# The mathematics of future computing s Budo



GRESHAM COLLEGE





Why is a Professor of Geometry giving a talk about computers?

Computers impact our lives in many ways:

- InternetSmart phonesCash machines
- •Washing machines ....

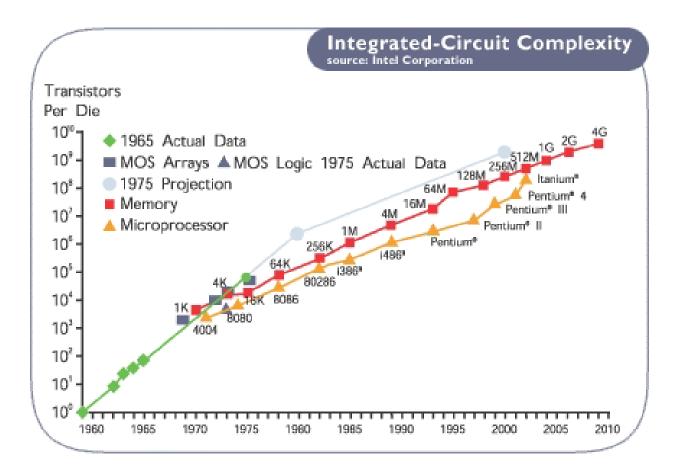


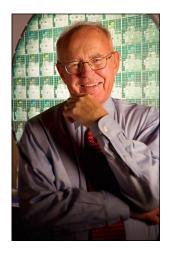
Computers were originally invented by mathematicians to solve mathematical and logical problems

Mathematics lies at the heart of the algorithms that give computers their power

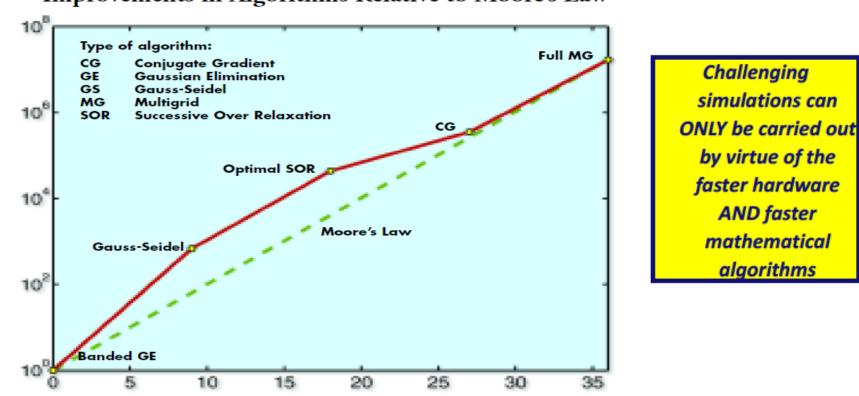
# Moore's law:

#### Computer hardware power doubles every 18 months





#### An even faster speed up is given by improvements in algorithms



Improvements in Algorithms Relative to Moore's Law

Number of Years

Rate of Increase in Processing Performance

Huge changes are coming in modern computation:

- Much faster supercomputers
- Huge increase in data
- Machine learning
- Quantum computing

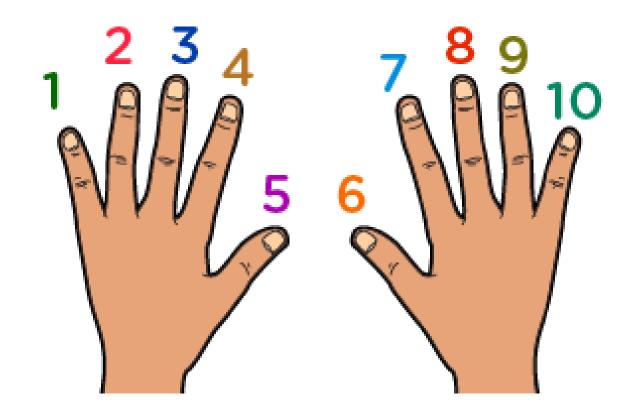


Only way to make sense of this and to exploit the full power is to use mathematics:

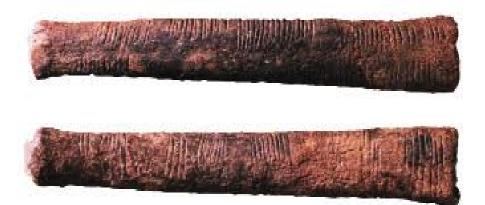
Taking full advantage of computing tools requires moremathematical sophistication not lessBoeing

A short history of computing

# The earliest digital computer



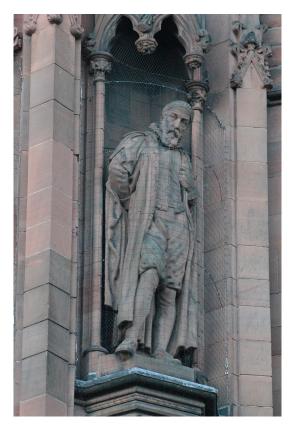




#### Abacus



# Napier and Logarithms 1614





His invention of logarithms was quickly taken up at Gresham College and prominent English mathematician Henry Briggs visited Napier in 1615. Among the matters they discussed were a re-scaling of Napier's logarithms, in which the presence of the mathematical constant now known as e (more accurately, e times a large power of 10 rounded to an integer) was a practical difficulty. Neither Napier nor Briggs actually discovered the constant e; that discovery was made decades later by Jacob Bernoulli.

