MAT1030 – Forelesning 9

Mengdelære

Dag Normann - 16. februar 2010

(Sist oppdatert: 2010-02-16 14:11)

Kapittel 5: Mengdelære

Oppsummering

- Vi er nå ferdige med kapittel 1–4.
- Vi har dekket
  - algoritmer og pseudokoder
  - tall og tallsystemer
  - litt om representasjon av data
  - utsagnslogikk
  - predikatlogikk
  - litt om hvordan bevis kan være strukturert.
- Noen spørsmål før vi går over til kapittel 5?

Mengder

- De fleste som tar MAT1030 har vært bort i mengder i en eller annen form tidligere.
- I statistikk og sannsynlighetsteori på VGS behandler man utfallsrom og studerer matematiske sannsynligheter eller sannsynlighetsfordelinger basert på eksperimenter på slike utfallsrom.
- Man ser på delmengder av slike utfallsrom og eksempelvis Bayes setning, som omhandler både det vi vil kalle komplement og det vi vil kalle snitt (og med de betegnelsene).
- Mengdebegrepet brukes også i de innledende emnene på universitetet, eksempelvis i form av løsningsmengder for ulikheter.

- En mengde er en samling objekter hvor det er entydig bestemt hvilke objekter som er med i mengden eller ikke.
- Bruken av mengder gjennomsyrer matematikk og andre teoretiske fag.
- Vi skal lære å bruke mengder slik at vi kan uttrykke oss presist om konstruksjoner og begreper av interesse i informatikk i en vid forstand.
- Mengdelære brukes på mange måter som matematikkens grunnlag, noe vi ikke skal legge så stor vekt på.

  Vi bruker klammeparenteser $\{$ og $\}$ for å beskrive mengder.

  Vi skal illustrere bruken ved eksempler.

Eksempel.

- $\{0, 1\}$ er mengden av digitale verdier en bit kan ha.
- $\{\mathbf{T}, \mathbf{F}\}$ er mengden av sannhetsverdier.
- $\{2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}$ er mengden av de 10 minste primtallene.
- $\mathbb{N} = \{1, 2, 3, \cdots\}$ er mengden av naturlige tall.

- I de tre første eksemplene har vi listet opp mengdens elementer.
- Mengdene er endelige og i disse tilfellene så små at vi kan beskrive mengden ved hjelp av en liste med klammeparenteser rundt.
- Vi sier at vi har gitt mengden på listeform.
- I en viss forstand har vi beskrevet $\mathbb{N}$ ved hjelp av en liste. I dette tilfellet er listen uendelig lang, og vi har brukt prikkene $\cdots$ for å antyde at opplistingen fortsetter.
- Denne mengden er også gitt på listeform (Læreboka: Enumerated form).
- Bruker vi listeform med prikker eller tilsvarende, må vi være sikre på at leseren vil oppfatte prikkene på samme måte som forfatteren.
- Hvordan ser mengden $\{1, 5, 15, 34, 65, \cdots\}$ ut?
  - Hjelper det å få vite at neste tall er 111?
  - Kunne vi startet med $\{0, 1, \ldots\}$ og forøvrig fått den samme tallmengden basert på den samme regelen?
- Vi trenger et eget symbol for å uttrykke at et objekt er et element i en mengde.

Definisjon.

a) Vi skriver $a \in A$ for å uttrykke at a er et element i mengden A.
b) Vi skriver $a \notin A$ for å uttrykke at a *ikke* er et element i A.

- Vi kunne ha skrevet $\neg(a \in A)$ i stedenfor.
- Det er ikke uvanlig å bruke $/$ som nektingssymbol, som i $3 \not< 2$ og $4 \neq 3$.
- Hvis vi skal beskrive noe mer kompliserte mengder, må vi bruke et litt annet format enn listeform, kalt predikatform i læreboka.
- Dette illustreres også best ved noen eksempler.

Eksempel.

