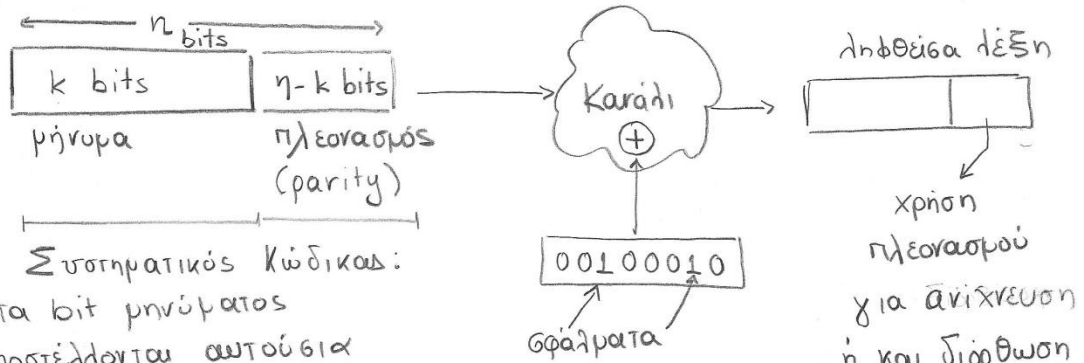
Κώδικας Διόρθωσης Σφαλμάτων (Block Code)



Συστηματικός Κώδικας: τα bit μηνύματος αποστέλλονται αυτούσια

Ρυθμός πληροφορίας: $r = \frac{k}{\eta}$

Βάρος Hamming λέξης x μήκους η bits

$$wt(x) = \sum_{i=1}^{\eta} (\text{ones}) \qquad \pi.\chi.\ wt(1011001) = 4$$

Απόσταση 2 λέξεων x, y

$$d(x,y) = \sum_{i=1}^{\eta} (\text{διαφορετικά bits στις αντίστοιχες θέσεις}) =$$

$$= wt(x + y)$$

↑ XOR

$\pi.\chi.$ $x = 1011001$

$y = 1101000$

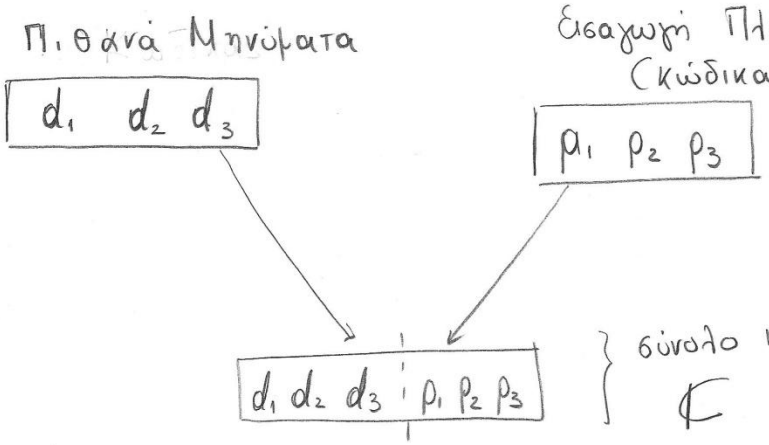
$x+y = 0110001 \rightarrow wt(x+y) = d(x,y) = 3$

διαφορετικά bits μεταξύ x, y

Πιθανότητα αποστολής λέξης x (η bits) και λήψης λέξης y με απόσταση $d(x,y) = d$ [δηλ. να συμβούν d λάθη]

Αν πιθανότητα επιτυχούς μετάδοσης 1 bit: p

τότε $P(d \text{ bits εσφαλμένα}, \eta-d \text{ bits σωστά}) = p^{\eta-d}(1-p)^{d}$ ⇐ Αξιοπιστία



Αν τα ψηφία πλεονασμού προκύπτουν ως γραμμικοί συνδυασμοί των ψηφίων μηνύματος ο κώδικας είναι γραμμικός.

