

Globale beschrijving van de drijvende arithmetiek van de EL X8.

Met "globaal" wordt bedoeld, dat deze beschrijving de lezer in staat stelle voor zich zelf uit te maken, wat het antwoord van een zekere operatie is, als hij de operanden kent. Het is niet gezegd dat deze beschrijving ook de beschrijving is van de manier, waarop de arithmetische circuits van de EL X8 dit antwoord construeren.

Voorstelbare getallen.

De drijvende arithmetiek opereert op getallen van de vorm

$$m \times 2^e$$

waarbij m (van mantisse) voldoet aan $0 \leq \text{abs}(m) \leq 2^{\uparrow 40} - 1$
 en e (van exponent) voldoet aan $0 \leq \text{abs}(e) \leq 2^{\uparrow 11} - 1 = 2047$.

Opm. 1. Voor e zijn 12 bits ter beschikking, voor m 41; in beide gevallen kan de waarde 0 op twee wijzen worden voorgesteld, nl. +0 en -0.

Opm. 2. Niet alleen e maar ook m neemt dus louter gehele waarden aan.

Afronding.

Optelling, aftrekking, vermenigvuldiging en deling kunnen alle een resultaat afleveren, dat, ongeacht de bovengrens van $\text{abs}(e)$ niet exact voorstelbaar is, omdat de afstand van meest en minst significante bit van $\text{abs}(m)$ meer dan 39 zou moeten zijn.

(Bij optelling en aftrekking kan dit ontstaan doordat de mantissen onder elkaar geschoven moeten worden, bij de vermenigvuldiging is de afstand van meest en minst significante bit van het exacte product in het algemeen 79 of 78, bij de doorgaans niet opgaande deling is deze afstand oneindig).

In al deze gevallen wordt exact afgerond, en in het twijfelgeval van de nul af.

Het twijfelgeval treedt op, wanneer de afstand van meest en minst significante bit van het exacte antwoord 40 is; m.a.w. als de exacte $m = 2^{\uparrow 40} + 1$ wordt dit tot $2^{\uparrow 40} + 2$ afgerond, zodat halvering tot $2^{\uparrow 39} + 1$ mogelijk is.

Aldus is de eventueel afgeronde m en exacte e (waarin netjes van alle halveringen is boekgehouden) bepaald. Van deze exacte e nemen we voorshands niet aan, dat hij aan $\text{abs}(e) \leq 2047$ voldoet.

Opm. 3. Door de conventie dat het twijfelgeval van de nul af wordt afgerond is gegarandeerd, dat $(-a) \times b = -(a \times b)$ is.

Vervolgens worden exacte e en (eventueel afgeronde) m aan de normering onderworpen.

De normering probeert de absolute waarde van e zo klein mogelijk te krijgen als zonder cijferverlies mogelijk is. (Elke verhoging van e moet gecompenseerd worden door halvering van m , wat alleen zonder cijferverlies mogelijk is, als m eerst even is; elke verlaging van e moet gecompenseerd worden door verdubbeling van m , wat zonder "cijferverlies aan kop" alleen mogelijk is als $\text{abs}(m) < 2^{\uparrow 39}$).

Opm. 4. Dit proces kan eindigen doordat $\text{abs}(e) = 0$ wordt; in dat geval wordt $e = +0$ afgeleverd. Als $m = \pm 0$, kan onbepikt zonder cijferverlies geschoven worden: in de hardware zit een $m = \pm 0$ detectie, die dit geval kortsluit en onmiddellijk $e = +0$ maakt.

Nadat aldus gepoogd is $\text{abs}(e)$ zo klein mogelijk te maken, wordt gekeken, of $\text{abs}(e)$ klein genoeg is.

Als $e \geq 2048$ is, dan wordt als antwoord afgeleverd $e = 2047$ en $m = \pm (2^{\uparrow 40} - 1)$, en wel met het teken van de oorspronkelijke m .

Als $e \leq -2048$ wordt (bij herhaling)

1) e met 1 opgehoogd

2) m afkappend gehalveerd, als $\text{abs}(m) > 1$ is;

als $\text{abs}(m) = 1$ is blijft m ongewijzigd.

Met andere woorden: in geval van overflow wordt het grootste getal met het goede teken afgeleverd, in het geval van underflow wordt door uitschuiven gepoogd de exponent alsnog binnen de perken te brengen. Het minste, wat overblijft is daarbij $\pm 2^{\uparrow (-2047)}$.

Underflow kan dus niet $m = 0$ genereren.

Opm. 5. Aftrekking van een getal " $m \times 2^{\uparrow e}$ " wordt gerealiseerd door " $-m \times 2^{\uparrow e}$ " op te tellen.

Opm. 6. Overflow kan optreden bij $+$, $-$, \times en $/$.

Underflow kan niet bij $+$ en $-$ optreden.

Opm. 7. Additieve bewerkingen leveren alleen nul op als het antwoord exact $= 0$ is; dit antwoord is alleen $m = +0$ wanneer de operanden van de effectieve optelling (zie opm. 5) beide $= +0$ zijn.

Opm. 8. De vermenigvuldiging levert alleen exact nul als antwoord, wanneer ten minste een factor $= 0$ is. Het teken van de productnul volgt de normale tekenregel voor producten.

Opm. 9. Voor de deling geldt:

deeltal $\neq 0$, deler $\neq 0 \rightarrow$ quotiënt $\neq 0$

deeltal = ± 0 , deler is $\neq 0 \rightarrow$ quotiënt = ± 0 (met in achtname van de normale tekenregels).

De situatie voor deler = ± 0 is momenteel het voorwerp van ernstige discussie.

Verwacht wordt:

deeltal $\neq 0$, deler $\pm 0 \rightarrow \pm (2^{40}-1) \times 2^{\uparrow 2047}$ met in achtname van de normale tekenregel.

deeltal = \pm deler = $\pm 0 \rightarrow$ een welgedefinieerd antwoord met in achtname van de tekenregel.

Eindhoven, 6 december 1965.