



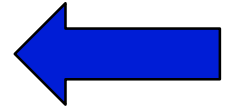
- Vittore Casarosa
 - Mail: casarosa@isti.cnr.it
 - Tel: 050 621 3115 (office) 348 397 2168 (mobile)
 - Skype: vittore1201
- “Ricevimento” at the end of the lessons or by appointment
- Final assessment
 - 70% oral examination
 - 30% project (development of a small digital library))
- Reference material:
 - Ian Witten, David Bainbridge, David Nichols, How to build a Digital Library, Morgan Kaufmann, 2010, ISBN 978-0-12-374857-7 (Second edition)
 - Material provided by the teacher
- **<http://cloudone.isti.cnr.it/casarosa/BDG/>**

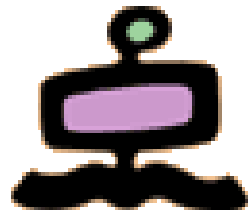


Modules

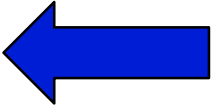
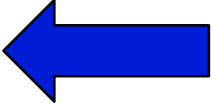


-
- Computer Fundamentals and Networking
 - A conceptual model for Digital Libraries
 - Bibliographic records and metadata
 - Information Retrieval and Search Engines
 - Knowledge representation
 - Digital Libraries and the Web
 - Hands-on laboratory: the Greenstone system

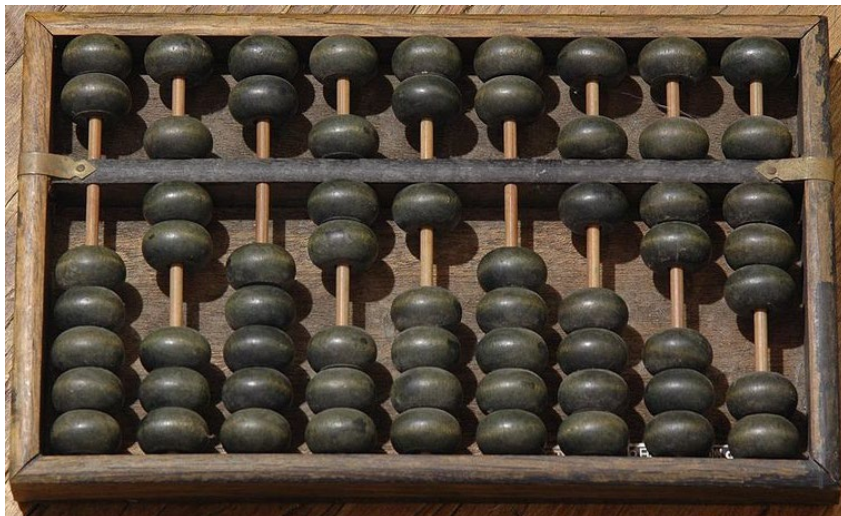
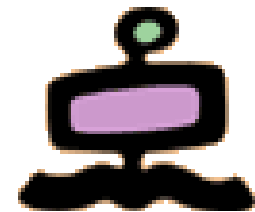




Refresher on Computer Fundamentals and Networking

- History of computers 
- Architecture of a computer 
- Computer networks and the Internet
- Data representation within a computer

Early devices (not computers)

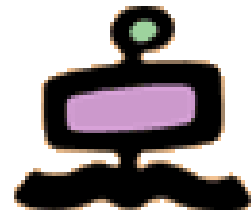


Abacus
2500 BC

Pascalina
1645



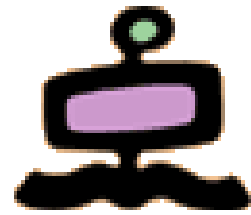
Early visions (of computers)



Charles Babbage
1791-1871

Professor of
Mathematics,
Cambridge University,
1827-1839

Babbage's engines



- *Difference Engine* 1823
- *Analytic Engine* 1833
 - The forerunner of modern digital computer