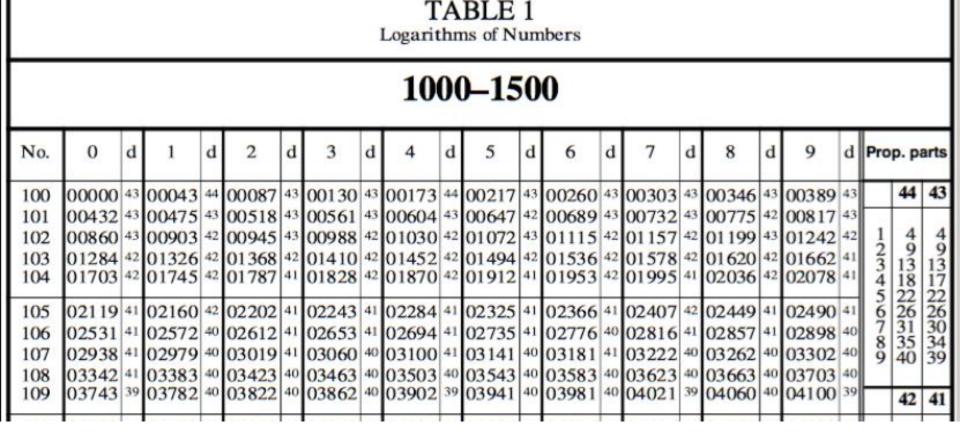
Wikipedia

$$x = 10^a$$
 a is the logarithm of x to base 10

Law of exponents

$$x \ y = 10^a \ 10^b = 10^{a+b}$$
$$x/y = 10^a/10^b = 10^{a-b}$$

Multiplication and division become addition and subtraction of logarithms



#### 13.45\*23.56 = ??

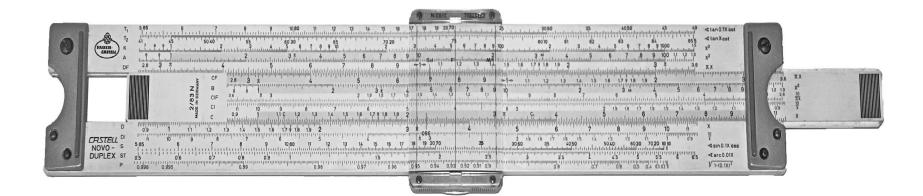
 $\log(13.45) = 1.1287$ 

log(23.56) = 1.3722

1.1287 + 1.3722 = 2.5009 antilog(2.5009) = 316.9

Exact product 316.882

#### Process automated by using a slide rule





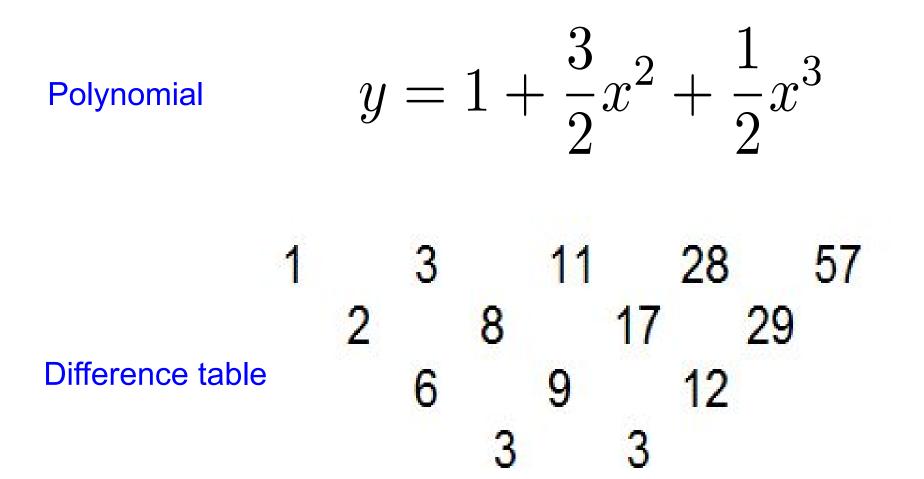
# Mechanical computers





#### Babbage's difference engine

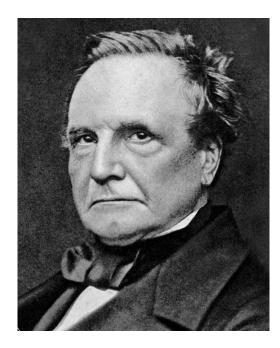
# **Divided differences**





# Answer is 101

- Method works for any polynomial
- And any function approximated by a polynomial
- Is very mechanical
- And can be mechanised



Charles Babbage 1823

**Difference Engine** 

£17 000

Developed many of the ideas used in modern computers

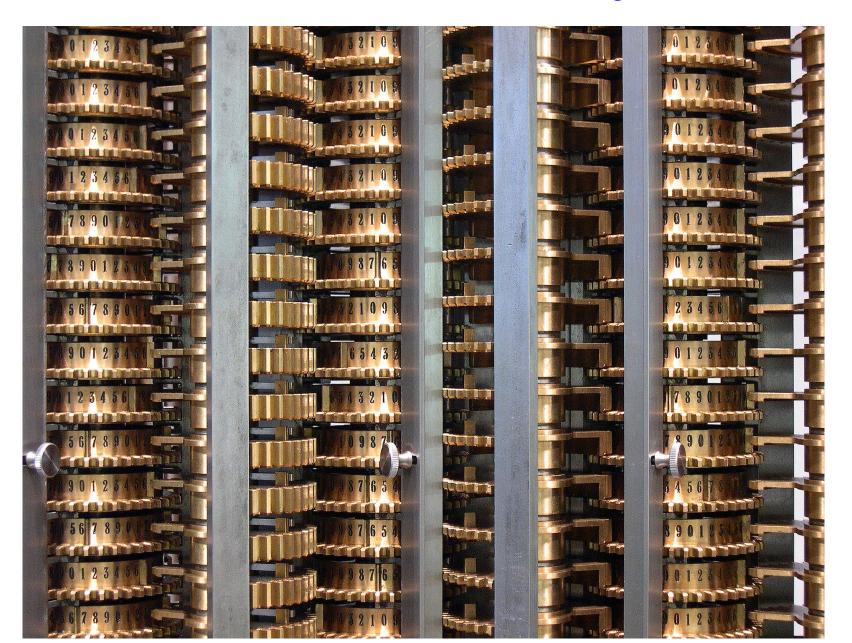
**Firstly** it had to be able to represent numbers to a certain precision. The more decimal digits we use the more accurate the calculation, but the harder it is to store them, and the slower the resulting calculation.

**Secondly** it must be able to store these numbers. To calculate an nth degree polynomial it must store (at least) n+1 numbers