Ιδιότητες γραμμικών κωδίκων

1) $\forall\ x, y \in \mathbb{C},\ x+y \in \mathbb{C}$

2) $\mathbf{0} = \underbrace{000\ldots0}_{\text{n bits}} \in \mathbb{C}$

3) Απόσταση κώδικα: (γενικά) η ελάχιστη απόσταση μεταξύ δύο κωδικολέξεων του

Αν $\mathbb{C}$ γραμμικός: Απόσταση = ελάχιστο των βαρών των κωδικολέξεων του $\mathbb{C}$

Μηνύματα

	d_1 d_2	Ελάχιστη απόσταση: 1
m_1	0 0	
m_2	0 1	Αρκεί 1 σφάλμα για να ληφθεί λάθος μήνυμα
m_3	1 0	
m_4	1 1	

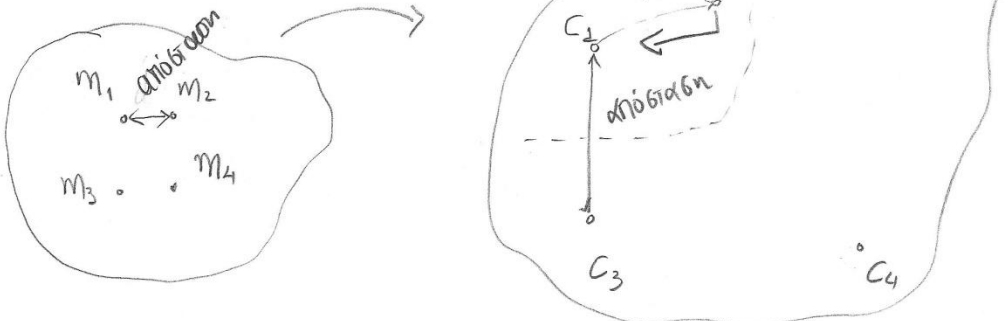
Έστω κώδικας με 2 parity bits

$$p_1 = d_1 + d_2$$

$$p_2 = d_1$$

	d_1 d_2 p_1 p_2	Ελάχιστη απόσταση: 2
C_1	0 0 0 0	
C_2	0 1 1 0	"Αυξάνει" η απόσταση μεταξύ 2 κωδικολέξεων
C_3	1 0 1 1	
C_4	1 1 0 1	

αρα αυξάνει η αξιοπιστία (Διόρθωση όλων των σφαλμάτων 1 bit και αρκετών 2 bit)



Κωδικοποίηση: Αύξηση Πλεονάσματος → { Αύξηση Αξιοπιστίας / Κόστος σε εύρος ζώνης

Παράδειγμα (Συνέχεια)

C_1 0 0 0 0

C_2 0 1 1 0

C_3 1 0 1 1

C_4 1 1 0 1

— Αποστολή 00 → (κωδ/ση) '0000' λήψη '0000'

Σύγκριση με κώδικα → C_1

Απόφαση: λανθασμένα λέξη '0000' → C_1 → m_1 → '00' (αποκωδικοποίηση)

— Αποστολή 11 → (κωδ/ση) '1101' λήψη '1100' (↑ bit error)

Σύγκριση με κώδικα → καμμία λέξη

ελάχιστη απόσταση από C_4

Απόφαση: λανθασμένα λέξη: '1101' → C_4 → m_4 → '11' (αποκωδικοποίηση)

(Αποκωδικοποίηση μέγιστης πιθανότητας) (διόρθωση)

→ Αποστολή 01 → (ΑΜΠ) '0110' λήψη '1111' (bit errors)

1111 ∉ C

ελάχιστη απόσταση { C_4, C_3 }

? → Π ΑΜΠ (Τυχαία Επιλογή C_4) (πλήρης)

Α ΑΜΠ (άδεια επανεκπομπής) (ατελής)

Θεωρήματα

Κώδικας $\mathcal{C}$ απόστασης d

$\Rightarrow$ Ανιχνεύει όλα τα σφάλματα ε με $wt(\varepsilon) < d-1$

$\Rightarrow$ Δεν ανιχνεύει ένα τουλάχιστον σφάλμα ε με $wt(\varepsilon) = d$

$\Rightarrow$ Διορθώνει όλα τα σφάλματα ε με

$$wt(\varepsilon) \leq \left\lfloor \frac{d-1}{2} \right\rfloor \leftarrow \text{ακέραιο μέρος}$$

$\Rightarrow$ Δεν διορθώνει ένα τουλάχιστον σφάλμα ε με

$$wt(\varepsilon) = 1 + \left\lfloor \frac{d-1}{2} \right\rfloor$$

Κατασκευή Κωδίκων.

