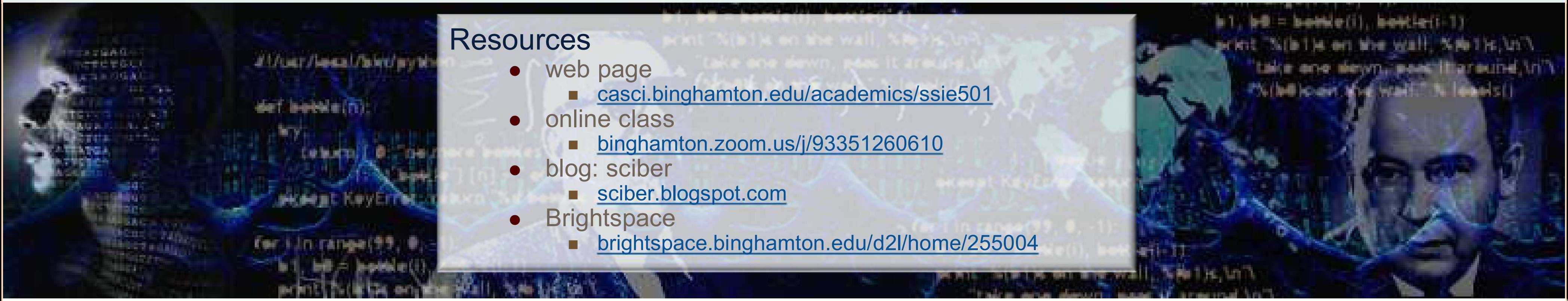


Resources

- 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/255004

hiroki sayama

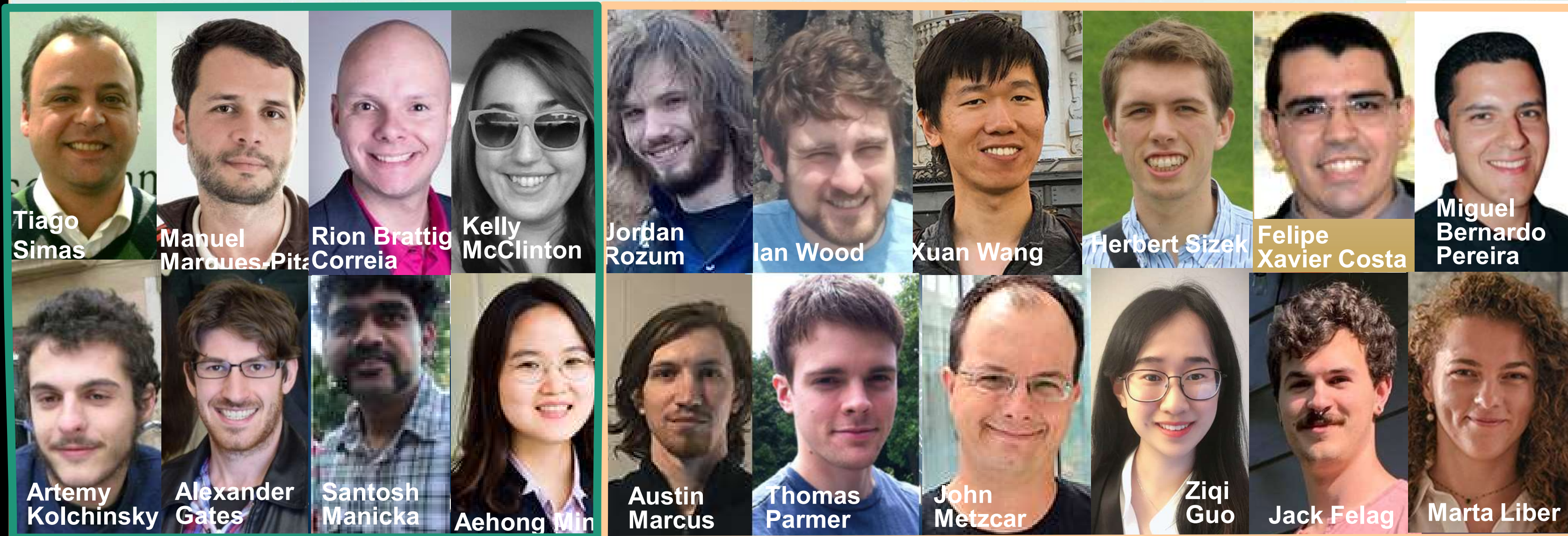
SSIE-501 - spring 2024

luis m. rocha

office hours:
Wednesdays: ???
binghamton.zoom.us/my/hirokisayama

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

interdisciplinary science



for understanding social and biomedical complexity

luis m. rocha

FULBRIGHT

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

cnrs

NSF

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

DEPARTMENT OF ENERGY · UNITED STATES OF AMERICA

FUNDAÇÃO LUSO-AMERICANA

IARPA BE THE FUTURE

PERSISTENT

careerbuilder

NIH NLM

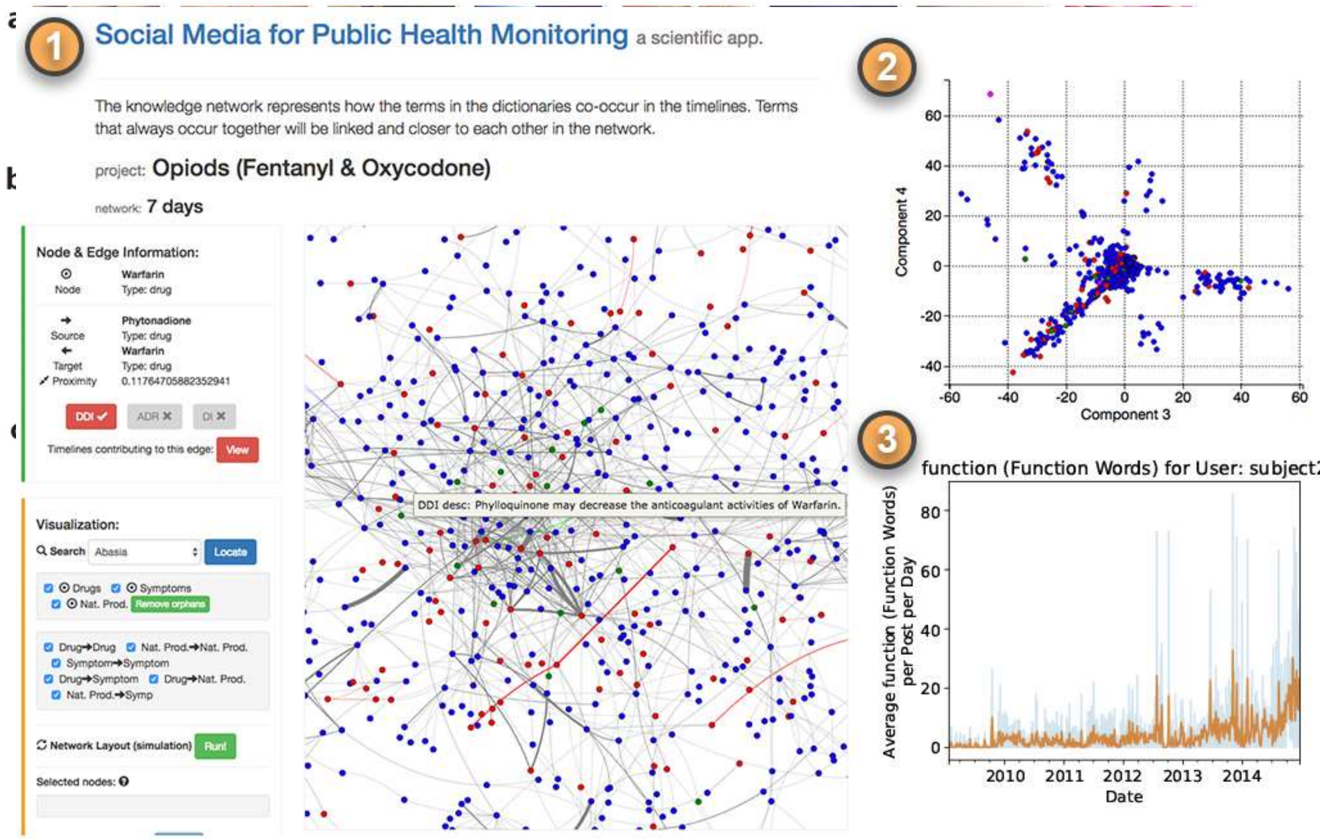
GULBENKIAN SCIENCE

rocha@binghamton.edu
casci.binghamton.edu


Lisbon


Luanda

social media data pipelines for biomedicine



Min et al [2023]. *CHI 2023*. **32.**

Wood, Correia, Miller, &Rocha [2022]. *Epilepsy & Behavior*. **128**: 108580.

