Error Correcting Codes

Stanley Ziewacz

22M:151

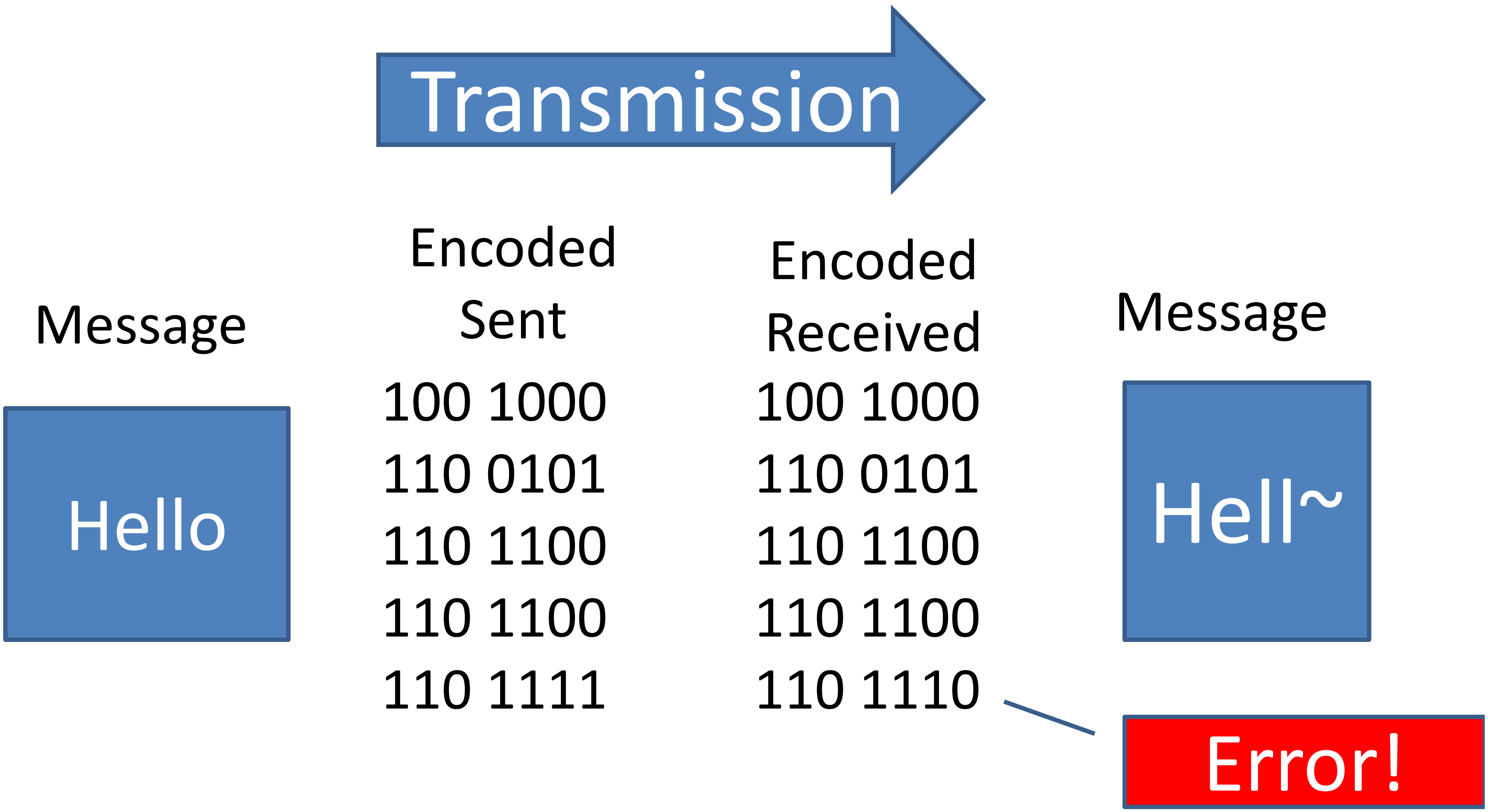
Spring 2009

Information Transmission

Transmission →

Message	Encoded Sent	Encoded Received	Message
Hello	100 1000	100 1000	Hell~
	110 0101	110 0101	
	110 1100	110 1100	
	110 1100	110 1100	
	110 1111	110 1110	