- Αν ο κώδικας $\mathcal{C}$ ορίζεται ως ανάπτυγμα ενός συνόλου κωδικολέξεων S, $\mathcal{C} = \langle S \rangle$ τότε

  ο $\mathcal{C}$ περιλαμβάνει:

  → Το $\mathbf{0}$

  → όλα τα στοιχεία του S

  → όλους τους γραμμικούς συνδυασμούς των στοιχείων του S

- Βάση Κώδικα $\mathcal{C}$: είναι ένα σύνολο κωδικολέξεων B για το οποίο ισχύουν τα εξής

  → $\mathcal{C} = \langle B \rangle$

  → Τα στοιχεία του B είναι γραμμικά ανεξάρτητα (κανένα δεν προκύπτει ως γραμμικός συνδυασμός των υπολοίπων)

Γεννήτορας Πίνακας G

Πίνακας Διαστάσεων $k \times \eta$ για

κωδικοποίηση μηνυμάτων k bits σε κωδικολέξεις η bits

Μήνυμα (k bits) × G ⟶ κωδική λέξη

π.χ.

$$G = \left[\begin{array}{cccc:ccc} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right]$$

$k = 4$ γραμμές ⇒ μήκος μηνυμάτων 4 bits

$\eta = 7$ στήλες ⇒ μήκος κωδικολέξης 7 bits

Έστω μήνυμα 0101

Κωδικοποίηση

$0101 \cdot G = 0 \times (1000111) + 1 \times (0100110) + 0 \times (0010101) +$ (απλό γινόμενο, XOR)

$+1 \times (0001011) = \underbrace{\underbrace{0101}_{\text{μήνυμα}}\underbrace{101}_{\text{πλεονασμός}}}_{\text{κωδικολέξη}}$

Στην ουσία προσθέτουμε τη 2η και 4η γραμμή του G

Αποκωδικοποίηση

Έλεγχος ισοτιμίας

Κατασκευάζεται πίνακας ισοτιμίας H $\eta \times (\eta - k)$

Αν η λαθεμένη κωδικολέξη × ανήκει στον κώδικα τότε

$\times \cdot H = 0 \leftarrow$ σύνδρομο

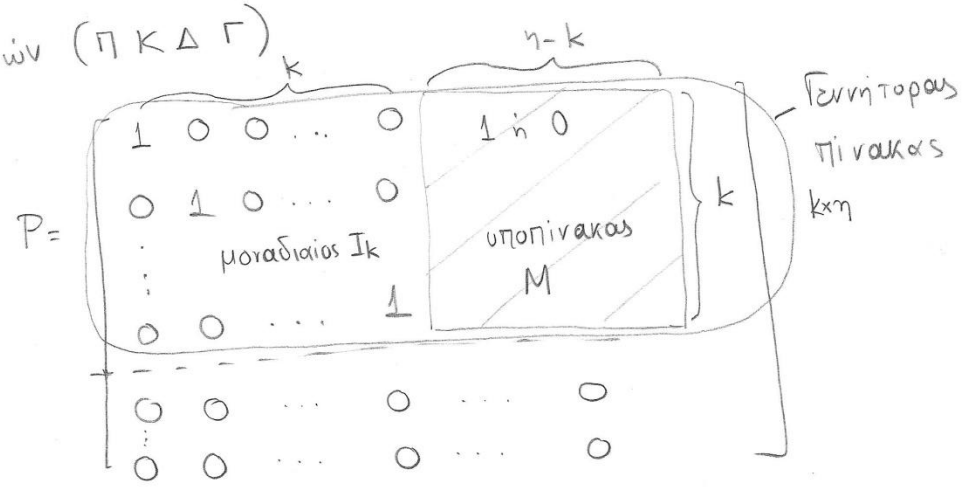
Αν το σύνδρομο είναι μη μηδενικό, χρησιμοποιείται για την ανίχνευση και πιθανή διόρθωση σφαλμάτων.

