Shannon Information from Borges to Shannon

Complex adaptive systems and computational intelligence (casci lab)

Resources
- web page
  - casci.binghamton.edu/academics/ssie501
- online class
  - binghamton.zoom.us/j/93351260610
- blog: sciber
  - sciber.blogspot.com
- Brightspace
  - brightspace.binghamton.edu/d2l/home/358842

SSIE-501/ISE-440 - Fall 2024

luis m. rocha

office hours:
Tuesdays 9:00- 11:30am
binghamton.zoom.us/my/luismrocha

Teaching Assistant

office hours:
Tuesdays: 7:00-8:00pm????
binghamton.zoom.us/my/

bit.ly/atBIC

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

- Background
- Interests
- Course expectations

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501

- Background
- Interests
- Course expectations

General Discussions, Questions or Suggestions ⌄

Topic	Threads	Posts	Last Post
General Discussion, Question or Suggestion ⌄ Aug 20 - Dec 22	1	1	Luis Rocha just now

bit.ly/atBIC

BINGHAMTON
UNIVERSITY
STATE UNIVERSITY OF NEW YORK

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501

evaluation

- **Participation: 20%.**
  - class discussion, everybody reads and discusses every paper
  - engagement in class
- **Paper Presentation and Discussion: 20%**
  - **SSIE501** students are assigned to papers individually or as group lead presenters and discussants
    - all students are supposed to read and participate in discussion of every paper.
  - Presenter prepares short summary of assigned paper (15 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.
  - **ISE440** students chose one of the presented papers to participate as lead discussant
    - not to present the paper, but to comment on points 2-3) above
  - Class discussion is opened to all
    - lead discussant ensures we important paper contributions and failures are addressed
- **Black Box: 60%**
  - Group Project (2 parts)
    - Assignment I (25%) and Assignment II (35%)

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501

key events coming up

- **Paper Presentation: 20%**
  - Present (501) and lead (501&440) the discussion of an article related to the class materials
    - Enginet students post/send video or join by Zoom synchronously
- **Module 1**: Cybernetics and the Information Turn
- **Today**
  - Borges, Jorge Luis. [1941]. *The Library of Babel*.
  - Borges, Jorge Luis. [1941]. *The Garden of Forking Paths* .
- **Next classes**
  - Discussion Set 1:
    - Kline, Ronald R [2015]. *The cybernetics moment, or, why we call our age the information age*. Johns Hopkins University Press. Chapters 1-2.
      - Optional: Heims, S.G. [1991]. *The Cybernetics Group*. MIT Press. Chapters: 1,2, 11, and 12
      - Optional: McCulloch, W. and W. Pitts [1943], "A Logical Calculus of Ideas Immanent in Nervous Activity". *Bulletin of Mathematical Biophysics* **5**:115-133.
    - Gleick, J. [2011]. *The Information: A History, a Theory, a Flood*. Random House. Chapter 8.
      - Optional: Prokopenko, Mikhail, Fabio Boschetti, and Alex J. Ryan. "An information theoretic primer on complexity, self-organization, and emergence." *Complexity* **15**.1 (2009): 11-28.
  - Discussion Set 2
    - Brenner, Sydney. [2012]. "History of Science. The Revolution in the Life Sciences". *Science* **338** (6113): 1427-8.
    - Brenner, Sydney. [2012]. "Turing centenary: Life's code script. *Nature* **482** (7386) (February 22): 461-461.
    - Cobb, Matthew. [2013]. "1953: When Genes Became 'Information'." *Cell* **153** (3): 503-506.
      - Optional: Searls, David B. [2010]. "The Roots of Bioinformatics". *PLoS Computational Biology* **6**(6): e1000809.
    - Weaver, W. [1948]. "Science and Complexity". *American Scientist*, **36**(4): 536-44. Also available in Klir, G.J. [2001]. *Facets of systems Science*. Springer, pp: 533-540.
  - **Discussion by all**

more upcoming readings (check brightspace)