- $\{n \in \mathbb{N} \ : \ n \text{ kan deles på } 5\}$
  Mengden av n i $\mathbb{N}$ slik at n kan deles på 5.
- $\{n \in \mathbb{N} \ : \ n \text{ er et primtall}\}$
  Mengden av n i $\mathbb{N}$ slik at n er et primtall.

Vi kan uttrykke dette mer formelt som følger.
Da blir det vanskeligere å uttrykke definisjonen med ord.

Eksempel.

- $\{n \in \mathbb{N} \ : \ \exists k \in \mathbb{N}(n = 5k)\}$

  Mengden av n i $\mathbb{N}$ slik at det fins en k i $\mathbb{N}$ slik at $n = 5k$.
- $\{n \in \mathbb{N} \ : \ n \geqslant 2 \wedge \forall m \in \mathbb{N} \forall k \in \mathbb{N}(n = km \rightarrow k = 1 \vee m = 1)\}$

  Mengden av de n i $\mathbb{N}$ som er slik at n er større eller lik 2, og slik at for alle m og k i $\mathbb{N}$ har vi at hvis $n = km$ så er $k = 1$ eller $m = 1$.

Vi ser at den formelle definisjonen faktisk er lettere å lese, mens definisjonen fra forrige side vel egentlig var den som var enklest å oppfatte.
Hvordan vi formulerer definisjonen av en mengde avhenger av hva vi vil kommunisere.

- Vi kaller ofte bruken
  $$\{\cdot : \cdots\}$$
  av parenteser for mengdebyggeren.
- Læreboka bruker kolon, :, i mengdebyggeren.
- Det er like vanlig, om ikke vanligere, å bruke en vertikal strek, $|$ i stedenfor.
- Da ser mengdebyggeren ut som
  $$\{\cdot \mid \cdots\}.$$
- Det er mulig at foreleser, av gammel vane, kommer til å bruke $|$ i stedenfor : noen ganger.
- Noen mengder vil vi referere til så ofte at vi bruker egne symboler for dem.
- Vi har allerede sett at vi bruker $\mathbb{N}$ for mengden av naturlige tall.
- Vi lar $\mathbb{J} = \{\cdots . -3, -2, -1, 0, 1, 2, 3, 4, \cdots\}$ være mengden av hele tall.

  Her følger vi boka, men internasjonalt (og i Norge) er det mer vanlig å bruke $\mathbb{Z}$ for denne mengden.
- Vi lar $\mathbb{R}$ stå for mengden av alle reelle tall.

  Det er så mange reelle tall at vi vanskelig kan liste dem opp, ikke engang ved hjelp av $\cdots$.
- Vi lar $\mathbb{Q}$ betegne mengden av rasjonale tall.
- Det er mulig å definere $\mathbb{Q}$ fra de andre mengdene.
  $$\mathbb{Q} = \{x \in \mathbb{R} \ : \ \exists p \in \mathbb{J} \ \exists q \in \mathbb{N} \ (x = \frac{p}{q})\}.$$
- Det er også mulig å skrive $\mathbb{Q}$ på listeform:
  $$\mathbb{Q} = \{0, -1, 1, -2, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, 2, -3, -\frac{8}{3}, -\frac{7}{3}, \cdots\}$$
- I en viss forstand er det umulig å definere $\mathbb{N}$ utfra $\mathbb{R}$ alene.
  - Den forstanden faller langt utenfor pensum i MAT1030.
- Vi har allerede truffet på endel mengder i dette emnet, uten å legge vekt på at de er mengder.

Eksempel.

- La VAL_k være mengden av fordelinger av sannhetsverdier **T** og **F** på k utsagnsvariable $p_1, \ldots, p_k$.
- La var være en uendelig mengde $\{p_1, p_2, \ldots\}$ av utsagnsvariable.

  Vi kan arbeide med denne mengden uten å vite nøyaktig hva disse variablene er, det holder at vi har navn for dem.
- La $\mathsf{REP}_{\mathbb{R}}$ være mengden av digitale representasjoner av reelle tall, og la $\mathsf{REP}_{\mathbb{J}}$ være mengden av digitale representasjoner av hele tall som ikke er den samme mengden som mengden av binære representasjoner av hele tall.
- Da vi snakket om hvordan vi skulle "forstå" kvantorer $\exists x$ og $\forall y$, sa vi at vi måtte presisere hvilke verdier x og y kan ta.

