

EXPECTED SURPRISE

the basis of information theory

20 QUESTIONS

I'm thinking of a number between 1 and 1024.

You can ask "is it greater than 900?" etc.

What's your best, first question?

Second?

How many questions will you need to be sure?

0	00	000	0000
			0001
		001	0010
			0011
	01	010	0100
			0101
		011	0110
			0111
1	10	100	1000
			1001
		101	1010
			1011
	11	110	1100
			1101
		111	1110
			1111

The total symbol code budget

ERROR CORRECTION IS POSSIBLE

C O N F I D E N T I A L

TEHRAN 1691/2

"GREAT CIVILIZATION" WERE DIRECTED PERSONALLY AND APPEARED
IN HIS MINDS. SOME HEADING, IT IS CLEAR TODAY THAT THEY
DEEPENED THE RESSENTMENT OF SHI'ITE ISLAM AND BROADENED
THE OPPOSITION TO HIS REGIME AMONG MORE AND MORE CLASSES
OF PEOPLE WELL BEYOND THOSE STRICTLY PRACTICING RELIGION.
TODAY, EVEN THE SHAH'S USE OF THE PERSIAN LANGUAGE IS
RIDICULED. IT SEEMS THE SHAH REGULARLY MAKES ERRORS IN
HIS USE OF THE LANGUAGE, WHICH ONLY REINFORCES THE CHARGE
THAT HE IS AN IMPLANT OF FOREIGN INTERESTS SERVING
FOREIGNERS RATHER THAN HIS OWN PEOPLE.

ERROR CORRECTION IS POSSIBLE

SEP 79 13 05z

SECRET 0125Z SEP 79 STAFF
HONG KONG 6279
IMMEDIATE DIRECTOR INFO PRIOR TO TEHRAN, TOKYO, BANGKOK.
L. VBAT AJAJI INTEL
A. DIRECTOR 505513
B. TOKYO 8622
SECRET CONFUSION IN LOCATION "KIBISI" WHICH WAS PLACE NAME
IN GRUND MAP AS SUBJECT REFS INDICATED AREA TO BE AVOIDED BY
PSI QLD PROOF SUBJECT CLAIMED SEPT 14 AT GROUND ASSAULT
IA / QLD F. AVOID ENTIRE BORDER REGION FROM MANDALA SOUTH TO THE
L. TOKYO GROUND FORCE INFANTRY DIVISION WOULD TAKE PLACE NORTH OF
L. TOKYO GROUND MOSTLY NO INTENSIVE TERRAIN
11-120-17 5 W FILE PPS DRW D9C J

"...the Iranians laid the shreds out on a floor and devised a sophisticated procedure for numbering, indexing and reassembling the individual shreds. Certainly it took a number of years for them to finish the process," Mr Byrne says.

<https://www.bbc.com/news/magazine-16036967>

https://en.wikipedia.org/wiki/Iran_hostage_crisis#/media/File:Shredded_1979-09-01_1305Z_CIA_cable_from_American_Embassy_Tehran.jpg

ALL JBULEMD UP

“fi yuo cna raed tihs, yuo hvae a sgtrane mnid too. Cna yuo raed tihs?
OLNY 55% of pliope can...

I cdnuolt bleveiee taht I cloud aulacity uesdnatnrd waht I was rdanieg.
The phaonmneal pweor of the hmuan mnid, aoccdrnig to a rscheearch at
Cmabrigde Uinervtisy, it dseno't mtaetr in waht oerdr the ltteres in a
wrod are, the olny iproamtnt tihng is taht the frsitr and lsat ltteer be
in the rghit pclae. The rset can be a taotl mses and you can sitll raed
it whotuit a pboerlm. Tihs is bcuseae the huamn mnid deos not raed ervy
lteter by istlef, but the wrod as a wlohe. Azanmig huh? yaeh, and I
awlyas tghuhot slpeling was ipmorantt!”

N ■ V ■ W ■ LS

“S ■ pp ■ s ■ y ■ ■ wr ■ t ■ ■ n ■ nt ■ r ■ st ■ ng st ■ ry.
S ■ pp ■ s ■ ■ t w ■ s th ■ b ■ st st ■ ry y ■ u h ■ d ■ v ■ r
wr ■ tt ■ n. N ■ w t ■ k ■ th ■ t spl ■ nd ■ d st ■ ry ■ nd
■ r ■ s ■ ■ ll th ■ v ■ w ■ ls. ■ r ■ s ■ th ■ m ■ ll ■ nt ■ l
th ■ st ■ ry m ■ k ■ s n ■ s ■ ns ■ b ■ c ■ us ■ , w ■ ll, wh ■ t is
■ w ■ rld w ■ th ■ ■ t v ■ w ■ ls? I s ■ pp ■ s ■ th ■ r ■ ■ r ■
s ■ m ■ w ■ rds w ■ th ■ ■ t ■ ny v ■ w ■ l... B ■ t ■ st ■ ry
w ■ th ■ ■ t v ■ w ■ ls w ■ ld b ■ c ■ nf ■ s ■ ng,
p ■ rpl ■ x ■ ng, c ■ mpl ■ t ■ ly b ■ ffl ■ ng. W ■ ld y ■ ■ b ■
■ bl ■ to r ■ rd ■ t? W ■ ld y ■ ■ b ■ ■ bl ■ t ■
c ■ mpr ■ h ■ nd wh ■ t th ■ st ■ ry w ■ s s ■ y ■ ng?”

