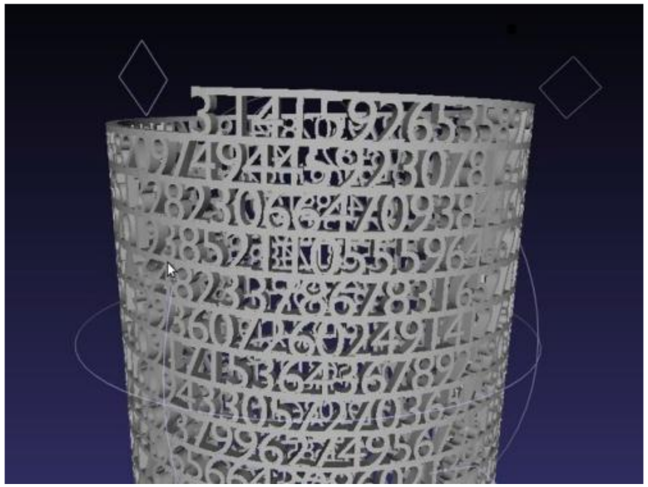
Numbers and Calculators

ARCHI 2019



Summit
2,397,824 Cores
2,801,664 GB mem
200,795 TFlop/s

13,000,000 W
200,000,000 $

TOP 500 The List.

The GREEN 500

Summit
2,397,824 Cores
2,801,664 GB mem
200,795 TFlop/s

13,000,000 W
200,000,000 $

MareNostrum P9 CTE
19,440 Cores
27,648 GB mem
1,018 TFlop/s

85,000 W
34,000,000 €

Summit
2,397,824 Cores
2,801,664 GB mem
200,795 TFlop/s

13,000,000 W
200,000,000 $

MareNostrum P9 CTE
19,440 Cores
27,648 GB mem
1,018 TFlop/s

85,000 W
34,000,000 €

Version N°3
38 PFlop/s

20 Watts
30,000 €

TOP 500 The List.

Summit
2,397,824 Cores
2,801,664 GB mem
200,795 TFlop/s

13,000,000 W
200,000,000 $

The GREEN 500

MareNostrum P9 CTE
19,440 Cores
27,648 GB mem
1,018 TFlop/s

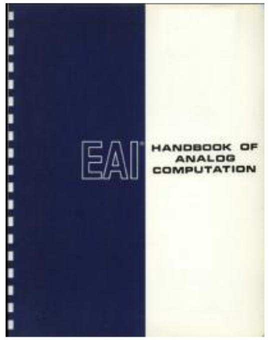
85,000 W
34,000,000 €

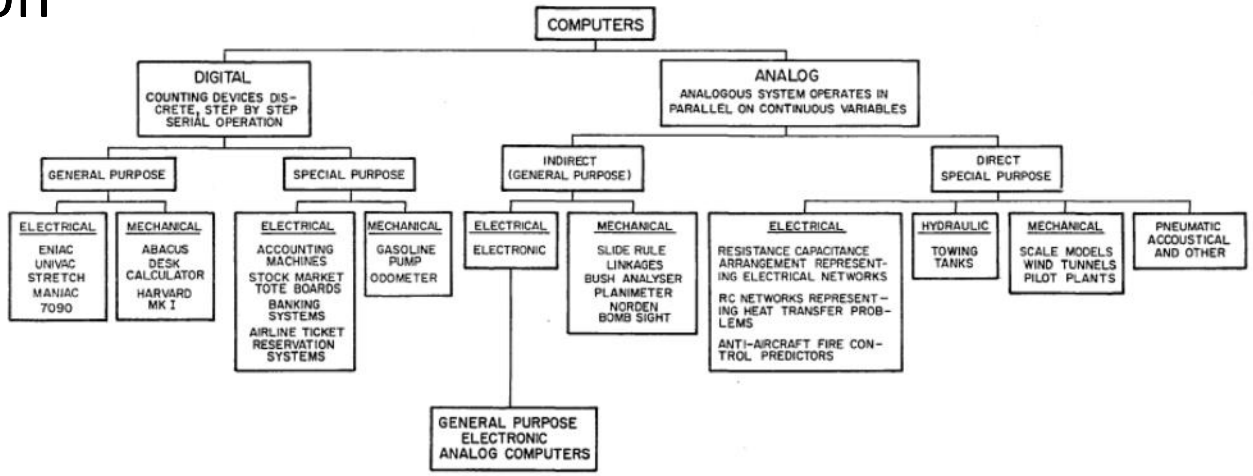
Version N°3
38 PFlop/s

20 Watts
30,000 €

Outline

1. Analogic computation
2. Computer's Zoo
3. A few words about gate
4. How to use transistors
5. Number's representation



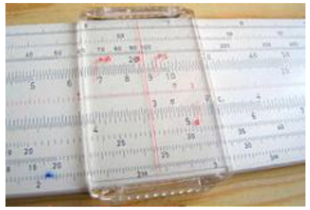


Outline

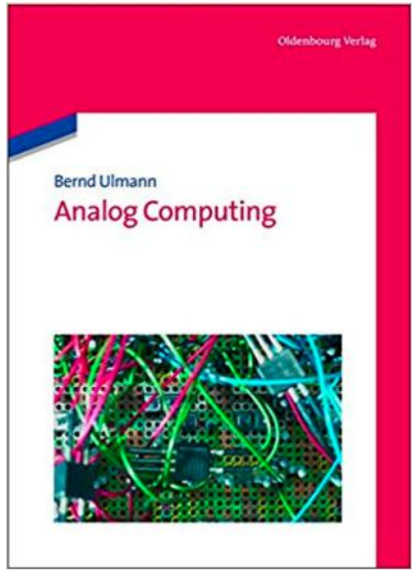
1. **Analogic computation**
2. Computer’s Zoo
3. A few words about gate
4. How to use transistors
5. Number’s representation

Numbers in continuous space: Analog computer

- Uses continuous property of physical phenomena (electrical, mechanical, hydraulic)
- Resilient to quantization noise
  But subject to physical noise (electronic, °C, …)
- Can be more powerful than digital computer
- Dominant in HPC until the 70s
  - Example:
    - Used for the Apollo 11 mission
    - Tide-Prediction machine (William & James Thomson)



Alternatives to numerical computer



- Human Brain
  - 38 PetaFlops, 20 Watts
- Analog computer
  - Not program by algorithm
    no need to store, fetch, decode instructions and operand
- Hybrid digital / analog computer
  - Complement their digital counterparts in solving equations relevant in:
    Biology, fluid dynamics, weather prediction, quantum chemistry, plasma physics, …

Toward an hybrid analog co-processor

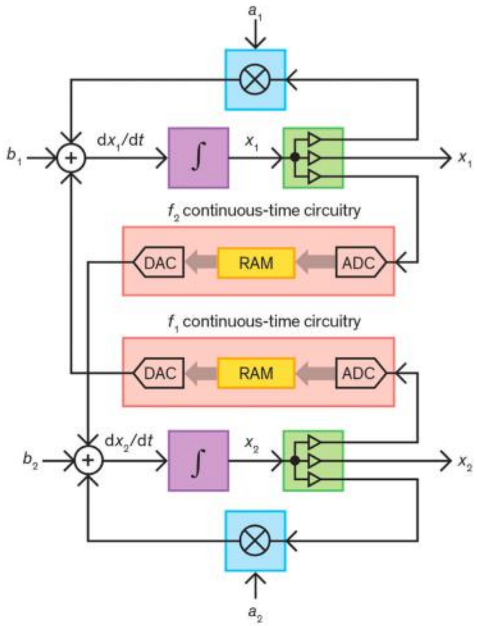
Telefunken RA770

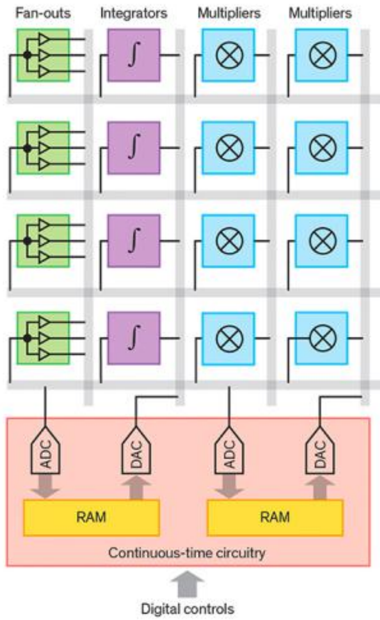
- System configured by the equations similar to the targeted system and allow variables in the analog computer to evolve with time
  - Analog computer to provide a quick approximation
  - Digital computer for programming, storage and precision
- Example

$$dx_1/dt = a_1x_1 + f_1(x_2) + b_1$$

$$dx_2/dt = a_2x_2 + f_2(x_1) + b_2$$

2 diff. equation with 2 variables





Cowan, G. E. R., Melville, R. C., & Tsividis, Y. P. (2006). A VLSI Analog Computer/Digital Computer Accelerator. IEEE Journal of Solid-State Circuits, 41(1), 42–53. doi:10.1109/jssc.2005.858618

Toward an hybrid analog co-processor

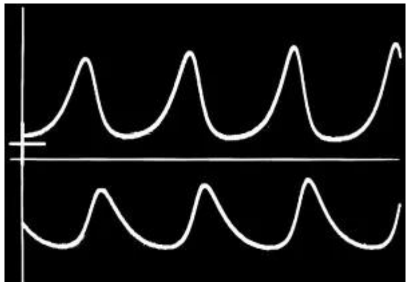
Predator-Prey model:

Rabbit evolution's

$$\frac{dr}{dt} = \alpha_1 r - \alpha_2 r f$$

Fox evolution's

$$\frac{df}{dt} = -\beta_1 f + \beta_2 r f$$



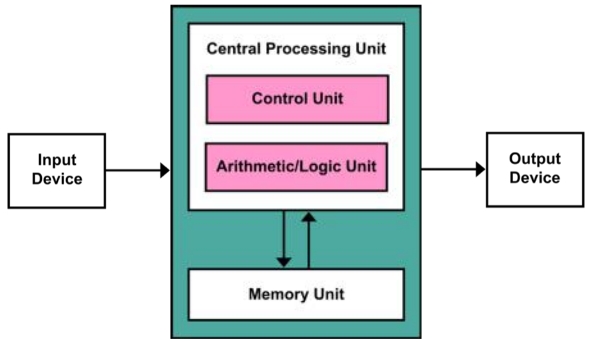
- Advantages
  - Voltages & currents evolves continuously
  - Highly parallel
  - No clock
  - Accuracy within a few percent
- Limitations
  - Large problem requires large numbers of analog computational blocks
    (or be able to divide by hand the problem => not parallel anymore)
  - Difficult to configure and connect distant analog blocks
  - Difficult in implementing multivalued functions
  - Approximate computing:
    Precision can't be increased by adding "bit", it requires larger chips
  - Difficult to program
    Compiler attempt

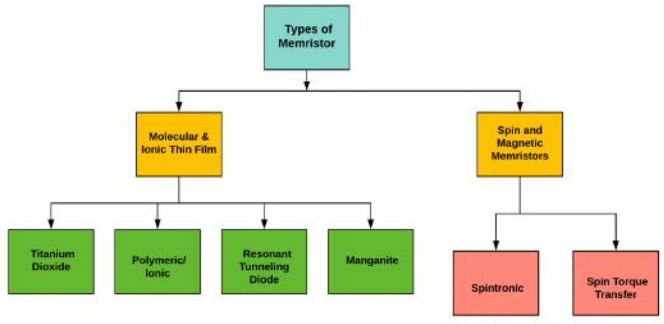
Neuromorphic computing

- Limit of the Turing-Von Neumann model

- Use analog circuit to mimic neuro-biological architectures
  - Memristors
  - Spintronic memories
  - Threshold switches

- Initiative to rethink the concept of computing
  - IEEE Rebooting Computing

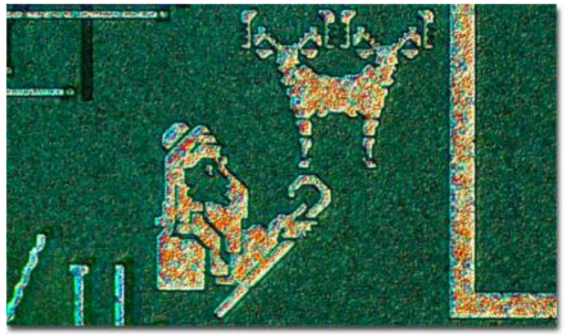




Outline

1. Analogic computation
2. **Computer's Zoo**
3. A few words about gate
4. How to use transistors
5. Number's representation

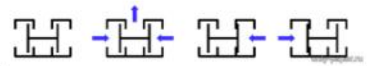
MicroVax 3000 (1987): in Cyrillic alphabet
" VAX—when you care enough to steal the very best".

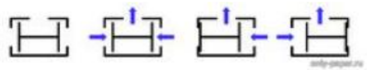


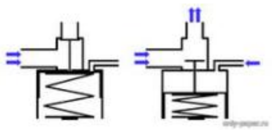
Magnified Intel 8207 controller dual-port RAM.
Shepherd with two headed ram.

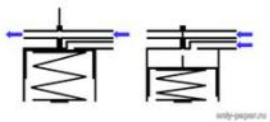
https://micro.magnet.fsu.edu/creatures/index.html

Binary pneumatic adder


Pneumatic AND Gate


Pneumatic OR Gate

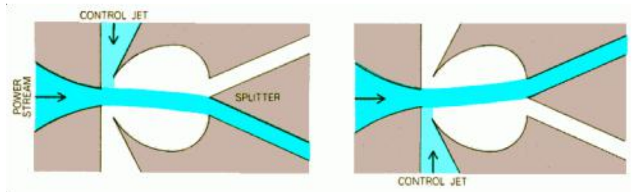

Transistor Amplifier


Pneumatic XOR Gate

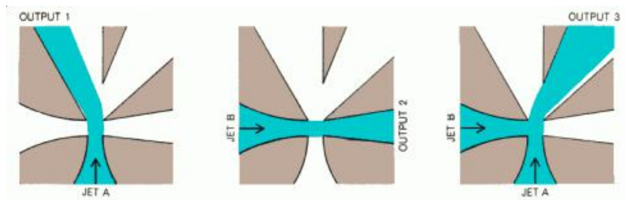
http://only-paper.ru/forum/38-23291-1

Fluidic calculator

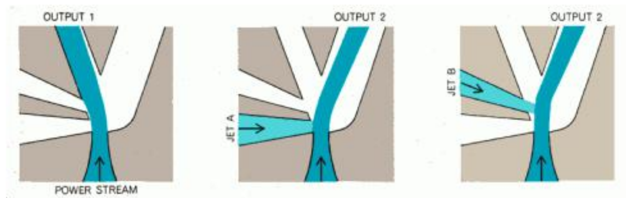
- Use of a Fluid to perform analog or digital operation
  - Pneumatics
  - Hydraulics
  - No moving part
    (paper computer is not consider fluidic)
- Use in environment where electronic digital is unreliable
  (electromagnetic interference, ionizing radiation…)



Fluidic Amplifier



Fluidic AND Gate

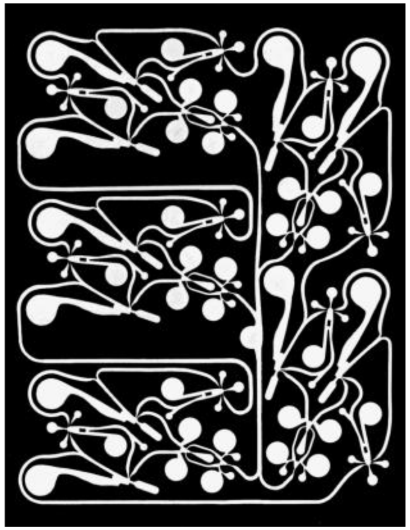


Fluidic OR Gate

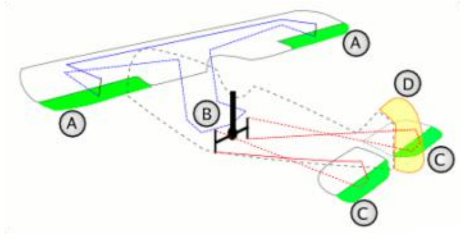
http://miriam-english.org/files/fluidics/FluidControlDevices.html

Fluidic calculator: Examples

MONIAC (1949) Computer to simulate economic principles at a time digital computer could not.



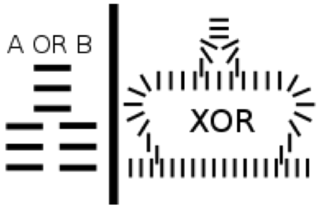
FLUID CIRCUIT performs the operation of dividing by 10 in an all-fluid digital computer: for every 10 input pulses circuit delivers one output pulse. Input pulses enter from above the plane of circuit through circular, bumplike hole attached to straight channel running from top to bottom just to right of center. The 10 identical logic elements, or modules, are arranged in a series of five pairs, three at left and two at right. Each pair contains two steady input streams (*small sausage shapes*), two outputs (*small circular shapes attached to short straight channels*), eight control jets (*small teardrop shapes*) and eight open vents (*large circular shapes*).



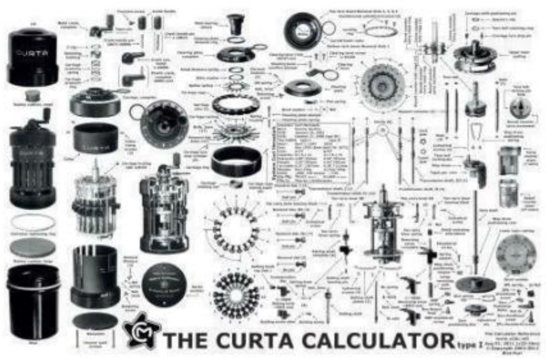
Aircraft Flight control systems, Valve in anesthesia machines for its advantages (lower mass, cost, drag, inertia, complexity)

Mechanical computers

- Various form
  - Build from levers & gears
  - Balls
  - Dominos



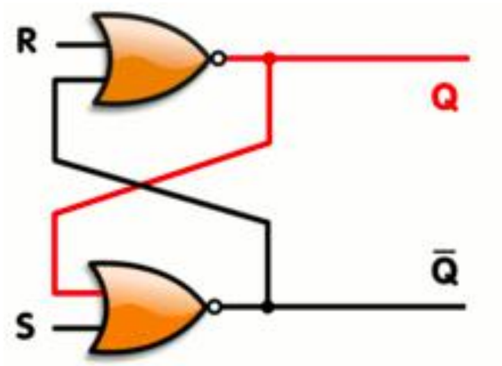
https://www.andrewt.net/maths/domputer/



Flip-Flop

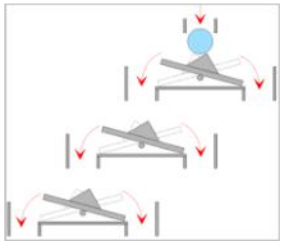
- Definition:

A Flip-Flop (latch) is a circuit that has 2 stable states and can be used to store state information.