**Thirdly** it must be able to add these numbers quickly and accurately.



#### Numbers stored as N columns of cog wheels



Prototype: 3 columns of 6 cog wheels

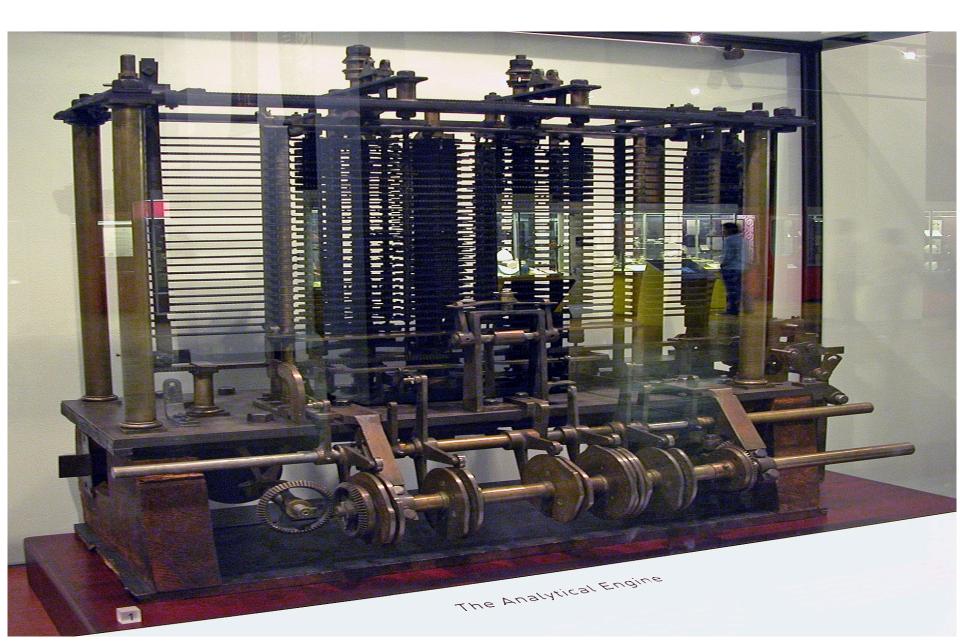
Science museum: 8 columns of 31 wheels

Lady Byron (mother of Ada Lovelace) reported on seeing the working prototype in 1833 that

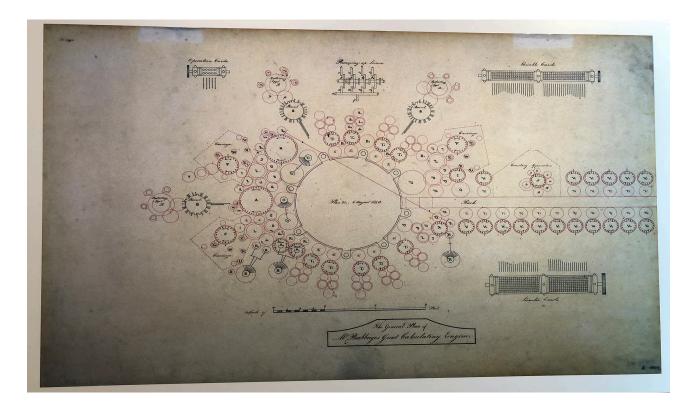
"We both went to see the thinking machine (for so it seems) last Monday. It raised several Nos. to the 2nd and 3rd powers, and extracted the root of a Quadratic equation."

The full project was not accomplished successfully, and it was abandoned in 1842

# Analytical engine



- Progammable
- Arithmetical unit
- Controlled flow
- Memory
- Punch cards



#### Ada Lovelace

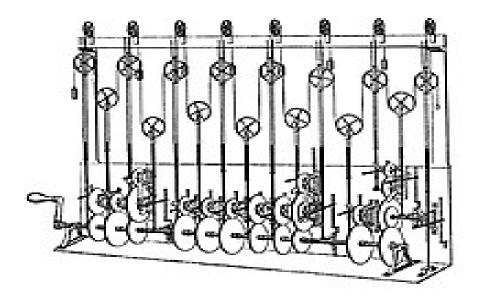


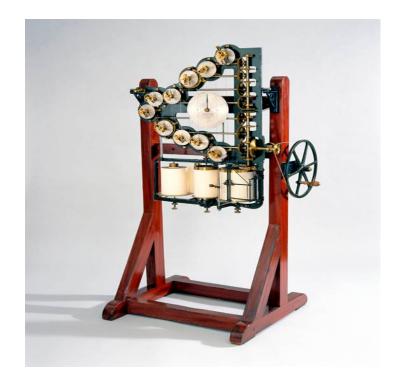
	Nature of Operation.	Variables acted upon.	Variables receiving results,	Indication of change in the value on any Variable,	Statement of Revolts.	Data					Working Variables.							Result Variables.				
Number of Operation						5000	5000m m	2000	5.0000	0000	1×0004	¥"0===	5,000 C	£0000	*****	*0 *	\$ <sup>9</sup> 0==0	Ψα Ο 0 0	F Bula Ort	T Control of Control o	F Balando S	20000 E
1 2 3 4 5 6	- + + +	$\frac{W_4 - W_1}{W_4 + W_4}$	ν <sub>τ</sub> , ν <sub>μ</sub> , ν <sub>η</sub> ν <sub>1</sub> , ν <sub>1</sub> , ν <sub>1</sub> , ν <sub>1</sub> ,	$\begin{cases} v_1 = v_2 \\ v_1 = v_1 \\ v_2 = v_1 \\ v_3 = v_1 \\ v_4 = v_4 \\ v_5 = v_4 \\ v_6 = v_6 \\ v_6 = v_6 \end{cases}$	$= \frac{2 \times \frac{1}{2 \times $	1 1	a       a		2x 7x - 1  0 	2x 2x+1 0 	2a 	1 1 1		1 1 1	1 1 1	$ \frac{\frac{2n-1}{2n+1}}{\frac{1}{2}} \frac{1}{7} \cdot \frac{2n-1}{2n+1} $ 0	Station .	$-\frac{1}{2}\cdot\frac{2n-1}{2n+1}-\lambda_{0}$		「日本」		A LESS STRATE
