Logikk og Mengdelære

Dag Normann
Universitetet i Oslo
Matematisk Institutt
Boks 1053 - Blindern
0316 Oslo

29. mars 2005

Innhold

Forord

Dette er et hefte skrevet for kurset

MA-EVU6, "Utsagnslogikk, mengdelære og matematiske bevisteknikker".

Heftet vil bli skrevet over tid, og alt blir ikke tilgjengelig med en gang.
Kapittelstrukturen ble endret underveis. Det opprinnelige Kapittel 5 - Relasjonslogikk går ut og de opprinnelige kapitlene 6-8 blir nå kapitler 5-7. Et nytt kapittel 8 vil inneholde fasit og veiledning til løsninger av noen oppgaver. Dette er resultatet av et ønske som kom frem under første samling. Kapittel 8 vil bli skrevet på i ledige stunder, og etter rimelig store utvidelser vil nye versjoner legges ut. Kapitlene 1-3 var temaet for første samling, kapitlene 4 - 5 vil bli temaet for andre samling og kapitlene 6 - 7 temaet for tredje samling.

Hvert kapittel er delt opp i avsnitt, hvorav det siste kalles "Utfordringer". Disse utfordringsavsnittene er ikke ment å inngå i pensum, men representerer nettopp utfordringer for lesere som har lyst på dem.
Pr. i dag er resten av stoffet planlagt som pensum, men det vil bli vurdert underveis i hvilken grad dette er rimelig.

Til hvert avsnitt er det gitt et sett med oppgaver. Dette vil være rene øvelsesoppgaver. Til slutt i hvert kapittel finner du en samling 'blandede oppgaver'. Dette var ment å være oppgaver av samme type som de som vil bli brukt til eksamen. Ettersom vi vil samle opp begreper og temaer underveis, vil oppgavene under dette punktet bare bli av realistisk vanskelighetsgrad etterhvert som vi kommer utover i stoffet. Det er for eksempel begrenset hvor mange eksamensoppgaver det er mulig å formulere på bakgrunn av stoffet i Kapittel 1 alene. Underveis ble det imidlertid bestemt at eksamen skal gis i form av tre hjemmeeksamener, slik at disse oppgavene også bare må sees på som øvingsoppgaver som bidrar til å øke forståelsen av stoffet.
Noen få av oppgavene vil bli merket med [u]. Disse må sees på som utfordringer som vil falle vanskeligere enn den typiske eksamensoppgaven.

Det vil bli utarbeidet en indeks for begreper som innføres. Den blir fortløpende oppdatert, og det kan være aktuelt å laste den inn på nytt hver gang det kommer et tillegg til manuskriptet.
Dette forordet vil også kunne bli oppdatert. Dette markeres ved at datoen endres.

Dette heftet må ikke oppfattes som en lærebok. Teksten og illustrasjonene er skrevet i LaTex, et tekstbehandlingsystem for matematikk, og stoffet er ikke viderebehandlet grafisk slik man vil gjøre med en moderne lærebok med bruk av farvede felter, margbemerkninger og andre virkemidler som gjør stoffet mer oversiktlig. Teksten blir nok heller ikke så rikholdig forsynt med eksempler som praksis er for moderne lærebøker, og tiden vil ikke strekke til for at utvalget av oppgaver skal bli så godtsom ønskelig.

Med disse forbehold ønskes leseren "Lykke til" !

Blindern 16/02 - 05

Dag Normann

Innledning

I en viss forstand er matematikk et statisk fag. Når først et matematisk resultat er bevist, så er det bevist for all fremtid. I den forstand skiller matematikken seg fra naturvitenskapene, hvor det som regnes som etablerte sannheter egentlig er teorier, teorier som kan etterprøves gjennom eksperimenter og hvor gyldighetstområdet for teoriene gjerne avgrenses etterhvert som man sprenger grensene mot de små detaljene og mot de store avstandene og hastighetene.

I en annen forstand er matematikk et dynamisk fag. Byggverket av matematisk kunnskap utvides stadig, og områdene hvor matematikk, til dels avansert matematikk, viser seg nyttig blir stadig større og bredere.

Høsten 1972 og våren 1973 vikarierte jeg en del på en realskole. Da var den moderne matematikken i skuddet, og en av de matematikk-klassene jeg underviste i skulle lære moderne matematikk. En vakker vårdag etter den skriftlige eksamen tok jeg, mot rektors instruks, klassen ut til en friluftstime. Der pratet vi litt løst og fast om matematikk, og et av spørsmålene elevene stilte meg var:

Når oppsto den moderne matematikken?

Elevene ble både overasket og litt skuffet da jeg fortalte at det meste de hadde lært som moderne matematikk oppsto i andre halvdel av 1800-tallet, en del til og med tidligere. Det dreide seg altså om matematikk som man først etter ca. 100 år fant ut at burde inn i skolematematikken i Norge og i store deler av resten av verden.

Nå er "moderne matematikk" som et vekkelsesfenomen utdødd. Det betyr ikke at det stoffet som den gang ble påtvunget elevene er unyttig, men hvor nytten lå ble nok litt misforstått.

I dette kurset skal vi ta for oss mengdelære og noe logikk. Vi skal også se på hvordan bevis føres i en normal matematisk tekst. Bruk av mengdebegrepet i passe doser er en god hjelp når man skal strukturere matematiske tankeganger, problemet oppstår når mengdelæren blir hovedsaken og ikke et hjelpemiddel. Kjennskap til litt logikk hjelper også til en bedre forståelse av matematikken. Det å kunne gjenkjenne riktige argumentasjoner er selvfølgelg viktig når man skal lese matematikk og eventuelt formulere argumenter på egen hånd.

I dette kurset er logikk og mengdelære hovedtemaet, og vi vil nødvendigvis etterlate det inntrykket at disse temaene er verd et studium for sin egen del. For den som har interesse for matematikkens mange spesialiseringer er dette inntrykket også riktig, både mengdelære og logikk er spennende grener av matematikken. Målet med å tilby et eget EVU-kurs i logikk og mengdelære er likevel ikke at disse emnene igjen skal bli noe mer enn støttefag i skolen, en god støtte for den som skal undervise og i moderate mengder en støtte for elevene der det kan hjelpe til med forståelsen av andre deler av matematikken.

Jeg sa innledningsvis at matematikken er dynamisk ved at den finner stadig nye anvendelsesområder. Det er også slik at disse anvendelsesområdene har en innflytelse på hvilke problemstillinger matematikerne finner interessante. Et fagområde som knapt er 60 år gammelt er *informatikken*, eller læren om datamaskiner, deres virkemåte og bruk. Mange av de temaene vi skal ta for oss i dette

heftet har fått fornyet interesse fordi de er nyttige i informatikk. Alt fagstoffet vi presenterer vil være en naturlig del av et universitetsstudium i informatikk overalt i verden, som en del av samlebegrepet *diskret matematikk*. Det vil føre alt for langt å gå inn på disse anvendelsene i detalj her, men det kommer noen antydninger underveis.

Det er ikke et mål at den som har fulgt dette kurset skal bli en kløpper i å skrive matematiske bevis. Det som vil bli prøvet til eksamen er noen tekniske ferdigheter som viser at man har lært seg det viktigste om de begrepene pensum omfatter. I en viss forstand er kapitlet om *induksjonsbevis* et untak. Det er meningen at man skal kunne føre et induksjonbevis, både i den tradisjonelle forstanden og i den mer generelle. Poenget med å innføre de generelle induksjonsbevisene er å avmystifisere bevisformen, ved at vi ser at den er nært knyttet til visse typer definisjoner av matematiske strukturer.

Kapittel 1

Mengdelære

1.1 Hva er en mengde?

Mengdebegrepet gjennomsyrer mye av matematikken i dag, både i skolematematikken og høyere opp i systemet. En *mengde* (engelsk: *Set*, tysk:*Menge*) er en samling av objekter. Begrepet må forståes på samme måte som vi må forstå tall, det er grunnbegreper i matematikken som vi ikke kan forklare ved hjelp av andre begreper. Faktisk er det slik at hvis vi godtar og forstår mengdebegrepet, så kan vi definere hva vi mener med *tall*. Vi skal ikke gå i detalj her, men se kort på hvordan tall kan 'defineres" fra mengder i avsnittet om utfordringer.

En mengde er en samling objekter. Når vi skal øve inn tallbegrepet for elever i grunnskolen kan vi tegne tre melkespann i en grønn sky og tre kyr i en blå sky, og så kan vi påstå at det er like mange kyr som melkespann. Samlingen av melkespann og samlingen av kyr er da to mengder som i en viss forstand er like store. Dette er selvfølgelig litt kontraintuitivt, ettersom tre kyr tar mye større plass enn tre melkespann. Likevel kan vi si at på en måte er mengdene like store når antall *elementer* er det samme.
I avsnittet over innførte vi et nytt begrep, elementene. En mengde er altså en samling objekter, og disse objektene kaller vi mengdens *elementer*.

I matematikken arbeider vi ikke så mye med mengder av melkespann og kyr, men heller med mengder av objekter som oppstår i matematikken. Vi kan snakke om mengden $\mathbb{N}$ av *naturlige tall* hvor objektene er tallene $1, 2, 3, \ldots$, om mengden $\mathbb{Z}$ av *hele tall* hvor objektene er $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$. Vi skal komme tilbake til andre tallmengder senere. Noen mengder, for eksempel mengden $\mathbb{R}$ av *reelle tall*, kan vi ikke forklare eller beskrive som over. De reelle tallene er alle tall som vi med rimelighet kan presse inn på tall-linjen, tall som $\sqrt{2}$, π, $\sqrt[17]{15 + 8\pi}$ etc. Denne forklaringen halter litt, men det er naturlig siden det å gi en mer presis definisjon av $\mathbb{R}$ krever en posjon generell mengdelære. Vi tar opp dette punktet igjen i avsnittet om utfordringer.

Hvis vi først aksepterer mengden $\mathbb{R}$ av reelle tall, kan det gi mening å definere

for eksempel *enhetssirkelen* som mengden av de punkter (x, y) i planet slik at

$$x^2 + y^2 = 1.$$

Finnes det så mengder av andre typer objekter enn melkespann, kyr, tall og tallpar? Selvfølgelig gjør det det. Hvis vi har gitt et dataprogram kan det være av interesse å se på mengden av for-løkker i programmet. Hvis vi studerer et programmeringsspråk er selvfølgelig mengden av programmer (som ikke inneholder skrivefeil) av interesse, og muligens mengden av programmer som kan kjøres på den tilgjengelige maskinvaren i løpet av en time.
Driver vi med spillteori er mengden av utfall i et forsøk med kast av 23 terninger av en mulig interesse, og for å kunne vurdere nytten av å spille lotto kan det være en fordel å se på mengden av de rekker som gir gevinst, og sammenlikne størrelsen på denne mengden med størrelsen av mengden av alle mulige lotto-rekker.
Vi vil se på mengder hvor elementene selv er mengder en god del i dette heftet.

Dette avsnittet hvor vi ramser opp mengder som er av interesse i en eller annen sammenheng kunne gjøres mye lengere, men det er på tide å stoppe her, og heller gå inn på mengdelæren som sådan.

Intuitivt består en mengde av objekter, og den skal helst ikke ha andre aspekter enn det. Dette uttrykker vi ved å definere når to mengder skal være like:

Definisjon 1.1 La A og B være to mengder.
$A = B$ hvis A og B har nøyaktig de samme elementene.

Dette betyr at om vi går tilbake til det innledende eksemplet med melkespann og kyr, flytter kyrne over på en grønn eng og melkespannene over på en melkerampe, så har vi fortsatt de to samme mengdene. Det betyr også at om vi gir to forskjellige beskrivelser av mengder, så er mengdene like såfremt elementene er de samme.

Eksempel 1.1 La A være mengden av naturlige tall mindre enn ti som er faktor i 105 og la B være mengden av oddetall mindre enn åtte som samtidig er primtall.
I begge tilfellene kan vi se at mengden består av tallene 3, 5 og 7 og ikke mer, så $A = B$.

Eksempel 1.2 La A være mengden av negative kvadrattall i $\mathbb{Z}$, og la B være mengden av reelle tall x som er løsning av likningen

$$x^2 + 2x + 5 = 0.$$

Nå vet vi at alle hele tall som samtidig er kvadrattall er ≥ 0, og vi vet at det ikke finnes noe reelt tall som er løsning av likningen. I begge tilfeller kan vi altså argumentere for at mengdene ikke har noen elementer. Det betyr faktisk at de har de samme elementene, nemlig ingen, så de må være like.

Eksempel 1.2 kan synes litt kunstig, vi snakker om mengder uten elementer. Med litt ettertanke ser vi imidlertid at definisjonen av mengde B er naturlig. Har vi gitt en likning er vi selvfølgelig interessert i mengden av løsninger. Vi ønsker ikke å begrense oss til tilfeller hvor likninger har løsninger, siden det for likninger av høyere grad enn 2 ikke alltid er opplagt om det finnes løsninger eller ikke (dette gjelder bare likninger hvor graden er et partall, oddegradslikninger har alltid løsninger). Det er altså bekvemmelig for matematikerne å snakke om mengder uten elementer, men alle mengder uten elementer må være like:

Definisjon 1.2 Den *tomme mengden* er mengden uten elementer. Vi bruker spesialsymbolet $\emptyset$ for den tomme mengden

Finnes det noen måte å uttrykke at en mengde er større enn en annen? Her må vi være litt forsiktige, for bruker vi ordet 'større', må det bety at vi har et mål på størrelse av mengder. Innledningsvis antydet vi at antall er et mulig mål på størrelse. Vi skal se på en annen naturlig tolkning av 'større enn'.

Eksempel 1.3 La A være mengden av primtall > 2 og la B være mengden av oddetall.
Siden alle primtall > 2 også er oddetall, ser vi at alle elementene i A også er elementer i B.
Siden det finnes oddetall som ikke er primtall, er B i en viss forstand større enn A. Vi vil si at A er ekte inneholdt i B.

Eksempel 1.4 La A være mengden av de tre kyrne på en blå sky vi så på i det innledende eksemplet, og la B være mengden av de samme tre kyrne på en grønn eng, nå sammen med en okse. Igjen har vi at alle elementene i A også er elementer i B, så A er inneholdt i B

Vi gir nå den formelle definisjonen.

Definisjon 1.3 La A og B være mengder.
Vi sier at A er *inneholdt* i B dersom alle elementene i A også er elementer i B.
Vi skriver $A \subseteq B$ for 'A er inneholdt i B'.
Hvis $A \subseteq B$ og B har minst et element som ikke er element i A, sier vi at A er *ekte inneholdt* i B og vi skriver $A \subset B$.

Bemerkning 1.1 Den symbolbruken vi har valgt er ganske vanlig, men ikke enerådende. I noen tekster vil vi finne $\subset$ brukt for 'inneholdt i' og $\subsetneq$ for 'ekte inneholdt i'.

En vanlig måte å uttrykke at en sammenheng ikke holder, er å sette en strek over symbolet som uttrykker sammenhengen. Derfor vil vi la $A \not\subseteq B$ bety at A ikke er inneholdt i B.
Hvis $A \subseteq B$ sier vi også at A er en *delmengde* av B, og hvis $A \subset B$ sier vi at A er en *ekte delmengde* av B.

Vi har innført noen symboler som skal hjelpe oss til å uttrykke visse former for sammenheng mellom mengder. I all bruk av mengdelære får vi selvfølgelig også

ofte bruk for å si at et objekt er element i en mengde. Derfor har vi innført et eget symbol for dette:

$$a \in A$$

uttrykker at objektet a er et element i mengden A. Vi har for eksempel at $17 \in \mathbb{N}$, $-17 \in \mathbb{Z}$ og $\pi \in \mathbb{R}$. Vi kan også uttrykke at vi vil snakke om et vilkårlig element i en mengde ved å skrive noe slikt som

$$\text{La } x \in \mathbb{Z} \text{ være gitt. Da er } x^2 \in \mathbb{N}.$$

Det som står over er ikke en riktig påstand, og vi kunne skrive et motargument som følger:

$$0 \in \mathbb{Z} \text{ og } 0^2 = 0 \notin \mathbb{N}, \text{ så } 0^2 \notin \mathbb{N}. \text{ Derfor er dette feil.}$$

Her har vi også brukt symbolet $\notin$ som betyr at objektet beskrevet foran $\notin$ ikke er element i mengden beskrevet etter $\notin$.

Oppgaver til avsnitt 1.1

Oppgave 1.1.1 Vi har gitt fem mengder under. Bestem hvilke som er like og hvilke som er forskjellige:

1. A er mengden av tallene -1, 0 og 1.
2. B er mengden av hele tall som er kvadratet av seg selv.
3. C er mengden av hele tall x slik at $0 \leq x^2 \leq 1$.
4. D er mengden av reelle tall x slik at $0 \leq x^2 \leq 1$.
5. E er mengden av reelle tall x slik at x^2 er et element i B.

Oppgave 1.1.2 For mengdene i oppgaven over, avgjør hvilke som er inneholdt i hvilke og avgjør hvilke som er ekte inneholdt i hvilke.

Oppgave 1.1.3 Vi sier at et helt tall n er *delelig* med et helt tall m hvis forholdet $\frac{n}{m}$ er et helt tall.
For de fire mengdene under, bestem hvilke som er inneholdt i hvilke:

1. A er mengden av hele tall som er delelige med 18.
2. B er mengden av hele tall som er delelig med 9.
3. C er mengden av hele tall som er delelig med 6.
4. D er mengden av hele tall som er delelig med 3.

Oppgave 1.1.4 La A, B, C og D være som i oppgaven over.
Bestem hvilke av disse påstandene som er sanne og hvilke som er usanne:
$36 \in A$, $21 \in B$, $27 \in B$, $6 \in D$, $3 \in C$, $9 \in C$ og $27 \in A$

Oppgave 1.1.5 Hvis $A \subseteq \mathbb{N}$ er mengden av positive kvadrattall, hva er de rette påstandene?

$$120 \in A \ , \ 121 \in A \ , \ 144 \notin A \ , \ 17 \notin A$$

Oppgave 1.1.6 Forklar hvorfor $A \subseteq B$ og $B \subseteq A$ medfører at $A = B$.

Oppgave 1.1.7 Forklar hvorfor $A \subseteq B$ og $B \subseteq C$ medfører at $A \subseteq C$.

Oppgave 1.1.8 La A være en vilkårlig mengde. Forklar hvorfor $\emptyset \subseteq A$.

1.2 Hvordan beskriver vi en mengde?

I det forrige avsnittet definerte vi endel mengder ved å gi en forklaring på hvilke elementer mengden har. Det burde ikke være noen tvil om hva som er elementer i disse mengdene. Hvis vi skal bruke mengdelære som et verktøy trenger vi imidlertid et mer egnet språk for å definere mengder, ettersom den direkte metoden vi har brukt til nå blir for omstendelig når vi skal beskrive mere komplekse mengder.

Universalmidlet for beskrivelser av mengder er bruk av klammeparentesene { og }.
Hvis vi i en matematisk tekst skriver $\{blablabla\}$ vil det alltid være underforstått at vi definerer en mengde, og at *blablabla* gir en nærmere beskrivelse av hvordan denne mengden ser ut. Vi kaller { } for *mengdebyggeren*.

Eksempler 1.5 Eksempler på hvordan vi kunne brukt mengdebyggeren i avsnittet over vil være:

1. $\{0, 1\}$ for mengden som har tallene 0 og 1 som elementer.
2. $\{-1, 0, 1\}$ for mengden som har tallene -1, 0 og 1 som elementer.
3. { Dagros, Litago, Selsdokka } for mengden av de tre kyrne.
4. $\mathbb{N} = \{1, 2, 3, \ldots\}$
5. $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$

Foreløpig har vi vunnet litt, men ikke så mye, vi har erstattet frasen 'mengden hvor elementene er $\cdots$' med $\{\cdots\}$. Vi har fortsatt en opplisting av elementene i mengden, enten fullstendig eller ved å angi et mønster for hvordan et vilkårlig element skal være.
Hvis vi skriver ned en fullstendig liste, er det selvfølgelig ingen problemer med å vite hvilke elementer mengden består av. Når vi bruker disse tre prikkene, forutsetter vi imidlertid at forfatter og leser har en felles forståelse av hva mønsteret er. Dette er ofte for usikkert til bruk i matematikk.
Når vi skriver opp en mengde på listeform, er det bare hvilke elementer som er listet opp som betyr noe, ikke hvilken rekkefølge vi har skrevet dem opp i og ikke hvor mange ganger vi eventuelt har skrevet opp et element. Det betyr at alle mengdene beskrevet under er like.

$$\{0,1,2\},\ \{1,2,0\},\ \{1,0,1,2,0,2,1\}$$

Det finnes situasjoner hvor vi bruker uttrykk for de enkelte elementene i en mengde, og hvor vi i utgangspunktet ikke vet hvilke uttrykk som beskriver like objekter, så derfor er det et poeng at vi tillater oss å beskrive mengder på listeform hvor samme objekt kan bli listet opp flere ganger. Sålenge det ikke finnes saklige grunner for å gjøre det, er det imidlertid en liten uskikk å ha repetisjon når vi lister opp elementene i en endelig mengde.

Definisjon 1.4 La A være en mengde og la $P(\ )$ uttrykke en egenskap.

a) Med $\{x \mid P(x)\}$ mener vi mengden av alle objekter x slik at egenskapen P holder for x.

b) Med $\{x \in A \mid P(x)\}$ mener vi mengden av alle objekter i A slik at x har egenskapen i P

Vi ser at vi bruker et uttrykk på formen $\{\cdot \mid \cdots\}$ hvor $\cdot$ er et uttrykk for et typisk element i mengden, eventuelt med ytre begrensninger på hva slags elementer vi vurderer, mens $\cdots$ representerer en nærmere presisering av mengden.
I bruk av mengdebyggeren er det viktig å bruke en avveining mellom presis og for rigid språkbruk, det viktigste er at leseren forstår hvilken mengde vi beskriver når vi bruker mengdebyggeren. Det er derfor på sin plass med mange eksempler:

Eksempler 1.6 a) En typisk definisjon kan se ut som følger:
La
$$\mathbb{R}^2 = \{(x,y) \mid x,y \in \mathbb{R}\}.$$
Det vi her mener er at et typisk element skal være et par, og at begrensningen skal være at paret utgjøres av reelle tall.

b) La $C = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Det vi her har definert er mengden av punkter i planet som ligger på sirkelen om origo med radius 1.

c) La $A = \{n \in \mathbb{Z} \mid n^4 + 7 = 23\}$
Denne mengden kunne vi lettere ha skrevet som $A = \{-2, 2\}$.

d) La $B = \{n^2 \mid n \in \mathbb{Z}\}$
Dette er mengden av kvadrattall, når vi regner 0 med som et av kvadrattallene.

e) La $C = \{x \in \mathbb{R} \mid x^3 - 3x^2 + 3x - 1 = 0\}$
Det forventes ikke at leseren uten videre skal innse det, men her kunne vi like gjerne skrevet $C = \{1\}$.

Vi kan bruke mengdebyggeren til å konstruere nye mengder fra gamle. Et viktig eksempel er potensmengden til en mengde:

Definisjon 1.5 La A være en mengde.
Med *potensmengden* $\mathcal{P}(A)$ mener vi mengden av alle delmengder av A, det vil si
$$\mathcal{P}(A) = \{B \mid B \subseteq A\}.$$

For små mengder er det mulig å skrive opp alle elementene i $\mathcal{P}(A)$, for eksempel er

$$\mathcal{P}(\{0,1,2\}) = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{1,2,3\}\}.$$

Vi ser at når A har tre elementer, vil $\mathcal{P}(A)$ ha $2^3 = 8$ elementer. Dette er ingen tilfeldighet. Generelt vil $\mathcal{P}(A)$ ha 2^n elementer når A har n elementer.

Bemerkning 1.2 Da vi definerte potensmengden til en mengde A, definerte vi den som mengden av alle mengder med en viss egenskap, i dette tilfellet å være delmengde av A. Som vi skal se i avsnittet om utfordringer, må man av og til være litt forsiktig når man definerer nye mengder på denne måten. Hvis A er en endelig mengde hvor vi kjenner elementene, vet vi i prinsippet at vi kan skrive ned elementene i potensmengden til A. dette med 'i prinsippet' er faktisk alvorlig ment. Hvis A har for eksempel 20 elementer, vil $\mathcal{P}(A)$ ha $2^{20} = 1038576$ elementer, og det krever mer enn god evne til ryddighet å liste opp alle elementene i en slik mengde på noen stykker papir.
Hvis A er uendelig, kan vi selvfølgelig ikke beskrive potensmengden på noen form for listeform. Likevel er det gunstig fra et matematisk synspunkt å kunne arbeide med potensmengder til uendelige mengder. I avsnittet om utfordringer skal vi se at hvis vi skal kunne konstruere de reelle tallene som en tallmengde, så trenger vi potensmengden til en uendelig mengde.

Oppgaver til avsnitt 1.2

Oppgave 1.2.1 Beskriv disse mengdene med ord:

1. $A = \{x^3 \mid x \in \mathbb{N}\}$
2. $B = \{z \in \mathbb{R} \mid z^3 + 3z^2 + 3z + 1 = 0\}$
3. $C = \{a^2 \in \mathbb{Z} \mid a^2 + a + 1 = 0\}$
4. $D = \{(x, y) \in \mathbb{R}^2 \mid x + y = 1\}$
5. $E = \{(x, x^2) \mid x \in \mathbb{R}\}$

Oppgave 1.2.2 Undersøk hvilke av påstandene under som er riktige:

1. $21 \in \{p \in \mathbb{N} \mid p + 2$ er et primtall $\}$.
2. $\{n^4 \mid n \in \mathbb{N}\} \subseteq \{a^2 \mid a \in \mathbb{Z}\}$
3. $\emptyset \in \mathcal{P}(\{1, 4, 6\})$
4. $\{n \in \mathbb{N} \mid n^2 < 24\} \subseteq \{m \in \mathbb{Z} \mid m^3 < 24\}$

Oppgave 1.2.3 Beskriv $\mathcal{P}(\emptyset)$. Forklar hvorfor denne mengden ikke er den samme som den tomme mengden.

Oppgave 1.2.4 Forklar hvorfor $\mathcal{P}(A) \subseteq \mathcal{P}(B)$ når $A \subseteq B$.
Forklar hvorfor omvendingen også holder, det vil si at hvis $\mathcal{P}(A) \subseteq \mathcal{P}(B)$, så må $A \subseteq B$.

Oppgave 1.2.5 Forklar hvorfor $\mathcal{P}(A)$ har 2^n elementer når A har n elementer.

1.3 Snitt, union og komplement

Etterhvert skal vi utvikle et språk og et begrepsapparatur som gjør det lettere å beskrive mengder som vi av forskjellige grunner kunne være interesserte i. I dette avsnittet skal vi se på litt av dette.

Eksempel 1.7 Vi har rotet oss opp i en situasjon hvor vi ønsker å se på de punktene i x, y, z-rommet som både ligger på en gitt elipsoide (leseren trenger ikke å vite hva en elipsoide er) og i et gitt plan. For å konkretisere kan vi være interesserte i punktene (x, y, z) som både ligger i mengden

$$\{(x, y, x) \in \mathbb{R}^3 \mid x^2 + 4y^2 + 9z^2 = 36\}$$

og i mengden

$$\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}.$$

Det vi er på jakt etter er det utsnittet av elipsoiden som ligger i planet, og vi vil kalle en slik fellesmengde for *snittet* av mengdene.

Eksempel 1.8 La T være mengden av naturlige tall a slik at både a og $a + 2$ er primtall. Vi ser at $11 \in A$ fordi både 11 og 13 er primtall. Andre tall vil være 17, 29 og 59. Disse tallene kaller vi *primtallstvillinger*, og vi vet ikke om denne mengden er endelig eller uendelig.
Mengden av primtallstvillinger utgjøres av de tall som ligger både i mengden

$$\{n \in \mathbb{N} \mid n \text{ er et primtall }\}.$$

og i mengden

$$\{n \in \mathbb{N} \mid n + 2 \text{ er et primtall }\}.$$

Mengden av primtallstvillinger er derfor, slik vi har definert dem, snittet av disse to mengdene.

Definisjon 1.6 La A og B være to mengder. Med *snittet* av A og B (engelsk: intersection) mener vi mengden av objekter som både er elementer i A og i B. Vi skriver $A \cap B$ for snittet av A og B.

En ting vi ser med en gang er at rekkefølgen på A og B ikke spiller noen rolle, for alle mengder A og B har vi

$$A \cap B = B \cap A.$$

I oppgavene under vil leseren få anledning til selv å bestemme snittet i endel tilfeller.
Vi skal nå se på hva vi mener med *unionen* av to mengder:

Eksempel 1.9 Av en eller annen grunn er vi blitt interesserte i fjerdegradslikningen

$$x^4 - 7x^3 + 17x^2 - 17x + 6,$$

og ved å regne på uttrykket har vi funnet ut at likningen kan skrives om til

$$(x^2 - 3x + 2)(x^2 - 4x + 3) = 0.$$

Dette gir oss to annengradslikninger,

$$x^2 - 3x + 2 = 0$$

med løsningsmengde $\{1, 2\}$ og

$$x^2 - 4x + 3 = 0$$

med løsningsmengde $\{1, 3\}$. Vi finner løsningsmengden til den opprinnelige likningen ved å slå disse to mengdene sammen til $\{1, 2, 3\}$. En slik sammenslåing av mengder kaller vi en *union*.

Definisjon 1.7 La A og B være to mengder. Med unionen av A og B mener vi mengden av objekter som ligger i minst en av A og B.
Vi skriver $A \cup B$ for unionen av$A$ og B.

Også for union ser vi at rekkefølgen ikke spiller noen rolle, $A \cup B = B \cup A$.
En annen ting vi ser er at vi kan utvide definisjonene av snitt og union til å omfatte mer enn to mengder.
Hvis $A_1, \ldots, A_n$ er mengder, lar vi

$$A_1 \cap \cdots \cap A_n$$

være mengden av de objektene som er i alle A_i'ene, og

$$A_1 \cup \cdots \cup A_n$$

være mengden av objekter som er i minst en av A_i'ene. Igjen spiller ikke rekkefølgen vi skriver opp mengdene i noen rolle.

Eksempel 1.10 La $A_1 = \{0, 1, 2\}$, $A_2 = \{0, 2, 3\}$ og $A_3 = \{0, 1, 3, 4\}$.
Da er $A_1 \cap A_2 \cap A_3 = \{0\}$ fordi 0 er det eneste tallet som er element i alle tre mengdene, mens $A_1 \cup A_2 \cup A_3 = \{0, 1, 2, 3, 4\}$ fordi alle disse tallene finnes i minst en av mengdene. 4 finnes i én av dem, 1,2 og 3 i to mengder hver og 0 i alle tre. Alle fem tallene kvalifiserer derfor for å være med i unionen.

Vi skal se på tre andre mengdeoperasjoner, *komplement*, *differens* og *symmetrisk differens*. Komplementet til en mengde A er intuitivt sett det som ikke er i mengden A. Dette gir imidlertid ikke mening hvis vi ikke har avgrenset oss til en sammenheng hvor det er klart hvilke objekter vi i det hele tatt interesserer oss for.

Eksempel 1.11 Hvis vi bedriver tallteori og har definert mengden P av primtall, vil komplementet være alle tall som ikke er primtall. Da er $\mathbb{N}$ det *universet* vi befinner oss i. Komplementet til P vil bestå av tallet 1 samt alle sammenstatte tall

Eksempel 1.12 Hvis vi er interesserte i heltallsaritmetikk, utgør $\mathbb{Z}$ et univers. Hvis A_3 er mengden av tall som er delelig med 3, vil komplementet til A_3 være mengden av heltall som ikke er delelig med 3.

Eksempel 1.13 La $\mathbb{Q}$ være mengden av rasjonale tall, det vil si alle tall som kan skrives på brøkform.
De *dyadiske tallene* er

$$D = \{\frac{a}{2^n} \mid a \in \mathbb{Z} \text{ og } n \in \mathbb{N} \cup \{0\}\}.$$

De dyadiske tallene er de som kan skrives på endelig desimalform i totallsystemet. Komplementet til D vil i dette tilfellet være alle rasjonale tall som ikke kan skrives på endelig desimalform i totallsystemet.

Vi er nå klar til å gi de formelle definisjonene:

Definisjon 1.8 a) Med et *univers* U mener vi en mengde som i den sammenhengen vi opererer i omfatter alle objekter av interesse.

b) Hvis U er et omforent univers og $A \subseteq U$ vil *komplementet* til A være mengden av elementer i U som ikke ligger i A.
Vi skriver A^c for komplementet til A.

Bemerkning 1.3 Definisjonen av et 'univers' er upresis. På den annen side gir det ikke noen mening i å snakke om komplementet til en mengde hvis vi ikke er presis på i forhold til hvilket univers vi tar komplementet.
Matematisk sett kan enhver mengde U benyttes som et univers, det er bare praksis som tilsier at et univers er mengden av objekter vi for tiden er interesserte i.

Eksempler 1.14 a) La A være mengden av partall og la B være mengden av tall som kan deles på fem. Vi får da mengden av partall som ikke kan deles på fem ved å ta A og 'trekke fra' B, det vil si at vi tar bort alle elementene i B. Dette er et eksempel på *mengdedifferens*.

b) La Z være mengden av punkter i planet som ligger innenfor sirkelen om origo med radius 1, men ikke innenfor sirkelen om $(1, 0)$ med radius 1. Vi finner Z ved å ta mengden X definert ved

$$X = \{(x, y) \mid x^2 + y^2 < 1\}$$

og så trekke fra mengden Y definert ved

$$Y = \{(x, y) \mid (x - 1)^2 + y^2 < 1\}.$$

Definisjon 1.9 Vi definerer *mengdedifferens* ved $X \setminus Y$ er mengden av de objekter som ligger i X men ikke i Y.

Vi ser at hvis vi arbeider innenfor et univers U vil vi ha at

$$X \setminus Y = X \cap Y^c.$$

Vi har nå gått igjennom de fire viktigste mengdeoperasjonene. Den siste operasjonen, som kalles *symmetrisk differens*, er definert ved

Definisjon 1.10 La A og B være to mengder. Med den *symmetriske differensen* mener vi

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

Oppgaver til avsnitt 1.3

Oppgave 1.3.1 Forklar med ord hvilke elementer den symmetriske differensen av A og B har.

Oppgave 1.3.2 Bestem $A \cap B$, $A \cup B$ og $A \setminus B$ for følgende par A og B av mengder:

1. $A = \{3, 5, 8, 9\}$ og $B = \{3, 6, 9, 11\}$.
2. $A = \{x \in \mathbb{R} \mid x + 1 > 0\}$ og $B = \{x \in \mathbb{R} \mid x < 0\}$.
3. A er mengden av kvadrattall og B er mengden av kubikktall i $\mathbb{Z}$.
4. $A = \{2k + 1 \mid k \in \mathbb{N}\}$ og $B = \{2k \mid k \in \mathbb{N}\}$.
5. $A = \{(x, y) \in \mathbb{R}^2 \mid x \geq 1\}$ og $B = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$. Formuler svaret med den blandingen av norsk og mengdenotasjon som du finner mest hensiktsmessig.

Oppgave 1.3.3 I hvert av punktene under har vi gitt et univers og fire mengder A, B, C og D. Bestem hvilken av de tre siste mengdene som er komplementet til A:

1. $U = \mathbb{N}$, A er mengden av partall, B er mengden av tall delelige med 3, C er mengden av oddetall og D er mengden av kvadrattall.
2. $U = \mathbb{R}^2$, $A = \{(x, y) \mid x^2 + y^2 = 1\}$, $B = \{(x, y) \mid x^2 + y^2 < 1\}$, $C = \{(x, y) \mid x^2 + y^2 > 1\}$ og $D = \{(x, y) \mid x^2 + y^2 \neq 1\}$.
3. $U = \mathbb{Z}$, $A = \{3a + 2 \mid a \in U\}$, $B = \{3a \mid a \in U\}$, $C = \{3a + 1 \mid a \in U\}$ og $D = C \cup B$.

Oppgave 1.3.4 Anta at A er en delmengde av et univers U. Gjelder det alltid at $A = (A^c)^c$?

Oppgave 1.3.5 Forklar hvorfor vi alltid har at

$$\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B).$$

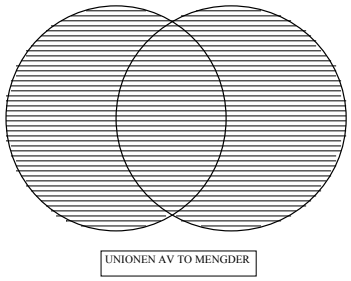
Oppgave 1.3.6 La A og B være mengder.
Når vil $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$?
Begrunn svaret.

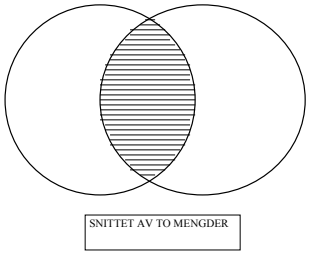
Oppgave 1.3.7 La A, B, C og D være fire mengder.
Forklar hvorfor vi alltid har at

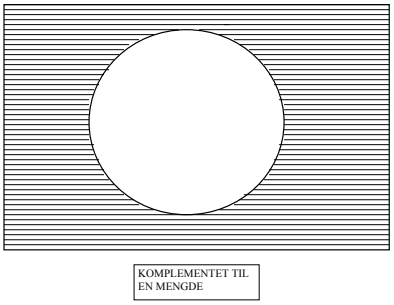
$$(A \cup B) \triangle (C \cup D) \subseteq (A \triangle C) \cup (B \triangle D).$$

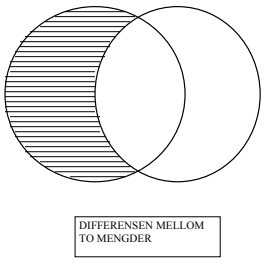
1.4 Venn-diagrammer

Et *Venn-diagram* er en tegning som gjør det oversiktlig å tolke sammensetninger av operasjonene vi så på over. Vi markerer hver mengde med en sirkel slik at alle mulige kombinasjoner av det å være med i noen mengder men ikke i andre, blir representert ved et område i planet. Venn-diagrammer kan være hensiktsmessige når vi skal se på to eller tre mengder. Hvis vi er i en situasjon med et univers, markerer vi gjerne det som et stort rektangel som omfatter alle sirklene vi har brukt til å markere mengdene.
Når vi skal markere resultatet av en operasjon, markerer vi det gjerne ved å skravere feltet vi kommer frem til. Illustrasjonene under er ment å være selvforklarende. De representerer fire av de mengdeoperasjonene vi har sett på til nå.





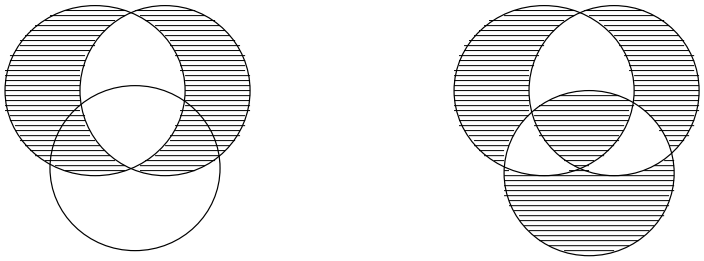




Som et eksempel skal vi tegne Venn-diagrammet til

$$(A \triangle B) \triangle C$$

i to skritt. A er representert av sirkelen øverst til venstre, B av sirkelen øverst til høyre og C av sirkelen nederst. Vi tegner først inn $A \triangle B$ i diagrammet, og deretter $(A \triangle B) \triangle C$.