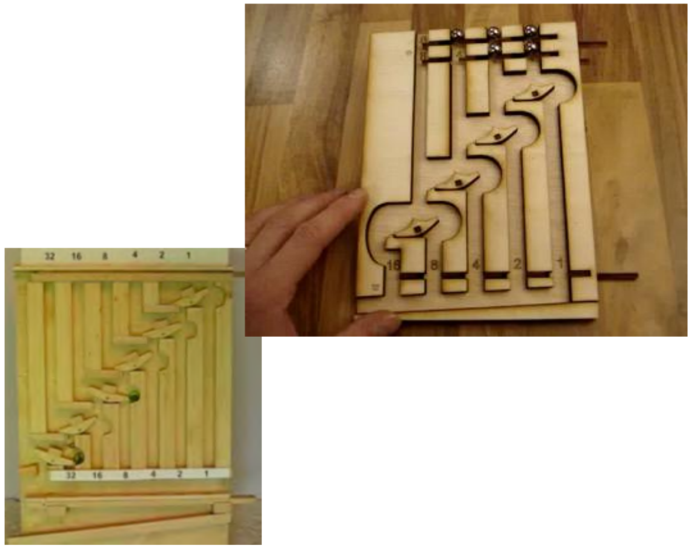


Marble's Flip-Flop

Objectives:
How to design a system with a minimum number of gate ?



Marble Flip-Flop

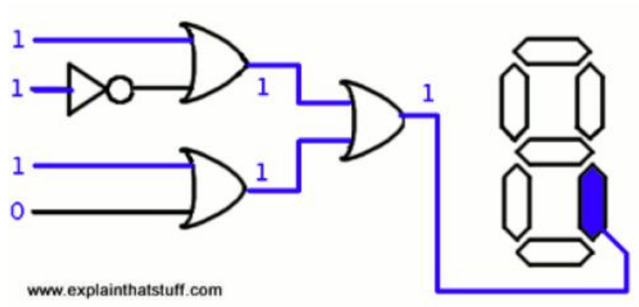


Marble's Adder

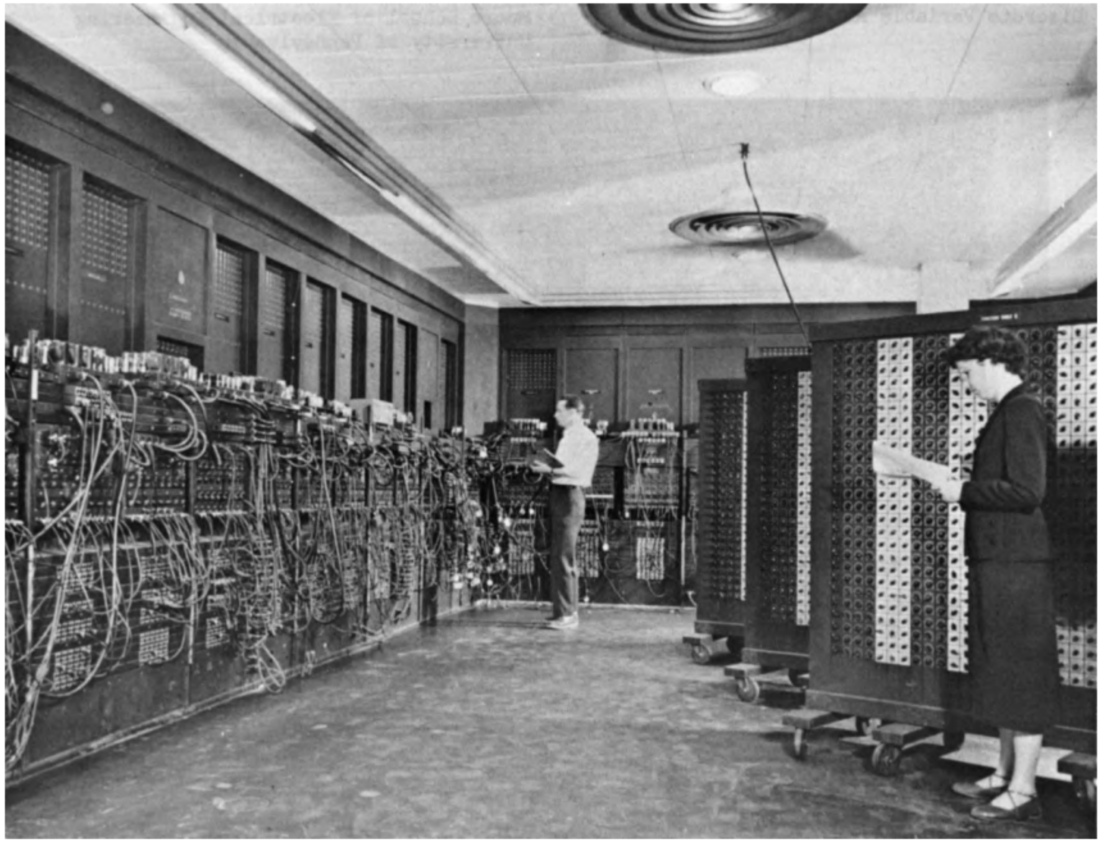
https://www.turingtumble.com/

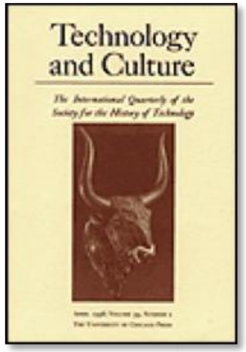
Outline

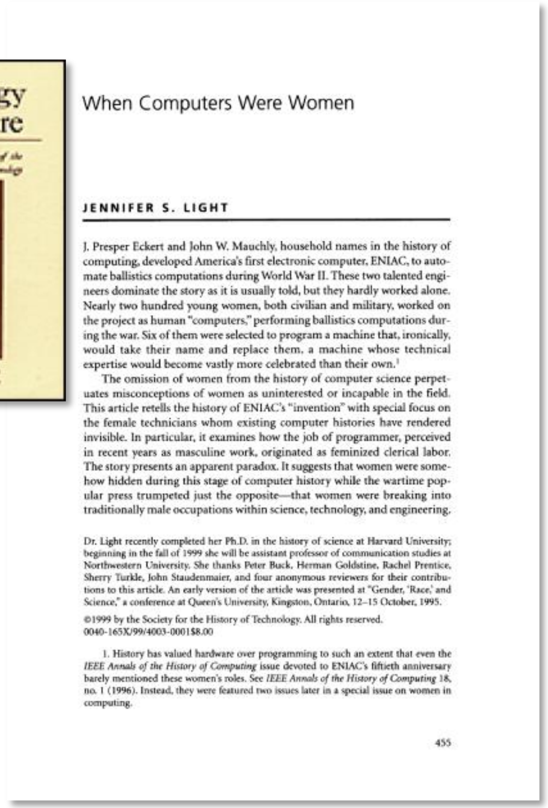
1. Analogic computation
2. Computer’s Zoo
3. **A few words about gate**
4. How to use transistors
5. Number’s representation



1945: ENIAC







When Computers Were Women

JENNIFER S. LIGHT

J. Presper Eckert and John W. Mauchly, household names in the history of computing, developed America's first electronic computer, ENIAC, to automate ballistics computations during World War II. These two talented engineers dominate the story as it is usually told, but they hardly worked alone. Nearly two hundred young women, both civilian and military, worked on the project as human "computers," performing ballistics computations during the war. Six of them were selected to program a machine that, ironically, would take their name and replace them, a machine whose technical expertise would become vastly more celebrated than their own.[1]

The omission of women from the history of computer science perpetuates misconceptions of women as uninterested or incapable in the field. This article retells the history of ENIAC's "invention" with special focus on the female technicians whom existing computer histories have rendered invisible. In particular, it examines how the job of programmer, perceived in recent years as masculine work, originated as feminized clerical labor. The story presents an apparent paradox. It suggests that women were somehow hidden during this stage of computer history while the wartime popular press trumpeted just the opposite—that women were breaking into traditionally male occupations within science, technology, and engineering.

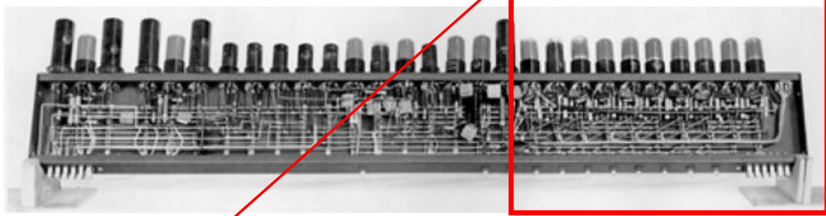
Dr. Light recently completed her Ph.D. in the history of science at Harvard University; beginning in the fall of 1999 she will be assistant professor of communication studies at Northwestern University. She thanks Peter Buck, Herman Goldstine, Rachel Prentice, Sherry Turkle, John Staudenmaier, and four anonymous reviewers for their contributions to this article. An early version of the article was presented at "Gender, 'Race,' and Science," a conference at Queen's University, Kingston, Ontario, 12–15 October, 1995.

