Foundations of Information Systems

Winter Semester 2024–25, Exercise 10

For discussion on Wednesday, January 15, 2025

1. In a RAID-5 array, one disk out of 5 has failed. The others contain the bit sequences 10001011 . . ., 11111000 . . ., 00000000 . . ., 01011010 . . .. Reconstruct the beginning of the bit sequence of the failed disk.

2. Decimal codes regularly entered by humans, such as IBANs, use a modulo-97 checksum. Direct computation of such a checksum for long sequences requires arbitrary-length integer arithmetic, which is not directly available in hardware. As an alternative, it is possible to compute the modulo-97 checksum of the first k digits, prepend the result to the remaining digits and repeat in chunks of k digits, until the entire sequence is consumed.

   Why does this work?

   Hint: If $a \equiv a' \mod 97$, then $10^k\, a + b \equiv 10^k\, a' + b \mod 97$. Why?

3. Check if the following IBANs are valid. Show all of the required steps explicitly.

   (a) DE51 7509 0300 0007 6333 00

   (b) DE51 7509 0300 0007 3633 00

4. Write out the $(12, 8)$ Hamming code for the binary word 01010111.

5. Which of the following $(7, 4)$ Hamming codes are correct, assuming that they contain at most a single-bit error? Correct any error you detect.

   (a) 0010111

   (b) 1111111

   (c) 0110101

6. Detect double-bit errors and correct single-bit errors in the following words originally encoded as a $(8, 4)$ Hamming code. (We are taking the convention from class that the overall parity bit is first.)

   (a) 00110011

(b) 10111100
(c) 10100001
(d) 00110101

7. Suppose each bit can flip with a probability of $1/100$, and that bits errors are statistically independent. What is the probability that a 4 bit message is received correctly when using
   (a) no error correction,
   (b) a $(7, 4)$ Hamming code,
   (c) an $(8, 4)$ Hamming code with retransmission? (You may make sensible simplifying assumptions.)