Information Transmission

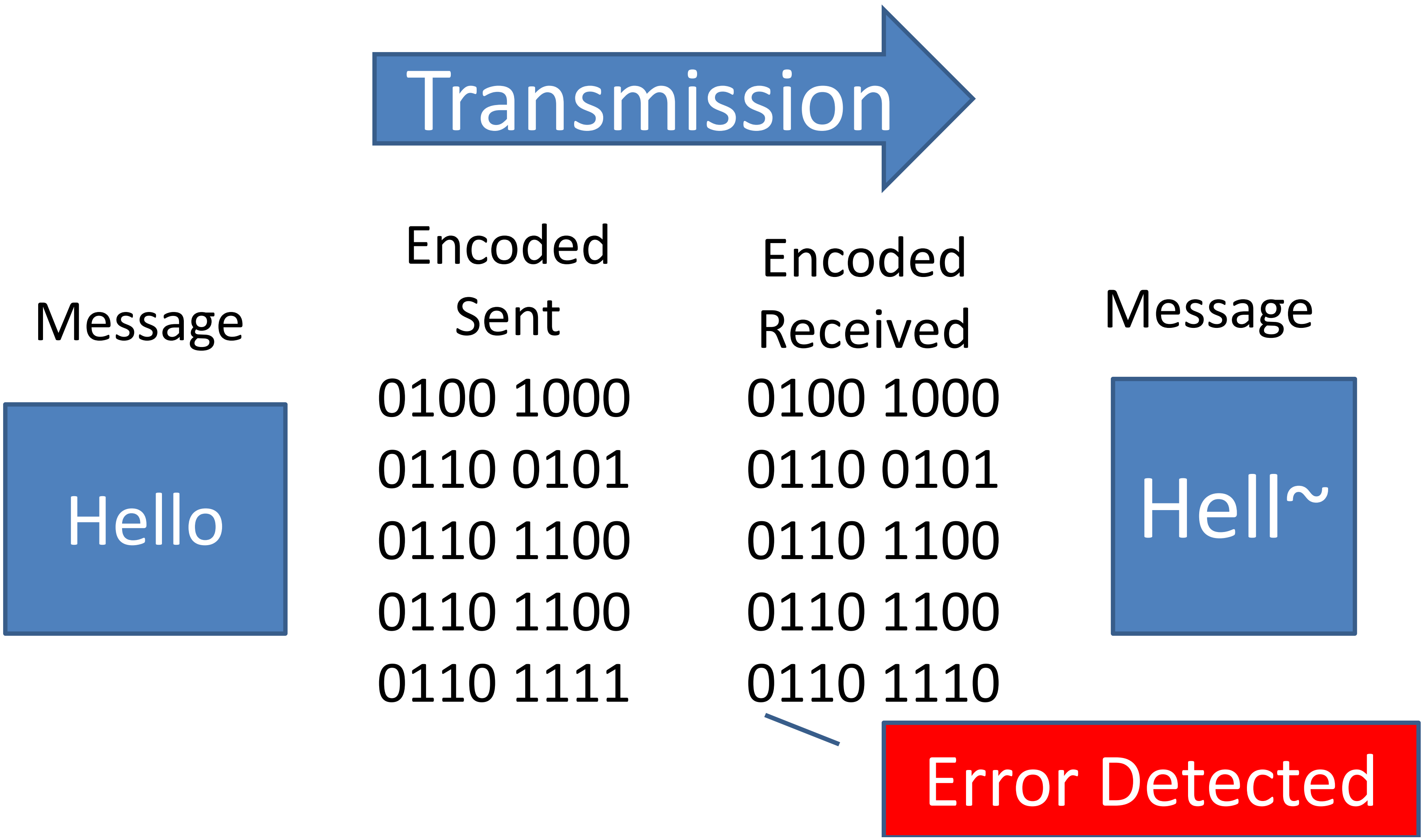


Information Transmission with Parity Bit

Transmission →

Message	Encoded Sent	Encoded Received	Message
Hello	0100 1000	0100 1000	Hell~
	0110 0101	0110 0101	
	0110 1100	0110 1100	
	0110 1100	0110 1100	
	0110 1111	0110 1110	

Information Transmission with Parity Bit



Definition of Code

Block code: all words are the same length.