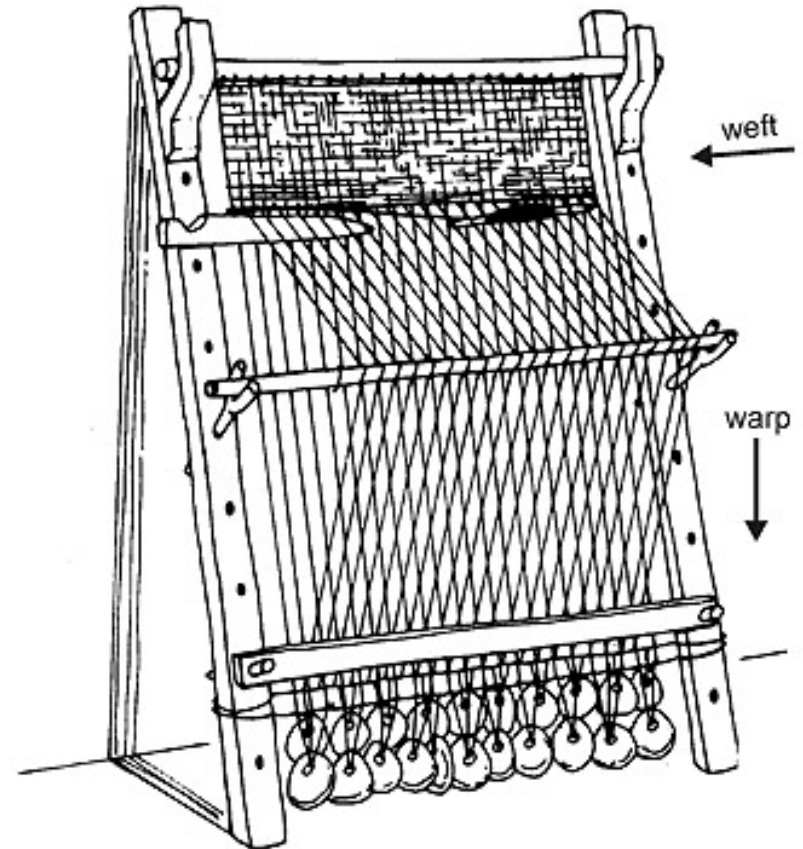
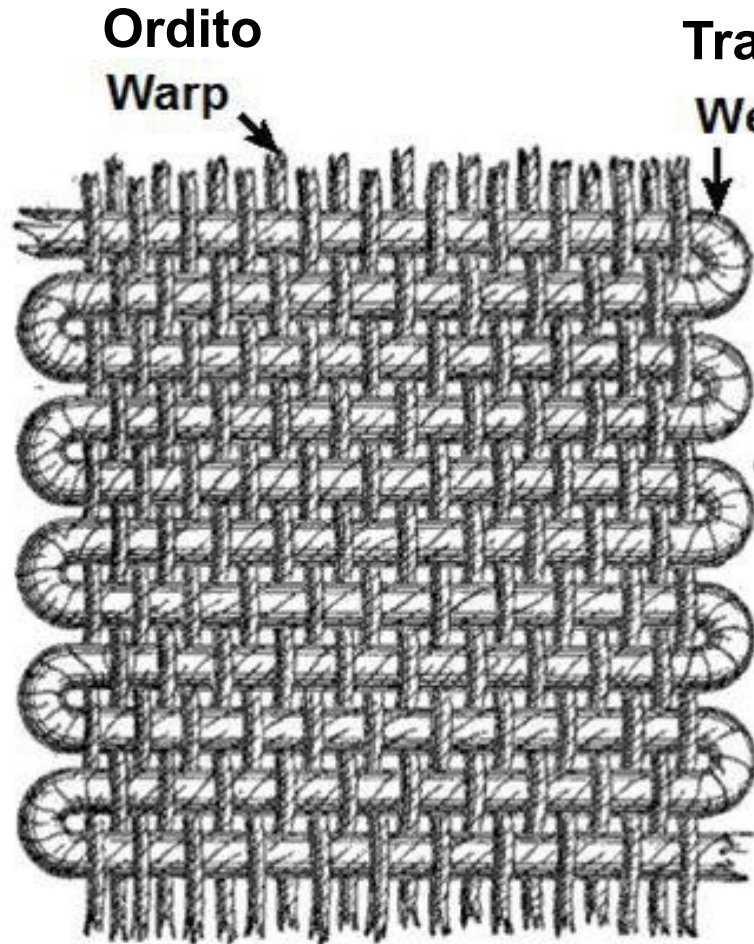
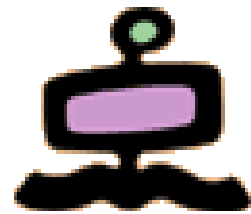
Application