Correia, Wood, Bollen, & Rocha [2020]. *Annual Review of Biomedical Data Science*, 3:1.

Wood, Varela, Bollen, Rocha & Sá [2017]. *Scientific Reports. 7: 17973* .

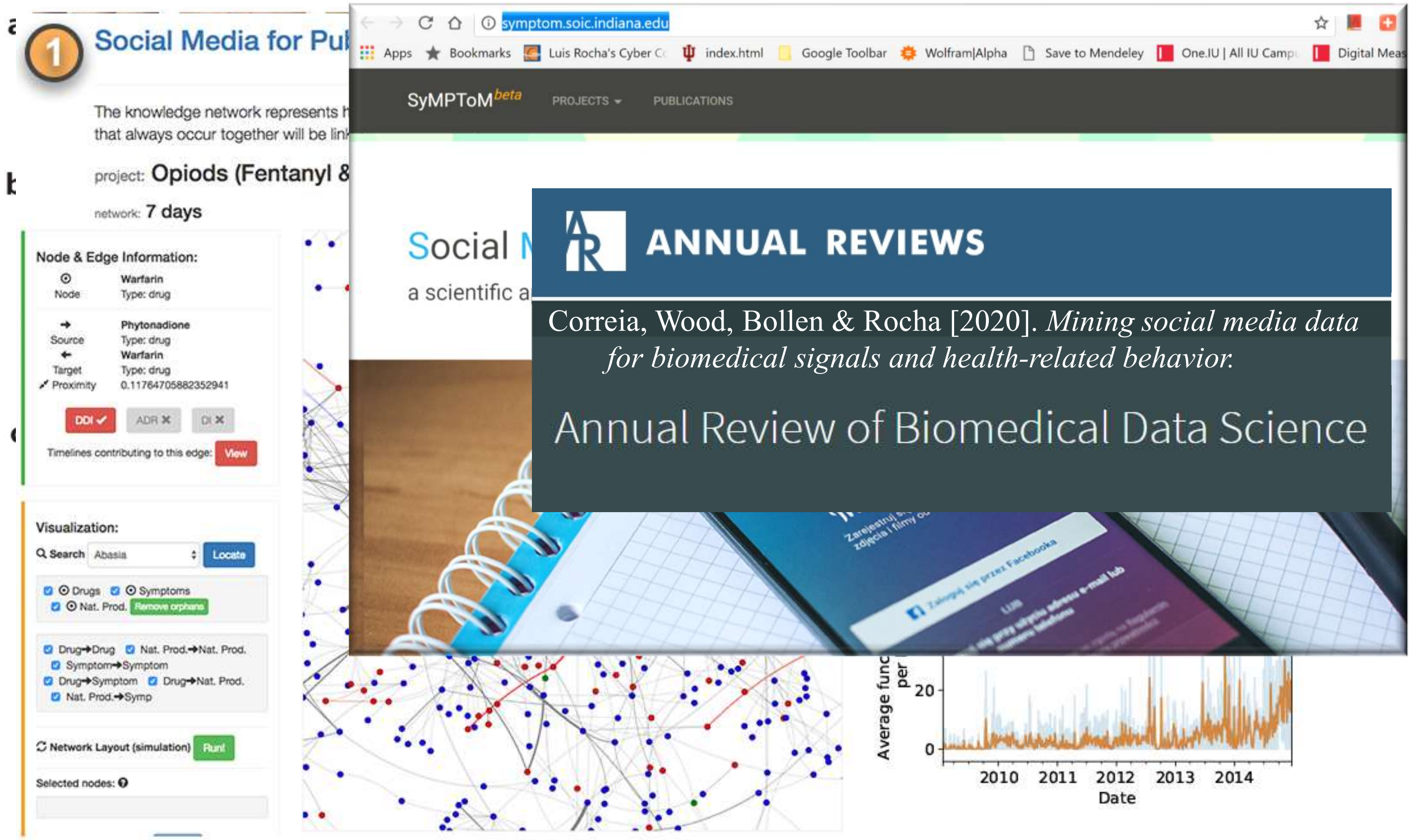
Correia, Li & Rocha [2016]. *PSB*: **21**:492-503.

Ciampaglia, et al [2015]. *PloS ONE*. **10**(6): e0128193.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK

rocha@binghamton.edu
informatics.indiana.edu/rocha/academics/i-bic

social media data pipelines for biomedicine



Min et al [2023]. *CHI 2023*. **32.**

Wood, Correia, Miller, &Rocha [2022]. *Epilepsy & Behavior*. **128**: 108580.

Correia, Wood, Bollen, & Rocha [2020]. *Annual Review of Biomedical Data Science*, 3:1.

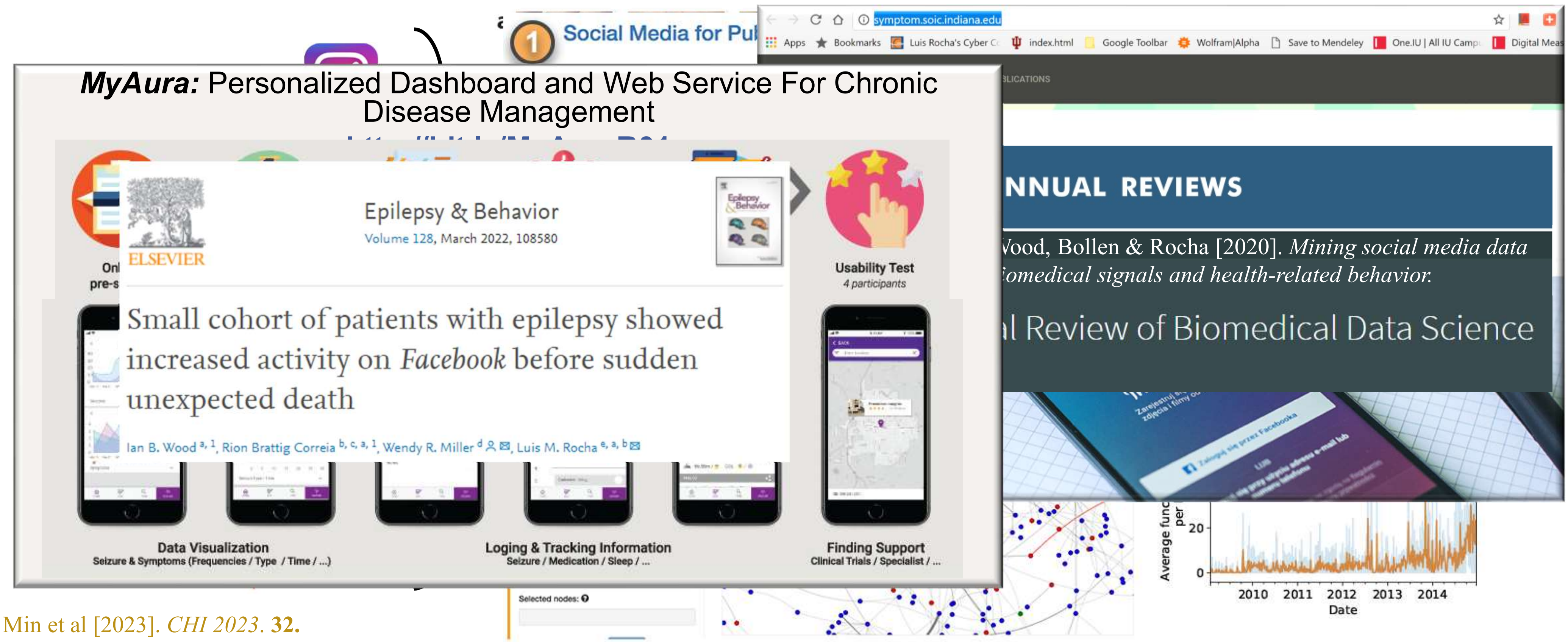
Wood, Varela, Bollen, Rocha & Sá [2017]. *Scientific Reports. 7: 17973* .