Den siste tegningen viser at rekkefølgen på A, B og C ikke spiller noen rolle, heller ikke hvordan vi setter parentesene. Dette gjelder også om vi har flere enn tre mengder, noe vi kommer tilbake til i oppgave 7.1.5.

Mange mengdeteoretiske identiteter kan vises ved å representere hva uttrykkene betyr via Venn-diagrammer. I oppgavene under utfordres leseren til å bekrefte noen slike påstander som vi kommer tilbake til i neste avsnitt.

Oppgaver til avsnitt 1.4

Oppgave 1.4.1 Bruk Venn-diagrammer til å vise at identitetene under holder for alle mengder A, B og C:

1. Første distributive lov:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

2. Andre distributive lov:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

3. $B \setminus (A \setminus B) = B$.
4. $B \setminus (A \cap B) = B \setminus A$.

Oppgave 1.4.2 Anta at vi er i en situasjon hvor vi har et mengdeunivers U. Bruk Venn-diagrammer til å vise at følgende identiteter holder:

1. DeMorgans første lov:

$$(A \cup B)^c = A^c \cap B^c.$$

2. DeMorgans andre lov:

$$(A \cap B)^c = A^c \cup B^c.$$

3. $A \setminus B = A \cap B^c$.

Oppgave 1.4.3 Bruk Venn-diagrammer til å bestemme om følgende påstander er riktige eller gale. Du kan anta at det finnes et univers U der det trengs for å gjøre spørsmålet meningsfylt.

1. $A \cup (B \setminus C) = B \cup (A \setminus C)$.
2. $(A \setminus B)^c = B^c \setminus A^c$.
3. $(A \setminus B) \setminus C = (A \setminus C) \setminus B$.
4. $(A \setminus B) \setminus C = (B \setminus A) \setminus C$.

1.5 Regneregler

I forrige avsnitt brukte vi Venn-diagrammer til å illustrere sammenhengen mellom de forskjellige mengdeoperasjonene som 'union', 'snitt', 'differens' og 'komplement'. Spesielt brukte vi oppgavene til å etablere noen viktige sammenhenger. I skolematematikken bruker vi gjerne ordet 'algebra' om regning med kjente og ukjente tall, men hvor vi hele tiden antar at tallene vi regner med er reelle. Det innebærer blant annet at vi bruker de fire regningsartene og i noen utstrekning kvadratrøtter. Regnereglene for oppløsning av parenteser, skifte av fortegn og liknende følger av at vi regner med tall og bokstaver som står for reelle tall.

Nå har vi begynt å regne med bokstaver som betegner vilkårlige mengder, og vi bruker tegnene $\cap$, $\cup$, $\setminus$ og ^{c} for å forme uttrykk. På samme måte som i vanlig algebra, finnes det her regler for å regne med slike uttrykk. De fleste reglene har vi sett på under avsnittet om Venn-diagrammer.

Utgangspunktet for disse regnereglene er at vi befinner oss i en *mengdealgebra* av delmengder av et felles univers, slik at vi kan forme snitt, union og komplement. Merk at mengdedifferens er definerbart ved

$$A \setminus B = A \cap B^c$$

så vi trenger ikke å gi spesielle regler for mengdedifferens.

Teorem 1.1 *Regning med mengder respekterer følgende regler:*

1. $(A^c)^c = A$
2. $U^c = \emptyset$
3. $A \cup A^c = U$ *og* $A \cap A^c = \emptyset$
4. $A \cap \emptyset = \emptyset$, $A \cup \emptyset = A$
5. $A \cap B = B \cap A$ *og* $A \cup B = B \cup A$
6. $A \cap (B \cap C) = (A \cap B) \cap C$ *og* $A \cup (B \cup C) = (A \cup B) \cup C$

7. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *og* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

8. $(A \cup B)^c = A^c \cap B^c$ *og* $(A \cap B)^c = A^c \cup B^c$

Selv om vi kaller dette for et teorem, kommer vi ikke til å bevise det. Noen av identitetene er gitt som oppgaver i forrige avsnitt, og resten er enten helt opplagte eller bevises lett ved å se på et Venn-diagram.

Vi skal se at det er enkelte generelle sammenhenger vi kan utlede fra disse regnereglene. Det betyr at vi ikke trenger å ta frem Venn-diagrammer for å få bekreftet at de er riktige. Det skal heller ikke være nødvendig å bruke definisjonene av mengder eller av snitt, union etc. Poenget her er ikke at identitetene er vanskelige å utlede fra definisjonene, men at de følger rent mekanisk fra reglene over. Det vil være mulig å programmere en datamaskin slik at den vil kunne utlede mange identiteter ved hjelp av disse reglene. Vi kan ikke på samme måte programmere en datamaskin til å forstå en innføring i mengdelære.

Eksempel 1.15 $A \cup U = ((A \cup U)^c)^c = (A^c \cap U^c)^c = (A^c \cap \emptyset)^c = \emptyset^c = U$

Eksempel 1.16 $(A \cap B) \cup (A \cap B^c) = A \cap (B \cup B^c) = A \cap U = A$

I begge disse eksemplene kan hvert regneskritt tilbakeføres til en av reglene vi ga i Teorem 1.1.
Det er ikke slik at alle reglene i Teorem 1.1 er innbyrdes uavhengige, det er mulig å fjerne et par av dem, og så regne seg frem til at de vi har fjernet holder likevel. Vi velger å ikke gå mere detaljert inn på dette.

Oppgaver til avsnitt 1.5

Oppgave 1.5.1 Bruk regnereglene til å vise at vi alltid har at $A \cap U = A$

Oppgave 1.5.2 Bruk regnereglene til å vise at vi alltid har at
$(A \cup B) \cap (A \cup B^c) = A$.

Oppgave 1.5.3 Regnereglene kan være nyttige når vi skal se på generelle sammenhenger mellom fire eller fler mengder. Da blir metoden med bruk av Venndiagrammer ofte uoversiktelig.
Bruk regnereglene til å vise at

$$(A \cup B) \cap (C \cup D) = (A \cap C) \cup (A \cap D) \cup (B \cap C) \cup (B \cap D).$$

Drøft hvor ille det er at vi droppet en del parenteser i uttrykket bak likhetstegnet.
Kan du skrive opp og vise en tilsvarende likning når vi bytter om på $\cap$ og $\cup$?

Oppgave 1.5.4 Det følger ved bruk av Venn-diagrammer at

$$(A \setminus B) \setminus C = (A \setminus C) \setminus B.$$

Kan du vise dette ut fra den algebraiske definisjonen av mengdedifferens og regnereglene vi har gitt?
Sagt på en annen måte: Kunne en datamaskin utledet dette?

1.6 Utfordringer

1.6.1 Russells paradoks

Da vi definerte potensmengden til en mengde, advarte vi mot at det ikke er bare-bare å definere en mengde som mengden av de mengder for hvilke noe holder. Russells paradoks er et eksempel på hvor galt det kan gå.
Anta at vi definerer en mengde X ved

$$X = \{Y \mid Y \notin Y\}.$$

Hvis X er en veldefinert mengde, må én av to holde:

1. $X \in X$
2. $X \notin X$

Vi ser at om 1. holder, så oppfyller ikke X kravet til å få være med i X, så da vil $X \notin X$, dvs. da holder 2.
På den annen side, hvis 2. holder, oppfyller X kravet til å få være med i X, så da holder 1.
Dette er en umulig situasjon, siden 1. og 2. ekskluderer hverandre. Konklusjonen må være at vi ikke kan konstruere X som en mengde. Det at definisjonen av X leder til en selvmotsigelse er kjent som *Russells paradoks* etter den britiske matematikeren og filosofen Bertrand Russell.

Tidlig på 1900-tallet jobbet man hardt for å finne et grunnlag for mengdelæren som tillater de mengdekonstruksjonene matematikere har bruk for, men som ungår situasjoner som i Russells paradoks. Den norske matematikeren Thoralf Skolem spilte en betydelig rolle i dette arbeidet. Det fører alt for langt å gå inn på hvordan dette grunnlaget ser ut. Leseren kan imidlertid være trygg på at alle de konstruksjonene vi vil gjøre i dette heftet er i tråd med dette grunnlaget.

1.6.2 De naturlige tallene

Vi har startet de naturlige tallene med 1, men i en del fremstillinger av mengdelære er det naturlig å starte med 0, fordi vi vil at de naturlige tallene skal svare til mulige antall elementer i mengder. Hvis vi ikke antar noe annet enn intuisjonen om mengder som gitt, er det først og fremst én mengde vi kan legge vår klamme hånd på, og det er den tomme mengden $\emptyset$. Hvis vi skal tolke 0 som en mengde, er da $\emptyset$ den naturlige kandidaten fordi det er den eneste mengden med null elementer. Siktemålet er da at vi for ethvert mulig antall på en naturlig måte skal finne en mengde som har det aktuelle antall elementer, og som vi derfor kan bruke som den mengdeteoretiske tolkningen av tallet.
Når vi først har overbevist oss om at $\emptyset$ er en mengde, har vi også en mengde med ett element, nemlig $\{\emptyset\}$, og siden dette er den første mengden med ett element vi kommer over, er det naturlig å la denne mengden svare til tallet 1. Men nå har vi funnet to objekter, nemlig $\emptyset$ og $\{\emptyset\}$. Dette er den første mengden med to elementer vi i en viss forstand kan konstruere fra prinsippene fra mengdelære

alene, så hvorfor ikke la denne mengden svare til 2.
Vi ser at $0 = \emptyset$, $1 = \{0\}$ og $2 = \{0, 1\}$. Vi kan følge denne tankerekken videre og sette $3 = \{0, 1, 2\}$, $4 = \{0, 1, 2, 3\}$ og så videre.
Vi ser at på denne måten finner vi en en naturlig måte å tolke alle ikke-negative heltall som mengder på, mengder som har nøyaktig det aktuelle antallet elementer.
Vi må nok en gang minne om at den eneste grunnen til å definere de naturlige tallene på denne måten er at vi vil vise at det er mulig å tufte matematikkem på mengdelærens prinsipper. For vanlig matematisk praksis kan vi betrakte naturlige tall som matematiske grunn-objekter som ikke trenger noen nærmere definisjon.

1.6.3 De reelle tallene

I skolematematikken problematiserer vi ikke eksistensen av de tallene vi snakker om. Vi bruker negative tall så snart vi får bruk for dem, og vi regner med rasjonale tall, for eksempel $\frac{5}{17}$, og tar eksistensen for gitt. Slik må det være, for skulle vi hefte oss med slike eksistensielle problemer i skolen, ville vi ikke fått tid til å lære den for de fleste elevene nyttige delen av matematikken.
I dette heftet har vi også tatt både naturlige tall og hele tall for gitt, og på samme måte vil vi ta de rasjonale tallene for gitt:

Definisjon 1.11 La de *rasjonale tallene* være definert ved

$$\mathbb{Q} = \{\frac{n}{m} \mid n \in \mathbb{Z} \text{ og } m \in \mathbb{N}\}.$$

For nå å definere den reelle tall-linjen ser vi strengt tatt på mengden av *tilnærminger* vi kan gjøre med rasjonale tall. For å gi en presis definisjon av hva vi mener med en tilnærming, bruket vi potensmengden til $\mathbb{Q}$:

Definisjon 1.12 Et Dedekind snitt er en delmengde A av $\mathbb{Q}$ slik at

1. $A \neq \emptyset$.
2. Hvis $p \in A$ og $q < p$, vil $q \in A$ (hvor $p, q \in \mathbb{Q}$ er underforstått).
3. A er begrenset, det vil si at det finnes $p \in \mathbb{Q}$ som er større enn alle $q \in A$.
4. A har ikke et største element.

Tanken til Dedekind er at hvis vi har et reelt tall x, så er det bestemt av $\{p \in \mathbb{Q} \mid p < x\}$. Da kan vi like gjerne bruke denne mengden til å representere x, for på denne måten å konstruere en mengde som har de egenskapene vi vil at $\mathbb{R}$ skal ha.
Det vil føre for langt å gå videre inn på hvor dette fører og på hvilke problemer vi må løse før vi har konstruert en fullstendig matematisk modell for den reelle tall-linjen. Det må likevel presiseres at de fleste matematikere ikke vil ha bruk for å bekymre seg om denne typen grunnlagsproblemer, vi har nok intuisjon om $\mathbb{R}$ til å utnytte den til vår matematiske fordel, uten å ha en konkret mengdeteoretisk modell å forholde oss til.

1.7 Blandede oppgaver

Oppgave 1.7.1 La A, B og C være mengder slik at $A \subseteq C$.

a) Vis at $A \cap B \subseteq B \cap C$.

b) La $D = A \cap B$ og $E = B \setminus A$. Vis at hvis $E \subseteq C$ så vil $B \subseteq C$.

c) Bestem $\mathcal{P}(D) \cap \mathcal{P}(E)$.

Oppgave 1.7.2 Anta for enkelthets skyld at $\mathbb{Q} \subseteq \mathbb{R}$ og la $\mathbb{R}$ være universet for våre betraktninger.

a) Vil $\mathbb{Q} \subseteq \{\frac{n}{m} \mid n \in \mathbb{N} \text{ og } m \in \mathbb{N}\}$?

b) Elementene i $\mathbb{R} \setminus \mathbb{Q}$ kalles *irrasjonale tall*.
La a være irrasjonal og la $\mathbb{Q} + a = \{b + a \mid b \in \mathbb{Q}\}$.
Vis at $\mathbb{Q} \cap (\mathbb{Q} + a) = \emptyset$.

c) La a og b være to irrasjonale tall.
Vis at enten har vi at $\mathbb{Q}+a = \mathbb{Q}+b$ eller så har vi at $(\mathbb{Q}+a) \cap (\mathbb{Q}+b) = \emptyset$.
Forklar med ord når vi har det ene og når vi har det andre.

Oppgave 1.7.3 Vi betrakter $\mathbb{N}$ som en delmengde av $\mathbb{Z}$.

a) La $A = \{a \in \mathbb{Z} \mid a^2 \in \mathbb{N}\}$.
Vil $A = \mathbb{Z}$?

b) La B_0 være mengden av tall i $\mathbb{Z}$ som kan deles på 3, B_1 mengden av tall hvor resten blir 1 og B_2 mengden av tall i $\mathbb{Z}$ hvor resten blir 2 om vi deler tallet på 3.
For $i, j \in \{0, 1, 2\}$, la $B_i \cdot B_j = \{a \cdot b \mid a \in B_i \text{ og } b \in B_j\}$.
Vis at $B_2 \cdot B_2 = B_1$.

c) Finn en generell regel for å "beregne" $B_i \cdot B_j$.

Oppgave 1.7.4 La X være en mengde og la $A \subseteq X$.
Vi sier at to delmengder C og D av X er *like på* A hvis $C \cap A = D \cap A$. Tilsvarende sier vi at de er *like utenom* A dersom $C \cap (X \setminus A) = D \cap (X \setminus A)$.

a) Vis at hvis C og D er like både på og utenom A, så er de like.

b) For hver $C \subseteq X$, la

$$[C] = \{D \subseteq X \mid D \text{ er lik } C \text{ på } A\}.$$

Vis at om C_1 og C_2 er to delmengder av X, så er $[C_1] = [C_2]$ eller $[C_1] \cap [C_2] = \emptyset$.

c) Vi definerer tilsvarende $[C]^c$ som mengden av de D som er lik C utenom A.
Bestem $[C] \cap [C]^c$ for en vilkårlig mengde C.

Kapittel 2

Utsagnslogikk

2.1 Hva er et utsagn?

I dagligtale kan vi ytre en masse forskjellige utsagn, som "Vinden løyer mot natten", "Det går tjue egg på et snes", "Lasse Lund er en fremragende fotballspiller", $4^2 = 16$, "Noen mennesker blir mindre verdsatt enn andre" og mange andre. Felles for disse eksemplene er at vi ytrer en påstand som andre enten kan være enige eller uenige i. På norsk kan et utsagn være så mangt, spørsmål, uhøvelige kraftuttrykk m.m., men i logikk vil vi begrense oss til å se på utsagn som uttrykker en påstand som enten er sann eller usann.

Det er normalt å være enda mere tilbakeholden enn eksemplene over tilsier. Vi er alle enige om at det går tjue egg på et snes, så dette er en påstand med en uomtvistelig sannhetsverdi, påstanden er sann.
Påstanden om at vinden løyer mot natten er av en litt annen kategori, dette er en påstand om at noe vil skje, og det er ingen problemer med at vi er enig i påstanden klokken 17 på ettermiddagen, i et optimistisk håp om at vi kan dra ut etter makrellen tidlig neste morgen, men innser at vi tok feil før vi går til sengs.
Utsagnet om Lasse Lund, som er syv år og spiller på knøttelaget til Tjuvholmen Sportsklubb, ble fremmet av moren hans. De andre foreldrene som følger med på kampen er ikke uten videre enige, men avslår høflig å kommentere påstanden. Dette er derfor et annet eksempel på en påstand hvor det ikke er klart om den er sann eller usann.
Det eneste eksemplet over som direkte egner seg for en logisk analyse er påstanden om størrelsen på et snes. Utsagnslogikk er et studium av utsagn og sammensetninger av utsagn, hvor vi antar at hvert enkelt utsagn er uomtvistelig sant eller uomtvistelig usant.

Begrensningene i hva vi på denne måten vil regne som utsagn gjør at de fleste utsagn fra dagligtale faller utenom. Likevel vil vi kunne behandle disse utsagnene som om de tilfredstilte disse begrensningene, og dermed studere hva som er sant og usant på et rent logisk grunnlag. Vi kan også studere hvor-

dan sannhetsgehalten i komplekse utsagn avhenger av hva den enkelte mener om delpåstander eller i hvilken grad et resonement kan sies å bygge på skjulte antagelser, representere feil tankegang eller utgjøre et logisk holdbart argument.

Matematiske påstander skal jo helst være enten sanne eller usanne. Eksempler på sanne påstander er

- $2 < 3$
- $5 + 7 = 12$
- *Det finnes uendelig mange naturlige tall*
- *Hvis* $X \subseteq Y$ *så vil* $\mathcal{P}(X) \subseteq \mathcal{P}(Y)$

Eksempler på usanne påstander kan være

- $3 < 2$
- $5 + 7 = 17$
- *Alle hele tall er positive*
- *For alle mengder* X *og* Y *gjelder det at* $\mathcal{P}(X) \cup \mathcal{P}(Y) = \mathcal{P}(X \cup Y)$

En annen type påstander som vi kan treffe på i matematikken er for eksempel likninger

$$x^2 - 3x + 2 = 0$$

eller ulikheter

$$\frac{x^2 - 3x + 2}{x^2 - 1} > 0.$$

I disse eksemplene har vi en *variabel* x, i likningsammenheng også kalt *en ukjent*, og sannhetsverdien til påstanden varierer med hvilken verdi vi gir x. Oppgaven vil gjerne gå ut på å finne de verdiene av x som gjør påstanden sann.

Oppgaver til avsnitt 2.1

Oppgave 2.1.1 Drøft i hvilken grad ytringene under kan betraktes som utsagn i den forstand vi har diskutert over:

1. Tror du at skøyteløperne våre vinner noen medaljer i Torino?
2. Det er flere egg i et snes enn i et dusin.
3. Måtte en vennlig sjel tildele din kjøttfulle bakende et realt spark!
4. $2^{10} = 2048$.
5. 7 er et primtall.
6. Hvis du går over bekken etter vann, blir du våt på beina.

7. $x^3 - 3x^2 + 3x - 1 = 0$.

Drøft også om utsagnene er sanne, usanne eller om sannhetsverdien er situasjonsavhengig.

Oppgave 2.1.2 For noen av utsagnene under er sannhetsverdien situasjonsbestemt, mens andre vil være enten sanne eller usanne. Bestem hvilke som er sanne, hvilke som er usanne og hvilke som har en situasjonsbestemt sannhetsverdi:

1. Per, Pål og Espen er tre brødre.
2. Per, Pål og Espen Askeladd er tre skikkelser som forekommer ofte i norske folkeeventyr.
3. Anders Jensen er far til Jens Andersen
4. Enten er månen en grønn ost eller så er den det ikke.
5. Det eneste tallet som tilfredstiller likningen

   $$2x - 3 = 0$$

   er 0.
6. Hvis $x^2 > 1$ er enten $x > 1$ eller $x < -1$.
7. 21 og 23 er to primtall.
8. $3 < 5$ og $7^2 > 29$.
9. $0 \leq 1$.

2.2 En matematisk modell

I det forrige avsnitet så vi på en del eksempler på utsagn, og den vesentlige egenskapen et utsagn har er at det enten er sant eller usant. For noen utsagn vil dette være uavhengig av den sammenhengen vi ser det i, mens sannhetsverdien for andre utsagn er situasjonsbestemt.

Når vi nå skal utvikle en matematisk analyse av sammensatte utsagn, må vi lage en *matematisk modell*. Matematiske modeller forekommer i omtrent alle situasjoner hvor vi vil bruke matematikk til å studere andre fenomener. Vi bruker for eksempel den reelle tall-linjen som en modell for tiden, og vi brukjer $\mathbb{R}^3$ som en modell for verdensrommet. Så lenge vi arbeider med Newtons mekanikk, og med hastigheter og avstander hvor Newtons mekanikk kan brukes til å forutsi fysiske fenomener, så er $\mathbb{R} \times \mathbb{R}^3$ en god matematisk modell for *tid-rommet*, hvor vi har fanget inn akkurat nok aspekter til at vi kan bruke matematikk på en nyttig måte.

Når vi skal lage en matematisk modell for utsagn, er det aspektet at et utsagn er sant eller usant som vi vil fange inn. Derfor vil vi la *mengden av sannhetsverdier* $\{\top, \bot\}$ som har et element $\top$ som representerer 'sant' og et element $\bot$

som representerer 'usant' være vår grunnmengde, eller vårt univers.
I dette avsnittet vil vi være opptatt av hvordan sannhetsverdien til et sammensatt utsagn avhenger av sannhetsverdiene til grunnutsagn. Vi vil innføre *utsagnsvariable* for slike grunnutsagn. Vi bruker bokstavene P, Q, A. B etc. som utsagnsvariable, og vi skal tenke oss at i denne sammenhengen vil for eksempel P stå for et utsagn som er sant eller usant i seg selv, og ikke som et utslag av sin oppbygging. Vi kan også la disse bokstavene stå for samensatte utsagn, men da regner vi bare med at vi kjenner til sannhetsverdien til dette utsagnet, ikke hvorfor denne sannhetsverdien er som den er.
Vi satser på at presisjonsnivået er godt nok for formålet med dette kurset. Det er mulig å holde et høyere presisjonsnivå, men erfaringen er at logikk som fag da blir unødig vanskelig.

Siden det bare er sannhetsgehalten i et utsagn som er av interesse for oss i denne sammenhengen, vil vi la sannhetsverdiene $\top$ og $\bot$ være de to mulige verdiene en utsagnsvariabel kan ha. Det betyr at utsagnsvariablene er variable over mengden $\{\top, \bot\}$, og kunne like gjerne vært kalt sannhetsverdivariable. Når vi kaller dem utsagnsvariable er det både fordi dette er fonetisk enklere og fordi det er i tråd med praksis i tilsvarende litteratur.

I endel eksempler vil vi oversette sammensatte utsagn i dagligtale til utsagn hvor vi ertatter grunnutsagn med utsagnsvariable og bruker bindeordene som vi skal se på i neste avsnitt. Det betyr at vi bruker disse bokstavene som variable over mengden av alle utsagn, noe som er vanlig. Problemet med denne praksisen er at den er uformell, vi har aldri definert noen mengde av utsagn, og vi har liten mulighet til å gjøre det. Vi må oppfatte slike oversettelser på samme måte som vi må oppfatte oversettelsen av et naturvitenskapelig fenomen til en likning mellom reelle parametre, vi utnytter den matematiske modellen.

Oppgaver til avsnitt 2.2

Oppgave 2.2.1 I noen sammenhenger er det aktuelt å tolke et utsagn som noe som er sant med en viss sannsynlighet p og usant med sannsynlighet $1 - p$. Et eksempel er utsagnet "jeg får en sekser når jeg kaster denne terningen" hvor sannsynligheten er $\frac{1}{6}$ for at utsagnet er sant og $\frac{5}{6}$ for at utsagnet ikke er sant.
Diskuter hva tolkningsrommet til utsagnsvariablene bør være hvis vi skal dekke et behov for å studere utsagn av denne typen.
Kan du finne utsagn fra andre områder enn terningspill hvor denne typen utsagn kan være aktuelle å studere.

Oppgave 2.2.2 I neste avsnitt vil vi tolke sammensatte utsagn, hvor grunnutsagnene er utsagnsvariablene $P_1, \ldots, P_n$, som funksjoner som gir en sannhetsverdi til hver kombinasjon av sannhetsverdier for $P_1, \ldots P_n$.
Forklar hvorfor det finnes 2^n forskjellige måter å fordele sannhetsverdier på utsagnsvariablene på og hvorfor det derfor finnes 2^{2^n} forskjellige slike funksjoner.
Diskuter hvorfor dette betyr at vi kan regne med å formulere høyst 256 essensielt forskjellige sammensatte utsagn hvis vi begrenser oss til utsagnsvariablene A, B og C.

2.3 Bindeord og sannhetsverditabeller

Hvis vi påstår at jorda er rund, så vil nok de fleste leserne være enige, men også være oppmerksome på at det finnes mennesker som ikke har akseptert dette ennå. Hvis vi imidlertid hevder at Norge er en ledende langrennsnasjon på herresiden, er trolig selv ikke alle leserne enige i om dette er sant eller usant.
Hvis vi hevder at jorda er rund eller jorda er ikke rund, så vil alle forhåpentligvis si seg enige, og de vil si seg enige av samme grunn hvis vi sier at Norge er en ledende langrennsnasjon på herresiden eller Norge er ikke en ledende langrennsnasjon på herresiden. Det eneste vi påstår med disse utsagnene er at et annet utsagn, vi trenger ikke å vite hvilket, er sant eller så er det ikke sant.
Hvis vi drister oss til å påstå at jorda er rund og jorda er ikke rund i samme utsagn, vil imidlertid de fleste være enige i at vi må ta feil, ettersom en påstand ikke kan være båd sann og usann på samme tid.

Ord og uttrykk som 'og', 'eller' og 'det er ikke slik at' kan vi bruke til å lage sammensatte utsagn fra grunnutsagn, og sannhetsverdien til det sammensatte utsagnet vil være bestemt av sannhetsverdiene til grunnutsagnene. Vi vil innføre noen spesialsymboler for slike ord og uttrykk, og vi vil kalle dem for *logiske bindeord* eller *konnektiver*.

2.3.1 Negasjon

Definisjon 2.1 Hvis A er et utsagn, er $\neg A$ et utsagn.
$\neg A$ leses som *'det er ikke slik at* A og kalles *negeringen* av A.
Hvis utsagnet A har verdien $\top$ vil utsagnet $\neg A$ få verdien $\bot$ og omvendt.
Dette uttrykker vi med følgende sannhetsverditabell:

A	$\neg A$
$\top$	$\bot$
$\bot$	$\top$

Vi kan bruke sannhetsverditabellen til å finne ut av hvordan sannhetsverdien til $\neg(\neg A)$ avhenger av sannhetsverdien til A. Det gjør vi ved å regne ut verdiene i en søyle ut fra verdiene i søylen til venstre, som i dette eksemplet:

Eksempel 2.1 Sannhetsverdien til et dobbelnegert utsagn finner vi ved

A	$\neg A$	$\neg(\neg A)$
$\top$	$\bot$	$\top$
$\bot$	$\top$	$\bot$

Ikke overaskende ser vi at dobbelnegasjon av A ikke endrer sannhetsverdien overhode.

2.3.2 Konjunksjon

Vi brukte bindeordet 'og' i eksemplene i innledningen til dette avsnittet, og betydningen må jo være at et og-utsagn er sant nøyaktig når begge delutsagnene er sanne. Dette uttrykker vi ved

Definisjon 2.2 Hvis A og B er utsagn, er $A \wedge B$ et utsagn.
Vi kaller $A \wedge B$ for *konjunksjonen* av A og B, og leser A *og* B.
Sannhetsverdien til $A \wedge B$ er bestemt fra sannhetsverdiene til A og B ved følgende tabell:

A	B	$A \wedge B$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$

Igjen kan vi bruke disse tabellene til å finne ut av hvordan sannhetsverdien til et sammensatt utsagn avhenger av sannhetsverdiene til enkeltutsagnene. I dette eksemplet er hver søyle regnet ut ved å bruke tabellen for $\neg$ eller tabellen for $\wedge$ på en eller to søyler til venstre:

Eksempel 2.2 Finn sannhetsverditabellen til $\neg(\neg A \wedge B)$:

A	B	$\neg A$	$\neg A \wedge B$	$\neg(\neg A \wedge B)$
$\top$	$\top$	$\bot$	$\bot$	$\top$
$\top$	$\bot$	$\bot$	$\bot$	$\top$
$\bot$	$\top$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$	$\bot$	$\top$

Eksempel 2.3 Anta at vi blir bedt om å finne den logiske formen til den sammensatte påstanden

Det er ikke slik at jorda ikke beveger seg og sola går i bane rundt jorda.

Det kan være delte meninger om hvordan denne setningen skal tolkes, men en vanlig språkfortolkning tilsier at vi benekter konjunksjonen mellom *jorda beveger seg ikke* og *sola går i bane rundt jorda*.
Hvis vi lar A stå for *Jorda beveger seg* og B stå for *Sola går i bane rundt jorda* får vi det sammensatte utsagnet

$$\neg(\neg A \wedge B).$$

2.3.3 Disjunksjon

Vi har ikke gitt så mange eksempler fra dagligtale på kamuflert bruk av $\neg$ og $\wedge$, og ingen fra matematikken så langt, fordi det ikke er noe flertydighet i bruken av disse ordene.
Nå skal vi se på et logisk bindeord som svarer til taleordet *eller*, og her har vi to muligheter.

Eksempel 2.4 Hvis en mor sier til sin tenåringsdatterdatter: *Du kan få penger til en ny bukse eller til et nytt skjørt*, vil hun normalt mene at datteren kan få penger til en ny bukse eller hun kan få penger til et nytt skjørt, men hun får ikke penger til begge deler. I dette eksemplet vil hun bruke et *eksluderende eller* hvor man kan velge mellom to alternativer.
Vi bruker ofte ordet 'eller' i den ekskluderende betydningen i dagligtale.

Eksempel 2.5 Når vi skal løse en 2. gradslikning

$$Ax^2 + Bx + C = 0$$

kan vi svare at

$$x = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

eller

$$x = \frac{-B - \sqrt{B^2 - 4AC}}{2A}.$$

Det vi mener her at en rot i likningen vil være på den ene formen eller på den andre formen. Alle som har løst annengradslikninger i noe omfang vet at vi av og til har sammenfallende løsninger. Det betyr at vi godkjenner svaret selv om de to løsningene faktisk er like. Dette er et eksempel på *inkluderende eller*, et eller-usagn er sant selv om begge leddene er sanne.

Eksempel 2.6 Under de klassiske Olympiske Leker opererte man ikke med andre og tredjeplasser, enten vant man eller så tapte man. La oss tenke oss at en tilskuer har to sønner A og B som deltar i en av løpsøvelsene, samtidig som sønnen C til hans verste uvenn også deltar. Uvennskapet er så sterkt at det vesentligste for tilskueren er at C ikke vinner. Hans fromme ønske er derfor at A slår C eller at B slår C.
Hvis både A og B slår C, må vi kunne si at mannens ønske er oppfylt.
Vi snakker derfor om inkluderende eller i dette tilfellet

Eksempel 2.7 Hvis n og m er heltall, kan vi med rimelighet hevde at hvis n er et partall eller hvis m er et partall, så er nm et partall.
I dette tilfellet bruker vi altså inkluderende'eller'.

Den inkluderende bruken av 'eller' er så enerådende i matematikken at man må presisere at det ikke er det man mener hvis man virkelig ikke gjør det.
Noen matematikklærere på begynnende universitets- og høyskolenivå har spurt sine studenter:

$$Er\ 0 \leq 1?$$

Ofte vil mange studenter svare at vi ikke kan ha at 0 er mindre eller lik 1 ettersom 0 *faktisk er mindre enn* 1.
Denne reaksjonen er forståelig fra et utgangspunkt i dagligtalen hvor et 'eller' gjerne åpner for muligheten for hvert av leddene kan være sanne, men i matematikken praktiserer vi at vi kan si at A *eller* B er sann også når vi vet at A er sann og vi vet at B er usann.

Definisjon 2.3 Hvis A og B er utsagn, er $A \vee B$ (leses A eller B) et utsagn. Sannhetsverdien til $A \vee B$ er bestemt av følgende tabell:

A	B	$A \vee B$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\top$
$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\bot$

Vi kaller $A \vee B$ for *disjunksjonen* mellom A og B.

Eksempel 2.8 Ved å sette opp en passende sannhetsverditabell, kan vi vise at $\neg A \vee \neg B$ og $\neg(A \wedge B)$ vil ha de samme verdiene når verdiene på A og B varierer:

A	B	$\neg A$	$\neg B$	$\neg A \vee \neg B$	$A \wedge B$	$\neg(A \wedge B)$
$\top$	$\top$	$\bot$	$\bot$	$\bot$	$\top$	$\bot$
$\top$	$\bot$	$\bot$	$\top$	$\top$	$\bot$	$\top$
$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\top$	$\top$	$\top$	$\bot$	$\top$

Vi ser at tredje siste og siste søyle er like, og de svarer til de to utsagnene vi ville sammenlikne

2.3.4 Hvis- så

Det er viktig både i dagligtale og i matematikk å kunne uttrykke at en påstand er en konsekvens av en annen påstand. Vi skal se på et typisk matematisk

Eksempel 2.9 De fleste som har vært borti et minimum av matematikk vil være enige i:

$$\textit{Hvis } x > 5 \textit{ så er } x > 3$$

Hvis vi lar $x = 6$, vil både antagelsen og konklusjonen være sanne.

Hvis vi lar $x = 4$ er antagelsen usann, men konklusjonen er sann.

Hvis vi lar $x = 2$ vil både antagelse og konklusjon være usanne.

Vi oppfatter den generelle påstanden for almengyldig, selv om den har som spesialtilfeller at noe usant medfører noe sant, eller at noe usant medfører noe usant.

Eksempel 2.10 Følgende påstand vil oppfattes som usann av de fleste:

$$\textit{Hvis } x > 0 \textit{ så er } x - 1 > 0.$$

Grunnen til at dette oppfattes som feil, er at påstanden tillater moteksempler. Hvis $x = 0,5$ har vi at antagelsen er sann, men konklusjonen er usann. Alle moteksempler til et 'hvis - så'-utsagn er av denne formen, antagelsen holder, mens konklusjonen ikke holder.

I de to siste eksemplene har vi sett på hva de fleste vil kalle *implikasjoner*, og vi hevdet at den første implikasjonen må være sann fordi den ikke tillater moteksempler, mens den andre ikke er sann fordi vi fant et moteksempel. Lesere som har vært borti implikasjoner, kjenner trolig til at vi bruker symbolet $\Rightarrow$ i den forbindelse.
Når vi nå skal innføre et bindeord for *hvis - så*, skal vi ikke uttrykke en generell sammenheng mellom påstander hvor det kan inngå en variabel, som $x > 5$

og $x > 3$, men som en måte å konstruere et nytt utsagn fra to gamle på. Sannhetsverdien på det nye utsagnet skal da være avhengig av sannhetsverdiene til de gamle utsagnene, men vi skal ikke la den være avhengig av noen "global" sammenheng. Derfor vil vi bruke et annet tegn for det utsagnslogiske 'hvis - så'. Vi kommer tilbake til implikasjon senere.

Definisjon 2.4 Hvis A og B er utsagn, er $A \to B$ et utsagn. Vi leser $A \to B$ som *Hvis A så B* eller som *A medfører B*.
Sannhetsverdien til $A \to B$ er bestemt fra sannhetsverdiene til A og B ved følgende tabell:

A	B	$A \to B$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\top$

Det siste bindeordet vi skal innføre er $\leftrightarrow$. På samme måte som $\to$ er bindeordsvarianten av $\Rightarrow$, så vil $\leftrightarrow$ være bindeordsvarianten av $\Leftrightarrow$. Vi kan følge vanlig praksis, og lese $\leftrightarrow$ som 'hvis - og bare hvis', men det vil være mere dekkende å lese det som 'har samme sannhetsverdi som', eller kortere, ' er likeverdig med'.

Eksempel 2.11 La oss se på det generelle utsagnet $x > 2$ hvis og bare hvis $x > 5$.
Ut fra vanlig forståelse av matematisk tekst, er dette meningsløst, $x = 4$ er jo et glimrende moteksempel. Imidlertid er både $x > 2$ og $x > 5$ utsagn hvor sanhetsverdien varierer med verdien på x. Da vil

$$x > 2 \textit{ har samme sannhetsverdi som } x > 5$$

også kunne oppfattes som et utsagn hvor sannhetsverdien varierer med x, utsagnet er sant når $x \leq 2$ eller $x > 5$ og usant ellers.

Definisjon 2.5 Hvis A og B er utsagn, er $A \leftrightarrow B$ et utsagn. Vi leser A *er likeverdig med* B for $A \leftrightarrow B$.
Sannhetsverdien til $A \leftrightarrow B$ er gitt ved følgende sannhetsverditabell:

A	B	$A \leftrightarrow B$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\top$

Bemerkning 2.1 En grunn til at man også bør være betenkt over å bruke leseformen 'er likeverdig med' er at det er et norsk uttrykk som i ikke-teknisk sammenheng betyr det samme som 'er ekvivalent med', et uttrykk vi vil bruke i en annen betydning senere. Gjennom et par oppgaver skal vi se at både det å bruke 'hvis-så' og 'hvis og bare hvis' som utsagnslogiske bindeord kan lede til utsagn som utsagnslogisk sett er sanne uansett omstendighet, men som fra et dagligtaleperspektiv må oppfattes som det rene nonsens.

Vi kunne spart oss å definere $\leftrightarrow$, ettersom vi allid kan skrive $(A \to B) \wedge (B \to A)$ i stedet for $A \leftrightarrow A$. Dette ser vi ved å sette opp sannhetsverditabellen for $(A \to B) \wedge (B \to A)$:

A	B	$A \to B$	$B \to A$	$(A \to B) \wedge (B \to A)$
$\top$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\top$	$\bot$
$\bot$	$\top$	$\top$	$\bot$	$\bot$
$\bot$	$\bot$	$\top$	$\top$	$\top$

2.3.5 Parentessetting

Hvis vi skriver $A \wedge B \vee C$ er det ikke klart hva vi mener, vi kan mene $(A \wedge B) \vee C$ og vi kan mene $A \wedge (B \vee C)$. Hvis vi skriver opp sannhetsverditabellene for disse to utsagnene ser vi at vi ikke får den samme søylen til slutt. I dette tilfellet ser vi at vi er nødt til å bruke parenteser for å presisere hva vi mener.
Det samme argumentet kan gøres gjeldende for $\neg A \vee B$, $A \wedge B \to C$ og $A \to B \vee \neg C$, ved å sette parenteser på forskjellige måter kan vi få forskjellige utsagn med forskjellige sannhetsverditabeller.
Det vil imidlertid bli et lite problem for oss hvis vi skal bruke alle de parentesene som skal til for å gi en entydig forståelse av utsagnene, det vil bli så mange parenteser at de vil skygge for arktekturen til utsagnet. Derfor vil vi benytte oss av noen standard *konvensjoner* for utelatelse av parenteser.