  Dette kan vi gjøre mer presist ved å snakke om variasjonsmengden eller tolkningsområdet til den enkelte variable som en mengde.

- Senere skal vi se på datastrukturer som mengder.
- En datastruktur vil bestå av de objektene en variabel i et program kan ta som verdi.
- Vi har tidligere nevnt at vi ofte må deklarere typen til en variabel.
- Dette betyr at vi må avgrense den delen av datastrukturen som den aktuelle variabelen kan hente sine verdier fra.
- For å kunne beskrive datastrukturer, datatyper og for å kunne diskutere hvordan, og i hvilket omfang, objektene i slike strukturer kan representeres som digitalisert informasjon, trenger vi et grunnlag i mengdelære.
- Det er dette grunnlaget vi skal få når vi gjennomgår stoffet i Kapittel 5.

- Vi skal ikke kaste oss inn i den filosofiske diskusjonen om hva en mengde *er* i noen større grad enn i diskusjonen om hva et tall er for noe.
- Det viktigste for oss er å forstå hvordan vi bruker mengder til å uttrykke oss presist.
- Hvis man noen gang har tenkt til å diskutere teoriproblemer med noen andre, er det viktig at man har den samme forståelsen.
- En viktig del av denne forståelsen er å vite når to mengder er like.

Definisjon.
To mengder A og B er like hvis de har nøyaktig de samme elementene.

- Det betyr at mengden er fullstendig bestemt av sine elementer, og det betyr ikke noe hvordan vi beskriver den.

Eksempel.

- $\{2,3,4\} = \{3,4,2\} = \{2,2^2,3\} = \{1+1,2,2+1,3,3+1,2+2,4\}$
- $\{3,4,7\} = \{\mathrm{III},\mathrm{IV},\mathrm{VII}\}$ så lenge det er klart at vi bruker arabiske og romerske måter å uttrykke tall på, mens hvis vi snakker om de konkrete symbolsekvensene er mengdene forskjellige.
- $$\{(x,y) \in \mathbb{R}^2 \ : \ 9x^2 + 16y^2 = 25\} =$$ $$\{(x,y) \in \mathbb{R}^2 \ : \ (\frac{3x}{5})^2 + (\frac{4y}{5})^2 = 1\}.$$

Eksempel (Tre "dumme" oppgaver).

- Finn $$\{x \in \mathbb{R} \ : \ x^4 + 4x^3 + 8x^2 + 8x + 4 = 0\}$$
- Bestem mengden av sannhetsverdifordelinger som gjør $$(p \vee q) \wedge (\neg p \wedge \neg q)$$ sann.
- Finn mengden av flytende-punkt representasjoner av det reelle tallet 0.

- Felles for alle disse tre oppgavene er at mengdene ikke har noen elementer.
- Alle løsningsmengdene har nøyaktig de samme elementene, nemlig ingen, og er derfor like.

Definisjon.
Vi lar $\emptyset$ betegne den tomme mengden, det vil si mengden som ikke har noen elementer.

- Symbolet $\emptyset$ er internasjonalt gjennomført, og man vil ikke finne det forklart andre steder enn i innføringstekster i mengdelære.
- Det er fremkommet ved en "null" med en strek over, og må ikke forveksles med noen bokstav i noe alfabet.
- Vi har at $\emptyset \neq \{\emptyset\}$; den første mengden har ingen elementer mens den andre har ett element, nemlig $\emptyset$ selv.
- Vi har tidligere sagt at når vi bruker en variabel i en programmeringssammenheng, må vi deklarere *typen* til variabelen.

  (Ikke alltid sant; det kan avhenge av programmeringsspråket, men for mange ikke-spesialiserte språk er det tilfelle.)