Correia, Li & Rocha [2016]. *PSB*: **21**:492-503.

Ciampaglia, et al [2015]. *PloS ONE*. **10**(6): e0128193.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
informatics.indiana.edu/rocha/academics/i-bic

social media data pipelines for biomedicine



Min et al [2023]. *CHI 2023*. **32.**
Wood, Correia, Miller, &Rocha [2022]. *Epilepsy & Behavior*. **128**: 108580.
Correia, Wood, Bollen, & Rocha [2020]. *Annual Review of Biomedical Data Science*, 3:1.
Wood, Varela, Bollen, Rocha & Sá [2017]. *Scientific Reports. 7: 17973* .
Correia, Li & Rocha [2016]. *PSB*: **21**:492-503.
Ciampaglia, et al [2015]. *PloS ONE*. **10**(6): e0128193.

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
informatics.indiana.edu/rocha/academics/i-bic

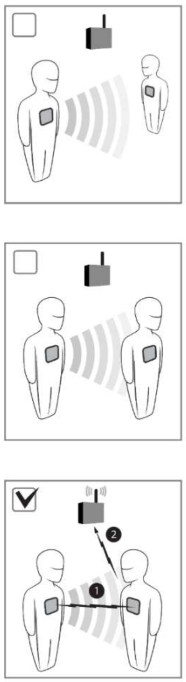
integrating and analyzing multiomic electronic health records with network science

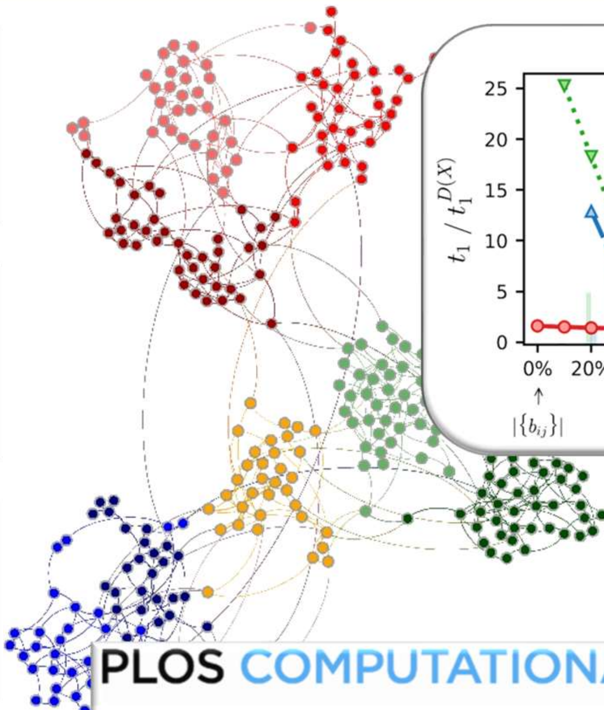
to predict comorbidity & drug interaction networks, disease factors & interventions

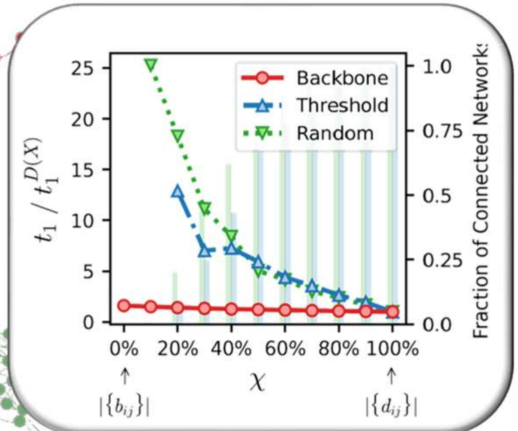
BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
informatics.indiana.edu/rocha/academics/i-bic

integrating and analyzing multilevel data sources with network science

to predict disease spread, information integration







PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

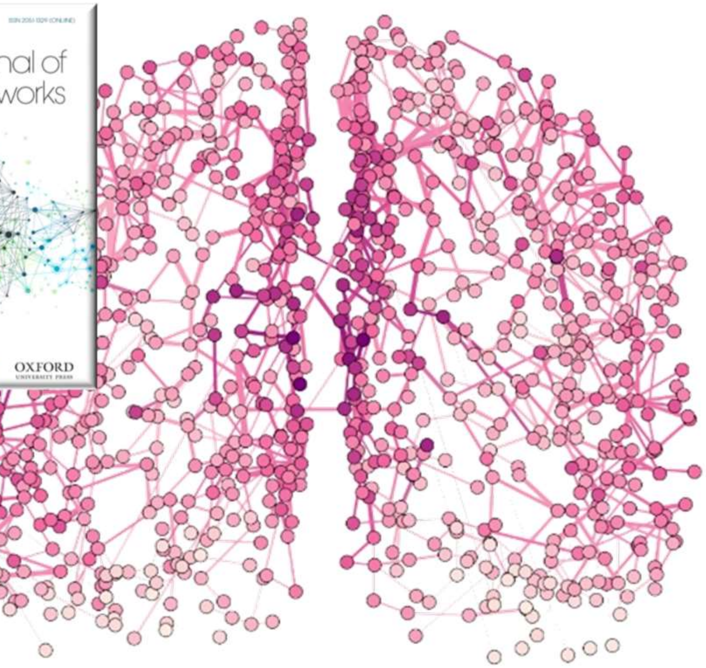
Contact networks have small metric backbones that maintain community structure and are primary transmission subgraphs

Rion Brattig Correia, Alain Barrat, Luis M. Rocha

frontiers in NEUROINFORMATICS

ORIGINAL RESEARCH ARTICLE
published: 24 July 2014
doi: 10.3389/fninf.2014.00066

Multi-scale integration and predictability in resting state brain activity



Simas & Rocha [2015]. *Network Science*. doi:10.1017/nws.2015.11
Simas, Correia & Rocha [2021]. *J Complex Networks*. **9** (6), cnab021.

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
informatics.indiana.edu/rocha/academics/i-bic

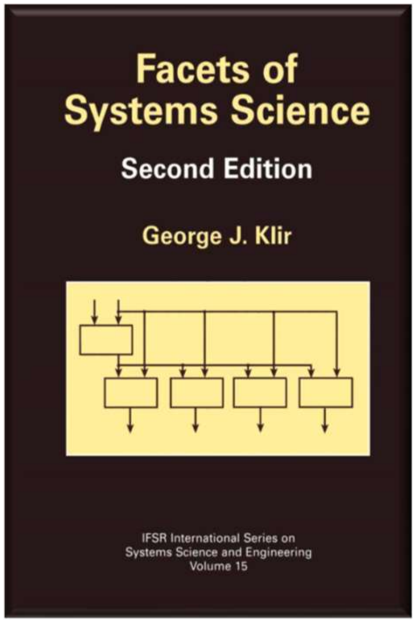

E-TRASH
RIOT BOOTIQU
DJ ANGST
FEB 21 - 10 PM - NO COVER
FRIDAY, AUGUST 19: 1AM (BASEMENT)


E-TRASH
LIVE IN
TikTok
@djetrash
TikTok
@djetrash
TikTok
@djetrash
NO COVER
FRIDAY, AUGUST 19: 1AM (BASEMENT)

what about you?