- **Paper Presentation: 20%**
  - Present (501) and lead (501&440) the discussion of an article related to the class materials
    - Enginet students post/send video or join by Zoom synchronously
- **Module 2: Systems Science**
  - **Discussion Set 3:**
    - Klir, G.J. [2001]. *Facets of systems Science*. Springer. Chapters 1 and 2.
      - Optional:
      - Rosen, R. [1986]. "Some comments on systems and system theory". *Int. J. of General Systems*, **13**: 1-3. Available in: Klir, G.J. [2001]. *Facets of systems Science*. Springer. pp: 241-243.
      - Wigner, E.P. [1960], "The unreasonable effectiveness of mathematics in the natural sciences". Richard courant lecture in mathematical sciences delivered at New York University, May 11, 1959. *Comm. Pure Appl. Math*, **13**: 1-14.
    - Klir, G.J. [2001]. *Facets of systems Science*. Springer. Chapter 3.
  - **Discussion Set 4:**
    - Klir, G.J. [2001]. *Facets of systems Science*. Springer. Chapter 8.
      - Optional: Klir, G.J. [2001]. *Facets of systems Science*. Springer. Chapter 11
    - Schuster, P. (2016). The end of Moore's law: Living without an exponential increase in the efficiency of computational facilities. *Complexity*. **21**(S1): 6-9. DOI 10.1002/cplx.21824.
    - Von Foerster, H., P. M. Mora and L. W. Amiot [1960]. "Doomsday: Friday, November 13, AD 2026." Science **132**(3436):1291-5.
- **Future Modules**
  - See brightspace

more upcoming readings (check brightspace)

- **Paper Presentation: 20%**
  - Present (501) and lead (501&440)
    - Enginet students post/send video or
- **Module 2**: Systems Science
  - **Discussion Set 3**:
    - Klir, G.J. [2001]. *Facets of system*
      - Optional:
      - Rosen, R. [1986]. "Some commen
        Klir, G.J. [2001]. *Facets of system*
      - Wigner, E.P. [1960], "The unreaso
        in mathematical sciences delivere
    - Klir, G.J. [2001]. *Facets of system*
  - **Discussion Set 4**:
    - Klir, G.J. [2001]. *Facets of system*
      - Optional: Klir, G.J. [2001]. *Facets*
    - Schuster, P. (2016). The end of
      of computational facilities. *Comp*
    - Von Foerster, H., P. M. Mora and
      Science **132**(3436):1291-5.
- **Future Modules**
  - See brightspace

Fall 2023 Intro to Systems Science (ISE-...) — BINGHAMTON UNIVERSITY — Luis Rocha

Course Home | Calendar | **Content** | Assignments | Quizzes | Discussions | Evaluation ⌄ | Classlist | Course Tools ⌄ | Help ⌄

Papers for Presentations ⌄   🖨 Print   ⚙ Settings

- Syllabus / Overview
- Bookmarks
- Course Schedule
- Table of Contents — 48
- Syllabus
- Office Hours
- Readings — 45
- Papers for Presentations
- Zoom — 2
- For EngiNet Students — 1

Add dates and restrictions...

All **SSIE501** Students are assigned to one paper as *lead presenters and discussants*, but all students are supposed to read and participate in the discussion of every paper. During class, the presenter prepares a short summary of the paper (10-15 minutes)---no formal presentations or PowerPoint unless figures are indispensable. The 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.

