Richard W. Hamming

Learning to Learn

The Art of Doing Science and Engineering

Session 10: Coding Theory I

Coding Theory Overview

Problem – Representation of Information

Standard Model of Information Transmission

- Analogy to Human Communication

Representation – Codes

- What kind of code do we want?

What is information?

Questions

What is information?

Can we recognize it from white noise?

- Hamming: "No."

How do we communicate ideas?

Is the idea contained in words?

How do machines differ from humans with respect to information?

Information Theory

Developed by C.E. Shannon at Bell Labs

Should it have been called "Communication Theory"?

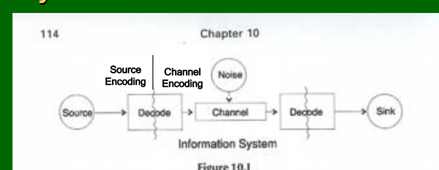


Figure 10.1

Information Theory

Nothing is assumed about source.

Communication over space is no different than communication over time.

Noise is different than Error.

- In physics theories, noise is not assumed to be part of the model.
- Error is in experiments, measurements

Human Information Theory

All you know is what comes out of the channel (hearing, reading, etc.)

How do you know what went in?

Design encoder to make this more probable

Analogy to human communication

- Thought – Words spoken – Words heard – Understanding
- Source – Encoding – Decoding – Sink

Why Information Theory

Precision of equipment is expensive for an additional bit.

- Too expensive to do it right!

Coding theory offers an alternative.

- Escape necessity of doing things exactly right, allow room for error without paying for more precision!

What codes to use?

Assume variable length codes

- Block codes (same length symbols) are a special case

Desired properties

- **Unique decodability** – each coded message represents a single source message if no noise added.
- **Instantaneous decodability** – look at each bit only once

Non-unique code

Consider the 4 symbol code:

$s_1 = 0; s_2 = 00; s_3 = 01; s_4 = 11$

If you receive:

0011

Is it $s_1 s_1 s_4$ or $s_2 s_4$?

Decoding Trees

Represents a unique and instantaneously decodable code

- No symbol is a prefix for another

Practical advice

- Have an exit symbol to know when entire decoding process done
- Like quotation marks in human language

Decoding Tree

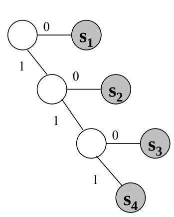
Decoding tree for:

$s_1 = 0;$

$s_2 = 10;$

$s_3 = 110;$

$s_4 = 111;$



Non-Instantaneous Code

Consider the 4 symbol code:

$s_1 = 0; s_2 = 01; s_3 = 011; s_4 = 111$

If you receive:

011111…….111

You have to count back by groups of 111 to determine if the first symbol is 0, 01, or 011, so you look at each bit more than once.

"Goodness" of Codes

McMillan's Theorem says if we have uniqueness we have instantaneousness.

- No need to settle for a code without both properties.

Average code length as reasonable "goodness" measure.

Good code design takes into account probabilities.

Average Code Length

Given an alphabet of q symbols having length l_i and probability p_i, the average code length L is:

$$L = \sum_{i=1}^{q} p_i l_i$$

Kraft Inequality

For an instantaneous code with symbol lengths l_i

$$K = \sum_{i=1}^{q} \frac{1}{2^{l_i}} \le 1$$

Holds for unique decodable codes as well (McMillan's theorem).

Proved via induction on branch lengths of decoding tree.

Proof of McMillan's Theorem (again)

Given a uniquely decodable code, the Kraft inequality holds, so the code is also instantaneously decodable.

Unique decodability requires $N_k \le 2^k$, so the following inequality must be true.

$$K^n = \left[\sum_{i=1}^{q} \frac{1}{2^{l_i}}\right]^n = \sum_{k=n}^{nl} \frac{N_k}{2^k} \le \sum_{k=n}^{nl} \frac{2^k}{2^k} = nl - n + 1$$

Prove by contradiction. Assume $K > 1$. For some large n, the exponential K^n exceeds the linear function $f(n) = nl - n + 1$

Intuitive Conclusions

If the Kraft sum is strictly less than one, you can add a shorter code, or shorten an existing code.

If it is exactly one, it is the best you can do.

Thus have another measure of a good code.

What is Information?

Words don't necessarily "contain" the idea. Evidence:

- Hamming explains an idea in lecture.
- If you go home to explain it to your spouse, you will use different words.

Words are an encoding.

We don't really know *what* is the idea behind words.

- Or if there is a single definite idea.

Is information like time?

- St. Augustine – I know what time is until you ask me.

Human Information Theory

- Information theory assumes symbols have consistent meaning.
- Human words do not have a fixed meaning.
- Lack of uniqueness is a challenge.
  - Accounts vary from different witnesses of the same event
- Once you encode an idea into words, that is what you believe.
- Coding theory is great for machines, not language.

Parting Thoughts

- Return to the AI question
  - How much of what we do (or want done) will machines do?
- Really asking “What can we program?”
- In order to program we need to understand our own ideas.
- Greatest contributions of mathematicians are often definitions.