- Background
- Interests
- Course expectations

course materials



- Lecture slides and notes
  - See course web page and brightspace
- Web links and general materials
  - Blog (**sciber.blogspot.com**) and brightspace
- Class Book
  - Klir, G.J. [2001]. *Facets of systems science*. Springer.
    - Available in electronic format for SUNY students.
- Various literature for discussion
  - Course web site and brightspace

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Overview and aims

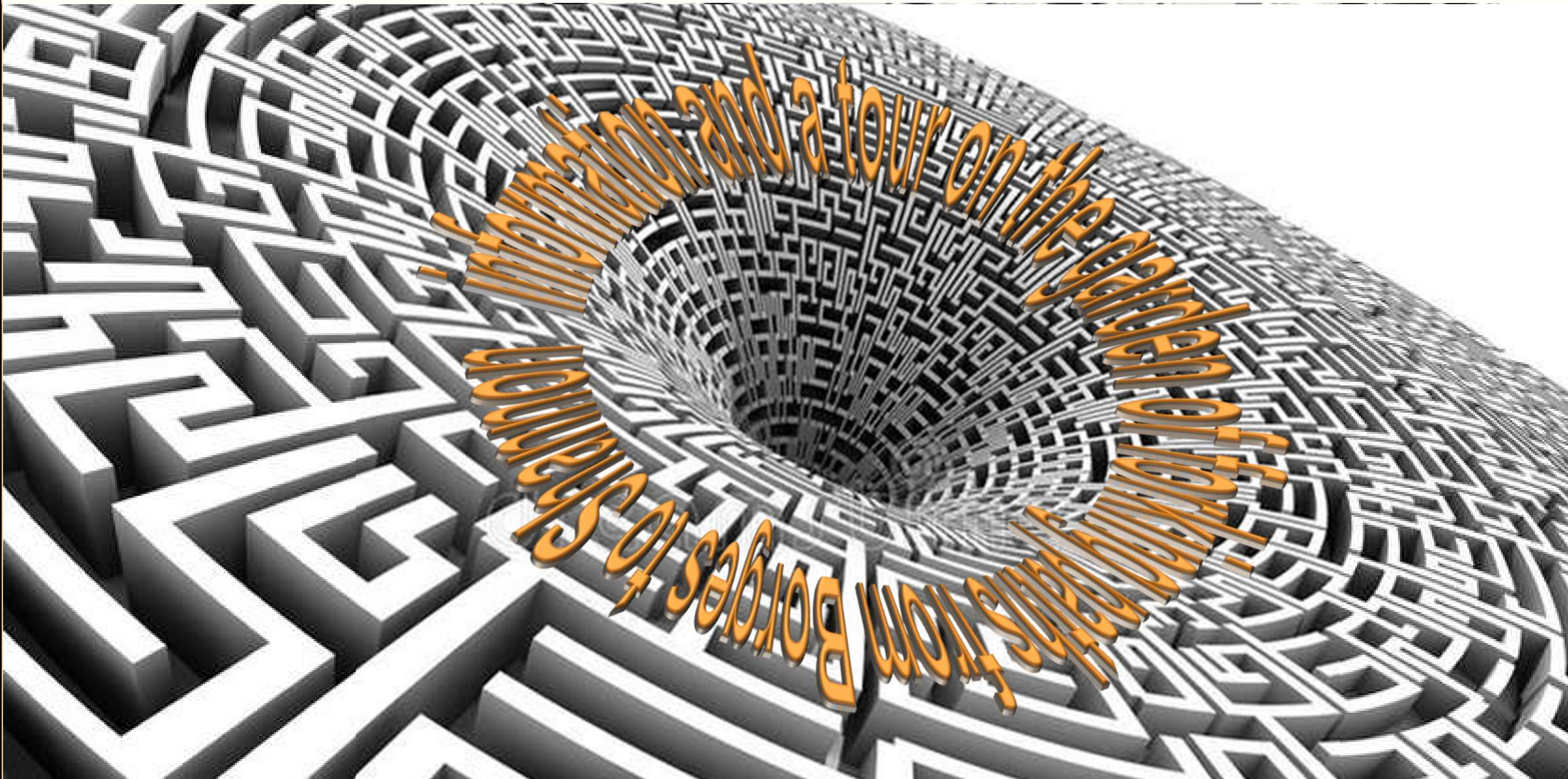
- The course deals with the foundations of Systems Science, as well as current advances in Complex Networks and Systems which is the modern expression of this interdisciplinary field. .
- Aims
  - Introduce and discuss the history, methodology and impact of complex systems science.
  - key literature, recent advances, and computational techniques in the field.
  - study concepts such as
    - Information, General Systems Theory, Networks, Modeling, Multi-Level Complexity, as well as their impact on science and society.
  - The course will also attempt to define and understand what systems thinking can bring to science and society.

rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

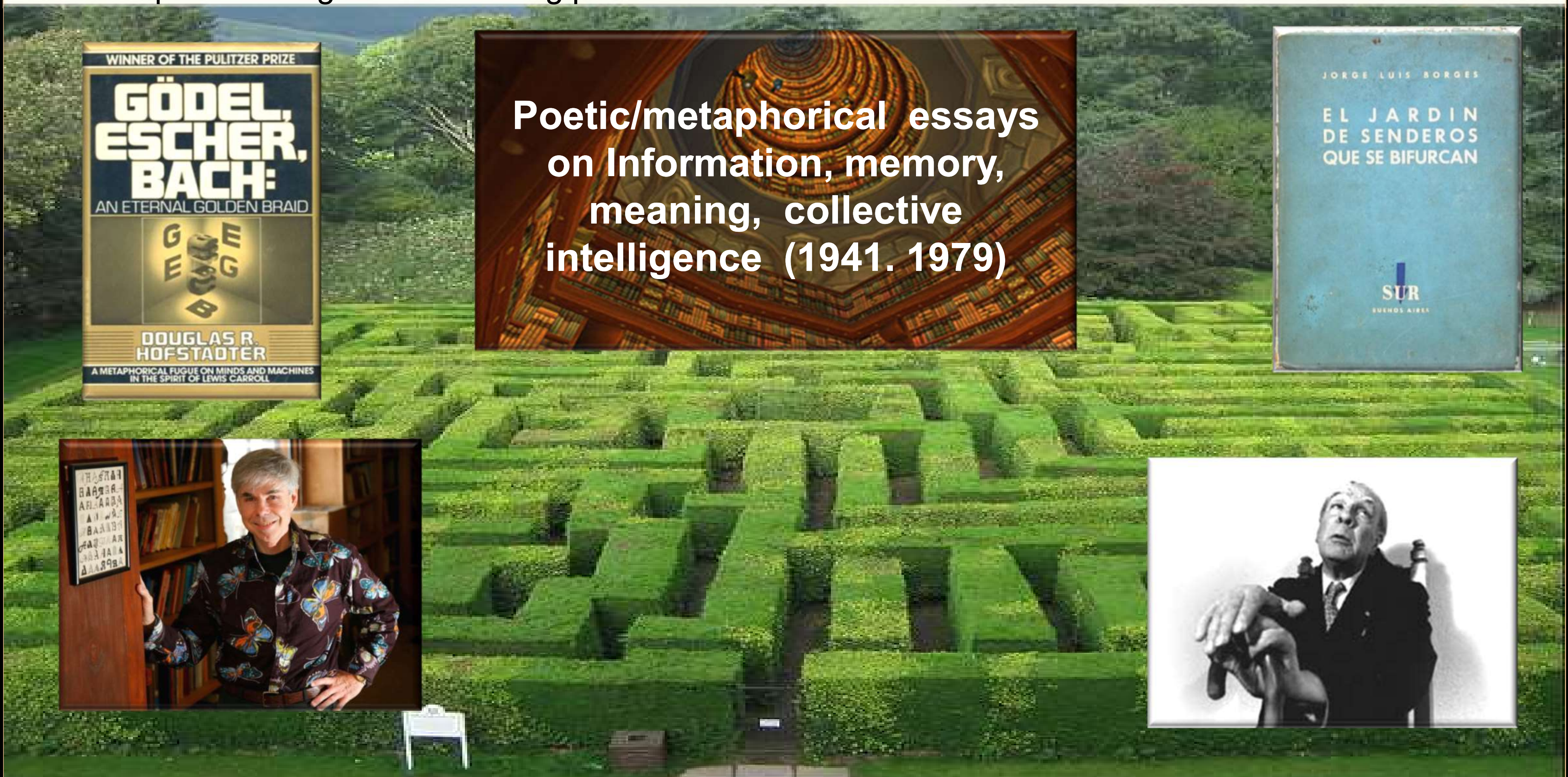
evaluation

- 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 addressed
- Python Homework: 25%
  - From Python workshop (3rd Session Prof. Sayama)
- Term Paper/Project proposal: 35%
  - A paper with an overview of the topics and literature covered, or a proposal for a project that uses complex systems thinking in your domain of expertise

lecture 1: information and a tour on the garden of forking paths from Borges to Shannon