After initial summary, discussion is opened to all, and role of presenter is to lead the discussion to make sure we address the important paper contributions and failures. **ISE440 students** will chose one of the presented papers to participate as lead discussant, whose role is not to present the paper, but to comment on points 2-3) above.
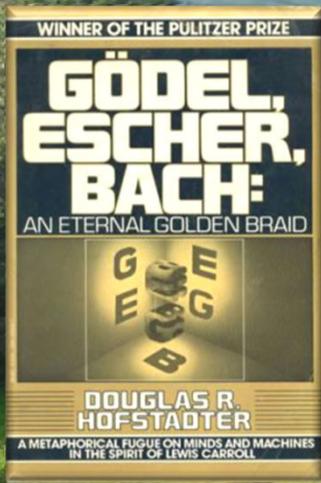
Next Presentations:

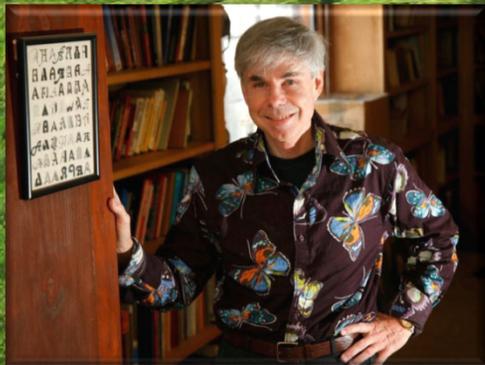Module 1 - Cybernetics and the Information Turn

Tuesday, August 29th

Presenter 1: Heims, S.G. [1991]. *The Cybernetics Group*. MIT Press. Chapters: 1 and 2.

Personal path in the garden of forking paths

WINNER OF THE PULITZER PRIZE

GÖDEL, ESCHER, BACH:
AN ETERNAL GOLDEN BRAID

DOUGLAS R. HOFSTADTER

A METAPHORICAL FUGUE ON MINDS AND MACHINES
IN THE SPIRIT OF LEWIS CARROLL

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

JORGE LUIS BORGES

EL JARDIN
DE SENDEROS
QUE SE BIFURCAN

SUR
BUENOS AIRES

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' autobiographies, the faithful catalogues of the Library, thousands and thousands of false catalogues, the demonstration of the fallacy of those catalogues, the demonstration of the fallacy of the true catalogue,[…] the true story of your death, the translation of every book in all languages…".
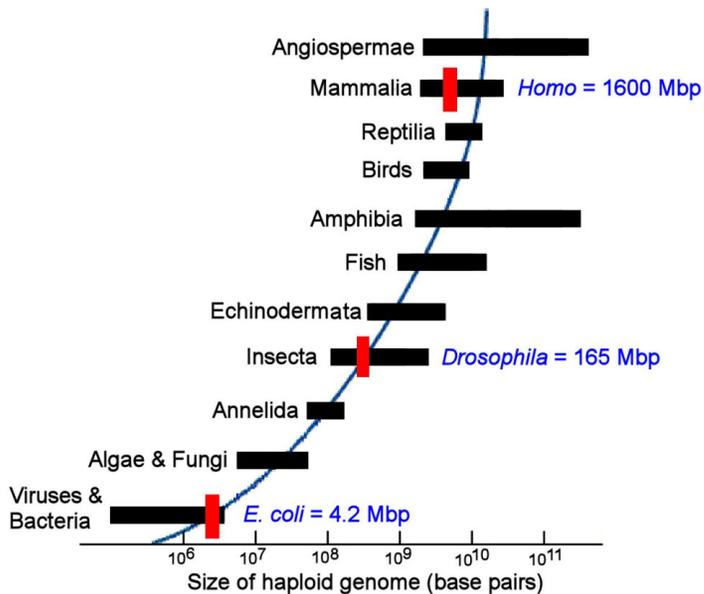
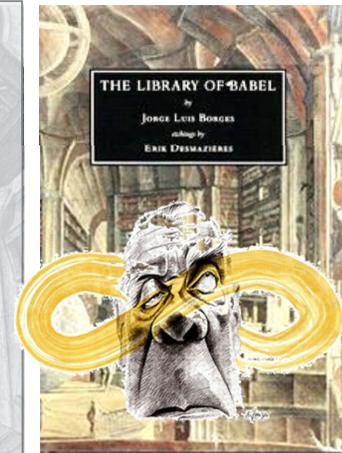"I have wandered in search of a book, perhaps the catalogue of catalogues"

