Informatics
luis rocha 2021

complex networks and systems

i609: advanced PhD seminar I

luis m. rocha

Luddy school of informatics, computing, & engineering

indiana university, bloomington, usa

and

instituto gulbenkian de ciência

oeiras, portugal

INDIANA UNIVERSITY

Indiana University **Network Science Institute**

Cognitive Science Program

INSTITUTO GULBENKIAN DE CIÊNCIA

rocha@indiana.edu
http://informatics.indiana.edu/rocha

what about you?

- Background
- Interests
- Course expectations

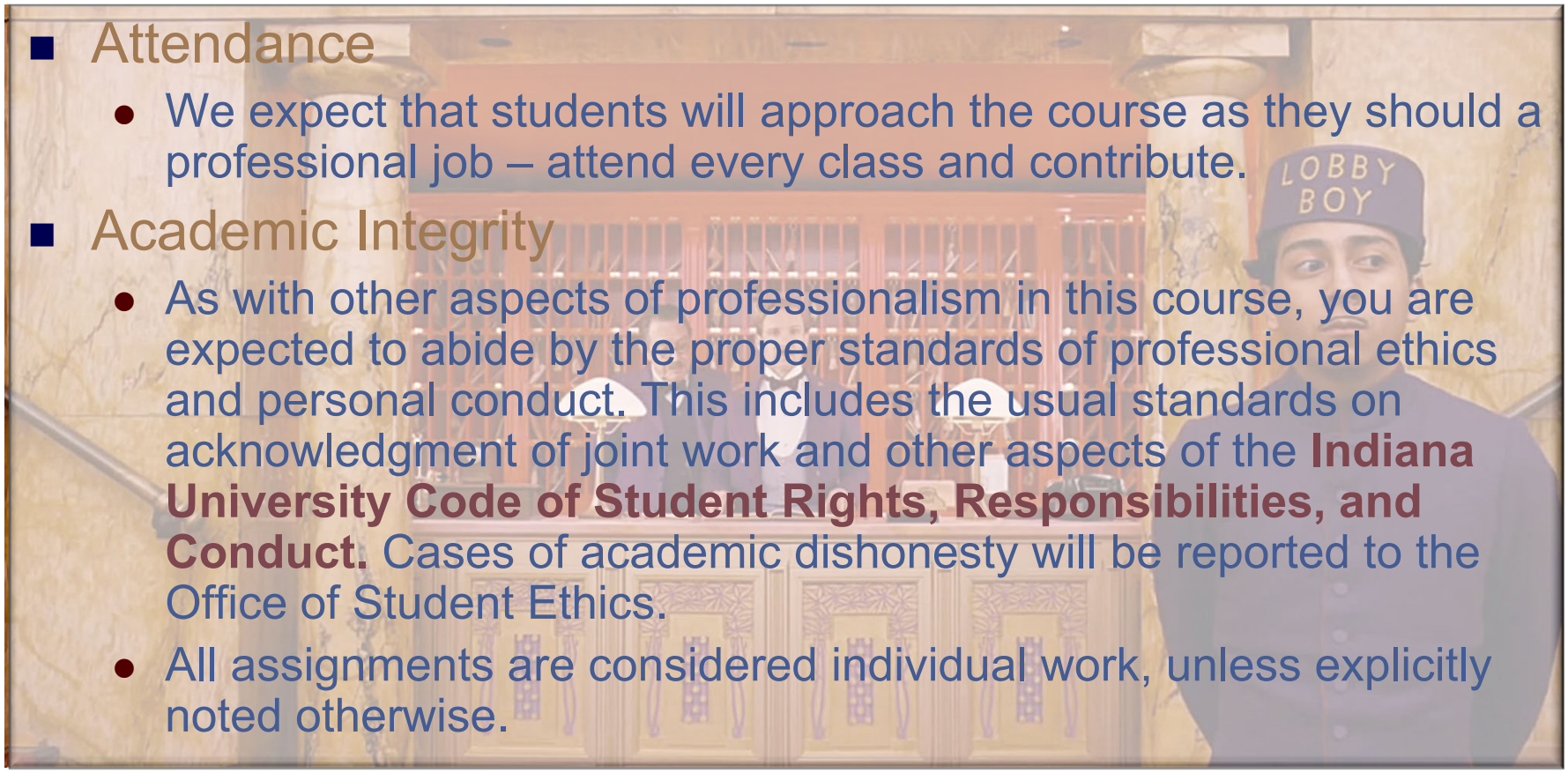
rocha@indiana.edu
http://informatics.indiana.edu/rocha

evaluation

- Participation and Discussion: 15%.
  - class discussion, everybody reads and discusses every paper
  - engagement in class
- Lead Discussions: 25%
  - Students are assigned to papers as lead discussants
    - all students are supposed to read and participate in discussion of every paper.
  - Lead discussant prepares short summary of assigned paper (10 minutes)
    - no formal presentations or PowerPoint unless figures are indispensable.
  - Summary should:
    - 1) Identify the key goals of the paper (not go in detail over every section)
    - 2) What discussant liked and did not like
    - 3) What authors achieved and did not
    - 4) Any other relevant connections to other class readings and beyond.
  - Class discussion is opened to all
    - lead discussant ensures we important paper contributions and failures are adressed
- Term Paper/Project: 60%
  - Outline agreement with instructor
    - **Due March 4**
    - Upload to Canvas
  - Final Paper/Project
    - **Due May 4**
    - Upload to Canvas

rocha@indiana.edu
http://informatics.indiana.edu/rocha

rules, rules, rules



- Attendance
  - We expect that students will approach the course as they should a professional job – attend every class and contribute.
- Academic Integrity
  - As with other aspects of professionalism in this course, you are expected to abide by the proper standards of professional ethics and personal conduct. This includes the usual standards on acknowledgment of joint work and other aspects of the **Indiana University Code of Student Rights, Responsibilities, and Conduct.** Cases of academic dishonesty will be reported to the Office of Student Ethics.
  - All assignments are considered individual work, unless explicitly noted otherwise.

rocha@indiana.edu
http://informatics.indiana.edu/rocha

even worse rules, rules, rules

- Incomplete Grade
  - only by prior arrangement in exceptional circumstances conforming to university and departmental policy which requires, among other things, that the student must have completed the bulk of the work required for the course with a passing grade, and that the remaining work can be made up within 30 days after the end of the semester.
- Technology in class
  - No laptops, tablets, or mobile phone
    - except of dedicated ebook readers and paper
- Deadlines
  - no late work, with the exception of force majeure.
    - Start early and consult with your professors on a timely basis if any issues emerge.
- Readings
  - All expected to read required list of readings and participate

rocha@indiana.edu
http://informatics.indiana.edu/rocha

GENES, COMPUTERS, AND CYBERNETICS

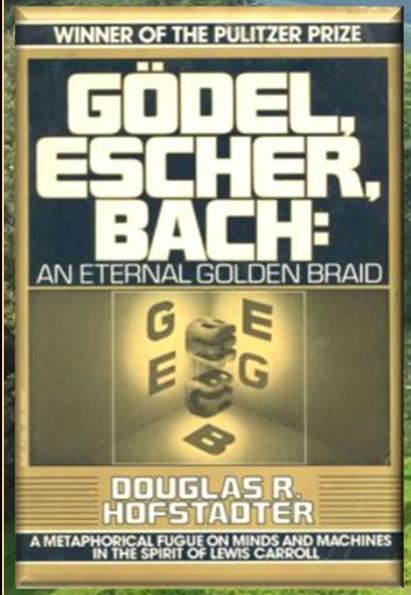
A TOUR ON THE GARDEN OF FORKING PATHS

INDIANA UNIVERSITY

rocha@indiana.edu
http://informatics.indiana.edu/rocha

the library of Babel

Personal path in the garden of forking paths



Poetic/metaphorical essays on Information, memory, meaning, collective intelligence (1941. 1979)

the library of Babel

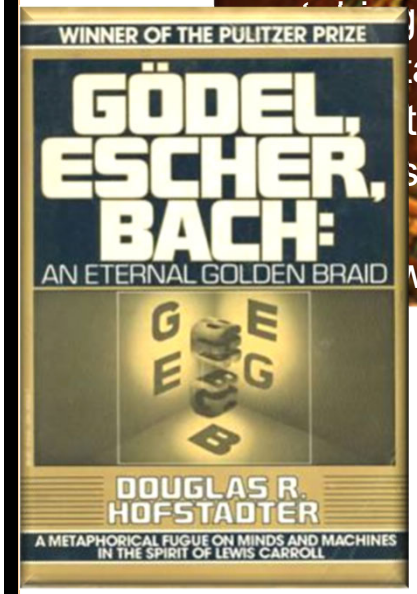
Jorge Luis Borges (1899 – 1986)

