CSE468
Information Conflict

Lecturer: Dr Carlo Kopp, MIEEE, MAIAA, PEng

Lecture 02

Introduction to Information Theory Concepts

Reference Sources and Bibliography

- There is an abundance of websites and publications dealing with basic information theory.
- Examples include:
  - http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html
  - http://okmij.org/ftp/Computation/limits-of-information.html
  - http://www.sveiby.com/articles/information.html
  - http://pespmc1.vub.ac.be/ASC/INFORMATION.html
  - http://www.mtm.ufsc.br/~taneja/book/node5.html
  - http://www.mtm.ufsc.br/~taneja/book/node6.html

Defining Information – Shannon and Weaver

- Shannon – '…that which reduces uncertainty…'
- Shannon & Weaver – 'The quantity which uniquely meets the natural requirements that one sets up for "information" turns out to be exactly that which is known in thermodynamics as entropy.'
- Shannon & Weaver - 'Information is a measure of one's freedom of choice in selecting a message. The greater this freedom of choice, the greater the information, the greater is the uncertainty that the message actually selected is some particular one. Greater freedom of choice, greater uncertainty greater information go hand in hand.' (Sveiby 1994)

Defining Information - Wiener

- Shannon and Wiener define information differently – Shannon sees it as measured by entropy, Wiener by the opposite of entropy.
- Wiener – ‘The notion of the amount of information attaches itself very naturally to a classical notion in statistical mechanics: that of entropy. Just as the amount of information in a system is a measure of its degree of organisation, so the entropy of a system is a measure of its degree of disorganisation.’ (Sveiby 1994)

Wiener Continued

- ‘One of the simplest, most unitary forms of information is the recording of choice between two equally probable simple alternatives, one or the other is bound to happen - a choice, for example, between heads and tails in the tossing of a coin. We shall call a single choice of this sort a decision. If we then ask for the amount of information in the perfectly precise measurement of a quantity known to lie between A and B, which may with uniform *a priori* probability lie anywhere in this range, we shall see that if we put A = 0 and B = 1, and represent the quantity in the binary scale (0 or 1), then the number of choices made and the consequent amount of information is infinite.’

Defining Information - Krippendorf

- Krippendorff -   - ‘Literally that which forms within, but more adequately: the equivalent of or the capacity of something to perform organizational work, the difference between two forms of organization or between two states of uncertainty before and after a message has been received, but also the degree to which one variable of a system depends on or is constrained by another. E.g., the DNA carries genetic information inasmuch as it organizes or controls the orderly growth of a living organism. A message carries information inasmuch as it conveys something not already known. The answer to a question carries information to the extent it reduces the questioner's uncertainty.’…

Krippendorf - Cont

- … 'A telephone line carries information only when the signals sent correlate with those received. Since information is linked to certain changes, differences or dependencies, it is desirable to refer to theme and distinguish between information stored, information carried, information transmitted, information required, etc. Pure and unqualified information is an unwarranted abstraction. information theory measures the quantities of all of these kinds of information in terms of bits. The larger the uncertainty removed by a message, the stronger the correlation between the input and output of a communication channel, the more detailed particular instructions are the more information is transmitter.' (Principia Cybernetica Web).