JORGE LUIS BORGES
EL JARDIN DE SENDEROS QUE SE BIFURCAN
SUR
BUENOS AIRES

WINNER OF THE PULITZER PRIZE
GÖDEL, ESCHER, BACH:
AN ETERNAL GOLDEN BRAID
G E G
E G
B
DOUGLAS R. HOFSTADTER
A METAPHORICAL FUGUE ON MINDS AND MACHINES IN THE SPIRIT OF LEWIS CARROLL
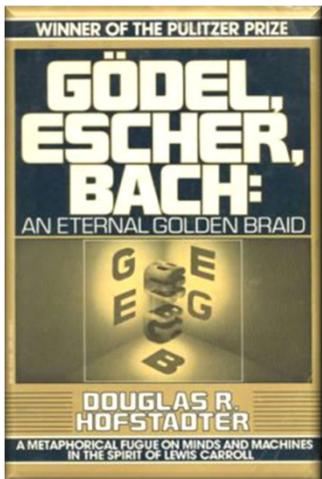
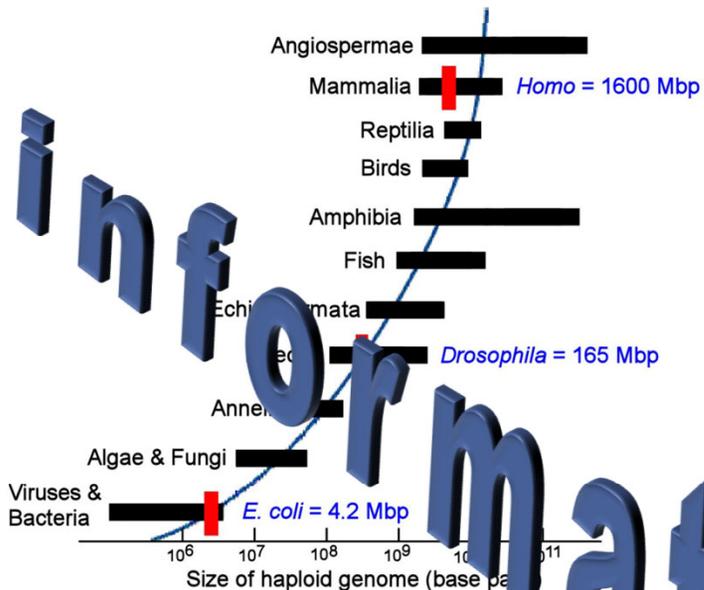Poetic essays on information and memory (1941)

numbers

- ■ Each book
  - 25 characters, written in any sequence for 410 pages of 40 lines of 80 characters
  - $410*40*80$ = sequence of $\approx 10^{6.1}$ characters
    - ■ $\approx 10^{7.2}$ base pairs (10 Mbp)
    - ■ $\approx$ 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!
    - ■ $\approx 1.956 \times 10^{1,834,097}$ 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)

THE LIBRARY OF BABEL
by Jorge Luis Borges
etchings by Erik Desmazières

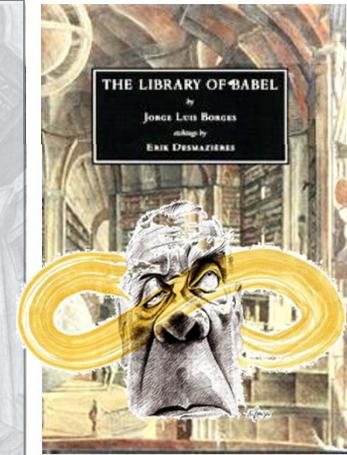**Information Space Is finite but larger than Physical space**

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

numbers

Angiospermae
Mammalia — *Homo* = 1600 Mbp
Reptilia
Birds
Amphibia
Fish
Echinodermata
Drosophila = 165 Mbp
Annelida
Algae & Fungi
Viruses & Bacteria — *E. coli* = 4.2 Mbp

10^6 10^7 10^8 10^9 10^{10} 10^{11}
Size of haploid genome (base pairs)

information

- **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!
    - ≈ $1.956 \times 10^{1,834,097}$ 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)

THE LIBRARY OF BABEL
by
JORGE LUIS BORGES
etchings by
ERIK DESMAZIÈRES

**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."

What to do in such information spaces to avoid becoming a Quixotic wanderer?