"The universe (which others call the Library) is composed of an indefinite and perhaps infinite number of hexagonal galleries, with vast air shafts between, surrounded by very low railings."

"……all the books, no matter how diverse they might be, are made up of the same elements: the space, the period, the comma, the twenty-two letters of the alphabet He also alleged a fact which travelers have confirmed: In the vast Library there are no two identical books."

"…Everything: the minutely detailed history of the future, the archangels' graphies, the faithful catalogues of the Library, thousands and thousands of alogues, the demonstration of the fallacy of those catalogues, the tration of the fallacy of the true catalogue,[…] the true story of your death, slation of every book in all languages…".

vandered in search of a book, perhaps the catalogue of catalogues"



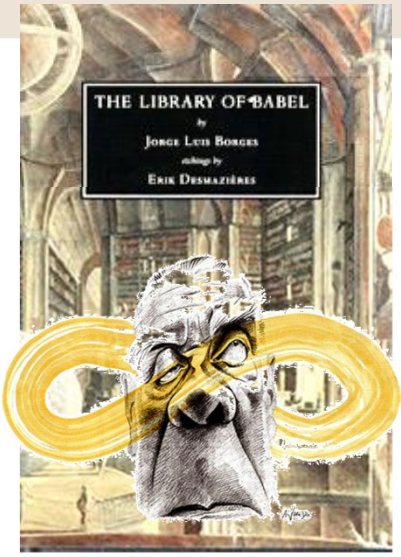
Poetic essays on information and memory (1941)

INDIANA UNIVERSITY

the library of Babel

numbers

- Each book
  - 25 characters, written in any sequence for 410 pages of 40 lines of 80 characters
  - 410*40*80 = sequence of ≈ $10^{6.1}$ characters
    - ≈ $10^{7.2}$ base pairs (10 Mbp)
    - ≈ 1 book to store E.Coli genotype, 10 for drosophila, and 100 for human
- How many possible books?
  - = $25^{(410*40*80)}$ combinations = $25^{1,312,000}$ books!
    - ≈ 1.956 x $\mathbf{10^{1,834,097}}$ books
  - Total number of atoms in the current, observable universe is about 10^{80}
    - If each book were the size of an atom, library would hold $10^{1,834,017}$ **universes**!
  - Yet finite!
    - Can also be reproduced with just two symbols (cf Quine, Turing, Leibniz)



Information Space Is finite but larger than Physical space

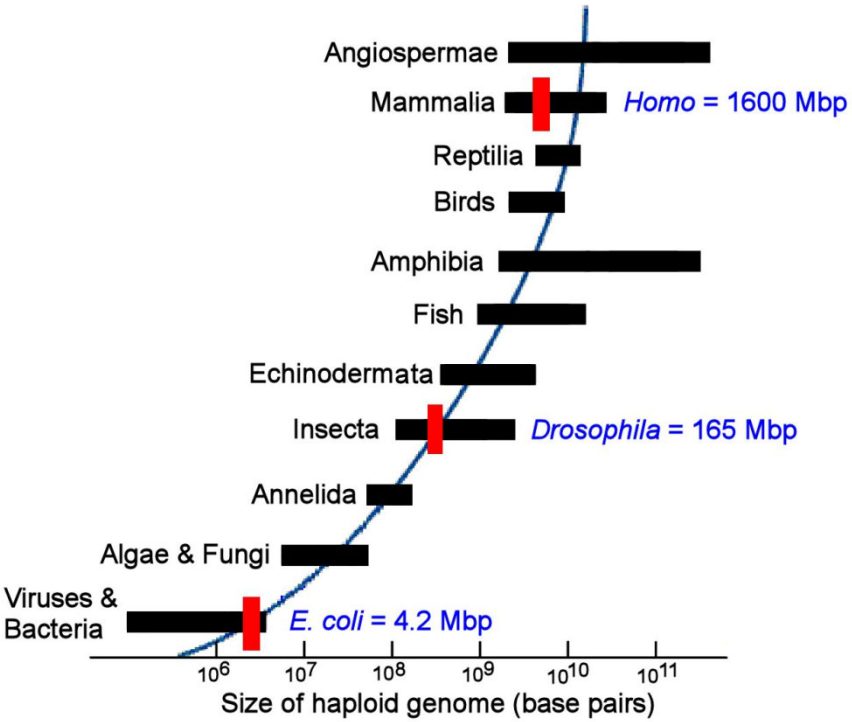
"the Library is so enormous that any reduction of human origin is infinitesimal." "every copy is unique, irreplaceable, but (since the Library is total) there are always several hundred thousand imperfect facsimiles: works which differ only in a letter or a comma."

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

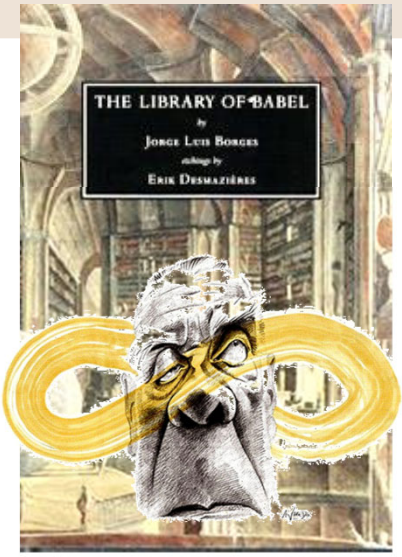
the library of Babel

numbers

- Each book
  - 25 characters, written in any sequence for 410 pages of 40 lines of 80 characters
  - 410*40*80 = sequence of ≈ $10^{6.1}$ characters
    - ≈ $10^{7.2}$ base pairs (10 Mbp)
    - ≈ 1 book to store E.Coli genotype, 10 for drosophila, and 100 for human



= $25^{1,312,000}$ books!
oks
he current, observable
size of an atom, library
niverses!
ed with just two symbols
niz)