N VWLS

“Spps y wrt n ntrstng stry. Spps t ws th bst stry y hd vr wrttn. Nw tk tht splndd stry nd rs ll th vwls. rs thm ll ntl th stry mks n sns bcs, wll, wht s wrd wtht vwl? spps thr r sm wrds wth n vwl... Bt stry wtht vwls wld b cnfsng, prplxng, cmpltly bfflng. Wld y b bl t rd t? Wld y b bl t cmprhnd wht th stry ws syng?”

REDUNDANCY → PREDICTABILITY → COMPRESSION

- if an encoding of any kind contains *redundancy*, then we can predict parts of it from other parts
- if you can predict part of a message, you don't need to actually send it
- we can compress information in this way
- the *compressed* encoding is going to **look random!**
 - can you see why?

Horace Barlow, 1921-2020, neuroscientist who argued we should view brains in this way



QUANTIFYING SURPRISE

ASCII encodes each letter in 8 bits (a “byte”).
The actual information conveyed by letters in English is less than 2 bits per character.

What does that (the word “information”) even mean?

how surprised are you at event **A** (happens with prob 1 in 1000)
compared to event **B**? (happens with prob 1 in 5)

It depends on the probabilities.

Lower prob \Rightarrow more surprise.

QUANTIFYING SURPRISE

Now, how surprised are you
at a compound event,

A and **B** ?



LOGS!

an
event

“surprise” is the
sound of doubt
collapsing: ie.
the gaining of
information

with independent events
the probabilities **MULTIPLY**

but total information
should **ADD...**

surprise(X) = log(prob(X)) ? (no: it needs to go *down* with prob)

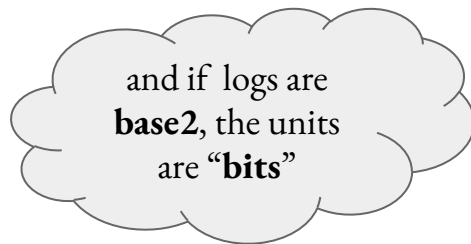
$$\text{surprise}(X) = \log \frac{1}{\text{Pr}(X)} \quad \checkmark$$

remind yourself about logs
<https://www.desmos.com/calculator>

A DEFINITION OF SURPRISE

(AND INFORMATION)

$$\text{surprise}(X) = \log \frac{1}{\text{Pr}(X)}$$



- “impossible” event \rightarrow infinitely surprising
- totally predictable event \rightarrow no surprise

Notice:

It depends on who you are – what you expected...

ENTROPY

e.g. Say I'm sending you a message, in the abcde... alphabet.

The “events” are arrivals of letters, one by one.

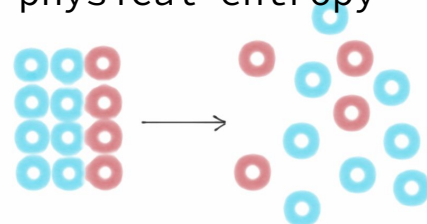
As the letters arrive, how surprised are you, on average?

Information Entropy:

$$\sum_{\text{events } X} \Pr(X) \log \underbrace{\frac{1}{\Pr(X)}}_{\text{surprise}}$$

statistics folk say
“in expectation”

cf. physical entropy



<https://fs.blog/entropy/>

Claude Shannon
(1916-2001)



ENTROPY IS "EXPECTED SURPRISE"

Shannon proved that this is the number of bits (per character) of the *optimal* code for sending information.

Information Entropy:

$$\sum_{\text{events } X} \Pr(X) \underbrace{\log \frac{1}{\Pr(X)}}_{\text{surprise}}$$

"...Among the great engineers of the 20th century, who contributed the most to our 21st-century technologies? I say: [Claude Shannon](https://spectrum.ieee.org/claude-shannon-in-formation-theory). " - <https://spectrum.ieee.org/claude-shannon-in-formation-theory>

Shannon's original paper, which created the field of Information Theory: <https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>