- Mathematical Tables – Astronomy
- Nautical Tables – Navy

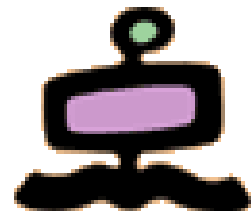
Technology

- mechanical gears, Jacquard's loom (1801), simple calculators

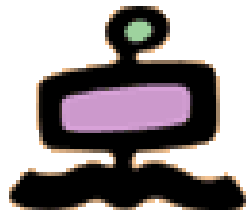
The loom



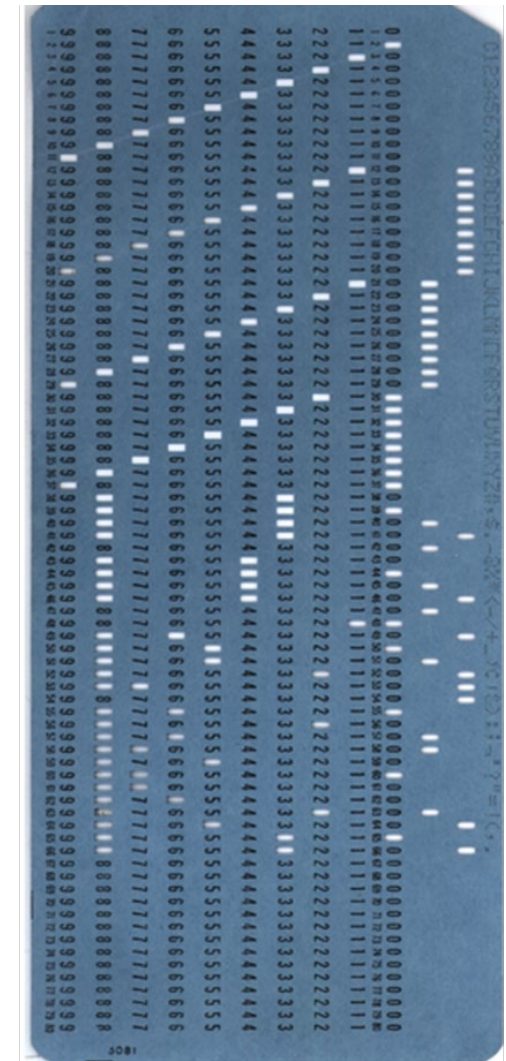
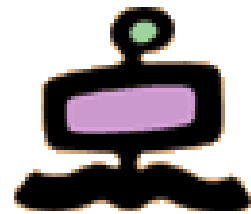
Use of punched paper tape



The organ grinder

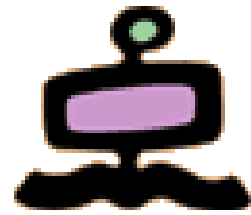


«Programmable» carillon and punched cards





Early experiments 100 years later



- Z1 machine (Konrad Zuse, private entrepreneur, 1936-1941)
- ABC (Atanasoff-Berry Computer, Iowa State University, 1937-1942)
- Mark I (Howard Aiken, MIT, 1937-1941)

1942 Second World War

Harvard Mark I



- Built in 1944 in IBM Endicott laboratories
 - Howard Aiken – Professor of Physics at Harvard
 - Essentially mechanical but had some electro-magnetically controlled relays and gears
 - Weighed *5 tons* and had *750,000* components
 - A synchronizing clock that beat every *0.015* seconds (66KHz)

Performance:

0.3 seconds for addition

6 seconds for multiplication

1 minute for a sine calculation

WW-2 Effort

Broke down once a week!

ENIAC

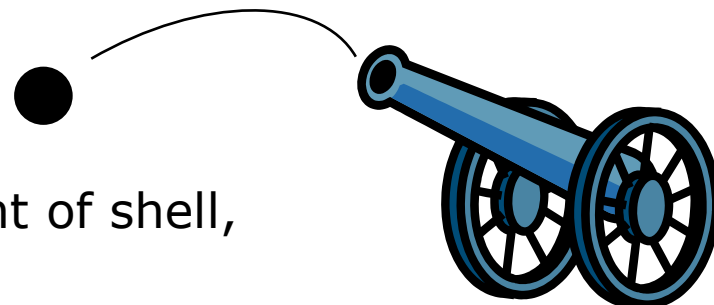


- Inspired by Atanasoff and Berry, Eckert and Mauchly designed and built ENIAC (1943-45) at the University of Pennsylvania
- The first, completely electronic, operational, general-purpose analytical calculator!
 - 30 tons, 72 square meters, 200KW
- Performance
 - Read in 120 cards per minute
 - Addition took 200 μ s, Division 6 ms
 - 1000 times faster than Mark I
- Not very reliable!