Information Space Is finite but larger than Physical space

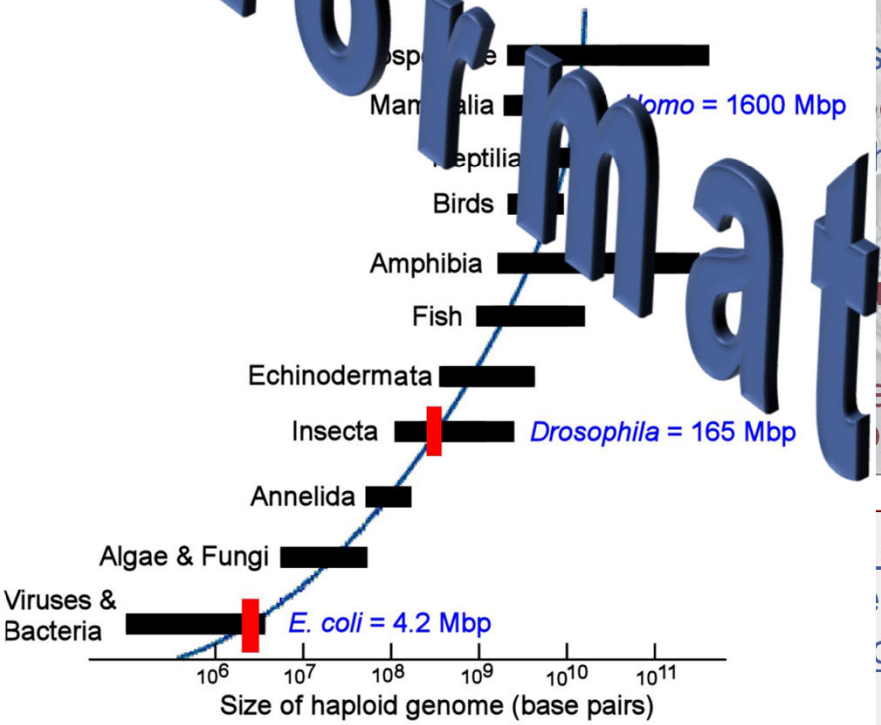
of human origin is infinitesimal."
the Library is total) there are
csimiles: works which differ only

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

the library of Babel

numbers

- Each book
  - 25 characters, written in any sequence for 410 pages of 40 lines of 80 characters
  - 410*40*80 = sequence of ≈ $10^{6.1}$ characters
    - ≈ $10^{7.2}$ base pairs (10 Mbp)
    - ≈ 1 book to store E.Coli genotype, 10 for drosophila, and 100 for human



information

s = $25^{1,312,000}$ books!

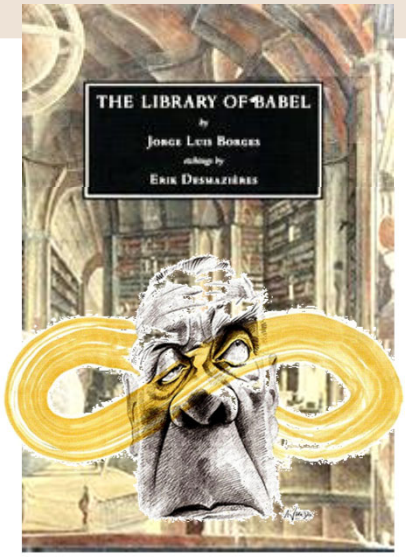
oks

he current, observable

siz of an atom, library
n erses!

ed wit t two symbols
n

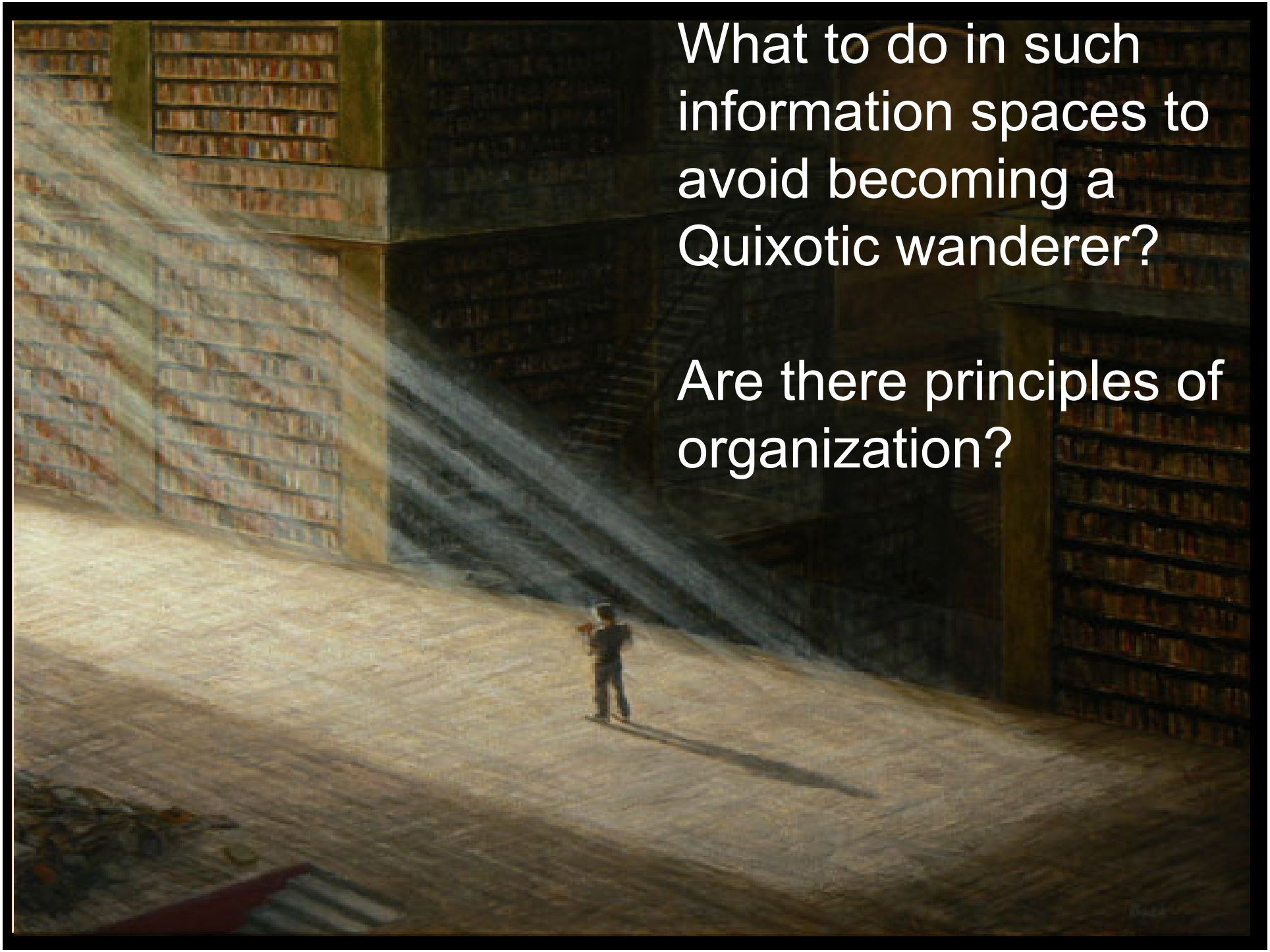
of hu or s infinitesimal."
the Lib ary tal) there are
csimiles: wo s which differ only