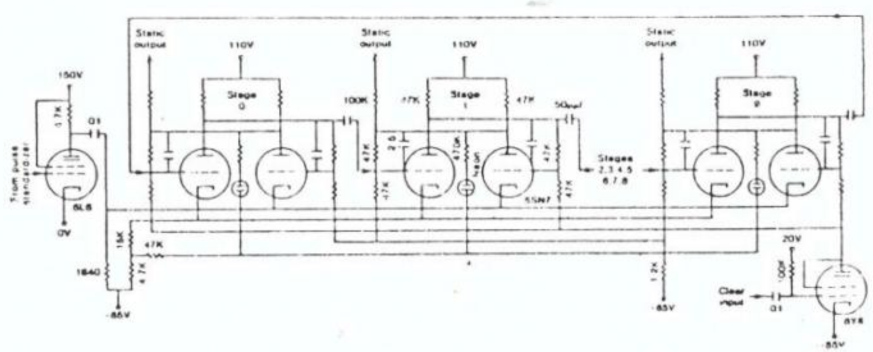
0040-165X/99/4003-0001$8.00

1. History has valued hardware over programming to such an extent that even the *IEEE Annals of the History of Computing* issue devoted to ENIAC's fiftieth anniversary barely mentioned these women's roles. See *IEEE Annals of the History of Computing* 18, no. 1 (1996). Instead, they were featured two issues later in a special issue on women in computing.

ENIAC: Numbers

- 20 accumulators / modules
  - ADD / SUBSTRACT
  - 10 Decimal Digit « Register » in
    - Serial decimal (10 bits per digit)
    - With a ring counter made of cascaded flip-flop (bit shift) (!= binary counter)
    - Arithmetic done by "counting" pulses with the ring counter and generate carry pulse if the counter wrapped around. (idea borrowed from digit wheels of mechanical adding machine)
  - Constant entered with punched cards
  - Numbers transferred between module through buses

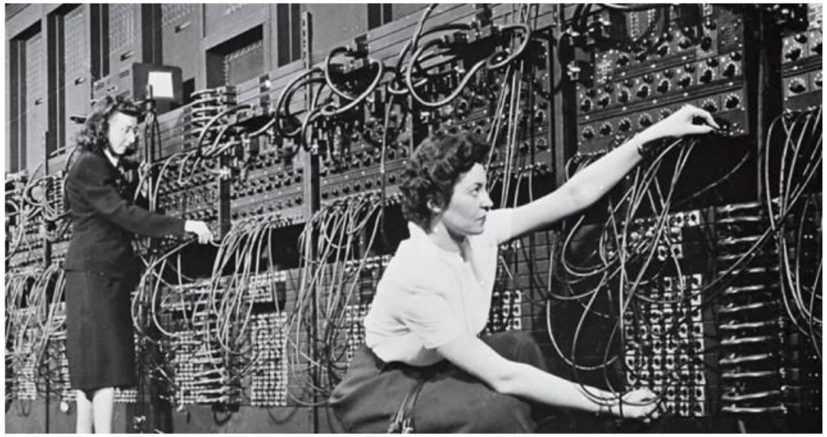




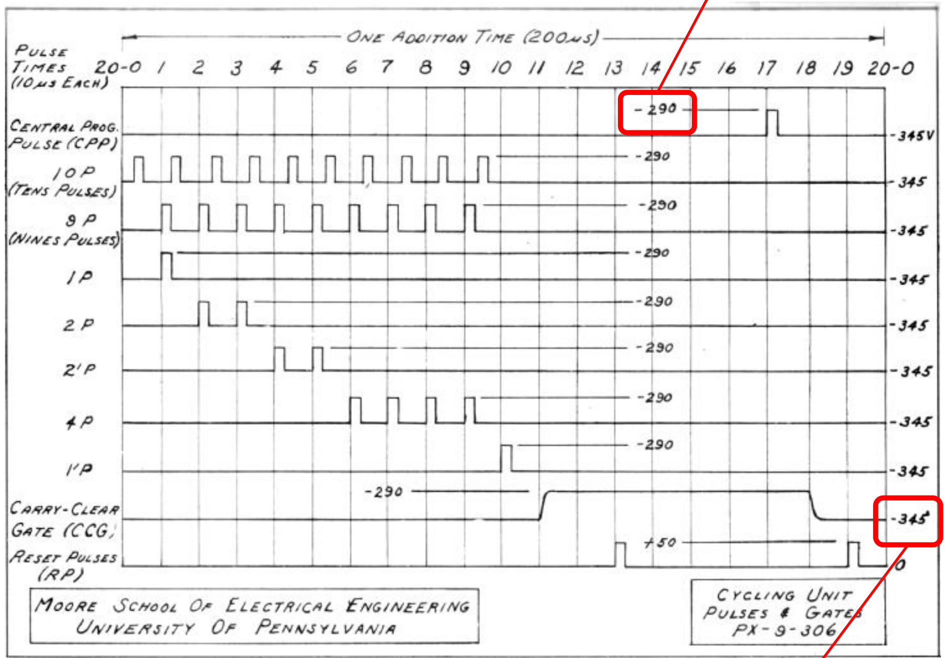
ENIAC decade ring counter. From Burks, "Electronic Computing Circuits of the ENIAC," *Proceedings of the IRE* 35: 746–767, © IRE (now IEEE), 1947. All rights reserved.

ENIAC: Operations

-290 V = Value 1



Programming done by women
(switches and cables)



-345 V = Value 1

https://archive.computerhistory.org/resources/text/Knuth_Don_X4100/PDF_index/k-8-pdf/k-8-r5367-1-ENIAC-circuits.pdf

ENIAC: Multiplication

Question:

How to decompose the set of numbers {0,1,2,3,4,5,6,7,8,9}
With a set of numbers of minimal size ?

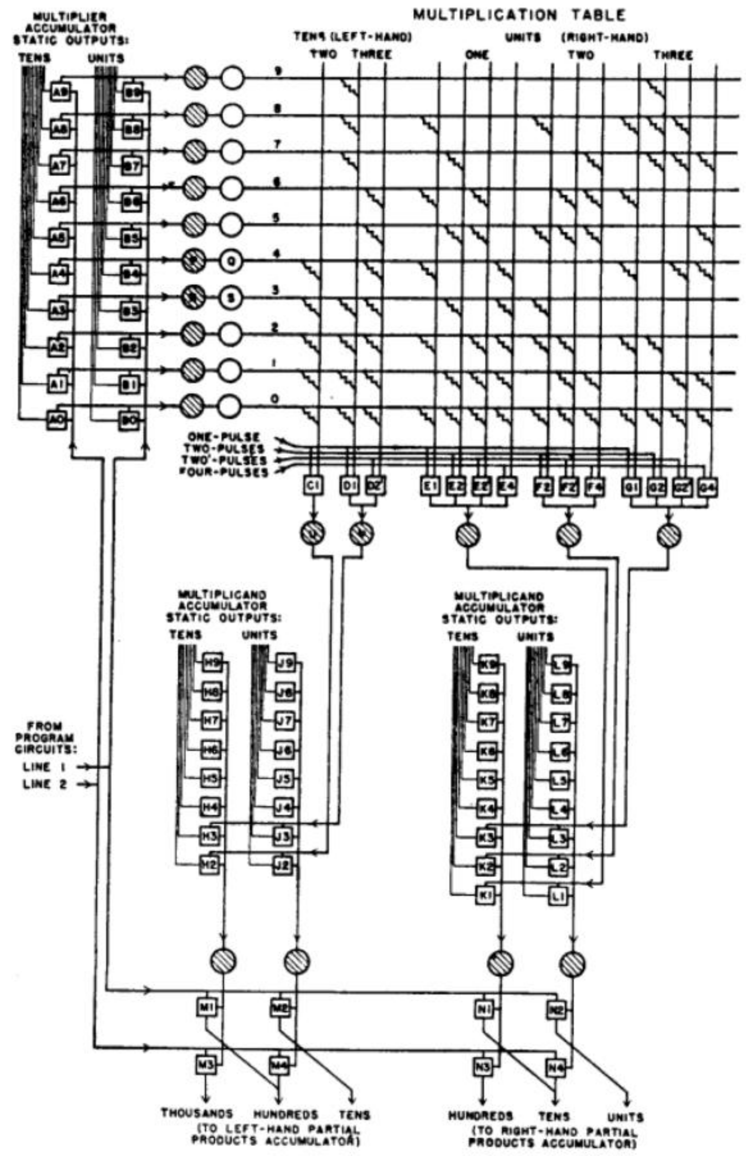


Fig. 4—Block diagram of multiplication circuits.

What is the best radix ?

