NumIntro - Uge 2 Onsdag

Nick Laursen

Københavns Universitet

12. september 2020

1. Opgaver

1) Skriv 18 og 119.3 på binæer form

Det er klart at 18 kan skrives som

$$16 + 2 = 2^4 + 2^1 = (10010)_2 \,.$$

For 119.3 vil vi dele tallet op, således vi kigge på 119 og 0.3 fordi $119 + 0.3 = 119.3$.

Herved ser vi så at $64 < 119 < 128 = 2^7$. Så vi ved nu at $119 = (1x_1x_2x_3x_4x_5x_6)_2$. Herved forsætter vi og får så $119 - 64 = 55 > 2^5 \Rightarrow x_1 = 1$. Ved at forsætte fås

$$(119)_{10} = (1110111)_2 \,.$$

1) Skriv 18 og 119.3 på binæer form

Så nu mangler vi bare at finde 0.3. Vi ser at $(0.3)_{10} = (0.a_1a_2a_3\dots)_2$. Følende algoritme kan bruges:

$$0.3 \cdot 2 = 0.6 \Rightarrow a_1 = 0 \tag{1.1}$$
$$0.6 \cdot 2 = 1.2 \Rightarrow a_2 = 1 \tag{1.2}$$
$$(1.2 - 1) \cdot 2 = 0.4 \Rightarrow a_3 = 0 \tag{1.3}$$
$$0.4 \cdot 2 = 0.8 \Rightarrow a_4 = 0 \tag{1.4}$$
$$0.8 \cdot 2 = 1.6 \Rightarrow a_5 = 1 \tag{1.5}$$

Vi ser nu at **(1.5)** giver os 1.6. Og derfor ved vi at $1.6 - 1 = 0.6$ som er **(1.2)**. Fra dette vil den forsætte ud i uendelig og derfor får vi

$$(119.3)_{10} = (1110111.01001110011\dots)_2 .$$

$\square$

2) Hvad sker der med et binær tal når man ganger det med 2^k

Vi husker at vi kan skrive et binær tal på følgende form

$$x = \sum_{i \in \mathbb{Z}}^{N} a_i 2^i,$$

hvor $a_i \in \{ 0, 1 \}$. Så det er klart, at med følgende udregning vil vi få svaret:

$$2^k x = 2^k \sum_{i \in \mathbb{Z}}^{N} a_i 2^i = \sum_{i \in \mathbb{Z}}^{N} 2^k a_i 2^i = \sum_{i \in \mathbb{Z}}^{N} a_i 2^{i+k}.$$

Herved ser vi nu at vi vil rykke x med k pladser. F.eks. $5 \cdot 2^3 = 101 \cdot 1000 = 101000 = 40.$ $\square$

3) Bevis at $\frac{4}{5}$ ikke kan repræsenteres i MARC-32. Hvad er det tætteste maskine-tal og den tætteste "*Roundoff error*"?

Vi kan bruge samme algoritme fa **Opgave 1** og får så:

$$(0.8)_{10} = (0.1100110011\ldots)_2 .$$

Hvorfor kan 0.8 ikke repræsenteres i MARC-32? (*p.40*)

Dette viser at hvis 0.8 skulle repræsenteres i MARC-32, så skulle vi have en uendelig stor "*Normalized Mantissa*", hvilket vi ikke har. Derfor kan 0.8 ikke repræsenteres i MARC-32.

3) Bevis at $\frac{4}{5}$ ikke kan repræsenteres i MARC-32. Hvad er det tætteste maskine-tal og den tætteste "*Roundoff error*"?