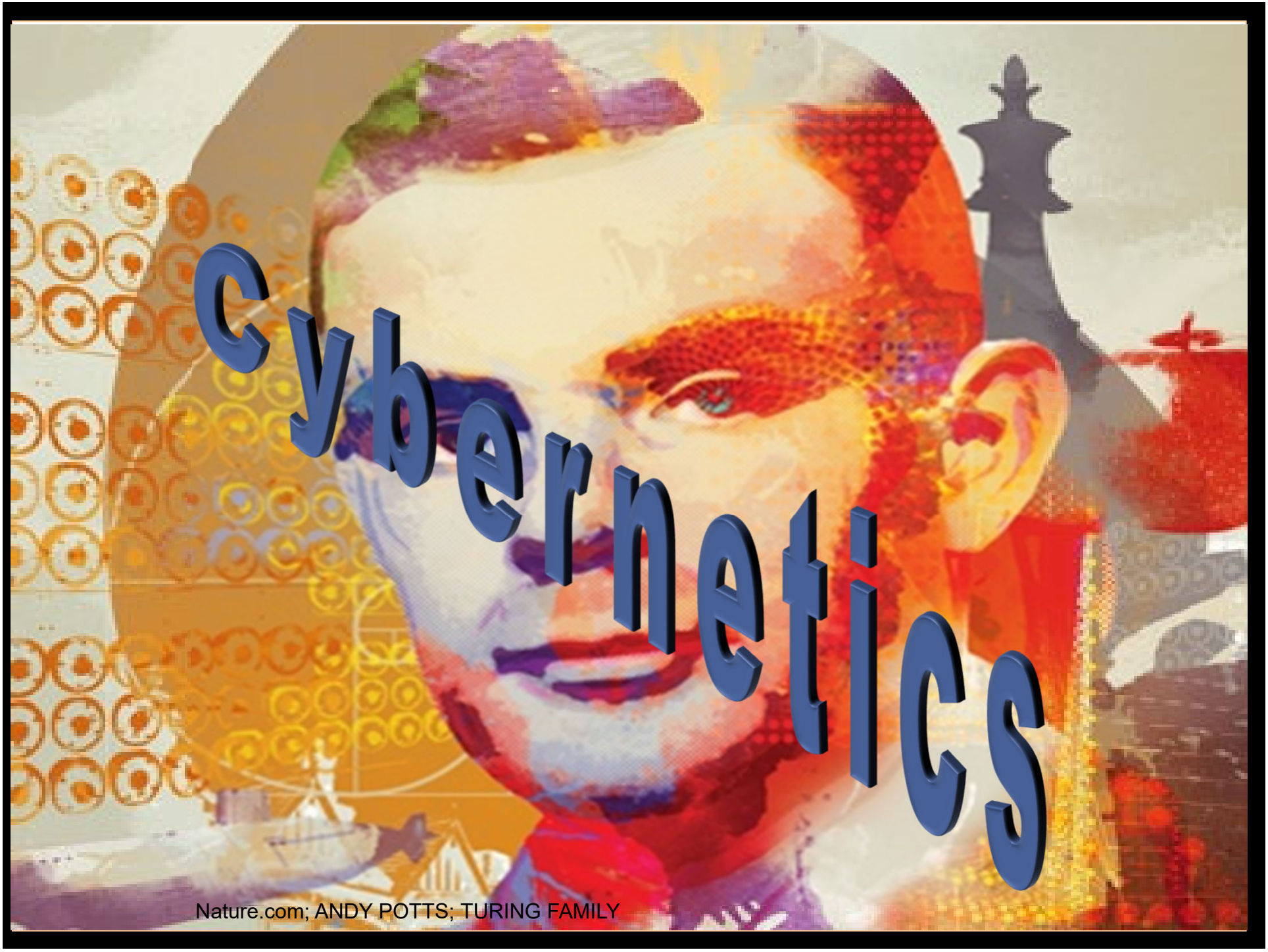


Information Space Is finite but larger than Physical space

INDIANA UNIVERSITY

rocha@indiana.edu
informatics.indiana.edu/rocha


What to do in such information spaces to avoid becoming a Quixotic wanderer?
Are there principles of organization?


Cybernetics
Nature.com; ANDY POTTS; TURING FAMILY

TURING'S TAPE AND COMPUTERS

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

Alan Turing (1912-1954)

key contributions (most relevant to biocomplexity)

- "The chemical basis of morphogenesis"
  - Turing, A. M. *Phil. Trans. R. Soc. Lond. B* **237**, 37–72 (1952).
    - Reaction-diffusion systems
- "Computing machinery and intelligence"
  - Turing, A. M. *Mind* **49**, 433–460 (1950).
    - The "Turing Test"
- "On computable numbers with an application to the *Entscheidungsproblem*"
  - Turing, A. M. *Proc. Lond. Math. Soc.* **s2–42**, 230–265 (1936–37).
    - Turing machine, universal computation, decision problem

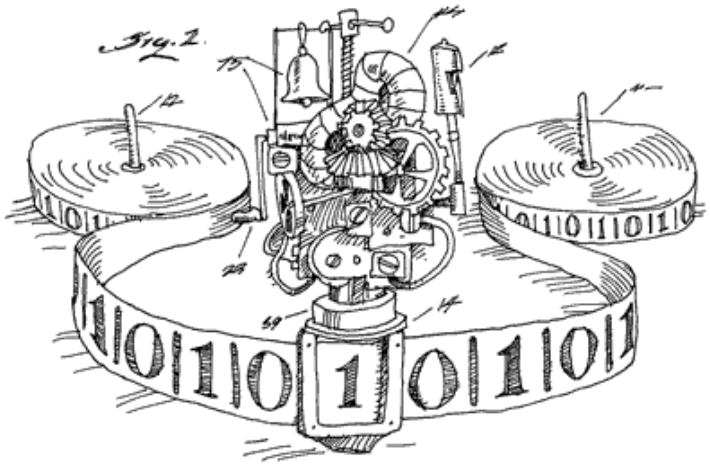
Brenner, Sydney. [2012]. "Life's code script." *Nature* **482** (7386): 461-461.

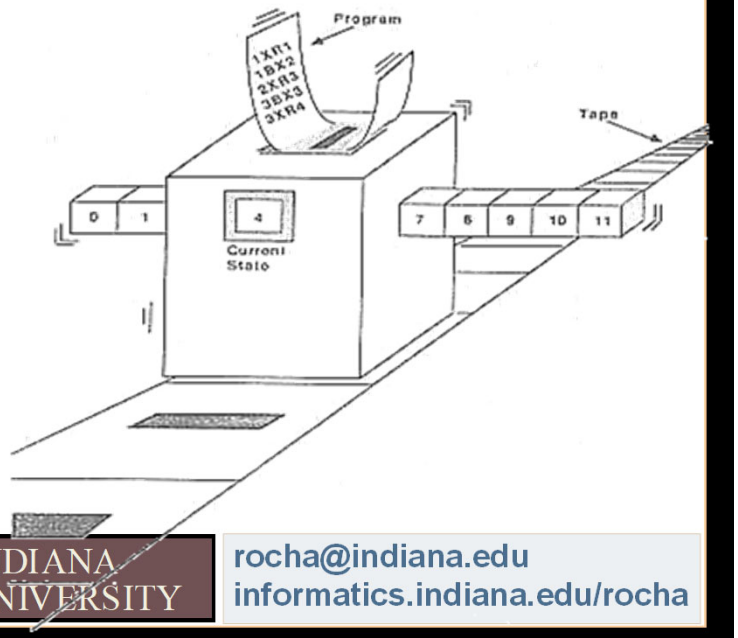
INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

Turing's tape

A fundamental principle of computation

- "On computable numbers with an application to the *Entscheidungsproblem*"
  - Turing, A. M. *Proc. Lond. Math. Soc.* **s2–42**, 230–265 (1936–37).
    - **Turing machine**, universal computation, decision problem
  - **Machine**'s *state* is *controlled* by a *program*, while *data* for program is on limitless **external tape**
    - every machine can be described as a **number** that can be stored on the tape (for itself or another machine)
      - Including a Universal machine
    - **distinction** between *numbers that mean things* (data) and *numbers that do things* (program)