7 8 9 10 11 12	- + + + -	$W_0 \neq W_0$ $W_0 \neq W_0$ $W_0 + W_0$	w <sub>o</sub>	$\begin{cases} v_{11}^{0} - v_{21}^{0} \\ v_{11}^{0} - v_{11}^{0} \\ w_{11}^{0} - w_{11}^{0} \\ w_{11}^{0} - w_{11}^{0} \\ v_{11}^{0} - w_{11}^{0} \\ w_{11}^{0} - w_{11}^{0} \\ w_{11}^{0} - w_{11}^{0} \\ v_{11}^{0} - w_{11}^{0} \\ v_{11}^{0} - w_{11}^{0} \\ v_{11}^{0} - w_{11}^{0} \\ w_{11}^{0} - w_{11}^{0} \\ w_{11}^{0} - w_{11}^{0} \end{cases}$	$= s - 1 (= 3)$ $= 2 + 0 - 2$ $= \frac{2s}{2} - \lambda_1$ $= B_1 \cdot \frac{2s}{2} = B_1 \lambda_1$ $= -\frac{1}{2} \cdot \frac{2s - 1}{2s + 1} + B_1 \cdot \frac{2s}{2} \dots$ $= s - 2 (= 2)$		1 1 1 1	•		1 1 1 1 1 1				1 1 1 1	n - 1	$\frac{2n}{2} = \lambda_1$ $\frac{2n}{2} = \lambda_2$ $\frac{2n}{2} = \lambda_2$	81. <del>21</del> - 51 - 51	$\left\{ = \frac{1}{2}, \frac{2n-1}{2n+1} + B_n, \frac{2n}{2} \right\}$	Ba	「日本の日本」		「日本」
13 ( 14 15 16 17 18 19	[+  ×  +	νı, + <sup>1</sup> ν,	94, 147,	$ \begin{array}{c} x_{1}^{2} + y_{2}^{2} \\ x_{1}^{2} + y_{1}^{2} $	$= 2 + 1 = 3$ $= 2 + 1 = 3$ $= \frac{2 + 1}{3}$ $= \frac{2 + 1}{3}$ $= \frac{2 + 2}{3}$ $= 2 + 3$ $= 3 + 1 = 4$ $= \frac{2 + 2}{4}$					1 1 1	2 n = 1  2 n = 1  2 n = 2  2 n = 2	3	2s - 1 3 0	 7.8 = 3 4		$\frac{\frac{2n}{7},\frac{2n-1}{3}}{\left\{\frac{2n}{2},\frac{2n-1}{3},\frac{2n-2}{3}\right\}}$			State of the second	あった。		A CONTRACTOR
2 7 7 1 1	× + -		т <sub>н</sub> т <sub>ы</sub> т <sub>ы</sub>	$\begin{cases} w_{1} - w_{2} \\ w_{21} - w_{21} \\ w_{22} - w_{22} \\ w_{23} - w_{23} \\ w_{23} \\ w_{23} - w_{23} \\ w$	$= \lambda_{0} + B_{1} \lambda_{1} + B_{2} \lambda_{2} \dots \dots$		11111		ere folk					0  	    	- A J •  y-three.	- Bj Aj 0	{ A <sub>2</sub> + B <sub>1</sub> A <sub>1</sub> + F <sub>2</sub> A <sub>2</sub> }	I I I I I I I I I I I I I I I I I I I			By
25	+	w <sub>a + m</sub>	·	$\left  \begin{cases} {}^{4}\Gamma_{15} = {}^{4}\Gamma_{15} \\ {}^{6}\Gamma_{25} = {}^{6}\Gamma_{25} \\$	= * + 1 = 4 + 1 = 5 by a Yariable-card. by a Yariable cost.	1		*+1	-	-	0	0						ñ				