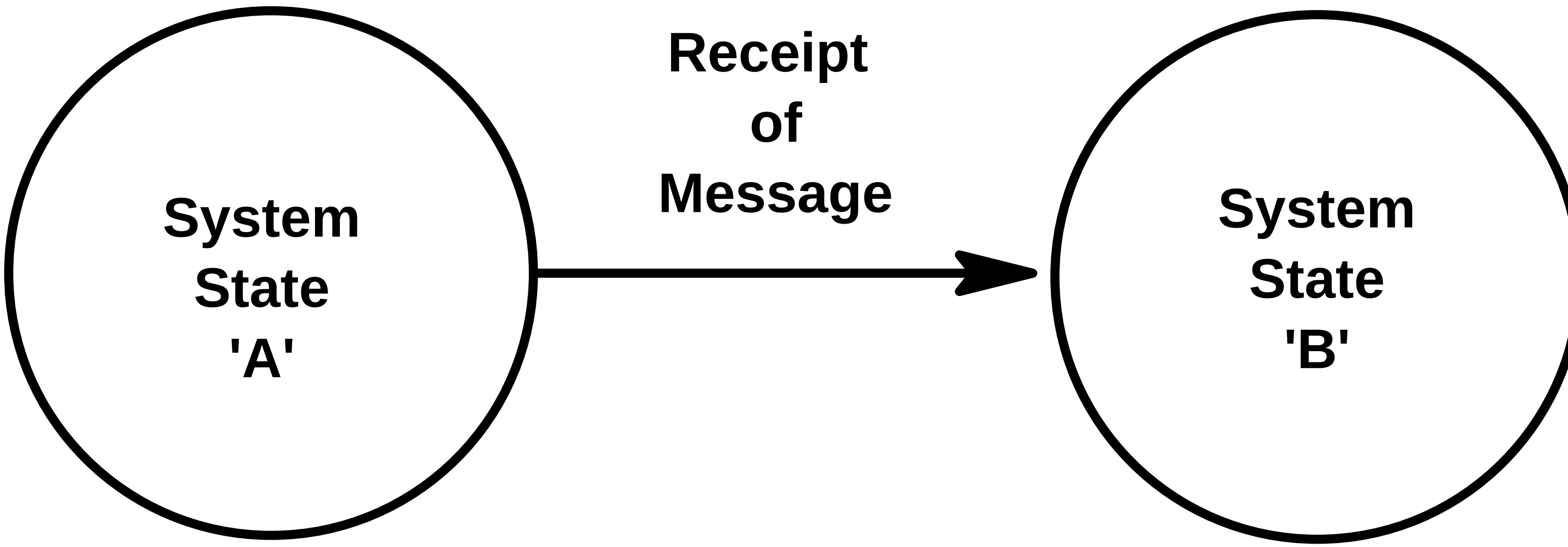
Defining Information - Hornung

- Bernd Hornung - ‘Information is the meaning of the representation of a fact (or of a message) for the receiver.’

Key Issues in Definition

- Information is a means via which the state of uncertainty in an entity can be reduced or changed.
- Entropy in thermodynamics is a measure of the state of disorder in a system; entropy in information theory is a measure of the state of disorder in an information processing system.
- For information to have effect the entity must understand the message in receives; if the message has no meaning to the entity receiving it, it cannot alter the state of uncertainty in that entity.
- If a message which is understood contains information, it will alter the system by changing the state of uncertainty.
- Information can be measured.

State Changes



Meaning in Information

- A key issue which is often not considered in definitions is the matter of *meaning* – can the message be understood?
- A work of Shakespeare written in English will be rich in information content, but only to an English speaker.
- An English speaker with good knowledge of Elizabethan English will perceive greater information content than a reader without; a reader with better knowledge of period history will perceive greater information content than a reader without; and so on …
- *In mathematical terms, the receiver of the message must be capable of decoding the message, to determine what information it contains.*

Example

- DNA encodes the definition of an organism's structure and function.
- Alter the DNA chain of an embryo and the resulting organism will be different, possibly in many ways.
- Does this mean that we can splice DNA in any manner we choose?
- For the DNA to be properly decoded, it needs to be inside a biological entity which can process (understand) what the DNA tells it to do.
- If the species between which the DNA is being spliced are too different, the DNA is not likely to be decoded in the manner intended, resulting in a non-viable organism.

Shannon's Entropy

- Shannon defines entropy as a measure of uncertainty, where H is entropy, and p_i is the probability of a symbol or message (Theorem 2):

$$H(X) = -\sum_{i=0}^{N-1} p_i \log_2 p_i$$

- The logarithm is base 2.
- Shannon's proof is well worth reading – refer:
  'Properties of Shannon's Entropy'
- This is based on the paper 'A Mathematical Theory of Communication'

Thermodynamics - Entropy

- The second law of thermodynamics ie '*the total entropy of any isolated thermodynamic system tends to increase over time, approaching a maximum value*' is often represented as:

$$S = -k_B \sum_i p_i \log_2 p_i$$

- Where k_B is Boltzmann's constant or $k = 1.3806505 \times 10^{-23}$ [joules/kelvin].
- This form is for all intents and purposes the same as that proven by Shannon in the Entropy Theorem.