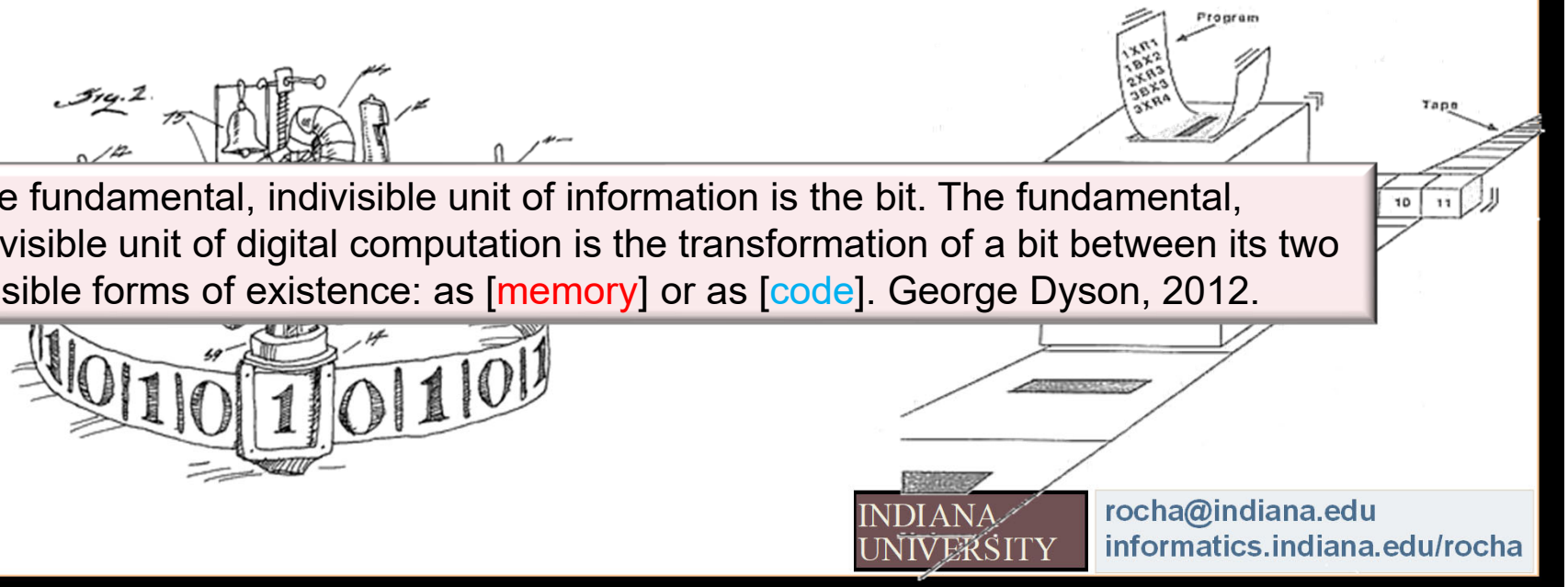


INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

Turing's tape

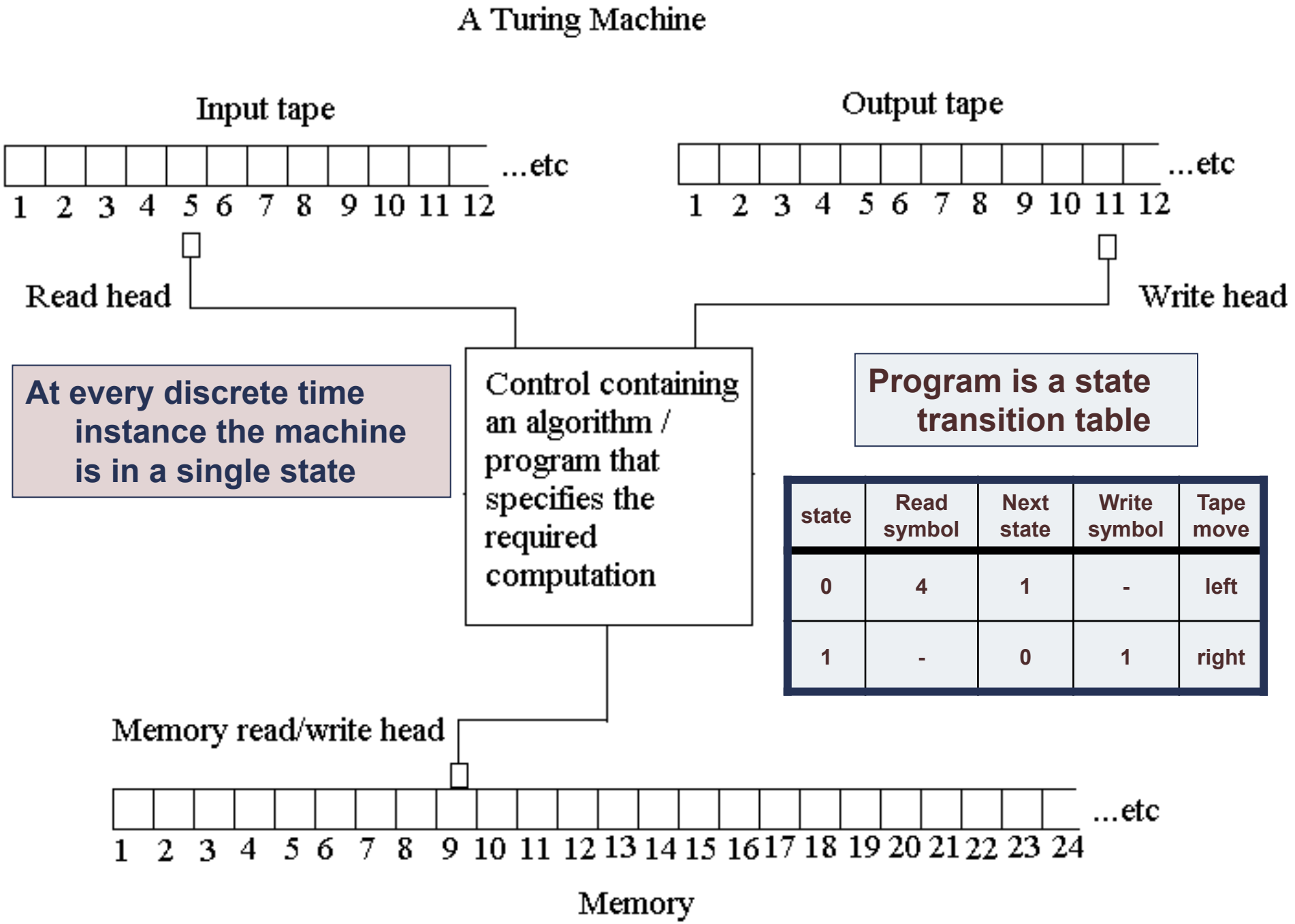
A fundamental principle of computation

- "On computable numbers with an application to the *Entscheidungsproblem*"
  - Turing, A. M. *Proc. Lond. Math. Soc.* **s2–42**, 230–265 (1936–37).
    - **Turing machine**, universal computation, decision problem
  - **Machine**'s *state* is *controlled* by a *program*, while *data* for program is on limitless **external tape**
    - every machine can be described as a **number** that can be stored on the tape (for itself or another machine)
      - Including a Universal machine
    - **distinction** between *numbers that mean things* (data) and *numbers that do things* (program)



"The fundamental, indivisible unit of information is the bit. The fundamental, indivisible unit of digital computation is the transformation of a bit between its two possible forms of existence: as [memory] or as [code]. George Dyson, 2012.