Diagram for the computation by the Engine of the Numbers of Bernoulli. See Note G. (page 722 et sep.)

Kelvin's tidal computer

# An analogue computer





The invention of the modern electronic computer

- Developments in mathematical logic in the 1930s
- Huge stimulus of WW2: code breaking, A bomb
- First programmable computers built by mathematicians in the late 1940s
- Transistor and integrated circuit led to the first commercial computers in the 1960s
- 1980s first practical home computers



# Alan Turing 1912-1954

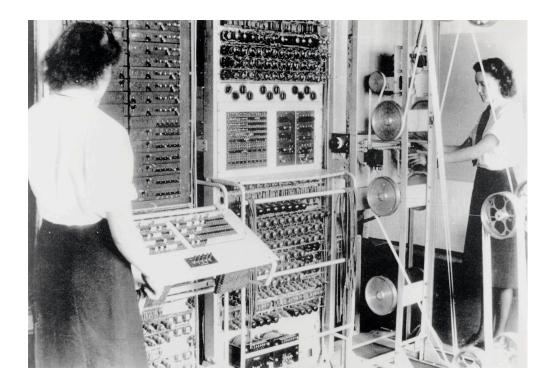
# Tommy Flowers 1905-1998

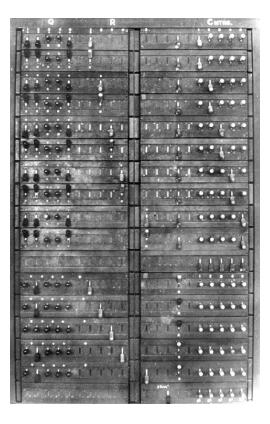




#### **Colossus computer**

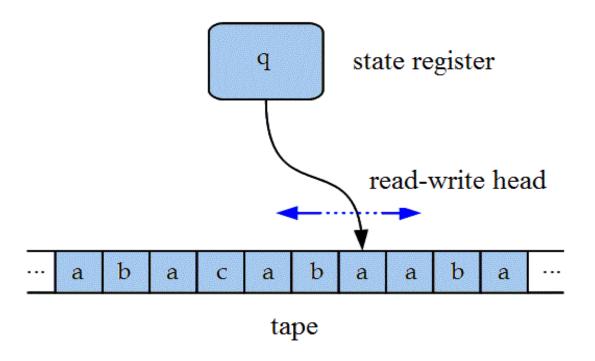
# Designed to break the Lorenz Cipher





#### Semi-programmable using a switch board

# Turing Machine 1936

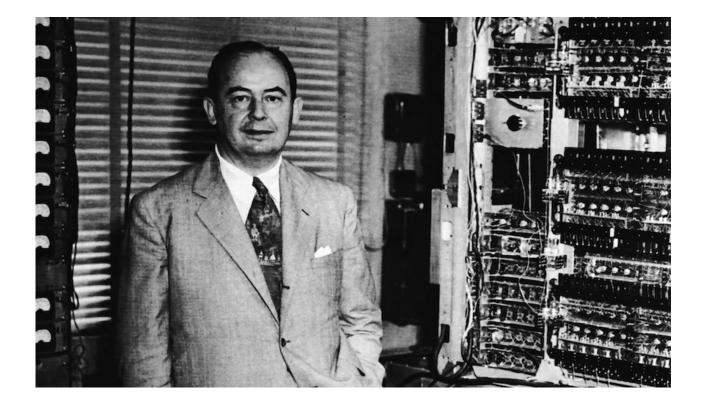


# Theoretical computer

Capable of doing arbitrary computations

Modern languages have to be Turing Complete

# John von Neumann 1903-1957



Samuel 7 algander

First Drift of a Report on the ENVAG

John von Neumann

Contract No. W-670-ORD-4926

Between the

United States Army Ordnance Department

and the

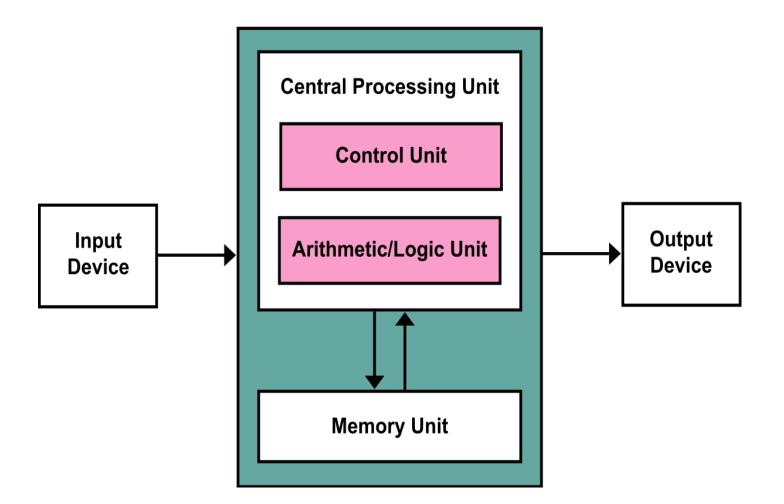
University of Pennsylvania

Moore School of Electrical Engineering University of Pennsylvania

June 30, 1945

National Bureau of Standards Division 12 Data Processing Systems

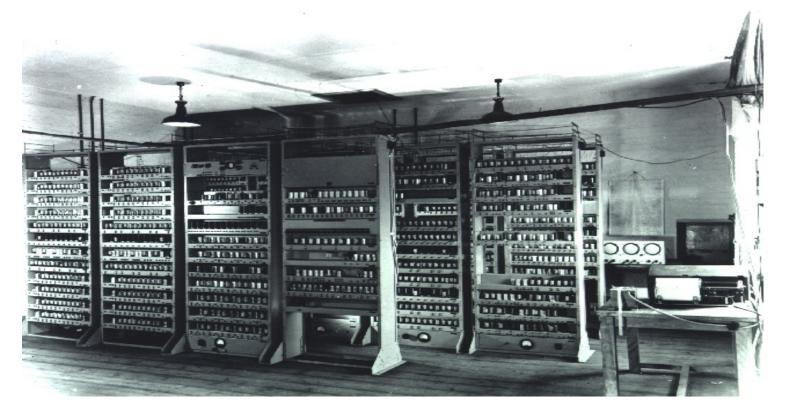
#### von Neumann architecture



Basis of all modern computers

#### Cambridge Maths EDSAC 1949 inspired by von Neumann

18 operation codes, 1024 words of memory, 650 instructions per second



The 'brain' [computer] may one day come down to our level [of the common people] and help with our income-tax and book-keeping calculations. But this is speculation and there is no sign of it so far

# 1970s: SUSIE and PDP10





# Software

Thanks in no small part to the work of (generations of) mathematical logicians, high level languages have been developed which allow computers to be programmed in a language like English.

Early examples: COBOL, FORTRAN, PASCAL and ALGOL.

Later came BASIC, the language of the BBC micro, and the object oriented language C++.

Other more sophisticated languages such as LISP and HASKELL are used for machine learning and AI applications.

Scientific computing and the future of large scale computing

My own field of research is *scientific computing:* Computing the solution to problems posed in scientific terms.

Applications in all areas of science: physics, chemistry, biology, neuro-science, cosmology and climate science

Plays a central role in engineering including aircraft and car design, in the drug industry, in medicine, and in genomics.

Also of major importance in film animation and gaming

Some of the 'biggest' calculations currently being undertaken are scientific

Numerical analysis:

Branch of mathematics behind the design of algorithms to solve mathematical problems accurately and efficiently

First challenge is how to represent real numbers such as

1/3 = 0.333333333333 ...

1.4142135623730950488 ....

(Babbage had the same problem)