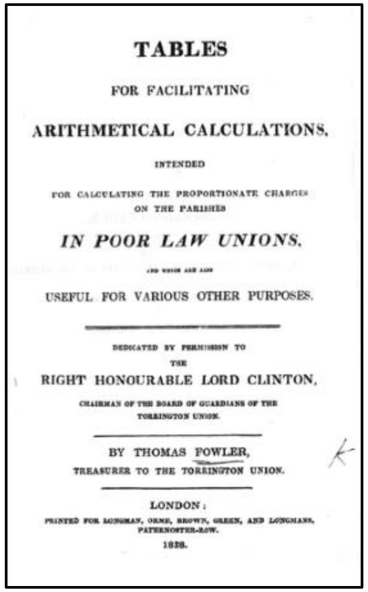
- A metric: Radix economy
  $E(b, N) = \lfloor log_b(N) + 1 \rfloor . b$
  - Number of digits (N) * Number of possible values each digit could have (b)
  - Example: 100_{10}
    - 100_{10}: 3 digits   E = 3 * 10 = 30
    - 1100100_2: 7 digits   E = 7 * 2 = 14
    - 10201_3: 5 digits   E = 5 * 3 = 15
- But binary system has greater noise immunity

Base	E(b,5329)
1	2665
2	22,9
e	**22,1**
3	22,2
4	23,9
5	26,3
6	28,3
7	31,3
8	33
9	34,6
10	37,9
16	50,9
20	58,4
30	84,8
40	107,7
60	138,8

Ternary calculation

- Fowler's machine (1840), Treasurer of the Poor Law Union
    - Used to compute the proportional fee for each parish of the Poor Law Union
    - Difficult task in the pre-decimal English currency
        - 20 shilling = 1 pound
        - 12 pence = 1 shilling
        - 4 farthings = 1 penny
    - Need to convert everything in farthing (=> large numbers)

St. Michaels Church, Devon, England


TABLES

FOR FACILITATING

ARITHMETICAL CALCULATIONS,

INTENDED

FOR CALCULATING THE PROPORTIONATE CHARGES
ON THE PARISHES

IN POOR LAW UNIONS,

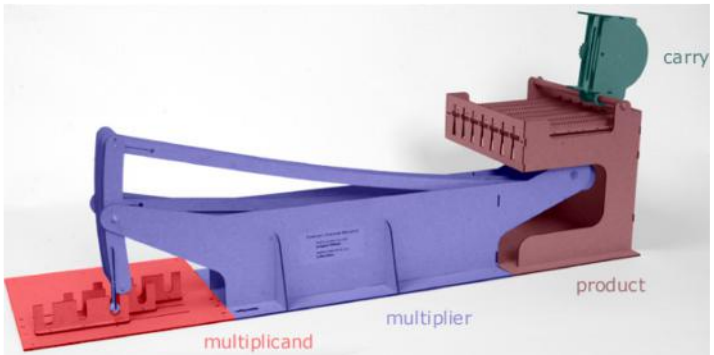
AND WHICH ARE ALSO

USEFUL FOR VARIOUS OTHER PURPOSES.

DEDICATED BY PERMISSION TO
THE
RIGHT HONOURABLE LORD CLINTON,
CHAIRMAN OF THE BOARD OF GUARDIANS OF THE
TORRINGTON UNION.

BY THOMAS FOWLER,
TREASURER TO THE TORRINGTON UNION.

LONDON:
PRINTED FOR LONGMAN, ORME, BROWN, GREEN, AND LONGMANS,
PATERNOSTER-ROW.
1838.



http://www.mortati.com/glusker/fowler/tables.htm

1958: SETUN

http://trinary.ru/projects/setunws/

- Ternary computer
  - Balanced ternary numeral system (-1 | 0 | +1)
  - Three-valued ternary logic
    (16 binary operators in Boolean logic & 19 683 in Ternary logic !)
  - Words made of 18 Trits
  - Each trits stored in a pair of magnetic cores, wired in tandem to represent 3 stable states
- 1960's
  - Attempt to build ternary logic gates and memory cells
  - 1973 TERNAC (G. Frieder)
- Possible usage:
  - Found in redundant binary representation to avoid carries propagation
  - 2009: quantum computer made of qutrit (instead of qubit)

... and according to

Perhaps the prettiest number system of all, is the balanced ternary notation

1958: SETUN

http://trinary.ru/projects/setunws/

- Ternary computer
  - Balanced ternary numeral system (-1 | 0 | +1)
  - Three-valued ternary logic
    (16 binary operators in Boolean logic & 19 683 in Ternary logic !)
  - Words made of 18 Trits
  - Each trits stored in a pair of magnetic cores, wired in tandem to represent 3 stable states

- 1960's
  - Attempt to build ternary logic gates and memory cells
  - 1973 TERNAC (G. Frieder)

- Possible usage:
  - Found in redundant binary representation to avoid carries propagation
  - 2009: quantum computer made of qutrit (instead of qubit)

… and according to

Perhaps the prettiest number system of all, is the balanced ternary notation

Balanced Number: Why is that such a nice system ?

- System discussed by
  A. Cauchy (1840), J. Leslie (1817), J. Colson's (1726), … Hindu Vedas (~1000 BCE)
- Balanced = arranged symmetrically around 0
  - Ex = { -1 , 0 , 1 }
  - 2-pan Balance to measure between 1-40 g,
    How many weights required, when weights can go in 1 / 2 pan?
  - (1,2,4,8,16,32 / 1,3,9,27)

- Properties
  - Ease comparison (3 possible states (< = >) vs 2 (Yes/No))
  - Test Odd / Even:
    - Radix2: Last digits
    - Radix3: Number of 1 in the numeral
  - Cut down the carry-rate in multi-digit multiplication
  - No carry for 1 digit mult.
  - Rounding = Truncation
  - No diff. Between + / - numbers

	27	9	3	1
1				1
2			1	-1
3			1	0
4			1	1
5		1	-1	-1
6		1	-1	0
7		1	-1	1
8		1	0	-1
9		1	0	0
10		1	0	1
11		1	1	-1
12		1	1	0
13		1	1	1

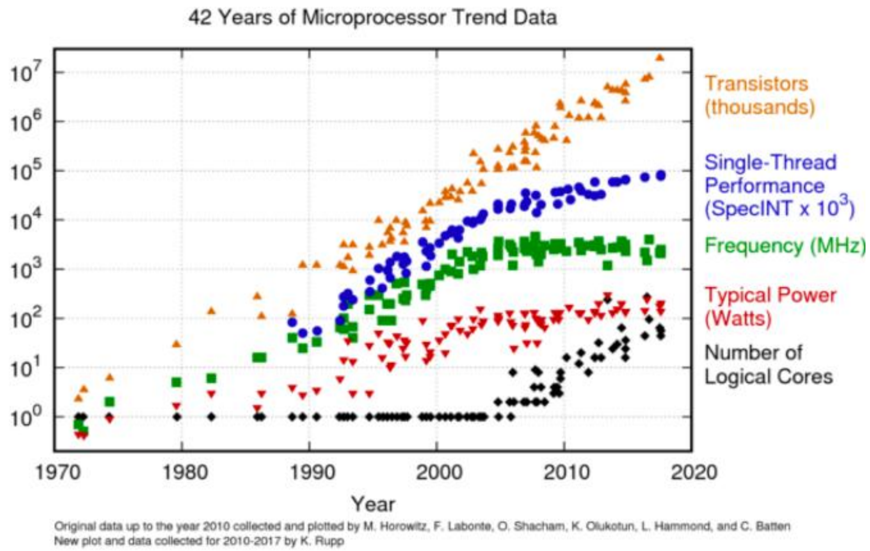
Outline

1. Analogic computation
2. Computer’s Zoo
3. A few words about gate
4. **How to use transistors**
5. Number’s representation

Ways to use available transistors

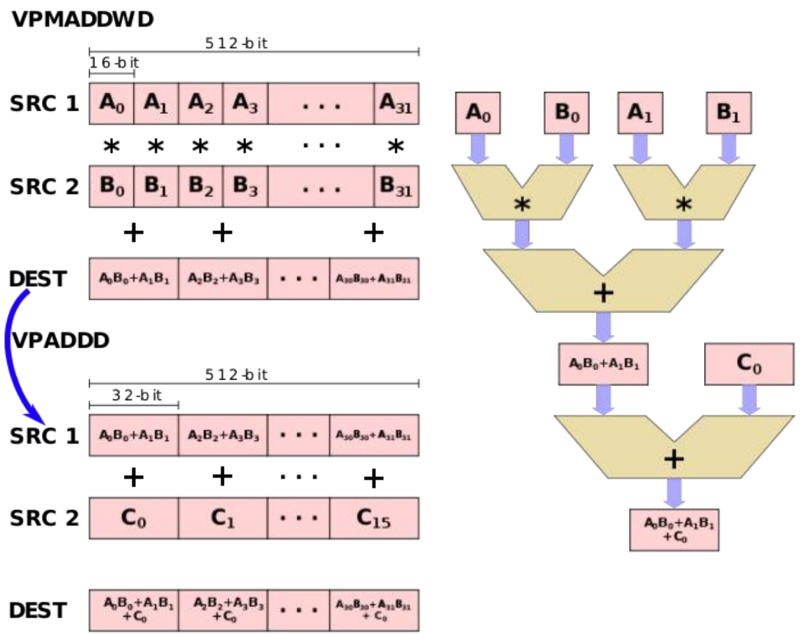
1. 2000’s: Vector Unit (Extension ISA)
2. 2010’s: More Cores (GPU)
3. 2020’s: Dedicated core (Neural Network Processing unit)