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha



state	Read symbol	Next state	Write symbol	Tape move
0	4	1	-	left
1	-	0	1	right

A Turing Machine

Input tape

1 2 3 4 5 6 7 8 9 10 ...etc

Read head

10 11 12

Write head

At every discrete time instance the machine is in a single state

a state
n table

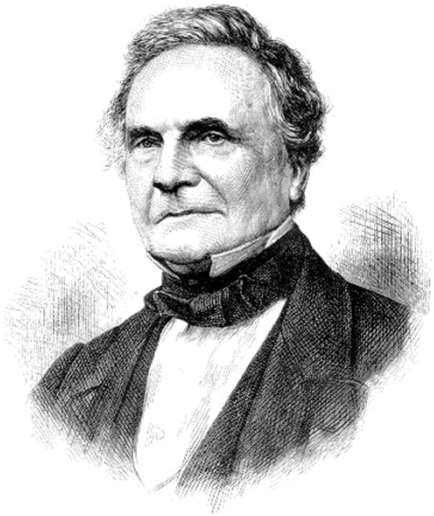
	Write symbol	Tape move
	-	left
	1	right

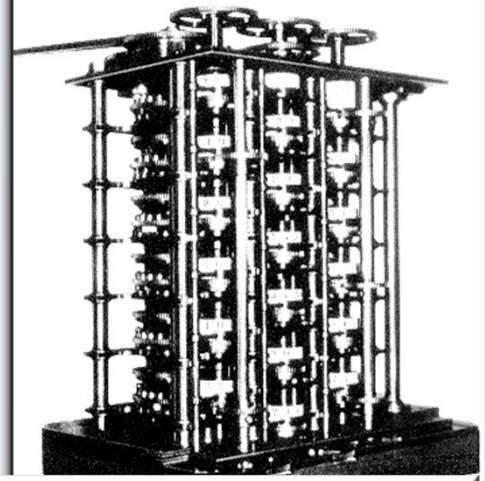
Memory read/write h

1 2 3 4 5 6 7 ... 21 22 23 24 ...etc

Charles Babbage (1791 – 1871)

- Difference Engine
  - Special-purpose digital computing machine for the **automatic** production of mathematical tables.
    - logarithm tables, tide tables, and astronomical tables
    - Steam-driven, consisted entirely of mechanical components - brass gear wheels, rods, ratchets, pinions, etc.
    - Numbers were represented in the decimal system by the positions of 10-toothed metal wheels mounted in columns.
  - Never completed the full-scale machine
    - Completed several fragments. The largest is on display in the London Science Museum. In 1990, it was built (London Science Museum)
  - The Swedes Georg and Edvard Scheutz (father and son) constructed a modified version of Babbage's Difference Engine.
  - For an interesting “what-if” scenario read “The Difference Engine” by Bruce Sterling and William Gibson

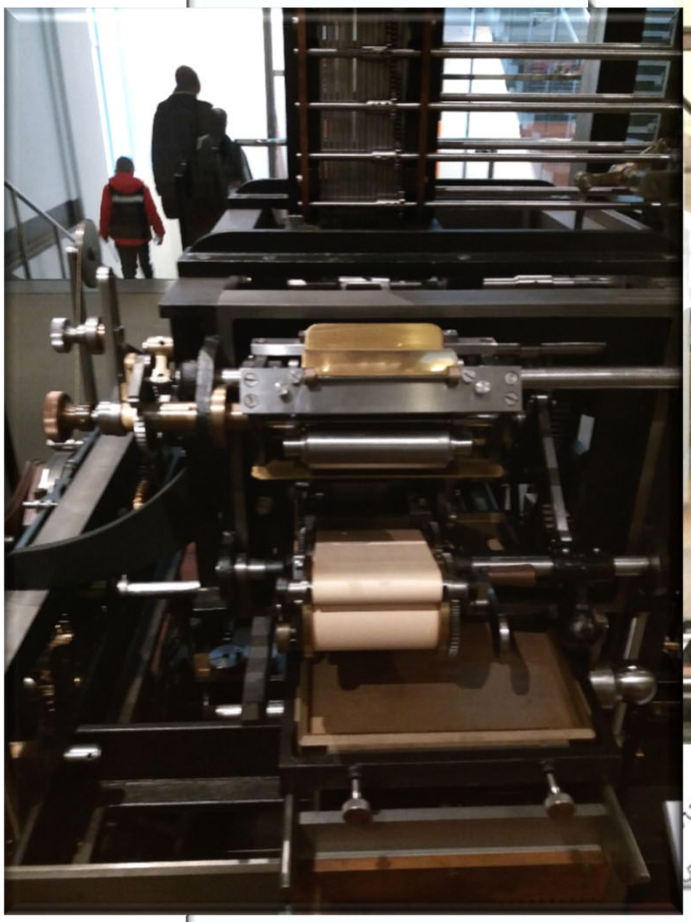


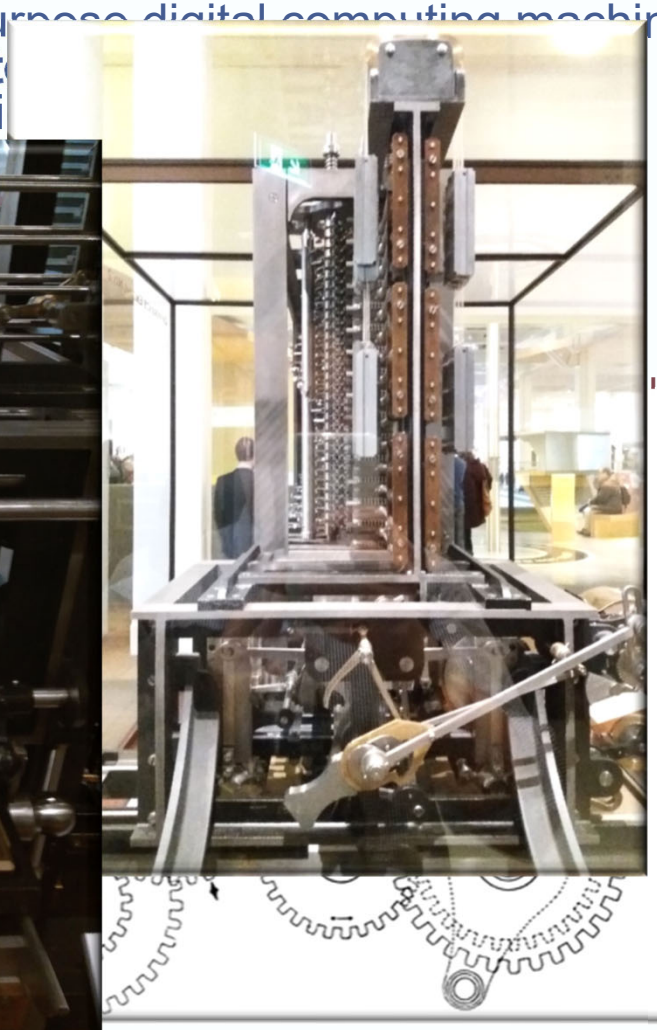


Not a universal Turing machine, but an analog computer

Charles Babbage (1791 – 1871)

- Difference Engine
  - Special-pu… digital computing machine… for the **aut**… mathemati…





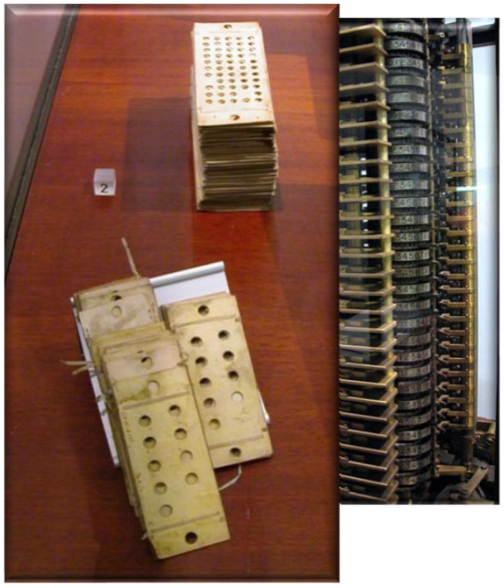
Not a universal Turing machine, but an analog computer