Early calculators very inaccurate

Modern computer: 20 digits



Ordinary differential equations

 $\frac{d\mathbf{u}}{dt} = \mathbf{f}(t, \mathbf{u})$ 

#### Partial differential equations

 $\frac{\partial^2 u}{\partial t^2} = c(x) \frac{\partial^2 u}{\partial x^2}$ 

Matrix eigenvalue problems

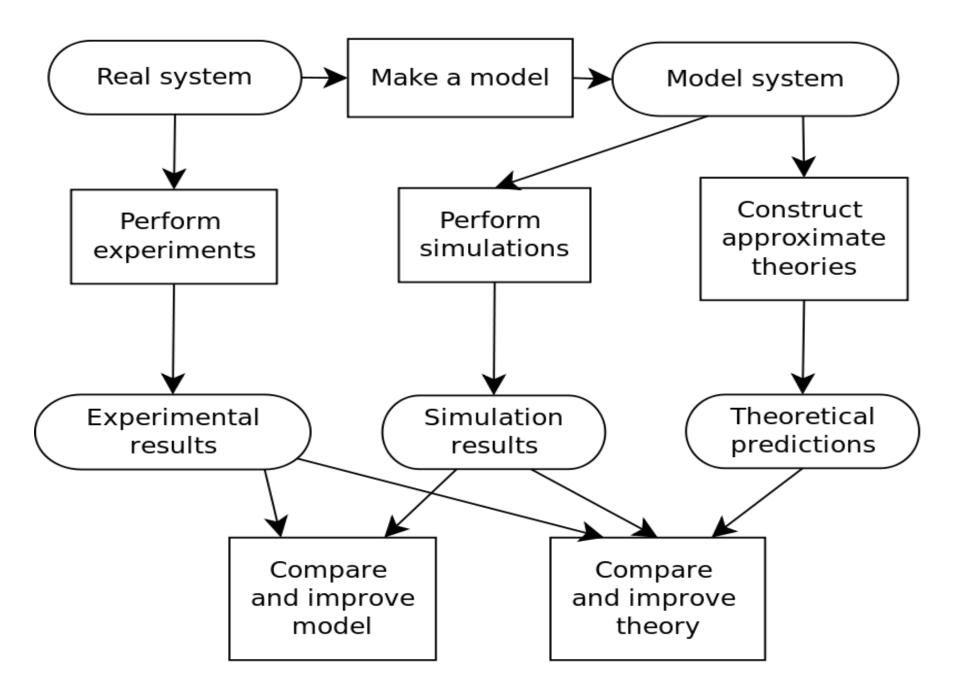
 $A \mathbf{x} = \lambda \mathbf{x}.$ 

# Modelling the world using scientific computing

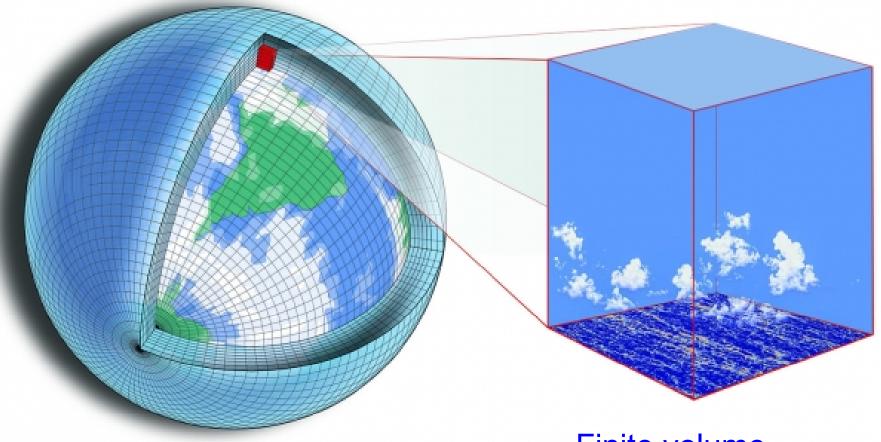
## Kings Cross Fire 1987





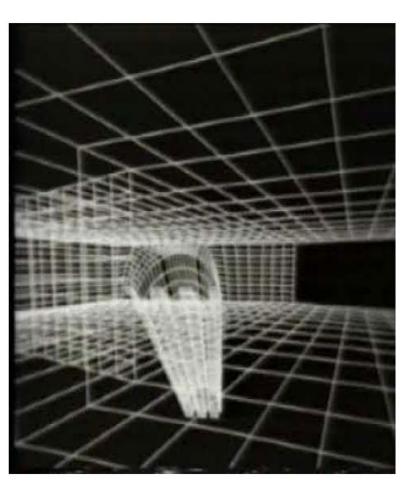


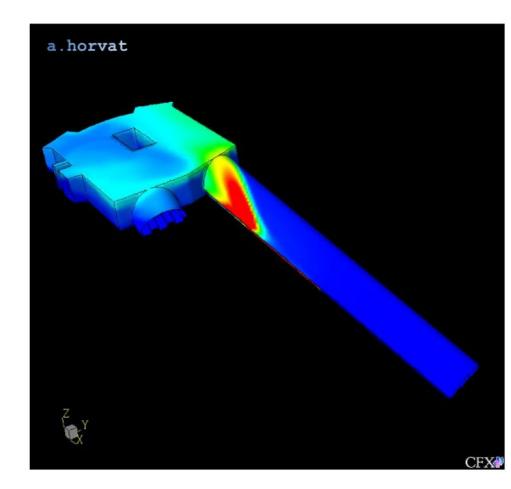
Simulation is performed by discretising the differential equations and solving the resulting algebraic systems

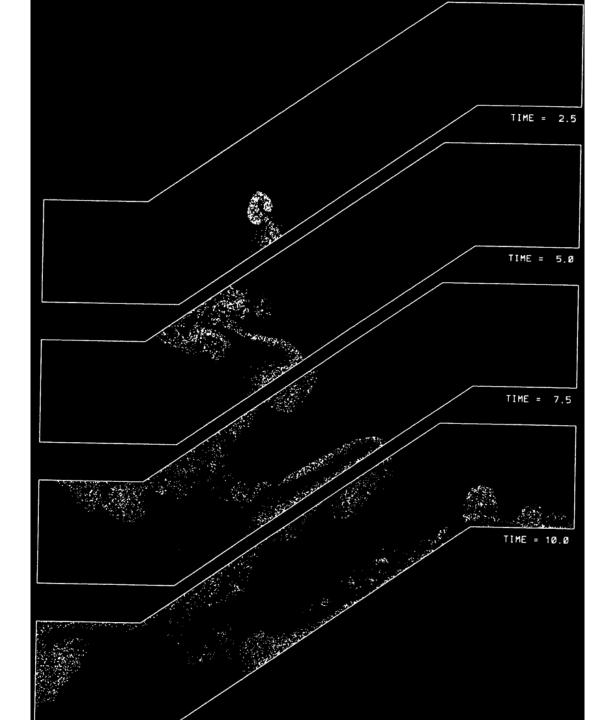


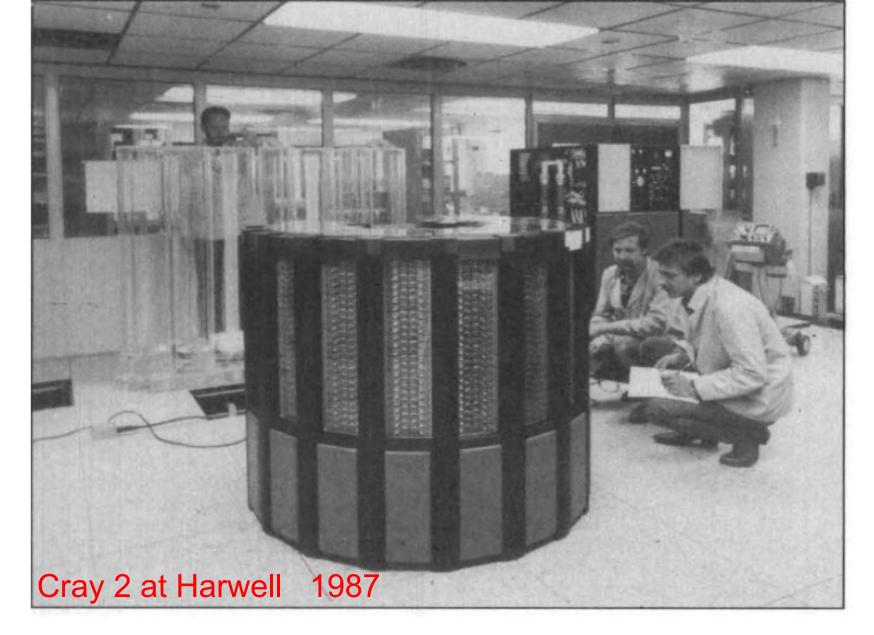
#### Finite volume

# Doing this for the Kings Cross Fire led to the discovery of the trench effect









This was one of the first ever super computers and had the (at the time) unheard of speed of 1.7 GFlops and a 2G byte RAM

#### Problems of scale (the maths of future computing)

In such a calculation there is a trade off between the number of finite volumes, and the accuracy speed of the calculation.

If there are  $\,N^3$  volumes (N in each spatial dimension), the error decreases at a rate

$$1/N^{2}$$

The computational time increases, often at a rate approaching

$$N^6$$

If N is large eg. 10 000 we pay a big time penalty for any increase in accuracy.

# Hence the need for supercomputers!

#### Increase in hardware

Moore said in 1965, that the number of transistors that would fit in a given circuit space was doubling every year.

Later this slowed down to every two years.

Electronic components based on Silicon can only shrink so much before they run into the limits of physics.