Vi grupperer bindeordene i tre grupper, $\{\neg\}$, $\{\wedge, \vee\}$ og $\{\to, \leftrightarrow\}$. $\neg$ har kortest rekkevidde, den virker bare på nærmeste utsagnsvariabel eller nærmeste parentesutrykk. Det betyr at vi kan skrive $\neg A \vee B$ hvis vi mener $(\neg A) \vee B$, mens vi må skrive $\neg(A \vee B)$ dersom det er det vi mener.
$\wedge$ og $\vee$ har den samme rekkevidden, og det betyr at vi vil bruke parenteser for å markere at utsagnet til venstre eller høyre omfatter $\to$, $\leftrightarrow$ eller en $\vee$ eller $\wedge$ av motsatt type. Det betyr at vi kan skrive

$$A \wedge \neg B \to \neg A \vee B$$

og mene

$$(A \wedge \neg B) \to (\neg A \vee B).$$

Siden parentessettingen i for eksempel $A \vee B \vee C$ ikke spiller noen rolle, vil vi tillate dette som et meningsfylt utsagn. Av tilsvarende grunn vil vi skrive $A \wedge B \wedge C$ uten parenteser. Det betyr at følgende eksempel er meningsfylt:

$$(A \wedge \neg B \wedge C) \vee (\neg A \wedge B \wedge C) \vee (\neg A \wedge \neg B \wedge C) \to A \vee (B \wedge C \wedge A)$$

mens vi ikke kan fjerne noen av disse parentesene og fortstt ha et utsagn som gir bare en mening.

Vi har **ikke** at $A \to (B \to C)$ og $(A \to B) \to C$ gir den samme sistesøylen i sannhetsverditabellen. Her spiller det altså en rolle hvordan vi skriver parentesene, så de må vi alltid ha med for å gruppere delutsagnene i et utsagn hvor $\leftrightarrow$ eller $\to$ forekommer flere ganger. Dette er nesten helt sant, se oppgave 2.3.5 for korrektiv.

Oppgaver til avsnitt 2.3

Oppgave 2.3.1 En plakat i et kafé-vindu reklamerer for daglig middag som består av

* Forrett: Dagens suppe eller dagens salat.
* Hovedrett: Dagens fiskerett, dagens kjøttrett eller dagens vegetarrett.
* Dessert: Dagens kake, blandet is eller dagens ostefat.
* Kaffe eller te.

La $R_1, \ldots, R_{10}$ være utsagnsvariable for "du kan få" og så de forskjellige rettene. Hvordan vil du uttrykke denne menyen som en utsagnslogisk formel.

Oppgave 2.3.2 Nedenunder har vi gitt noen sammensatte utsagn i dagligtale. Erstatt grunnutsagnene med utsagnsvariable og finn et utsagnslogisk utsagn som svarer til det som er gitt.

1. Hvis det er midt på dagen og det er overskyet, kan jeg ikke lese boka mi utendørs.
2. Maleren kommer i morgen og hvis du ikke har ryddet rommet ditt mister du lommepengene for en måned.
3. Hvis jeg får med nok penger skal jeg kjøpe meg en jakke eller et par sko.
4. Tro kan flytte fjell, men TNT egner seg bedre.
5. Hvis du er snill og hjelper meg, får du penger til kino, men hvis du ikke er snill og hjelper meg, skal du gå til sengs uten mat.
6. Hvis solen står opp i vest, havnet jeg på flyet til Brisbane og ikke på flyet til Vancouver.

Oppgave 2.3.3 Sett opp sannhetsverditabellen til følgende samensatte utsagn:

1. $(A \wedge \neg B) \rightarrow \neg A$
2. $A \rightarrow (B \rightarrow A)$
3. $(A \rightarrow B) \wedge (B \rightarrow C) \rightarrow (A \rightarrow C)$
4. $\neg A \wedge (A \rightarrow B) \rightarrow B$
5. $((A \leftrightarrow B) \wedge C) \rightarrow (B \leftrightarrow C)$
6. $(A \rightarrow B) \wedge A \wedge \neg B$

Oppgave 2.3.4 Bestem hvilke uttrykk under som er skrevet i overenstemmelse med våre konvensjoner for parentessetting. Der hvor uttrykket ikke kan leses entydig, finn minst to måter å sette parenteser på.

1. $A \wedge B \rightarrow B \vee A$
2. $\neg A \vee B \wedge C \leftrightarrow \neg B \wedge C$
3. $(A \rightarrow \neg B \vee \neg C) \leftrightarrow \neg A \vee \neg B$
4. $A \rightarrow (\neg B \vee C) \rightarrow A$
5. $\neg B \wedge \neg C \leftrightarrow \neg(B \wedge C)$

Oppgave 2.3.5 Ved å se på sannhetsverditabellene, vis at utsagnene $A \leftrightarrow (B \leftrightarrow C)$ og $(A \leftrightarrow B) \leftrightarrow C$ uttrykker det samme.
Forklar hvorfor dette betyr at om vi har et utsagn hvor $\leftrightarrow$ er det eneste bindeordet vi har brukt, så spiller ikke parentessetting eller rekkefølgen på utsagnsvariablene noen rolle, bare hvor mange forekomster vi har av hver enkelt utsagnsvariabel.

2.4 Tautologier og kontradiksjoner

Etter å ha løst endel oppgaver i å skrive sannhetsverditabeller, vil man oppdage at noen tabeller ender med at det står $\top$ på alle plassene til høyre, mens andre får $\bot$ på alle linjene i søylen til høyre.
Når det første skjer, betyr det at utsagnet må være sant uansett hvilke sannhetsverdier vi gir til grunnutsagnene. Et slikt utsagn kaller vi en *tautologi*. Tautologier representerer utsagn som er sanne på et rent logisk grunnlag, de er i en viss forstand selvfølgeligheter.

Hvis vi derimot har et utsagn som får verdien $\bot$ uansett verdiene på utsagnsvariablene, har vi en *kontradiksjon*, eller en *selvmotsigelse*.

Eksempler 2.12 Eksempler på tautologier kan være $A \rightarrow (B \rightarrow A)$, $\neg(A \vee B) \leftrightarrow \neg A \wedge \neg B$ og $A \vee \neg A$.
Eksempler på kontradiksjoner kan være $A \wedge \neg A$, $A \wedge B \leftrightarrow \neg A \vee \neg B$.

Vi ser at dersom et utsagn er en tautologi, så er negasjonen en kontradiksjon og omvendt.

Vi vil nå bruke de store greske bokstavene Φ (uttales fi) og Ψ (uttales psi) for sammensatte utsagn.

Definisjon 2.6 Vi sier at Φ *impliserer* Ψ dersom $\Phi \rightarrow \Psi$ er en tautologi. Vi skriver $\Phi \Rightarrow \Psi$ i dette tilfellet.
Vi sier at Φ og Ψ er *ekvivalente* dersom $\Phi \leftrightarrow \Psi$ er en tautologi. Vi skriver $\Phi \Leftrightarrow \Psi$ i dette tilfellet.

Det at $\Phi \leftrightarrow \Psi$ er en tautologi, betyr at Φ og Ψ vil få den samme sannhetsverdien uansett hvilke sannhetsverdier vi gir til utsagnsvariablene. En annen måte å si dette på er at sannhetsverditabellene for Φ og Ψ får den samme søylen lengst til høyre.

En viktig egenskap ved ekvivalens er at hvis vi har et komplekst uttrykk, og

så erstatter vi en del av uttrykket med noe ekvivalent, så er resultatet også ekvivalent. Vi skal se på et par eksempler, men ikke bevise dette foreløpig. Etter at vi har gått igjennom kapitlet om generell induksjon, vil vi ha det matematiske verktøyet som skal til for et fullstendig bevis.

Eksempler 2.13 *Vi skal se på to eksempler på omskrivninger av utsagn til ekvivalente former:*

1. *La Φ være utsagnet $A \vee (\neg\neg B \to A)$*
   Siden $\neg\neg B \Leftrightarrow B$ kan vi erstatte $\neg\neg B$ med B, og vi får det ekvivalente utsagnet $\Phi_1 = A \vee (B \to A)$.

2. *Vi fortsetter eksemplet. Vi har at $B \to A \Leftrightarrow \neg B \vee A$, så Φ_1 er ekvivalent til $\Phi_2 = A \vee (\neg B \vee A)$ som igjen er ekvivalent til $A \vee \neg B$.*

En forklaring på hvorfor dette er riktig kan være som følger: Når vi bygger opp sannhetsverditabellen til et uttrykk, starter vi med utsagnsvarablene og søyle for søyle finner vi sannhetsverditabellen for deluttrykk. Hvis vi erstatter et deluttrykk med et annet som gir den samme søylen, vil fortsettelsen i de to tilfellene se helt like ut, ettersom det ikke er hvilke uttrykk vi har men hvilke søyler de gir i sannhetsverditabellen vi bruker.
Dette er ikke et fullstendig bevis, men forhåpentligvis en relativt overbevisende forklaring.

Oppgaver til avsnitt 2.4

Oppgave 2.4.1 Vi skal gi tre påstander uttrykt på norsk.
Vis at hvis vi omformer disse påstandene til utsagnslogiske formler, blir resultatene tautologier.

a) Hvis jeg vrir om tenningsnøkkelen og trår på gasspedalen, vil bilen starte. Følgelig, hvis jeg vrir om tenningsnøkkelen vil bilen starte eller hvis jeg trår på gasspedalen vil bilen starte.

b) Hvis jeg er i Paris er jeg i Frankrike og hvis jeg er i London er jeg i England. Følgelig, hvis jeg er i Paris er jeg i England eller hvis jeg er i London er jeg i Frankrike.

c) Jeg klipper plenen hvis og bare hvis jeg får penger eller jeg klipper plenen hvis og bare hvis jeg ikke får penger.

2.5 Regneregler

I forrige avsnitt så vi at vi kan erstatte et delutsagn i et sammensatt utsagn med et annet som er ekvivalent med delutsagnet. Dette kan sammenliknes med at vi kan regne på deluttrykk i algebraiske uttrykk, vi kan erstatte et deluttrykk med et annet som gir de samme verdiene.
I dette avsnittet skal vi se på noen utvalgte ekvivalenser, og vi vil oppfatte disse

som regneregler. Vi vil regne på uttrykk hvor sannhetsverdiene $\top$ og $\bot$ også kan forekomme. En måte å vise at Φ er en tautologi på er å vise at $\Phi \Leftrightarrow \top$, og tilsvarende vil Ψ være en kontradiksjon dersom vi kan regne oss frem til $\bot$ fra Ψ.

Når vi formulerer regnereglene, vil vi bruke utsagnsvariable, men det er meningen at hvilke som helst utsagn kan stå på plassene til utsagnsvariablene. Vi må bare passe på å erstatte flere forekomster av den samme variabelen med det samme utsagnet. Vi gir regnereglene i grupper:

Eliminering/introduksjon av $\top$ og $\bot$.

1. $\top \vee A \Leftrightarrow \top$ og $\bot \vee A \Leftrightarrow A$.
2. $\top \wedge A \Leftrightarrow A$ og $\bot \wedge A \Leftrightarrow \bot$
3. $\top \rightarrow A \Leftrightarrow A$, $A \rightarrow \top \Leftrightarrow \top$, $\bot \rightarrow A \Leftrightarrow \top$ og $A \rightarrow \bot \Leftrightarrow \neg A$.
4. $\top \leftrightarrow A \Leftrightarrow A$ og $\bot \leftrightarrow A \Leftrightarrow \neg A$.
5. $\neg\top \Leftrightarrow \bot$ og $\neg\bot \Leftrightarrow \top$.
6. $A \vee \neg A \Leftrightarrow \top$ og $A \wedge \neg A \Leftrightarrow \bot$.

Vi tar ikke sikte på å gi et minimalistisk sett av regler, så flere av disse reglene kan erstattes av gjentatt bruk av andre regler. Vår neste liste dreier seg om

Eliminering av $\rightarrow$ og $\leftrightarrow$

1. $A \leftrightarrow B \Leftrightarrow (A \rightarrow B) \wedge (B \rightarrow A)$.
2. $A \rightarrow B \Leftrightarrow \neg A \vee B$.

Hvorvidt det lønner seg å bruke disse reglene til eliminering av $\leftrightarrow$ og $\rightarrow$, avhenger av formålet med regningen. Hvis formålet er å sjekke om vi har en tautologi eller ikke, lønner det seg gjerne ikke å fjerne dem, noe vi vil se i avsnitt 2.6. I avsnitt 2.7 skal vi se på en anvendelse av utsagnslogikk hvor formlene ikke bare må være fri for forekomster av $\leftrightarrow$ og $\rightarrow$, men hvor negasjonstegnet bare må brukes i tilknytning til utsagnsvariable. Vårt neste sett av regler dreier seg om

Flytting av $\neg$ innover

1. $\neg\neg A \leftrightarrow A$.
2. $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$.
3. $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$.

De to siste ekvivalensene er kjent som *DeMorgans lover*. Vi hadde også DeMorgans lover i mengdelæren. Det er en sammenheng, som vi vil utdype i kapittel 3

Vi skal gi to sett til med regler. Det første settet dreier seg om oppløsning av

parenteser og omvendingen, det å sette noe utenfor en parentes. Disse to prosessene er motsatte av hverandre og gis i sammenfattende regler. Det andre settet gir regler for at parentessetting, rekkefølge og antall ganger vi gjentar et utsagn er uvesentlig i rene $\wedge$-utsagn og i rene $\vee$-utsagn.

Distributive lover

1. $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$
2. $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$

Resten

1. Kommutative lover: $A \wedge B \Leftrightarrow B \wedge A$ og $A \vee B \Leftrightarrow B \vee A$.
2. Assosiative lover: $A \wedge (B \wedge C) \Leftrightarrow (A \wedge B) \wedge C$ og $A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$.
3. Sammentrekning: $A \wedge A \Leftrightarrow A$ og $A \vee A \Leftrightarrow A$.

Eksempel 2.14 Som et eksempel på hvordan vi kan bruke disse reglene, skal vi regne på uttrykket

$$(A \to B) \wedge (B \to C) \to (A \to C)$$

og vise ved regning at dette er en tautologi. For å forenkle regnestykket litt, kommer vi ikke til å ta med flere parenteser enn det som kreves for at uttrykkene blir meningsfylte eller for å illustrere hvilke regler vi har brukt.
Med untak av rest-reglene, vil vi ikke bruke mer enn en regel av gangen. Leseren utfordres til å finne ut av hvilken regel vi har brukt.

1. $(A \to B) \wedge (B \to C) \to (A \to C)$
2. $\neg((A \to B) \wedge (B \to C)) \vee (A \to C)$
3. $\neg((A \to B) \wedge (B \to C)) \vee \neg A \vee C$
4. $\neg((\neg A \vee B) \wedge (B \to C)) \vee \neg A \vee C$
5. $\neg((\neg A \vee B) \wedge (\neg B \vee C)) \vee \neg A \vee C$
6. $\neg(\neg A \vee B) \vee \neg(\neg B \vee C) \vee \neg A \vee C$
7. $(\neg\neg A \wedge \neg B) \vee \neg(\neg B \vee C) \vee \neg A \vee C$
8. $(\neg\neg A \wedge \neg B) \vee (\neg\neg B \wedge \neg C) \vee \neg A \vee C$
9. $(A \wedge \neg B) \vee (\neg\neg B \wedge \neg C) \vee \neg A \vee C$
10. $(A \wedge \neg B) \vee (B \wedge \neg C) \vee \neg A \vee C$
11. $(A \vee (B \wedge \neg C) \vee \neg A \vee C) \wedge (\neg B \vee (B \wedge \neg C) \vee \neg A \vee C)$
12. $(\top \vee (B \wedge \neg C) \vee C) \wedge (\neg B \vee (B \wedge \neg C) \vee \neg A \vee C)$

13. $\top \wedge (\neg B \vee (B \wedge \neg C) \vee \neg A \vee C)$
14. $\neg B \vee (B \wedge \neg C) \vee \neg A \vee C$
15. $(\neg B \vee B \vee \neg A \vee C) \wedge (\neg B \vee \neg C \vee \neg A \vee C)$
16. $\top \vee \neg A \vee C) \wedge (\neg B \vee \neg C \vee \neg A \vee C)$
17. $\top \wedge (\neg B \vee \neg C \vee \neg A \vee C)$
18. $(\neg B \vee \neg C \vee \neg A \vee C)$
19. $\top \vee \neg B \vee \neg A$
20. $\top$

Dette eksemplet er ikke enestående, enhver tautologi kan regnes om til $\top$. Dette er noe vi skal komme tilbake til i avsnittet om normalformer under Utfordringer i dette kapitlet. Som vi så, ble regnestykket langt, og det er heller ikke noe enestående. For utsagn med få utsagnsvariable, er det overkommelig på en systematisk måte å regne seg frem til om det er en tautologi eller ikke. I noen teknologiske anvendelser av utsagnslogikk kan det imidlertid forekomme mange utsagnsvariable (typisk en variabel for hvert relevant grunnutsagn om objekter registrert i en database, eksempelvis bilregisteret), og da vil det være en uoverkommelig oppgave selv for raske datamaskiner å teste om et gitt utsagn er en tautologi eller ikke. Vi skal ikke komme nærmere inn på dette her.

Oppgaver til avsnitt 2.5

Oppgave 2.5.1 Bruk regnereglene over til å vise at hvert av utsagnene under er tautologier.

1. $A \to (B \to A)$
2. $(A \to B) \to ((B \to C) \to (A \to C))$
3. $(A \to B) \wedge (\neg A \to C) \to (B \vee C)$
4. $A \leftrightarrow \neg\neg A$
5. $(A \vee B) \wedge (A \to C) \wedge (B \to C) \to C$

Oppgave 2.5.2 Bruk regnereglene til å vise at følgende utsagn er kontradiksjoner:

1. $(A \to \neg A) \wedge (\neg A \to A)$
2. $(A \to C \wedge B) \wedge (C \to B) \wedge (B \to \neg A) \wedge A$
3. $\neg(A \vee B \vee C) \wedge \neg(\neg A \vee \neg B \vee \neg C)$

2.6 Teknikker

Regnereglene vi så på i forrige avsnitt er komplette i den forstand at vi kan regne oss fra et hvilket som helst sammensatt utsagn til et ekvivalent et ved hjelp av disse reglene. Hvis vi blir bedt om å bestemme om et gitt utsagn er en tautologi eller ikke, finnes det imidlertid andre teknikker enn streng bruk av disse reglene som i de fleste tilfellene er mere effektive. Vi skal se på et par slike teknikker, uten å gå dypt inn i materien. Vi demonstrerer teknikkene på eksemplet

$$(A \to B) \to ((B \to C) \to (A \to C)).$$

En viktig grunn til at bruken av regnereglene ukritisk gir lange utregninger er at når vi bruker de distributive lovene til å løse opp parenteser, forlenger vi samtidig uttrykket. Er vi uheldige får vi veldig lange uttrykk å regne på.

2.6.1 Teknikk 1

Vi setter inn de to sannhetsverdiene for en bokstav av gangen og forenkler uttrykket uten bruk av distributive lover før vi (om nødvendig) setter inn verdier for flere variable.
Utfordringen er å holde orden på hvilke variable vi har satt inn hvilke verdier for, og det gjøres best ved en forstandig layout med nytt innrykk hver gang vi setter inn en verdi for en ny variabel.
Det kan også være greit å avgjøre om vi får en tautologi ved å sette inn $\top$ for A (eller den utsagnsvariabelen vi tror det hjelper best å sette inn verdier for) før vi setter inn $\bot$ for A
La oss se på eksemplet.

$\top$ for A:
$(\top \to B) \to ((B \to C) \to (\top \to C))$
$B \to ((B \to C) \to C)$
$\top$ for B:
$\top \to ((\top \to C) \to C)$
$C \to C$
som er en tautologi.
$\bot$ for B:
$\bot \to ((\bot \to C) \to C)$
som er ekvivalent med $\top$.

$\bot$ for A:
$(\bot \to B) \to ((B \to C) \to (\bot \to C))$
$\top \to ((B \to C) \to \top)$
$\top \to \top$
som er ekvivalent med $\top$.

Erfaringsmessig vil man kunne gjennomføre store forenklinger i det minste i noen av tilfellene. En annen fordel er at metoden kan brukes til å finne sannhetsverditabellen til kompliserte utsagn uten å sette opp hele sannhetsverditabellen. Hvis

man etter å ha satt inn sannhetsverdier for noen få av variablene reduserer uttrykket til $\top$ eller $\bot$, kan man fylle inn alle de tilsvarende linjene i tabellen. I eksemplet over så vi ved rask regning at alle linjene som starter med $\bot$ for A har $\top$ lengst til høyre. På tilsvarende måte kan vi ofte fylle ut store deler av sannhetsverditabellen om gangen ved å sette inn sannhetsverdier for en eller to utsagnsvariable, og så forenkle uttrykket.

For lesere med litt kjennkap til programmering kan vi definere følgende kvasiprosedyre **P** hvor $\mathbf{P}(Q)$ agjør om Q er en tautologi eller ikke når Q er et sammensatt utsagn hvor $\top$ og $\bot$ kan forekomme. Vi kan beskrive **P** ved følgende rekke instruksjoner:

1. Finn en enklere form av Q ved enten å eliminere alle forekomster av $\top$ og $\bot$ eller ved å forenkle Q til $\top$ eller til $\bot$.
2. Hvis Q forenkles til $\top$, konkluder med at Q er en tautologi.
3. Hvis Q forenkles til $\bot$, konkluder med at Q ikke er en tautologi.
4. Hvis ingen av delene gjelder, finn en utsagnsvariabel A i Q og la $Q[\top]$ være det vi får når vi erstatter A med $\top$ overalt i Q.
5. Bruk **P** på $Q[\top]$. Hvis vi da konkluderer med at $Q[\top]$ ikke er en tautologi, konkluderer vi med at Q ikke er en tautologi, og avslutter prosedyren. Hvis vi konkluderer med at $Q[\top]$ er en tautologi, går vi videre til neste punkt.
6. La $Q[\bot]$ være det vi får når vi setter inn $\bot$ for A i Q.
7. Bruk **P** på $Q[\bot]$ og la **P** trekke samme konklusjon om Q som **P** trekker om $Q[\bot]$.

2.6.2 Teknikk 2

Denne teknikken egner seg best hvis utsagnet inneholder mange $\rightarrow$, og hvis målet er å undersøke om utsagnet er en tautologi eller ikke. I anvendelsenes verden er dette ofte tilfelle. Vi ønsker å undersøke om en konklusjon kan trekkes fra en gitt mengde av forutsetninger. Forutsetningene er ofte på formen "hvis ditt så datt" eller de er fakta som kan erstattes med utsagnsvariable, og vi lurer da på om visse enkle påstander er konsekvenser.

Strategien går ut på å undersøke om det er mulig å gjøre utsagnet usant, og finner vi ut at det er umulig, har vi vist at utsagnet er en tautologi. La oss gyve løs på eksemplet vårt:

Er

$$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$$

en tautologi?
For at utsagnet skal bli usant må forutsetningen $(A \rightarrow B)$ være sann, mens

konklusjonen $(B \to C) \to (A \to C)$ må være usann.
Dette siste medfører at $B \to C$ er sann, mens $(A \to C)$ er usann, noe som igjen betyr at $A = \top$ og $C = \bot$.
Men setter vi $A = \top$ inn i $(A \to B) = \top$ får vi $B = \top$, og setter vi inn denne verdien for B inn i $(B \to C) = \top$ får vi $C = \top$.
Det følger at for at utsagnet skal bli usant må vi både ha at $C = \top$ og $C = \bot$, hvilket selvfølgelig er umulig. Derfor er det opprinnelige utsagnet en tautologi.

En fordel med denne metoden er at hvis vi ikke kommer frem til en selvmotsigelse, har vi ofte bestemt verdier på utsagnsvariableme som gjør utsagnet usant. De som løser Oppgave 2.6.2 bør prøve å få til dette der hvor det er aktuelt.

Oppgaver til avsnitt 2.6

Oppgave 2.6.1 Bruk Teknikk 1 til å finne sannhetsverditabellene til følgende utsagn:

1. $(A \vee B) \wedge (C \vee D) \to (A \vee D)$
2. $\neg(A \to B) \wedge \neg(B \to C) \to (A \wedge B \wedge C)$
   (Hint: det kan lønne seg å sette inn sannhetsverdiene for B først.)
3. $(A \to B) \to ((A \to C) \to (B \to C))$

Oppgave 2.6.2 Bruk Teknikk 2 til å bestemme om følgende utsagn er tautologier:

1. $(A \to B) \to ((A \to C) \to (B \to C))$
2. $(A \to B) \wedge (C \to D) \to ((B \to C) \to (A \to D))$
3. $(A \wedge (A \to B) \to B) \to B$

2.7 Strømkretser

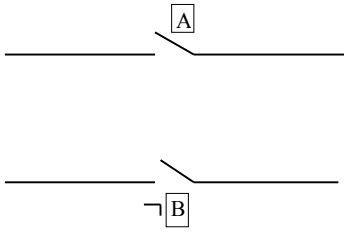
En tradisjonell anvendelse av utsagnslogikk er konstruksjon av strømkretser. La oss se på et eksempel.
I en trappeoppgang er det tre lysbrytere, en for hver etasje. Hver bryter er i en opp eller ned posisjon, og vi ønsker at man skal kunne tenne lyset om det er avslått eller slå det av hvis det er på ved å endre posisjonen til hvilken som helst av bryterne.
La oss kalle bryterne for A, B og C, og la oss ta utgangspunkt i at lyset i trappeoppgangen skal være på dersom alle bryterne peker oppover. Da må også lyset være på dersom nøyaktig én av bryterne peker oppover, ellers må det være avslått.
La oss nå betrakte bryterne som utsagnsvariable, og at det at en bryter peker

oppover representerer at variabelen får verdien $\top$. Da får vi et sammensatt utsagn som uttrykker at det skal være lys i trappeoppgangen:

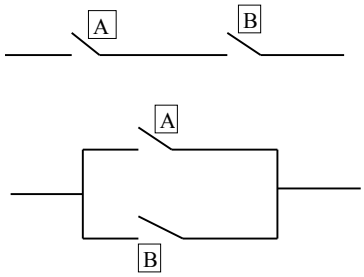
$$(A \wedge B \wedge C) \vee (A \wedge \neg B \wedge \neg C) \vee (\neg A \wedge B \wedge \neg C) \vee (\neg A \wedge \neg B \wedge C).$$

Hvordan skal dette hjelpe oss til å trekke ledningene? Vi skal se at hvis vi antar at hver bryter kan forbinde flere sett ledninger, noen når bryteren peker oppover og andre når bryteren peker nedover, så er det mulig å trekke ledninger gjennom disse tre bryterne slik at trappeoppgangen fungerer som den skal.

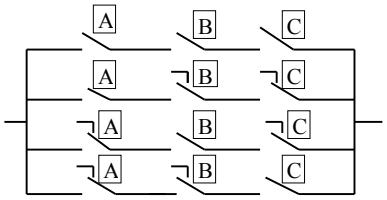
Vi skal tegne opp ledningsnett. Vi markerer ledningene med streker, og tegner inn bryterne i ledningsnettet. Vi skriver en utsagnsvariabel ved bryteren hvis vi vil at det skal gå strøm gjennom ledningen når bryteren peker oppover, og vi skriver negasjonen av en utsagnsvariabel hvis vi vil at det skal gå strøm gjennom ledningen når bryteren peker nedover:



Ved å seriekoble to brytere kan vi uttrykke konjunksjoner, og ved å parallellkoble bryterne kan vi uttrykke disjunksjon:



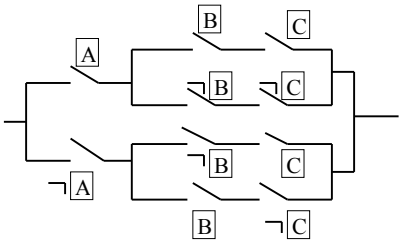
Det er selvfølgelig mulig å seriekoble eller parallellkoble strømkretser også. Det betyr at vi faktisk kan lage oss en strømkrets som svarer til behovet i trappeoppgangen:



Dette ser jo ganske komplisert ut. Vi må trekke fire ledninger i parallell gjennom alle bryterne. Heldigvis kan vi nå bruke det vi har lært om utsagnslogikk, og forenkle uttrykket noe. Hvis vi trekker sammen to og to ledd får vi

$$(A \wedge ((B \wedge C) \vee (\neg B \wedge \neg C))) \vee (\neg A \wedge ((B \wedge \neg C) \vee (\neg B \wedge C))).$$

Dette gir oss en litt enklere strømkrets, ved at vi bare trenger å trekke to ledninger gjennom bryter A:



Det er ikke alle sammensatte utsagn som kan omgjøres til strømkretser. Vi er avhengige av at negasjonstegnet bare står rett foran en utsagnsvariabel, og vi kan bare simulere $\wedge$ og $\vee$. På den annen side kan alle utsagn skrives om til et ekvivalent utsagn som oppfyller disse kravene. Derfor isolerer vi denne familien av sammensatte utsagn gjennom en definisjon:

Definisjon 2.7 a) En *literal* er en utsagnsvariabel A eller negasjonen $\neg A$ av en utsagnsvariabel.

b) Et utsagn er på *svak normalform* hvis det er bygget opp fra literaler ved bruk av $\wedge$ og $\vee$, men ingen andre bindeord.

Hvis et utsagn er på svak normalform, kan vi lage en strømkrets slik at vi for alle fordelinger av sannhetsverdier på utsagnsvariablene har at det går strøm

gjennom kretsen hvis og bare hvis sannhetsverdiene bestemt av bryternes stilling (opp betyr $\top$ og ned betyr $\bot$) gjør at utsagnet blir sant.

Oppgaver til avsnitt 2.7

Oppgave 2.7.1 En gruppe på tre personer skal avgi stemme for en sak ved å la en bryter peke oppover.
Konstruer en strømkrets slik at det går strøm gjennom kretsen hvis og bare hvis et flertall stemmer for saken.
Gjør kretsen så enkel du kan.

Oppgave 2.7.2 I en toetasjers skolebygning er det en bryter i hver etasje, som på dagtid skal virke slik at man kan skru av/på lyset fra begge etasjene ved å endre bryterens stilling. Samtidig har vaktmesteren en tredje bryter slik at hun kan slå av lyset helt om natten.
Kall de to etasjebryterne A og B og vaktmesterens bryter C og konstruer en strømkrets gjennom disse bryterne slik at de virker som de skal.

Oppgave 2.7.3 Vi har gitt noen sammensatte utsagn som ikke er på svak normalform. Finn ekvivalente utsagn som er på svak normalform ved å erstatte $\rightarrow$ med $\neg$ og $\vee$, ved å bruke DeMorgans lover til å skyve negasjonstegnet innover og ved å fjerne forekomster av $\neg\neg$ som måtte oppstå. Det er ikke meningen at du skal falle for fristelsen til å forenkle uttrykket for eksempel ved å erstatte $A \vee \neg A$ med $\top$ eller trekke sammen eller løse opp parenteser.

1. $\neg(A \rightarrow B)$
2. $A \rightarrow \neg(B \vee \neg A)$
3. $\neg(\neg B \vee (A \wedge \neg(B \vee \neg A)))$
4. $(A \rightarrow B) \rightarrow (B \rightarrow C)$
5. $A \rightarrow (B \rightarrow A)$

2.8 Utfordringer

2.8.1 Adekvate sett av konnektiver

Vi har sett hvordan vi kan bruke sannhetsverditabeller til å finne hvordan sannhetsverdien til et sammensatt utsagn vil variere når sannhetsverdiene til grunnutsagnene varierer. Nå skal vi stille problemet på hodet. Anta at vi har gitt en tabell som gir en sannhetsverdi til alle mulige verdier av grunnutsagnene, kan vi da finne et sammensatt utsagn som har en sannhetsverditabell som slutter med den gitte søylen?
Svaret er positivt, og vi skal vise det ved å se på et eksempel som er tilstrekkelig generelt til at vi ser sammenhengen. La oss starte med følgende tilfeldige utvalgte tabell:

A	B	C	Φ
$\top$	$\top$	$\top$	$\bot$
$\top$	$\top$	$\bot$	$\bot$
$\top$	$\bot$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\bot$
$\bot$	$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$

Problemet er om vi kan finne et sammensatt utsagn Φ som gir oss denne tabellen?
Vi ser at tabellen skal være sann i rad 3 og i rad 6. For å være sann i rad 3, må Φ være sann når A er sann, B er usann og C er sann. $A \wedge \neg B \wedge C$ har denne egenskapen, og alle disjunksjoner hvor $A \wedge \neg B \wedge C$ inngår som et av leddene har denne egenskapen.
Videre skal Φ være sann når A er usann, B er sann og C er usann. $\neg A \wedge B \wedge \neg C$ har denne egenskapen.
Vi ser at for hver linje i sannhetstabellen kan vi finne et utsagn ved å bruke $\neg$ og $\wedge$ som er sann nøyaktig for den linjen. Hvis vi nå tar disjunksjonen av alle disse utsagnene for de linjene hvor vi vil at den formelen vi søker etter skal være sann, oppnår vi det vi vil. I vårt eksempel blir formelen

$$(A \wedge \neg B \wedge C) \vee (\neg A \wedge B \wedge \neg C).$$

Denne formelen vil da ha tabellen i eksemplet som sin sannhetsverditabell.

Definisjon 2.8 En mengde konnektiver eller bindeord kalles *adekvat* hvis alle sannhetsverditabeller kan realiseres ved hjelp av bindeord fra mengden.

Vi har vist at enhver sannhetsverditabell med minst en $\top$ i kan uttrykkes ved bruk av $\neg$, $\wedge$ og $\vee$. Det er ingen heksekunst å lage en kontradiksjon, så tabeller med bare $\bot$ i siste søyle kan realiseres. Vi har derfor at $\{\neg, \vee, \wedge\}$ er en adekvat mengde.
Videre vet vi at vi kan kvitte oss med alle forekomster av for eksempel $\vee$ ved å erstatte $A \vee B$ med $\neg(\neg A \wedge \neg B)$ systematisk. Derfor er også $\{\neg, \wedge\}$ en adekvat mengde, og $\{\neg, \vee\}$ vil være en adekvat mengde ut fra et tilsvarende argument.

Det finnes to "bindeord" som er adekvate hver for seg, det ene svarer til *ikke både A og B* og det andre til *ikke A eller ikke B*. Leseren utfordres til å definere disse to bindeordene ved sannhetsverditabeller, og åuttrykke negasjon, konjunksjon og disjunksjn i hvert av disse to bindeordene.

2.8.2 Normalformer

I avsnittet om elektriske kretser, så vi på utsagn hvor vi ikke har forekomster av $\rightarrow$ eller $\leftrightarrow$, og hvor negasjonstegnet bare forekommer i tilknytning til utsagnsvariable. Et slikt utsagn sa vi at var på *svak normalform*. Vi minner om at en *literal* er en utsagnsvariabel A eller negasjonen $\neg A$ av en utsagnsvariabel. Da

vi argumenterte for at enhver sannhetsverditabell kan realiseres av et utsagn, startet vi med å lage en konjunksjon av literaler for hver linje i tabellen. Hvis vi for eksempel har fire utsagnsvariable A, B, C og D, og ser på linje 7, som ser ut som $\top \mid \bot \mid \bot \mid \top$ vet vi at utsagnet $A \wedge \neg B \wedge \neg C \wedge D$ vil gi oss en tabell som gir verdien $\top$ i linje 7, men $\bot$ i alle de andre linjene.
For å konstruere et utsagn som svarer til en gitt tabell, ser vi derfor på alle de linjene hvor vi ønsker $\top$, og tar disjunksjonen av de utsagnene som er sanne på hver enkelt av de aktuelle linjene.

Definisjon 2.9 Et utsagn Q er på *disjunktiv normalform* hvis Q er på formen

$$P_1 \vee \cdots \vee P_n$$

hvor hver P_i er en konjunksjon av literaler

Eksempler på formler på disjunktiv normalform kan være

$$(A \vee \neg B \vee \neg C) \wedge (\neg A \vee B \vee C) \wedge (\neg A \vee negB \vee C)$$

og

$$(A \wedge B) \vee (B \wedge \neg C) \vee (C \wedge \neg D).$$

I noen lærebøker kreves det at alle leddene (disjunktene) må inneholde en forekomst av hver utsagnsvariabel. Vi vil ikke la det være et krav her.
En måte å finne en disjunktiv normalform til et utsagn på er å regne ut sannhetsverditabellen, og så bruke den konstruksjonen vi har repetert over. En annen og mer algebraisk metode er å bruke de distributive lovene til å skyve $\wedge$ innover og forbi alle forekomster av $\vee$. Da er det greit å starte med et utsagn på svak normalform. La oss vise hva som skjer ved et eksempel:

Eksempel 2.15 Vi skal finne den disjunktive normalformen til

$$(A \vee B) \wedge ((A \wedge C) \vee (\neg C \wedge \neg A)).$$

Vi gjennomfører følgende regning, hvor vi hver gang enten trekker sammen og fjerner ledd som blir trivielle eller bruker at $\wedge$ kan distribueres over $\vee$. Vi gjør enkelte steder flere operasjoner samtidig:

1. $(A \vee B) \wedge ((A \wedge C) \vee (\neg C \wedge \neg A))$
2. $(A \wedge ((A \wedge C) \vee (\neg C \wedge \neg A))) \vee (B \wedge ((A \wedge C) \vee (\neg C \wedge \neg A))$
3. $(A \wedge A \wedge C) \vee (A \wedge \neg C \wedge \neg A) \vee (B \wedge A \wedge C) \vee (B \wedge \neg C \wedge \neg A)$
4. $(A \wedge C) \vee \bot \vee (A \wedge B \wedge C) \vee (\neg A \wedge B \wedge \neg C)$
5. $(A \wedge C) \vee (\neg A \wedge B \wedge \neg C)$

I den siste overgangen foretok vi oss to ting. Vi fjernet $\bot$ fra disjunksjonen fordi den ikke bidrar til noe. Vi fjernert også det tredje leddet. Det er fordi det første leddet er en konsekvens av det tredje leddet, så å ta med ledd 3 representerer bare en overpresisering av hva som trengs.

Vi finner den disjunktive normalformen ved systematisk å bruke at $\wedge$ kan distribueres over $\vee$. Hvis vi gjør det motsatte, systematisk bruker at $\vee$ kan distribueres over $\wedge$, kan vi regne oss frem til den *konjunktive normalformen*:

Definisjon 2.10 La Q være et utsagn. Vi sier at Q er på *konjunktiv normalform* hvis Q er på formen $Q = P_1 \wedge \cdots \wedge P_n$, hvor hver P_i er en disjunksjon av literaler.