Year	1996		1998	1999		2001			2004	2005	2006	2007	2008	2009			2012	2013	2014	2015
SIMD	MMX		3DNow!	SSE		SSE2			SSE3		SSE4		AVX	F16C / XOP			FMA	AVX2		AVX-512
SIMD Length	64			128									256							512
Bit Manipulation												ABM					BMI1	BMI2	ADX	
Compressed Inst																				
Crypto / Sec													AES-NI			CLMU	RdRand	SHA		MPX / SGX
Trans. mem																		TSX		
Virtualization										VT-x										
# inst.	57			70		144			13		54		12	60			24	30		



https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/

Example: AVX512 VNNI



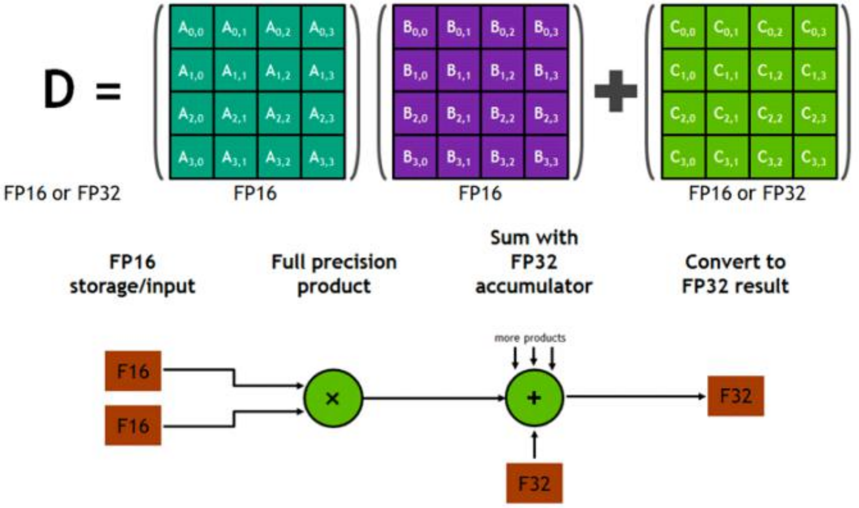
Trend: Fused operators

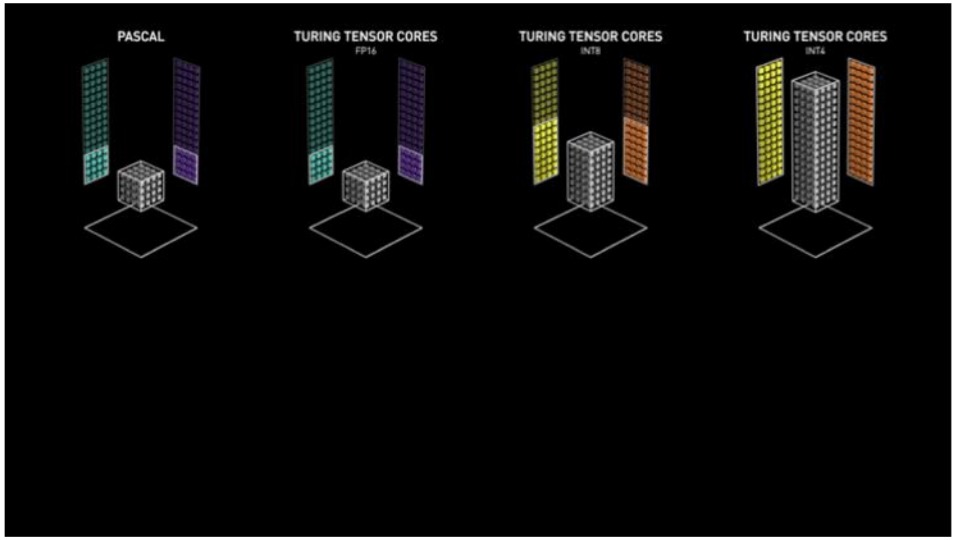
Nb of fused operators

Name	++	+*(FMA)	+/	*+	**	*/	/+	/*	//
168.wupwise	38	**74**	0	42	27	0	7	4	0
171.swim	180	**184**	0	162	48	0	10	7	0
172.mgrid	**249**	98	0	91	11	0	2	0	0
173.applu	**1826**	1783	86	1797	782	7	69	109	0
177.mesa	116	**274**	0	164	71	0	10	10	0
178.galgel	358	**856**	26	415	443	21	57	36	10
179.art	73	98	0	**100**	34	0	8	3	0
183.equake	52	**119**	3	75	74	8	10	8	1
188.ammp	260	**418**	2	291	205	5	12	19	0
189.lucas	80	**402**	0	87	54	2	2	2	0
191.fma3d	215	**461**	12	266	144	11	24	32	0
301.apsi	514	**1015**	175	881	673	83	192	302	14
Total	3961	**5782**	304	4371	2566	137	403	532	25

D. Defour, « Collapsing dependent floating-point operations », *IMACS World Congress Scientific Computation, Applied Mathematics and Simulation*, 2005

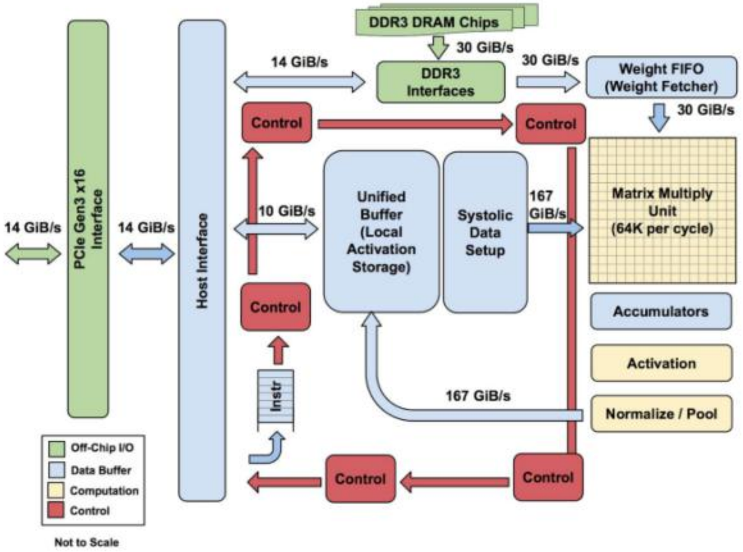
Nvidia Tensor Core

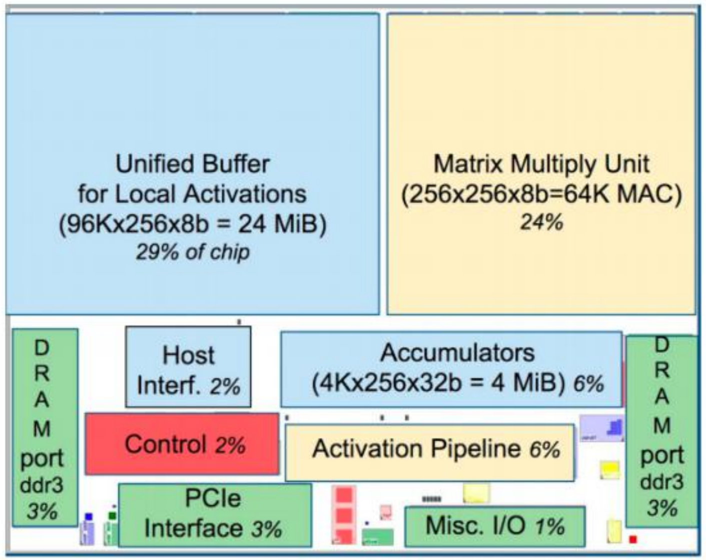




https://www.anandtech.com/show/12673/titan-v-deep-learning-deep-dive/3
https://arxiv.org/pdf/1804.06826.pdf

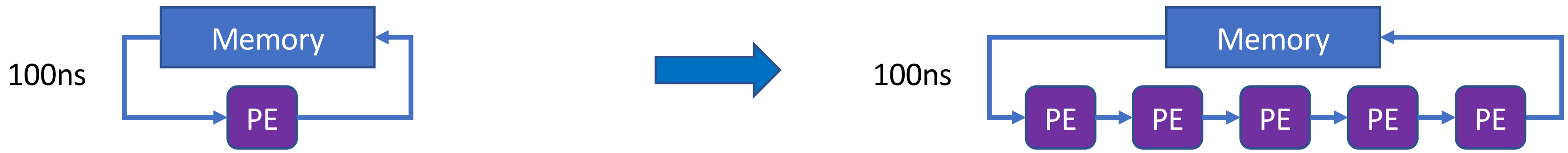
Google Tensor Processing Unit

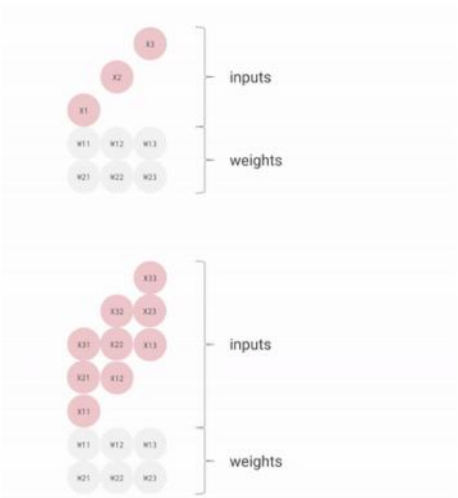


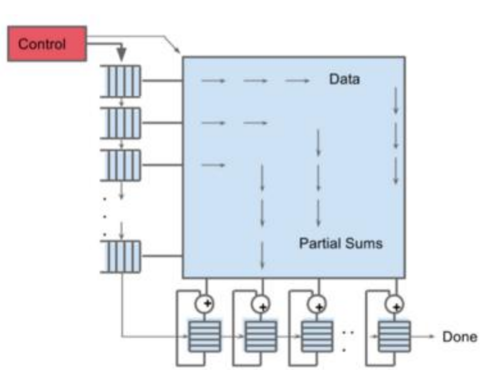


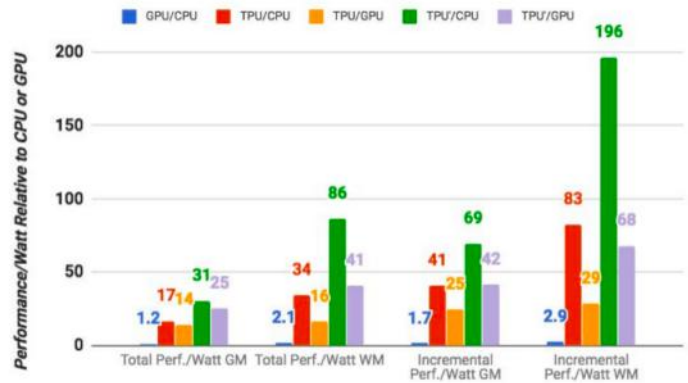
Google Tensor Processing Unit: A systolic array

