MI1100 - Forelesning 14/9-2016

Sist: representasjon av heltall og brøktall

I dag: Mer om representasjon av tall på datamaskin

Kompendiet 3.4-3.5 og 4.1-4.2 + litt kap 5 hvis tid

BRØKTALL

Eks: $0.14 = 1 \cdot \frac{1}{10} + 4 \cdot \frac{1}{100} = 1 \cdot 10^{-1} + 4 \cdot 10^{-2}$

$0.45928 = 4 \cdot 10^{-1} + 5 \cdot 10^{-2} + 9 \cdot 10^{-3} + 2 \cdot 10^{-4} + 8 \cdot 10^{-5}$

Generelt er brøktall på formen $(0.d_{-1} d_{-2} d_{-3} \cdots)_\beta$

$= d_{-1} \cdot \beta^{-1} + d_{-2} \beta^{-2} + d_{-3} \cdot \beta^{-3} \cdots$

Også her er d'ene tall mellom 0 og $\beta - 1$

Brøktall ligger i intervallet $[0,1)$

Hvordan kan vi representere en brøk i β-tallsystem

Eks: $a = 1/5$ med $\beta = 8$

Må finne $d_{-1}, d_{-2}, d_{-3} \cdots$ så

$1/5 = d_{-1} 8^{-1} + d_{-2} 8^{-2} + d_{-3} 8^{-3} \cdots$

Ganger med 8 på begge sider

$1 + 3/5 = 8/5 = \underbrace{d_{-1}}_{\text{Heltall}} + \underbrace{d_{-2} 8^{-1} + d_{-3} 8^{-2} + \cdots}_{\text{Brøkdel}}$

$d_{-1} = 1$

Trekke fra heltallsdel på begge sider

Får

$3/5 = d_{-2} 8^{-1} + d_{-3} 8^{-2} + \cdots$

Ganger med 8 og får

$4 + 4/5 = 24/5 = \underbrace{d_{-2}}_{\text{Heltall}} + \underbrace{d_{-3} \cdot 8^{-1} + d_{-4} 8^{-2} + \cdots}_{\text{BRØKDEL}}$

$\Rightarrow d_{-2} = 4$

Fortsetter på samme måte for å finne flere siffer.

$6 + 2/5 = 32/5 = d_{-3} + d_{-4} 8^{-1} + d_{-5} 8^{-2} + \cdots$

$\Rightarrow d_{-3} = 6$

$3 + 1/5 = 2/5 \cdot 8 = d_{-4} + \underbrace{d_{-5} 8^{-1} + d_{-6} 8^{-2} + \cdots}$

$\Rightarrow d_{-4} = 3$

$1/5 = d_{-5} 8^{-1} + d_{-6} 8^{-2} + d_{-7} 8^{-3} + \cdots$

Har sett denne ligningen før, vet at $d_{-5} = 1$, $d_{-6} = 4$ osv.

$1/5 = 0.\underbrace{1463}\underbrace{1463}\underbrace{1463}\cdots$

Teorem 3.15: Ethvert tall i $(0,1)$ kan skrives i β-system med $\beta \in \mathbb{N}$, $\beta > 1$
entydig hvis vi ikke tillater uendelig mange siffere $\beta - 1$ til slutt.
$(0.99999\ldots = 1)$

Eks:

BRØKDEL	$8x$	HELTALL
$1/5$	$8/5$	1
$3/5$	$24/5$	4
$4/5$	$32/5$	6
$2/5$	$16/5$	3
$1/5$	$8/5$	1
		4
		6
		3

Skal se på en algoritme for dette

Trenger „floor"-operasjonen:

$\lfloor x \rfloor$ = nærmeste heltall som er mindre $\overbrace{\text{enn}}^{\text{eller lik}}$ x

Eks $\lfloor 1.3 \rfloor = 1$ $\lfloor 3.1415\cdots \rfloor = 3$ $\lfloor 3.1 \rfloor = 3$

$\lfloor -1.3 \rfloor = -2$

Algoritme 3.16. La a være et brøktall i $(0,1)$.
Da kan vi finne sifferne $d_{-1}, d_{-2}, d_{-3}, \cdots$ til a i β-tallsystem slik:

$a_{-1} = a$
for $i = -1, \cdots, -k$
$d_i = \lfloor a_i \cdot \beta \rfloor$ (Heltallsdel av $a_i \cdot \beta$)
$a_{i-1} = a_i \cdot \beta - d_i$ (Brøkdelen av $a_i \cdot \beta$)

Eks

1/5	2/5	0
2/5	4/5	0
4/5	8/5	1
3/5	6/5	1
1/5		

$1/5 = (0.\underbrace{0011}\underbrace{0011}\underbrace{0011}\ldots)_2$

Lemma 3.21 La $a \in (0,1)$.
a er et rasjonalt tall hvis og bare hvis sifferne til a i β-system gjentar seg i sykler til slutt

Hva med $1/5 = 0.2000\ldots$ 0 gjentar seg ...

