16/10/17

Άσκηση: Να βρεθεί το σύνολο αριθμών μηχανής $M = M(2,3,-2,2)$

Λύση

$M = M(2,3,-2,2)$: ο απόλυτα μικρότερος:

$$.100 \cdot 2^{-2} = 2^{-1} \cdot 2^{-2} = 2^{-3} = \frac{1}{8} = (0.125)_{10}$$

Μεγαλύτερος αριθμός: $1.11 \cdot 2^{2} = (2^{-1}+2^{-2}+2^{-3})\, 2^{2}$

$$= 2+1+2^{-1}$$

$$= 3.5$$

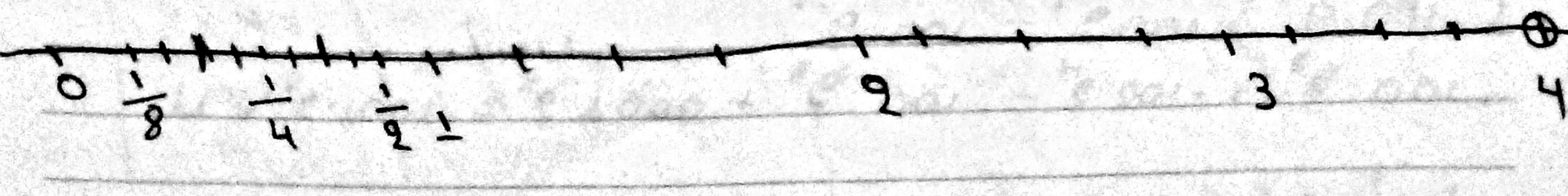
• $e=-2$: $.100 \cdot 2^{-2} = \frac{1}{8}$, $.101 \cdot 2^{-2} = \frac{1}{8}+\frac{1}{32}$, $.110 \cdot 2^{-2} = \frac{1}{8}+\frac{1}{16}$, $.111 \cdot 2^{-2} = \frac{1}{8}+\frac{1}{16}+\frac{1}{32}$.

• $e=-1$: $.100 \cdot 2^{-1} = \frac{1}{4}$, $.101 \cdot 2^{-1} = \frac{1}{4}+\frac{1}{16}$, $.110 \cdot 2^{-1} = \frac{1}{4}+\frac{1}{8}$, $.111 \cdot 2^{-1} = \frac{1}{4}+\frac{1}{8}+\frac{1}{16}$.

• $e=0$: $.100 \cdot 2^{0} = \frac{1}{2}$, $.101 \cdot 2^{0} = \frac{1}{2}+\frac{1}{8}$, $.110 \cdot 2^{0} = \frac{1}{2}+\frac{1}{4}$, $.111 \cdot 2^{0} = \frac{1}{2}+\frac{1}{4}+\frac{1}{8}$

• $e=1$: $.100 \cdot 2^{1} = 1$, $.101 \cdot 2^{1} = 1+\frac{1}{4}$, $.110 \cdot 2^{1} = 1+\frac{1}{2}$, $.111 \cdot 2^{1} = 1+\frac{1}{2}+\frac{1}{4}$,

• $e=2$: $.100 \cdot 2^{2} = 2$, $.101 \cdot 2^{2} = 2+\frac{1}{2}$, $.110 \cdot 2^{2} = 2+1=3$, $.111 \cdot 2^{2} = 2+1+\frac{1}{2} = 3{,}5$



$$\begin{array}{r} .101 \\ +\ .001 \\ \hline 110 \end{array} \qquad \begin{array}{r} .111 \\ +\ .001 \\ \hline 1.000 \end{array} = .100 \cdot 2$$

Αν x ανήκει στο εύρος των αριθμών μηχανής $M=(B,t,L,U)$
(Υποθέτω χωρίς βλάβη γενικότητας, ότι: $x>0$),
αλλά x δεν είναι αριθμός μηχανής, τότε
ο x στρογγυλεύεται στον πλησιέστερο αριθμό μηχανής.