Εύρεση Πινάκων G, H

Αν δίνεται ο κώδικας $\mathbb{C}$ ή κάποιο υποσύνολό του S του οποίου το ανάπτυγμα δίνει τον $\mathbb{C}$, $\mathbb{C} = \langle S \rangle$ με τη μορφή συνόλου κωδικολέξεων $\{S_1, S_2, \ldots, S_m\}$:

→ Σχηματίζουμε τον πίνακα $P = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_m \end{bmatrix}$

Με αντιμετάθεση γραμμών ή αντικατάσταση γραμμής από το άθροισμά της με άλλη γραμμή μετασχηματίζουμε τον P σε μορφή Περιορισμένης Κλιμακωτής Διάταξης Γραμμών (ΠΚΔΓ)



Γραμμές του G: βάση του $\mathbb{C}$

Κατασκευή πίνακα ισοτιμίας

$$H = \begin{bmatrix} M_{k\times(n-k)} \\ \hline I_{n-k} \end{bmatrix}_{n\times(n-k)}$$

Από τον H προκύπτει και η βάση του δυϊκού κώδικα $\mathcal{C}^{\perp}$ (Σημείωση: Αν $\mathcal{C} = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ τότε ο δυϊκός κώδικας $\mathcal{C}' = \{\beta_1, \dots, \beta_m\}$ είναι δυϊκός του $\mathcal{C}$ δηλ. $\mathcal{C}' = \mathcal{C}^{\perp}$ εφόσον

$$\forall\, i, j \text{ ισχύει } \alpha_i \cdot \beta_j = 0 \quad , \begin{matrix} i = 1, \dots, n \\ j = 1, \dots, m \end{matrix}$$

Οι στήλες του H δίνουν μια βάση για τον $\mathcal{C}^{\perp}$

Εύρεση Απόστασης Κώδικα

1) Αν γνωρίζουμε όλες τις κωδικολέξεις ⇒ ελάχιστο βάρος.
2) Αν γνωρίζουμε ένα υποσύνολο S τέτοιο ώστε $\mathcal{C} = \langle S \rangle$
   ↳ είτε υπολογίζουμε το πλήρες ανάπτυγμα $\langle S \rangle$ και εφαρμόζουμε το βήμα 1)
   ↳ είτε προσδιορίζουμε τους G, H και πηγαίνουμε στο βήμα 3

3) Όριο Singleton: $d \leq n-k$ για γραμμικούς κώδικες (n, k)

Α. Θέτουμε $d_0 = 2$.

Β. Παρατηρούμε στον G αν υπάρχουν d_0 γραμμές των οποίων τα μέρη ίσου μήκους να αθροίζονται σε 000... Αν υπάρχουν τότε απόσταση $d = d_0$. Αν όχι, τότε θέτουμε $d_0 = d_0 + 1$ και επαναλαμβάνουμε το (Β) μέχρι $d_0 = n-k$.

Όριο συστηματικό αντιστοιχείται με τις γραμμές του πίνακα H.

π.χ. $H = \begin{bmatrix} 110 \\ 101 \\ 110 \\ 101 \\ 100 \\ 010 \\ 001 \end{bmatrix}$ →

$$G = \underbrace{\left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right]}_{n=7} \} k=4$$

$d \leq n-k = 3$

Όμως υπάρχουν { 2 γραμμές του G με κοινό τμήμα ίσου μήκους ή 2 όμοιες γραμμές του H } που αθροίζονται σε '000'

Άρα $d = 2$.