The smallest transistors today consist of about a hundred atoms. Silicon transistors are down to 10 nano meters, and there is now talk of 5-nanometer electronics,

But that may be the limit, in part due to the effects of quantum physics

New developments in optics and graphene

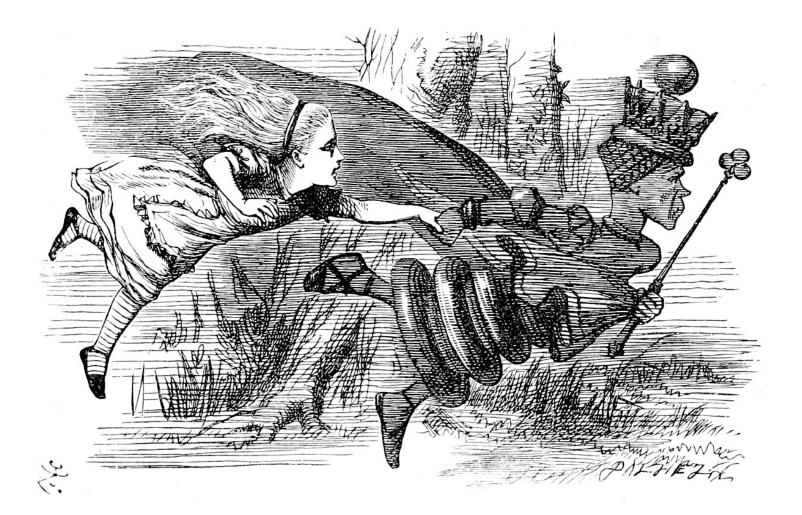
The big problem in making things smaller is heat dissipation A real problem with modern computing is keeping them cool

At present this is a major limitation to their usage, as is the sheer amount of energy that they need to use. This is in the order of Mega Watts.

One good way to reduce the amount of energy is to do calculations more efficiently. This can involve developing better mathematical algorithms, using reduced precision when you can get away with it, or replacing the solution of a physical problem by using a machine learning alternative

All of these are the subject of intensive research

#### Red Queens Race of Software vs. Hardware



A very significant extra are of advance will be in the use of parallel processing methods in which a lot of operations are carried out at the same time

Modern computers are multi-core machines, with a lot of processors all working together simultaneously.

The latest Met Office Cray XC40 computer, has 460,000 separate cores



Eg. Calculating  $n! = n(n-1)(n-2)(n-3) \dots 3.2.1$ 

Serial method:  $n! = n \cdot (n-1)!$ 

Takes n operations

Parallel method (divide and conquer)

n! = (n(n-1))((n-2)(n-3))((n-4)(n-5)).... (2.1)

Takes ONE operation

Repeat  $m = \log_2(n)$  times

Eg. n = 542 288 m = 19

# Petascale computers and beyond



# A Peta is $10^{15}$

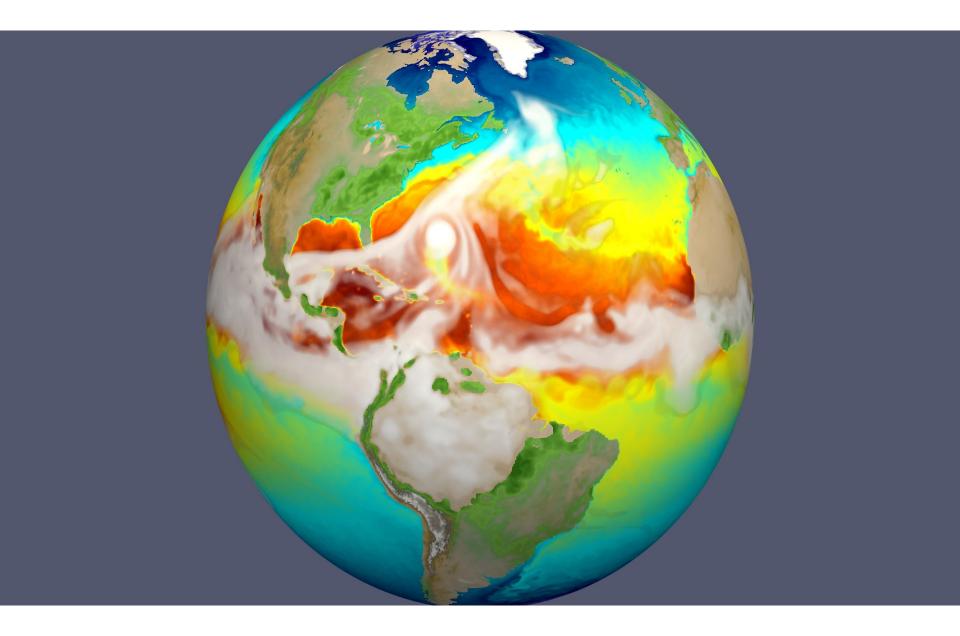
The three Met Office Cray computers are Peta scale computers

Capable of over 14 Peta arithmetic operations per second

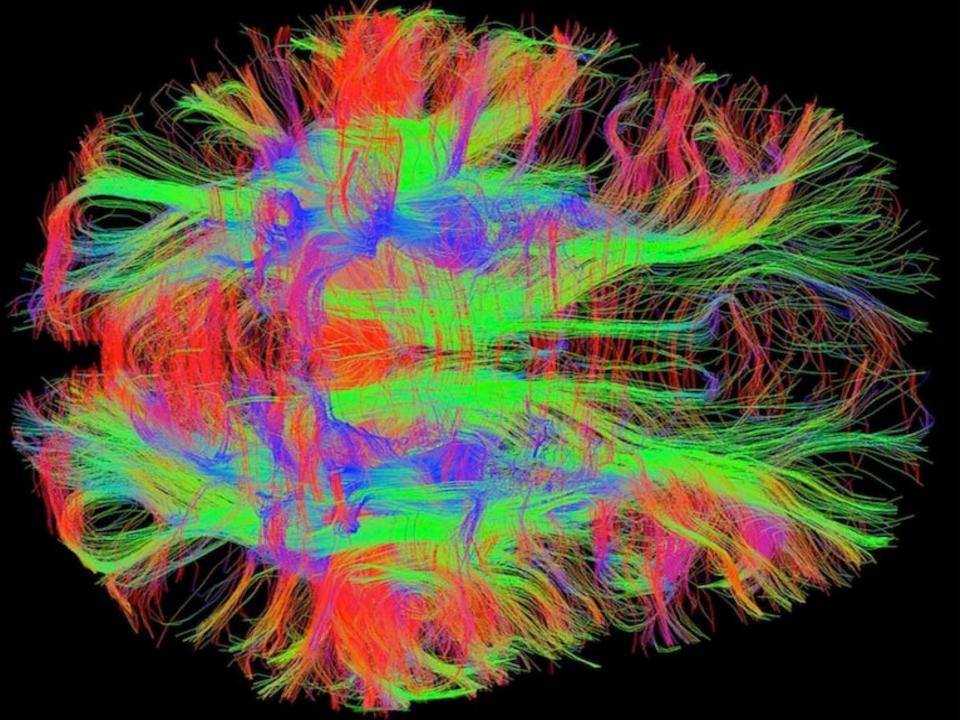
It has 2 Peta bytes of RAM memory

The first Peta scale computer went on line in 2008 and there are around twenty Peta scale computers currently in use.