Vi vil nu finde det tætteste maskine tal. Vi skiver vores binær tal op på $(1.f)$ modellen. Vi gør dette, da vi kan representere vores binær tal med 1 bit mere. Så vi har
$(0.11001100\ldots)_2 = (1.10011001100\ldots)_2 \cdot 2^{-1}$. Herved finder vi nu dens "*chop-off*" og "*roundup*" maskine tal. Dette giver os

$$x_- = \left(1.\underbrace{1001100\ldots1100}_{\text{23-bit}}\right)_2 \cdot 2^{-1} \quad (\textbf{Chop})$$

$$x_+ = \left(1.\underbrace{1001100\ldots1101}_{\text{23-bit}}\right)_2 \cdot 2^{-1} \quad (\textbf{Round})\,.$$

Lad os nu begynde udregningerne for det tætteste maskine tal.

3) Bevis at $\frac{4}{5}$ ikke kan repræsenteres i MARC-32. Hvad er det tætteste maskine-tal og den tætteste "*Roundoff error*"?

Det ses at

$$\begin{aligned} x - x_- &= (0.1100\dots)_2 - (0.1100\dots1100)_2 \\ &= (0.110011\dots)_2 \cdot 2^{-24} = \frac{4}{5} \cdot 2^{-24}, \end{aligned}$$

og

$$x_+ - x = (x_+ - x_-) - (x - x_-) = 2^{-24} - \frac{4}{5} \cdot 2^{-24} = \frac{1}{5} \cdot 2^{-24}.$$

Herved ser vi at x_+ er det tætteste maskine tal og derfor er $\mathrm{fl}(x) = x_+$. Dette er den abselute "roundoff error".

3) Bevis at $\frac{4}{5}$ ikke kan repræsenteres i MARC-32. Hvad er det tætteste maskine-tal og den tætteste "*Roundoff error*"?

Dens relative roundoff-error fås til at være

$$\frac{|fl\,(x) - x|}{|x|} = \frac{\frac{1}{5} \cdot 2^{-24}}{\frac{4}{5}} = \frac{1}{4} \cdot 2^{-24} = 2^{-26}.$$

$\square$

2.1.39) *Evaluer:* $0.1818\ldots$, $2.702702\ldots$ *og* $98.198198\ldots$

Lad $x = 0.1818\ldots$. Så er det klart at $100x = 18.1818\ldots$. Herved får vi

$$100x - x = 99x = 18.1818\ldots - 0.1818\ldots = 18.$$

Vi isolerer nu x og får

$$99x = 18 \Rightarrow x = \frac{18}{99}.$$

På ligende måde fås:

$$2.702702\ldots = \frac{2700}{999},\ 98.198198\ldots = \frac{98100}{999}.$$

□

2.2.8) *Giv en metode sådan at vi kan undgå problemer med andengrads ligningen*

Hvad er det første vi skal være opmærksom på? Svaret er værdien af b. Hvis $b < 0$ eller $b > 0$, så vil fortegnet for den ene værdi af b ikke have samme inflydelse for det andet b.

Jeg vil foreslå at kigge på **Example 1, p.56**. Vi laver følgende opskrivning:

$$x = \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a}\right)\left(\frac{b + \sqrt{b^2 - 4ac}}{b + \sqrt{b^2 - 4ac}}\right) = \frac{-2c}{b + \sqrt{b^2 - 4ac}}.$$

Gælder denne ligning for $b > 0$ eller $b < 0$? Svaret er for $b > 0$, da $-b + |b| = 0 \Rightarrow b \geq 0$.
For $b < 0$ bruges samme teknik og vi får:

$$x = \left(\frac{-b - \sqrt{b^2 - 4ac}}{2a}\right)\left(\frac{b - \sqrt{b^2 - 4ac}}{b - \sqrt{b^2 - 4ac}}\right) = \frac{-2c}{b - \sqrt{b^2 - 4ac}}$$

2.2.8) *Giv en metode sådan at vi kan undgå problemer med andengrads ligningen*

Herved kan vi finde rødderene. Vi får følgende:

For $b > 0:$

$$x_1, x_2 = \left\{ \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \frac{-2c}{b + \sqrt{b^2 - 4ac}} \right\},$$

og for $b < 0:$

$$x_1, x_2 = \left\{ \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \frac{-2c}{b - \sqrt{b^2 - 4ac}} \right\}.$$