Ethvert utsagn kan bringes over på konjunktiv normalform. Vi skal ikke stresse det noe mer.

2.8.3 Kompleksitet

Hvis leseren har løst oppgave 2.7.3 som foreskrevet vil hun/han innse at man ganske rutinemessig kan skrive om et utsagn til svak normalform. Vi kan sågar formulere en algoritme for hvordan det skal gjøres. For enkelthets skyld antar vi at vi ikke har brukt $\leftrightarrow$:

1. Erstatt etter tur alle forekomster av $\rightarrow$ med $\neg$ og $\vee$
2. Let fra venstre mot høyre etter en forekomst av $\neg$ som ikke står rett foran en utsagnsvariabel.
   Hvis du ikke finner noen, smil fornøyd, mens hvis du finner en, merk de første du finner og fortsett til 3.
3. Hvis den markerte $\neg$ står rett foran en annen $\neg$, fjern begge og gå tilbake til 2.
   Hvis den markerte $\neg$ står rett foran et $\wedge$- eller et $\vee$-utsagn, bruk DeMorgans lov på stedet og fortsett med 2.

Spørsmålet er hvor mange ganger vi kan risikere å gå igjennom denne prosessen før vi har funnet en svak normalform.
Antall operasjoner under punkt 1. er bestemt av antall forekomster av $\rightarrow$.
Hver gang vi bruker DeMorgans lov øker vi antall forekomster av $\neg$ med 1. Vi vil likevel ikke innføre flere slike forekomster enn antall symboler i det opprinnelige utsagnet. Dette kan vi se på mere formelt under avsnittet om generell induksjon. En konsekvens er at to ganger antall symboler i det opprinnelige uttrykket er en øvre grense for hvor mange ganger vi må gjennomføre 2. og 3.
Kompleksiteten til en oppgave er et mål på hvor mange regneskritt vi må utføre for å løse oppgaven. Dette målet er selvfølgelig avhengig av omfanget av oppgaven selv, i dette tilfellet hvor mange symboler vi har i utsagnet som skal omformes.
Det tar tid å gjennomføre hvert av punktene, og det varierer fra matematisk modell til matematisk modell hva som omfattes av et regneskritt. Hvis vi bruker en av standardmodellene vil det ta under $10(n+1)^3$ regneskritt å skrive om et utsagn med n symboler til svak normalform.

Hvis vi imidlertid ønsker å skrive utsagnet om til disjunktiv eller konjunktiv normalform (se avsnitt 2.8.2) vil vi øke lengden av utsagnet betraktelig hver gang vi løser opp en parentes. Vi kan finne et tall k slik at vi kan skrive om et

utsagn med n symboler til disjunktiv normalform med mindre enn $k \cdot 2^n$ antall skritt, men ingen har klart å forbedre dette vesentlig. Det er heller ingen som har klart å lage en algoritme som avgjør om et utsagn er en tautologi eller ikke på vesentlig kortere tid.
Det forholder seg imidlertid slik at det muligens finnes k og r slik at vi kan avgjøre om et utsagn med n symboler er en tautologi eller ikke med færre enn $k \cdot (n+1)^r$ regneskritt, men ingen vet om det er tilfelle eller ikke.
Lesere som løser dette problemet vil samtidig ha løst et av de såkalte **milleniumsproblemene** i matematikken. Dette er syv problemer som det ble utlovet dusør på, $ 1.000.000 for hvert av dem, i forbindelse med årtusenskiftet.

2.9 Blandede oppgaver

Oppgave 2.9.1 a) La Φ være formelen $A \vee (B \to (B \wedge A))$.
Sett opp hele sannhetsverditabellen til Φ.

b) Bruk sannhetsverditabellen til å finne et enklere utsagn Ψ som er ekvivalent til Φ.

c) Bruk regnereglene til å regne deg frem til Ψ fra Φ.

Oppgave 2.9.2 Undersøk om følgende utsagn er en tautologi. Du kan bruke den metoden du selv finner mest egnet:

$$(A \wedge \neg B \wedge C) \vee (\neg A \wedge B \wedge C) \to ((A \wedge B) \to C).$$

Oppgave 2.9.3 a) Undersøk om følgende sammensatte utsagn er tautologier:

$$P_1 = (A \vee B \to C) \to (A \to C) \vee (B \to C)$$
$$P_2 = (\neg A \to (\neg B \to A))$$
$$P_3 = (C \to A \wedge B) \wedge (B \to \neg C) \wedge (A \to C)$$

b) Vis at $P_3 \Rightarrow P_2$.

c) Vis at $P_2 \not\Rightarrow P_3$.

d) Er det mulig for to sammensatte utsagn P og Q å være ekvivalente selv om de ikke har de samme utsagnsvariablene?

Oppgave 2.9.4 a) Vis at $P_1 \Rightarrow P_2$ hvor

$$P_1 = (C \to B) \wedge (\neg C \to A) \to (A \to B)$$

og

$$P_2 = (A \wedge \neg B) \to (B \to D).$$

b) Finn et sammensatt utsagn P_3 med bare A og B som utsagnsvariable slik at $P_1 \Rightarrow P_3$ og $P_3 \Rightarrow P_2$.

c) [u] Del b) er et spesialtilfelle av et generelt teorem som heter *Interpolasjonsteoremet for utsagnslogikk*, som sier.

La Q_1 og Q_2 være to sammensatte utsagn slik at $Q_1 \Rightarrow Q_2$.
Da finnes det et utsagn Q_3 som bare inneholder utsagnsvariable som finnes i både Q_1 og Q_2 og som er slik at

$$Q_1 \Rightarrow Q_3 \Rightarrow Q_2.$$

Vis interpolasjonsteoremet for utsagnslogikk.
[Hint: Bruk at alle sannhetsverditabeller kan realiseres av et utsagn. La X være mengden av utsagnsvariable som finnes både i Q_1 og i Q_2. La Q_3 være sann for en sannhetsfordeling på X hvis og bare hvis den kan utvides til en fordeling av sannhetsverdier på alle utsagnsvariable i Q_1 som gjør Q_1 sann.]

Oppgave 2.9.5 La P være det sammensatte utsagnet

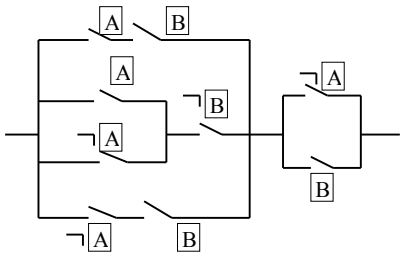
$$((A \wedge B) \rightarrow C) \wedge \neg(C \rightarrow A \wedge B).$$

a) Skriv P på svak normalform.

b) Forenkle den svake normalformen så mye som mulig.

c) Konstruer en strømkrets som svarer til den forenklede normalformen.

Oppgave 2.9.6 En elektrikerlærling oppdager at mesteren har montert en strømkrets som ser slik ut:



a) Finn det sammensatte utsagnet som svarer til denne kretsen.

b) Finn et enklere ekvivalent utsagn, og vis hvordan lærlingen kan forbedre mesterens montasje.

Oppgave 2.9.7 (u) La Φ være et utsagn i bindeordene $\neg$, $\vee$, $\wedge$, $\rightarrow$ og $\leftrightarrow$.
La *det kritiske tallet* $K(\Phi)$ være antall forekomster av utsagnsvariable i Φ.

a) Vis at $K(\Phi)$ er én større enn samlet antall forekomster av $\vee$, $\wedge$, $\rightarrow$ og $\leftrightarrow$.

b) Forklar hvorfor vi må ha to negasjonstegn rett etter hverandre hvis negasjonene utgjør mer enn halvparten av symbolene i Φ.

c) Drøft hva dette betyr for kompleksiteten av algoritmen for å finne en svak normalform til et sammensatt utsagn.

Kapittel 3

Boolesk algebra

3.1 Hva har mengdealgebra og utsagnslogikk felles?

I de to foregående kapitlene har vi innført mengdelære og utsagnslogikk. Den observante leser kan ikke ha ungått å se at det finnes likhetstrekk mellom deler av mengdelæren og utsagnslogikken. I mengdelæren opererer vi med snitt, union og komplement, og regnereglene for disse minner til forveksling om de regnereglene vi fant for konjunksjon, disjunksjon og negasjon. Vi har til og med gått så langt at vi har brukt den samme betegnelsen, *DeMorgans lover*, på regneregler for både mengdelære og utsagnslogikk:

$$(A \cap B)^c = A^c \cup B^c$$

$$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$$

$$(A \cup B)^c = A^c \cap B^c$$

$$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B.$$

Hvis vi ser på definisjonene av $\wedge$, $\vee$, $\cap$, $\cup$, ^{c} og $\neg$ ser vi sammenhengen direkte. Hvis A er en mengde i et univers U, og $a \in U$, kan vi betrakte $a \in A$ som et grunnutsagn. Da har vi følgende ekvivalenser:

$$a \in A \cap B \Leftrightarrow a \in A \wedge a \in B.$$

$$a \in A \cup B \Leftrightarrow a \in A \vee a \in B.$$

$$a \in A^c \Leftrightarrow \neg(a \in A).$$

Eksempler 3.1 a) Vi ser på tre delmengder A, B og C av $\mathbb{N}$ definert ved

A er mengden av oddetall.

B er mengden av tall på formen $2^n + 1$.

C er mengden av primtall.

Vi er utfordret til å beskrive mengden $(B \cap C) \setminus A$.
Det første vi gjør er å skrive om mengdeuttrykket til

$$(B \cap C \cap A^c)$$

og deretter bruke sammenhengen mellom mengdeoperasjonene og logikk til å slå fast at for alle tall n gjelder det at

$$n \in (B \cap C) \setminus A \Leftrightarrow n \in B \wedge n \in C \wedge n \notin A.$$

Dette betyr at n skal være et primtall, n skal være på formen $2^m + 1$ og n skal ikke være et oddetall. Tallet 2 er det eneste som oppfyller alle disse betingelsene, så vår analyse viser at den mengden vi definerte faktisk er $\{2\}$.

b) En ondsinnet kollega har bedt oss om å løse fjerdegradslikningen

$$(x^2 - 1)^2 + (x^2 + x - 2)^2 + (x - 2)^2 = 0.$$

Vi ser at siden venstre side er en sum av kvadrater, og siden ingen har sagt noe om at vi skal regne med komplekse tall her, så svarer likningen til konjunksjonen

$$x^2 - 1 = 0 \wedge x^2 + x - 2 = 0 \wedge x - 2 = 0,$$

som igjen svarer til

$$(x = 1 \vee x = -1) \wedge (x = -2 \vee x = 1) \wedge (x = 2).$$

Omsatt til mengdenotasjon ser vi at løsningsmengden er

$$(\{1\} \cup \{-1\}) \cap (\{1\} \cup \{-2\}) \cap \{2\},$$

eller skrevet litt tydeligere

$$\{1, -1\} \cap \{1, -2\} \cap \{2\} = \emptyset.$$

Vi ser at likningen til kollegaen vår ikke har noen løsning.

c) Et styre i et borettslag vil undersøke hvor omfattende kjæledyrsholdet er blant medlemmene, og sender ut et spørreskjema hvor medlemmene skal krysse av på om de har hund, om de har katt og på om de har en annen form for kjæledyr.
Styrets nestleder har fått det for seg at en hund som bor bare sammen med familien og eventuelt andre hunder ikke representerer noe problem, men hvis den deler tilværelsen med andre kjæledyr, vil den bli bråkete og til plage for andre beboere. Hvis vi lar H være mengden av beboere med hund, K mengden av beboere med katt og A mengden av beboere med andre kjæledyr, er mengden av beboere som ikke har et kjæledyrshold etter nestlederens smak $H \cap (K \cup A)$, og beboer x er suspekt hvis $x \in H \wedge (x \in K \vee x \in A)$.

Definisjon 3.1 La $A_1, \ldots, A_n$ være delmengder av et univers U, og la ϕ være en mengde definert fra $A_1, \ldots, A_n$ ved hjelp av operasjonene $\wedge$, $\vee$ og ^{c} på $A_1, \ldots, A_n$. La x være en variabel.
Vi finner den *utsagnslogiske versjonen* $UV(\phi)$ av ϕ ved å erstatte A_i med $x \in A_i$ i definisjonen av ϕ, alle forekomster av $\wedge$ med $\cap$, alle forekomster av $\vee$ med $\cup$ og $\neg()$ med $()^c$ overalt.

Vi har da

Teorem 3.1 *a) Hvis $UV(\phi)$ er en tautologi, vil $\phi = U$.*

b) Hvis $UV(\phi)$ er en kontradiksjon vil $\phi = \emptyset$.

c) Hvis ϕ og ψ er slik at $UV(\phi) \Rightarrow UV(\psi)$, vil $\phi \subseteq \psi$.

Bemerkning 3.1 Den kritiske leseren vil ha oppdaget at vi definerte $UV(\phi)$ ut fra hvordan mengden ϕ er definert, ikke bare ut fra mengden selv. Det vil ikke skape noen vansker her, og vi skal ikke problematisere denne unøyaktigheten nærmere. I Oppgave 6.4.1 vil vi komme tilbake til hvordan dette kan gjøres mer stringent.

Bevis for Teorem 3.1
Vi beviser a) i detalj, og overlater bevisene for b) og c) til leseren som en oppgave.
Anta at $UV(\phi)$ er en tautologi. La $a \in U$ være vilkårlig. a definerer en tilordning av sannhetsverdier til grunnutsagnene $x \in A_i$ ved at $x \in A_i$ settes til $\top$ om $a \in A_i$ og $x \in A_i$ settes til $\bot$ ellers.
Denne tilordningen vil gjøre $UV(\phi)$ sann, siden $UV(\phi)$ er en tautologi. Men sammenhengen mellom mengdelæren og utsagnslogikken forteller oss at $a \in \phi$ hvis og bare hvis $UV(\phi)$ blir sann under denne tilordningen, helt uavhengig av at ϕ er en tautologi, så det følger at $a \in \phi$.
Siden a var vilkårlig valgt, må vi ha at $U = \phi$.

Bemerkning 3.2 Lesere som sitter med en følelse at noe mangler i beviset over, gjør trolig rett i det. Påstanden om at $a \in \phi$ hvis og bare hvis $UV(\phi)$ er sann under denne tilordningen av sannhetsverdier, kan begrunnes videre ved hjelp av et induksjonsbevis. Dette gir vi som oppgaver i kapitlene 6 og 7.

Vi kan ikke trekke noen generelle konklusjoner den andre veien, ettersom vi ikke har lagt noen føringer på manglende sammenhenger mellom de forskjellige A_i-ene. Hvis vi for eksempel lar U være mengden $\mathbb{Z}$ av hele tall, A_1 mengden av tall som kan deles på 3, A_2 mengden av tall som gir 1 til rest når vi deler på 3 og A_3 mengden av tall som gir 2 til rest når vi deler på 3, så har vi at $U = A_1 \cup A_2 \cup A_3$ uten at $x \in A_1 \vee x \in A_2 \vee x \in A_3$ er en tautogi.
Vi skal se nærmere på hva som kreves for at omvendingen skal gjelde i avsnittet om utfordringer.

Oppgaver til avsnitt 3.1

Oppgave 3.1.1 For hver definisjon ϕ av en mengde, finn $UV(\phi)$:

a) $\phi = (A \cup (B \cap A^c)) \setminus (A^c \cap B))$.

b) $\phi = ((A \cup B)^c \cap (B \cup C)^c)^c \cap A$.

c) $\phi = (C^c \cap A^c)^c \cup (B \cap C)$.

Oppgave 3.1.2 Vi har sett på sammenhengen mellom noen logiske bindeord og mengdeteoretiske operasjoner som snitt, union og komplement.
Ved hjelp av denne sammenhengen kan disse bindeordene illustreres via Venn-diagrammer.

a) Finn Venn-diagrammer som illustrerer bindeordene $\rightarrow$ og $\leftrightarrow$

b) Bruk Venn-diagrammene du fant i a) til å forklare sammenhengen mellom $\leftrightarrow$ (hvis og bare hvis) og $\triangle$ (symmetrisk differens).

Oppgave 3.1.3 Bevis punktene b) og c) i Teorem 3.1

3.2 Booleske algebraer

I avsnittet over så vi på sammenhenger mellom mengdealgebra og utsagnslogikk. I begge tilfeller hadde vi en algebra med et minste element ($\emptyset$ eller $\bot$) og et største element (U eller $\top$). Vi opererte også med to operatorer som svarte til 'og' og 'eller', og en operator som svarte til negasjon eller komplement.

Vi kan oppfatte utsagnslogikk i sin snevreste forstand som en algebra over mengden $\mathbb{B} = \{\top, \bot\}$ hvor $\neg : \mathbb{B} \rightarrow \mathbb{B}$, $\wedge : \mathbb{B}^2 \rightarrow \mathbb{B}$ og $\vee : \mathbb{B}^2 \rightarrow \mathbb{B}$ er funksjoner definert via sannhetstabeller. (Vi skal diskutere hva vi mener med funksjoner i detalj i kapittel 4. Leseren kan la teksten over være en appetittvekker for Kapittel 4 om hun/han har problemer med formuleringene her.)
De som har lest litt om databehandling bør ha fått med seg at data i sin enkleste form består av 0'er og 1'ere, av positive og negative ladninger eller på annen måte av to distinkte objekter. Det hadde ikke vært noe i veien for å erstatte $\top$ med 1 og $\bot$ med 0 og utvikle utsagnslogikken på det grunnlaget. Da kan vi til og med definere $\neg$, $\wedge$ og $\vee$ på en alternativ måte:

$$\neg x = 1 - x$$

$$x \wedge y = min\{x, y\}$$

$$x \vee y = max\{x, y\}$$

Arkitekturen i moderne datamaskiner er slik at sekvenser av 0'er og 1'ere av bestemte lengder blir behandlet som informasjonsenheter. Den gang man så på tre og tre slike tall, regnet databehandlere i 8-tallsystemet, eller *oktalt*. Brukes fire og fire siffer som enheten svarer det til heksadesimal regning, et finere navn

på regning i 16-tallsystemet. I moderne dataspill samler man minst seks og seks siffre som enheter. Dette svarer da til utsagnslogikk med seks utsagnsvariable. Vi har tidligere vist (i et avsnitt om utfordringer) at alle sannhetsverdifunksjoner kan defineres via et sammensatt utsagn i utsagnslogikken. Konsekvensen er at alle operasjoner vi måtte ønske å gjøre på datasekvenser av lengde 6 kan defineres utsagnslogisk. Det er trukket to konsekvenser av dette, hvorav vi skal forfølge den andre nå og den første i neste avsnitt:

1. Arkitekturen i datamaskiner er bygget på innsikt fra utsagnslogikk.
2. Det som er felles for mengdelæren, utsagnslogikken, regning med datainformasjon og deler av sannsynlighetsteorien er samlet i et begrep, Booleske algebraer.

Definisjon 3.2 En *Boolesk algebra* består av en mengde $\mathbf{B}$, to forskjellige utvalgte elementer $\mathbf{0} \in \mathbf{B}$ og $\mathbf{1} \in \mathbf{B}$ og funksjoner

$$\sim\colon \mathbf{B} \to \mathbf{B}$$

$$\sqcap : \mathbf{B}^2 \to \mathbf{B}$$

$$\sqcup : \mathbf{B}^2 \to \mathbf{B}$$

slik at følgende holder for alle x, y og z i $\mathbf{B}$:

1. (Identiteter)
   $x \sqcap \mathbf{1} = x$
   $x \sqcup \mathbf{0} = x$
2. (Kommutativitet)
   $x \sqcap y = y \sqcap x$
   $x \sqcup y = y \sqcup x$
3. (Distributivitet)
   $x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z)$
   $x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z)$
4. (Komplement)
   $x \sqcap \sim x = \mathbf{0}$
   $x \sqcup \sim x = \mathbf{1}$
5. (Assosiativitet)
   $x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z$
   $x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z$
   (Egentlig kan assosiativitet vises ut fra de andre reglene, men vi tar det med fordi beviset er langt og lite illustrerende.)

Bemerkning 3.3 Vi har valgt å bruke symbolene $\sqcap$ og $\sqcup$ ettersom de likner på $\cap$ hhv. $\wedge$ og på $\cup$ hhv. $\vee$.
Det er ikke uvanlig å bruke $\cdot$ og $+$ i stedet. Det viktigste er at disse tegnene står for operasjoner som oppfyller de reglene vi har satt opp.

Eksempler 3.2 Vi skal se på noen eksempler på Booleske algebraer.

1. La $\mathbf{B} = \mathcal{P}(U)$, det vil si potensmengden til en mengde U, hvor vi tolker $\mathbf{1}$ som U, $\mathbf{0}$ som $\emptyset$, $\sim$ som komplementdannelse, $\sqcap$ som snitt og $\sqcup$ som union.
   Da er alle egenskapene til en Boolesk algebra oppfylt.

2. La $\mathbf{B}$ være mengden av alle funksjoner fra $\mathbb{N}$ til $\{0, 1\}$. La

   - $\mathbf{1}(n) = 1$ for alle n og $\mathbf{0}(n) = 0$ for alle n.
   - $(f \sqcap g)(n) = min\{f(n), g(n)\}$ for alle n
   - $(f \sqcup g)(n) = max\{f(n), g(n)\}$ for alle n

   Da er alle egenskapene til en Boolesk algebra oppfylt.
   Slike funksjoner kan oppfattes som følger $\{f(n)\}_{n\in\mathbb{N}}$ av data, og omtales gjerne som *datastrømmer*. Datamaskiner som blir satt til å utføre i prinsippet evigvarige oppgaver vil håndtere datastrømmer i den forstand at de til stadighet vil tilføres nye data. Et eksempel vil være en maskin som overvåker nivået i kraftmagasiner og regulerer utslippet av vann. Den vil basere seg på en strøm av data vedrørende volum i magasinet og vannføring i tilstøtende vassdrag, snedybde i fjellet o.l.

3. La ϕ være et sammensatt utsagn i utsagnsvariablene A_1, A_2 og A_3.
   La $[\phi]$ være mengden av slike sammensatte utsagn som er ekvivalente til ϕ.
   Hvis $\phi \Leftrightarrow \psi$ vil $[\phi] = [\psi]$, ellers vil $[\phi] \cap [\psi] = \emptyset$.
   Vi lar $\mathbf{B}$ være mengden av slike mengder $[\phi]$.
   La $\mathbf{1}$ være mengden av tautologier og $\mathbf{0}$ være mengden av kontradiksjoner.
   La $\sim [\phi] = [\neg\phi]$, $[\phi] \sqcap [\psi] = [\phi \wedge \psi]$ og $[\phi] \sqcup [\psi] = [\phi \vee \psi]$.
   (Leseren bør reflektere over hvorfor dette er lovlige definisjoner.) Da er alle reglene for Booleske algebraer oppfylt.

Et av poengene med å studere Booleske algebraer og andre algebraiske teorier er at det er mange ytterlige egenskaper som vi kan regne oss frem til, og som da er felles for alle Booleske algebraer, alternativt de andre algebraene. Eksempelvis la vi stor vekt på DeMorgans lover både i mengdealgebra og i utsagnslogikk, men vi har ikke tatt med DeMorgans lover her. Grunnen er at vi faktisk kan bevise dem.

I resten av dette avsnittet skal vi bevise en rekke egenskaper ved Booleske algebraer. Det er ikke meningen at man skal kunne gjennomføre slike utregninger selv til eksamen, og det får være opp til hver enkelt hvor langt og detaljert man vil lese dette avsnittet. Vi kaller påstandene våre lemmaer. Der et lemma består av to deler, viser vi den ene, når den andre delen vises på helt tilsvarende måte, se Oppgave 3.2.1.
Alle likheter i bevisene under baserer seg på en av reglene eller på en likhet som er vist før. Leseren utfordres til å finne ut av hvilke regler eller likheter som er brukt.

Lemma 3.1 *(Idempotens)*
$x \sqcap x = x$
$x \sqcup x = x$

Bevis
$x = x \sqcap \mathbf{1} = x \sqcap (x \sqcup \sim x) = (x \sqcap x) \sqcup (x \sqcap \sim x) = (x \sqcap x) \sqcup \mathbf{0} = x \sqcap x.$

Lemma 3.2 *(Anihilering)*
$x \sqcap \mathbf{0} = \mathbf{0}$
$x \sqcup \mathbf{1} = \mathbf{1}$

Bevis
$x \sqcap \mathbf{0} = (x \sqcap \mathbf{0}) \sqcup \mathbf{0} = (x \sqcap \mathbf{0}) \sqcup (x \sqcap \sim x) = x \sqcap (\mathbf{0} \sqcup \sim x) = x \sqcap \sim x = \mathbf{0}.$

Lemma 3.3 *(Dobbelnegasjon)*
$\sim\sim x = x.$

Bevis
$x = x \sqcap \mathbf{1} =$
$x \sqcap (\sim x \sqcup \sim\sim x) =$
$(x \sqcap \sim x) \sqcup (x \sqcap \sim\sim x) =$
$\mathbf{0} \sqcup (x \sqcap \sim\sim x) =$
$(\sim x \sqcap \sim\sim x) \sqcup (x \sqcap \sim\sim x) =$
$(x \sqcup \sim x) \sqcap \sim\sim x =$
$\mathbf{1} \sqcap \sim\sim x =$
$\sim\sim x.$

Lemma 3.4 *(Topp og bunn)*
$\sim \mathbf{0} = \mathbf{1}$ *og* $\sim \mathbf{1} = \mathbf{0}.$

Bevis
$\sim \mathbf{0} = \sim \mathbf{0} \sqcup \mathbf{0} = \mathbf{1}.$

Lemma 3.5 *(Likhetskriterium)*
Hvis $x \sqcup y = x \sqcap y$ *vil* $x = y.$

Bevis
Vi viser at $x = x \sqcap y$. Tilsvarende regning gir at $y = y \sqcap x = x \sqcap y$, så det holder.

$x = x \sqcup \mathbf{0} =$
$x \sqcup (y \sqcap \sim y) =$
$(x \sqcup y) \sqcap (x \sqcup \sim y) =$
$(x \sqcap y) \sqcap (x \sqcup \sim y) =$
$(x \sqcap y \sqcap x) \sqcup (x \sqcap y \sqcap \sim y) =$
$(x \sqcap y) \sqcup \mathbf{0} = x \sqcap y.$

Vi brukte den assosiative loven for $\sqcap$ her. Lesere som har lyst til å utlede den assosiative loven fra de andre reglene må passe på ikke å bruke dette lemmaet eller noe lemma som bygger på det.

Lemma 3.6 *(Komplementærkriterium)*
Hvis $x \sqcup y = \mathbf{1}$ *og* $x \sqcap y = \mathbf{0}$*, vil* $x = \sim y$.

Bevis
I følge foregående lemma er det tilstrekkelig å vise at

$$x \sqcup \sim y = x \sqcap \sim y$$

ut fra antagelsen. Det gjøres ved følgende utregninger:

$$x \sqcup \sim y = (x \sqcup \sim y) \sqcap (x \sqcup y) = x \sqcup (\sim y \sqcap y) = x$$

$$x \sqcap \sim y = (x \sqcap \sim y) \sqcup (x \sqcap y) = x \sqcap (\sim y \sqcup y) = x.$$

Lemma 3.7 *(DeMorgans lover)*

$$\sim (x \sqcap y) = \sim x \sqcup \sim y$$

$$\sim (x \sqcup y) = \sim x \sqcap \sim y$$

Bevis
Vi viser bare den første delen, og bruker Lemma 3.6. Da er følgende utregninger tilstrekkelige:

$$(\sim x \sqcup \sim y) \sqcup (x \sqcap y) = (\sim x \sqcup \sim y \sqcup x) \sqcap (\sim x \sqcup \sim y \sqcup y) = \mathbf{1} \sqcap \mathbf{1} = \mathbf{1}$$

$$(\sim x \sqcup \sim y) \sqcap (x \sqcap y) = (\sim x \sqcap x \sqcap y) \sqcup (\sim y \sqcap x \sqcap y) = \mathbf{0} \sqcup \mathbf{0} = \mathbf{0}$$

Oppgaver til avsnitt 3.2

Oppgave 3.2.1 Fullfør bevisene av Lemmaene 3.1 - 3.4.

Oppgave 3.2.2 Forklar hvorfor eksempel 2 i 3.2 egentlig bare er et spesialtilfelle av eksempel 1.
Hint: Erstatt f med $\{n \mid f(n) = 1\}$

Oppgave 3.2.3 Vis følgende likninger for Booleske algebraer uten å bruke assosiativitet.

1. $x \sqcap (x \sqcup y) = x$.
2. $x \sqcup (x \sqcap y) = x$.
3. $\sim x \sqcup (x \sqcup y) = \mathbf{1}$.
4. $\sim x \sqcap (x \sqcap y) = \mathbf{0}$.

Hint: Bruk 1. og 2. i 3. og 4.

Oppgave 3.2.4 [u] Vis de assosiative reglene fra de andre reglene.
Hint: Du vil få bruk for 3. og 4. fra foregående oppgave. Denne oppgaven er ikke lett likevel.

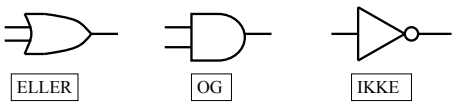
Oppgave 3.2.5 La $\mathbf{B}$ bestå av alle delmengder X av $\mathbb{N}$ som enten er endelig eller slik at komplementet er endelig.

a) Vis at hvis X og Y er i $\mathbf{B}$, så er $X \cap Y \in \mathbf{B}$, $X \cup Y \in \mathbf{B}$ og $X^c \in \mathbf{B}$.

b) Vis at $\mathbf{B}$ på en naturlig måte kan betraktes som en Boolesk algebra.

3.3 Digitale kretser

Vi har sett hvordan problemer i tilknytning til konstruksjon av strømkretser kan løses ved hjelp av utsagnslogikk. Der ble utsagnene representert ved brytere, og posisjonen på bryteren skal angi sannhetsverdien til grunnutsagnet.
En annen måte å tolke strømkretser som utsagn på er å la hver ledningsbit representere et utsagn. Dette utsagnet er da sant om det går strøm gjennom ledningen, mens det er usant om det ikke går strøm. Arkitekturen til datamaskiner baserer seg faktisk på denne grunnidéen, og byggestenene er det vi vil kalle *digitale kretser*. En digital krets er et ledningsnett hvor vi kan ha forgreninger og hvor ledningene kan bindes sammen via det vi vil kalle *logiske porter*. Vi skal se på tre logiske porter her:

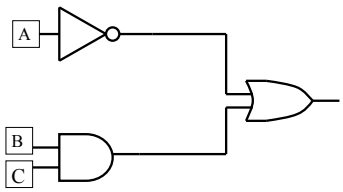


I *Eller-porten* kommer det strøm ut til høyre hvis det kommer strøm inn fra minst en av ledningene til venstre, i *Og-porten* kommer det strøm ut til høyre om det kommer strøm inn i begge ledningene til venstre, og i *Ikke-porten* kommer det strøm ut til høyre hvis og bare hvis det ikke kommer strøm inn fra venstre. I reelle digitale nettverk forekommer det også andre porter med sine faste skrivemåter, porter for *hverken eller* og for *ikke både og*. Disse tegner vi ved å tegne en liten ring (det er den som markerer negasjon) rett bak tegnet for eller-porten eller og-porten. Grunnen til at vi ikke bare bruker den lille ringen ved negasjon er at vi vil markere strømretningen i hver port når vi tegner en digital krets.

Eksempel 3.3 Vi ønsker å lage en digital krets som representerer utsagnet $A \to (B \wedge C)$. Ettersom vi ikke har laget noen port for $\to$, må vi først skrive om utsagnet slik at bare bindeordene $\vee$, $\wedge$ og $\neg$ forekommer. Dette blir

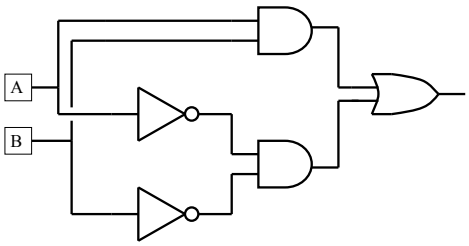
$$\neg A \vee (B \wedge C).$$

Kretsen vil bli seende slik ut:



Vi ser at vi må nøste opp utsagnet innenfra og utover når vi tegner den digitale kretsen fra venstre mot høyre.

Eksempel 3.4 Da vi tegnet logiske kretser tegnet vi inn en bryter for hver forekomst av utsagnsvariablene. Her ønsker vi bare å tegne en ledning inn fra venstre for hver variabel. Hvis vi har flere forekomster av samme variabel, må vi derfor la dette komme til uttrykk ved at ledningene deler seg. Ikke rent sjelden må vi også la ledninger krysse hverandre (dette er ikke nødvendig i teorien, men i praksis). Det vil vi markere med et lite opphold i den vertikale ledningen.
La oss se på $(A \wedge B) \vee (\neg A \wedge \neg B)$, som gir opphav til følgende integrerte krets:



Ved å se på den ekvivalente formen

$$(A \wedge B) \vee \neg(A \vee B)$$

kan vi finne en enklere krets som har de samme funksjonene som kretsen over, se Oppgave 3.3.1

Eksempel 3.5 Vårt neste og siste eksempel skal delvis illustrere hvordan vi kan bruke innsikten fra utsagnslogikk til å konstruere digitale kretser som er tilpasset arbeidsoppgaver og delvis illustrere hvordan vi kan innarbeide former for hukommelse av data i integrerte kretser.
Vi har tidligere nevnt at en datastrøm kan sees på som en følge $\{a_n\}_{n\in\mathbb{N}}$ hvor hver a_n er 0 eller 1. Det er da av interesse å lage kretser som produserer en

ny datastrømm fra en eller flere andre strømmer. De eksemplene vi har sett på til nå kan sees på på denne måten, ved at de gir ut Booleske kombinasjoner av dataene fra strømmene A, B og i det ene tilfellet C. For å holde oss med en teknologiaktig terminologi vil vi kalle tallene i en datastrøm for *pulser*.
Det er ikke unaturlig at vi vil at den strømmen vi får ut skal være avhengig av flere pulser i de strømmene vi får inn enn de sist ankomne. Da er vi nødt til å innarbeide en form for hukommelse. En måte å gjøre dette på er å ha en eller flere hukommelsesceller som også innvirker på sluttresultatet og som oppdateres hver gang vi sender et sett pulser gjennom kretsen. En annen, og teoretisk sett likeverdig måte er å la kretsen sende pulser til seg selv, men da med ett eller flere skritts forsinkelse.
Vi innfører symbolet

som markerer at strømmen som kommer inn fra høyre fortsetter i samme form mot venstre, men da forsinket med en enhet (tidsavstanden mellom pulsene i datastrømmen).
Dette er en symbolbruk forfatteren har innført for anledningen. Klokken skal markere at ting tar tid.
La oss se på et enkelt eksempel, hvor vi får inn en datastrøm A og ønsker at den strømmen B vi får ut skal bestå av en 1 om antall 1'ere lest fra A er et partall og av en 0 ellers. For å få til dette må vi lagre informasjon om pariteten til antall 1'ere lest så langt.

La A være en utsagnsvariabel som representere datastrømmen vi får inn ved at $\top$ står for 1 og $\bot$ står for 0, B tilsvarende representere datastrømmen vi vil sende ut, N er en utsagnsvariabel hvor verdien $\bot$ svarer til at antall 1'ere lest før pulsen kommer er et partall og M en utsagnsvariabel som på samme måte tilsvarer antall 1'ere etter at pulsen er lest.
Vi kan da sette opp en sannhetsverditabell for hvordan B og M avhenger av A og N:

A	N	B	M
$\top$	$\top$	$\top$	$\bot$
$\top$	$\bot$	$\bot$	$\top$
$\bot$	$\top$	$\bot$	$\top$
$\bot$	$\bot$	$\top$	$\bot$

Tabellen kan vi komme frem til på følgende måte:

- $A = \top$ og $B = \top$ betyr at vi leser 1 og har på forhånd lest et oddetall 1'ere.
  Det betyr at vi samlet har lest et partall 1'ere, noe som markeres ved at $B = \top$ og $M = \bot$.

- $A = \top$ og $B = \bot$ betyr at vi leser 1 og har på forhånd lest et partall 1'ere. Det betyr at vi samlet har lest et oddetall 1'ere, noe som markeres ved at $B = \bot$ og $M = \top$.
- $A = \bot$ betyr at vi ikke leser noen 1'er. Da må M settes lik N og B settes lik det motsatte av N.

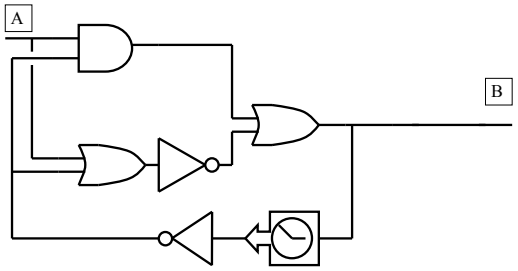
Fra tabellen ser vi at $M = \neg B$. Hvis vi bruker konstruksjonen av en formel fra sannhetsverditabellen, ser vi at vi kan sette

$$B = (A \wedge N) \vee (\neg A \wedge \neg N).$$

I dette tilfellet er det lurt å bruke DeMorgans lov motsatt av det vi pleier. Da får vi

$$B = (A \wedge N) \vee \neg(A \vee \mathbb{N}).$$

Tar vi med at $M = \neg B$, ser vi at vi kan bruke følgende digitale krets til å løse problemet:



Vi ser at vi har speilvendt ikke-porten nederst. Dette er gjort for å angi strømretningen.

Bemerkning 3.4 I eksemplet over kan det synes litt kunstig at vi lot $N = \bot$ svare til at antall 1'ere lest var et partall. Hadde vi latt det svare til at antall 1'ere lest er et oddetall, kunne vi forenklet kretsen ved å la $M = B$, det vil si at vi både lar verdien på B gå ut som utgangspulsen og som minnepulsen.
Grunnen til at vi gjørde det slik vi gjorde er at vi lar $\bot$ stå for at det ikke er strøm i ledningen, og når vi starter prosessen har vi lest et partall, nemlig ingen, 1'ere og det vil heller ikke være strøm i ledningen som kommer fra klokken.

Oppgaver til avsnitt 3.3

Oppgave 3.3.1 Finn en enklere digital krets som gjør den samme nytten som kretsen i Eksempel 3.4.

Oppgave 3.3.2 Finn en digital krets som kan brukes som logisk port for $\rightarrow$.
Foreslå en figur som kan brukes til å markere en slik port.
Drøft hvorfor en slik port kan være upraktisk i tegning av kretser.

Oppgave 3.3.3 Finn digitale kretser som representerer utsagnene under. Ved å bruke passende omskrivninger av utsagnene, prøv å få kretsene så enkle som mulig.

a) $A \vee \neg(B \wedge A)$

b) $(\neg A \wedge \neg B) \rightarrow (C \rightarrow B)$

c) $(A \wedge B \wedge C) \vee (\neg A \wedge \neg B \wedge \neg C)$.