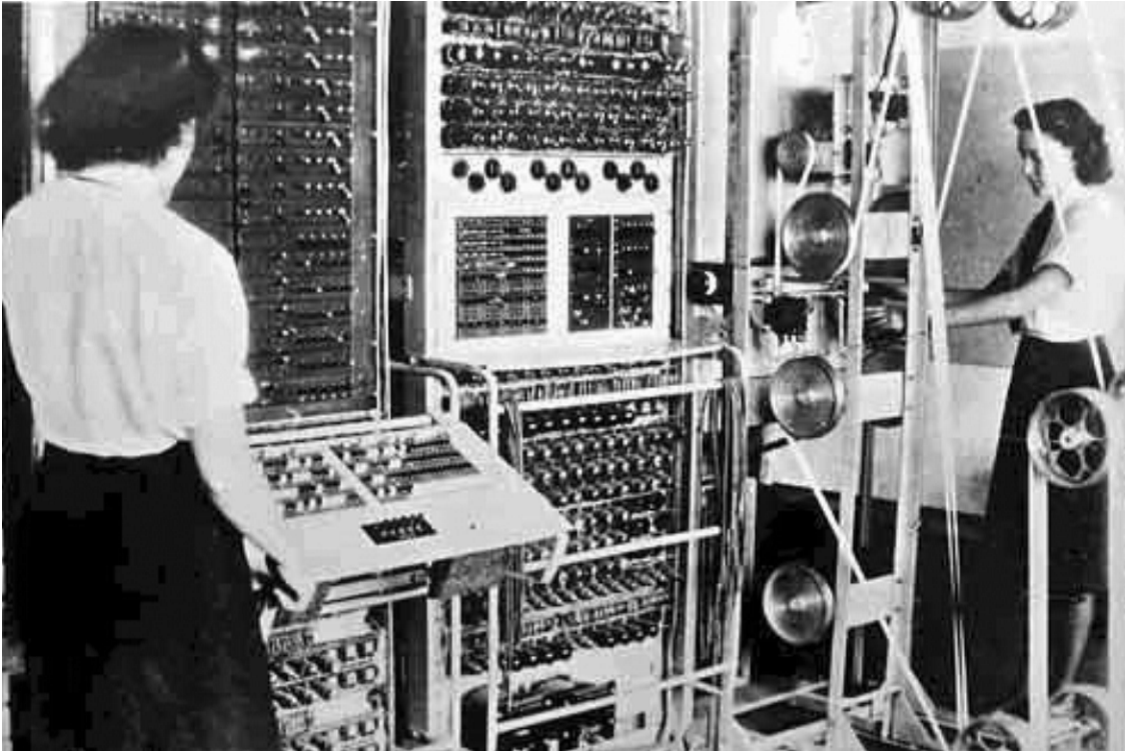
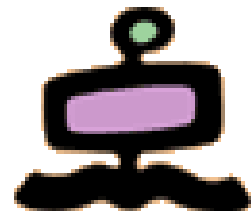
WW-2 Effort

Application: Ballistic calculations

angle = f (location, tail wind, cross wind,
air density, temperature, weight of shell,
propellant charge, ...)