- Alternativt arbeider vi med en datastruktur hvor vi henter alle variabelverdier fra.
- Vi har ofte definert mengder som $\{n \in \mathbb{N} : \cdots\}$ eller $\{x \in \mathbb{R} : \cdots\}$. I slike tilfeller er det ofte klart fra sammenhengen hvilke mengder n eller x skal hentes fra, og det kan være brysomt å måtte presisere det hver gang.
- Hvis A er en mengde, og vi vil se på mengden av objekter som ikke er med i A, vil vi normalt ønske å avgrense oss til de objektene som er av interesse i den aktuelle sammenhengen.
- Alt dette gjør det aktuelt å innføre et eget symbol for en universell mengde, uten at denne universelle mengden trenger å være den samme i enhver sammenheng.

Definisjon.

a) Vi lar $\mathcal{E}$ stå for den universelle mengden.

b) $\mathcal{E}$ vil betegne forskjellige mengder i forskjellige sammenhenger, men skal ligge fast i enhver gitt sammenheng.

c) I noen sammenhenger (definisjoner, oppgaver) vil $\mathcal{E}$ bare betegne en vilkårlig universell mengde, mens i andre sammenhenger vil betydningen av $\mathcal{E}$ bli presisert.

Vi bruker $\mathcal{E}$ fordi læreboka gjør det. Det er imidlertid ingen fast notasjon for den universelle mengden uavhengig av forfatter og anvendelsesområde slik det er med $\emptyset$ for den tomme mengden.

Mengdealgebra

- Vi skal nå se på noen grunnleggende operasjoner på mengder.
- Disse kalles Boolske operasjoner.
- Det er en nær sammenheng mellom Boolske operasjoner og utsagnslogiske bindeord.
- Utsagnslogikk og Boolsk mengdelære er begge instanser av det som kalles Boolsk algebra.
- Et mer systematisk studium av Boolsk algebra er relevant for retninger innen teoretisk informatikk.

Mengdealgebra – union

Eksempel.

- Vi har fått i oppgave å finne en digital representasjon av funksjonen

$$f(x) = \sqrt{x^2 - 1}$$

  og vet at denne funksjonen bare er definert når $|x| \geqslant 1$.
- Det betyr at vi må finne mengden av 32-bit representasjoner av tall $\geqslant 1$ og av tall $\leqslant -1$ og så slå disse mengdene sammen.
- En slik sammenslåing kaller vi en union.

Definisjon.

- Hvis A og B er to mengder, definerer vi unionen av A og B som

$$A \cup B = \{x \ : \ x \in A \vee x \in B\}.$$

- Unionen av A og B består altså av de objektene som er element i minst en av de to mengdene A og B, men også gjerne i begge.

Eksempel.

- $\mathbb{N}$ er unionen av mengdene av positive partall og positive oddetall.
- Løsningsmengden til ulikheten $x^2 > 1$ er

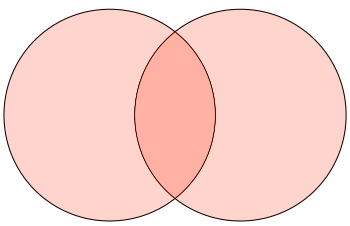
$$\{x \ : \ x < -1\} \cup \{x \ : \ x > 1\}.$$

- Mengden av 32-bitsekvenser som representerer tall x med $|x| \geqslant 1$ kan skrives som

$$\{01\sigma \ : \ \sigma \in \{0,1\}^{30}\} \cup \{11\sigma \ : \ \sigma \in \{0,1\}^{30}\}$$

  hvor $\{0,1\}^{30}$ er en vanlig måte å uttrykke mengden av 30-bits sekvenser.

- En vanlig måte å illustrere Boolske operasjoner på er ved å bruke Venndiagrammer.
- Et Venndiagram for en Boolsk kombinasjon av to eller tre mengder vil bestå av en sirkel for hver mengde, slik at de overlapper hverandre.
- Hvis vi har bruk for å markere den universelle mengden $\mathcal{E}$, gjør vi det ved et rektangel som omslutter alle sirklene.
- Ved å bruke forskjellige skraveringer, kan vi illustrere hvilke punkter som ligger i mengden og hvilke som ikke gjør det.