Oppgave 3.3.4 Konstruer en digital krets med forsinket tilbakeføring slik at hvis den mates med en datastrøm så skriver den ut 1'ere så lenge inngangsstrømmen bare har 1'ere, men skriver ut 0'er etter at den har mottatt første 0.

Oppgave 3.3.5 [u] Hvis vi har to datastrømmer kan vi oppfatte dem som uendelige binære tall (tall i 2-tall systemet) ved at første tall svarer til siste siffer, neste tall til nest siste siffer og så videre. Da kan vi legge strømmene sammen som binære tall og få en ny datastrøm.
Finn en integrert krets som gjør dette for oss.
Hint: Vi trenger en forbindelse med forsinkelse som representerer 'mente' eller 'minne'. Du må da finne to sammensatte utsagn en for utgangssymbolet og en for den nye menten, i tre variable, to for strømmene og en for den gammel menten. Advarsel: Kretsen blir forholdsvis kompleks.

3.4 Utfordringer

3.4.1 Valuasjoner

Vi har sett på hvordan vi kan oversette mengdealgebra til utsagnslogikk. Vi skal nå ta for oss sammensatte utsagn og se hvordan vi kan tolke hvert utsagn som en mengde slik at vi får sammenhengen den "andre" veien.

Definisjon 3.3 La P_1, P_2, ... være en liste av utsagnsvariable. Vi skal se på sammensatte utsagn hvor alle utsagnsvariablene er fra denne listen.
En *valuasjon* er en tilordning v av en sannhetsverdi $v(P_i)$ i $\{\top, \bot\}$ til hver utsagnsvariabel P_i.

I dette avsnittet lar vi V være mengden av alle valuasjoner, og betrakter V som et univers.
Vi lar hver utsagnsvariabel P_i bestemme en mengde $A(P_i) \subset V$ ved

$$A(P_i) = \{v \in V \mid v(P_i) = \top\}.$$

Enhver valuasjon vil bestemme sannhetsverdien også til sammensatte utsagn Φ, og vi kan da utvide definisjonen av A til $A(\Phi)$ som mengden av alle valuasjoner som gjør Φ sann.
Vi ser at en valuasjon gjør $\Phi \wedge \Psi$ sann hvis den gjør både Φ og Ψ sann. Det betyr at

$$A(\Phi \wedge \Psi) = A(\Phi) \cap A(\Psi).$$

Av tilsvarende grunner ser vi at

$$A(\Phi \vee \Psi) = A(\Phi) \cup A(\Psi)$$

og at

$$A(\neg\Phi) = (A(\Phi))^c.$$

Etter beviset for Teorem 3.1 bemerket vi at omvendingen ikke gjelder, og vi ga et eksempel.
Vi minner om at ϕ er en mengde definert fra grunnmengder $A_1, \ldots, A_n$ ved hjelp av snitt, union og komplement, og at $UV(\phi)$ er det tilsvarende sammensatte utsagnet hvor vi erstatter A_i med $x \in A_i$ osv.
Det er mulig at $\phi = U$ selv om ikke $UV(\phi)$ er en tautologi, fordi det kan finnes valuasjoner som gjør $UV(\phi)$ usann, men som ikke kan defineres ut fra noen a i U. Det er jo tilstrekkelig at alle valuasjoner definert fra vilkårlige $a \in U$ gjør $UV(\phi)$ sanne for at beviset skal holde.

Definisjon 3.4 La $A_1, \ldots, A_n$ være delmengder av et univers U.
Vi sier at $A_1, \ldots A_n$ er en *generisk* hvis alle valuasjoner på utsagnsmengden $\{x \in A_1, \ldots, x \in A_n\}$ kan defineres ut fra en $a \in U$.

Hvis $A_1, \ldots, A_n$ er generisk gjelder omvendingen av Teorem 3.1.
Ved å oversette en utsagnsvariabel til mengden av valuasjoner som gjør utsagnsvariabelen sann, sørger vi nettopp for å tolke utsagnsvariablene som en generisk familie av mengder. Det er gjennom slike generiske familier at vi ser den fullstendige sammenhengen mellom mengdekonstruksjoner og utsagnslogikk.

3.4.2 Representasjon av Booleske algebraer

Vi har stort sett brukt mengdealgebraer og utsagnslogikk som eksempler på Booleske algebraer, og det er ingen tilfeldighet. Det er nemlig mulig å vise at alle Booleske algebraer kan betraktes som mengdealgebraer. Det vil føre for langt å gå i detalj på hva vi mener her, og alt for langt å gi noen form for bevis. Vi skal se på spesialtilfellet hvor den Booleske algebraen er endelig, og forklare fenomenet da.

Gjennom resten av dette avsnittet skal vi la $\mathbf{B}$ med $\mathbf{0}$, $\mathbf{1}$, $\sqcap$, $\sqcup$ og $\sim$ være en Boolesk algebra hvor $\mathbf{B}$ er endelig.
Vi definerer en ordning på $\mathbf{B}$ ved

$$x \sqsubseteq y \Leftrightarrow x \sqcap y = x.$$

Det er lett å se at vi også har at

$$x \sqsubseteq y \Leftrightarrow x \sqcup y = y$$

og at

$$x \sqsubseteq y \Leftrightarrow \sim y \sqsubseteq \sim x.$$

Ved å bruke assosiativitet av $\sqcap$ ser vi at

$$x \sqsubseteq t \wedge y \sqsubseteq x \Rightarrow x \sqsubseteq z.$$

Vi sier at $a \in \mathbf{B}$ er *minimal* hvis $a \neq \mathbf{0}$ og vi har

$$x \sqsubseteq a \Rightarrow x = \mathbf{0} \vee x = a.$$

Hvis $x \in \mathbf{B}$ er forskjellig fra $\mathbf{0}$ er enten b binimal, eller så finnes det $b_1 \sqsubseteq b$ slik at $b_1 \neq \mathbf{0}$ og $b_1 \neq b$. Hvis b_1 heller ikke er minimal, kan vi fortsette. Siden $\mathbf{B}$ er endelig, kan vi ikke fortsette i det uendelige, så før eller siden treffer vi på et minimalt element.

La A være mengden av minimale elementer.
For $x \in \mathbf{B}$, la

$$P(x) = \{a \in A \mid a \sqsubseteq x\}.$$

For $B \subseteq A$, la $b(B) = \sqcup B$, det vil si at vi skriver elementene i B i rekkefølge og skriver $\sqcup$ mellom dem. Hvis vi pr. definisjon lar $b(\emptyset) = \mathbf{0}$, kan vi vise (betrakt det som en overkommelig utfordring) at vi har samsvar mellom den gitte Booleske algebraen og potensmengden til A ved

$$b(P(x)) = x \text{ for alle } x \in \mathbf{B}$$

og

$$P(b(B)) = B \text{ for alle } B \subseteq A.$$

Den som prøver seg på dette bør starte med å vise at $b(A) = \mathbf{1}$.

3.5 Blandede oppgaver

Oppgave 3.5.1 Betrakt det sammensatte utsagnet

$$\phi = (A \to B) \wedge (B \to C) \to (A \vee C).$$

a) Undersøk om ϕ er en tautologi eller ikke.

b) Illustrer mengden

$$((X^c \cup Y) \cap (Y^c \cup Z))^c \cup (X \cup Z)$$

i et Venn-diagram og avgjør om denne mengden alltid utgjør hele universet.

c) Forklar hvorfor vi kan bruke resultatet i b) til å vise a) og hvordan Venn-diagrammet gir oss informasjon om når ϕ eventuelt kan være usann.

Oppgave 3.5.2 a) Vis at følgende utsagn ikke er en tautologi:

$$((A \to (A \to B)) \to (B \to (B \to A)).$$

b) Finn et ekvivalent utsagn som er så kort som mulig.

c) Finn en digital krets som representerer utsagnet.

Oppgave 3.5.3 La

$$U = \{1, 2, 3, 4, 5, 6, 7\}$$

og la $A \subseteq U$, $B \subseteq U$ og $C \subseteq U$.

a) Forklar hvorfor deter umulig å velge A, B og C slik at alle feltene i Venn-diagrammet er representert ved et element i U.

b) Finn eksempler på A, B og C slik at

$$U = A \cup B \cup C$$

og slik at alle andre felt i Venn-diagrammet er representert ved et element i U.

c) Bruk de eksemplene du finner til å finne en naturlig forbindelse mellom U og $\mathcal{P}(\{A, B, C\})$.

Oppgave 3.5.4 La $\mathbf{B}$ med $\mathbf{0}$, $\mathbf{1}$, $\sim$, $\sqcup$ og $\sqcap$ være en Boolesk algebra.

a) La $t(x, y, z) = x \sqcup (y \sqcap \sim z)$ og la $s(x, y, z) = (\sim y \sqcup z) \sqcap \sim x$.
Bruk aksiomene for Booleske algebraer og de reglene vi har vist i teksten til å vise at $t(x, y, x) = \sim s(x, y, z)$ holder for alle $x, y, z \in \mathbf{B}$.

b) Finn tilsvarende sammensatte utsagn T og S i utsagnslogikk i utsagnsvariablene X, Y og Z og vis at $T \leftrightarrow \neg S$ er en tautologi.

c) Finn enklest mulige sammensatte utsagn som er ekvivalente til T og S og konstruer digitale kretser som representerer disse.

Kapittel 4

Relasjoner og funksjoner

4.1 Ordnede par og liknende

Når vi ga eksempler på mengder og på hvordan vi kunne definere mengder, brukte vi $\mathbb{R}^2$ og løsningsmengder til likninger i noen av eksemplene. Vi gjorde det som er helt vanlig i matematikken, vi problematiserte overhode ikke hva disse elementene i $\mathbb{R}^2$ egentlig er. I så godt som alle fremstillinger av matematikkens mengdeteoretiske grunnlag er *ordnede par* et matematisk grunnbegrep som vi mener å forstå like godt som for eksempel tallene 3 og 17, vi forklarer hva vi mener og vi bare godtar at disse objektene finnes.
Dette skal være vår tilnærming også i dette kurset. Hvis a og b er to objekter, lar vi (a, b) være det *ordnede paret* av a og b. Det vesentlige er at rekkefølgen spiller en rolle, slik at om a og b er forskjellige, så er $(a, b) \neq (b, a)$. Dette kan vi uttrykke ved en regel for når to ordnede par er like:

$$(a, b) = (c, d) \text{ hvis og bare hvis } a = c \text{ og } b = d.$$

I oppgave 4.1.1 skal vi se på hvordan ordnede par kan defineres i mengdelæren, men forsøk på å bringe denne typen definisjoner til elever uten helt spesielle interesser for matematikk resulterer trolig i ødsling av tid.

Når vi har sagt hva vi mener med ordnede par, kan vi også definere $\mathbb{R}^2$ som mengden av ordnede par av reelle tall, $\mathbb{N}^2$ som mengden av ordnede par av naturlige tall, $\mathbb{Z}^2$ og $\mathbb{Q}^2$. Vi kan også definere *Kartesiske produkter* generelt:

Definisjon 4.1 La A og B være to mengder.
Med produktet $A \times B$ mener vi $\{(a, b) \mid a \in A \wedge b \in B\}$, altså mengden av ordnede par hvor det første elementet er fra A og det andre fra B.

I endel anvendelser av matematikk opererer vi ikke bare med ordnede par, men med ordnede trippler, ordnede kvadrupler, kvintupler m.m. Noen har vært utsatt for regning i x, y, z-rommet. Lesere av fysikk vil ha hørt om det firedimensjonale tid-rommet hvor et punkt er bestemt av fire koordinater: Plassering i et tredimensjonalt rom samt tiden punktet befinner seg i.

Vi vil definere begrepet *ordnet sekvens* med den samme grad av presisjon som vi brukte da vi definerte ordnet par:

Definisjon 4.2 La n være et naturlig tall. Med et *ordnet n-tuppel* eller en *ordnet sekvens av lengde n* mener vi et objekt $(a_1, \ldots, a_n)$. To ordnede sekvenser er like hvis de har samme lengde og er like punkt for punkt.

Eksempler 4.1 a) $(1, 3, 4, 2)$ og $(1, 3, 2, 4)$ er to 4-tupler (det vi innledningsvis kalte kvadrupler) eller ordnede sekvenser av lengde 4. De er ikke like fordi vi har byttet om på de to siste leddene.

b) $(1, \frac{1}{2}, \frac{1}{3})$ og $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4})$ er to ordnede sekvenser, men de er ikke like fordi de har forskjellig lengde.

c) Det å skulle løse to likninger med tre ukjente kan uttrykkes som å skulle finne alle ordnede trippler (x, y, z) slik at for eksempel

$$x^2 + y^2 = z^2$$
$$x + y = z$$

I oppgave 4.1.2 skal vi se på hvordan vi kan utvide Kartesisk produkt til å omfatte mer enn to mengder. I oppgave 4.1.3 ser vi på hvordan vi også kan tufte definisjonen av endelige sekvenser på et mengdeteoretisk grunnlag.

Oppgaver til avsnitt 4.1

Oppgave 4.1.1 Hvis a og b er to objekter, kan vi definere $(a, b) = \{\{a\}, \{a, b\}\}$. Vis at hvis vi bruker denne definisjonen, så gjelder det at

$$(a, b) = (c, d) \Leftrightarrow a = c \wedge b = d.$$

Oppgave 4.1.2 Hvis A, B og C er mengder, lar vi $A \times B \times C$ være mengden av alle ordnede trippler (a, b, c) slik at $a \in A$, $b \in B$ og $c \in C$.

a) Gi en direkte definisjon av $\mathbb{N} \times \mathbb{R} \times \mathbb{N}$.

b) Prøv å formuler en definisjon av $A_1 \times \cdots \times A_n$ generelt.

c) Se på definisjonene av $A \times B \times C$, $A \times (B \times C)$ og $(A \times B) \times C$.
Drøft hvorfor disse mengdene strengt tatt er forskjellige og hva sammenhengen mellom dem er.

Oppgave 4.1.3 I oppgave 4.1.1 så vi på hvordan vi kan betrakte et ordnet par som en mengde.
Vi kan definere $(a, b, c) = ((a, b), c)$, $(a, b, c, d) = ((a, b, c), d)$ og så videre.

a) Forklar hvordan disse definisjonene kan utvides til en definisjon av *ordnet sekvens av lengde n* for enhver n, slik at karakteriseringen av når to ordnede sekvenser av lengde n er like holder.

b) [u]Definisjonen av en ordnet sekvens av lengde n som vi la opp til i a) har en svakhet. Hvis vi ser på $(((a,b),c),d)$ kan dette oppfattes både som et ordnet par, en ordnet sekvens av lengde tre og som en ordnet sekvens av lengde fire. Diskuter hvordan vi kan endre definisjonen av en ordnet sekvens slik at to ordnede sekvenser er like hvis og bare hvis de har samme lengde og er punktvis like.

Oppgave 4.1.4 Bestem hvilke ordnede par som er like av $(1,1)$, $(4,1)$, $((-1)^2, 2^2)$, $(1^2, (-1)^2)$, $(1,-1)$ og $(2^2,1)$.

Oppgave 4.1.5 Undesøk om noen av disse ordnede sekvensene av lengde 3 er like: $(1,4,9)$, $(4,9,1)$, $(1,2+2,3+3+3)$, $(1+3,1+3+5,3-2)$, $(9,4,1)$ og $(3^2,2^2,1^2)$.

Oppgave 4.1.6 La (a,b,c) og (d,e,f) være to ordnede sekvenser av reelle tall av lengde 3 slik at

- $a+b=d+e$
- $a+c=d+f$
- $b+c=e+f$

Vis at de to sekvensene må være like.

4.2 Relasjoner

Ordet 'relasjon' er et fremmedord på norsk som kan bety noe slikt som 'forbindelse', 'sammenheng' og 'forhold til'. Vi kan snakke om mellomfolkelige relasjoner når vi diskuterer hvordan land og folkeslag trives sammen, og vi kan snakke om at 'utgiftene må sees i relasjon til nytteverdien' når vi skal argumentere for kjøp av en dyr bil. Vi bruker den samme ordstammen når vi sprer om oss med det intelektuelle 'alt er relativt', og engelsmennene bruker ordet 'relative' om en så vanlig forbindelse som en slektning.

I matematikken vil vi bruke ordet 'relasjon' om en matematisk størrelse som beskriver en sammenheng mellom to objekter.
La oss se på noen eksempler før vi gir den formelle definisjonen.

Eksempler 4.2 a) Når vi regner med tall og ulikheter bruker vi symboler som $<$ og $\leq$. Disse uttrykker en sammenheng mellom to tall, nemlig at det ene er mindre enn det andre, eller at det er mindre eller lik det andre.

b) I tallteori er det at et tall er en faktor i et annet tall en viktig egenskap. Vi skriver ofte $a|b$ for å uttrykke at a er en faktor i b.

c) I et slektsregister er det viktig å markere hvem som er far eller mor til hvem, hvem som er gift med hvem, hvilket kjønn den enkelte har og muligens andre relasjoner.

Hvis vi kjenner til 'barn-foreldre'-relasjonen i slekten, kan vi også finne ut av hvem vi er etterkommere av. Dette er et eksempel på en relasjon hvorfra vi kan utlede andre relasjoner.

d) I utsagnslogikken definerte vi hva vi mener med at et sammensatt uttrykk impliserer et annet eller at det er ekvivalent med et annet. Den presise definisjonen gjør $\Rightarrow$ og $\Leftrightarrow$ til relasjoner på mengden av sammensatte utsagn i utsagnslogikk.

d) Det er fullt mulig å snakke om relasjoner mellom objekter av vidt forskjellig karakter. På de fleste skoler er mengden av lærere disjunkt fra mengden av elever, men fortsatt kan vi snakke om lærer-elev-relasjonen i betydningen A er læreren til B i minst et fag.
Et annet eksempel på en relasjon som er av interesse for eiendomsregisteret, men hvor de to "aktørene" er vidt forskjellige er

Person A er eier av eiendom B.

Når vi skal lage en matematisk modell for relasjonsbegrepet er det viktig å få med seg de viktige aspektene, men ellers gjøre begrepet så generelt som mulig. Det er ingen grunn til å tro at vi kan begrense oss til å studere de relasjonene som er eller vil bli av interesse for noen. Det som er viktig er at en relasjon representerer en mulig sammenheng mellom to objekter, og at sammenhengen kan være asymmetrisk og mellom forskjellige sorter objekter. (Symmetri ville betydd, for eksempel i d) over, at huset og tomten min eier meg, og det er ikke tilfelle.

Definisjon 4.3 La A og B være mengder. En (binær) relasjon fra A til B er en delmengde $R \subseteq A \times B$.
Vi sier at R er en relasjon *på* A dersom R er en relasjon fra A til A.

Vi skrev ordet 'binær' i parentes for å antyde at definisjonen kan utvides til å omfatte at flere enn to objekter kan stå i et forhold til hverandre.
Hvis R er en relasjon fra A til B vil vi kalle A for *argumentområdet* til R og B for *verdiområdet* til R. Denne språkbruken vil falle oss mer naturlig etter at vi har lest avsnitt 4.3, men den antyder også at det kan være forskjell i viktighetsgrad mellom objektene i argumentområdet og objektene i verdiområdet. For en lærer er det av interesse hvilke elever hun/han har, så lærerne vil nok oppfatte seg som argumentområdet i lærer-elev-relasjonen. For elevene er det nok mere viktig å ha oversikt over hvilke lærere de har, og de vil nok heller se på elev-lærer-relasjonen.
I et personregister kan det være aktuelt å liste opp hvilke eiendommer den enkelte personen besitter, mens i et eiendomsregister vil det være viktig å liste opp den (eller de) som står som eier(e).
Det at det ofte er naturlig å se en relasjon fra to kanter, har ledet til at vi definerer den *inverse* relasjonen:

Definisjon 4.4 La R være en relasjon fra en mengde A til en mengde B.
Med den *inverse* relasjonen mener vi relasjonen R^{-1} fra B til A definert ved

$$R^{-1} = \{(b, a) \mid (a, b) \in R\}.$$

Eksempler 4.3 Vi så hvordan vi kan konstruere den inverse til en relasjon. Nå skal vi se på noen eksempler hvor vi konstruerer nye relasjoner fra gamle på andre interessante måter:

a) Hvis vi har gitt relasjonene 'mor til' og 'far til' har vi også indirekte gitt de fire relasjonene 'mormor til', 'morfar til', 'farmor til' og 'farfar til'.

b) Vi har tidligere antydet at 'lærer-elev'-relasjonen er av interesse. Det er ganske opplagt at for elevene er 'barn-foresatt'-relasjonen av interesse. Det innebærer imidlertid at lærerne også vil ha en gruppe foresatte de må forholde seg til. 'Lærer-foresatt'-relasjonen er avledet av de to foregående.

Vi skal nå gi en definisjon som fanger opp denne typen konstruksjoner:

Definisjon 4.5 La R være en relasjon fra A til B og S en relasjon fra B til C.
Vi definerer *sammensetningen*

$$T = S \circ R$$

av R og S som en relasjon fra A til C ved

$$(a, c) \in T \text{ hvis det finnes } b \in B \text{ slik at } (a, b) \in R \text{ og } (b, c) \in S.$$

Hvis vi i eksemplet over lar M være 'mor til'-relasjonen og F være 'far-til'-relasjonen, vil $M \circ F$ være 'farmor', $M \circ M$ være 'mormor'. For å få Lærer-foresatt-relasjonen må vi se på $L \circ B$, hvol L er 'lærer til' og B er 'barn av'.

Oppgaver til avsnitt 4.2

Oppgave 4.2.1 La $A = \{1, 2\}$ og $B = \{1, 2, 3, 4\}$.
Forklar hvorfor det finnes 256 relasjoner fra A til B.

Oppgave 4.2.2 La $A = \{1, 2\}$, $B = \{3, 4\}$ og $C = \{5, 6\}$.
La $R = \{(1, 3), (2, 4)\}$ og $S = \{(3, 6), (4, 6), (3, 5)\}$.
Bestem argumentområdene og verdiområdene til S og T.
Finn S^{-1}, R^{-1} og $S \circ R$.
Kan du finne en relasjon som kan kalles $S^{-1} \circ R^{-1}$?
Vis at $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$.
Tror du dette er en tilfeldighet?

Oppgave 4.2.3 Relasjonen $<$ på $\mathbb{N}$ er definert som

$$\{(n, m) \in \mathbb{N}^2 \mid n < m\}.$$

Vis at $(<)^2 \subset <$, det vil si at det er en ekte delmengde, hvor $(<)^2 = < \circ <$.
Er $\leq^2 = \leq$?

Oppgave 4.2.4 La R være relasjonen på $\mathbb{N}$ definert ved at $(n, m) \in R$ hvis $\frac{m}{n} \in \mathbb{N}$. Dette er den samme relasjonen som relasjonen $|$ vi har sett på tidligere.

a) Forklar hvorfor $R^{-1} \circ R$ består av alle par av naturlige tall.

b) Beskriv $R \circ R^{-1}$ og forklar hvorfor du mener den er slik den er.

Oppgave 4.2.5 Finn et eksempel på en relasjon R på en mengde A slik at $R^{-1} \circ R \neq R \circ R^{-1}$.
Forklar hva som er sammenhengen mellom de to sammensatte relasjonene.

4.3 Funksjoner

Funksjonsbegrepet er viktig i matematikken, og egentlig mer fundamentalt enn relasjonsbegrepet. På barnetrinnet møter elevene funksjonene $+$, $-$, $\cdot$ og $:$ og lærer å regne med dem. I Grunnskolen lærer også elevene formler for arealet av en sirkel, volumet av en pyramide, overflaten av en kule og andre formler for areal og overflater. Alle disse formlene beskriver egentlig funksjoner.
På matematikk-intensive retninger på Videregående Skole blir elevene utsatt for en mer systematisk funksjonslære, gjennom kjennskap til trigonometriske funksjoner, logaritmefunksjonen og eksponensialfunksjonen, og sammenhengen mellom disse funksjonene er en viktig del av det de skal lære.
Det er naturlig å forestille seg at hvis man skal lage en matematisk modell for funksjonsbegrepet, så bør man prøve å fange opp hva slags sammenheng det kan være mellom det vi putter inn i funksjonen og det vi får ut. De fleste vil ha en forestilling av at det foregår en prosess fra argument til verdi. Denne forestillingen har mye for seg, og når man prøver å lage matematiske modeller for hvordan datamaskiner behandler inngangsdata og kommer med et svar, er det nettopp de viktigste aspektene ved en slik prosess man prøver å fange opp.

Når vi nå skal gi en mengdeteoretisk tolkning av hva vi mener med en funksjon, skal vi legge oss på et lavere ambisjonsnivå. Vi skal nevne to grunner til det:

1. Vi bør begrense oss til det ene grunnleggende aspektet ved funksjoner, nemlig at hvis vi putter noe inn får vi en og bare en ting ut. Hvis vi får bruk for å modellere flere aspekter får vi heller gjøre det ved å legge på tillegsinformasjon, ikke endre grunnbegrepet.

2. Det finnes mange situasjoner hvor vi med rimelighet kan snakke om funksjoner, men hvor det ikke er noen påtakelig sammenheng mellom inngangsdata og utgangsdata.

Definisjon 4.6 La A og B være mengder.
Med en *funksjon* $f : A \to B$ mener vi en relasjon f fra A til B slik at

$$\text{For alle } a \in A \text{ finnes det en og bare en } b \in B \text{ slik at } (a, b) \in f.$$

Når $f : A \to B$ er en funksjon fra A til B, skriver vi $f(a)$ for den ene b som er slik at $(a, b) \in f$.

Det vi her har gjort er å identifisere en funksjon med sin *graf*, nemlig relasjonen $f(a) = b$ som vi får ut fra vanlig språkbruk.

Eksempel 4.4 La $A = \{0, 1\}$ og $B = \{0, 1, 2, 3\}$. Da finnes det nøyaktig 16 forskjellige funksjoner $f : A \to B$, vi har fire valg for $f(0)$ og fire valg for $f(1)$. Disse valgene kan gjøres uavhengige av hverandre, så det totale antall friheter er $4 \cdot 4 = 16$.

Eksempel 4.5 Vi kan definere en funksjon $f : \mathbb{R} \to \mathbb{R}$ ved at $f(x) = x^3 - 17$ om x er et rasjonalt tall, mens $f(x) = 0$ ellers.
Poenget her er at f ikke er særlig kontinuerlig, og det kan bli svært vanskelig å tegne grafen til f. Likefullt er f en funksjon.

Eksempel 4.6 La $X = \{P_1, \ldots, P_n\}$ være en mengde utsagnsvariable.
La

$$V = \{f \mid f : X \to \{\top, \bot\}\}$$

og la

$$TAB = \{F \mid F : V \to \{\top, \bot\}\}.$$

V er mengden av fordelinger av sannhetsverdier til de gitte utsagnsvariablene, og det svarer til alle linjene i en sannhetsverditabell. TAB svarer da til alle mulige sannhetsverditabeller for de gitte utsagnsvariablene, når vi ikke tar hensyn til mellomregningene, men bare til sluttsøylen.
Dette illustrerer at vi kan fange opp mye mer enn funksjoner på tallmengder ved det generelle funksjonsbegrepet.

Da vi så på relasjoner generelt, definerte vi sammensetning av relasjoner. Den som er vant til å arbeide med funksjoner har trolig også sett på sammensetning av funksjoner:

$$(g \circ f)(x) = g(f(x)).$$

Hvis vi nå ser på vår definisjon av en funksjon som en relasjon, skal vi ha at $(x, z) \in g \circ f$ hvis og bare hvis det finnes en y slik at $(x, y) \in f$ og $(y, z) \in g$.
Skriver vi dette på den vanlige måten, er dette det samme som at det finnes en y slik at $y = f(x)$ og $z = g(y)$. Siden vi arbeider men funksjoner finnes det en og bare en y slik at $y = f(x)$, så $(x, z) \in g \circ f$ betyr det samme som at $z = g(f(x))$.

Hvis R er en relasjon fra A til B definerte vi R^{-1} som en relasjon fra B til A. Det betyr at hvis $f : A \to B$ er en funksjon har vi hjemmel for å snakke om f^{-1} som en relasjon fra B til A. Det er imidlertid en betenkelighet med den språkbruken, og det er at f^{-1} ikke alltid er en funksjon. Det kan være to grunner til at en relasjon S fra B til A ikke er en funksjon:

1. Det finnes en $b \in B$ slik at vi ikke har $(b, a) \in S$ for noen $a \in A$.
2. Det finnes en $b \in B$ slik at det finnes forskjellige a og c i A hvor vi har både $(b, a) \in S$ og $(b, c) \in S$.

Vi skal se på eksempler på funksjoner f hvor disse problemene oppstår for f^{-1}. La oss starte med problem 1:

Eksempel 4.7 La $f : \mathbb{N} \to \mathbb{N}$ være definert ved $f(x) = 2x$. Da er

$$f^{-1} = \{(2x, x) \mid x \in \mathbb{N}\}.$$

Hvis vi setter $x = 1$ er ikke $\frac{x}{2} \in \mathbb{N}$, så det finnes ingen $y \in \mathbb{N}$ slik at $(x, y) \in f^{-1}$.

Vi ser at vi får dette problemet hver gang $f : A \to B$ og det finnes en $b \in B$ som ikke er på formen $f(a)$ for noen $a \in A$.

Definisjon 4.7 La $f : A \to B$.
Vi sier at f avbilder A *på* B, eller at f er *surjektiv* dersom vi for alle $b \in B$ har en $a \in A$ slik at $f(a) = b$.

Eksempler 4.8 Følgende er eksempler på surjektive funksjoner:

a) La $\mathbb{R}^+$ være mengden av positive reelle tall. La $f : \mathbb{R} \to \mathbb{R}^+$ være definert ved
$$f(x) = 2^x.$$

b) La $f : \mathbb{N} \to \{0, 1, 2\}$ være definert ved at $f(x)$ er resten vi får når vi deler x på 3.

c) La $f : (\mathcal{P}(\mathbb{N}) \setminus \{\emptyset\}) \to \mathbb{N}$ være definert ved $f(A)$ er det minste tallet i A.

Følgende funksjoner er eksempler på funksjoner som ikke er surjektive:

d) $f : \mathbb{R} \to \mathbb{R}$ definert ved $f(x) = 2^x$.

e) $f : \mathbb{N} \to \mathbb{N}$ definert ved $f(n) = n + 1$.

Ser vi på eksemplene a) og d) ser vi at vi kan gjøre en funksjon som ikke er surjektiv om til en funksjon som er surjektiv ved å endre på verdimengden, men egentlig ikke forandre funksjonen selv.

Det er ikke vanskelig å vise at sammensetningen av surjektive funksjoner er surjektiv. Dette overlates leseren som oppgave 4.3.2.

La oss nå se på et eksempel hvor problem 2 oppstår:

Eksempel 4.9 La $f : \mathbb{R} \to \mathbb{R}$ være definert ved

$$f(x) = x^2.$$

f er ikke surjektiv, men det er ikke noen stor hindring for å studere den inverse, det er bare å begrense den til ikke-negative tall (eller å utvide tallområdet til de komplekse tallene hvis man er fortrolig med dem). Det er et større problem at hvis $y > 0$ finnes det to tall x slik at $f(x) = y$.
Alle som har lært å løse 2. gradslikninger vet at vi bruker $\sqrt{x}$ som den inverse, og at vi bruker $\pm\sqrt{x}$ i formelen for løsningen til en 2.gradslikning.

Definisjon 4.8 La $f : A \to B$ være en funksjon. Vi sier at f er *enentydig* eller *injektiv* hvis vi for alle a og b i A har at om $f(a) = f(b)$, så er $a = b$.

Teorem 4.1 *La $f : A \to B$ være en funksjon som både er surjektiv og injektiv. Da er f^{-1} en funksjon fra B til A.*

Vi har allerede kommenter at det at f er surjektiv og injektiv svarer til de to betingelsene for at f^{-1} er en funksjon, så det er egentlig ikke noe å vise.

Selv om uttrykket f^{-1} gir mening for alle funksjoner f fordi de er relasjoner, skal man være forsiktig med å bruke det i andre tilfeller enn der f^{-1} faktisk er en funksjon, i alle fall bør man presisere at man fraviker denne anstendighetsregelen når man gjør det. I oppgave 4.3.4 skal vi se på en situasjon hvor en generell bruk av inverse funksjoner gir mening.

Et viktig aspekt ved inversdannelser i matematikken er at hvis man gjør noe og så gjør det motsatte (bruker den inverse) så skal man komme tilbake til utgangspunktet. Hvis man legger til et tall og så trekker det fra igjen er man tilbake til utgangspunktet og hvis man først multipliserer med et tall forskjellig fra null og så dividerer med det samme tallet er man også tilbake til utgangspunktet. Dette gjelder generelt for sammensetningen av en funksjon som har en invers med sin inverse, vi har alltid at $f^{-1}(f(x)) = x$.

Oppgaver til avsnitt 4.3

Oppgave 4.3.1 Undersøk hvilke av disse funksjonene som er surjektive, injektive og bestem f^{-1} der f har en invers:

a) $A = B = \mathbb{N}$ og $f(n) = n^2$.

b) $A = \mathbb{R}$, $B = \mathbb{R}^+$ og $f(x) = e^x$.

c) $A = B = \{1, 2, 3, 4, 5\}$ og $f(x) = 6 - x$.

d) $A = \{0, 1, 2, 3\}$, $B = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ og $f(x) = x^2$.

e) $A = \mathbb{R}$, $B = [-1, 1]$ og $f(x) = \sin x$.

Oppgave 4.3.2 Vis at hvis $f : A \to B$ og $g : B \to C$ er surjektive funksjoner så er $(g \circ f) : A \to C$ også surjektiv.

Oppgave 4.3.3 Vis at hvis $f : A \to B$ og $g : B \to C$ er injektive funksjoner så er $(g \circ f) : A \to C$ også injektiv.

Oppgave 4.3.4 La $f : A \to B$ være en vilkårlig funksjon.
Hvis $C \subseteq B$ definerer vi det *inverse bildet* $f^{-1}[C]$ av C under f ved

$$f^{-1}[C] = \{a \in A \mid f(a) \in C\}.$$

Vis at det inverse bildet respekterer mengdealgebraen på B ved å vise

a) Hvis $C \subseteq B$ og $D \subseteq B$ er

$$f^{-1}[C \cap D] = f^{-1}[C] \cap f^{-1}[D]$$

og

$$f^{-1}[C \cup D] = f^{-1}[C] \cup f^{-1}[D].$$

b) Hvis $C \subseteq B$ er

$$f^{-1}[B \setminus C] = A \setminus f^{-1}[C].$$

Vi har brukt uttrykket $f^{-1}[\cdot]$ for å markere at dette er noe annet enn f^{-1} selv om det er et visst slektskap.
Drøft hva som er argumentområdet og verdiområdet til $f^{-1}[\cdot]$.

Oppgave 4.3.5 I lys av Oppgave 4.3.4 er det naturlig å definere det *direkte bildet* av en mengde under en funksjon som

$$f[C] = \{f(a) \mid a \in C\}.$$

a) Vis at hvis $f : A \to B$ $C \subseteq A$ og $D \subseteq A$, såvil

$$f[C \cup D] = f[C] \cup f[D].$$

b) La $A = \{1, 2, 3, 4\}$ og $B = \{1, 2\}$. La $f(1) = f(2) = 1$ og $f(3) = f(4) = 2$.
La $C \subset A$ være mengden $\{1, 3\}$ og $D \subset A$ være mengden $\{2, 4\}$.
Vis at $f[C \cap D] \neq f[C] \cap f[D]$ og at $f[A \setminus C] \neq B \setminus f[C]$.

c) Vis at vi for alle funksjoner $f : A \to B$, $C \subseteq A$ og $D \subseteq A$ har at

$$f[C \cap D] \subseteq f[C] \cap f[D].$$

d) [u] Finn naturlige egenskaper som sikrer at det direkte bildet respekterer snitt og komplement.

Oppgave 4.3.6 [u] La $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.
La $P : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ være definert ved

$$P(n, m) = \frac{1}{2}((n + m)^2 + 3m + n).$$

Vis at P er både injektiv og surjektiv.
[Hint: Regn ut $P(0, 0)$, $P(1, 0)$, $P(0, 1)$, $P(2, 0)$, $P(1, 1)$ og $P(0, 2)$ og se om du finner et mønster. Prøv å argumenter for at det mønsteret du finner fortsetter.]
Eksistensen av P viser at det i en viss forstand finnes like mange par av tall som tall. Cantor definerte to mengder A og B som like store hvis det finnes en $F : A \to B$ som er både injektiv of surjektiv. Hvis noen synes at dette strider mot sunn fornuft, så må man huske at dette bare er en teknisk definisjon av et matematisk begrep. Den stemmer overens med hva vi mener med 'like mange' for endelige mengder og viderfører noe (men ikke alt) av intuisjonen omkring 'like stor' til uendelige mengder.

4.4 Algoritmer

Da vi innførte funksjonsbegrepet innrømmet vi at det finnes mange sider ved vår intuitive oppfatning av hva en funksjon er som ikke blir fanget opp av definisjonen vår. Den viktigste intuisjonen er at en funksjon faktisk **gjør** noe med de argumentene vi gir den før den gir oss en verdi. I endel lærebøker blir funksjonsbegrepet forklart ved hjelp av en boks hvor vi putter x inn på den ene siden og så kommer $f(x)$ ut på den andre siden. Vi hevdet imidlertid også at det finnes så mange tilfeller hvor det kan være greit å snakke om funksjoner, men hvor det ikke er noen slik underliggende "prosess".

I dette avsnittet skal vi snakke om noen slike prosesser, om det vi kaller *algoritmer*. Det vil føre for langt å gå inn på noen systematisk behandling av algoritmebegrepet, dette er noe som hører inn under et dypere studium av informatikkens og matematikkens grunnlag. Siktemålet er gjennom noen eksempler å gi en intuisjon om hva en algoritme er og hvorfor det kan være interesssant å studere fenomenet mer systematisk.
Vi skal definere hva vi mener med en algoritme, men definisjonen er så upresis at algoritmebegrepet vårt ikke kan brukes til å formulere matematiske teoremer.

Definisjon 4.9 En *algoritme* er et sett regler som forteller oss hvordan vi mekanisk skal kunne finne en verdi $F(x)$ fra et argument x.

Eksempel 4.10 Når vi skal multiplisere to store tall, for eksempel $328 \cdot 542$ har vi lært at vi først multipliserer 2 med 328. På linjen under, og forskjøvet en plass til venstre, skriver vi 4 multiplisert med 328. Slik fortsetter vi til alle sifrene i tallet til høyre er multiplisert med tallet til venstre og svarene er skrevet under hverandre forskjøvet en plass for hver linje. Til slutt legger vi sammen alle linjene slik de står og får svaret.
Alle som kan multiplisere ensifrede tall med flersifrede tall og legge sammen flere flersifrede tall kan greie å multiplisere flersifrede tall med hverandre uten å forstå hvorfor det de gjør er riktig. de kan følge algoritmen mekanisk.
Det finnes selvfølgelig underliggende algoritmer for å multiplisere ensifrede og flersifrede tall med hverandre og for å summere flersifrede tall, og det eneste elevene virkelig må kunne er den lille multiplikasjonstabellen samt hvordan ensifrede og flersifrede tall legges sammen.