Personal path in the garden of forking paths



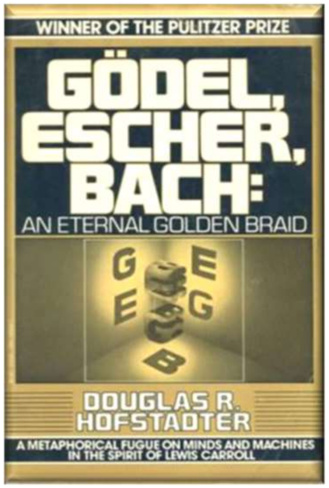
Jorge Luis Borges (1899 – 1986)

"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"

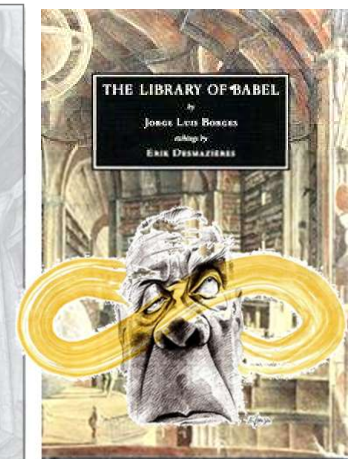


Poetic essays on information and memory (1941)

numbers

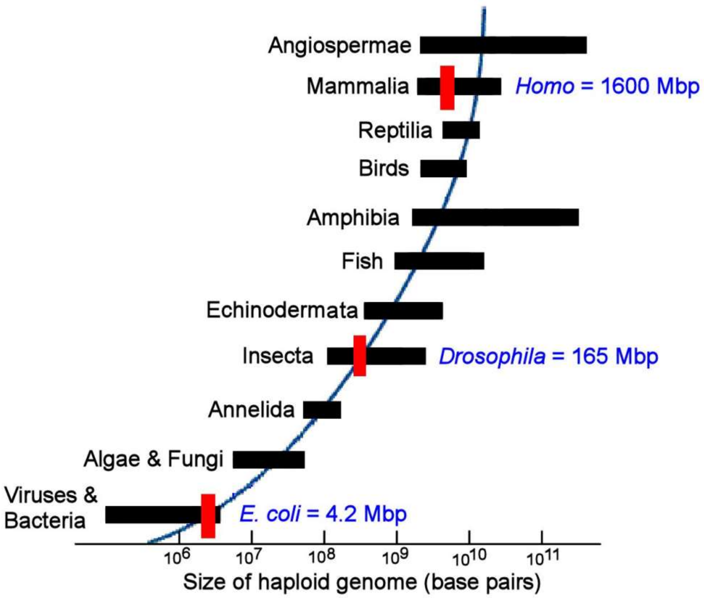
information

- 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 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)

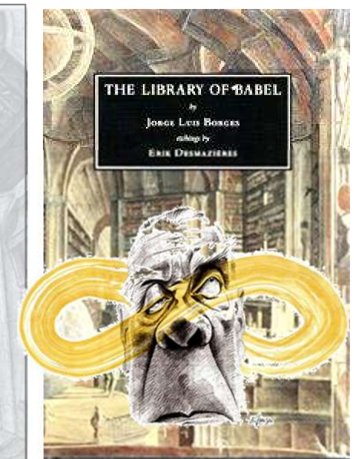


Information Space Is finite but larger than Physical space

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 <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."

rocha@binghamton.edu
informatics.indiana.edu/rocha/academics/i-bic

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

Are there principles of organization?

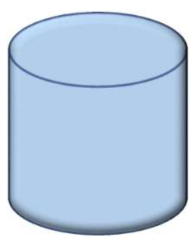

How did we get here?
MEDIA
NETWORK SEARCH
WORLD
BUSINESS

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

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

Fathers of uncertainty-based information

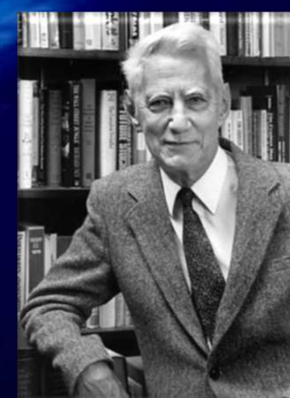
- Information is transmitted through noisy communication channels
  - Ralph Hartley and Claude Shannon (at Bell Labs), the fathers of Information Theory, worked on the problem of efficiently transmitting information; i. e. ***decreasing the uncertainty*** in the transmission of information.

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

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

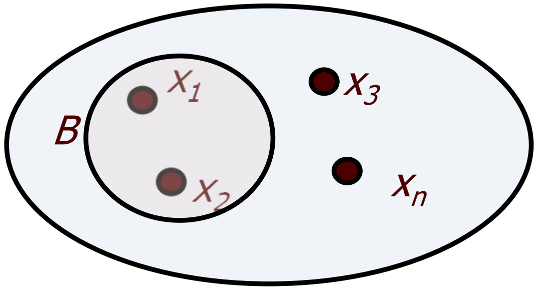
Combinations quickly grow with long sequences of variables (and state choices)

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

- 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**

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

A = Set of Alternatives



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

Measured in bits

Number of Choices

Elementary Choice is between 2 alternatives: 1 bit

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

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

$$\log_2(16) = 4 \qquad 2^4 = 16$$

$$\log_2(1) = 0$$

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

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

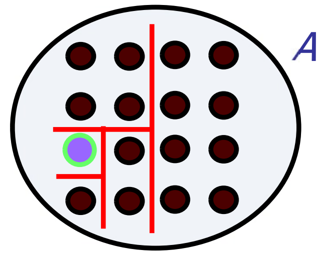
$$H(A) = \log_2 |A|$$

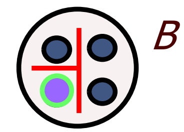
Measured in bits

Number of Choices

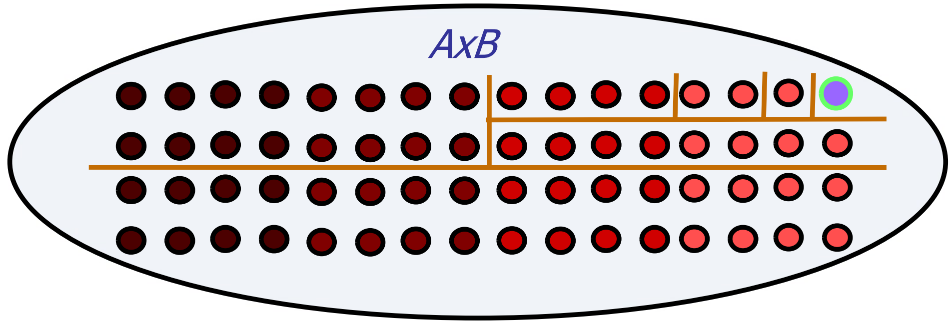
Quantifies how many yes-no questions need to be asked to establish what the correct alternative is