Μέθοδοι Αποκωδικοποίησης

Α Μέθοδος Συνοράδων: κώδικα $C(n,k)$

Συνοράδα κώδικα C για το στοιχείο x: το σύνολο $C+x$.

Πλήθος διαφορετικών συνοράδων: 2^{n-k}

1. Λήψη λέξης y
2. Υπολογισμός συνοράδας $C+y$
3. Αν $y \in C$ τότε $0 \in C+y$
4. Αν $y \notin C$ τότε επιλέγουμε το στοιχείο του $C+y^{\varepsilon}$ με το μικρότερο βάρος $\Rightarrow$ πρότυπο σφάλματος
5. Διόρθωση λέξης: $y' = y + \varepsilon$

Αν περισσότερα του 1 δυνατά πρότυπα σφάλματος ελαχίστου βάρους $\Rightarrow$ { ΠΑΝΠ τυχαία επιλογή ενός; ΑΑΝΠ αίτηση για επανεκπομπή

(B) Με τυπική Διάταξη Αποκωδικοποίησης ΤΔΑ.

1. Υπολογισμός πίνακα H

2. Για πρότυπα σφάλματος ελάχιστου βάρους ε_i υπολογίζουμε τα γινόμενα (σύνδρομα) $\varepsilon_i H$ και τα βάζουμε σε πίνακα

πρότυπο σφάλματος	σύνδρομο

3. Για τη ληφθείσα λέξη y υπολογίζουμε το γινόμενο yH

4. Αν $yH = 0$ $y \in C$

5. Αν $yH \neq 0$ αναζητούμε το μη μηδενικό σύνδρομο yH στον ανωτέρω πίνακα και το αντιστοιχίζουμε στο σχετικό πρότυπο σφάλματος ε_i

6. Διόρθωση λέξης $y' = y + \varepsilon_i$

7. Αν περισσότερα πρότυπα σφάλματος αντιστοιχούν στο σύνδρομο, ισχύουν οι προαναφερθείσες επιλογές ΠΑΜΠ, ΑΑΜΠ

Θέμα 4 Ενδεικτικές Ασκήσεις

Λέξη Πληροφορίας $d_1\ d_2$	Κωδικολέξη $d_1\ d_2\ p_1\ p_2\ p_3$
00	00000
01	01101
10	10111
11	11010

? Συστηματικός κώδικας

Όλες οι κωδικολέξεις περιέχουν στα πρώτα 2 ψηφία τους τις αντίστοιχες λέξεις πληροφορίας

$$p_1 = a_1 d_1 + a_2 d_2$$

2η γραμμή $p_1 = 1, \quad d_1 = 0 \quad d_2 = 1$

$$\Rightarrow 1 = a_2$$

3η γραμμή $p_1 = 1, \quad d_1 = 1, \quad d_2 = 0$ → $p_1 = d_1 + d_2$

$$\Rightarrow 1 = a_1$$

$$p_2 = d_1$$

$$p_3 = d_1 + d_2$$

ο κώδικας είναι γραμμικός $\quad \eta = 5$

$k = 2$

Είρεση G

$$G = \left[\begin{array}{cc|ccc} d_1 & d_2 & p_1 & p_2 & p_3 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{array}\right]$$

$$H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Απόσταση Κώδικα $d \leq n-k = 3$ (όριο Singleton)

$d=3$ (ελάχιστο βάρος κώδικα -που μας δίνεται)

ικανότητα διόρθωσης λαθών $\left\lfloor \frac{d-1}{2} \right\rfloor = \left\lfloor \frac{3-1}{2} \right\rfloor = 1$. bit error

Τυπική Διάταξη αποκωδικοποίησης (ΤΔΑ)

Πρότυπο ελαχίστου βάρους (x)	Σύνδρομο $x \cdot H$
1 0 0 0 0	1 1 1
0 1 0 0 0	1 0 1
0 0 1 0 0	1 0 0
0 0 0 1 0	0 1 0
0 0 0 0 1	0 0 1
0 0 0 1 1	0 1 1
ή 1 0 1 0 0	0 1 1
0 0 1 1 0	1 1 0
ή 1 0 0 0 1	1 1 0