Eksempel 4.11 Elevene lærer også en algoritme for hvordan man deler et flersifret tall med et annet. Leseren vil kjenne denne algoritmen like godt som multiplikasjonsalgoritmen, så poenget her er bare at vi erkjenner at også divisjon er noe man lære å utføre mekanisk. Man trenger ikke å tenke seg om for å dividere riktig.

Eksempel 4.12 En kjent klassisk algoritme for å finne største felles faktor til to (store) tall er Euklids algoritme.
Den benytter seg av to observasjoner:

1. Hvis $n < m$ er to naturlige tall og n er en faktor til m, så er n den største felles faktoren til n og m.

2. Hvis $n < m$, n ikke er en faktor til m og k er resten vi får når vi deler m med n, så er den største felles faktoren til n og m den samme som den største felles faktoren til k og n.

Den første observasjonen er helt triviell, mens den andre trenger litt tankevirksomhet. Vi overlater observasjonen til leseren i ubevist tilstand, men leseren bør ikke oppfatte det som en avskrekking; det er lett.
Dette gir oss en algoritme for å finne den største felles faktoren til to tall, en algoritme som vi illustrerer med et eksempel, vi vil finne største felles faktor til 2331 og 171.
Ideen er at vi deler det største tallet på det minste, og fortsetter deretter med det minste og med resten, inntil divisjonen går opp. Da har vi funnet største felles faktor:

1. $2331 = 13 \cdot 171 + 108$
2. $171 = 1 \cdot 108 + 63$
3. $108 = 1 \cdot 63 + 45$
4. $63 = 1 \cdot 45 + 18$
5. $45 = 2 \cdot 18 + 9$
6. $18 = 2 \cdot 9$

Vi har på en helt mekanisk måte bestemt at den største felles faktoren til 2331 og 171 er 9.

Eksempel 4.13 Når jeg skriver dette kompendiet, bruker jeg et tekstbehandlingsystem som heter *Latex*. Den teksten jeg skriver på skjermen ser ganske anderledes ut enn den som kommer på trykk. For eksempel vil jeg skrive '\alpha' hvis jeg ønsker at det skal stå 'α'.
Poenget her er at det finnes en underliggende algoritme, et *program* som omformer det jeg skriver p{\aa} skjermen til det jeg ønsker at leseren skal se.
Programmer må oppfattes som algoritmer fordi de egentlig er instruksjoner som forteller en datamaskin hvordan den skal finne utgangsdata fra inngangsdata

Eksempel 4.14 En algoritme som begynnende datastudenter ofte får i oppdrag å skrive et program for er algoritmen for *Hanois Tårn*.
Hanois tårn består av tre pinner. På den ene pinnen ligger det n ringer i forskjellig størelse slik at det aldri ligger en større ring over en mindre. Poenget er å flytte alle ringene over til en av de andre pinnene ved bare å flytte én og én ring, og slik at vi aldri har en stor ring oppå en mindre.
Hvis $n = 1$ er det bare å flytte ringen.
Hvis $n > 1$ sier algoritmen at vi først må bruke den til å flytte de $n - 1$ øverste ringene til den tredje pinnen. Deretter flytter vi den største ringen dit den skal og til sist bruker vi algoritmen til å flytte de $n - 1$ mindre ringene fra den tredje pinnen til dit de skal.

Her har vi beskrevet algoritmen ved å henvise til at vi skal bruke algoritmen på en mindre mengde ringer. Siden vi vet hva vi skal gjøre når det bare er én ring, fungerer dette. Prøv selv med tre eller fire spillkort med forskjellige valører, 'puttebokser' eller liknende, og se at det virker.

Eksempel 4.15 Et annet eksempel på en selvkallende algoritme kan være følgende sorteringsalgoritme. Anta at vi har gitt en lang liste av navn og vi ønsker å sortere den alfabetisk. Hvis listen ikke var lang likevel, for eksempel bare inneholdt ett navn, så var den ferdigsortert. Algoritmen for å sortere en liste med ett navn er altså kjent, vi sier "i orden". Hvis listen er lenger, tar vi utgangspunkt i det øverste navnet på listen. Vi søker nedover i listen, og alle navn som kommer foran det første i alfabetet flyttes oppover. Samtidig sjekker vi om listen er ferdig sortert. Hvis listen er ferdig sortert, skriver vi "i orden" og avslutter. Hvis listen ikke er ferdig sortert bruker vi algoritmen på den delen av listen som nå er kommet over det navnet som opprinnelig sto øverst og vi bruker algoritmen på de navnene som ble stående igjen under. Når dette er gjort er listen ferdig sortert.

Sorteringsalgoritmen beskrevet i dette siste eksemplet er ikke den mest effektive og er mere egnet til å illustrere algoritmebegrepet enn til teknologiske anvendelser.

Det finnes prosesser som følger faste mønstre, men som ikke kan regnes som utførelse av algoritmer:

Eksempel 4.16

- Man putter 34 nummererte kuler i en beholder.
- Man blander kulene godt.
- Man trekker ut 7 av kulene, mens man blander mellom hver gang.
- Man deler et betydelig beløp på de som har krysset av syv nummererte ruter på et ark, og nummerne som er krysset av svarer til nummerene på de kulene som blir trukket ut.

Erfaring tilsier at resultatet av denne prosessen varierer fra gang til gang, selv om den blir riktig utført hver lørdag. Resultatet av en algoritme skal ikke være situasjonsavhengig så lenge utgangspunktet er det samme.

Oppgaver til avsnitt 4.4

Oppgave 4.4.1 Et populært eksempel blant matematikere som besøker grunnskolen er et spill hvor to spillere sitter med en felles bunke fyrstikker. Etter tur kan spillerne trekke en, to eller tre fyrstikker, og den som forsyner seg med den siste fyrstikken har vunnet.
Beskriv en algoritme som du kan benytte deg av slik at du i tre av fire spill er sikker på å vinne hvis du får lov til å begynne og som alltid sikrer at du vinner hvis motstanderen minst en gang ikke følger algoritmen.

Oppgave 4.4.2 La $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ være en funksjon, hvor vi antar at vi har en algoritme for å beregne f.
Definer h ved hjelp av følgende algoritme:

- $h(x, y) = x$ hvis $f(x, y) = 0$
- $h(x, y) = h(x, y + 1) + 1$ hvis $f(x, y) \neq 0$.

La $g(x) = h(x, 0)$.
Vis at $g(x)$ er den minste y slik at $f(x, y) = 0$ hvis det finnes en slik y, mens algoritmen for $g(x)$ aldri vil lede til noe svar hvis det ikke finnes noen slik y.
[Hint: h er slik at hvis $h(x, y)$ er definert, vil $f(x, y + h(x, y)) = 0$.]

4.5 Noen relasjonstyper

I dette avsnittet skal vi se på noen av de vanlige fremmedordene som brukes i tilknytning til binære relasjoner, og på noen av de viktigste relasjonstypene.

Definisjon 4.10 La R være en relasjon på en mengde A. Vi sier at R er

1. *refleksiv* hvis aRa for alle $a \in A$
2. *irrefleksiv* hvis ikke aRa for noen $a \in A$
3. *symmetrisk* hvis aRb medfører bRa for alle a og b i A
4. *antisymmetrisk* hvis aRb og bRa medfører at $a = b$ for alle a og b i A
5. *transitiv* hvis aRb og bRc medfører at aRc for alle a, b og c i A
6. *total* hvis vi for alle a og b i A har at aRb, bRa eller $a = b$

Alt i alt skulle vi ha 64 mulige kombinasjoner av disse egenskapene. Vi ser imidlertid at hvis R er både refleksiv og irrefleksiv vil $A = \emptyset$ og hvis R er både symmetrisk og antisymmetrisk kan A ha høyst ett element. Det er også anndre kombinasjoner som ikke vil oppstå eller som bare oppstår i trivielle og uinteressante tilfeller. Vi skal først se på noen lett gjenkjennelige relasjoner og se på hvilke egenskaper de har. Deretter vil vi definere noen klasser av relasjoner som er av spesiell interesse i matematikken og i den teoretiske informatikken.

Eksempler 4.17 a) La $A = \mathbb{N}$ og $R =<$. Da er R irrefleksiv, antisymmetrisk, transitiv og total. Dette vil også være tilfelle i mange andre situasjoner hvor vi snakker om 'mindre enn'.

b) La $A = \mathbb{N}$ og $R =\leq$. Da er R refleksiv, antisymmetrisk, transitiv og total. Dette vil også være tilfelle i mange andre situasjoner hvor vi snakker om 'mindre eller lik'.

c) La $A = \mathcal{P}(\{1, 2, 3\})$ og $R =\subseteq$. Da er R refleksiv, antisymmetrisk og transitiv. R vil ikke være total, for det finnes par av mengder slik at ingen av dem er inneholdt i den andre, eksempelvis $\{1, 2\}$ og $\{2, 3\}$.

d) La $A = \mathbb{Z}$ og aRb hvis vi får samme rest når vi deler a og b på 17. R er refleksiv, symmetrisk og transitiv. Dette er en relasjonstype vi ofte får når vi vil uttrykke at elementene deler noen egenskaper.

e) La $A = \mathbb{R}^2$ og la $(x, y)R(u, v)$ hvis $(x - u)^2 + (y - v)^2 < 1$, det vil si hvis avstanden mellom punktene er mindre enn 1.
R vil være refleksiv, symmetrisk men ikke transitiv.

f) La $A = \mathbb{R}^2$ og la $(x, y)R(u, v)$ hvis $(x - u)^2 + (y - v)^2 = 1$. Da er R symmetrisk og irrefleksiv.

Vi skal nå se på noen relasjonstyper.

Definisjon 4.11 La R være en relasjon på en mengde A.

a) Vi kaller R en *ordning* om R er transitiv, refleksiv og antisymmetrisk.

b) Vi kaller R en *strikt ordning* hvis R er transitiv og irrefleksiv.

c) En ordning eller strikt ordning kalles *total* hvis den er total som relasjon.

Noen ganger kan vi bruke betegnelsen *partiell ordning* hvis vi vil understreke at den ikke trenger å være total.

Lemma 4.1 *En strikt ordning er også antisymmetrisk.*

Bevis
Anta at R er en strikt ordning, og la aRb og bRa. Ved transitivitet vil aRa, noe som er i konflikt med at R er irrefleksiv. Vi kan derfor ikke ha at aRb og bRa samtidig.

I kapittel 5 vil vi se på bevisformer generelt, og da vil vi gjenkjenne dette som et kontrapositivt bevis.
I avsnittet om utfordringer skal vi se på en situasjon hvor det er naturlig å utvide en relasjon til en som er refleksiv og transitiv.

Definisjon 4.12 La R være en binær relasjon på en mengde A.
Vi definerer den *refleksive og transitive tillukningen* R^* av R som følger:
aR^*b hvis $a = b$ eller hvis det finnes $c_1, \ldots, c_n$ i A slik at $a = c_1$, $b = c_n$ og $c_i R c_{i+1}$ for alle $i < n$.

Vi har at aR^*b hvis $a = b$ eller hvis vi kan komme fra a til b ved endelig mange skritt slik at vi beveger oss langs relasjonen R i hvert skritt.

Lemma 4.2 *a) Hvis R er en binær relasjon på en mengde A, er R^* en refleksiv og transitiv relasjon på A.*

b) Hvis $R \subseteq S$ er binære relasjoner på en mengde A og S er refleksiv og transitiv, så vil $\mathbb{R}^* \subseteq S$.

Bevis
For å gi et formelt bevis trenger vi å ha lest Kapittel 7 om induksjonsbevis. Her vil vi basere oss på den intuisjonen som ligger til grunn for induksjonsbevis.

a) R^* er refleksiv fordi vi pr. definisjon har at aR^*a for alle $a \in A$.
Hvis aR^*b og bR^*c har vi at $a = b$, $b = c$ eller at vi kan komme fra a til b og videre fra b til c ved endelig mange R-hopp. I de to første tilfellene har vi opplagt aR^*c, og i det siste tilfellet setter vi sammen de to seriene av endelig mange R-hopp fra a til b og fra b til c til en endelig serie R-hopp fra a til c.

b) Hvis aR^*b fordi $a = b$ vil aSb fordi S er refleksiv.
Hvis aR^*b på grunn av rekken $c_1, \ldots, c_n$ av R-hopp, har vi at c_iSc_{i+1} for alle $i < n$ fordi $R \subseteq S$. Ettersom S er transitiv vil c_iSc_j hver gang $1 \leq i < j \leq n$ (det er her vi egentlig trenger induksjon) så vi har spesielt $a = c_1Sc_n = b$.

Eksempler 4.18 Vi skal se på følgende eksempler på den refleksive, transitive tillukningen:

a) La $A = \mathcal{P}(\mathbb{N})$ og la XRY hvis $X \subset Y$ og Y har ett element mer enn X.
Da vil XR^*Y hvis $X \subseteq Y$ og $Y \setminus X$ er endelig.

b) La A være en gjeng soldater stilt opp i en rad og la aRb hvis b står bak a i en armlengdes avstand og uten noen mellom dem. Da vil aR^*b hvis b ikke står foran a og alle soldatene mellom er stilt opp slik at alle har armlengdes avstand til soldaten foran.

c) La A være en samling dominobrikker som er stilt på høykant. Vi er interesserte i å vite hvilke brikker som vil falle dersom vi dytter på en av dem. Hvis vi lar aRb bety at b blir dyttet over ende i det a faller, vil R^* være den relasjonen vi egentlig er interessert i.

Definisjon 4.13 La R være en relasjon på en mengde A. Vi kaller R en *ekvivalensrelasjon* hvis R er refleksiv, symmetrisk og transitiv.

Vi så på et eksempel på en ekvivalensrelasjon, nemlig at vi får samme rest når vi deler på 17. Det er ikke noe spesielt med tallet 17, så vi har allerede uendelig mange eksempler her. La oss se på noen andre

Eksempler 4.19 a) La A være mengden av sammensatte utsagn i utsagnsvariablene P_1, P_2, P_3. Vi lar $\phi R\psi$ hvis $\phi \Leftrightarrow \psi$, det vil si hvis $\phi \leftrightarrow \psi$ er en tautologi.
Da er R en ekvivalensrelasjon.

b) En *vektor* er et par (x, y) av punkter i planet eller rommet. x kalles *roten* til vektoren og y kalles *spissen* til vektoren.
Det er vanlig å si at to vektorer er like hvis de har samme lengde og retning. Det uttrykker vi ved

$$(x, y)R(u, v) \Leftrightarrow y - x = v - u.$$

Dette er en ekvivalensrelasjon.

c) La A og B være mengder, $f : A \to B$ være surjektiv.
Da kan vi definere R på A ved $aRb \Leftrightarrow f(a) = f(b)$. Da er R en ekvivalensrelasjon.

Vi skal nå se at det siste eksemplet egentlig dekker alle ekvivalensrelasjoner. Hvis vi derfor tenker oss at f er en funksjon som avbilder a på mengden av egenskaper vi synes er interessante for øyeblikket, vil den tilhørende ekvivalensrelasjonen uttrykke at to objekter i A har de samme interessante egenskapene.

Definisjon 4.14 La R være en ekvivalensrelasjon på en mengde A og la $a \in A$. Vi lar *ekvivalensklassen* til a være

$$[a] = \{b \in A \mid aRb\}.$$

Teorem 4.2 *La R og A være som over.*

a) For alle $a \in A$ vil $a \in [a]$.

b) Hvis aRb vil $[a] = [b]$.

c) Hvis $\neg(aRb)$ vil $[a] \cap [b] = \emptyset$

Dette teoremet sier at hvis R er en ekvivalensrelasjon på A, så kan vi dele A opp i disjunkte klasser av parvis ekvivalente elementer.

Bevis
Siden aRa vil $a \in [a]$. Dette viser a).
La aRb. Siden vi da også har at bRa er det nok å vise at $[b] \subseteq [a]$ for å vise b). Så la $c \in [b]$. Da vil $aRb \wedge bRc$ så aRc. Det betyr at $c \in [a]$.
Til sist, anta at $c \in [a] \cap [b]$. Da vil $aRc \wedge bRc$. Ved symmetri og transitivitet for R følger det at aRb. Snur vi dette argumentet på hodet, får vi at $\neg(aRb) \Rightarrow [a] \cap [b] = \emptyset$. Dette viser c), og teoremet er bevist.

Oppgaver til avsnitt 4.5

Oppgave 4.5.1 La R være en relasjon på en mengde A slik at R er antisymmetrisk og irrefleksiv. Forklar hvorfor vi ikke kan finne a og b i A slik at aRb og bRa.

Oppgave 4.5.2 La $A = \{1, 2, 3, 4, 5\}$ og la R være relasjonen som består av tallparene $(1, 3)$, $(3, 2)$, $(3, 5)$ og $(5, 4)$.
Bestem hvilke av de 25 parene i A^2 som er i R^*.

Oppgave 4.5.3 La R være ekvivalenrelasjonen på $\mathbb{Z}$ definert i Eksempel 4.17 d).

a) Vis at om aRb og cRd så vil $(a + c)R(b + d)$.

b) Vi definerer addisjon mellom ekvivalensklasser ved $[a] + [b] = [a + b]$. Bruk a) til å forklare hvorfor dette er en lovlig definisjon.

c) Forklar hvorfor vi kan gi en tilsvarende definisjon av produktet av to ekvivalensklasser.

d) Vis at $[4] \cdot [11] = [1]$.

Ved å bruke at 17 er et primtall er det mulig å vise at hvis $0 < a < 17$ så finnes det et tall b slik at $[a] \cdot [b] = [1]$. Ta utfordringen og prøv å vis dette. Stikkord: Euklids algoritme. 1 er største felles faktor til a og 17. Regn bakover.

Oppgave 4.5.4 Vi definerer en relasjon R på $\mathbb{R}^2$ ved $(x, y)R(u, v)$ hvis $x + y = u + v$.
Vis at R er en ekvivalensrelasjon.
Beskriv ekvivalensklassene til R.
Finn en naturlig funksjon f som avbilder $\mathbb{R}$ injektivt og surjektivt på mengden av ekvivalensklasser.

Oppgave 4.5.5 [u] La R være en binær relasjon på en mengde A. Vis at

$$(R \cup R^{-1})^*$$

er en ekvivalensrelasjon og at den er den minste ekvivalensrelasjonen som utvider R.

4.6 Utfordringer

En modell for programmering.

Vi skal se på et veldig enkelt programmeringsspråk for regning med ikke-negative hele tall.
I en viss forstand snakker vi om programmering for regning med tallerkenstabler, hvor vi kan legge til en tallerken i en stabel, fjerne en tallerken fra en ikke-tom stabel og bestemme oss for hva vi vil gjøre avhengig av om en stabel er tom eller ikke.

Vi vil ha et programmeringsspråk med variable x_1, x_2, ... og uttrykk for alle tall $a \in \mathbb{N}_0$.
Et *program* skal være en instruksjon for hvordan vi kan endre verdiene på variablene. Grunnprogrammene vil være

- $x := a$ hvor $a \in \mathbb{N}_0$
- $x := x + 1$ hvor vi øker verdien av variabelen x med 1.
- $x := x \dot{-} 1$ hvor vi trekker fra 1 hvis mulig, men ellers lar verdien fortsatt være 0.
- Hvis P og Q er programmer lar vi $P; Q$ være et program. Ideen er at vi først kjører P og så kjører Q.

- Hvis P og Q er programmer og x er en variabel, er

  $$if\ x > 0\ then\ P\ else\ Q\ fi$$

  et program.
  Intuisjonen her er at dersom verdien på variabelen x er positiv, så skal vi bruke programbiten P, ellers skal vi bruke programbiten Q.

- Hvis x er en variabel og P er et program, er

  $$while\ x > 0\ do\ P\ od$$

  et program.
  Intuisjonen er at vi skal gjenta P, som vil endre verdiene på de variable, inntil variabelen x får verdien 0, og så er vi ferdige.

Det er to måter vi kan gi en presis matematisk mening til hva et slikt program 'gjør'. La oss for enkelthets skyld anta at vi bare har tre varible. En *valuasjon* vil da være et sett verdier $v(x_1)$, $v(x_2)$ og $v(x_3)$ på de tre variablene. En *situasjon* vil bestå av et par (v, P) hvor v er en valuasjon og P er et program. Vi lar $\emptyset$ betegne det tomme programmet, som vi tillater oss å regne som et program. En situasjon på formen $(v, \emptyset)$ kaller vi da et *svar*, ettersom det er den type situasjon vi skal stå igjen med når programmet er ferdig kjørt.
Vi definerer relasjonen $\vdash$ mellom situasjoner som følger:

- $(v, x_i := a; P) \vdash (v', P)$ hvor $v'(x_i) = a$ og $v'(x_j) = v(x_j)$ for $j \neq i$.
- $(v, x_i := x_i{+}1; P) \vdash (v', P)$ hvor v' kommer fra v ved å la $v'(x_i) = v(x_i){+}1$, men ellers ikke gjøre noen endringer.
- $(v, x_i := x_i \dot{-} 1; P) \vdash (v', P)$ hvor v' kommer fra v ved å redusere $v'(x_i)$ med 1 om mulig, men ellers ikke gjøre noen endringer.
- $(v,\ if\ x_i > 0\ do\ \mathrm{P}\ else\ Q\ fi\ ; R) \vdash (v, P; R)$ om $v(x_i) > 0$
- $(v,\ if\ x_i > 0\ do\ \mathrm{P}\ else\ Q\ fi\ ; R) \vdash (v, Q; R)$ om $v(x_i) = 0$
- $(v,\ while\ x_i > 0\ do\ P\ od\ ; Q) \vdash (v, Q)$ om $v(x_i) = 0$.
- $(v,\ while\ x_i > 0\ do\ P\ od\ ; Q) \vdash (v, P;\ while\ x_i >\ do\ P\ od\ ; Q)$ ellers.

Hvis vi i tillegg setter $\emptyset; P = P$, så får vi her regler for hvordan vi fra en valuasjon og et program får en ny valuasjon og et restprogram etter ett regnetrinn.

Dette kalles den *operasjonelle tolkningen* av programmeringsspråket, den forteller oss hvordan 'regningen' foregår skritt for skritt. Hvis vi bruker den refleksive og trasitive tillukningen $\vdash^*$ får vi en presis definisjon av hva vi mener med at vi vil komme fra en situasjon (u, P) til en annen situasjon (v, Q) ved å regne ingen, ett eller flere skritt.

Dette kan igjen brukes til å definere den *denotasjonelle tolkningen*. Hvis u og v er valuasjoner sier vi at

$$v\langle P\rangle u$$

hvis $(v, P) \vdash^* (u, \emptyset)$.
På denne måten bruker vi relasjoner og begreper knyttet til relasjoner til å gi en matematisk definisjon av hvilken funksjon et program definerer.

Bemerkning 4.1 Det er svært bevisst at dette avsnittet står under overskriften *Utfordringer*. Det er ingen grunn til fortvilelse hvis man synes at dette var vanskelig eller sågar umulig å forstå. En lærebokfremstilling av dette stoffet ville nok krevd en 4-5 sider med forklarende tekst. Hensikten med å ta med dette eksemplet er å illustrere at den typen matematikk vi tar opp i EVU 6 faktisk blir brukt for å legge det teoretiske grunnlaget for informasjonsteknologien.
Hvis man arbeider med mer kompliserte programmeringsspråk, så vil en matematisk behandling av hva disse programmene virkelig betyr innebære en operasjonell og en denotasjonell tolkning som minner mye om det vi har gjort her.

Vi sier at en funksjon $f : \mathbb{N}_0 \to \mathbb{N}_0$ er *beregnbar* hvis det finnes et program P slik at vi for alle a har at $v_a\langle P\rangle v_{f(a)}$ hvor $v_b(x_1) = b$ og $v_b(x_j) = 0$ når $i \neq j$.
Dette virker som en beskjeden definisjon av når en funksjon er beregnbar, men faktum er at ingen har greid å komme opp med en funksjon som er beregnbar i en eller annen form, men som ikke faller inn under vår definisjon.

Kvotienter

En viktig bruk av ekvivalensrelasjoner og ekvivalensklasser er at hvis vi tar utgangspunkt i en matematisk modell med en ekvivalensrelasjon, så vil vi ofte få en nedarvet struktur på mengden av ekvivalensklasser. I oppgave 4.5.3 så vi et eksempel på dette. Der definerte vi addisjon og multiplikasjon mellom ekvivalensklasser av hele tall. Noe av poenget med den konstruksjonen er at vi får en rikere matematisk modell, ved at vi også kan utføre divisjon mellom visse ekvivalensklasser.

Her skal vi se hvordan vi kan konstruere en partiell ordning fra en binær relasjon som er transitiv og refleksiv:
La R være en binær relasjon på en mengde A slik at R er refleksiv og transitiv.
La $\hat{R}$ være definert ved

$$a\hat{R}b \Leftrightarrow aRb \wedge bRa.$$

Vi skal se at $\hat{R}$ er en ekvivalensrelasjon. Det er tre egenskaper vi må vise:

1. $\hat{R}$ er refleksiv:
   Dette følger direkte ved at R er refleksiv siden vi for alle $a \in A$ har at $aRa \wedge aRa$.

2. $\hat{R}$ er symmetrisk:
   Hvis $a\hat{R}b$ har vi $aRb \wedge bRa$, så vi har at $bRa \wedge aRb$, hvilket betyr at $b\hat{R}a$.

3. $\hat{R}$ er transitiv:
   Anta $a\hat{R}b \wedge b\hat{R}c$. Da har vi på den ene siden at $aRb \wedge bRc$, så aRc siden R er transitiv.
   På den andre siden har vi at $cRb \wedge bRa$, så cRa.
   Til sammen gir dette oss at $a\hat{R}c$.

La $B = A/\hat{R}$ være mengden ac $\hat{R}$-ekvivalensklasser $[a]$. Vi definerer $S = R/\hat{R}$ på B ved

$$[a]S[b] \Leftrightarrow aRb.$$

Vi må strengt tatt vise at hvis $[a]S[b]$, $[a] = [c]$ og $[b] = [d]$, så vil $[c]S[d]$.
Dette er en konsekvens av transitiviteten til R, og overlates leseren.
S er antisymmetrisk, for hvis $[a]S[b]$ og $[b]S[a]$, vil aRb og bRa, så $a\hat{R}b$, dvs $[a] = [b]$.
Det at S er transitiv følger fra definisjonen og fra at R er transitiv, så S er en partiell ordning på B.

Det vi har gjort er å slå samen par av objekter i A hvis de utgjør et moteksempel til antisymmetri .
La oss knytte denne konstruksjonen til studiet av programmer, og la oss bruke det eksemplet vi så på i forrige avsnitt. Et problem med å forstå det eksemplet er at det ikke er hver gang at et gitt program med en gitt valuasjon som utgangspunkt gir oss et svar, noen av disse programmene vil 'regne' uendelig lenge uten å komme til noe svar.Vi kan derfor ikke bruke mengden av funksjoner fra $\mathbb{N}_0$ til $\mathbb{N}_0$ som tolkningsområde for slike programmer.

Definisjon 4.15 La P og Q være programmer. Vi sier at $P \preceq Q$ hvis vi for alle valuasjoner u og v har at

$$u\langle P\rangle v \Rightarrow u\langle Q\rangle v,$$

det vil si at hver gang P regner ut v fra u så vil Q gjøre det samme.

Hvis vi lar $\equiv \; = \; \hat{\preceq}$ blir $\equiv$ en ekvivalensrelasjon på mengden av programmer, og ekvivalensklassene blir ordnet va $\preceq \; / \equiv$ slik at det svarer til at et program er mindre nyttig enn et annet.

Dette er bare et forsøk på å gi et eksempel på hvordan elementær mengdelære og læren om relasjoner brukes til å lage matematiske modeller med relevans for informatikk-faget. Skulle vi gjort dette på en seriøs måte ville det krevd adskillig flere detaljer og også større generalitet.

4.7 Blandede oppgaver

Oppgave 4.7.1 La $A = \{1, 2, 3\}$ og $B = \{1, 2, 3, 4\}$. La R, S og T være de tre relasjonene fra A til B:

$$R = \{(1, 3), (2, 1), (1, 2), (3, 4)\}$$

$S = \{(2,2),(3,1),(1,4)\}$

$T = \{(1,2),(1,3),(3,4),(2,4)\}$

a) Bestem hvilke av disse relasjonene som er funksjoner.

b) Finn R^{-1}, S^{-1} og T^{-1} og bestem hvilke av disse relasjonene som er funksjoner.

c) Finnes det en relasjon U fra A til B slik at både U og U^{-1} er funksjoner? Begrunn svaret.

Oppgave 4.7.2 La A være en mengde, R en partiell ordning på A. La B være mengden av alle endelige sekvenser $a_1, \ldots, a_n$ fra A.
Vi definerer en relasjon S på B som følger:
$(a_1, \ldots, a_n)S(b_1, \ldots, b_m)$ hvis en av to holder:

1. $n \leq m$ og $a_i = b_i$ for alle $i \leq n$.
2. Det finnes en $i \leq min\{n, m\}$ slik at $a_i R b_i$, slik at $a_i \neq b_i$ og slik at $a_j = b_j$ når $j < i$.

a) Vis at S er en partiell ordning.

b) Vis at hvis R er en total ordning, så er også S total.

c) Ordningen S kalles den *leksikografiske ordningen*. Diskuter hvorfor dette er en naturlig betegnelse, ut fra hvordan man ordner artikler i et leksikon, navn i en telefonkatalog m.m.

Oppgave 4.7.3 I avsnitt 3.3 så vi på digitale kretser med forsinkelse. En slik krets kan ta i mot en eller flere datastrømmer i form av strømmer av sannhetsverdier, og den kan sende ut en eller flere datastrømmer. I den forstand kan en digital krets med forsinkelse tolkes som en funksjon som avbilder n datastrømmer på m datastrømmer.
Hvis vi tolker en slik krets som en funksjon uten å ta hensyn til hvordan vi kommer fra inngangs-strømmene til utgangs-strømmene, snakker informatikerne om den *denotasjonelle tolkningen*. Hvis vi derimot finner hvordan algoritmen som ligger under utføres trinn for trinn, snakker vi om den *operasjonelle tolkningen*. Ta utgangspunkt i de eksemplene vi hadde i avsnitt 3.3 med tilhørende oppgaver, og formuler en algoritme som beregner den funksjonen som de digitale kretsene definerer.

Kapittel 5

Bevisformer

5.1 Direkte bevis

Et *bevis* er en måte å kommunisere en forståelse for hvorfor en matematisk påstand er riktig fra en person til en annen, eventuelt til mange andre. Når vi skal skrive et bevis, er dette det viktigste aspektet, vi må argumentere for hvorfor det vi påstår stemmer på en slik måte at leseren/tilhøreren forstår at det må være riktig. Et bevis vil gjerne bestå av en serie enkeltsluttninger og utregninger, og en måte å forstå et bevis på er å forstå at de enkelte delene av beviset er riktige. En dypere måte å forstå et bevis på er å forstå hvordan den som har utformet beviset har tenkt, hvorfor enkeltutledningene og de enkelte regnestykkene er utført i den rekkefølgen og på den måten som det er gjort.
Alle som skal lære å skrive godt, må lese mye. Dette gjelder norsk, det gjelder fremmedspråk og det gjelder også matematikk. Alle som skal lære seg det håndverket det er å skrive leselige bevis, bør lese en rekke eksempler. Det vi skal gjøre i dette heftet er å gi noen eksempler på bevis. Vi kan dele argumenter opp i visse hovedtyper, og vi skal illustrere disse hovedtypene gjennom noen kommenterte eksempler. En viktig klasse av bevis, induksjonsbevisene, er tilegnet et helt kapittel, Kapittel 7.
I dette avsnittet skal vi konsentrere oss om det som kalles *direkte bevis*. I mange av eksemplene vil vi bevise påstander som leseren trolig ikke synes trenger ytterligere bevis. Hensikten med dette er å bli kjent med nytt farvann ved å starte med det kjente.

Eksempel 5.1 Vi skal vise at differensen mellom to kvadrattall som kommer etter hverandre i tallrekken er et oddetall.
Vi kan formulere dette mer matematisk som en setning:

$$\text{For alle tall } n \text{ er } (n+1)^2 - n^2 \text{ et oddetall.}$$

Bevis
Ved 1. kvadratsetning er $(n+1)^2 = n^2+2n+1$, så $(n+1)^2-n^2 = n^2+2n+1-n^2 =$

$2n + 1$.
Siden $2n + 1$ alltid er et oddetall er setningen vist.

I dette eksemplet formulerte vi først det vi skulle vise i en matematisk språkdrakt, deretter regnet vi litt på den differensen vi skulle bevise var et oddetall, og endte opp med at det var akkurat et oddetall det var.
Hvis vi analyserer beviset litt nærmere ser vi at alle oddetallene kan fremkomme som en slik differens, for å få $2n + 1$ som verdi, kan vi velge kvadrattallene $(n + 1)^2$ og n^2.

Når vi først har funnet et bevis, kan vi undersøke om samme metode kan gi oss mere innsikt. Det vil ofte være tilfelle, men kan kreve ekstra innsats. La oss se om vi kan bruke samme resonnement til å si noe om differensen mellom kubikktall.

Eksempel 5.2 For alle naturlige tall n er $(n + 1)^3 - n^3$ et oddetall.

Bevis
Vi har at $(n + 1)^3 = n^3 + 3n^2 + 3n + 1$, så

$$(n + 1)^3 - n^3 = 3n^2 + 3n + 1.$$

Hvis n er et partall, er $3n^2 + 3n$ også et partall, så $(n + 1)^3 - n^3$ er et oddetall.
Hvis n er et oddetall, er både $3n^2$ og $3n$ oddetall, så $3n^2 + 3n$ er fortsatt et partall, og også i dette tilfellet er $(n + 1)^3 - n^3$ et oddetall.
Dermed er påstanden bevist.

Som vi ser, brukte vi akkurat samme resonement i starten av disse to bevisene. I det andre beviset måtte vi imidlertid etterhvert dele argumentet opp i to tilfeller, ett for at n er et oddetall og ett for at n er et partall. Siden dette dekker alle mulighetene, er beviset fullstendig.
La oss gi et tredje eksempel på et bevis hvor vi må dele argumentet opp i tilfeller:

Eksempel 5.3 La oss bevise følgende påstand:

Hvis n er et helt tall, kan $n^2 - n$ deles på 6 eller $n^2 + n$ kan deles på 6.

Bevis
Vi har at $n^2 - n = (n - 1)n$ og at $n^2 + n = n(n + 1)$.
Nøyaktig ett av tallene $n - 1$, n eller $n + 1$ kan deles på 3.
Hvis $n - 1$ kan deles på 3 er ett av tallene $(n - 1)$ eller n et partall, og da er $(n - 1)n$ delelig med 6.
Hvis $n + 1$ er delelig med 3 er ett av tallene n eller $n + 1$ partall, så $n(n + 1)$ er delelig med 6.
Hvis n er delelig med 3 ser vi ved samme argument at både $(n-1)n$ og $n(n+1)$ er delelige med 6.
Tilsammen beviser dette påstanden.

Dette eksemplet viser hvordan man enkelte ganger må dele et argument opp i tilfeller. Det viser imidlertid også at man av og til må få anta at leseren henger

med i noen av svingene, uten at alle detaljene som ligger til grunn for beviset blir tatt med. Vi har for eksempel tatt det for gitt at leseren er med på at 6 er faktor i $(n-1)n$ når $n-1$ kan deles på 3 og $n-1$ eller n er et partall.
Vi har heller ikke minnet om det er fordi $n^2 - n = (n-1)n$ at vi har vist påstanden når vi i realiteten viser at $(n-1)n$ eller $n(n+1)$ kan deles på 6.
Hvor mange detaljer man tar med er en vurderingsak, og vil være avhengig av målgruppen. Matematisk erkjennelse er bygget opp over mange år og formidling av ny kunnskap forutseter at man har skaffet seg kunnskap om det de nye bevisene bygger på. I dette heftet er det meningen at leseren skal være fortrolig med grunnskolens pensum, og da er det kunnskap fra det pensumet, samt det som har vært behandlet tidligere i heftet argumentasjonen kan bygge på.

Enkelte ganger kan bevis være ufullstendige fordi den som skriver beviset har latt være å ta hensyn til vansker som leseren burde ha, og gjennom det muligens forleder leseren til å tro på at en påstand er riktig på sviktende grunnlag.

Eksempel 5.4 La P være en mangekant med n kanter. Da er summen av vinklene i P lik $(n-2) \cdot 180°$.

Bevis
La $Q_1, \ldots, Q_n$ være hjørnene i P og la Q være et punkt inne i P.
Hvis vi trekker linjen fra Q til hver av Q_i'ene deler vi opp P i n trekanter. Vinkelsummen i hver av disse er 180°. Når vi gjør det, får vi med oss alle vinklene i P samt en hel omdreining på 360° rundt Q.
Det er dette siste som er for mye, så resten, $(n-2)\cdot 180°$ blir summen av vinklene i P.

Problemet med dette beviset er ikke at det er litt knapt skrevet, men at det finnes mangekanter som beviset ikke holder for. Vi forutsetter for eksempel at vi kan finne et punkt Q inne i P slik at vi får delt P opp i trekanter ved å trekke linjene fra Q til hvert av hjørnene. Hvis n er passe stor er dette ikke alltid mulig. Tenk for eksempel på en barneoppgave hvor en figur skal tegnes ved å tegne streker mellom nummererte punkter på et ark. Hvis resultatet for eksempel blir en tilnærmet månesigd, tegner barnet en mangekant hvor forutsetningen i beviset ikke gjelder. Idéen bak beviset er forsåvidt god nok, men gjennomføringen krever at man tar hensyn til at det punktet vi velger kan ligge utenfor P, og da må man passe på fortegn på vinklene og liknende.
Det kan også være en svakhet at vi ikke har gitt en tilstrekkelig presis definisjon av hva vi mener med en mangekant, og hva vi mener med vinkelsummen i en mangekant. Hvor viktig dette er avhenger av hva vi forventer av leseren når det gjelder å kjenne til de begrepene vi arbeider med.

Det er viktig at man ikke underslår vanskeligheter som leseren måtte kunne få. Det eksemplet vi ga var et mangelfullt bevis, og mangelfulle bevis er ikke bra.

Vi har gitt noen enkle eksempler fra tallregning og ett eksempel fra geometrien. Vi skal nå gi noen eksempler på direkte bevis uten kommentarer. Vi skal bevise satser med relevans for andre deler av dette heftet. Leseren bør legge vekt på å se på hvordan argumentasjonen er bygd opp, og bør reflektere over hvor det

kunne vært ønskelig med flere detaljer. Et bevis kan godt stille krav til leseren, men bør ikke inneholde tankesprang som leseren ikke har forutsetninger for å følge med på.

Eksempel 5.5 La A være en mengde slik at $A \cup B = B$ for alle mengder B. Da er $A = \emptyset$.

Bevis
Vi har spesielt at $A \cup \emptyset = \emptyset$, samtidig som vi vet at $A \cup \emptyset = A$. det følger at $A = \emptyset$.