- Example
  - Menu Choices
    - A = 16 Entrees
    - B = 4 Desserts
  - How many dinner combinations?
    - 16 x 4 = 64

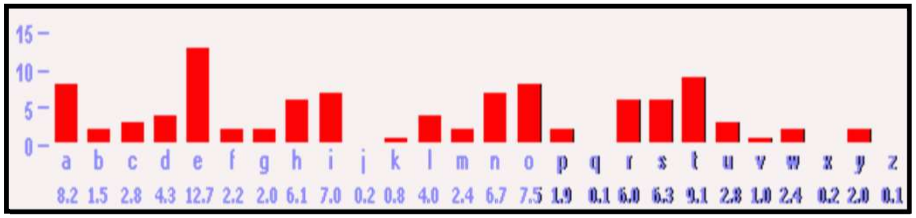




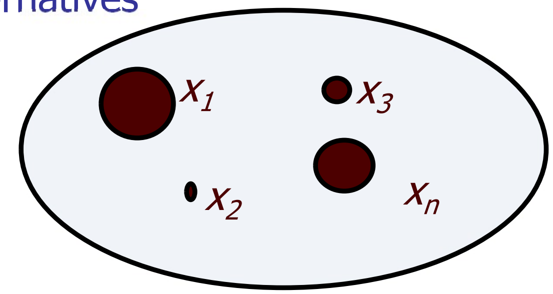
$$H(A \times B) = \log_2(16 \times 4) =$$
$$= \log_2(16) + \log_2(4) = 6$$



uncertainty-based information



A = Set of weighted Alternatives



- **Shannon's measure**
  - The ***average*** amount of uncertainty associated with a set of ***weighted*** alternatives (e.g. messages) is measured by the ***average*** amount of information needed to remove the uncertainty

$$H_S(A) = -\sum_{i=1}^{n} p(x_i)\log_2\big(p(x_i)\big)$$

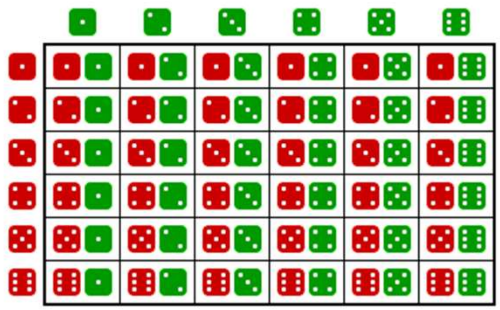
Probability of alternative

Measured in bits

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

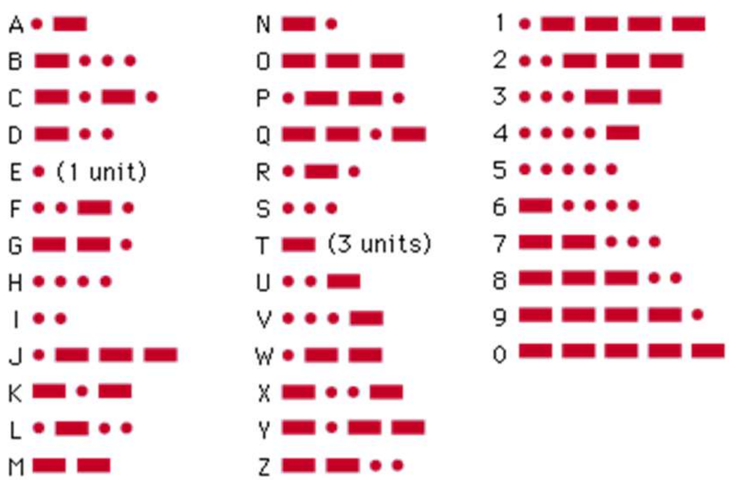
entropy of a message

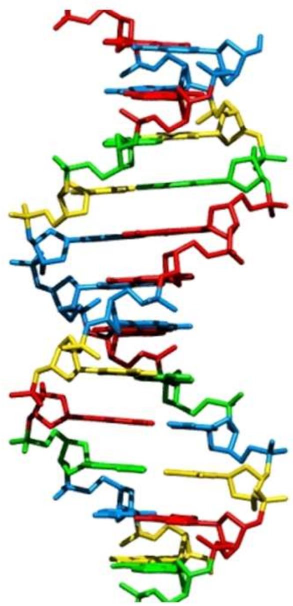
alphabet examples

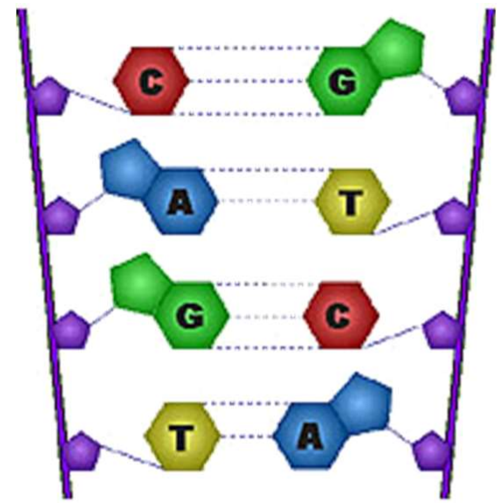


Message encoded in an alphabet of ***n*** symbols, for example:

- English (26 letters + space + punctuations)
- Morse code (dot, dash, space)
- DNA (A, T, G, C)
- Two dice (11 integers)







BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

5-letter “english”

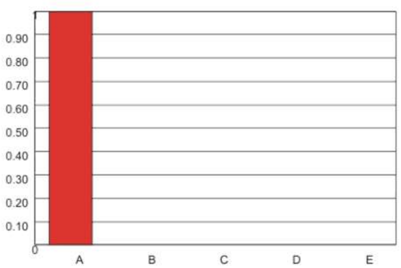
- Given a symbol set {A,B,C,D,E}
  - And occurrence probabilities P_A, P_B, P_C, P_D, P_E,
- The Shannon entropy is
  - The average minimum number of bits needed to represent a symbol

$$H_S = -\left(p_A \log_2(p_A) + p_B \log_2(p_B) + p_C \log_2(p_C) + p_D \log_2(p_D) + p_E \log_2(p_E)\right)$$

$$H_S = -(1.\log_2(1) + 0.\log_2(0) + 0.\log_2(0) + 0.\log_2(0) + 0.\log_2(0)) = -\log_2(1)$$

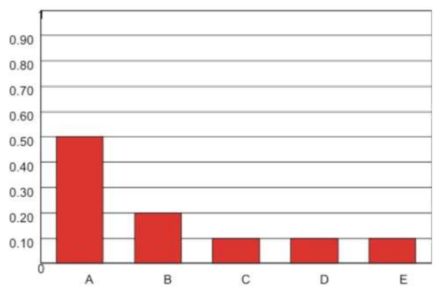
$$H_S = -5.\left(\frac{1}{5}\right).\log_2\left(\frac{1}{5}\right) = -(\log_2(1) - \log_2(5)) = \log_2(5)$$

$$H_S = -(\tfrac{1}{2}.\log_2\left(\tfrac{1}{2}\right) + \tfrac{1}{5}.\log_2\left(\tfrac{1}{5}\right) + 3.\left(\tfrac{1}{10}\right).\log_2\left(\tfrac{1}{10}\right))$$



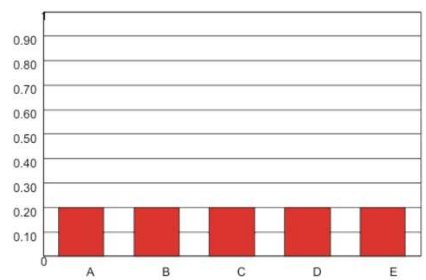
H_S = 0 bits

0 questions



H_S = 1.96

≈2 questions



H_S = 2.32 bits

information is surprise

what it measures

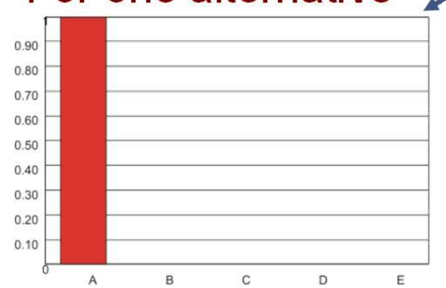
uncertainty, about outcome. How much information is gained when symbol is known