$A \cup B$

- Vi har en sirkel for hver mengde.
- Vi har fire felter, et for hver kombinasjon av $x \in A$ og $x \in B$.
- Hele det fargede området markerer $A \cup B$.

Mengdealgebra – snitt

Eksempel.

- La A være mengden av naturlige tall n slik at n kan deles på både 6 og 8.
- La B være mengden av reelle tall x slik at

$$((x-1)^2-1)^2+(x^4-16)^4=0$$

- La C være mengden av digitale representasjoner av ikke-negative partall.
- I tilfelle A gir vi to krav direkte.
- I tilfelle B krever vi at vi både må ha at $(x-1)^2-1=0$ og at $x^4-16=0$.
- I tilfelle C krever vi at bitsekvensen må ha den foreskrevne lengde, og både må starte med 0 og ende med 0.

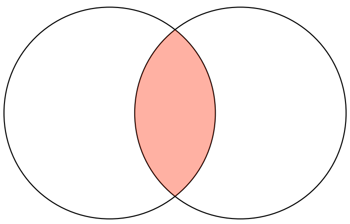
Definisjon.

- La A og B være to mengder.
- Med snittet av A og B mener vi

$$A \cap B = \{x \;:\; x \in A \land x \in B\}.$$

- $A \cap B$ består altså av de objektene som er elementer både i A og i B.

Vi kan også illustrere snitt ved et Venndiagram.



$A \cap B$

- Vi farger eller skraverer det feltet som markerer fellesdelen av de to mengdene.

Eksempel.

- Hvis

$$A = \{0,1,2,3,4,5\}$$
$$B = \{1,3,5,7,9\}$$

er $A \cup B = \{0,1,2,3,4,5,7,9\}$, mens $A \cap B = \{1,3,5\}$.
- Hvis

$$A = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \geqslant 1\}$$
$$B = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leqslant 1\}$$

er $A \cup B$ hele planet, mens $A \cap B$ er mengden av punkter på enhetssirkelen.

Mengdealgebra – komplement

- Vi har ikke hatt bruk for den universelle mengden $\mathcal{E}$ i definisjonen av union og snitt.
- Vi har formulert definisjonen av $\cup$ og $\cap$ slik at sammenhengen med $\vee$ og $\wedge$ skal komme klart frem.
- Den neste mengdeoperasjonen vi skal se på er komplement.
- Den vil ha et nært slektskap til $\neg$.
- For å definere komplement må vi ha tilgang til en universell mengde.

Eksempel.

1. La P betegne mengden av primtall.

   Et sammensatt tall er et naturlig tall $\neq 1$ som ikke er et primtall.

   Vi kan definere
   $$S = \{n \in \mathbb{N} \ : \ n \notin P \cup \{1\}\}.$$
   Det hadde vært enklere om vi kunnet skrive dette på en kortere måte som vi alle kunne forstå.

Eksempel (Fortsatt).

2. Vi definerer de irrasjonale tallene som mengden av de reelle tallene som ikke er rasjonale.

   Dette kunne vi skrevet som
   $$\{x \in \mathbb{R} \ : \ x \notin \mathbb{Q}\}.$$
   Hvis alle forstår oss om vi skriver $\overline{\mathbb{Q}}$ istedenfor, ville det vært greiere.

Eksempel (Fortsatt).

3. Vi har ord både for partall og for oddetall, men vi har ikke et eget ord for tall som ikke kan deles på 13 (eller 17 for den saks skyld).

   I denne sammenhengen var det klart at vi snakker om hele tall, dvs
   $$\mathcal{E} = \mathbb{J}.$$

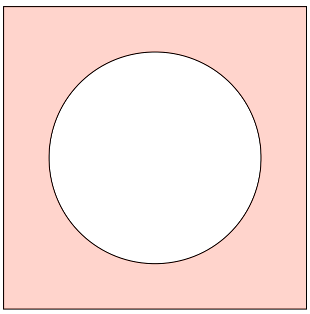
Vi vil si at mengden av tall som ikke kan deles på 13 er komplementet til mengden av tall som kan deles på 13, og vi markerer komplementet til en mengde ved å sette en strek over uttrykket for mengden.