Eksempel 5.6 La R være en binær relasjon på en mengde A slik at R er symmetrisk og transitiv. La B være mengden av de $a \in A$ sik at aRb for minst en b. Da er R begrenset til B en ekvivalensrelasjon.

Bevis
Siden R er transitiv og symmetrisk, er også R begrenset til B transitiv og symmetrisk. Derfor er det nok å vise at R begrenset til B er refleksiv.
Så la $a \in B$. Pr. definisjon av B finnes det en $b \in A$ slik at aRb. Siden R er symmetrisk vil også bRa, og siden R er transitiv har vi at $aRb \wedge bRa \Rightarrow aRa$. Dette viser at R er refleksiv på B.

Eksempel 5.7 La R være en relasjon på en mengde A. La $a\tilde{R}b$ hvis aSb hver gang S er en ekvivalensrelasjon på A slik at $R \subseteq S$.
Da er $\tilde{R}$ en ekvivalensrelasjon slik at $R \subseteq \tilde{R}$.

Det følger at det finnes en minste ekvivalensrelasjon som utvider R.

Bevis
Vi må bevise fire egenskaper: Refleksivitet, symmetri og transitivitet for $\tilde{R}$ og at $R \subseteq \tilde{R}$.

1. $\tilde{R}$ er refleksiv:
   La $a \in A$. Hvis $R \subseteq S$ og S er en ekvivalensrelasjon, har vi at aSa. Siden S var vilkårlig følger det at $a\tilde{R}a$.

2. $\tilde{R}$ er symmetrisk:
   Anta $a\tilde{R}b$. La S være en ekvivalensrelasjon som utvider A. Fra definisjonen av $\tilde{R}$ følger det at aSb, og siden S er en ekvivalensrelasjon følger det at bSa. Siden S var vilkårlig må $b\tilde{R}a$.

3. $\tilde{R}$ er transitiv:
   Anta $a\tilde{R}b$ og $b\tilde{R}c$. La S være som før. Da vil aSb og bSc, så aSc. Det følger at $a\tilde{R}c$.

4. $R \subseteq \tilde{R}$:
   La aRb. Hvis $R \subseteq S$, vil aSb, og det gjelder spesielt når S i tillegg er en ekvivalensrelasjon. Derfor vil $a\tilde{R}b$.

Alle de eksemplene vi har sett på er eksempler på direkte bevis. Vi har tatt utgangspunkt i det vi vet, og så har vi bevist flere og flere fakta inntil vi har et

fullstendig bevis for det vi ønsker å bevise.

I beviset over beviste vi at det finnes en minste ekvivalensrelasjon som utvider en gitt binær relasjon ved først å definere den, og så bevise at relasjonen vi definerte hadde de ønskede egenskapene. Når vi skal bevise at noe finnes, er det selvfølgelig best med en konstruksjon eller en definisjon, for da får vi mere informasjon. Av og til er det imidlertid slik at vi beviser at det finnes noe med en ønsket egenskap, uten at vi har gitt en fullstendig beskrivelse av objektet. Vi skal se på to eksempler:

Eksempel 5.8 Det finnes uendelig mange primtall.

Bevis
Det vi skal bevise er at det for ethvert tall n finnes et primtall $p > n$. Beviset er klassisk.
La $n \in \mathbb{N}$ være gitt. La $m = n! + 1$ hvor vi minner om at $n! = 1 \cdot 2 \cdot \ldots \cdot n$.
La $p \leq m$ være et primtall slik at m kan deles på p. Hvis $p \leq n$ vil vi få 1 til rest hvis vi deler m på p, så vi må ha at $n < p$.

I dette beviset har vi vist at det finnes et primtall $p > n$, men vi har bare gitt et øvre estimat for hvor lite det neste primtallet kan være, ikke gitt noen direkte definisjon. Dette kaller matematikerne et rent eksistensbevis. Nå er det jo teknisk sett mulig, etter at n er gitt, å teste ut for alle tall q slik at $n < q \leq n! + 1$ om de er primtall eller ikke, og på den måten finne det minste primtallet større enn n. Vi skal gi et annet eksempel på et eksistensbevis i form av et bevis ved tilfeller, hvor vi ikke kan trekke mer informasjon ut av beviset.

Eksempel 5.9 Det finnes to irrasjonale tall a og b slik at $a^b \in \mathbb{Q}$.

Bevis
Vi deler beviset opp i to tilfeller.

Tilfelle 1: $\sqrt{2}^{\sqrt{2}} \in \mathbb{Q}$.
Da lar vi $a = b = \sqrt{2}$ og $a^b \in \mathbb{Q}$.

Tilfelle 2: $\sqrt{2}^{\sqrt{2}} \notin \mathbb{Q}$.
Da lar vi $b = \sqrt{2}$ og $a = \sqrt{2}^{\sqrt{2}}$. Da får vi

$$a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{(\sqrt{2} \cdot \sqrt{2})} = \sqrt{2}^2 = 2 \in \mathbb{Q}.$$

Dette eksemplet blir ofte brukt til å illustrere forskjellen på konstruktive bevis og bevis basert på klassisk logikk. I klassisk logikk kan vi gjøre uhemmet bruk av antagelsen at enten gjelder en påstand P eller så gjelder negasjonen $\neg P$. I konstruktiv matematikk kreves det at vi i tillegg har noe informasjon om hvilken av de to som gjelder, eller i det minste en metode for å avgjøre hvilken av de to som gjelder i en gitt situasjon. Ut fra det beviset vi har sett på kan vi ikke si noe sikkert om hvilket par a, b av irrasjonale tall det er som er slik at $a^b \in \mathbb{Q}$, bare at det finnes et slikt par.

Konstruktiv matematikk har vært sammenlikned med kappgang, det kan være en krevende idrett, men trenger man å rekke toget, så løper man likevel. Fremveksten av datateknologien gjør det imidlertid i mange tilfeller aktuelt å kunne trekke algoritmer ut av bevis, og da gir konstruktive bevis mer nyttig informasjon. Derfor har det i de senere årene vært en dreining fra at noen idealister har jobbet med konstruktiv matematikk fordi bevisene er "riktigere" enn tradisjonelle matematiske bevis til at flere bruker konstruktive metoder fordi de får mer informasjon ut av bevisene sine på den måten. Det må likevel innrømmes at konstruktiv matematikk er nisjepreget og blir drevet av spesielt interesserte.

Det siste eksemplet vi skal gi er et eksempel på en annen type konstruksjon, hvor vi påstår at det finnes et reelt tall med en gitt egenskap ved å bestemme desimalutviklingen bit for bit:

Eksempel 5.10 La $A \subseteq [0,1]$ være uendelig. Da finnes det et reelt tall $x \in [0,1]$ slik at $A \cap (x - \frac{1}{n}, a + \frac{1}{n})$ er uendelig for alle n.
Vi kaller x et *opphopningspunkt* for A.

Bevis
Alle tallene $a \in A$ kan skrives på desimalform som $0, a_1 a_2 \cdots$ hvor hver $a_i \in \{0, 1, \ldots, 9\}$.
La $A_0 = A$. La x_1 være slik at vi for uendelig mange $a \in A$ har at $a_1 = x_1$, og la $A_1 = \{a \in A \mid a_1 = x_1\}$. Det må finnes en slik x_1 fordi vi bare har 10 mulige førstedesimaler og uendelig mange $a \in A$. Derfor må det være uendelig mange $a \in A$ med den samme førstedesimalen.
Nå kan vi fortsette. La x_2 være slik at for uendelig mange $a \in A_1$ så er $a_2 = x_2$, og la $A_2 = \{a \in A_1 \mid a_2 = x_2\}$.
Vi kan finne en slik x_2 av tilsvarende grunn som vi kunne finne x_1.
Slik kan vi fortsette, og vi ser at vi kan finne en $x = 0, x_1 x_2 \cdots$ slik at vi for alle n har uendelig mange $a \in A$ med de samme første n desimalene som x.
Da er x et opphopningspunkt for A.

Dette beviset er ikke konstruktivt i den forstand at vi baserer oss på at om vi deler en uendelig mengde opp i endelig mange deler, så må den ene delen være uendelig, uten at vi kan si hvilken. Det viser seg da også at selv om vi kan skrive et program som lister opp alle tallene i A, så kan vi normalt ikke finne et program som gir oss desimalutviklingen til noen x som er et opphopningspunkt for A. Dette vil det føre alt for langt å utdype nærmere.

Oppgaver til avsnitt 5.1

Oppgave 5.1.1 Gå tilbake til oppgavene 1.1.6 og 1.1.7 og skriv bevis for de påstandene som fremmes der.

Oppgave 5.1.2 Bruk formelen $(n+1)^4 = n^4 + 4n^3 + 6n^2 + 4n + 1$ til å vise at $(n+1)^4 - n^4$ alltid er et oddetall. Trenger du å bruke bevis ved forskjellige tilfeller her?

Leseren begynner muligens å ane et mønster. Hvis leseren er fortrolig med formelen for $(a+b)^k$ uttrykt via binomialkoeffisienter eller ved hjelp av Pascals talltrekant, er det nå en passe tilleggsoppgave å vise at $(n+1)^k - n^k$ alltid er et oddetall når $k \geq 1$.

Oppgave 5.1.3 Vi skal gi et alternativt bevis for at vinkelsummen i en mangekant er bestemt av antall kanter. Bestem om dette beviset også er mangelfullt og gjør i så fall rede for hva forfatteren ikke har tatt høyde for:

Bevis
La P være en mangekant med n kanter, og la Q være et av hjørnene i P. Trekk linjer fra Q til de $n-3$ hjørnene i P som ikke er nabohjørner til Q. Dette deler P opp i $n-2$ trekanter, og vinkelsummen i disse trekantene vil samlet bli summen av vinklene i P. Dette er $(n-2) \cdot 180°$.

Oppgave 5.1.4 Som et eksempel på hvordan vi kan utforme et bevis, ga vi et bevis for at enhver binær relasjon kan utvides til en ekvivalensrelasjon på en minimal måte. Nå skal vi gi et bevis for at enhver partiell ordning kan utvides til en total ordning. Påstanden er faktisk sann, men det beviset vi gir er ikke helt bra.
Diskuter "beviset" og finn ut av hvor feilen(e) er:

Bevis
La R være en partiell ordning på A, og la R^T være relasjonen definert ved

$$aR^Tb \text{ hvis } aSb \text{ hver gang } S \text{ er en total ordning som utvider } R.$$

Bevisene for at R^T er refleksiv og transitiv er som bevisene for at $\tilde{R}$ er refleksiv og transitiv i Eksempel 5.7, så vi konsentrerer oss om totalitet.
La a og b være i A, og la S være en vilkårlig total ordning som utvider R. Da har vi aSb, $a = b$ eller bSa. Hvis $a = b$ er alt i orden. Hvis aSb bruker vi at S var vilkårlig og får at aR^Tb. Hvis bSa bruker vi samme argument og får at bR^Ta.
Derfor er R^T også total.

5.2 Indirekte bevis

Indirekte bevis baserer seg på den utsagnslogiske ekvivalensen

$$A \to B \Leftrightarrow \neg B \to \neg A.$$

Det innebærer at hvis vi ønsker å vise at B er en konsekvens av A, så kan vi like gjerne vise at $\neg A$ er en konsekvens av $\neg B$. En måte å vise $\neg A$ fra $\neg B$ på er å vise at A og $\neg B$ sammen leder til en selvmotsigelse.
Vi skal se på noen eksempler på *indirekte bevis*. Felles for disse bevisene er at vi antar det motsatte av det vi vil bevise, og så arbeider vi oss frem til at det må være noe i veien med forutsetningene. Slike bevis kalles også ofte for *kontrapositive bevis*.
Vi starter med det trolig mest brukte eksemplet av dem alle.

Eksempel 5.11 $\sqrt{2}$ er ikke et rasjonalt tall.

Bevis
Anta at $\sqrt{2} \in \mathbb{Q}$. Da finnes det to tall n og m slik at $\sqrt{2} = \frac{n}{m}$, og vi kan anta at $\frac{n}{m}$ er maksinmalt forkortet, det vil si at n og m ikke har noen felles faktorer.
Ved å kvadrere på begge sider har vi at $\frac{n^2}{m^2} = 2$, eller skrevet på en annen måte, at $n^2 = 2 \cdot m^2$.
Det betyr at n må være et partall, det vil si at $n = 2k$ for en $k \in \mathbb{N}$, og dermed har vi at $4 \cdot k^2 = 2 \cdot m^2$.
Nå kan vi forkorte med 2 og får at $2 \cdot k^2 = m^2$
Men det innebærer at m også er et partall, og det er i konflikt med at n og m ikke har noen felles faktorer. Siden alle tall i $\mathbb{Q}$ kan skrives på en maksimalt forkortet form må feilen ligge i antagelsen om at $\sqrt{2} \in \mathbb{Q}$.

Vi henter vårt neste eksempel fra stoffet i Kapittel 4:

Eksempel 5.12 La A være en endelig mengde, og la R være en ikketom delvis (partiell) ordning på A. Da har A et R-minimalt element a, det vil si at $\neg(bRa)$ for alle $b \in A$.

Bevis
Anta at vi ikke har noe slikt R-minimalt element. La $a \in A$. Siden a ikke er R-minimal, kan vi finne $a_1 \neq a$ slik at $a_1 Ra$. Siden a_1 heller ikke er R-minimal, kan vi finne $a_2 \neq a_1$ slik at $a_2 Ra_1$. Slik kan vi fortsette med å finne stadig R-mindre a_i i det uendelige. Men det at vi kan fortsette i det uendelige er i konflikt med at A er endelig, så derfor er dette umulig i lengden.
Derfor må antagelsen om at A ikke har noe R-minimalt element være feil.

I dette beviset har vi gitt et eksistensbevis i form av et indirekte bevis, vi antok at det ikke fantes noe R-minimalt element og utledet at forutsetningen måtte være feil. Dette er den minst konstruktive måten vi kan bevise at noe finnes på, beviset gir ingen informasjon om hvordan vi kan finne det vi påstår at finnes.

I vårt neste eksempel skal vi gå tilbake til en av utfordringene i mengdelæren, Cantors diagonalargument. I dette tilfelle vil vi bevise at en viss type funksjon ikke kan finnes, ved å utlede en motsigelse fra antagelsen om at den finnes:

Eksempel 5.13 La X være en mengde. Da finnes det ikke noen surjektiv funksjon $f : X \to \mathcal{P}(X)$.
Vi minner om at $\mathcal{P}(X)$ er potensmengden til X, altså mengden av alle delmengder av X.

Bevis
Anta at $f : X \to \mathcal{P}(X)$ er surjektiv. La $A \subseteq X$ være definert ved

$$x \in A \Leftrightarrow x \notin f(x).$$

Siden f er surjektiv, finnes det en $y \in X$ slik at $A = f(y)$.
Hvis $y \in A$ har vi at $y \in f(y)$, så pr. definisjon av A vil $y \notin A$.

Hvis på den annen side $y \notin A$, vil $y \notin f(y)$, så pr. definisjon av A vil $y \in A$.
Vi ser at vi har at

$$y \in A \Leftrightarrow y \notin A.$$

Dette er en umulig situasjon, så antagelsen om at f er surjektiv må være feil.

I vårt neste eksempel skal vi kombinere resultatet fra Eksempel 5.10 med et indirekte bevis, og vise et viktig resultat i reell analyse. Lesere som ikke er helt fortrolig med definisjonen av kontinuitet får stole på intuisjonen sin her.

Eksempel 5.14 La $f : [0, 1] \to \mathbb{R}$ være kontinuerlig. Da er f begrenset.
Det at f er begrenset vil si at det finnes et tall N slik at $f(x) \leq N$ for alle $x \in [0, 1]$.
Det er viktig at vi har et lukket interval her, for $f(x) = \frac{1}{x}$ definert på $(0, 1]$ er både kontinuerlig og ubegrenset.

Bevis
Vi skal gi et indirekte bevis, så vi starter med å anta at f ikke er begrenset. For hvert tall $n \in \mathbb{N}$ kan vi da finne x_n slik at $0 \leq x_n \leq 1$ og $f(x_n) > n$.
La $A = \{x_n \mid n \in \mathbb{N}\}$. Da er A en uendelig mengde. Fra teoremet vist i Eksempel 5.10 følger det at A har et opphopningspunkt $x \in [0, 1]$. La $N > f(x)$. Ut fra konstruksjonen av A og ut fra det at x er et opphopningspunkt for A finnes det $x_n \in A$ vilkårlig nært x slik at $n > N$. Siden f er kontinuerlig finnes det imidlertid en k slik at om $|x-y] < 10^{-k}$ så er $f(y) < N$. Dette gir en motsigelse siden alle x_n slik at x_n og x har de $k+1$ første desimalene felles vil oppfylle at $|x - x_n| < 10^{-k}$, mens det finnes uendelig mange slike x_n i A, og for de fleste av disse vil $f(x_n) > N$.
Antagelsen var at f er ubegrenset, så det må være den som er feil. Dermed er setningen bevist.

Vårt siste eksempel skal være innenfor det vi kaller *spillteori*, og vi skal se på spillet "Bondesjakk".

Definisjon 5.1 En *strategi* for en spiller i et spill er en funksjon som til en stilling hvor spilleren er i trekket gir neste trekk.
En *vinststrategi* er en strategi slik at spilleren vil vinne hvis hun/han følger strategien i alle trekk.

Disse begrepene har mening for spill hvor begge spillerne har full informasjon, som sjakk, bondesjakk, kinasjakk, GO og andre, men ikke for sjansespill som yatzy, monopol og poker.

La oss kalle spillerne i bondesjakk for I og II. I begynner med et kryss på et ruteark, II fortsetter med en ring i en annen rute og slik fortsetter spillerne ved annenhver gang å sette et kryss eller en ring i en ny rute. Den som først får fem av sine tegn på rad, opp-ned, høyre-venstre eller på skrå, har vunnet. Hvis ingen får fem på rad noen gang, fortsetter spillet teoretisk sett i det uendelige, og vi kaller resultatet for uavgjort. I praksis er spillet uavgjort når arket er utskrevet uten at noen har fått fem på rad.

Lemma 5.1 *Spiller II kan ikke ha en vinststrategi*

Bevis
Hvis spiller II har en vinststrategi, kan spiller I late som om han er spiller II ved å tenke seg første trekket til spiller II, deretter følge strategien til II i sine trekk, og hver gang II gjør et trekk som I allerede hadde basert seg på at II hadde gjort, så later I som at II gjør et annet trekk. I tankene bruker derfor I strategien til II, og ender opp med å vinne.

Dette beviset var indirekte. Vi antok at II hadde en vinststrategi, og brukte den til å beskrive en vinststrategi for I. Siden begge spillerne ikke kan ha noen vinststrategi, må antagelsen være feil.
Det beste II kan oppnå er en strategi for å spille uavgjort. Hvis II følger en eventuell uavgjortstrategi kan II håpe på at I gjør en feil, og deretter finne seg en vinststrategi et stykke ute i spillet.

Teorem 5.1 *I bondesjakk vil enten spiller I ha en vinststrategi, eller II ha en uavgjortstrategi.*

Bevis
Hvis I har en vinststrategi, holder teoremet, så anta at I ikke har en vinststrategi. Hvis stillingen er slik at I er i trekket, men fortsatt ikke har noen vinststrategi, må II ha et mottrekk til alle trekk I gjør slik at I heller ikke etter de to trekkene har noen vinststrategi. Hvis ikke ville jo trekket I kunne gjort uten mottrekk kunnet være en del av en vinststrategi. Det betyr at II kan følge følgende plan: Hver gang II er i trekket setter hun/han sin ring slik at I fortsatt ikke har noen vinststrategi fra det punktet i spillet av. Hvis II følger denne planen, vil spillet ende som uavgjort, eller II kan vinne hvis hun/han er heldig. Beviset for dette er igjen indirekte. Hvis I vinner har han/hun en vinststrategi før sitt siste trekk. II spiller slik at I aldri har noen vinststrategi. Altså kan ikke I vinne.

Bemerkning 5.1 Det kan stilles noen kritiske kommentarer til dette beviset. Den viktigste er at vi har snakket om vinststrategier gitt en stilling i spillet, uten at vi har definert hva vi mener med det. Under utskrivningen av beviset antok forfatteren her at leseren henger med, og selv generaliserer begrepet 'vinststrategi' til 'vinststrategi fra en stilling'. Poenget er selvfølgelig at strategien begrenses til spill som har utviklet seg fra den stillingen det er snakk om, og at den bare gir vinst for slike spill.
En annen kritikk av bevisene både for lemmaet og for teoremet kan være at litt for mye er basert på at leseren følger med på de komplekse setningene, og at beviset derfor kan virke noe ustrukturert.

Vi har vist at I har en vinststrategi eller at II har en uavgjortstrategi, men vi vet ikke hvilken av de to mulighetene det er som gjelder. En mulighet som forfatteren synes er sannsynlig er at II har en uavgjort-strategi, men at hver gang vi programmerer en datamaskin til å spille for II, så kan I finne en motstrategi som virker. I spillteori kan vi lage "spill" som minner om bondesjakk, og som er slik at II har en uavgjortstrategi i teorien, men at ingen slik strategi kan

programmeres. Disse eksemplene er kunstige og baserer seg på teori som taes opp i et av hovedfagskursene i matematikk. Det ville vært interessant om et "naturlig" spill som bondesjakk hadde den samme egenskapen.

Oppgaver til avsnitt 5.2

Oppgave 5.2.1 I beviset under Eksempel 5.8 ga vi et direkte bevis for at det finnes uendelig mange primtall. Ved nærmere ettersyn, se om du finner et lite snev av et indirekte bevis i et delargument.

Oppgave 5.2.2 Vi skal gå tilbake til utsagnslogikk, og bevise en påstand om *adekvate bindeord*. Forklar hvor vi bruker indirekte bevis i delargumenter i teksten under. For den som leser matematikk skrevet av forskjellige forfattere, kan det ofte være en utfordring at den ene forfatteren formulerer seg litt anderledes enn den man er kjent med. I denne oppgaven har vi brukt skrivemåter som ved første øyekast kan virke ukjente, men som det skulle være mulig å forstå. Prøv å skriv om teksen slik at den blir mer i tråd med skrivemåten i kapittel 2. Oppgaven slutter etter at beviset for Teorem 5.2 er avsluttet.

Definisjon 5.2 Et binært bindeord er en funksjon $F : \{\top, \bot\}^2 \to \{\top, \bot\}$

Eksempelvis ser vi at $\wedge$ er et binært bindeord, når vi definerer $\wedge$ som

- $\wedge(\top, \top) = \top$
- $\wedge(\top, \bot) = \wedge(\bot, \top) = \wedge(\bot, \bot) = \bot$

Definisjon 5.3 Et binært bindeord er *adekvat* hvis vi kan definere alle sannhetsverdifunksjoner ved hjelp av dette bindeordet alene.

Teorem 5.2 *Det finnes nøyaktig to adekvate binære bindeord.*

Bevis
La F være et binært adekvat bindeord. Hvis $F(\top, \top) = \top$ eller $F(\bot, \bot) = \bot$ vil det være umulig å definere $\neg$ ved F alene, så vi må ha at $F(\top, \top) = \bot$ og at $F(\bot, \bot) = \top$.
Det etterlater oss med fire muligheter. Hvis $F(\top, \bot) = \top$ og $F(\bot, \top) = \bot$ vil vi ha at $F(A, B) \Leftrightarrow A$, og hvis $F(\top, \bot) = \bot$ og $\bot, \top) = \top$ vil vi ha at $F(A, B) \Leftrightarrow B$. I ingen av disse tilfellene vil F kunne være adekvat fordi ethvert uttrykk vil bli ekvivalent til første, henholdsvis siste utsagnsvariabel.
Det gjenstår to muligheter som vi skriver på tabellform:

A	B	$A \downarrow B$
$\top$	$\top$	$\bot$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\top$

A	B	$A \mid B$
$\top$	$\top$	$\bot$
$\top$	$\bot$	$\top$
$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\top$

Vi ser at $A \downarrow A \Leftrightarrow A \mid A \Leftrightarrow \neg A$. Videre har vi at

$$(A \downarrow B) \downarrow (A \downarrow B) \Leftrightarrow A \vee B$$

og

$$(A \mid B) \mid (A \mid B) \Leftrightarrow A \wedge B.$$

Ved DeMorgans lover ser vi at $\wedge$ kan uttrykkes ved $\neg$ og $\vee$ og at $\vee$ kan uttrykkes ved $\neg$ og $\wedge$. Til sammen gir dette oss at alle bindeordene vi har sett på kan uttrykkes ved $\downarrow$ alene eller ved $\mid$ alene, og i avsnittet om utfordringer så vi at alle sannhetsverdifunksjoner kan uttrykkes ved $\wedge$, $\vee$ og $\neg$ Dette viser at $\downarrow$ og $\mid$ hver seg er adekvate. Vi har vist at ingen andre bindeord kan være det. Derfor er det nøyaktig disse to som er adekvate.
Dette avslutter oppgaven.

Oppgave 5.2.3 Forklar hvorfor dette indirekte "beviset" for en uriktig påstand er feil:

Alle hele tall er positive.

Bevis
Anta at påstanden er feil, og la $x < 0$ være et helt tall. Da er $x^2 > 0$, og $x = \sqrt{x^2}$. Siden kvadratroten av et positivt tall alltid er positivt, får vi at $x > 0$, noe som var mot antagelsen. Antagelsen om at det fantes negative heltall må derfor være feil.

Oppgave 5.2.4 Vurder om følgende bevis for at det finnes uendelig mange primtall inneholder mangler eller direkte feil. Skriv et bevis som følger opp den samme tankegangen.

Bevis
Anta at det bare er endelige mange primtall, $p_1, \ldots, p_k$ La m være en mer enn produktet av alle primtallene. Siden m ikke kan ha noen p_i som faktor er m selv et primtall. Men det er umulig siden $p_i < m$ for alle i.

Oppgave 5.2.5 Vi argumenterte for bruk av regnereglene for mengdealgebra med at det er vanskelig og uoversiktlig å bruke Venn-diagrammer når vi har fire eller flere mengder. Vi skal argumentere for at det er umulig å tegne Venn-diagrammer for fire mengder hvis vi skal bruke sirkler. Ved å bruke andre og mer uformelige figurer er det selvfølgelig mulig.
Forklar hvorfor det beviset vi gir i prinsippet er riktig, og fyll ut detaljene:

Påstand
Fire sirkler i planet kan maksimalt dele planet inn i 14 deler.

Bevis
Anta at vi har tegnet inn tre sirkler, og at vi så tegner inn en fjerde. Denne vil krysse hver av de andre tre i høyst to punkter hver, så vi får høyst seks skjæringspunkter. Sirkelsegmentet mellom to skjæringspunkter vil dele ett av de områdene de tre første sirklene delte planet i i to. Det betyr at antall områder

disse fire sirklene deler planet i er høyst 6 mer enn det de tre greide, så vi får maksimalt 14 felt.

Konsekvens
Vi kan ikke tegne Venn-diagram for fire mengder ved å bruke sirkler.

5.3 Hvordan fører vi et bevis?

Etter å ha gått igjennom en del eksempler på bevis skal vi ta en sluttdiskusjon om hva som kreves av et matematisk bevis.
Det viktigste er at man aldri må basere seg på egenskaper ved de matematiske begrepene man arbeider med som ikke er en konsekvens av de tekniske definisjonene. Hvis vi for eksempel gir en definisjon av at

> to mengder A og B er *like store* hvis det finnes en $f : A \to B$ som er både injektiv og surjektiv.

så kan vi ikke tillegge dette begrepet flere egenskaper, for eksempel ved å si at det kan umulig være like mange naturlige tall som hele tall siden det er så mange flere hele tall enn naturlige tall. Da bruker vi en intuisjon som vi ikke har grunnlag for ut fra de tekniske definisjonene.
Ellers er det som tidligere nevnt viktig at man ikke foretar raske enkeltslutninger som leseren ikke kan henge med på eller unnlater å nevne problemer som leseren kan mene bør drøftes. En vanlig feil, som vi illustrerte gjennom beviset for formelen for vinkelsummen av en trekant, er at man ubevisst forenkler problemet ved at man ikke tar hele kompleksiteten i det man studerer inn over seg. I det tilfellet arbeider vi med mangekanter, hvor en mangekant er en lukket kurve i planet som består av endelig mange rette linjestykker som ikke krysser hverandre, og vi var interessert i vinklene mot det indre av figuren. I beviset tilla vi mangekantene flere egenskaper enn det som følger ut fra definisjonen.

Som hovedregel kan vi si at et bevis skal forklare en leser, og en selv, hvorfor en matematisk påstand er riktig. I tillegg til å gi et logisk holdbart argument for at påstanden er riktig, bør beviset derfor også være skrevet slik at det hjelper leseren til å se hvor problemene ligger og til å vise hvordan vi takler dem. Hvis vi gir et indirekte bevis som en del av et argument, i det minste en del som strekker seg over flere linjer, kan det være naturlig å starte med å presisere hva vi antar, og når vi kommer til en form for motsigelse, så bør vi referere tilbake til hva det var som ledet til denne motsigelsen, og som derfor er motbevist. Ellers, hvis beviset er litt langt, og man bruker påstander vist tidligere i beviset, eller endog i en lærebok, et kompendium eller i matematikklitteraturen generelt, så er det en god skikk å referere tilbake til det man bruker.
Den siste formaningen vi skal komme med er at man bør tenke på leseren og leserens kompetanse, men ut over det bare ta sikte på å uttrykke seg så klart og forståelig som mulig.

Kapittel 6

Rekursjon

6.1 De naturlige tallene

La oss ta utgangspunkt i følgende problem: Hvis vi trekker n linjer i planet slik at tre linjer aldri krysser hverandre i det samme punktet, kan vi da si noe om hvor mange felter vi deler planet inn i?
La oss ta en sjanse, og gjette på at dette antallet er fast, uavhengig av hvordan vi trekker disse linjene. La $A(n)$ være antall felter vi får med n linjer.
Det er lett å se at $A(0) = 1$, $A(1) = 2$ og $A(2) = 4$. Nå er det selvfølgelig fristene å gjette på at $A(n) = 2^n$, men en rask skisse på et papir viser at $A(3) = 7$, så denne gjettingen kan avvises med en gang.
Hvis vi ser på hva som skjer når vi trekker linje nummer tre, så ser vi at den nye linjen vil skjære de gamle linjene i to punkter, ett for hver linje. Da blir den nye linjen delt i tre segmenter, og hvert segment deler et av de gamle feltene i to. Det betyr at vi får et ekstra felt for hvert segment av den nye linjen, altså er $A(3) = 4 + 3 = 7$.
Vi kan fortsette tankegangen for å finne $A(4)$. Når vi trekker linje nr. 4, vil de tre første linjene dele denne nye i fire segmenter, hvert av disse segmentene deler et av de gamle feltene i to, så vi får fire nye felter. Det betyr at $A(4) = 7 + 4 = 11$.
Det er ikke noe magisk ved tallene 3 eller 4, vi kan tenke på samme måte for alle tall n. Resonementet vårt sier at $A(n) = A(n-1) + n$ fordi linje nummer n blir delt i n segmenter av de $n-1$ linjene som er på plass allerede, så n av feltene blir delt i to når vi trekker linje nummer n.

Vi ser at følgende to betingelser vil definere $A(n)$ for alle heltall $n \geq 0$:

- $A(0) = 1$
- $A(n) = n + A(n-1)$ når $n > 0$.

Dette er et eksempel på en funksjon $A : \mathbb{N}_0 \to \mathbb{N}_0$ definert ved *rekusjon*. Vi skal se på noen flere eksempler før vi gir en mer formell definisjon, men idéen bak rekursive definisjoner er at vi definerer en funksjon f ved på den ene siden å

bestemme verdien av f for ett eller noen få argumenter i starten, og for andre argumenter a lar vi $f(a)$ avhenge av verdier $f(b)$ hvor b kommer før a.

Eksempel 6.1 Fibonacci-tallene utgjør et historisk viktig eksempel.
Vi definerer $F(n)$ for $n \geq 1$ ved:

- $F(1) = F(2) = 1$.
- $F(n) = F(n-1) + F(n-2)$ når $n > 2$.

Det følger at $F(1) = 1$, $F(2) = 1$, $F(3) = 2$, $F(4) = 3$, $F(5) = 5$, $F(6) = 8$, $F(7) = 13$, $F(8) = 21$, $F(9) = 34$ og $F(10) = 55$.
Den som har lyst, kan regne ut $F(13)$ på egen hånd.
Vi finner igjen par av Fibonaccitall som står etter hverandre i naturen, for eksempel som antall høyre og venstredreide spiraler i en solsikke, og det har lenge vært kjent at forholdet $\frac{F(n+1)}{F(n)}$ har forholdet i det gyldne snittet som grenseverdi. Vi skal ikke se på det matematiske grunnlaget for denne observasjonen, ettersom det faller utenom våre temaer.

Eksempel 6.2 Vi kan ta utgangspunkt i eksemplet med deling av planet ved hjelp av linjer, og spørre om det finnes noen øvre grense for hvor mange felter vi kan dele planet i ved hjelp av n sirkler. Dette har relevans for hvor nyttige Venn-diagrammer kan være når vi skal studere mange mengder. Vi definerer $B(n)$ for $n \geq 0$ ved

- $B(0) = 1$
- $B(1) = 2$
- $B(n) = B(n-1) + 2n - 2$ når $n > 1$

Motivasjonen bak denne definisjonen er som følger: Hvis vi ikke bruker noen sirkel, er planet ikke delt, altså har vi én del. $B(0) = 1$.
Hvis vi har tegnet $n-1$ sirkler, og så tegner en ny, vil den nye skjære hver av de gamle i høyst to punkter. Dette deler den nye sirkelen opp i høyst $2n-2$ segmenter, og hvert segment deler en av de gamle delene i to. Det gir høyst $2n-2$ nye segmenter. Dette argumentet er greit bortsett fra for $n = 1$, hvor en sirkel uten krysspunkter fortsatt må regnes som et segment, slik at den faktisk deler planet i to, "innenfor" og "utenfor".

Hvis vi regner litt på dette får vi

1. $B(0) = 1$
2. $B(1) = 2$
3. $B(2) = B(1) + 2 \cdot 2 - 2 = 4$
4. $B(3) = B(2) + 2 \cdot 3 - 2 = 8$
5. $B(4) = B(3) + 2 \cdot 4 - 2 = 14$

6. $B(5) = B(4) + 2 \cdot 5 - 2 = 22$
7. $B(6) = 32$
8. $B(7) = 44$
9. $\cdots$
10. $\cdots$

Eksempel 6.3 Vi har sett på noen eksempler hvor vi definerer funksjoner fra naturlige tall til naturlige tall ved rekursjon. Det er imidlertid ikke vesentlig at verdiområdet for disse funksjonene er hele tall eller naturlige tall.
La oss som et eksempel anta at vi disponerer et fond på én million til tannløse hunders livsopphold. Fondet gir en avkastning på 10%, men fondets gode formål tilgodesees med kr. 50.000 pr. år. Hvis vi lar $A(0)$ være fondets størrelse i startåret, $A(n)$ være fondets størrelse etter n år og vi kun betaler ut fra fondet ved slutten av hvert år, så får vi følgende rekursive definisjon av $A(n)$:

- $A(0) = 1.000.000$
- $A(n) = 1,1 \cdot A(n-1) - 50.000$

Hvis vi regner litt på dette ser vi at de første leddene vil bli

- $A(1) = 1,1 \cdot 1.000.000 - 50.000 = 1.050.000$
- $A(2) = 1,1 \cdot 1.050.000 - 50.000 = 1.105.000$
- $A(3) = 1,1 \cdot 1.100.500 - 50.000 = 1.165.500$
- $A(4) = 1,1 \cdot 1.160.550 - 50.000 = 1.232.050$

Vi ser at for $A(7)$ vil vi være nede på 5-øresnivå, så da kan det hende at vi bør ta med avrunding til nærmeste 50-øring som en del av formelen.

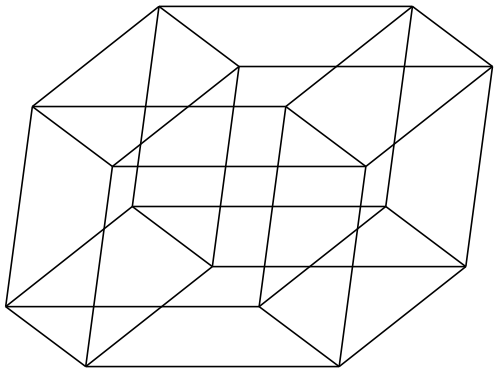
Eksempel 6.4 En rekursiv definisjon behøver ikke en gang å definere en funksjon. Den kan like gjerne være en instruksjon om hvordan man starter og hvordan man utfører et skritt. Vi skal se på et eksempel.

De fleste av oss kan tegne en terning på et stykke papir. Hvis man tar med skjulte streker, vil tegningen gjerne bestå av to kvadrater som er parallellforskjøvet i forhold til hverandre, og hvor vi trekker linjer mellom punktene i kvadratene som svarer til hverandre. En slik tegning kan oppfattes som en projeksjon av kantene i en terning ned i papirplanet.
(Dette er ikke helt sant, hvis vi tegner projeksonen av en terning slik at to av sidene fremstår som kvadrater, betyr det at disse flatene er parallelle med planet vi projiserer ned i, og da ville vi ikke sett noen av de andre linjene. Denne formen for juks vil gjennomsyre det som står under i dette eksemplet.)
Et punkt kan oppfattes som en null-dimensjonal terning, et linjestykke som en én-dimensjonal terning, et kvadrat som en to-dimmensjonal terning og terningen selv som en tre-dimensjonal terning. Matematikere snakker gjerne om

fire-dimensjonale terninger og generelt om n-dimensjonale terninger. Det vi skal gi nå er en metode til å tegne projeksjoner av n-dimensjonale terninger ned i papirplanet. Metoden er godt egnet for elektroniske tegneprogrammer

- For å tegne en 0-dimensjonal terning, tegn et punkt på arket.
- Anta at du har tegnet en $n-1$-dimensjonal terning.
  For å tegne en n-dimensjonal terning tar du en kopi av tegningen din, parallellforskyver den og trekker så linjer mellom tilsvarende punkter.

Hvis du følger denne oppskriften vil du få et linjestykke når $n = 1$ og et parallellogram når $n = 2$. For ikke å oppfatte dette som en feil, er det viktig å huske at projeksjonen av et kvadrat som ligger å flyter et sted oppe i rommet vil være et parallellogram, og ikke nødvendigvis et kvadrat, et rektangel eller en rombe. Nedenunder gir vi et eksempel på bildet av en fire-dimensjonal terning tegnet på denne måten:



En viktig årsak til at vi kan definere funksjoner ved rekursjon over $\mathbb{N}$ eller $\mathbb{N}_0$ er at disse mengdene er det vi kaller *induktivt oppbygget*. La oss se på $\mathbb{N}_0$ som er den vanlige mengden i logikk og programmeringsteori. N_0 er et eksempel på en *datatype* og er som sådan definert ved

- $0 \in \mathbb{N}_0$
- Hvis $n \in \mathbb{N}_0$ så er $S(n) \in N_0$.