- ***on average***, how many *yes-no* questions need to be asked to establish what the symbol is
- "structure" of uncertainty in situations

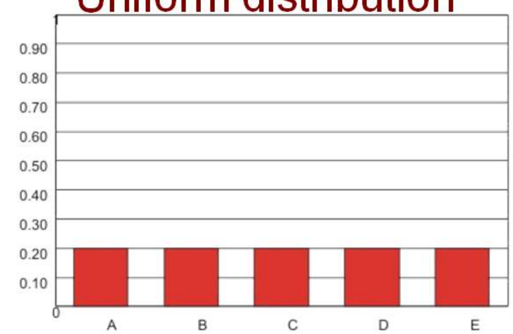
$$H_S \in = -\sum_{i=1}^{n} p(x_i)\log_2\big(p(x_i)\big)$$

$$H_S \in [0, \log_2|X|]$$

For one alternative



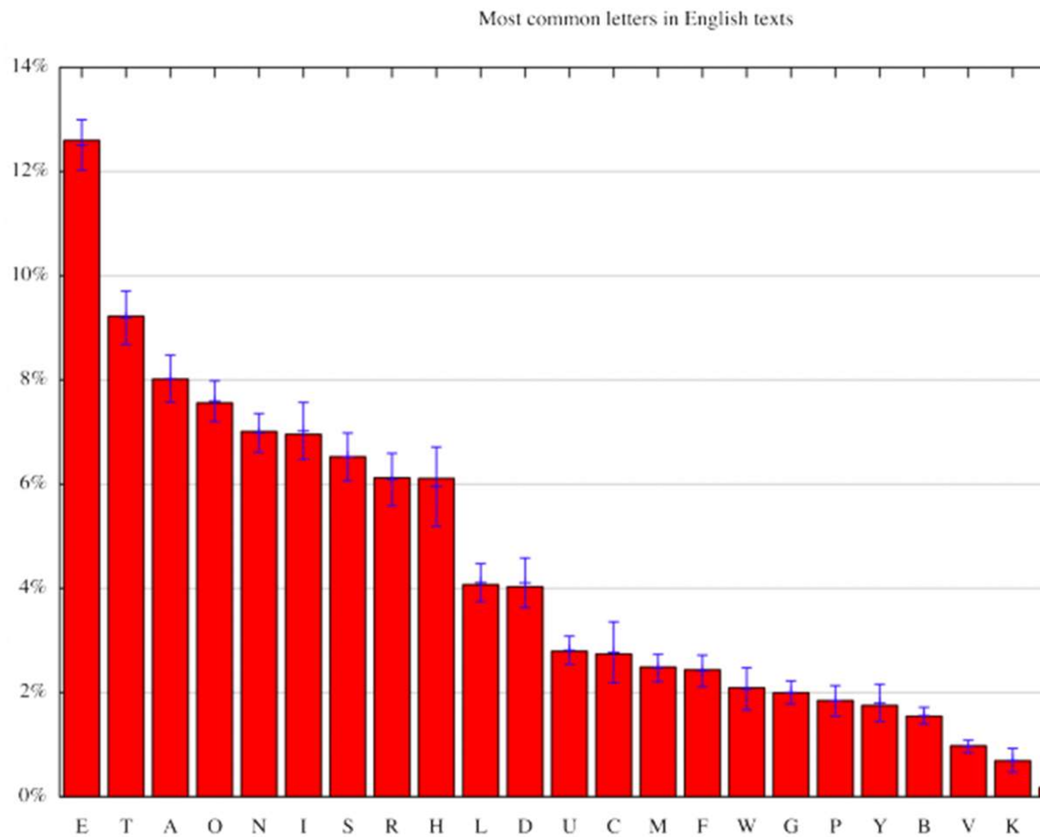
Uniform distribution



from letter frequency

	p(x)	log2(p(x))	-p(x).log2(p(x))
e	0.124167	-3.0096463	0.373698752
t	0.096923	-3.3670246	0.326340439
a	0.082001	-3.6082129	0.295877429
i	0.076805	-3.7026522	0.284382943
n	0.076406	-3.7101797	0.283478135
o	0.07141	-3.8077402	0.271908822
s	0.070677	-3.8226195	0.270170512
r	0.066813	-3.903723	0.260820228
l	0.044831	-4.4793659	0.200813559
d	0.036371	-4.7810716	0.173891876
h	0.035039	-4.8349111	0.169408515
c	0.034439	-4.8598087	0.167367439
u	0.028777	-5.11894	0.147307736
m	0.028[illegible]	[illegible]	[illegible]094755
f	0.023[illegible]	[illegible]	[illegible]220629
p	0.020[illegible]	[illegible]	[illegible]205704
y	0.018918	-5.7240814	0.108289316
g	0.018119	-5.7863688	0.104842059
w	0.013523	-6.2084943	0.083954364
v	0.012457	-6.3269343	0.078812722
b	0.010658	-6.5519059	0.069830868
k	0.00393	-7.9911852	0.031406876
x	0.002198	-8.8294354	0.019409218
j	0.001998	-8.9669389	0.017919531
q	0.000933	-10.066609	0.009387113
z	0.000599	-10.705156	0.006412389
		Entropy	**4.14225193**

Hartley Measure
H(|26|) 4.7004397



	p(x)	log2(p(x))	-p(x).log2(p(x))
Space	0.18288	-2.4509943	0.448249175
E	0.10267	-3.2839625	0.337152952
T	0.07517	-3.7336995	0.280662128
A	0.06532	-3.9362945	0.257125332
O	0.06160	-4.0210249	0.247678132
N	0.05712	-4.1298574	0.235897914
I	0.05668	-4.1409036	0.234724772
S	0.05317	-4.2332423	0.225081718
R	0.04988	-4.3254212	0.215748053
H	0.04979	-4.3281265	0.215478547
L	0.[illegible]	[illegible]	[illegible]63015644
D	0.[illegible]	[illegible]	[illegible]61811184
U	0.[illegible]	[illegible]	[illegible]124201198
C	0.02234	-5.4844363	0.122504535
M	0.02027	-5.6248177	0.113990747
F	0.01983	-5.6561227	0.112164711
W	0.01704	-5.8750208	0.100104113
G	0.01625	-5.9435013	0.096576215
P	0.01504	-6.0547406	0.091082933
Y	0.01428	-6.1301971	0.087518777
B	0.01259	-6.3117146	0.079456959
V	0.00796	-6.9728048	0.055511646
K	0.00561	-7.4778794	0.041948116
X	0.00141	-9.4709063	0.013346416
J	0.00098	-10.001987	0.009754119
Q	0.00084	-10.222907	0.008554069
Z	0.00051	-10.929184	0.005604998
		Entropy	**4.0849451**

Hartley Measure
H(|27|) 4.7548875

http://www.macfreek.nl/memory/Letter_Distribution

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

- entropy quantifies information (surprise), but it does not consider information content
  - semantic aspects of information are irrelevant to the engineering problem in Shannon's conception

We were good, we were gold
Kinda dream that can't be sold
We were right 'til we weren't
Built a home and watched it burn

Mm, I didn't wanna leave you
I didn't wanna lie
Started to cry, but then remembered I
I can buy myself flowers
Write my name in the sand
Talk to myself for hours
Say things you don't understand
I can take myself dancing
And I can hold my own hand
Yeah, I can love me better than you can

$$H_S \in = -\sum_{i=1}^{n} p(x_i)\log_2\big(p(x_i)\big)$$

wdeo eog geWI ewr e deorw
aainhmta d rettoeKandl dsbc
eeeier ntw hWttr ewrgliwe
oriaedatmht ndc lwn thuaBeuib

eanm dtal vewdi nl o unMay
al indn nItawde i
cI rettedtebrmSrb reemntuy da oth e
uolrawe blnffmsyylc es
niWe dty ne rsehmntiama
arem TlI ssytrfu fkoooh
nyoh e gdodudtnaraustsi tnynS
atf Ik emcnegyn snlicad a
hmhydcndAwannoo n dl I a
tlhl eatta nom Ybrueny h ee oaavn cce

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

entropy according to probabilistic model

0th order model: equiprobable symbols

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

Hartley Measure
H(|27|) 4.7548875

XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGXYD QPAAMKBZAACIBZLHJQD

1st order model: frequency of symbols

$$H_S(A) = -\sum_{i=1}^{n} p(x_i)\log_2(p(x_i))$$

$H_S = 4.08$

OCRO HLI RGWR NMIELWIS EU LL NBNESBEYA TH EEI ALHENHTTPA OOBTTVA NAH BRL

2nd order model: frequency of digrams

Most common *digrams*: th, he, in, en, nt, re, er, an, ti, es, on, at, se, nd, or, ar, al, te, co, de, to, ra, et, ed, it, sa, em, ro.