Used to do computations in many fields: weather and climate simulation, nuclear simulations, cosmology, genomics, quantum chemistry, medical imaging, remote sensing, space flight, lowerlevel brain simulation, molecular dynamics, drug design, aerodynamics, fusion reactors.









Exa scale computing: 1000 times faster



Estimated to be the order of processing power of the brain

Achieved in 2018 at Oak Ridge National Laboratory which performed a 1.8 Exaflop calculation on the Summit OLCF-4 Supercomputer while analysing genomic information

China will develop an Exa scale computer during the 13th Five-Year-Plan period (2016–2020) project, which is planned to be named Tianhe-3

Zetta scale computing will the next 1000 fold increase in computing power. It is estimated that this may happen in 2030. Watch this space!

# **Machine Learning and Algorithms**

Main impact of the computer and algorithms on our lives is more domestic

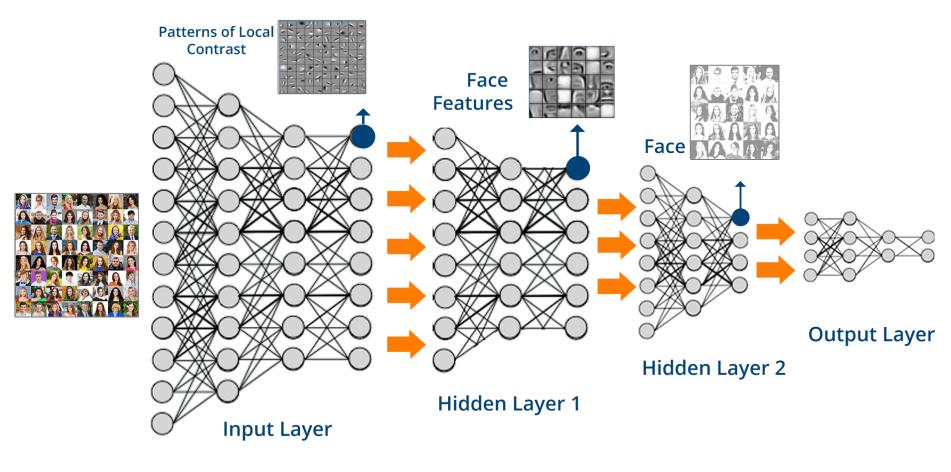
Already very significant:

- Internet ... Queuing theory
- Google ... Matrix eigenvalue problem
- Mobile phone ... Error correcting codes
- Credit cards ... Encryption

# But a Tsunami is about to hit us!



## Machine (deep) learning



A machine learning algorithm is trained to do a task by looking at past examples, and used to do similar tasks in the future

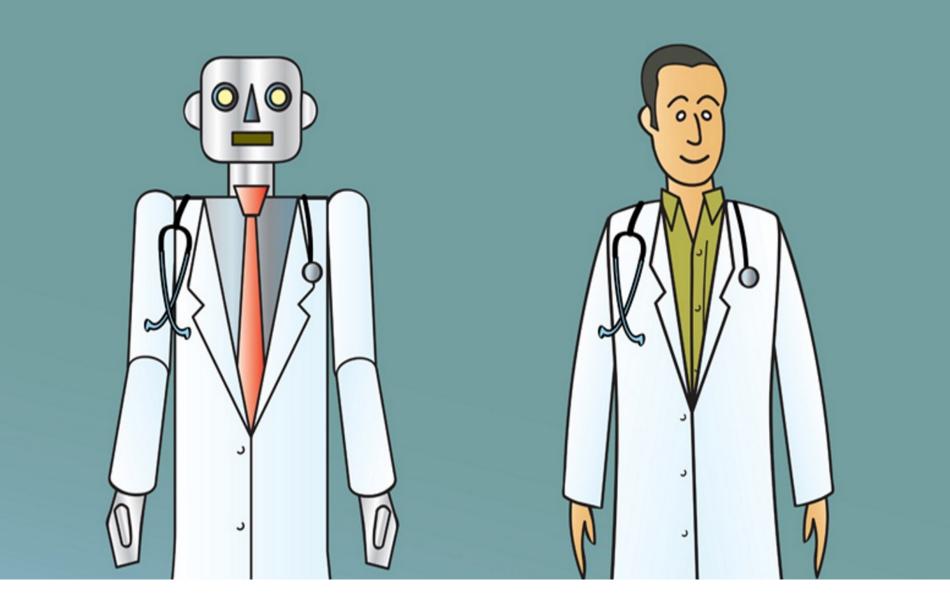
Examples of machine learning include:

Chess and Go, speech recognition, image recognition, weather forecasting, and recommendations for books on Amazon

Also include:

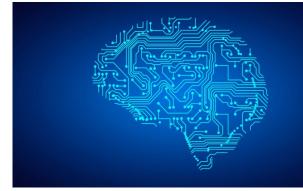
Medical diagnosis, dating sites, personnel recruitment, driverless cars, and even making legal judgments

Potentially we are looking at a future without doctors, car drivers or judges. This will radically change our society.



All based on mathematical algorithms

But these are still poorly understood



Need to understand the mathematics better to have

Explainable AI Towards a Fundamental Theory of Deep Learning

Trustworthy AI Determining the Limits of Deep Learning Technology

**New avenues** Bringing Deep Learning to New Horizons

## **Quantum Computing**

The field of quantum computing was initiated by the work of Benioff and Manin 1980, Feynman 1982, and Deutsch 1985.

Predicted that quantum computers could come to dwarf the processing power of today's conventional computers, by harnessing the effects of quantum theory

Quantum computers could eventually allow work to be done at a speed almost inconceivable today.

Quantum computers operate on qubits

These allow many operations to be carried out simultaneously

Main problem is to keep these in a state of coherence whilst the algorithm is running

Many national governments are funding quantum computing research to develop quantum computers for civilian, business, and national security purposes

A small 20-qubit quantum computer exists and is available for experiments via the *IBM-Q quantum experience* 



Exciting times certainly lie before us with the increased power of computing

But in case we ever get complacent I leave you with one final quote from the great John von Neumann who said in the late 1940s

It would appear that we have reached the limits of what it is possible to achieve with computer technology, although one should be careful with such statements, as they tend to sound pretty silly in 5 years.