Are there principles of organization?

How did we get here?
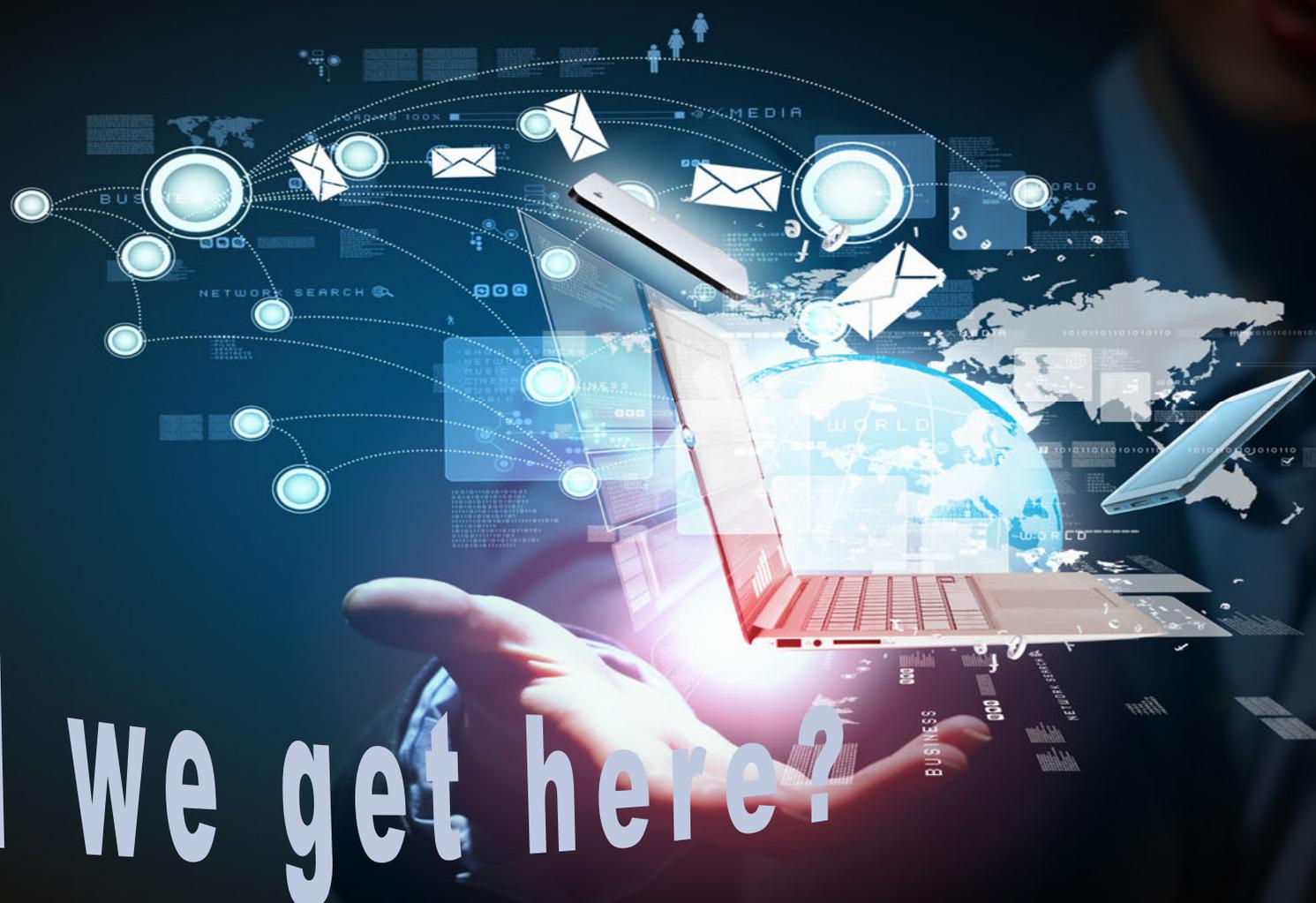
observer and choice