Colossus



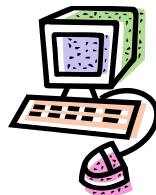
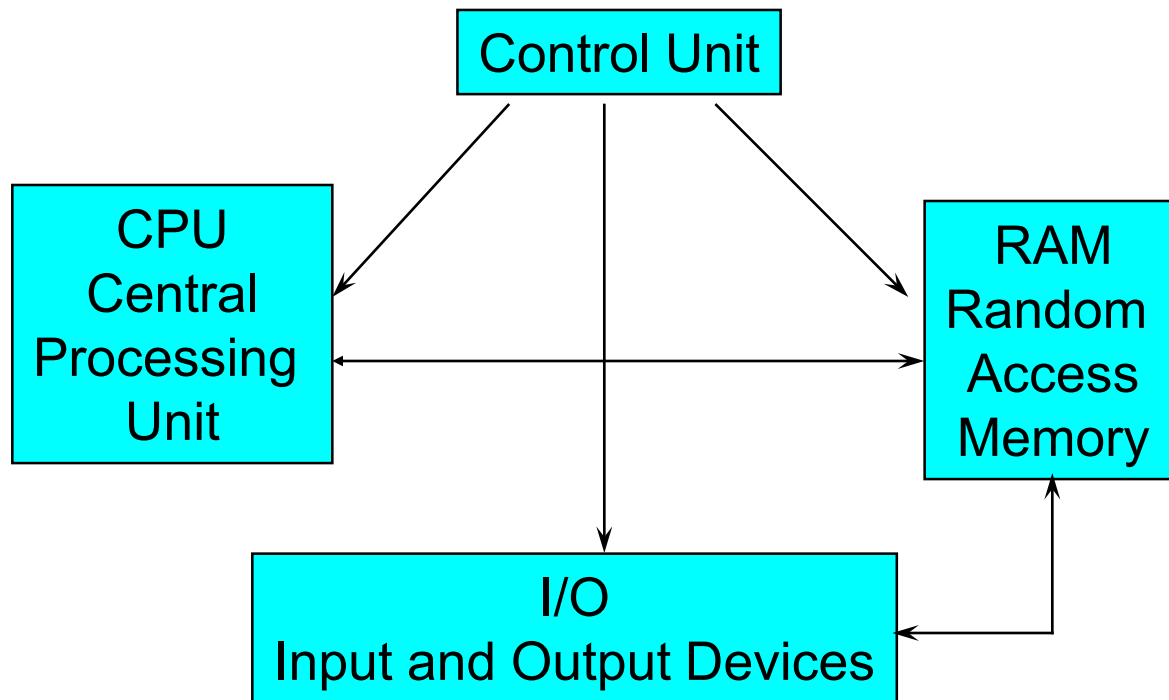
Colossus (derived from Mark 1 and Mark 2) was used in London during the second World War to decipher secret German messages (Enigma machine)

EDVAC - Electronic Discrete Variable Automatic Computer



- ENIAC's programming system was external
 - Sequences of instructions were executed independently of the results of the calculation
 - Human intervention required to take instructions “out of order”
- Eckert, Mauchly, John von Neumann and others designed EDVAC (1944) to solve this problem
 - Solution was the *stored program computer*
 - ⇒ “*program can be manipulated as data*”
- *First Draft of a report on EDVAC* was published in 1945, but just had von Neumann's signature
- In 1973 the court of Minneapolis attributed the honor of *inventing the computer* to John Atanasoff

Basic components of a computer



Von Neuman architecture

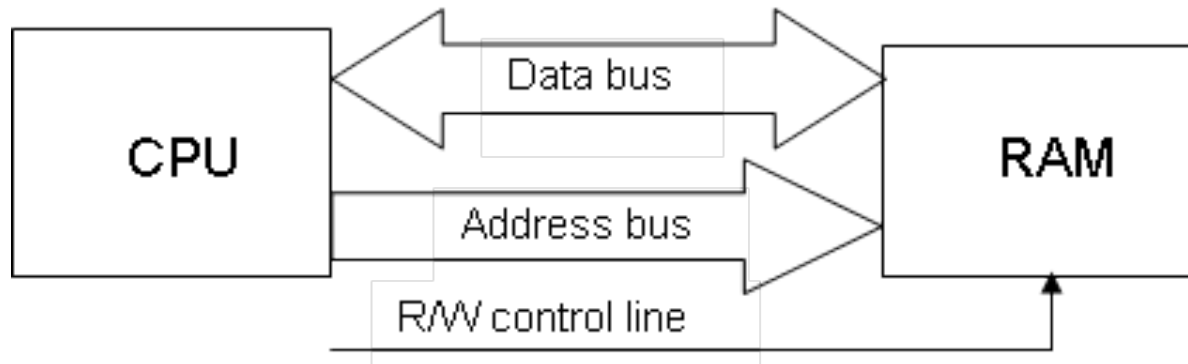


- The RAM contains both the program (machine instructions) and the data
- The basic model is “sequential execution”
 - each instruction is extracted from memory (in sequence) and executed
- Basic execution cycle
 - Fetch instruction (from memory) at location indicated by LC
 - Increment Location Counter (to point to the next instruction)
 - Bring instruction to CPU
 - Execute instruction
 - Fetch operand from memory (if needed)
 - Execute operation
 - Store result
 - in “registers” (temporary memory in the CPU)
 - in memory (RAM)

