MATINF1100 - dag 8

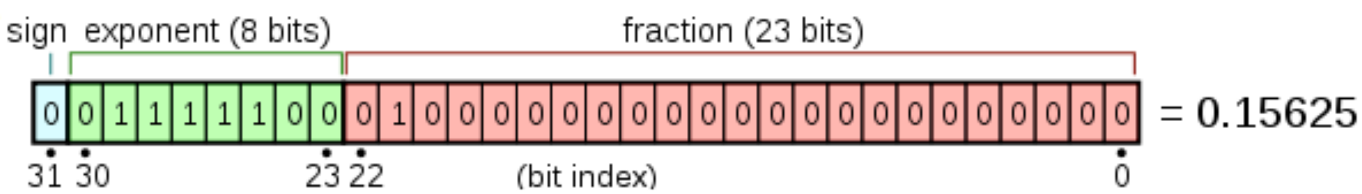
Denne uken:

- Tekst og andre karakterer 4.3-4.5
- Aritmetikk på datamaskin: 5.1-5.4

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:

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

REPRESENTASJON AV TEKST

BRUKER TABELLER : OVERSETTE FRA TALL TIL SYMBOL

ASCII : KODER FOR ENGELSK SPRÅK , 0, ..., 127
TRENGER 7 BITS

ISO LATIN : UTVIDELSE AV ASCII MED 128 TEGN
LEGGER TIL 1 BIT, TOTALT 8 BITS
256 TEGN
KAN TA MED Æ,Ø,Å OSV
ER MED I ISO LATIN 1, SOM HAR MED DE FLESTE VANLIGE VESTLIGE TEGN

UNICODE : UTVIDELSE AV ASCII, OG AV ISO LATIN 1
KJEMPETALL : 100.000 TEGN
MULTINASJONALE TEGN.
TRENGER I UTG. PUNKT MINST 3 BYTES
ULIKE ENKODINGER

UTF-8 : VARIABEL LENGDE ENKODING
ASCII KARAKTERER TRENGER 1 BYTE
85 % AV ALLE WEBSIDER (AUG 2015)
TILBAKEKOMPATIBELT MED ASCII

ANDRE : UTF-16 : 2 ELLER 4 BYTES PR KARAKTER
UTF-32 : ALLTID 4 BYTES (FAST LENGDE)

EKS : Knut
Mørken

	K	n	u	t	↓ (Linjeskift)	M	ø	r	k	e	n
ISO LATIN 1 :	4b	6e	75	74	0a	4d	f8	72	6b	65	6e
UTF-8 :	"	"	"	"	"	"	$\underbrace{\text{c3b8}}_{ø}$	"	"	"	"

KAP 5: ARITMETIKK OG AVRUNDING

HELTALL MED ENDELIG ANTALL BITS

FEKS: 1 BYTE (8 BITS) $-127 \leq a \leq 128$

HVA BLIR $128 \cdot 10 = 1280$

KAN IKKE REPRESENTERES MED 8 BITS

SJELDEN PROBLEM I PRAKSIS

ARITMETIKK MED FLYTTALL (5.2)

FLYTTALL SKRIVES PÅ NORMALFORM

$a = b \cdot 10^n$ hvor $\frac{1}{10} \leq |b| < 1$ (DESIMAL)

Eks: $42.8 = 0.428 \cdot 10^2$

$a = b \cdot 2^n$ hvor $\frac{1}{2} \leq |b| < 1$ (BINÆR)

Eks: $101.001 = 0.101001 \cdot 2^3$

MED 32-bits flyttallsrepresentasjon, kan representere tall mellom 10^{-45} og 10^{38}, med ca 7 desimale siffers nøyaktighet.

POENG: MED ET ENDELIG ANTALL SIFFER FÅR VI BEGRENSET STØRRELSE, OG PRESISJON

$\Rightarrow$ MÅ GJØRE TILNÆRMINGER

ØVELSE: REGN $(\sqrt{2})^2 - 2 = 0$

PÅ NOEN MASKINER FÅR VI FEKS 10^{-16}

MÅ APROKSIMERE! TO METODER

- TRUNKERING : 3.1415... → 3.141 med 4 siffer
- AVRUNDING : 3.1415... → 3.142 med 4 siffer
  BRUKER DET NÆRMESTE TALLET MED 4 SIFFER

EN ENKEL FLYTTALLSMODELL (normalform)

$$a = \alpha \cdot 10^n$$

med $\alpha \in \left[\frac{1}{10}, 1\right)$

4 desimale siffer for α (+ fortegn)

1 — " — — " — n (+ fortegn)

⇒ kan representere $\overbrace{\text{tall}}^{\text{positive}}$ mellom $0.1 \cdot 10^{-9}$ og $0.9999 \cdot 10^{9}$

Overflow kan forekomme, men er enkelt å håndtere

Størst problem med flyttallsaritmetikk er tap av nøyaktighet signifikante siffer

Størst problemer/feil kan oppstå ved addisjon/subtr

Addisjon av flyttall

Skal beregne $a+b$ hvor begge er flyttall, $|a| \geq |b|$,

① Skriv a på normalform
$$a = \alpha \cdot 10^n$$
② Skriv b på tilsvarende form
$$b = \beta \cdot 10^n \quad \text{(samme eksponent)}$$
③ Regn ut $r = \alpha + \beta$, $a+b = r \cdot 10^n$

④ Finn normalform til dette tallet (avrundet)

Eks 5.7: $a = 5.645$ og $b = 7.821$

$a = 0.5645 \cdot 10^1 \quad b = 0.7821 \cdot 10^1$

$\alpha + \beta = 1.3466$

Resultat $0.1347 \cdot 10^2$

Eks : 5.12 : $a = \frac{10}{7} \approx 0.1429 \cdot 10^1 \quad b = -0.142 \cdot 10^1$

$\alpha + \beta = 0.1429 - 0.142 = 0.0009$

Resultatet blir $a+b = 0.9000 \cdot 10^{-3}$

Nøyaktig resultat er $0.8571 \cdot 10^{-3}$

Ingen siffer er riktige!

Grunnen er avrunding og subtraksjon av to nesten like tall