Week 4: Information Theory

Jay Daigle

Occidental College

September 20, 2018

Kerckhoffs's Principle

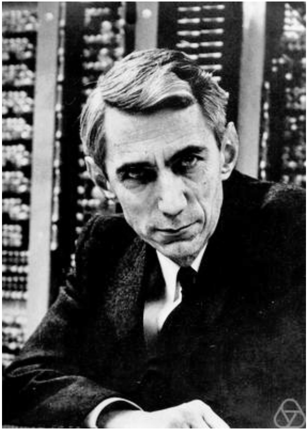
A system for encryption "should not require secrecy, and it should not be a problem if it falls into enemy hands."

Kerckhoffs's Principle

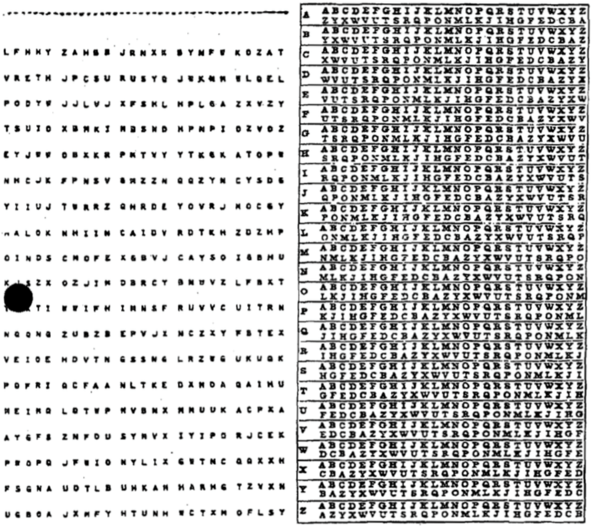
A system for encryption "should not require secrecy, and it should not be a problem if it falls into enemy hands."

Shannon's Maxim

"The enemy knows the system."



Claude Shannon



A one-time pad setup used by the NSA, codenamed DIANA.

Definition

Let X be a random variable that takes on finitely many possible values $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$. Then the entropy of X is given by

$$H(X) = H(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

(adopting the convention that if $p = 0$ then $p \log_2 p = 0$).

Definition

Let X be a random variable that takes on finitely many possible values $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$. Then the entropy of X is given by

$$H(X) = H(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

(adopting the convention that if $p = 0$ then $p \log_2 p = 0$).

Proposition (Shannon)

1. *H is continuous in each variable.*
2. *If X_n is a random variable uniformly distributed over n possibilities, then $H(X_n)$ is monotonically increasing as a function of X.*
3. *If X can be broken down into consecutive subchoices, then $H(X)$ is a weighted sum of H for the successive choices.*

Definition

Let X be a random variable that takes on finitely many possible values $x_1, \ldots, x_n$ with probabilities $p_1, \ldots, p_n$. Then the entropy of X is given by

$$H(X) = H(p_1, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

(adopting the convention that if $p = 0$ then $p \log_2 p = 0$).

Proposition (Shannon)

1. *H is continuous in each variable.*
2. *If X_n is a random variable uniformly distributed over n possibilities, then $H(X_n)$ is monotonically increasing as a function of X.*
3. *If X can be broken down into consecutive subchoices, then $H(X)$ is a weighted sum of H for the successive choices.*

Further, any function with these three properties is a constant multiple of H.

Aoccdrnig to rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit a porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.