Random Access Memory



- The RAM is a linear array of “cells”, usually called “words”
- The words are numbered from 0 to N, and this number is the “address” of the word
- In order to read/write a word from/into a memory cell, the CPU has to provide its address on the “address bus”
- A “control line” tells the memory whether it is a read or write operation
- In a read operation the memory will provide on the “data bus” the content of the memory cell at the address provided on the “address bus”
- In a write operation the memory will store the data provided on the “data bus” into the memory cell at the address provided on the “address bus”

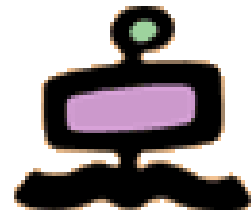


Data within a computer



- The Control Unit, the RAM, the CPU and all the physical components in a computer act on electrical signals and on devices that (basically) can be in only one of two possible states
- The two states are conventionally indicated as “zero” and “one” (0 and 1), and usually correspond to two voltage levels
- The consequence is that all the data within a computer (or in order to be processed by a computer) has to be represented with 0s and 1s, i.e. in “binary notation”

Evolution of computer technology



- First Generation
mechanical/electromechanical
- Second Generation
vacuum tubes
- Third Generation
discrete transistors (solid state devices)
SSI, MSI, LSI integrated circuits
- Fourth Generation
VLSI integrated circuits

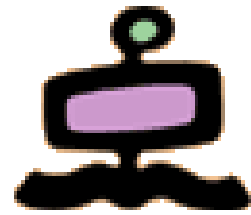
VLSI = Very Large Scale Integration

Evolution of computer components



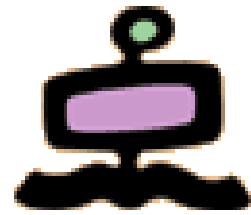
- Computer technology
 - CPU on integrated chips
 - From KHz to MHz to GHz
 - Random Access Memories
 - RAM – from KB to MB to GB
 - External memories
 - Tapes, hard disks, floppy disks
 - Memory sticks
 - CDs
 - DVDs
 - from MB to GB to TB to PB to EB

Size of digital information



1000	k	kilo
1000 ²	M	mega
1000 ³	G	giga
1000 ⁴	T	tera
1000 ⁵	P	peta
1000 ⁶	E	exa
1000 ⁷	Z	zetta
1000 ⁸	Y	yotta

Evolution of the software



- Operating systems
 - Multi user
 - Multi tasking
- Applications
 - Client-Server
 - Multimedia
- Communication

