Information in Communication Systems:
Redundancy, Source Coding and Noise

Notes

Chapter Contents

Notes

Redundancy

- N symbol vocabulary has maximum possible entropy $\log_2(N)$ bits.
- Normally entropy is less than maximum entropy.
  - $\Rightarrow$ built-in **Redundancy**:

$$\text{Redundancy} = 1 - \frac{\text{actual entropy}}{\text{maximum entropy}}.$$

Notes

Redundancy Example

Notes

A source has 5 symbols, $U = \{q, w, e, r, t\}$ with probabilities:

symbol	q	w	e	r	t
probabilities	0.3	0.25	0.2	0.15	0.1

Solution

- Maximum Entropy= $\log_2(5) = 2.32$ bits.
- Actual entropy= $0.3 \times 1.74 + 0.2 \times 2 + 0.2 \times 2.32 + 0.15 \times 2.74 + 0.1 \times 3.3 = 2.23$
- $\therefore$ Redundancy=$1 - \frac{2.23}{2.32} = 0.04 \approx 4\%$.
- Probabilities close to uniform distribution
  - $\Rightarrow$ low redundancy.

Redundancy

Notes

Redundancy can be useful

- Redundancy can help to overcome errors.

Redundancy increases amount of information sent or stored.

- Parity bit (addition of single bit)
  - Makes code 89% efficient.
  - And able to detect single bit error for every byte
- Hamming code $(4, 7)$
  - 43% redundancy.
  - Corrects single error in group of 4 bits

Source or Entropy Coding

Notes

- Some signals might be **highly redundant** before any coding:
  - Signal can be converted into another less redundant form.
  - Process known as **Source Coding**.

Very old example of source coding: **Morse Code**

- *Short codes for common letters*, e.g. E is

  ●

- *Long codes for uncommon letters*, e.g. Z is

  —— —— ● ●

Source Coding: Code Capacity

Code Capacity=Amount of information code can carry per source symbol

(on average)

Code Capacity=
average number of string symbols per symbol $\times$ how much information individual string symbol can carry

---

Efficiency of coding scheme:

$$\eta = \frac{\text{source entropy}}{\text{code capacity}}$$

Notes

Source Code Example

A source uses 5 symbols $\{q, w, e, r, t\}$ with probabilities and codes:

symbol	q	w	e	r	t
probability	0.3	0.25	0.2	0.15	0.1
codes	1111	1110	110	10	0

Determine efficiency of the code.

Solution

Average string length =

$$0.3 \times (4 \text{ digits }) + 0.25 \times (4 \text{ digits }) + 0.2 \times (3 \text{ digits }) + 0.15 \times (2 \text{ digits }) + 0.1 \times (1 \text{ digits }) = 3.2 \text{binary digits.}$$

Maximum entropy of a binary symbol is $\log_2(2) = 1$ bit / digit

$\therefore$ Code capacity is $3.2 \times 1 = 3.2$bits per source symbol

Entropy of source is $\sum_{k \in U} p_k h_k = 2.2228$ bits.

$\therefore$ Efficiency is:

$$\frac{2.228}{3.2 \times 1} = 69.6\% \text{ efficient}$$

Equivalent to 30.4% redundancy (not very efficient).

Notes

Source Code Example II

Modify the previous source to use shorter codes for more likely symbols, i.e.

symbol	q	w	e	r	t
probability	0.3	0.25	0.2	0.15	0.1
codes	00	01	11	100	101

Determine efficiency of the code.

Solution

Average string length =

$$0.3 \times (2 \text{ digits }) + 0.25 \times (2 \text{ digits }) + 0.2 \times (2 \text{ digits }) + 0.15 \times (3 \text{ digits }) + 0.1 \times (3 \text{ digits }) = 2.25 \text{binary digits.}$$

Maximum entropy of a binary symbol is $\log_2(2) = 1$ bit / digit

$\therefore$ Code capacity is $2.25 \times 1 = 2.25$bits per source symbol

Entropy of source is $\sum_{k \in U} p_k h_k = 2.2228$ bits.

$\therefore$ Efficiency is:

$$\frac{2.228}{2.25 \times 1} = 99.0\% \text{ efficient}$$

Equivalent to 1.0% redundancy (much more efficient than previous codes).

Notes

Limiting Effects of Noise[1]

Noise and bandwidth not considered so far...

- Signals are usually affected by **noise** N
- Signal has power S and bandwidth B Hz then,

Channel Capacity upper limit on bits per second that can be sent:

$$C = B \log_2 \left(1 + \frac{S}{N}\right) \quad \text{bits/ second.}$$

[1] Shannon-Hartley channel capacity theorem: maximum rate of information transmission over a channel with bandwidth B and SNR=S/N in bits/s.

Notes

Shannon-Hartley Channel Capacity

As B increases, the SNR= S/N can decrease.

$$C = B \log_2 \left(1 + \frac{S}{N}\right) \quad \text{bits/ second.}$$

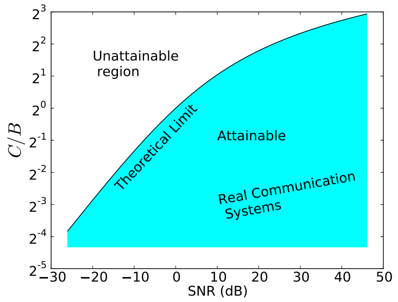
Spectral Efficiency

$$\frac{C}{B} = \log_2 \left(1 + \frac{S}{N}\right) \quad \text{bits/ second/ Hz}$$

Useful benchmark - but very difficult to realise in practice because it is a theoretical optimal value.

Notes

Limiting Effects of Noise



$$\frac{C}{B} = \log_2(1 + \text{SNR})\text{bits/second/Hz.}$$

Notes

Channel Capacity Example

Example
An audio channel has 4kHz bandwidth and 26dB SNR what is the upper limit of the channel capacity and the spectral efficiency?

Solution

Convert SNR $= 10^{(26/10)} = 398$.

Channel capacity:

$$C = 4000 \times \log_2(1 + 398) = 34.6\text{kbits/s}.$$

Spectral efficiency is $C/4$kHz$= 8$ bits/s/Hz.

Notes

Summary

- Information is communicated by communication systems.
- Smallest unit of information is a bit:
  - 1 and 0
  - YES and NO
  - ON and OFF
- Self information measures amount of information for one symbol
- Entropy measures amount of information across all symbols:
  - Entropy is the expectation of the self informatoin for all symbols
- We also looked at maximum Entropy, efficiency, redundancy and capacity of communication systems.

Notes

Notes