A q-ary code C of length n is a set of n-character words over an alphabet of q elements.

Examples:

C_1 = {000, 111} binary code of length 3

C_2 = { 00000, 01100, 10110} binary code of length 5

C_3 = {0000, 0111, 0222, 1012, 1020, 1201, 2021, 2102, 2210} ternary code of length 4

Error Correcting Code

- An error is a change in a symbol
- Want to detect and correct up to t errors in a code word
- Basic assumptions
  - If $i < j$ then i errors are more likely than j errors
  - Errors occur randomly
  - Nearest neighbor decoding
    - Decode y to c, where c has fewer differences from y than any other codeword

Hamming Distance

- The Hamming distance between two words over the same alphabet is the number of places where the symbols differ.
- Example : d(100111, 001110) = 3
  - Look at 100111

    001110
- For a code , C, the minimum distance d(C) is defined by $d(C) = \min\{d(c_1,c_2), \mid c_1, c_2 \in C, c_1 \neq c_2\}$

Hamming Distance Properties

- Let x and y be any words over the alphabet for C; x and y may or not be codewords.
- $d(x, y) = 0$ iff $x = y$
- $d(x, y)= d(y, x)$ for all x, y
- $d(x, y) \leq d(x, z) + d(z, y)$ for all x, y, and z

Detection and Correction

- A code C can detect up to s errors in any codeword if $d(C) \geq s + 1$
- A code C can correct up to t errors if $d(C) \geq 2t + 1$
  - Suppose: c is sent and y is received, $d(c,y) \leq t$ and $(c' \neq c)$
  - Use triangle inequality $2t + 1 \leq d(c, c') \leq d(c, y) + d(y, c') \leq t + d(y,c')$

(n, M, d) q-ary code C

- Codewords are n characters long
- $d(C) = d$
- M codewords
- q characters in alphabet
- Want n as small as possible with d and M as large as possible
- These are contradictory goals

Hard Problem

Maximize the number of codewords in a q-ary code with given length n and given minimum distance d.

We'll use Latin squares to construct some codes.

(4, 9, 3) ternary code

0 0 0 0

0 1 1 1

0 2 2 2

1 0 1 2

1 1 2 0

1 2 0 1

2 0 2 1

2 1 0 2

2 2 1 0

Latin square

- A Latin square of order n is an n x n array in which n distinct symbols are arranged so that each symbol occurs once in each row and column.
- Examples:

0	1	2		0	1	2
1	2	0		2	0	1
2	0	1		1	2	0

Orthogonal Latin Squares

- Two distinct Latin squares $A = (a_{ij})$ and $B =(b_{ij})$ are orthogonal if the n x n ordered pairs (a_{ij}, b_{ij}) are all distinct.
- Example:

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \qquad \begin{pmatrix} (0,0) & (1,1) & (2,2) \\ (1,2) & (2,0) & (0,1) \\ (2,1) & (0,2) & (1,0) \end{pmatrix}$$

(4, 9, 3) ternary code constructed from orthogonal Latin squares

0 0 0 0

0 1 1 1

0 2 2 2

1 0 1 2

1 1 2 0

1 2 0 1

2 0 2 1

2 1 0 2

2 2 1 0

0	1	2
1	2	0
2	0	1

0	1	2
2	0	1
1	2	0

Theorem

- There exists a q-ary $(4, q^2, 3)$ code iff there exists a pair of orthogonal Latin squares of order q.
- Proof:

Look at the following 6 sets

$\{(i, j)\}$ $\{(i, a_{ij})\}$, $\{(i, b_{ij})\}$, $\{(j, a_{ij})\}$, $\{(j, b_{ij})\}$, $\{(a_{ij}, b_{ij})\}$

References

- Colbourn, Charles J. and Jeffrey H. Dinitz, Handbook of Combinatorial Designs, Second Edition, Chapman & Hall/CRC, Boca Raton, FL, 2007
- Laywine, Charles F. and Gary L. Mullen, Discrete Mathematics Using Latin Squares, John Wiley and Sons, New York, 1998
- Pless, Vera, Introduction to the Theory of Error-Correcting Codes, John Wiley and Sons, New York, 1982
- Roberts, Fred S. and Barry Tesman, Applied Combinatorics, 2nd Edition, Pearson Education, Upper Saddle River, NJ , 2005