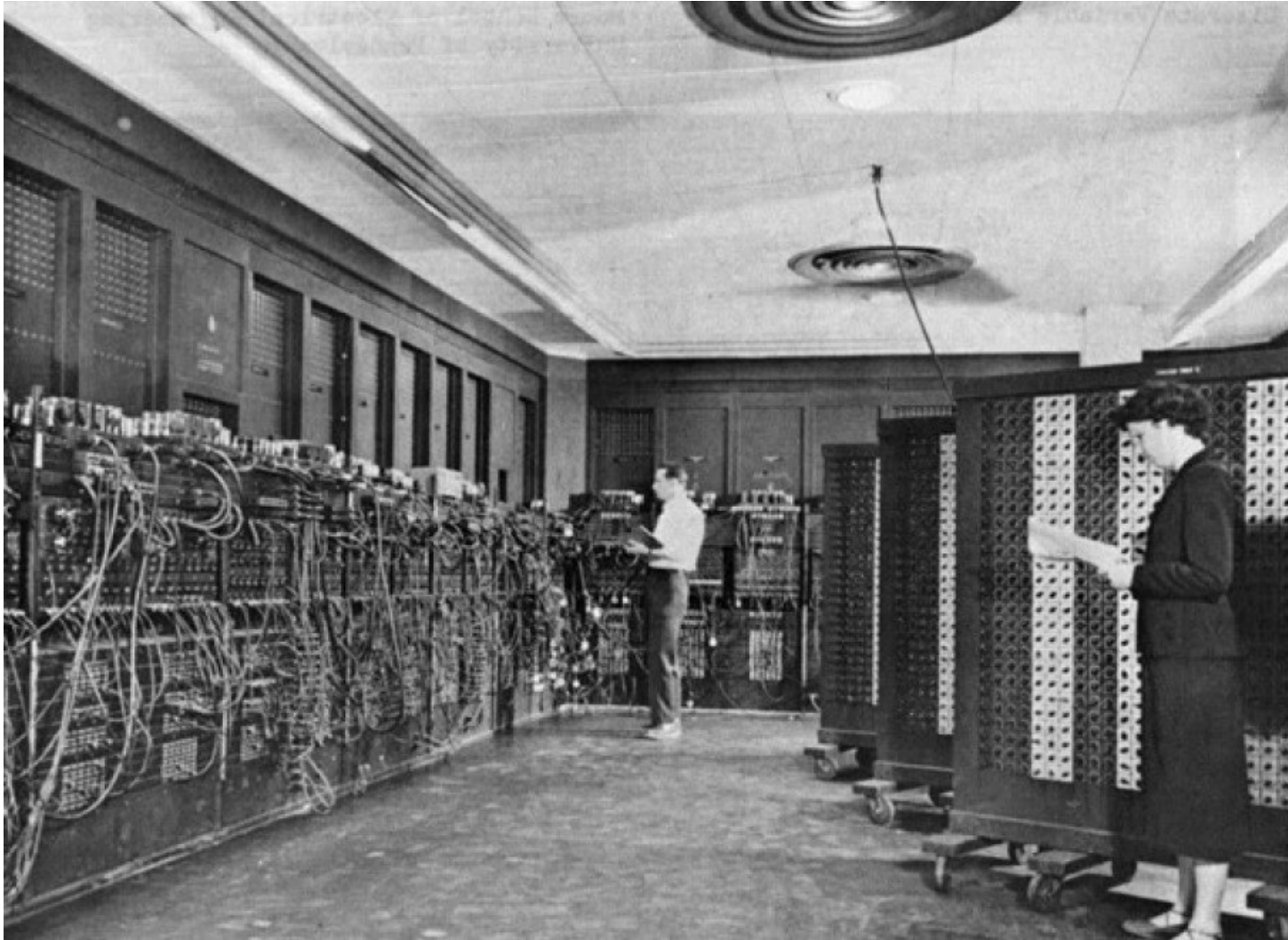
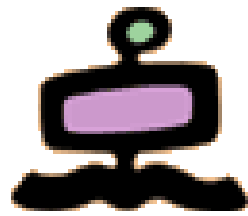
Evolution of computer market



- Military applications in early 40s
- Scientific/research applications in late 40s
- Commercial applications appear in early 50s
- Monopoly of IBM starts with 650, 701, 702
- Monopoly of IBM continues with 7070, 7090 and the 360 series, starting the “mainframe era” (in the 60s)
- Arrival of the “minicomputers” in the 70s
- Arrival of the PC in the 80s
- Arrival of the Internet in the 90s
- Arrival of the Web in the 90s

