Datarepresentation

Daniel Bosk[1]

Avdelningen för informations- och kommunikationssytem (IKS),
Mittuniversitetet, Sundsvall.

datarep.tex 1278 2013-09-09 17:22:51Z danbos

[1] Detta verk är tillgängliggjort under licensen Creative Commons Erkännande-DelaLika 2.5 Sverige (CC BY-SA 2.5 SE). För att se en sammanfattning och kopia av licenstexten besök URL http://creativecommons.org/licenses/by-sa/2.5/se/.

Översikt

1. Talsystem
   - Olika talsystem
   - Positionssystem
   - Byte av talbas
   - Additionsalgoritmen
2. Hur gör datorn?
   - Boolesk algebra
   - Grindar
   - Representation
   - Additionsalgoritmen igen
3. Icke-numeriska data
   - Teckenkodning
   - Komprimering
   - Datakommunikation

Föreläsningen tar upp kapitel 1 ”Data Storage” i [Bro12], som handlar om fysisk lagring av data på olika media.
Innehållet kommer även att kompletteras med delar från kapitlen 2 och 6 i [Bos11] som förklarar den logiska representationen av data.
Kapitel 2 tar logik och bevis. Då en dator är strikt baserad på logiska operationer är detta en viktig grund. Bevis är helt enkelt tillämpning av de logiska reglerna och detta behövs för att förstå bevisen i kapitel 6.
Kapitel 6 täcker talsystem och förklarar varför vi kan använda datorns logiska system för att räkna med godtyckliga tal. Logiska operationer samt representation och beräkningar av tal är vad en dator gör när program körs.

Översikt

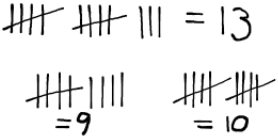
Mittuniversitetet
MID SWEDEN UNIVERSITY

Olika talsystem

Tabell : Olika representationer av etthundratjugotre i olika talsystem.

Binära talsystemet	1111011
Decimala talsystemet	123
Hexadecimala talsystemet	$7B$
Romerska talsystemet	CXXIII

Olika talsystem



Figur : Tecknen i ett enkelt teckenvärdessystem.

Olika talsystem

Tabell : De romerska siffrorna.

I	V	X	L	C	D	M
1	5	10	50	100	500	1000

Olika talsystem

Tabell : Talen 1-10 i det decimala och det romerska talsystemen.

Decimala talsystemet	1	2	3	4	5	6	7	8	9	10
Romerska talsystemt	I	II	III	IV	V	VI	VII	VIII	IX	X

Mittuniversitetet
MID SWEDEN UNIVERSITY

Olika talsystem

Tabell : Några tal skrivna med det romerska talsystemet.

2011	MMXI	$1000 + 1000 + 10 + 1$
1999	MCMXCIX	$1000 + (1000 - 100) + (100 - 10) + (10 - 1)$
1998	MCMXCVIII	$1000 + (1000 - 100) + (100 - 10) + 5 + 1 + 1 + 1$
587	DLXXXVII	$500 + 50 + 10 + 10 + 10 + 5 + 1 + 1$
487	CDLXXXVII	$(500 - 100) + 50 + 10 + 10 + 10 + 5 + 1 + 1$

Positionssystem

Exempel

I 111 betyder den första ettan 100 medan den andra ettan betyder 10 och den sista betyder 1. Det vill säga, samma siffra har olika betydelse beroende på vilken position den har i representationen som den befinner sig i.

Positionssystem

Exempel

Det decimala talsystemet har basen 10. Talet etthundratjugotre representeras i detta system som 123. Vi har då

$$1 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0 = 100 + 20 + 3 = 123.$$

Exempel

Det binära talsystemet har basen 2. Således representeras talet etthundratjugotre som 1111011. Vi har då

$$1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 123.$$

Positionssystem

Definition (Positionssystem[a])

[a] *Eng.* positional system, place-value system

Ett *positionssystem*, eller positionsvärdesystem, har en *talbas* $b \in \mathbb{N} \setminus \{0, 1\}$, siffrorna $S = \{s \in \mathbb{N} \colon s < b\}$ och representerar ett tal $x \in \mathbb{N}$ som $d_1 d_2 \cdots d_n$, där $d_i \in S$ är siffran på position i, och

$$x = d_1 b^{n-1} + d_2 b^{n-2} + \ldots + d_{n-1} b^1 + d_n b^0.$$

Mittuniversitetet
MID SWEDEN UNIVERSITY

Positionssystem

Exempel

Det decimala talsystemet har basen $b = 10$ och använder siffrorna $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Exempel

Det binära talsystemet har basen $b = 2$ och använder siffrorna $S = \{0, 1\}$.