Definisjon.

La $\mathcal{E}$ være en universell mengde og la A være en mengde hvor alle elementene ligger i $\mathcal{E}$. Med komplementet til A mener vi

$$\overline{A} = \{x \in \mathcal{E} \,:\, x \notin A\} = \{x \in \mathcal{E} \,:\, \neg(x \in A)\}.$$

Når vi skriver $\overline{A}$ skal det alltid være klart at $\mathcal{E}$ er kjent
(eller at vi arbeider generelt med en vilkårlig $\mathcal{E}$.)
Vår notasjon for komplement er ikke uvanlig, men på ingen måte enerådende.
Vi kan beskrive komplementet ved et Venndiagram.



$\overline{A}$

- Det fargede/skraverte feltet markerer komplementet.

Eksempel.

- Hvis $\mathcal{E} = \{0, 1, 2, 3, 4, 5, 6, 7\}$ og $A = \{2, 3, 5\}$, er $\overline{A} = \{0, 1, 4, 6, 7\}$.
- Hvis $\mathcal{E}$ er som over og $B = \{1, 3, 5, 7\}$, er $\overline{B} = \{0, 2, 4, 6\}$
- Hvis $\mathcal{E} = \mathbb{J}$ og $B = \{1, 3, 5, 7\}$ er

$$\overline{B} = \{\cdots, -3, -2, -1, 0, 2, 4, 6, 8, 9, 10, \cdots\}.$$

Mengdealgebra – mengdedifferens

- Den siste mengdeoperasjonen vi skal innføre er mengdedifferens.

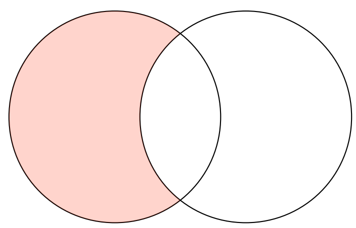
Definisjon.
Hvis A og B er to mengder, er differansen A minus B definert ved

$$A - B = \{x \,:\, x \in A \wedge x \notin B\}$$

Vi bruker ofte betegnelsen mengdedifferens

En alternativ skrivemåte mye brukt i litteraturen er $A \setminus B$.

Vi kan også illustrere mengdedifferens ved et Venndiagram.



$A - B$

- Det fargede/skraverte området markerer differensen
- Vi har ikke hatt bruk for $\mathcal{E}$ her.
- $A - B = A \cap \overline{B}$

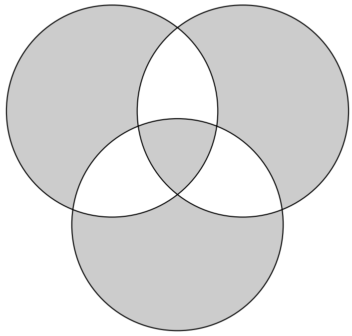
Venndiagrammer

- Den virkelige nytten av Venndiagrammer ligger i at de kan brukes til å studere sammenhengen mellom forskjellige Boolske uttrykk.
- I en viss forstand er bruk av Venndiagrammer en parallell til bruk av sannhetsverditabeller.
- Vi kommer til å vise noen eksempler på tavlen på hvordan vi kan etablere mengdeteoretiske identiteter ved hjelp av Venndiagrammer.

Oppgave.
Vi definerer ofte symmetrisk differens ved

$$A \triangle B = (A - B) \cup (B - A).$$

a) Illustrer $A \triangle B$ ved et Venndiagram.
b) Vis at $(A \triangle B) \triangle C$ kan illustreres ved Venndiagrammet på neste side.
c) Drøft hvorfor dette viser at vi kunne skrevet $A \triangle B \triangle C$ uten bruk av parenteser.



$(A \triangle B) \triangle C$