Lemma 3.22. La $a \in (0,1)$. Da vil a ha et endelig antall siffer $\neq 0$ i grunntallet β hvis og bare hvis a er et rasjonalt tall $a = \frac{b}{c}$ og alle primtallsfaktorene i c er faktorer i β!

Eks

$1/5 = (0.2)_{10}$ $\quad \beta = 2\cdot 5$ $\quad c = 5$

$1/3 = (0.333\ldots)_{10}$ $\quad c = 3$

$1/3 = (0.2)_6$ $\quad \beta = 2\cdot 3$ $\quad c = 3$

$1/10 = (0.1)_{10} = (0.00011001100\ldots)_2$

$1/8 = \frac{125}{1000} = (0.125)_{10}$ $\quad c = 2\cdot 2\cdot 2$

Kap 3.4 Aritmetikk i β-systemer

Addisjon i 10-tallsystem:

$$\begin{array}{r} 457_{10} \\ +\ 325_{10} \\ \hline 782_{10} \end{array}$$

Det samme gjør vi i β-tallsystem

en over

$$\begin{array}{r} 457_{8} \\ +\ 325_{8} \\ \hline 1004_{8} \end{array}$$

$7+5=12_{10}=14_8$

$1+5+2=8_{10}=10_8$

$1+4+3=10_8$

Subtraksjon:

$$\begin{array}{r} \cancel{1}4_{8} \\ -\ 7_{8} \\ \hline 5_{8} \end{array}$$

$8+4-7=5$

$14=1\cdot 8^1+4$

Multiplikasjon: $(2\cdot 3)_4 = 6_{10} = 12_4$

Kapittel 4: Representasjon av tall på datamaskin

4.1 HELTALL

SEKVENS AV 8 SIFFER I 2-TALLSYSTEM KALLES „BYTE"

Eks: $10000001 = 129_{10}$

32 BITS (BINÆRE SIFFER 0/1) TALL DVS 4 BYTES

BRUKER 1 SIFFER (BIT) TIL FORTEGN

31 — " — TIL TALLVERDI

Kan representere heltallene mellom $-2^{31}, \ldots, 2^{31}-1$

64 BITS HELTALL kan representere $-2^{63}, \ldots, 2^{63}-1 \approx 9 \cdot 10^{18}$

4.2 Reelle tall

Vi bruker normalform.

I desimaltall

$a = b \cdot 10^{n}$ der $\frac{1}{10} \leq |b| < 1$

eksponent (størrelse) heltall

signifikand (nøyaktighet)

F.eks $\pi \approx 0.31415 \cdots 10^{1}$

$1/7 \approx 0.1429 \cdot 10^{0}$

$\frac{10\,000\,000}{23} \approx 0.4348 \cdot 10^{7}$

Normalform binærtall

$a = b \cdot 2^{n}$ der b er brøktall $\frac{1}{2} \leq |b| < 1$

Flyttall på IEEE 754 standard 32 bits flyttall

$a = b \cdot 10^{n}$ der b er på binær normalform med 23 bits inklusive fortegn

n er 9-bits heltall ink. fortegn

kan representere tall med tallverdi $1.4 \cdot 10^{-45}$ til $3.4 \cdot 10^{38}$

med 7 desimale siffers nøyaktighet

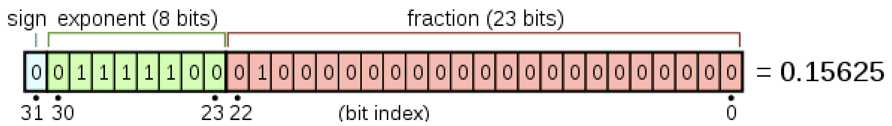
poeng: Med et endelig antall får vi begrenset størrelse og presisjon

$\Rightarrow$ Må gjøre tilnærminger

32-bits flyttall - IEEE 754

Binær normalform: $a = b * 2^n$ der

- b er et 23 bits brøktall 1/2 <= |b| < 1
- n er et 8 bits heltall -127<=n<=128



The real value assumed by a given 32 bit **binary32** data with a given biased exponent **e** (the 8 bit unsigned integer) and a **23 bit fraction** is

$= (-1)^{\text{sign}}(1.b_{22}b_{21}...b_0)_2 \times 2^{e-127}$ where more precisely we have $\text{value} = (-1)^{\text{sign}} \left(1 + \sum_{i=1}^{23} b_{23-i}2^{-i}\right) \times 2^{(e-127)}$.

In this example:

- $\text{sign} = 0$
- $1 + \sum_{i=1}^{23} b_{23-i}2^{-i} = 1 + 2^{-2} = 1.25$
- $2^{(e-127)} = 2^{124-127} = 2^{-3}$

thus:

- $\text{value} = 1.25 \times 2^{-3} = 0.15625$