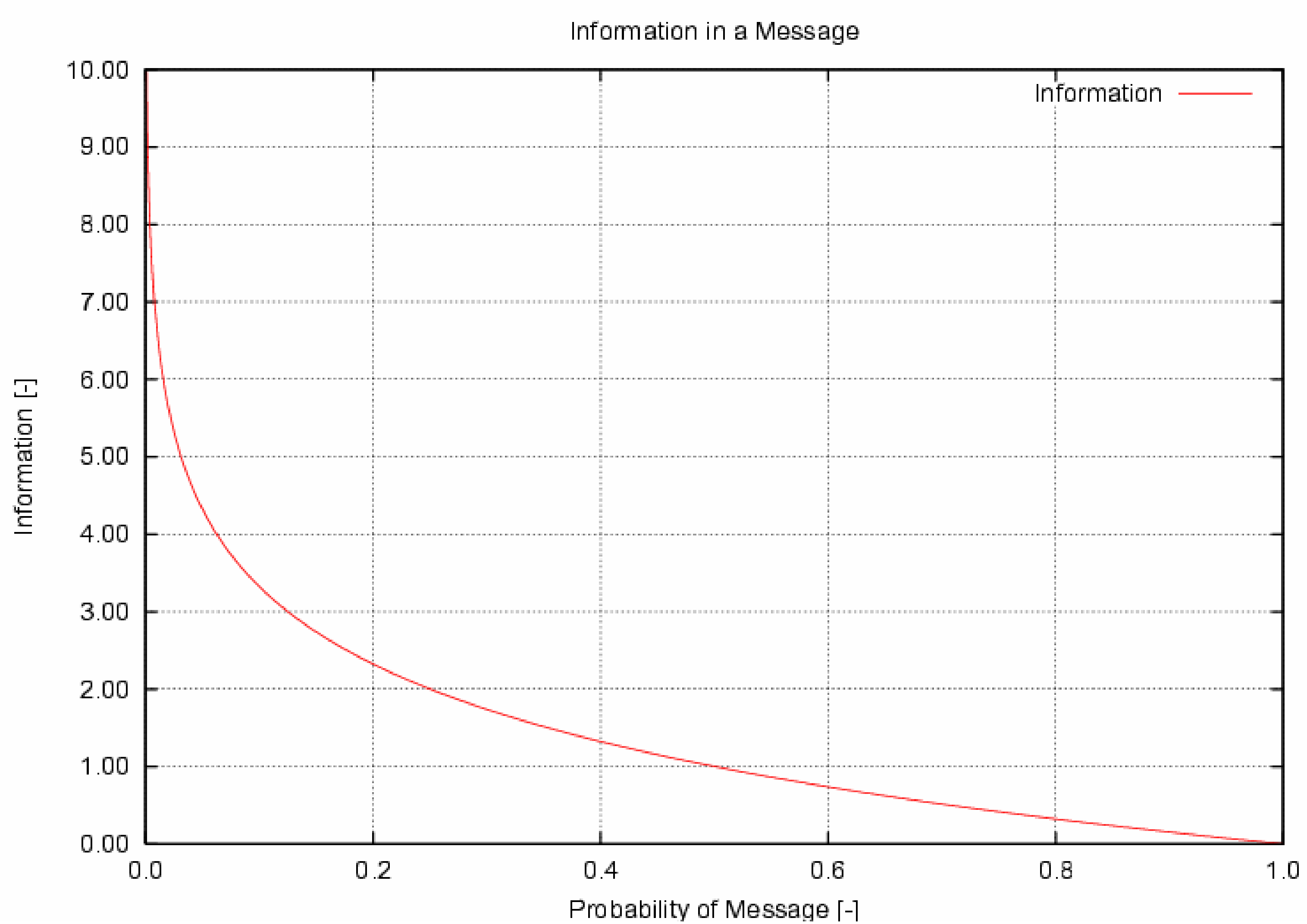
Information in a Message

- The Entropy Theorem presents the total entropy (information) in the system, across all of the possible messages in the system.
- For many problems we are interested in the information in one of the N messages. This can be represented thus:

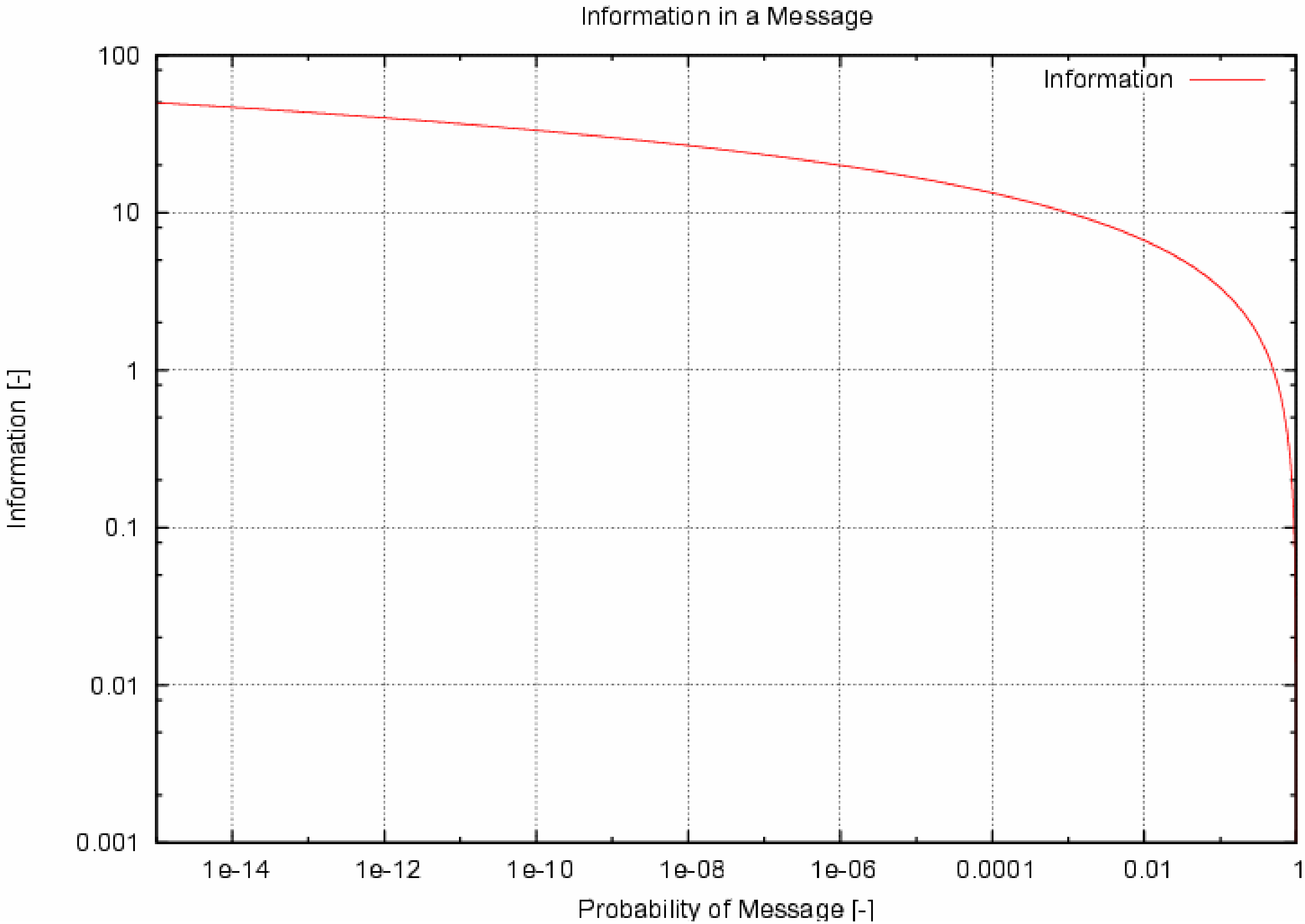
$$I(m) = -\log_2(p(m))$$

- As is evident, highly probable messages contain little information and vice versa.