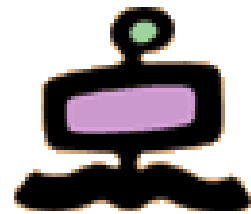


ENIAC - Electronic Numerical Integrator And Computer

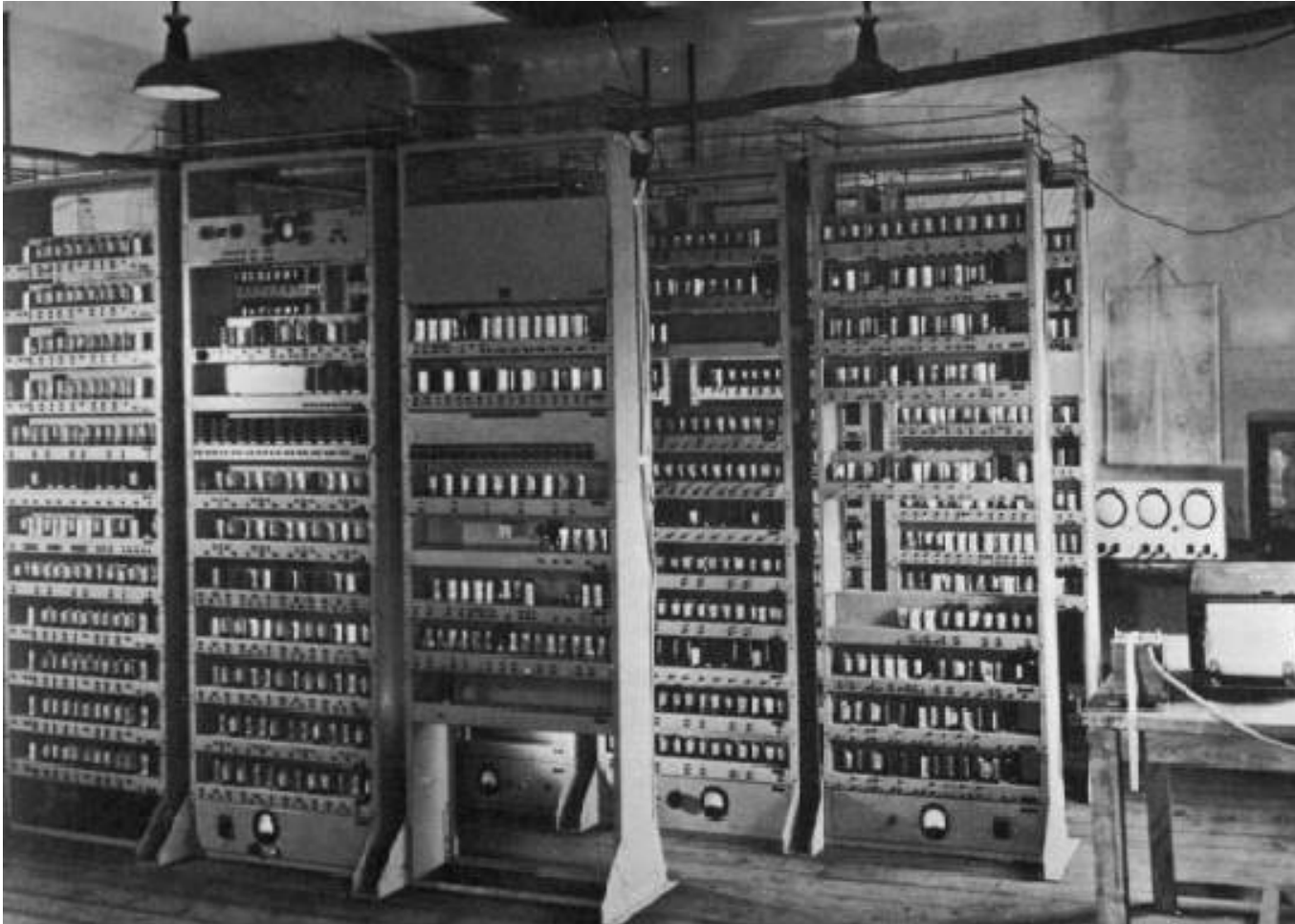




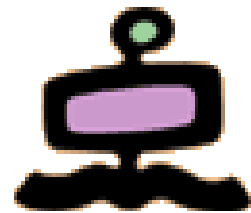
EDSAC - Electronic Delay Storage Automatic Calculator



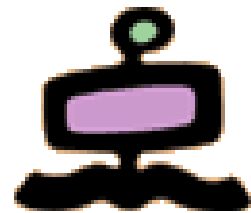
EDSAC, University of Cambridge, UK, 1949



A “mainframe” in the 60’

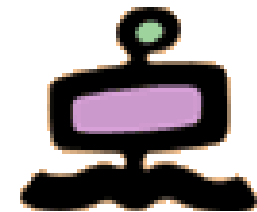


A “mainframe” in the 70’



Photograph: Dominic Hart/NASA Ames

Minicomputers



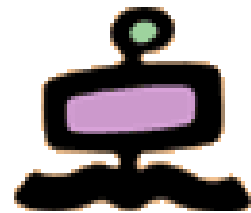
UNIPI BDG 24-25



Vittore Casarosa – Biblioteche Digitali

Refresh Computers - 29

Early PCs



Evolution of technology



- Computer technology
 - CPU and integrated chips
 - Random Access Memories
 - RAM – from KB to GB
 - External memories
 - Tapes, hard disks, floppy disks
 - Memory sticks
 - CDs
 - DVDs
 - from MB to GB to TB to PB to EB
- Communication technology (networks)
 - (Telephone) line speed
 - Point to point (leased lines)
 - Local Area Networks
 - Inter-networking (TCP/IP)