- *Information* is defined as "a measure of the freedom from *choice* with which a message is *selected* from the set of all possible messages"
- Bit (short for *binary digit*) is the most elementary **choice** one can make
  - Between two items: "0' and "1", "heads" or "tails", "true" or "false", etc.
  - Bit is equivalent to the choice between two equally likely alternatives
    - Example, if we know that a coin is to be tossed, but are unable to see it as it falls, a message telling whether the coin came up heads or tails gives us one bit of information

1 Bit of *information*

uncertainty **removed**,
information gained
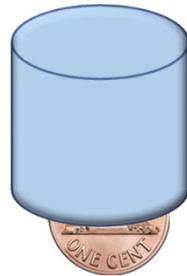
1 Bit of uncertainty

H,T?

choice between 2 symbols
recognized by an observer

Hartley, R.V.L., "Transmission of Information", *Bell System Technical Journal*, July 1928, p.535.

- **Information is transmitted through noisy communication channels**
  - Ralph Hartley and Claude Shannon (at Bell Labs), the fathers of Information Theory, worked on the problem of efficiently transmitting information; i. e. **decreasing the uncertainty** in the transmission of information.

C. E. Shannon [1948], "A mathematical theory of communication". *Bell System Technical Journal*, **27**:379-423 and 623-656

C. E. Shannon, "A Symbolic analysis of relay and switching circuits" .*MS Thesis*, (unpublished) MIT, 1937.

C. E. Shannon, "An algebra for theoretical genetics." *Phd Dissertation*, MIT, 1940.
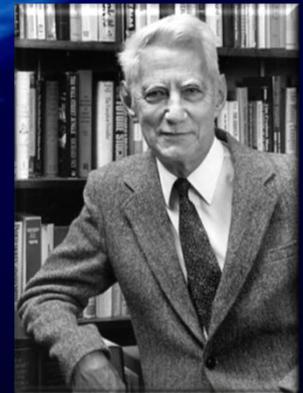
■ Multiplication Principle

- "If some choice can be made in M different ways, and some subsequent choice can be made in N different ways, then there are M x N different ways these choices can be made in succession" [Paulos]

  ■ 3 shirts and 4 pants = 3 x 4 = 12 outfit choices

Htr gnsfytsx%vzmpq-2wt| % nym%tsl%
xjvzjshjx%tk%(fwfggjx%fsi%xyfyj%mtmhjx.

- ## Nonspecificity
  - ### Hartley measure
    - The amount of uncertainty associated with a set of alternatives (e.g. messages) is measured by the **amount of information needed to <u>remove</u> the uncertainty**
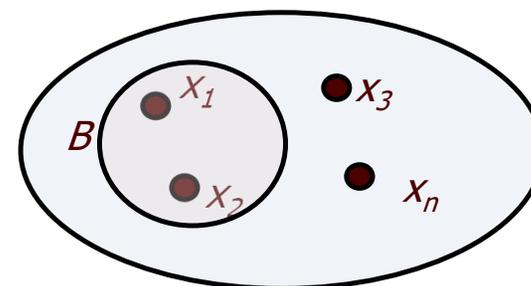
A = Set of Alternatives

Quantifies how many yes-no questions need to be asked to establish what the correct alternative is