- H.T. Kung, “Why systolic Architectures ?” IEEE Computer 1982.







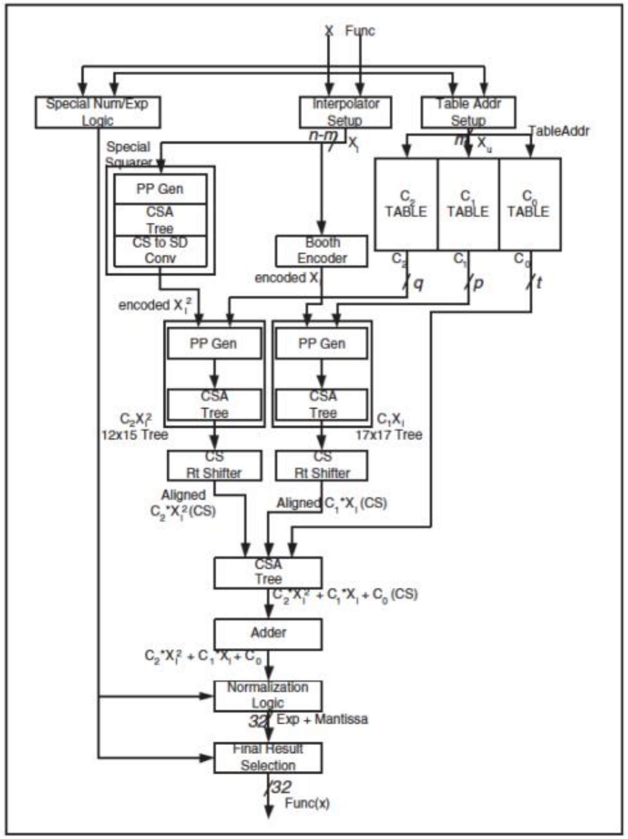


N. P. Jouppi *et al.*, "In-datacenter performance analysis of a tensor processing unit," *2017 ACM/IEEE 44th Annual International Symposium on Computer Architecture (ISCA)*, Toronto, ON, 2017, pp. 1-12.
https://cloud.google.com/blog/products/gcp/an-in-depth-look-at-googles-first-tensor-processing-unit-tpu

Nvidia Special Function Unit

$$F(x) = C_0 + C_1 x + C_2 x^2$$

Function	Input Interval	m	Configuration	Accuracy (good bits)	ulp error	% exactly rounded	monotonic	Lookup table size
$1/X$	[1,2)	7	26,16,10	24.02	0.98	87%	yes	6.50Kb
$1/\sqrt{X}$	[1,4)	6	26,16,10	23.40	1.52	78%	yes	6.50Kb
2^X	[0,1)	6	26,16,10	22.51	1.41	74%	yes	3.25Kb
$\log_2 X$	[1,2)	6	26,16,10	22.57	n/a	n/a	yes	3.25Kb
sin/cos	$[0, \pi/2)$	6	26,15,11	22.47	n/a	n/a	no	3.25Kb
Total								22.75Kb



S. F. Oberman and M. Siu. A high-performance area-efficient multifunction interpolator. IEEE ARITH 17, pages 272–279, July 2005.

Rise of a new era

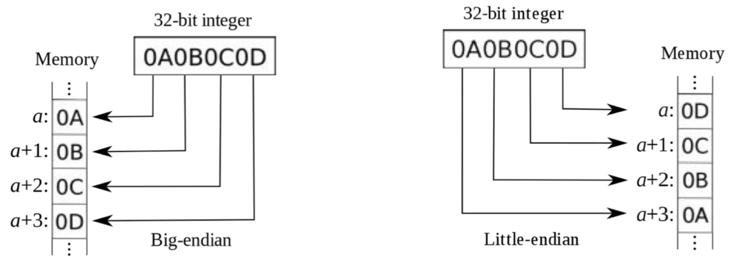
- Uses transistors for specialized Hardware
  - Nvidia SFU, Tensor Core
  - Google Tensor Processing Unit
  - Intel Loihi
  - Microsoft Catapult
  - Movidius Myriad 2 VPU
  - Neuromorphic processor, IBM TrueNorth, SpiNNaker

- Quantized format
  - FP16
  - BFLOAT
  - Flexpoint
  - Unum
  - FPANR

Outline

1. Analogic computation
2. Computer’s Zoo
3. A few words about gate
4. How to use transistors
5. **Number’s representation**

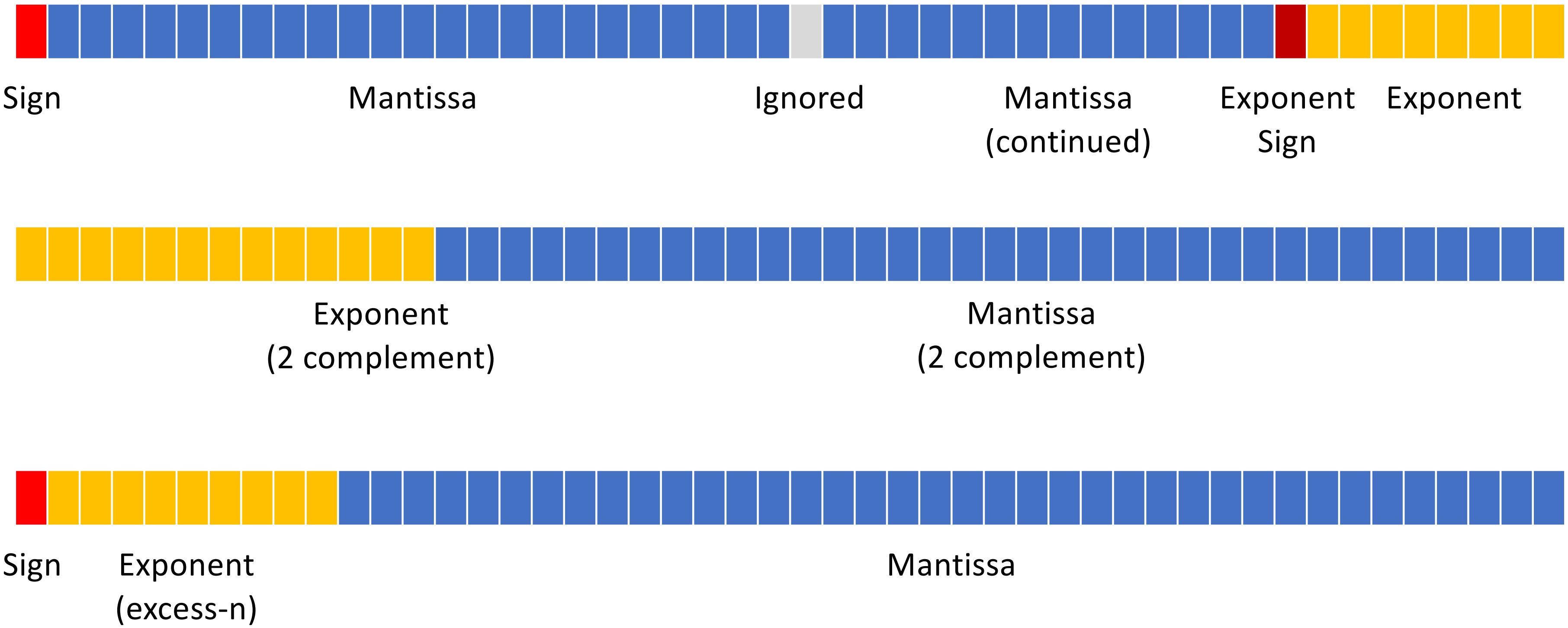
Endianness



- Little-endian
  - Ease multi-byte addition due to carry propagation (reason Datapoint & Intel selected it)
  - Allow to read a numerical value independently of the length as long as it fits into memory
- Big-endian
  - Allow to get an approximation by reading the first byte (sloppy arithmetic)
  - Ease division (MSB)
- Middle-endian
  - PDP-11