Charles Babbage (1791 – 1871) and Ada Lovelace (1815-1852)

The analytical engine had an “external tape”

Turing on programs (numbers as instructions) : “[Babbage] had all the essential ideas [and] planned such a machine, called the ***Analytical Engine***. […]

- general-purpose mechanical digital computer.
  - Separated **memory store** from a **central processing unit** (or ‘mill’)
  - able to select from among **alternative actions** consequent upon the outcome of its previous actions
    - Conditional branching: Choice, information
  - Mechanical cogs not just numbers
    - **Variables** (states/configurations)
- Programmable
  - Data and instructions on distinct **punched cards**



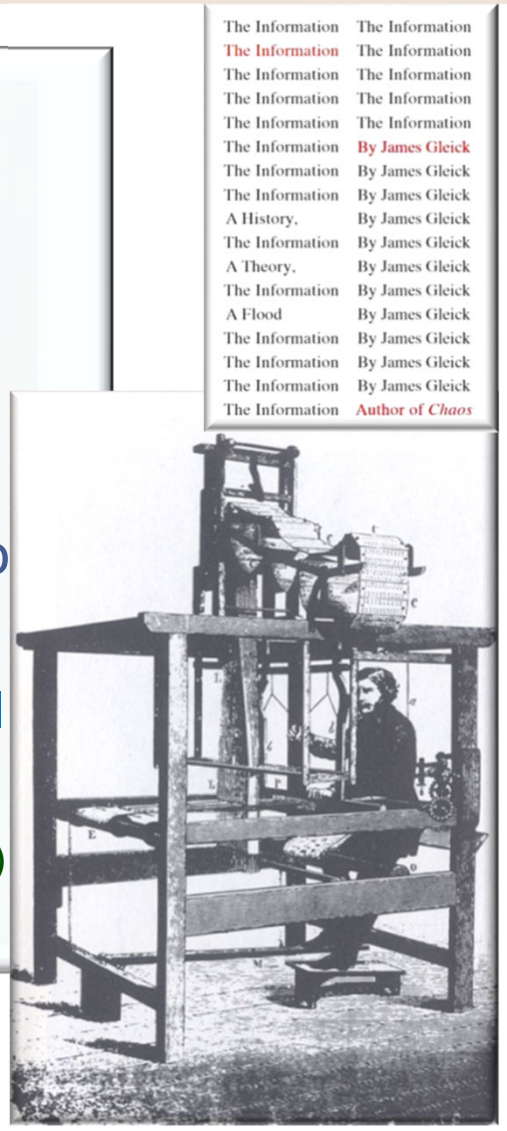
"It is only a question of cards and time, […] and there is no reason why (twenty thousand) cards should not be used if necessary, in an Analytical Engine for the purposes of the mathematician”. Henry Babbage (1888)

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

Charles Babbage (1791 – 1871) and Ada Lovelace (1815-1852)

The external tape as a general design principle (system) of universal computation

- Analytical engine
  - Separated **memory store** from a **central processing unit** (or 'mill')
  - Cogs not just numbers
    - **variables**
- Programmable
  - instructions on **punched cards**
    - Inspired by the Jacquard Loom
  - Ada Lovelace: the science of operatio
    - Set of (recursive) rules for producing Bernoulli numbers (a program)
    - **Separation** of variable and operational (data) cards
    - would punch out cards for later use
      - "the Engine eating its own tail." (Babbage)



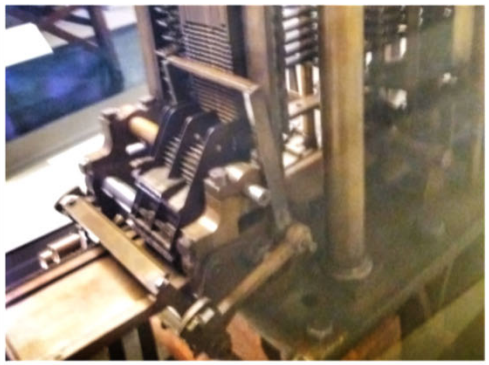
distinction between *numbers that mean things* and *numbers that do things*.

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

Charles Babbage (1791 – 1871) and Ada Lovelace (1815-1852)

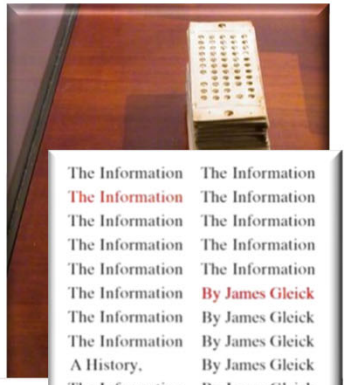
The external tape as a general design principle (system) of universal computation

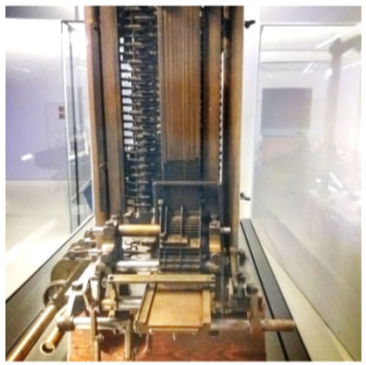
- A
- P



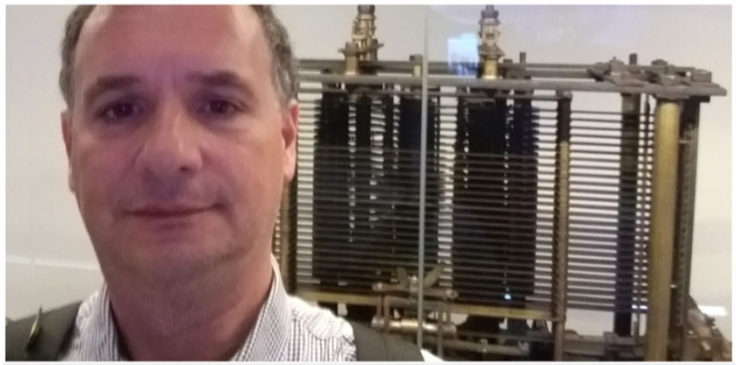
'We are not aware that anything in the nature of the Analytical Engine has been hitherto proposed, or even thought of, as a practical possibility, any more than the idea of a thinking or of a reasoning machine.'

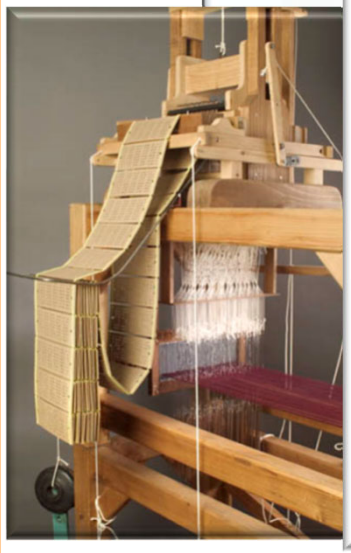
Ada Lovelace, mathematician, 1843





Analytical engine trial model
1834–71





The Information
The Information By James Gleick
A History, A Theory, A Flood
Author of *Chaos*

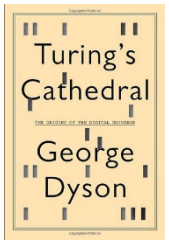
distinction between *numbers that mean things* and *numbers that do things*.

