

08-05-20

ΠΑΡΤΗ ΣΤΡΑΤΗΣ 2020

Σημειώσεις για τον Τύπο Βαβελίδη
Ε' του Σπύρο ΜαρωνιέρηΑριθμοί Μηχανής

Ορισμός. Ένας αριθμός $x \neq 0$ μπορεί να παρουσιαστεί ως αριθμός κινητής υποδιαστολής (floating point number) ως εξής:

$$fl(x) = \pm (0.d_1d_2d_3 \dots d_t) \cdot b^e, \quad d_1 \neq 0 \quad \begin{array}{l} \text{πχ } 17 = +0.17000000 \cdot 10^2 \\ -347.14 = -0.34714000 \cdot 10^3 \end{array}$$

Ορισμός. Το σύνολο των αριθμών μηχανής που διαφέρει ένας υπολογιστής παρουσιάζεται ως εξής:

$$M(b, t, L, U) \quad \begin{array}{l} \text{πχ } M(10, 8, 3, 12) \\ M(2, 9, 1, 10) \end{array}$$

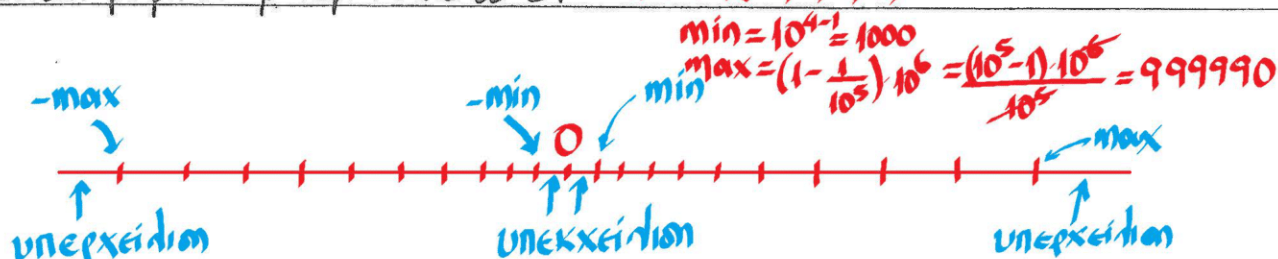
όπου: $\begin{cases} b \text{ η βάση του συστήματος αριθμών} \\ t \text{ η ακρίβεια, δηλ. το πλήθος των επιτρεπόμενων ψηφίων} \\ U \text{ το άνω φράγμα του εκθέτη } e \\ L \text{ το κάτω φράγμα του εκθέτη } e, \text{ δηλ. } L \leq e \leq U \end{cases}$

Ιδιότητες του συνόλου $M(b, t, L, U)$

i) Για δεδομένες τιμές των b, t, L, U το σύνολο $M(b, t, L, U)$ είναι πεπερασμένο & έχει κατ'απόφαση τιμή ένα ελάχιστο & ένα μέγιστο σημείο. Ισχύει λοιπόν ότι:

$$\min_{(x) \in M} |(x)_b| = b^{L-1} \quad \& \quad \max_{(x) \in M} |(x)_b| = \left(1 - \frac{1}{b^t}\right) \cdot b^U$$

ii) Οι αριθμοί μηχανής είναι πυκνά κατανομημένοι κοντά στο 0 & αραιά κατανομημένοι μακριά από το 0. πχ $M(10, 5, 4, 6)$



-2-

- ii) Το πλήθος των αριθμών με τον ίδιο εκθέτη e , είναι ανεξάρτητο του e & ίσο με $2(b-1) \cdot b^{t-1}$.
- iii) Το πλήθος όλων των στοιχείων του M είναι ίσο με $2(V-L+1) \cdot (b-1) \cdot b^{t-1}$.
- iv) Εάν θέλουμε να παραστήσουμε έναν αριθμό $x \in \mathbb{R}$ με έναν αριθμό μηχανής & ισχύει ότι $|x| < \min |(x)_b|$, τότε δώ υπάρχει αριθμός μηχανής για να τον παραστήσουμε $(x)_b^{EM}$ & η κατάσταση χαρακτηρίζεται ως υπερχείλιση (underflow).
- vi) Εάν ομοίως ισχύει ότι $|x| > \max |(x)_b|$, τότε πάλι δώ υπάρχει κατάλληλος αριθμός μηχανής $(x)_b^{EM}$ για να τον παραστήσει & η κατάσταση χαρακτηρίζεται ως υπερχείλιση (overflow).

Σφάλματα στη παραίστηση αριθμών μηχανής

Διαδεχόμε το σύνολο $M(b, t, L, V)$ & τον αριθμό ενντηώς υποδιαστολής μη αναγνωρίσιμα ψηφία

$$fl(x) = (0. \underbrace{d_1 d_2 d_3 \dots d_t d_{t+1} d_{t+2} \dots}_{\text{πλήθος αναγνωρίσιμων ψηφίων}}) \cdot b^k$$

Ο μέγιστος αριθμός αναγνωρίσιμων ψηφίων ισούται με t , με συνέπεια ο αριθμός x να μην μπορεί να παρασταθεί στη μνήμη του υπολογιστή. Έτσι, ο $x \notin M$ πρέπει να προσεγγιστεί από έναν αριθμό, ο οποίος συμβολίζεται με $fl(x)$. Η προσέγγιση αυτή μοιάζει προφανώς ένα σφάλμα. Οπότε:

- i) εάν η προσέγγιση που εφαρμόζουμε είναι προϊόν στρογγύλων, για το απόλυτο σχετικό σφάλμα ισχύει ότι:

$$\frac{|f(x) - x|}{|x|} \leq \frac{1}{2} b^{1-t}$$

(απόλυτο) σχετικό
μοναδιαίο σφάλμα
στρογγυλεύσεως.

ii) εάν η προσέγγιση που εφαρμόζουμε είναι πρόσθ αποκοπή, για
το απόλυτο σχετικό σφάλμα ισχύει ότι:

$$\frac{|f(x) - x|}{|x|} \leq b^{1-t}$$

(απόλυτο) σχετικό
μοναδιαίο σφάλμα
αποκοπής.

A

-4-