Choose Ohio First

KENT STATE UNIVERSITY

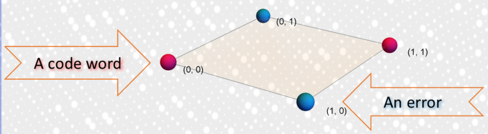
Finite Arbitrary Length Error Detecting Code Words

Wayne Fincher, Advisor: Dr. R. Aron

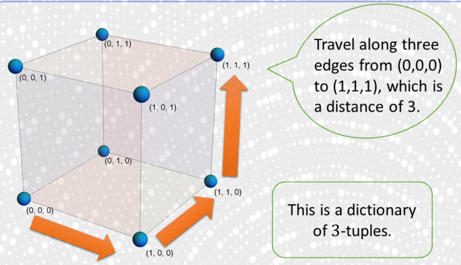
Can Errors Be Detected?

The questions we seek to answer are:

- What is the **maximum** number of **detectable code words**?
- What is the **minimum** number of **detectable code words**?
- What is the **maximum** number of **correctable code words**?



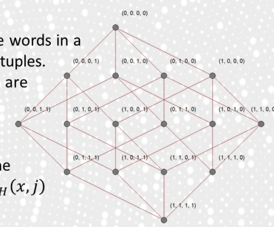
A Dictionary as a Unit Hypercube

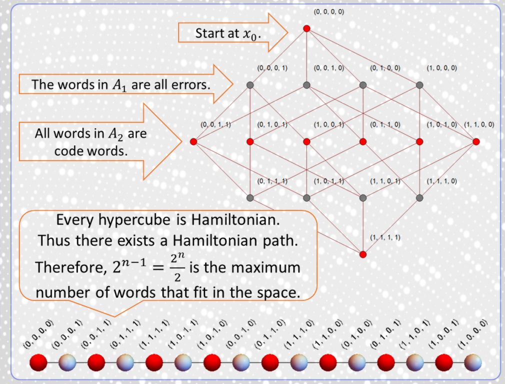


How Many Words Are In a Ball?

- The image shows all of the words in a dictionary consisting of 4-tuples.
- And it shows which words are a distance of one away.
- The pattern is binomial coefficients.
- The general formula for the number of words in the $B_H(x,j)$ where $1 \leq j \leq n$ is

$$\sum_{k=0}^{j} \binom{n}{k}.$$



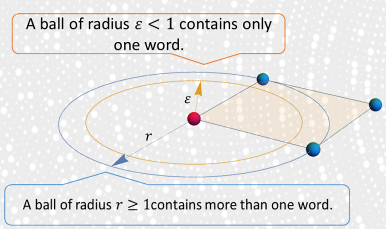


What are the code words?

- A **word** is a binary n-tuple such as $(0,1,1)$ or $(0,0,0,0)$.
- A **dictionary** is the set of all words of a given length, n. For example, a dictionary with $n = 2$ is the set
  $$\{(0,0),(0,1),(1,0),(1,1)\}.$$
- A set of **code words** form a subset of a dictionary that can be sent in a message from one person to another.
- The **distance** between two words is defined as
  $$H(x,y) = \sum_{i=1}^{n} |x_i - y_i|.$$
- The **ball** centered at x with radius r is denoted $B_H(x,r)$.

Properties of the Space

Let the DICTIONARY be called X. Then the space (X, H) is a **metric space** with the **discrete topology**.



Words in an Annulus Have Even Distances

Theorem: For any $x_0 \in X$, the set $A_j = \{y: H(x_0,y) = j\}$ does not contain any two words an odd distance apart.

Proof. Starting from x_0 words that are a distance of one from x_0, denoted A_1, each have one component that is different from x_0. Each of those words in A_1 has two components different from the other words in A_1.

Assume that the words in A_j all have an even distance from the other words in A_j. Let us consider the set A_{j+1}, and let a be an element of A_{j+1} and let a' be the element in A_j such that one component of a' is transposed to obtain a. By symmetry, we only need to consider the distance between a and the elements of $\{A_{j+1} - a\}$, which we shall denote as B.

Consider the distance between a and some element $b \in B$ such that $b' \in A_j$ and one component of b' is transposed to obtain b. If the component of a' that is transposed is the same component of b' that is transposed, then $H(a,b) = H(a',b')$. Otherwise, $H(a,b) = H(a',b') + 2$. □

Minimum Detectable Words

Lower bound: The maximum number of disjoint balls $B_H(x,1)$ that could fit into the space is

$$\left\lfloor \frac{2^n}{n+1} \right\rfloor = \left\lfloor \frac{\text{size of dictionary}}{\text{size } B_H(x_0,1)} \right\rfloor.$$

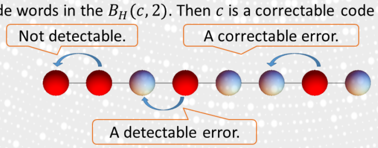
Upper bound: The set of words a distance j from x_0 is denoted A_j. Their exists a set of code words such that every word in $\bigcup_{j\geq 0}\{A_{3j}\}$ is a code word and all the remaining words are errors. This gives an upper bound of

$$\sum_{k=0}^{\lfloor n/3 \rfloor} \binom{n}{3k}$$

code words.

What are the code words (cont.)?

- An **error** is when a component of a code word that was sent in a transmission is transposed. The set of error words are the compliment of the code words with respect to the dictionary.
- An error is **detectable** when a code word c has no other code words in the $B_H(c,1)$. Then c is a detectable code word.
- An error is **correctable** when a code word c has no other code words in the $B_H(c,2)$. Then c is a correctable code word.



Symmetry of the Space

Lemma: For any $x_0 \in X$, there exists a homeomorphism $f(x): X \to X'$ that maps x_0 to the origin of X'.

Proof. Let f be defined as

$$f(x) = \begin{bmatrix} \begin{cases} x_1, & x_{0_1} = 0 \\ |x_1 - 1|, & x_{0_1} = 1 \end{cases} \\ \vdots \\ \begin{cases} x_n, & x_{0_n} = 0 \\ |x_n - 1|, & x_{0_n} = 1 \end{cases} \end{bmatrix}.$$

It is clear that $f(f(x)) = x$ for any $x \in X$. A map from one discrete space to another discrete space always maps open sets to open sets. That proves the lemma. □

Maximum Detectable Code Words

- Pick an x_0 to be the first code word.
- No words in the annulus A_1 can be a code word.
- We showed that all words in A_2 are at least a distance of two from each other. Chose all of them as code words.
- Every word in A_3 is a distance of one from some code word in A_2, so there are no code words here.
- Keep choosing every word in every other annulus.
- This gives 2^{n-1} detectable code words.
- But can we fit even more code words into the space?

No! Why? See the next slide.

Future Research

- Our future research will focus on narrowing the upper and lower bounds of the minimum detectable code words until they converge.
- We have upper and lower bounds for the maximum number of correctable code words that we could not fit onto the poster. We will attempt to improve those bounds until they converge as well.

Acknowledgements

- We thank Dr. Aron for his guidance on this project.
- Adams, Colin Conrad, and Robert David Franzosa. *Introduction to Topology: Pure and Applied*. Upper Saddle River, NJ: Pearson Prentice Hall, 2008. 150-152.