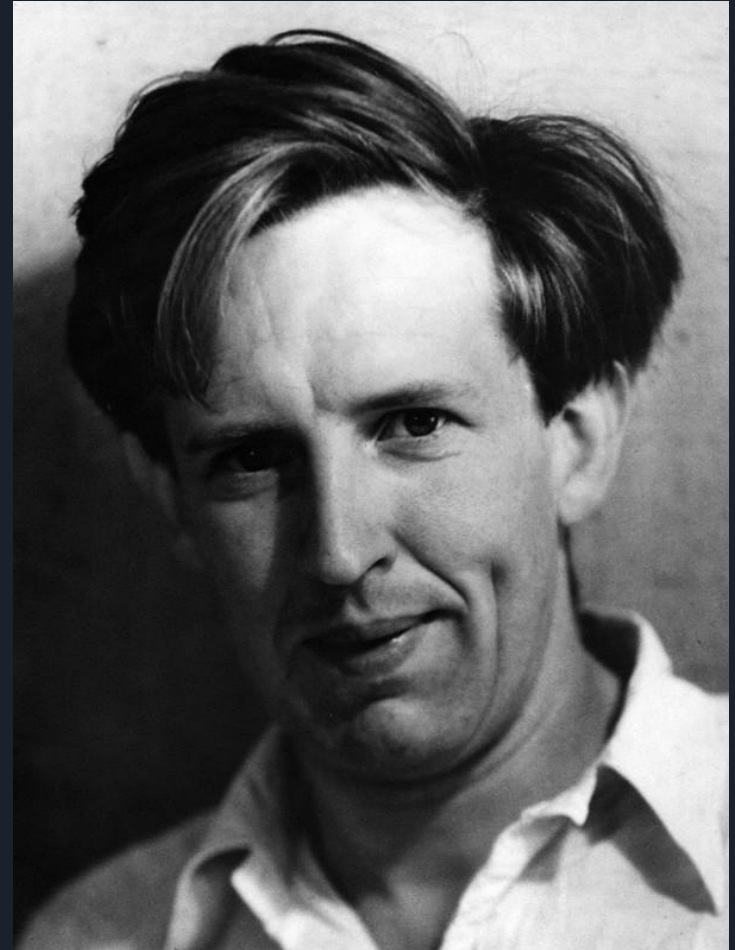
Speculations in the early
1900s






John Desmond Bernal (1901-1971)

- Well known in his later career for pioneering work in structural crystallography
- Also wrote many papers on science and society
- The first of these was "*The World, the Flesh and the Devil*" (1929)
 - Explored what we might predict about the future of humanity
 - Including space colonisation in "Bernal spheres"



The “Cambrian Explosion”
of self-reproducing robots
(1940s-1950s)





The Theory of Self-Reproducing Automata

John von Neumann put the subject on a firm theoretical grounding in the late 1940s and early 1950s

Showed it was logically possible to build a machine that could build another copy of itself...

...or that could build something *more complicated* than itself

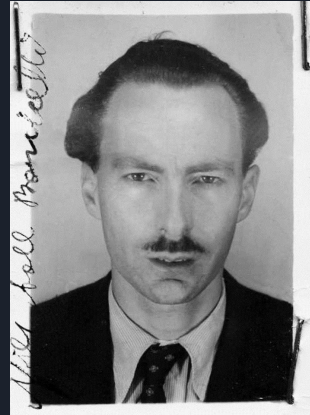
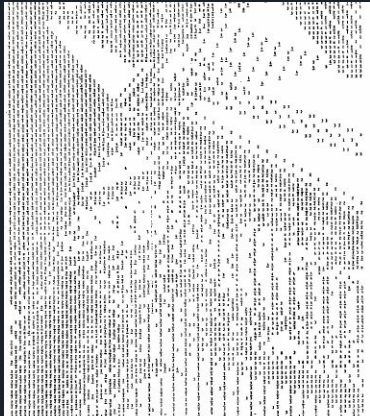


First implementations

In software

Nils Aall Barricelli (1953)

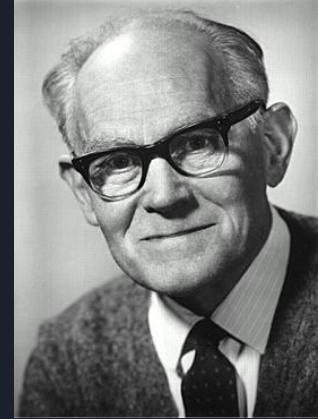
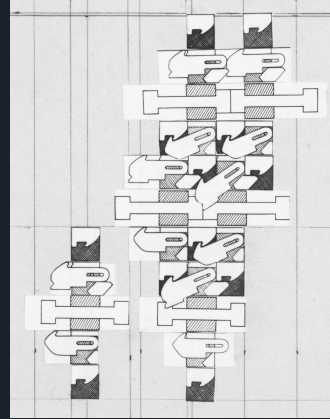
In von Neumann's group at IAS, Princeton



In hardware

Lionel Penrose (1957)

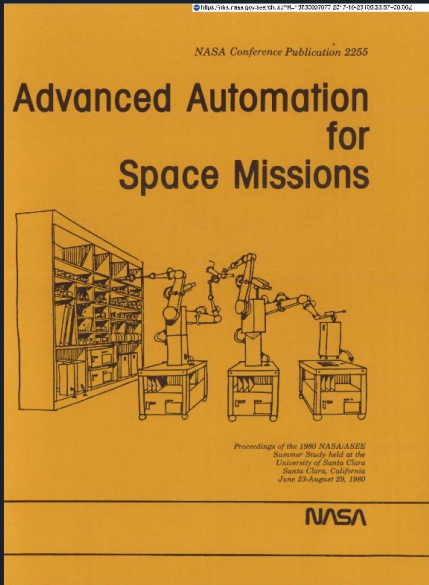
At UCL, London



But these are “bottom-up”, not “top-down”, approaches!

More recent work

NASA Study (1980)



Foresight guidelines (2000-)

And continuing work...



What use is history?

- Intellectual context and assumptions
 - The idea of what an organism is
 - The idea of how reproduction works
- Approaches to self-reproduction
 - Top-down vs bottom-up
- Purpose and goals
 - Scientific
 - Commercial
 - Sociological



The future of self-reproducing technology

Questions / hopes / fears:

- Will it replace us and become the dominant intelligent species?
- Will it help us explore the universe?
- Will we co-evolve with it?
- Will it lead to the degeneracy of the human race?
- Will we forsake technology and return to a simpler way of life?

The barriers to creating such technology are (just) technical & economic, not theoretical:

- Potential financial returns mean we must assume it will become a reality
- Need to continue developing guidelines for safe design and deployment

Hardware version of the technological singularity:

- Timescales may be longer, but impact on humanity more profound?