Information and Entropy

September 6, 2013

- **Amount of Information I**

There are lots of messages, for the receiver, each message has it's own probability for being received

For a message m that the probability for being received as P, the amount of information it contained is

$$I = -\log_2 P = \log_2 \frac{1}{P}$$

This is a "surprisal" definition : **More information is being conveyed for a message with lower probability being correctly received.** That means, a highly improbable outcome is very surprising

The log base is 2, so the unit is bit

For example, *there will be rain in the desert tomorrow* contains A LOT of information since this statement is a BIG surprise (if this event really happened)

For a group of messages $\{m_1, m_2, ..., m_n\}$ with probability for being received as $\{P_1, P_2, ..., P_n\}$, then the information content of message k and h are

$$I_k = \log_2 \frac{1}{P_k} \qquad I_h = \log_2 \frac{1}{P_h}$$

And the amout of information when both messages are received is

$$\begin{aligned} I_{h,k} &= -\log_2 (P_{h,k}) \\ &= -\log_2 (P_h P_k) \\ &= -\log_2 P_h - \log_2 P_k \\ &= \log_2 \frac{1}{P_h} + \log_2 \frac{1}{P_k} \\ &= I_h + I_k \end{aligned}$$

Thus

$$I_{1,2,...,n} = \sum_{i=1}^{n} I_i$$

- **Entropy**

For example, there are totally N message being sent

On average, there should be NP_1 message m_1 being sent , NP_2 message m_2 being sent.

Therefore the total amount of information transmitted is

$$
\begin{aligned}
I_{Tot} &= NP_1I_1 + NP_2I_2 + ...NP_nI_n \\
&= NP_1 \log_2 \frac{1}{P_1} + NP_2 \log_2 \frac{1}{P_2} + ... + NP_n \log_2 \frac{1}{P_n} \\
&= N\left(P_1 \log_2 \frac{1}{P_1} + P_2 \log_2 \frac{1}{P_2} + ... + P_n \log_2 \frac{1}{P_n}\right) \\
&= N \sum_{i=1}^{n} P_i \log_2 \frac{1}{P_i}
\end{aligned}
$$

The **average information content** , called **entropy** H **,** is defined by

$$
H = \frac{I_{Tot}}{N} = \sum_{i=1}^{n} P_i \log_2 \frac{1}{P_i} = \sum_{i=1}^{n} P_i I_i
$$

The entropy is a measure of "average information" per message interval , for example , if $H = 4$ and 5 messages are being received by the receiver, then the "expected" amount of total information is thus

$$
I_{Tot} = HN = 4 \cdot 5 = 20
$$

- **Information Rate**

The information rate is defined by the average amount of message per second times average entropy

$$
R = rH
$$

The information rate can be used to judge the coding scheme is efficient or not.

For example, there are 2 event, A, B with $P_A = 0.8$, $P_B = 0.2$

Then $H = P_AI_A + P_BI_B = 0.8 \log_2 \frac{1}{0.8} + 0.2 \log_2 \frac{1}{0.2} = 0.72$

Suppose we receive 2 message per second , then $R = rH = 2 \cdot 0.72 = 1.44$, that is , we use 2 bit to transmitte information , but only 1.44 information is being transmitted. That's why we need some coding algorithm to increase the amount of information being transmitted.

$-END-$