Exempel

Det hexadecimala talsystemet har basen $b = 16$. Det hexadecimala talsystemet använder vanligtvis siffrorna

$$S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F\},$$

där $A = 10$, $B = 11$, ..., $F = 15$.

Byte av talbas

Exempel

Talet $123_{10} = 7B_{16}$, detta finner vi genom följande:

$$123 = 7 \cdot 16 + 11$$
$$7 = 0 \cdot 16 + 7$$

Således får vi siffrorna 7 och B samt att

$$(0 + 7) \cdot 16 + 11 = 7 \cdot 16^1 + 11 \cdot 16^0 = 7B_{16} = 123_{10}.$$

Detta betyder att $7B_{16} = 123_{10}$ och därför är $7B$ hexadecimalt samma tal som 123 är decimalt.

Byte av talbas

Exempel

Talet $123_{10} = 1111011_2$, detta finner vi genom följande:

$$
\begin{aligned}
123 &= 61 \cdot 2 + 1 \\
61 &= 30 \cdot 2 + 1 \\
30 &= 15 \cdot 2 + 0 \\
15 &= 7 \cdot 2 + 1 \\
7 &= 3 \cdot 2 + 1 \\
3 &= 1 \cdot 2 + 1 \\
1 &= 0 \cdot 2 + 1
\end{aligned}
$$

Byte av talbas

Exempel (Fortsättning)

Då får vi

$$\begin{aligned} 1+2\cdot(1+2\cdot(0+2\cdot(1+2\cdot(1+2\cdot(1+2\cdot(1+0)))))) \\ = 1\cdot 2^6+1\cdot 2^5+1\cdot 2^4+1\cdot^3+0\cdot 2^2+1\cdot 2^1+1\cdot 2^0 \\ = 1111011_2 = 123_{10}. \end{aligned}$$

Additionsalgoritmen

Översikt

Boolesk algebra

Möjliga värden Falskt (0), Sant (1).

Operationer och (and, $\wedge$), eller (or, $\vee$), icke (not, $\neg$), exklusivt eller (xor, $\oplus$).

Boolesk algebra

Tabell : Sanningstabell för konjunktionen och disjunktionen. S betecknar sant och F betecknar falskt.

P	Q	$P \wedge Q$	$P \vee Q$
S	S	S	S
S	F	F	S
F	S	F	S
F	F	F	F

Mittuniversitetet
MID SWEDEN UNIVERSITY

Grindar

- Elektriska komponenter som tar två insignaler och ger ifrån sig en utsignal.
- Dessa finns bland annat i formerna "och" samt "eller".
- Finns även så att vi kan får "icke".

Grindar

Olika grindar

Representation

Att lagra en bit

Representation

- Åtta bitar lagras tillsammans och formar en *cell*.
- En *bitsträng* av längd åtta kallas för en *byte*.
- Minnet består av massor av celler.
- Vår bild av minnet är en lång rad med celler adresserade från noll.
- Mest signifikanta bit: high-order end, den biten som påverkar mest om vi byter värde (första sifran i ett tal).
- Minst signifikanta bit: low-order end, den biten som påverkar minst om vi byter värde (sista siffran i ett tal).

Additionsalgoritmen igen

Med logiska operationer och grindar

Översikt

Teckenkodning

- ASCII
- Latin-1
- UTF-8
- Unicode
- ...

Komprimering

Huffmankodning

Komprimering
Lempel–Ziv–Welsh

Datakommunikation

Kommunikationsfel

Referenser

[Bos11] Daniel Bosk. Matematik 1c. I: *En formalisering av matematiken i svensk gymnasieundervisning*, Trita-MAT. MA, ISSN 1401-2278, kapitel Bilaga B. Institutionen för matematik, Kungl. Tekniska högskolan, Stockholm, 2011.

[Bro12] J. Glenn Brookshear. *Computer Science: An Overview*. Pearson Addison-Wesley, Boston, 11, internationella utgåvan, 2012.