Elementary Choice is between 2 alternatives: 1 bit

$$H(A) = \log_2 |A|$$

Measured in bits

Number of Choices

$$H(B) = \log_2(2) = 1$$

$$\log_2(4) = 2 \qquad 2^2 = 4$$

$$\log_2(16) = 4$$

$$2^4 = 16$$

$$\log_2(1) = 0$$

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501

$$H(A) = \log_2(16) = 4$$

$$H(B) = \log_2(4) = 2$$

$$H(A) = \log_2|A|$$

Measured in bits

Number of Choices

Quantifies how many yes-no questions need to be asked to establish what the correct alternative is
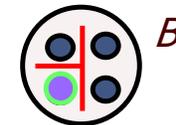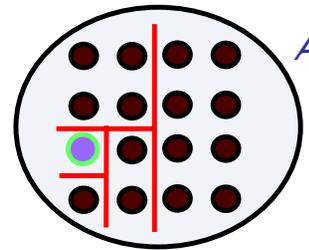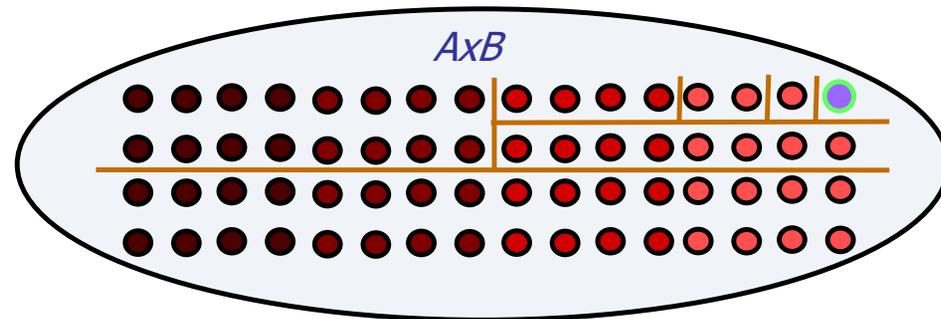
- Example
  - Menu Choices
    - A = 16 Entrees
    - B = 4 Desserts
  - How many dinner combinations?
    - 16 x 4 = 64

A

B

$$H(A \times B) = \log_2(16 \times 4) =$$
$$= \log_2(16) + \log_2(4) = 6$$

AxB

readings

■ **Class Book**
- Klir, G.J. [2001]. *Facets of systems science*. Springer.

■ **Papers and other materials**
- **Discussion Set 1**
  - Kline, Ronald R [2015]. *The cybernetics moment, or, why we call our age the information age*. Johns Hopkins University Press. Chapters 1-2.
    - Heims, S.G. [1991]. *The Cybernetics Group*. MIT Press. Chapters: 1,2, 11, 12.
    - Optional: McCulloch, W. and W. Pitts [1943], "A Logical Calculus of Ideas Immanent in Nervous Activity". *Bulletin of Mathematical Biophysics* **5**:115-133.
  - Gleick, J. [2011]. *The Information: A History, a Theory, a Flood*. Random House. Chapter 8.
    - Optional: Prokopenko, Mikhail, Fabio Boschetti, and Alex J. Ryan. "An information theoretic primer on complexity, self-organization, and emergence." *Complexity* **15**.1 (2009): 11-28.

- **Discussion Set 2**
  - Brenner, Sydney. [2012]. "History of Science. The Revolution in the Life Sciences". *Science* **338** (6113): 1427-8.
  - Brenner, Sydney. [2012]. "Turing centenary: Life's code script. *Nature* **482** (7386) (February 22): 461-461.
  - Cobb, Matthew. [2013]. "1953: When Genes Became 'Information'." *Cell* **153** (3): 503-506.
    - Optional: Searls, David B. [2010]. "The Roots of Bioinformatics". *PLoS Computational Biology* **6**(6): e1000809.
  - Weaver, W. [1948]. "Science and Complexity". *American Scientist*, **36**(4): 536-44. Also available in Klir, G.J. [2001]. *Facets of systems Science*. Springer, pp: 533-540.

BINGHAMTON
UNIVERSITY
STATE UNIVERSITY OF NEW YORK