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

From mathematical generality to physical implementation constraints

- Process of rewriting strings in a formal system according to a program of rules
  - Operations and states are syntactic
  - Symbols follow syntactical rules
  - Rate of computation is irrelevant
    - Program determines result, not speed of machine
  - Physical implementation is irrelevant for result
- Computer
  - **Physical device** that can reliably execute/approximate a formal computation
    - Errors always exist
    - Design aims to make **rate** and **dynamics** irrelevant

"[…] essential elements in the machine are of a binary […] nature. Those whose state is determined by their history and are time-stable are memory elements. Elements of which the state is determined essentially by the existing amplitude of a voltage or signal are called 'gates'". Bigelow et al, 1947



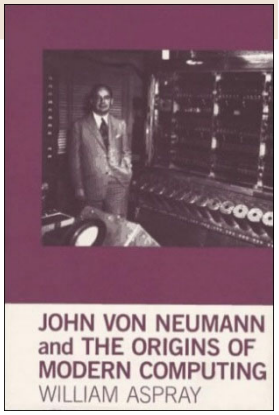
INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

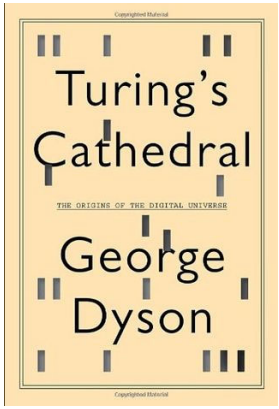
John Von Neumann (1903-1957)

Turing machines beyond the decision problem

" 'Words' coding the orders are handled in the memory just like numbers" --- distinction between *numbers that mean things* and *numbers that do things*.

- realizing the power of Turing's tape
  - physical (electronic) computers
  - emphasized the importance of the ***stored-program concept*** (the external tape)
    - EDVAC
  - allows machine to modify its own program
    - von Neumann architecture: The functional **separation** of storage from the processing unit.
      - programs can exist as data (two roles)
    - Converts tape to fixed-address memory (random-access memory)
  - Ultimate ***general-purpose*** machines





"Let the whole outside world consist of a long paper tape".
—John von Neumann, 1948

INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

John Von Neumann (1903-1957)

Turing machines beyond the decision problem

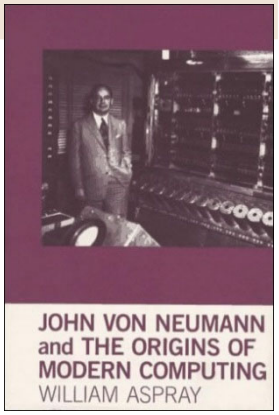
" 'Words' coding the orders are handled in the memory just like numbers" --- distinction between *numbers that mean things* and *numbers that do things*.

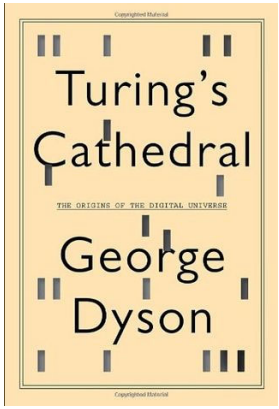
- realizing the power of Turing's tape
  - physical (electronic) computers
  - emphasized the importance of the ***stored-program concept*** (the external tape)
    - EDVAC

"Since Babbage's machine was not electrical, and since all digital computers are in a sense equivalent, we see that this use of electricity cannot be of theoretical importance.… The feature of using electricity is thus seen to be only a very superficial similarity." (Alan Turing)

    - programs can exist as data (two roles)
    - Converts tape to fixed-address memory (random-access memory)
  - Ultimate ***general-purpose*** machines

"Let the whole outside world consist of a long paper tape".
—John von Neumann, 1948





INDIANA UNIVERSITY
rocha@indiana.edu
informatics.indiana.edu/rocha

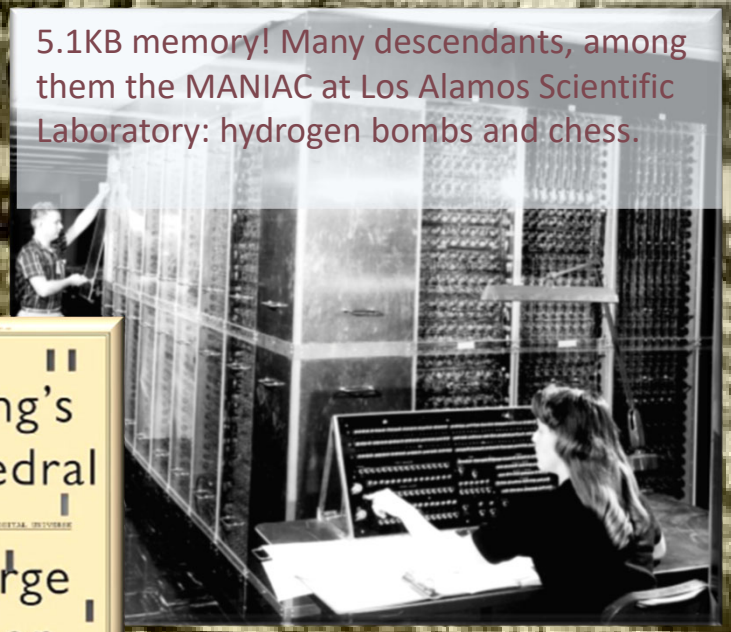
IAS Machine (1952)

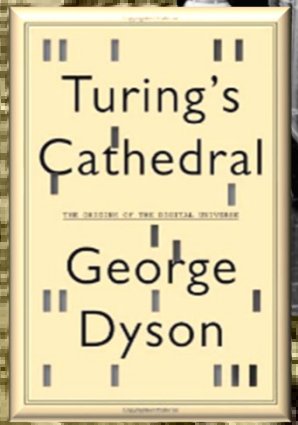
First electronic digital computer with 40 bit word (IAS, Princeton)

http://www.cs.uakron.edu/~margush/465/01_intro.html

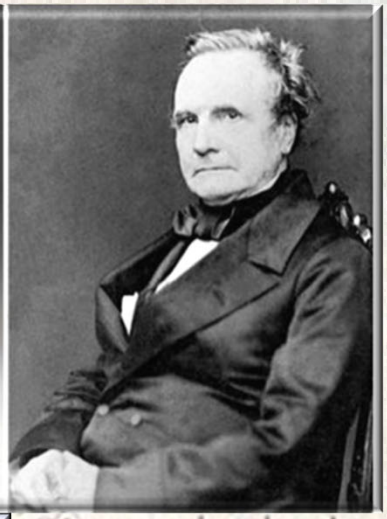
Klara and John Von Neumann

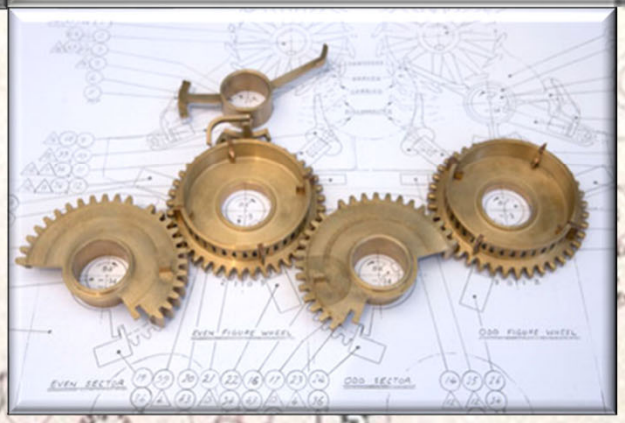
5.1KB memory! Many descendants, among them the MANIAC at Los Alamos Scientific Laboratory: hydrogen bombs and chess.





Babbage/Lovelace first to try to build it (before Turing)





distinction between *numbers that mean things* and *numbers that (do things) move matter*