Floating-Point formats

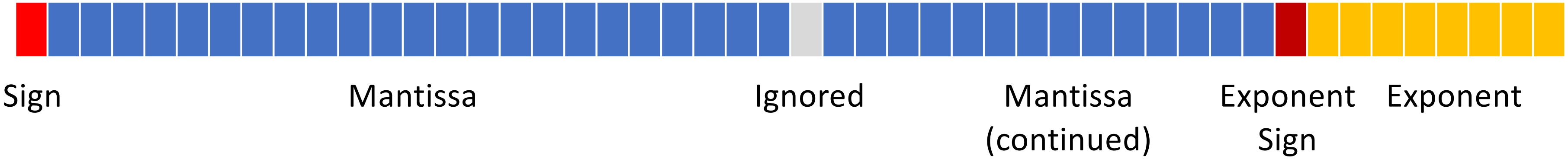
- 3 basic formats



http://www.quadibloc.com/comp/cp0201.htm

Floating-Point formats (1)

Konrad Zuse



- FP format found on computer with 24-bit word length
  - HP 2114-2116, SDS-9 series
- No hardware support
  - Hardware integer multiplication
- 1st bit of the second word ignored
  (or sign copy to easy Fixed-Point implementation)
- FP numbers represented by 2 signed binary numbers

Floating-Point formats (2)

PDP-8

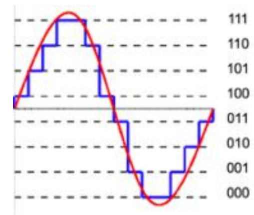
- FP format available on PDP-8 (1965) & RECOMP II (1958)
- FP numbers represented by 2 signed binary numbers

Floating-Point formats (3)

William Kahan

- FP compare instruction == Integer compare for positive nb
- Popular on architecture with FP hardware support
- PDP-11
  1970: came with the hidden first bit trick
- IEEE-754
  1985: W. Kahan

Quantized Format



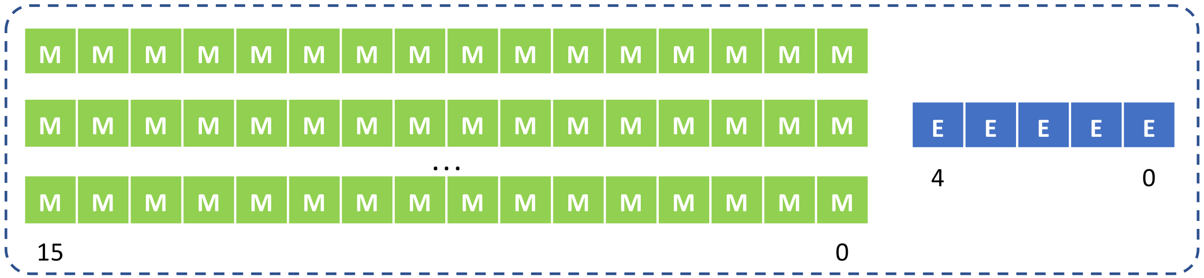
Range: ~0.00006 ; 65 504

IEEE-754 FP16

Range: ~1e^{-38} ; 3e^{+38}

BFLOAT16

FlexPoint: Flex16+5



FPANR FP16

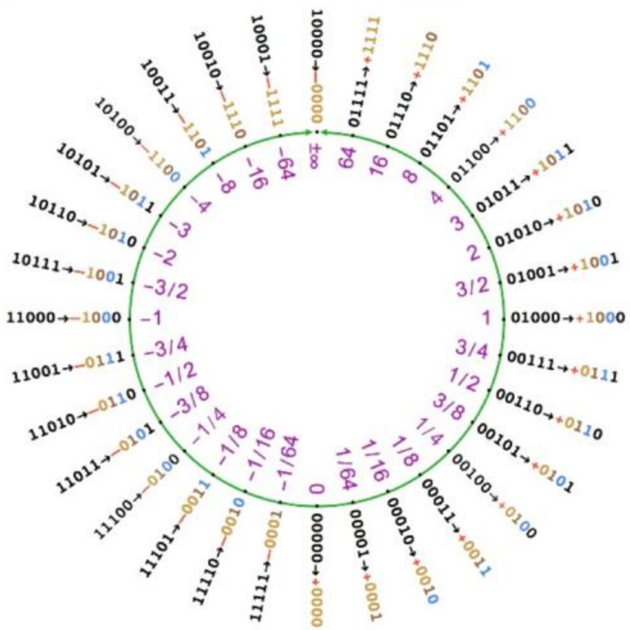
D. Defour, "FP-ANR: A representation format to handle floating-point cancellation at run-time," *2018 IEEE 25th Symposium on Computer Arithmetic (ARITH)*, Amherst, MA, 2018, pp. 76-83.

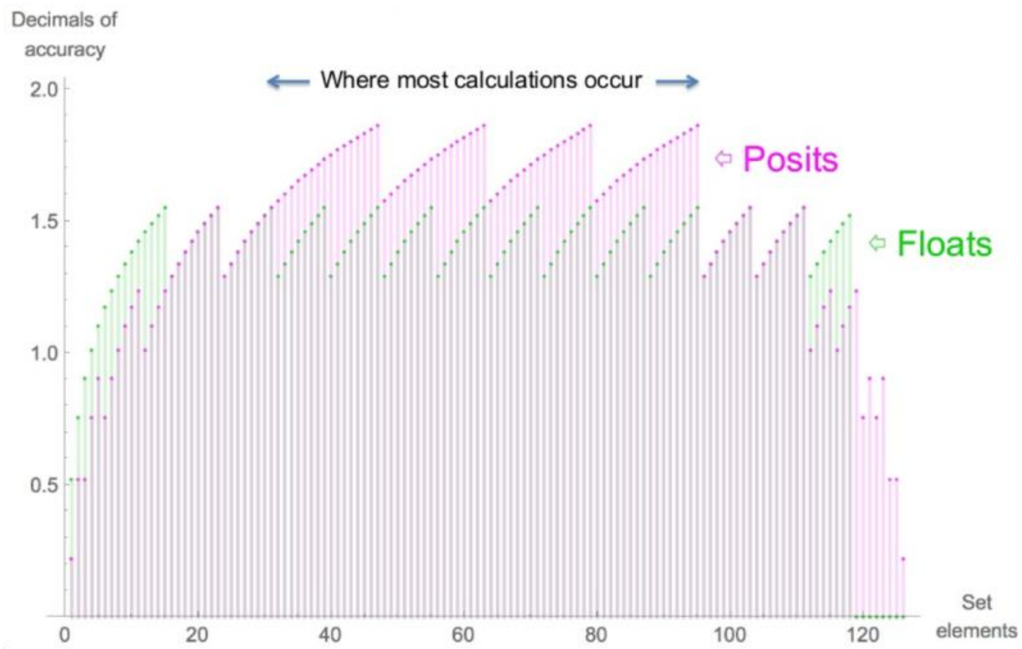
Posit aka Unum III

POSIT (ES=2)

S	R	R	R	R	R	$\bar{R}$	E	E	M	M	M	M	M	M	M
15					10										0

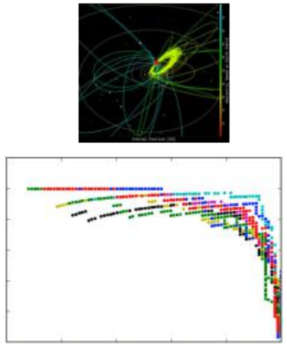
$0\_001\_11\_1 = +(2^4)^{-3} \,.\, 2^3 \,.\, (1+1/2)$



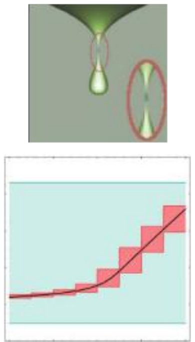


Number’s representation

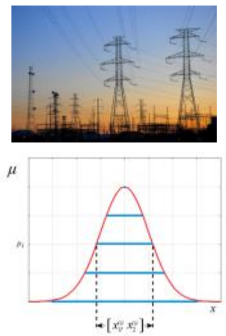
- Fixed-point
- **Floating-point**
- Logarithmic number systems
- **Tapered floating-point representation**
- Significance arithmetic
- Arbitrary-precision
- Floating-point expansions
- Rational arithmetic
- Interval arithmetic
- Algebraic system



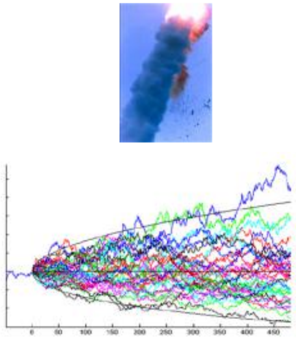
Multiprecision



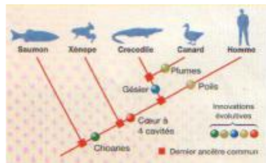
Interval



Floue



Stochastique



Gogoplex/minex
$10^{gogol} = 10^{10^{100}}$

LNS

… but always remember that according to

Recurring attempts to invent cheaper substitutes for Interval Arithmetic have all failed in the end after enough local limited success initially to tantalize their inventors with **dreams of glory**.

Trend in computation

- Limit of the traditional Von Neuman –Turing model
  - Power consumption
  - Care about data transfer (Bandwidth, Latency)
  - New ISA
- Rise of alternative representation format
  - Each bit count
  - Low precision & Approximate computing
  - Mixed-precision
  - Analogic computation (neuromorphic processor)
- New way of processing data
  - Fused operators
  - Self-correcting algorithms