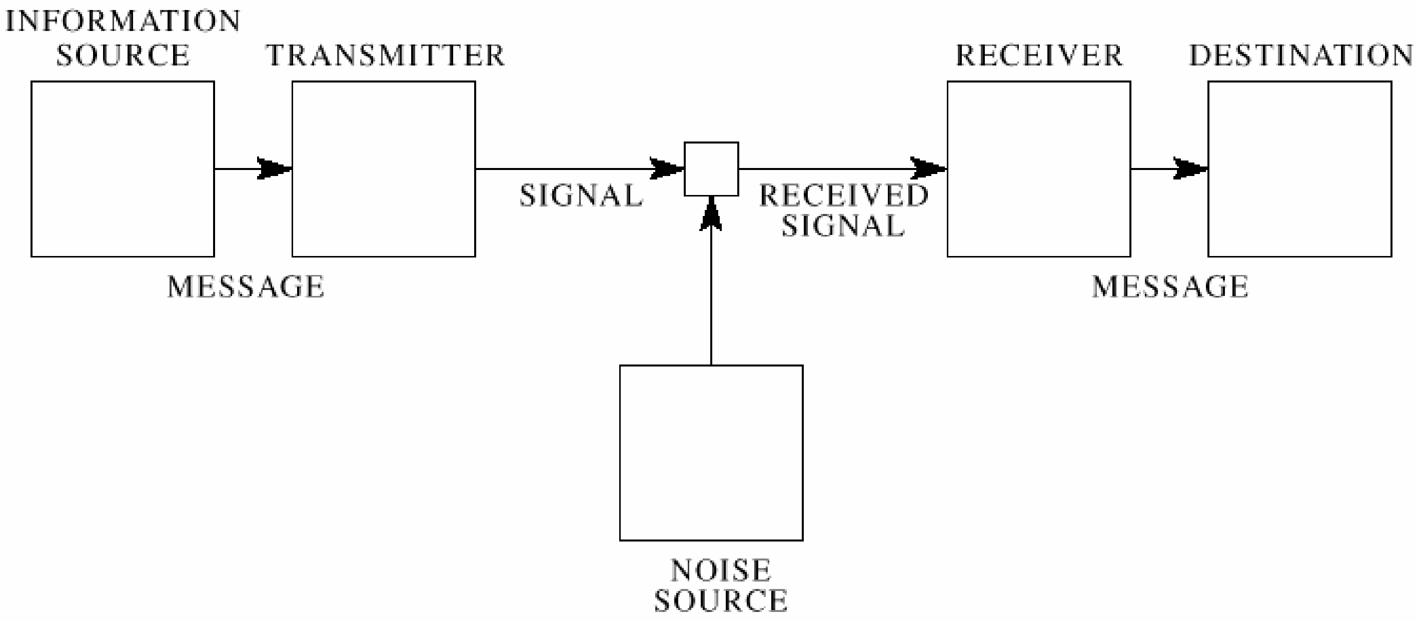
Information vs Likelihood of Message



Information vs Likelihood of Message



Shannon’s Channel Capacity Model (1)



Shannon’s Channel Capacity Model (2)

- The model have five key components:
- The ‘information source’ which generates messages containing information
- The ‘transmitter’ which sends messages over the ‘channel’.
- The ‘channel’ and associated ‘noise source’, this could be any number of physical channel types including copper or optical cable, radio link or acoustic channel.
- The ‘receiver’ which detects and demodulates messages received over the ‘channel’.
- The ‘destination’ or ‘information sink’ which responds to messages by changing its internal state.
- It is implicitly assumed that messages sent by the ‘information source’ can be understood by the ‘sink’.

Shannon’s Channel Capacity Model (3)

- Shannon demonstrated that for a ‘noisy’ channel, ie one in which random noise could additively contaminate the messages flowing across the channel, the capacity of the channel (amount of information it could carry) is defined by (Theorem 17):

$$C = W \log_2\left(\frac{(P+N)}{N}\right)$$

- Where *C* is channel capacity, *W* is channel bandwidth, *P* is signal power, and *N* is noise power.
- This equation is most commonly used in the following form, as *P/N* is the widely used measure of ‘signal to noise ratio’ or ‘SNR’:

$$C = W \log_2\left(1 + \frac{P}{N}\right)$$

Shannon's Channel Capacity Model (4)