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE

3rd order model: frequency of trigrams

Most common *trigrams*: the, and, tha, ent, ing, ion, tio, for, nde, has, nce, edt, tis, oft, sth, men

IN NO IST LAT WHEY CRATICT FROURE BERS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE

4th order model: frequency of tetragrams

$H_S = 2.8$

THE GENERATED JOB PROVIDUAL BETTER TRAND THE DISPLAYED CODE ABOVERY UPONDULTS WELL THE CODERST IN THESTICAL IT DO HOCK BOTHE MERG INSTATES CONS ERATION NEVER ANY OF PUBLE AND TO THEORY EVENTIAL CALLEGAND TO ELAST BENERATED IN WITH PIES AS IS WITH THE

including more structure reduces surprise

http://pages.central.edu/emp/LintonT/classes/spring01/cryptography/letterfreq.html

http://everything2.com/title/entropy+of+English

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

other measures to infer structure and organization in nature and society

- Mutual Information
  - Amount of information about one variable that can be gained (uncertainty reduced) by observing another variable
- Information Gain (Kullback-Leibler Divergence)
  - Difference between two probability distributions *p* and *q*,
    - average number of bits per data point needed in order to represent *q* (model approximation) as it deviates from *p* ("true" or theoretical distribution)
- Transfer Entropy
  - transfer of information between two random processes in time
    - Amount of information (in bits) gained, or uncertainty lost, in knowing future values of Y, knowing the past values of X and Y.

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log_2 \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

$$IG(p(X), q(X)) = \sum_{i=1}^{n} p(x_i)\log_2 \frac{p(x_i)}{q(x_i)}$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$T_{X\rightarrow Y} = H(Y_t|Y_{t-1:t-L}) - H(Y_t|Y_{t-1:t-L}, X_{t-1:t-L})$$

BINGHAMTON UNIVERSITY STATE UNIVERSITY OF NEW YORK
rocha@binghamton.edu
casci.binghamton.edu/academics/ssie501m

other measures to infer structure and organization in nature and society

- Mutual Information
  - Amount of information about one variable that can be gained (uncertainty reduced) by observing another variable
- Information Gain (Kullback-Leibler Divergence)
  - ons *p* and *q*,
    - order to represent *q* (model approximation) as it
    -)
  - processes in time
    - ainty lost, in knowing future values of Y, knowing the past values of X and Y

Optional Readings: Golan, Amos, and John Harte. "Information theory: A foundation for complexity science." *Proceedings of the National Academy of Sciences* **119**.33 (2022): e2119089119.

James, R., and Crutchfield, J. (2017). "Multivariate Dependence beyond Shannon Information". *Entropy*, **19**(10), 531.

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) \log_2 \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

$$IG(p(X), q(X)) = \sum_{i=1}^{n} p(x_i) \log_2 \frac{p(x_i)}{q(x_i)}$$

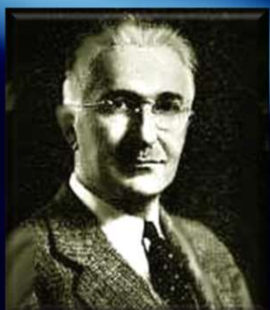
$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$$T_{X\to Y} = H(Y_t | Y_{t-1:t-L}) - H(Y_t | Y_{t-1:t-L}, X_{t-1:t-L})$$

Optional Reading: Prokopenko, Mikhail, Fabio Boschetti, and Alex J. Ryan. "An information theoretic primer on complexity, self organization, and emergence." *Complexity* **15**.1 (2009): 11-28.

uncertainty-based information

information as decrease in uncertainty .



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

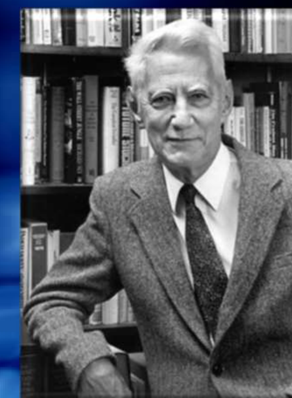
Measured in bits

Number of Choices

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

including more structure reduces surprise

information is surprise

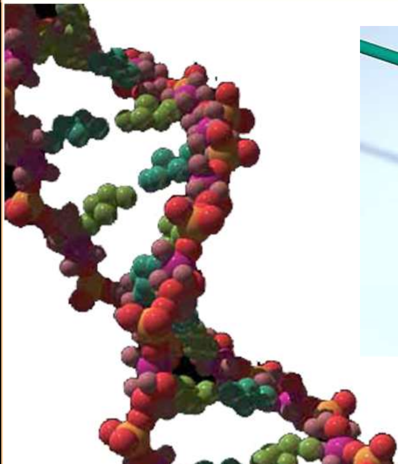


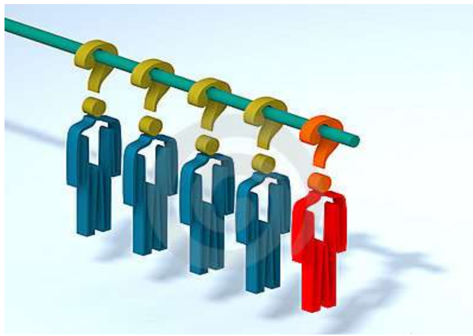
$$H_S(A) = -\sum_{i=1}^{n} p(x_i)\log_2\big(p(x_i)\big)$$

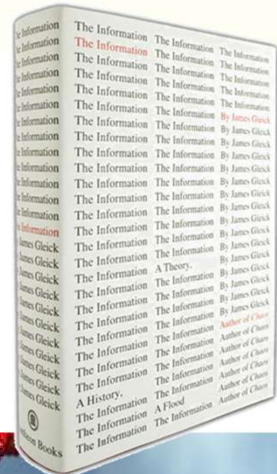
Measured in bits

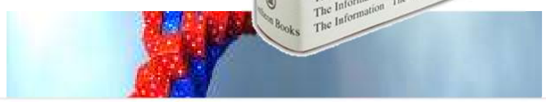
Probability of alternative

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







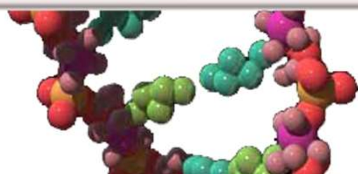


Holdin' me back
Gravity's holdin' me back
I want you to hold out the palm of your hand
Why don't we leave it at that?
Nothin' to say
When everything gets in the way
Seems you cannot be replaced
And I'm the one who will stay, oh

Optional Readings:

Prokopenko, Mikhail, Fabio Boschetti, and Alex J. Ryan. "An information theoretic primer on complexity, self organization, and emergence." *Complexity* **15**.1 (2009): 11-28.

James, R., and Crutchfield, J. (2017). "Multivariate Dependence beyond Shannon Information". *Entropy*, **19**(10), 531.



"syntactic" surprise But what about function and meaning (semantics)?