→ Το απόλυτο σχετικό σφάλμα εκτιμάται ως

$$\left| \frac{fl(x)-x}{x} \right| \leq \frac{1}{2} B^{1-t}$$

<u>Απόδειξη</u>

Έστω $x = q \cdot B^k$, $L \leq k \leq U$, $q = .d_1 d_2 \ldots d_t d_{t+1} \ldots$
Έστω x', x'' οι πλησιέστεροι αριθμοί μηχανής
$(x' < x < x'')$
Τότε το $|fl(x)-x| \leq \frac{1}{2}|x'-x''|$

$x' = .d_1 d_2 \ldots d_t \cdot B^k$
$x'' = (.d_1 d_2 \ldots d_t + B^{-t}) \cdot B^k$
$x'' - x' = B^{k-t}$
$x = q \cdot B^k \geq .d_1 \cdot B^k \geq .1 \cdot B^k$

$$\left| \frac{fl(x)-x}{x} \right| \leq \frac{1}{2} \frac{|x'-x''|}{|x|} \leq \frac{1}{2} \frac{B^{k-t}}{B^{k-1}} = \frac{1}{2} B^{1-t} = u$$

και λέγεται μοναδιαίο σφάλμα.

$(.100 \cdot 2^{-2}) \cdot .100 \cdot 2^{-2} = .100 \cdot 2^{-5}$
$.100 \cdot 2^{2} + .100 \cdot 2^{-1} = .100 \cdot 2^{2} + .0001 \cdot 2^{2} = .1001 \cdot 2^{2} \notin M$

▶ Έστω $*$ μια από τις πράξεις $\{+,-,\cdot,\div\}$ για τον υπολογισμό της πράξης $x*y$, όπου x,y είναι στο εύρος αριθμών μηχανής αλλά δεν είναι αριθμοί μηχανής, υπολογίζουμε την ποσότητα $z = fl(fl(x) * fl(y))$

Έστω ε το σχετικό σφάλμα ενός αριθμού x.

$$\varepsilon = \frac{fl(x)-x}{x} \Leftrightarrow fl(x) = x(1+\varepsilon).$$

Πολλαπλασιασμός

$$z = fl(fl(x)\cdot fl(y)) = fl(x(1+\varepsilon_1)\cdot y(1+\varepsilon_2))$$
$$= xy(1+\varepsilon_1)(1+\varepsilon_2)(1+\varepsilon_3),$$
$$|\varepsilon_i| \le u,\ i=1,2,3.$$

$$\left|\frac{z-xy}{x\cdot y}\right| = \left|\frac{xy(1+\varepsilon_1)(1+\varepsilon_2)(1+\varepsilon_3)-xy}{xy}\right|$$

$$= |(1+\varepsilon_1)(1+\varepsilon_2)(1+\varepsilon_3)-1|$$
$$= |(1+\varepsilon)^3-1|,\ |\varepsilon|\le u.$$

Έστω $\lambda = \prod_{i=1}^{m}(1+\varepsilon_i) = (1+\varepsilon)^m,\ |\varepsilon|\le u,\ |\varepsilon_i|\le u.$

$(1-u)^m \le \lambda \le (1+u)^m$. Από θεώρημα ενδ. τιμής: $\exists\, \varepsilon \in [-u,u]$ τέτοιο ώστε $\lambda = (1+\varepsilon)^m$.

$$\phi(x) = 1+x,\quad \phi(-u) \le \underset{\substack{\uparrow \\ \phi(\varepsilon)}}{\lambda} \le \phi(u).$$

$$\left|\frac{z-xy}{xy}\right| = |(1+\varepsilon)^3-1| = |3\varepsilon+3\varepsilon^2+\varepsilon^3|$$

$$\le 3|\varepsilon| + 3|\varepsilon|^2 + |\varepsilon|^3 \le 3u + \underbrace{3u^2+u^3}_{\text{τα θεωρώ αμελητέα}} \simeq 3u$$

Το απόλυτο σχετικό σφάλμα τριπλασιάζεται κατά του πολλαπλασιασμό.