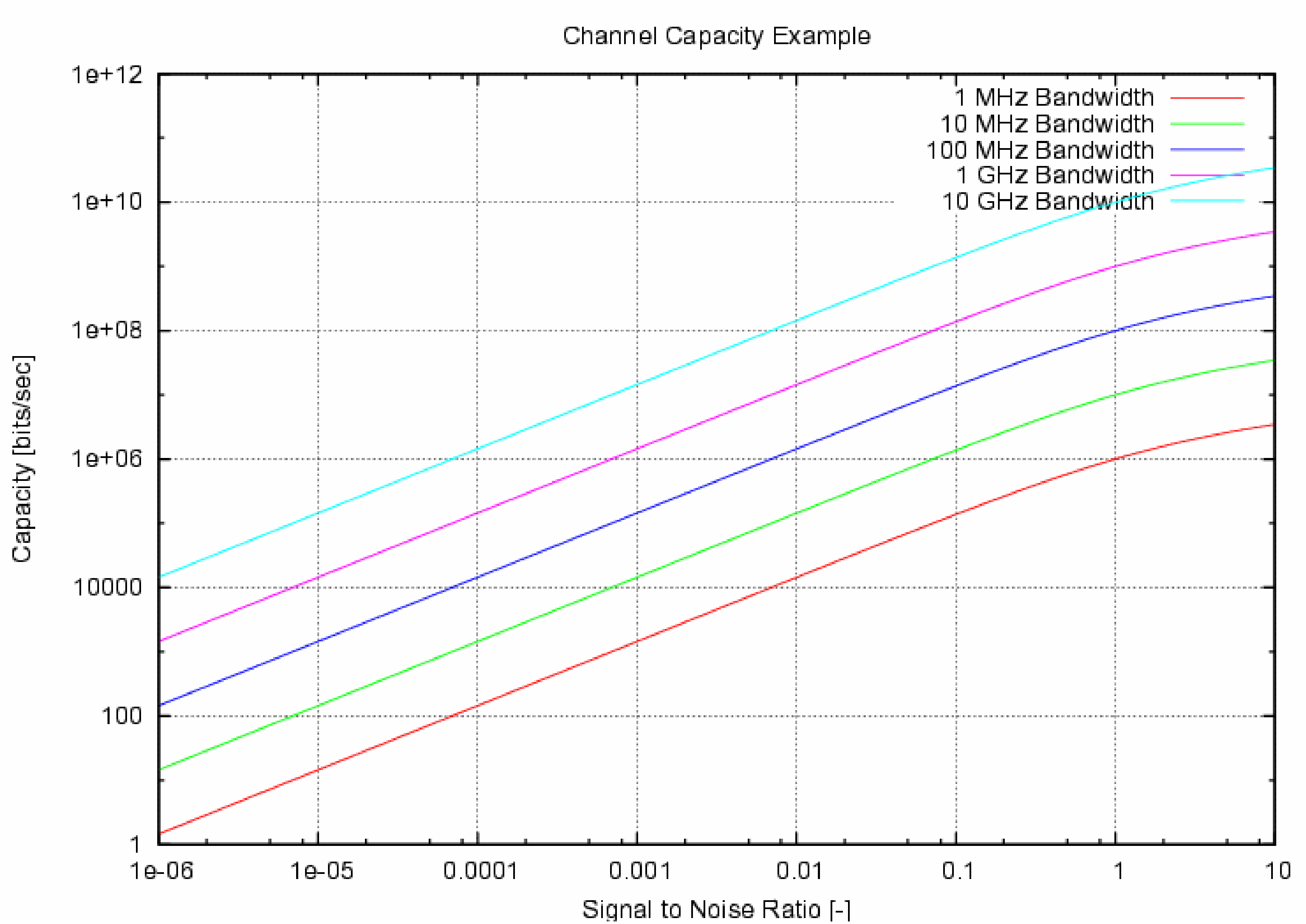
- Assumption (1) – the additive noise is 'white thermal noise' ie it has a normal distribution in time/power;
- Assumption (2) – the power of the signal (message) in the channel is the average, rather than peak power;
- Assumption (3) – the power is not limited by the transmitter's peak power rating;
- Metrics (1) – channel capacity is defined in bits/sec;
- Metrics (2) – bandwidth is defined in Hertz (cycles/sec);
- Metrics (3) – signal and noise power are defined in Watts;
- In numerical applications which compute capacity, it is customary to use this form as $\log_2$ is often not available:

```
C=B*(1/log(2))*log( 1 + S/N);
```

What Does This Model Tell Us?

- It is possible to trade between bandwidth and signal to noise ratio to achieve an intended capacity – an example is in spread spectrum communications;
- Channels with severely limited bandwidth but very high signal to noise ratio can still achieve high capacity – example are voiceband modems running over digital switches;
- Where the SNR >> 1, the second term approximates the logarithm of SNR;
- Where the SNR << 1, the second term -> 0, and bandwidth becomes the dominant means of improving capacity;
- *By manipulating the bandwidth and SNR of a channel we can manipulate its capacity, and thus how much information it can carry.*

Channel Capacity Example



Key Points

- What constitutes information in a message depends on the ability of an entity to understand that information.
- If a message contains information, an entity receiving it and understanding it will experience a state change which alters its level of uncertainty.
- The less likely the message, the greater its information content (Entropy theorem).
- The capacity of a channel to carry information depends on the magnitude of interfering noise in the channel, and the bandwidth of the channel (Capacity theorem).

Tutorial

- Discussion and debate
- Applications of Shannon
- (1) review Entropy proof
- (2) review Capacity proof