ΠΑΜΠ: επιλογή ενός προτύπου

ΑΑΜΠ: κενό πρότυπο αίτηση επανεκπομπής

Παράδειγμα Αποκωδικοποίησης

Λήψη λέξης $y = 0\ 0\ 0\ 0\ 1 \notin C$

→ Μέθοδος με συνομάδες

Σύνολο $C + 0,0001 = \{\underline{00001}, 01100, 10110, 11011\}$

↑ πρότυπο σφάλματος ελαχίστου βάρους

Άρα διορθωμένη λέξη $y' = y + 00001 = \underline{00000}$

→ Μέθοδος με ΤΔΑ.

Σύνδρομο $yH = 00001 \cdot H = \underline{001}$

Από πίνακα ΤΔΑ πρότυπο σφάλματος → 00001.

άρα $y' = y + 00001 = 00000$

Με βάση τον πίνακα ΤΔΑ και υποθέτοντας ΠΑΜΠ.

ο κώδικας διορθώνει:

→ 5 σφάλματα 1 bit

→ 2 σφάλματα 2 bit

Πιθανότητα σφάλματος

$P_F = 1 - P(\text{καμένα σφάλμα}) - P(\text{σφάλματα 1bit}^{5}) -$

$- P(2 \text{ σφάλματα } 2 \text{ bit}) = 1 - (1-\varepsilon)^5 - 5\varepsilon(1-\varepsilon)^4 -$

$- 2\varepsilon^2(1-\varepsilon)^3$

Κώδικας Hamming:

Χαρακτηριστικά:

→ Μήκος της μορφής $\eta = 2^r - 1 \quad r \geq 2$

→ Πίνακας ελέγχου ισοτιμίας H με όλες τις μη μηδενικές λέξεις μήκους r

→ Διάσταση $k = \eta - r = 2^r - 1 - r$

→ Απόσταση $d = 3$

→ Ικανότητα διόρθωσης 1 σφάλματος $\left(\left\lfloor \frac{d-1}{2} \right\rfloor = 1\right)$

→ Στην ΤΔΑ ο πίνακας συνδρόμων περιλαμβάνει όλες τις γραμμές του H [όλες τις δυνατές λέξεις μήκους r]

· Όριο Hamming.

Αν έχουμε κώδικα C με πλήθος κωδικών λέξεων $|C|$ μήκος κωδικολέξης η και απόσταση $d = 2t+1$ ή $d = 2t+2$ τότε ισχύει ότι $|C| \cdot \left[\binom{\eta}{0} + \binom{\eta}{1} + \binom{\eta}{2} + \ldots + \binom{\eta}{t}\right] \leq 2^n$

$$\binom{\eta}{i} = \frac{\eta!}{i!(\eta - i)!}$$

Τέλειοι κώδικες

Αν $d = 2t+1$ και ισχύει η αυστηρώς σχέση με το σύμβολο της ισότητας, ο κώδικας είναι τέλειος

Παράδειγμα κώδικα Hamming

$$H = \left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$$

← όλες οι μη μηδενικές δυνατές δέξεις 3 bit

$n = 7$

$$G = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{array}\right]$$

$k = 4$

$d \leq 7 - 4 = 3$ (όριο Singleton)

$d = 3$ (ιδιότητα Hamming)

Πλήθος κωδικολέξεων

$$|C| = 2^k = 2^4$$

Υπολογισμός ορίου Hamming $d = 2 \cdot 1 + 1$, $t = 1$

$$|C| \cdot \left[\binom{n}{0} + \binom{n}{t}\right] =$$

$$= 2^4 \cdot \left[\binom{7}{0} + \binom{7}{1}\right] = 2^4 \cdot \left[\frac{7!}{0!\,7!} + \frac{7!}{1! \cdot 6!}\right] =$$

$$= 2^4 \left[1 + 7\right] = 2^4 \cdot 8 = 2^4 \cdot 2^3 = 2^7 = 2^n.$$

Άρα, τέλειος κώδικας