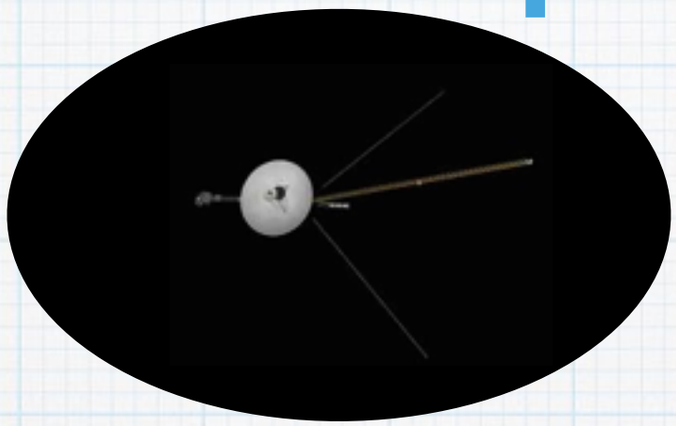
Entropy is Your Friend

Information Content

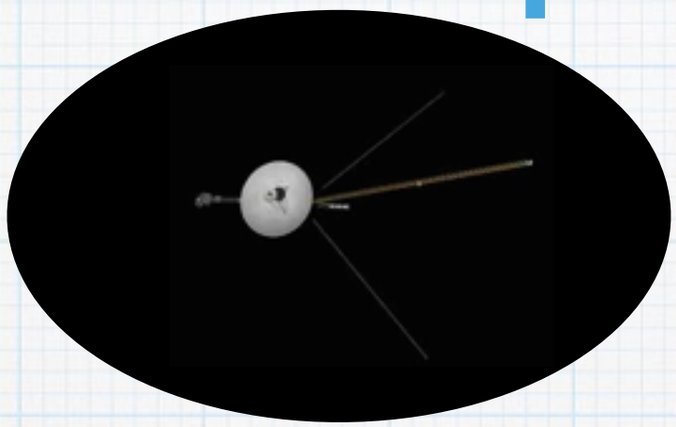
* How much “real” information in a message
* If you think about receiving a message - how surprised are you by it

Example



No Aliens	00
Friendly Aliens	01
Hostile Aliens	10
Aloof Aliens	11

Example



No Aliens	00	0
Friendly Aliens	01	100
Hostile Aliens	10	101
Aloof Aliens	11	110

Entropy

* Information for a series of messages

* $- \Sigma P(x_i) \mathrm{Log}_n P(x_i)$

Message	Code	Probability	$Log_2P(x)$	$P(x)Log_2P(x)$
No Aliens	00	0.85	-0.234	-0.199
Friendly Aliens	01	0.05	-4.321	-0.216
Hostile Aliens	10	0.05	-4.321	-0.216
Aloof Aliens	11	0.05	-4.321	-0.216

$$P(x_i)Log_nP(x_i) = 0.847$$

Bits per **Message**

Message	Code	Probability	Bits	P(x) * Bits
No Aliens	0	0.85	1	0.85
Friendly Aliens	100	0.05	3	0.15
Hostile Aliens	101	0.05	3	0.15
Aloof Aliens	110	0.05	3	0.15

Mean bits / day = 1.3

Toss Two Coins

Message	Code	Probability	$Log_2P(x)$	$P(x)Log_2P(x)$
Two Heads	00	0.25	-2	-0.5
Head Tail	01	0.25	-2	-0.5
Tail Head	10	0.25	-2	-0.5
Tail Tail	11	0.25	-2	-0.5

$$P(x_i)Log_nP(x_i) = 2$$

Bits per **Message**

Not Just Bits

* Log_{10} gives Digits
* Log_{26} shows entropy per letter
* $\text{Log}_{?}$ for words in a language

Your Friend Because

* Entropy in passwords is what makes them hard to guess
* Understanding entropy is key to compression
* Useful in encryption to understand unpredictability and redundancy