"S" står for "successor" eller "etterfølger" som vi vil si på norsk. Dette er en helt formell definisjon som kan sees som løsrevet fra barneskolens intuisjon om hele tall. På den annen side er det ikke noe i veien for å kalle det formelle elementet $S(S(S(S(S(0)))))$ for 'fem' eller 5, så sammenhengen er klar. Poenget her er at når en matematisk struktur kan genereres på denne måten, så er den tilgjengelig for rekursive definisjoner. Eksempelvis kan vi definere addisjon og multiplikasjon ved rekursjon som følger:

- $x + 0 = x$
- $x + S(y) = S(x + y)$
- $x \cdot 0 = 0$
- $x \cdot S(y) = x \cdot y + x$

Vi kan godt oppfatte teksten over som *program* for addisjon og multiplikasjon over datastrukturen generert av 0 og S. I begge tilfellene har vi brukt *rekursjon i variabelen* y.
I avsnittene under skal vi se på andre strukturer som også aksepterer konstruksjoner ved rekursjon. Hensikten med å ta med disse eksemplene er å forklare de vesentlige aspektene ved rekursjon, og i neste kapittel, induksjon, slik at det blir lettere å forstå fenomenene for de naturlige tallene.

Oppgaver til avsnitt 6.1

Oppgave 6.1.1 Hvis vi bruker andre figurer enn sirkler til å tegne Venn-diagrammer, har vi selvfølgelig muligheten til å dele planet opp i flere felter. Eksempelvis kan to elipser skjære hverandre i fire punkter. Finn en rekursiv definisjon av en funksjon $C(n)$ definert for $n \geq 0$ slik at $C(n)$ gir en øvre grense for hvor mange felter vi kan dele planet inn i ved hjelp av n elipser.
Finn det minste tallet n slik at $C(n) < 2^n$ og forklar hvorfor vi heller ikke kan bruke elipser til å tegne Venn-diagrammer for n mengder.

Oppgave 6.1.2 Vis hvordan du kan definere x^y ved rekursjon over variabelen y.
Du kan anta at vi har definert sum og produkt, og du skal definere x^0 og $x^{S(y)}$.

Oppgave 6.1.3 Et enkelt fyrstikkspill for barn består i at spillerne etter tur forsyner seg med én eller to fyrstikker fra en bunke på n fyrstikker.
La $F(n)$ være antall måter dette spillet kan spilles på.

a) Forklar hvorfor $F(0) = 1$, $F(1) = 1$ og

$$F(n) = F(n-1) + F(n-2)$$

når $n \geq 2$. Kjenner du igjen funksjonen n fra et annet sted i heftet?

b) I det egentlige spillet skal man ta én, to eller tre fyrstikker. Bruk innsikten fra a) til å finne en rekursiv definisjon av funksjonen G hvor $G(n)$ er antall måter man kan spille dette spillet med n fyrstikker.

6.2 Eksempler

I det forrige avsnittet så vi på rekursive funksjoner defrinert over $\mathbb{N}$ eller over $\mathbb{N}_0$. Grunnen til at dette er en lovlig definisjon er at vi fanger opp alle tallene

på en og bare en måte ved å starte med 0 og så bruke etterfølgeren S.
Dette er imidlertid ikke den eneste situasjonen hvor vi konstruerer en matematisk struktur ved å starte med ett eller flere *grunnobjekter*, og så bruker forskjellige konstruksjoner til å lage oss mer komplekse objekter. Når dette er tilfellet, gir det også mening å definere funksjoner ved rekursjon.
Det vil føre for langt å gi en fullstendig inføring i induktivt definerte strukturer og rekursive definisjoner. I dette avsnittet vil vi gi noen eksempler på hvor dette fenomenet dukker opp.

6.2.1 Ord

Vi sier gjerne at når vi uttrykker oss på norsk, bruker vi et alfabet med 29 tegn. I en viss forstand er dette riktig, ettersom vi har 29 tegn som uttrykker lydbilder, og som vi derfor bruker når vi skal beskrive sammensatte lyder som vi kan tenkes å si til hverandre. Når vi uttrykker oss skriftlig, bruker vi imidlertid mange ekstra tegn. For det første kommer alle bokstavene våre i to varianter, som små og store bokstaver. For det andre bruker vi tegn, komma, spørsmålstegn mm. for å gi setningene våre mening. På tastaturet på datamaskinen har vi også egne tegn for prosent, dollar, euro og mye annet. Datamaskinen skiller ikke mellom bokstaver som representerer lydbilder og tegn som vi bruker for andre formål. I den teknologiske verden er alle disse tegnene likeverdige. Alfabetet vi bruker når vi kommuniserer med datamaskinen omfatter derfor mye mer enn de 29 bokstavene vi vanligvis sier at det norske alfabetet består av.
For den som kjenner litt til representasjon av data, så er egentlig alle disse tegnene vi omtalte over sekvenser av 0 og 1. For den som forsker på datamaskinens innerste vesen, er all informasjon lagret i ord over alfabetet $\{0, 1\}$. En teori for formelle språk vil måtte forholde seg til begge disse situasjonene, samt til 10-tallsystemet med sitt alfabet $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Derfor bruker vi en generell definisjon:

Definisjon 6.1 Et *alfabet* er en endelig mengde Σ av tegn.

Vi skal ikke problematisere hva et tegn er, men vi skal tenke på et tegn som noe som kan være grunndata for en datamaskin.

Definisjon 6.2 La Σ være et alfabet (som ikke inneholder det spesielle symbolet e).

- e er et ord over Σ, vi kaller e det *tomme ordet*.
- Hvis w er et ord over Σ og s er et tegn i Σ, så er ws et ord over Σ.

Vi vil normalt droppe e når vi skriver ned ord med tegn i, så hvis $\Sigma = \{0, 1\}$, vil vi skrive 01 i stedet for det "korrekte" $e01$.

Hvis målet er å studere ord over et alfabet matematisk, er denne definisjonen ganske rigid, det hadde holdt å si at et ord er en endelig ordnet sekvens av elementer i alfabetet. Hvis vi imidlertid skal behandle mengden av ord over et alfabet som en datatype, spiller definisjonen en stor rolle i forhold til hvordan

vi vil representere ord, hvordan vi vil hente ut informasjon om et ord og i det hele hvordan vi kan skrive programmer som regner på ord. Husk at data i seg selv er å betrakte som ord over $\{0,1\}$ så selv om vi kan gjøre forskjellige valg for forskjellige formål er forståelsen av ord over et alfabet av stor teknologisk betydning.
Vårt valg av definisjon av begrepet 'ord' er selvfølgelig gjort for å kunne illustrere induktivt oppbygde datatyper, men det representerer også en av de vanlige måtene å definere ord på.

Hvis v og w er to ord, setter vi dem intuitivt sammen til et ord vw ved først å skrive tegnene i v og deretter tegnene i w. I vår formelle verden, gir ikke dette mening, og da er det praktisk at vi kan definere sammensetning ved *rekursjon på oppbyggingen av* w.

- $ve = v$
- $v(ws) = (vw)s$

Av og til kan det være oppklarende å være enda mer formell. Vi tar utgangspunkt i et alfabet Σ, og definerer mengden Σ^* som den som er generert fra $e \in \Sigma^*$ ved hjelp av $P : \Sigma^* \times \Sigma \to \Sigma^*$. På samme måten som vi klarte oss med $O \in \mathbb{N}_0$ og antagelsen om en $S : \mathbb{N}_0 \to \mathbb{N}_0$ for å definere $\mathbb{N}_0$, så klarer vi oss med $e \in \Sigma^*$ og antagelsen om en $P : \Sigma^* \times \Sigma \to \Sigma$ for å definere Σ^*. Hvis $\Sigma = \{a, b\}$ skal vi riktignok formelt sett skrive $P(P(P(e, a), b), a)$ i stedet for aba av samme grunn som vi formelt sett skal skrive $S(S(S(0)))$ i stedet for 3. Så formelle bør vi ikke være untatt når omstendighetene tvinger oss til det (for eksempel når vi skal programmere).

Den funksjonen P vi har brukt kalles ofte "PUSH". En annen funksjon som brukes når man programmerer med ord er "POP". POP(e) er udefinert, mens POP$(ws) = s$ for alle ord w og symboler s. I avsnittet om utfordringer skal vi se nærmere på dette.
Vi skal gi ett eksempel til på en rekursiv definisjon over Σ^*, et eksempel hvor vi ikke kan oversette bokstav for bokstav.

Definisjon 6.3 La $\Sigma = \{a, b\}$

a) Vi definerer funksjonen lh ved

   - $lh(e) = 0$.
   - $lh(ws) = lh(w) + 1$.

   Denne definisjonen gir oss bare at $lh(w)$ er lengden av ordet w.

b) Vi definerer *tallkoden* $c(w)$ rekursivt ved

   - $c(e) = 0$.
   - $c(wa) = c(w) + 3^{lh(w)}$.
   - $c(wb) = c(w) + 2 \cdot 3^{lh(w)}$.

Et poeng her er at hvis $v \neq w$, så er $c(v) \neq c(w)$, slik at vi har en injektiv funksjon fra mengden Σ^* til $\mathbb{N}_0$. I mengdeteoriterminologi betyr dette at det er minst like mange ikke-negative hele tall som det er ord over alfabetet $\{a, b\}$.

6.2.2 Utsagnslogiske formler

Vi kan definere et formelt språk for utsagnslogikk med tre utsagnsvariable A, B og C ved

- A, B og C er formler.
- Hvis P og Q er formler, vil $(\neg P)$, $(P \vee Q)$, $(P \wedge Q)$, $(P \to Q)$ og $(P \leftrightarrow Q)$ være formler.
- Alle formlene fremkommer ved reglene over.

På sett og vis har denne definisjonen det samme formatet som de induktive definisjonene av $\mathbb{N}_0$ og av Σ^*, vi har valgt ut noen basisobjekter, i dette tilfelle utsagnsvariablene, og vi har fortalt hvordan vi konstruerer nye objekter fra gamle. Valg av notasjon er pare et spørsmål om lesbarhet, vi kunne like gjerne skrevet de sammensatte formlene som $\neg(P)$, $\vee(P, Q)$, $\wedge(P, Q)$, $\to (P, Q)$ og $\leftrightarrow (P, Q)$. De utsagnslogiske formlene er derfor generert fra utsagnsvariablene ved hjelp av en funksjon med én variable og fire funksjoner med to variable.
Dette er igjen en situasjon hvor vi kan definere funksjoner ved rekursjon, ved at vi definerer funksjonen direkte på basisobjektene og ved hjelp av hvordan funksjonen virker på forgjengerne for sammensatte objekter. Vi skal se på to eksempler:

Eksempel 6.5 Vi er interessert i å skrive om en formel til en ekvivalent formel som ikke inneholder bindeordene $\to$ og $\leftrightarrow$. Hvis P er en formel, definerer vi $F(P)$ som følger:

- $F(A) = A$, $F(B) = B$ og $F(C) = C$.
- $F((\neg P)) = (\neg F(P))$
- $F((P \vee Q)) = (F(P) \vee F(Q))$ og $F((P \wedge Q)) = (F(P) \wedge F(Q))$.
- $F((P \to Q)) = ((\neg F(P)) \vee F(Q))$
- $F((P \leftrightarrow Q)) = (((\neg F(P)) \vee F(Q)) \wedge ((\neg F(Q)) \vee F(P)))$

Det kan muligens virke noe forvirrende med alle parentesene, men skal dette gjøre formelt riktig, må de være der.

I vårt neste eksempel vil vi anta at formelen ikke inneholder noen $\to$ eller $\leftrightarrow$. Vi kan da velge om vi vil si at vi nå ser på en mengde definert ved induksjon, men med færre *konstruktører*, eller om vi vil si at vi begrenser den rekursive definisjonen til noen tilfeller, og lar den være udefinert i andre tilfeller.

Eksempel 6.6 I avsnittet om strømkretser og utsagnslogikk så vi at vi kan tegne strømkretser for formler hvor negasjonstegnet bare står i tilknytning til utsagnsvariable. Enhver formel kan skrives om til en ekvivalent formel på den ønskede formen, og vi finner den ekvivalente formelen ved hjelp av en rekursiv definisjon av en funksjon G (vi dropper unødige parenteser her):

- $G(A) = A$, $G(B) = B$ og $G(C) = C$.
- $G(\neg A) = \neg A$, $G(\neg B) = \neg B$ og $G(\neg C) = \neg C$
- $G(P \vee Q) = G(P) \vee G(Q)$ og $G(P \wedge Q) = G(P) \wedge G(Q)$
- $G(\neg\neg P) = G(P)$
- $G(\neg(P \wedge Q)) = G(\neg P) \vee G(\neg Q)$
- $G(\neg(P \vee Q)) = G(\neg P) \wedge G(\neg Q)$

6.2.3 Programmeringsspråk

I avsnitt 4.6 beskrev vi et enkelt programmeringsspråk. Klassen av programmer er definert ut fra noen *grunninstruksjoner* av formen $x := 17$ og $x := x \dot{-} 1$ (og andre) ved hjelp av sammensetning av programmer og mere komplekse programmer som

$$if \;\; x > 0 \;\; then \;\; P \;\; else \;\; Q \;\; fi.$$

Vi skal ikke repetere hele definisjonen her. Poenget er at klassen av programmer er definert ved hjelp av to sett av prinsipper. Det ene angir grunnprogrammene og det andre angir måter vi kan konstruere komplekse programmer fra enklere programmer på.
Når vi da skal definere hva et program "gjør", skjer det ved rekursjon på oppbyggingen av programmet. Hvis vi tolker et program som en funksjon som sender visse inngangsdata til utgangsdata, tolker vi sammensatte programmer ved hjelp av sammensetning av funksjoner, og løkker eller hvis-så-ellers programmer blir tolket ut fra tolkningen av de umiddelbare delprogrammene. Den matematiske analysen av programmeringsspråk, både forskjellige matematiske tolkninger av programmer og spørsmål om logikk i tilknytning til hva programmer "gjør" er gjennomsyret av rekursive definisjoner.

Oppgaver til avsnitt 6.2

Oppgave 6.2.1 La $\Sigma = \{a, b\}$. Vis at følgende funksjoner definert på Σ^* kan defineres rekursivt:

a) $f(w)$ er differensen mellom antall a'er i w og antall b'er i w.

b) $g(w)$ er antall a'er i w.

c) $h(w)$ er ordet i alfabetet $\{c, d\}$ vi får hvis vi bytter ut alle forekomster av a i w med ordet cd og alle forekomster av b i w med ordet ddc.

Hensikten her er ikke å gjøre noe matematisk interessant, men å vise at man har forstått formatet på definisjoner ved rekursjon.

Oppgave 6.2.2 La F og G være som i underavsnitt 6.2.2. La

$$P = (A \leftrightarrow B) \rightarrow C$$

(hvor vi dropper noen parenteser).

a) Finn $F(P)$ og forklar hvorfor vi får en ekvivalent formel.

b) Finn $G(F(P))$.

I begge tilfellene, skriv opp hele utregningen

6.3 Utfordringer

6.3.1 Gramatikker

Vi så på mengden av ord over et alfabet som en induktivt definert mengde som tillater konstruksjoner ved rekursjon. En alternativ tilnærming til teorien for formelle språk er studiet av *formelle gramatikker*. Hvis vi ser på familien av utsagnslogiske formler i utsagnsvariablene A, B og C, ser vi at enhver formel er et ord i alfabetet

$$\{A, B, C, \neg, \vee, \wedge, \rightarrow, \leftrightarrow, (,)\}.$$

Denne mengden av formler er da et eksempel på et *formelt språk*, det vil si en mengde av ord i et alfabet. Vi skiller mellom formelle språk og naturlige språk som norsk, engelsk, sanskrit og noen til. Alle programmeringsspråk er formelle språk, i den forstand at de er definert gjennom en teknisk definisjon og det kan aldri være tvil om et ord (et forsøk på program) tilhører språket eller ikke.
Både mengden av utsagnslogiske formler og det programmeringsspråket vi har sett på som eksempel er definert innenfra og utover, i den forstand at vi har noen basisord, og vi har noen regler for hvordan vi kan generere mere komplekse uttrykk fra enkere uttrykk. For å studere denne typen konstruksjoner har man utviklet teorien for *formelle gramatikker*. Nordmannen Axel Thue var en foregangsmann i dette feltet.
Det vil føre alt for langt å utvikle noen teori for gramatikker her, så la oss nøye oss med kort å se på en viktig klasse, de kontekstfrie gramatikkene.

Definisjon 6.4 En *konteksfri gramatikk* består av

1. Et alfabet Σ
2. Et alfabet Γ disjunkt fra Σ
3. Et startsymbol $S \in \Gamma$

4. En endelig mengde R av regler på formen

$$G \to w$$

hvor w er et ord i $\Sigma \cup \Gamma$.

Eksempel 6.7 La $\Sigma = \{A, B, C, \neg, \vee, \wedge, (,)\}$
La $\Gamma = \{S\}$.
La R bestå av reglene

- $S \to A$
- $S \to B$
- $S \to C$
- $S \to (\neg S)$
- $S \to (S \vee S)$
- $S \to (S \wedge S)$

Idéen bak en slik gramatikk er at vi kan bruke den til å *utlede* ord fra S ved å bruke reglene på følgende måte:
Hvis $G \to w$ er en regel i R og $u = vGy$ er et ord i $\Sigma \cup \Gamma$, sier vi at

$$xGy \vdash xwy,$$

det vil si at vi erstatter en forekomst av G i et ord med ordet w.
En *utledning* vil være en rekke av ord $w_0 = S, w_1, \ldots, w_n$ slik at

$$w_i \vdash w_{i+1} \text{ for } i < n.$$

Målet er å finne utledninger for ord i Σ. I dette tilfelle vil vi kunne utlede nøyaktig alle utsagnslogiske formler i variablene A, B og C, hvor vi bare benytter bindeordene $\neg$, $\vee$ og $\wedge$.
Bokstavene i Γ kan sies å stå for gramatikalske former. Bokstaven S står for "setning", og antyder at vi genererer en mengde setninger i en eller annen forstand.

Selv om vi ikke kan komme inn på flere eksempler, er kontekstfrie gramatikker hyppig forekommende i informatikk, så hyppig at de har laget seg sin egen notasjon.
Eksempelvis vil vårt eksempel skrives som

- Formel S
- $S ::= A|B|C|(\neg S)|(S \vee S)|(S \wedge S)$

6.3.2 Borellmengder

I de eksemplene vi har sett på har vi hatt endelige basismengder, og nye objekter har vært konstruert fra endelig mange gamle. Dette er egentlig ikke noen nødvendig begrensning.

Eksempel 6.8 Vi definerer mengden B ved

- Hvis x og y er reelle tall, lar vi $\{x, y\} \in B$.
- Hvis $x_n \in B$ for hver $n \in \mathbb{N}$, lar vi *følgen* $\{x_n\}_{n\in\mathbb{N}} \in B$.
- Hvis $x \in B$, lar vi *det ordnede paret* $(c, x) \in B$ hvor c er et på forhånd utplukket objekt.
- B består nøyaktig av de objektene som fremkommer ved bruk av reglene over.

Poenget her er at elementene i mengden B definerer delmengder av $\mathbb{R}$ på en naturlig måte. For hver $x \in B$ definerer vi C_x ved rekursjon på x som følger:

- Hvis $x \in \mathbb{R}$ og $y \in \mathbb{R}$, lar vi $C_{\{x,y\}}$ være det åpne intervallet mellom x og y.
- Hvis $x = \{x_n\}_{n\in\mathbb{N}}$ hvor hver $x_n \in B$, lar vi $C_x = \bigcup_{n\in\mathbb{N}} C_{x_n}$
- Hvis $x = (c, y)$ hvor $y \in B$, så lar vi $C_x = \mathbb{R} \setminus C_y$.

Selv om dette eksemplet er mer komplekst og vanskeligere å gjennomskue enn de foregående, så er vi i samme situasjon som tidligere: Vi har generert en mengde fra noen grunnelementer ved hjelp av visse funksjoner eller konstruksjoner, og da kan vi gjennomføre rekursive definisjoner som gir mening for alle de objektene vi har laget oss.
Det vi har gjort er å definere *Borell-mengdene*. En vanlig karakterisering av Borellmengdene er at de utgjør den minste klassen av delmengder av $\mathbb{R}$ som inneholder mengden av åpne intervaller og som er lukket under komplement og unioner av opptellbare delmengder (elementer i en følge). Vår tilnærming er ekvivalent, men tar mere vare på dynamikken i konstruksjonen av Borellmengdene.
Borellmengdene er av interesse i matematikk og sannsynlighetsteori, fordi vi på en naturlig måte kan beskrive hvor "lang" en Borell-mengde er uten å ta i bruk avanserte matematiske definisjoner. Definisjonen krever likevel at man behersker teorien for summer av uendelige følger, og det faller utenom læringsmålet for dette heftet.

6.3.3 Programmering

Vi har sett på et enkelt programmeringsspråk basert på regning med registre, hvor vi kunne sjekke om verdien av et register (eller en variabel) er null eller ikke, og hvor opsjonene var å legge til et tall i et register eller ta bort et. Dette

gir selvfølgelig meget trege programmer.
Det er imidlertid naturlig å utvide denne formen for programmer til også å regne på ord i alfabetet $\{0,1\}$. Vi vil da ha tre måter å forandre en variabel på:

- $PUSH(0)$: $x := x0$
- $PUSH(1)$: $x := x1$
- $TAIL(x)$ som vil gi y hvis $x = y0$ eller $x = y1$

Vi vil også ha en "mens $x \neq e$" kommando og en utnyttelse av $POP(x0) = 0$ og $POP(x1) = 1$ til å skille mellom forskjellige underprogrammer.
Vi går ikke i detalj her.

6.4 Blandede oppgaver

Oppgave 6.4.1 (u) I Definisjon 3.1 definerte vi $UV(\phi)$ for en mengde ϕ definert fra en familie grunnmengder $A_1, \ldots, A_n$ i et univers U.
Forklar hvordan mengden av *definisjoner* av slike mengder ϕ kan oppfattes som en rekursivt definert struktur, og hvordan UV kan defineres ved rekursjon over denne strukturen.

Oppgave 6.4.2 Det er mulig å definere funksjoner av flere variable ved rekursjon over noen eller alle variablene, og det er mulig å definere flere funksjoner samtidig, såkalt *simultanrekursjon*. Vi skal se på et eksempel på dette:
La $\Sigma = \{0,1\}$. Vi skal oppfatte ordene i Σ^* som binære tall, og vi skal se hvordan vi kan definere addisjon av binære tall ved rekursjon. For å få til det, må vi samtidig definere binær addisjon når vi har med mente. Det betyr at vi vil definere to funksjoner $v + w$ og $v +' w$.
Vi vil la $v + w$ være resultatet av binær addisjon, mens $v +' w$ skal bli $v + w + 1$, altså det vi bør få hvis vi har med en mente fra en tidligere del av regnestykket. Det tomme ordet oppfattes best som 0. Vi ser på følgende definisjon:

1. $e + e = e0$
2. $e +' e = e1$
3. $e + ws = ws + e = ws$ for $s \in \{0,1\}$
4. $e +' w0 = w0 +' e = w1$
5. $e +' w1 = w1 +' e = (e +' w)0$
6. $v0 + w0 = (v + w)0$
7. $v1 + w0 = v0 + w1 = (v + w)1$
8. $v1 + w1 = (v +' w)0$
9. $v0 +' w0 = (v + w)1$

10. $v0 +' w1 = v1 +' w0 = (v +' w)0$

11. $v1 +' w1 = (v +' w)1$

a) Finn $e101 + e11$ ved å bruke definisjonen over.

b) Sammenlikn utregningen av $e101 + 11$ med addisjon av de binære tallene for 5 og 3.

c) Gi en skriftlig forklaring på sammenhengen mellom denne rekursive definisjonen og addsjon av binære tall.

Det er et poeng at denne formen for addisjon har et tidsforbruk som er proposjonalt med lengden av binærrepresentasjonen, mens addisjon slik vi definerte den over $\mathbb{N}_0$ generert fra 0 og S har et tidsforbruk som er proposjonalt med tallet selv. Dette er en enkel illustrasjon på hvordan en fornuftig representasjon av data kan spare oss for mye regnetid.

Kapittel 7

Induksjonsbevis

7.1 Tradisjonell induksjon

Det mest tradisjonelle bildet som skal illustrere induksjonsbevis er dominobrikkene som er satt opp etter hverandre i en lang rekke. Idéen bak en slik oppstilling er at hvis en brikke velter bakover, vil den ta med seg den neste. Konsekvensen er at hvis vi velter den første brikken, så vil brikke nr. 2 velte fordi den felles av den første. Deretter velter brikke nr. tre fordi den felles av den andre, brikke nr. fire velter deretter, så nr. 5 osv. Det er dette selvfølgelige osv. som er kjernen i induksjonsbevis.

La oss se på et annet eksempel. For noen år siden oppdaget en gammel hunnape i Japan at det gikk an å varme seg om vinteren ved å bade i den varme kilden kolonien bodde i nærheten av. Hver ny generasjon lærer denne ferdigheten av foreldregenerasjonen, så siden den gang har alle etterkommerne av den gamle apen levd et behagelig vinterliv. Fenomenet reflekterer det samme som de fallende dominobrikkene, når først et objekt i en rekke har en egenskap (faller bakover eller kjenner gleden ved å bade i varme kilder) og det er slik at hvis et objekt har denne egenskapen, så vil også neste ledd i rekken ha denne egenskapen (ved å bli veltet eller ved å lære av foreldrene), så vil alle objektene i rekken fra og med det første individet dele egenskapen.

Vi skal se på et tredje eksempel: På et lagledermøte i den lokale idrettsforeningen er det sekretærens oppgave å innhente en navneliste over alle laglederne som er møtt opp. Han gir et tomt ark til den første laglederen (første rad til høyre) og ber om at hver lagleder skriver navn og lag på arket og så sender det videre. Resultatet er at dette arket passerer alle deltagerne på møtet og sekretæren får innhentet sine opplysninger. Mønsteret er igjen det samme. Hvis den første utfører oppgaven og vi sørger for at når én har utført oppgaven så vil også den neste gjøre det, så har vi i realiteten sørget for at alle utfører oppgaven.

Vi har tatt med disse eksemplene i håp om å avmystifisere induksjonsbevis før vi behandler dem matematisk. Det er også på tide å innrømme at vi har antatt

at intuisjonen rundt induksjon er så enkel at vi har tatt teoremer som egentlig krever bevis ved induksjon for god fisk tidligere i heftet. Vi kommer i noen grad tilbake til dette. Problemet med induksjonsbevis er ikke at de er vanskelige å forstå, men at med en gang vi isolerer induksjonsbevis som en egen bevisform, føler mange at de er bundet av formalismen rundt, og begynner å føle seg usikre. Det er nå på tide å se på noen matematiske eksempler:

Eksempel 7.1 Vi starter med et eksempel som alle innføringer i induksjonsbevis synes å ta med:

Påstand
Summen

$$1 + 2 + \ldots + n$$

av de n første naturlige tallene er $\frac{n(n+1)}{2}$.

Bevis
Vi beviser to delpåstander som vi vil kalle *induksjonstart* og *induksjonskritt*.

Induksjonstart: Påstanden holder når $n = 1$.

Induksjonstarten vises ved å sette 1 inn for n i uttrykket $\frac{n(n+1)}{2}$ og se at vi får 1 som også er summen av det første tallet.

Induksjonskritt: Anta at påstanden holder for et tall n. Da holder den også for det neste tallet $n + 1$.

Vi beviser induksjonskrittet ved å regne på

$$1 + \cdots + n + (n + 1).$$

Vi kan skrive dette som

$$(1 + \cdots + n) + (n + 1),$$

og hvis vi bruker antagelsen om at påstanden holder for n, blir dette det samme som

$$\frac{n(n+1)}{2} + (n + 1).$$

Nå setter vi dette uttrykket på felles brøkstrek og får

$$\frac{n(n+1)}{2} + (n + 1) = \frac{n(n+1) + 2(n+1)}{2}$$

$$= \frac{(n+1)(n+2)}{2} = \frac{(n+1)((n+1)+1)}{2}.$$

Dette er jo nettopp påstanden vår for $n + 1$.

Kan vi så trekke noen konsekvenser av dette?
La $P(n)$ være påstanden vår. Vi har vist at $P(1)$ holder, og vi har vist $P(n) \Rightarrow P(n + 1)$.
Fra utsagnslogikk har vi at $P(1) \wedge (P(1) \rightarrow P(2)) \Rightarrow P(2)$

Det betyr at $P(2)$ må være sann.
Fra utsagnslogikk har vi også at $P(2) \wedge (P(2) \rightarrow P(3)) \Rightarrow P(3)$.
Det betyr at $P(3)$ må være sann.
Fra $P(3)$ og induksjonskrittet følger $P(4)$, så følger $P(5)$, så $P(6)$, $P(7)$ osv.
Vi ser at vi kan gi et direkte bevis for hver $P(n)$ ved å bruke induksjonstarten, induksjonskrittet for alle tall $k < n$ og ren utsagnslogikk. Denne innsikten sier oss at $P(n)$ må være sann for alle n. Induksjonsbevis formaliserer denne innsikten.

Eksempel 7.2 I innledningen til kapittel 6 diskuterte vi hvor mange felt n linjer i planet vil dele planet i, når ingen punkt er kryssningspunkt for tre eller flere linjer.
Vi fant ut at dette antallet er gitt av en funksjon $A(n)$ som er definert ved rekursjon, hvor $A(0) = 1$ og $A(n+1) = A(n) + n + 1$ for vilkårlige n.
Vi så at disse to fakta vil bestemme $A(n)$ for alle n, og argumentet for det er akkurat det samme som argumentet for at beviset i eksemplet over faktisk beviser at påstanden i det eksemplet holder for alle n.
Hvis omfanget av dette emnet hadde vært større enn 5 studiepoeng, kunne vi gått inn på endel teknikker for å utvikle eksplisitte beskrivelser av funksjoner som er definert ved visse former for rekursjon. I dette og i et senere eksempel skal vi se på slike eksplisitte beskrivelser, og de er funnet ved å bruke standard metoder kjent for studenter av diskret matematikk.

Påstand
La $A(n)$ være antall felter vi kan dele planet inn i ved hjelp av n rette linjer. Da er

$$A(n) = \frac{(n+2)(n-1)+4}{2}.$$

Bevis
Vi vet at $A(0) = 1$. Setter vi inn 0 for n i uttrykket over, vil vi også få 1.
La nå k være fast og anta at

$$A(k) = \frac{(k+2)(k-1)+4}{2}.$$

Da er

$$A(k+1) = A(k) + k + 1 = \frac{(k+2)(k-1)+4}{2} + k + 1.$$

Setter vi dette på felles brøkstrek og regner på uttrykket får vi

$$A(k+1) = \frac{(k+2)(k-1)+4+2k+2}{2} = \frac{k^2+k-2+4+2k+2}{2}$$

$$= \frac{k^2+3k+4}{2} = \frac{(k+3)k+4}{2}.$$

Dette er det samme som vi får når vi setter inn $k+1$ for n i uttrykket

$$\frac{(n+2)(n-1)+4}{2}.$$

Vi kan nå resonere som i eksemplet over. Vi har vist ved direkte innsetning at formelen vår for $A(n)$ holder når $n = 0$, og vi har vist at hvis den holder for et tall k så holder den også for tallet $k + 1$. Da tilsier dominoeffekten at vi har funnet en formel som virker for $n = 1$, $n = 2$, $n = 3$, osv. Det er dette 'osv.' som fortsatt er kjernen i induksjonsbevis.

Før vi oppsummerer denne første delen av innføring i induksjonsbevis, skal vi se på et mer komplekst eksempel:

Eksempel 7.3 Vi har tidligere vist at vinkelsummen i en mangekant med n hjørner vil være $(n - 2)180°$, men bevisene har ikke holdt vann fordi de har ubevisst antatt at mangekanten er konveks. Så la oss først være tilstrekkelig presis på hva vi mener med en mangekant:
La $P_1, \ldots, P_n$ være n forskjellige punkter i planet slik at $n \geq 3$ og slik at hvis vi trekker alle linjene P_iP_{i+1} for $i < n$ og linjen P_nP_1, så vil ingen av disse krysse hverandre. En *mangekant* er da et område av planet avgrenset av slike linjer.

Vi vil kalle $P_1, \ldots, P_n$ for *hjørnene* i mangekanten. Det betyr at vi holder muligheten åpen for at et hjørne kan ligge på linjen mellom nabohjørnene, men vi holder for eksempel ikke muligheten åpen for at figuren vi tegner får et kryss i noen av punktene P_i.

Vi vet at formelen stemmer for mangekanter med tre hjørner.
Vi skal nå se at hvis formelen stemmer for mangekanter med n hjørner hvor $n \geq 3$, så stemmer den for mangekanter med $n + 1$ hjørner også. Det kan være en fordel å tegne en figur for hver av delene i beviset, som er et bevis ved tilfeller.

Lemma
Anta at vinkelsummen i enhver mangekant med n hjørner er $(n - 2)180°$*, hvor* $n \geq 3$.
Da er vinkelsummen i enhver mangekant med $n + 1$ *hjørner* $((n + 1) - 2)180°$.

Bevis
La $P_1, \ldots, P_{n+1}$ være hjørnene i mangekanten M, listet opp i rekkefølge med urviseren.
Hvis vi fjerner linjene P_1P_{n+1} og $P_{n+1}P_n$ og trekker linjen P_1P_n i stedet, får vi en ny mangekant M^- med n hjørner.
Vi må se på tre muligheter:

1. P_{n+1} ligger på linjen mellom P_1 og P_n.
   Da er tilsynelatende M og M^- den samme mangekanten, men vi har tegnet inn et ekstra "hjørne" på en av linjene i M^-. Det vil legge $180°$ til vinkelsummen, fordi vi da tar med vinkelen på $180°$ som ligger i P_{n+1}.

2. P_{k+1} ligger utenfor M^-.
   Da får vi vinkelsummen i M ved å legge sammen vinkelsummen i M^- med vinkelsummen i trekanten $P_nP_{n+1}P_1$, som jo er $180°$.

3. P_{k+1} ligger innenfor M^-.
   For å finne vinkelsummen i M fra vinkelsummen i M^- legger vi først til 360° i punktet P_{k+1}, og så trekker vi fra vinkelsummen til trekanten $P_nP_{n+1}P_1$, som fortsatt er 180°

Når vi først har bevist dette lemmaet kan vi bruke det til å vise at vinkelsummen i en firkant er 360°.
Når vi vet at vinkelsummen i en firkant er 360°, kan vi bruke lemmaet til å vise at vinkelsummen i en femkant er 540°.
Når vi vet at at vinkelsummen i en femkant er 540°, kan vi bruke lemmaet til å vise at vinkelsummen i en sekskant er 720°.
Slik kan vi fortsette og fortsette, og vi ser at ved å bruke dette lemmaet tilstrekkelig mange ganger kan vi bestemme vinkelsummen i enhver mangekant, såfremt vi kjenner antall hjørner. Det er derfor like greit å trekke den slutningen at vi fra det faktum at vinkelsummen i en trekant er 180° og lemmaet kan konkludere med at vinkelsummen i en n-kant er $(n-2)180°$.

Vi har nå sett på tiltrekkelig mange eksempler til at vi kan formulere *Induksjonsprinsippet* i sin grunnleggende form:

Induksjonsprinsippet I
La $P(n)$ være en påstand om et vilkårlig naturlig tall.
Anta at vi kan bevise

- *Induksjonstarten* $P(1)$, det vil si at P holder for $n = 1$.
- *Induksjonskrittet* $P(k) \Rightarrow P(k+1)$ for et vilkårlig naturlig tall k.

Da kan vi konkludere med at $P(n)$ holder for alle naturlige tall n.

Hvis man vil ha en aksiomatisk tilnærming til matematikken, det vil si at vi legger noen aksiomer og noen faste regler for hvordan vi kan trekke nye satser ut fra gamle til grunn for våre beviser, så vil induksjonsprinsippet gjerne være en av disse faste reglene, alternativt inngå som et aksiom.
Problemet med dette vil imidlertid ofte være at forfattere av bevis blir for bundet opp av formuleringen av induksjonsprinsippet, mens alternative bevis som baserer seg på innsikten bak prinsippet vil være lettere både å skrive og å forstå.
I resten av dette avsnittet og i det neste skal vi se på noen eksempler hvor vi benytter oss av innsikten, uten at de bevisene vi gir holder seg strengt til formuleringen av induksjonsprinsippet.

Eksempel 7.4 Vi har definert Fibonacci-tallene $F(n)$ ved følgende rekursive definisjon:

- $F(0) = 1$
- $F(1) = 1$
- $F(n) = F(n-1) + F(n-2)$ når $n \geq 2$

I kapitlet om rekursjon argumenterte vi for at $F(n)$ er bestemt for alle $n \geq 0$ på denne måten.
Hvis vi nå er i stand til å finne en formel for $F(n)$ og vi vil bevise at den er riktig, så må vi gjøre tre ting:

- Vi må vise at formelen holder for $n = 0$.
- Vi må vise at formelen holder for $n = 1$.
- Vi må vise at formelen holder for n ut fra en antagelse om at den holder for $n - 1$ og $n - 2$, siden $F(n)$ er bestemt av $F(n - 1)$ og $F(n - 2)$.

Vi vil ha at

$$F(n) = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^{n+1} - (\frac{1-\sqrt{5}}{2})^{n+1}).$$

Dette er ikke basert på gjetning, men på en generell teori for å finne formler som beskriver leddene i rekursivt definerte følger, en generell teori som vi ikke skal utvikle her.
La

$$G(n) = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^{n+1} - (\frac{1-\sqrt{5}}{2})^{n+1}).$$

og la $F(n)$ være Fibonaccitallet slik det er definert rekursivt. Vår påstand er da at

$$F(n) = G(n)$$

for alle tall $n \geq 0$.
Ved innsetting og litt regning får vi at $G(0) = 1$.
Ved innsetting og noe mer regning får vi at $G(1) = 1$. Den noe mere regningen vil omfatte bruk av kvadratsetningene.
Tilslutt vil vi vise at $G(n) = G(n - 1) + G(n - 2)$ for alle n. Når vi har vist det, følger det at $G(n) = F(n)$ fra premissene $F(n - 2) = G(n - 2)$, $F(n - 1) = G(n - 1)$ og $F(n) = F(n - 1) + F(n - 2)$.
Vi skal se på litt av regningen som ligger bak induksjonskrittet, men la oss først se på hvorfor det vi gjør faktisk beviser at formelen vår er riktig:
Vi viser $P(0)$ og $P(1)$, og vi viser $P(k - 2) \wedge P(k - 1) \Rightarrow P(k)$ for vilkårlige tall k.
Dette gir oss umiddelbart $P(3)$. Men fra $P(2) \wedge P(3)$ og induksjonskrittet får vi $P(4)$. Slik kan vi fortsette og vise $P(5)$, $P(6)$ osv. Riktignok må vi vise både $P(15)$ og $P(16)$ før vi kan konkludere $P(17)$ når vi setter det opp på denne måten, men vi vet at hvis vi vil, så byr ikke det på noen problemer. Det betyr at når vi har etablert $P(0)$, $P(1)$ og induksjonskrittet, så kan vi være trygge på at $P(n)$ holder for alle n.
Dette utvidede induksjonsprinsippet kan sammenliknes med dominobrikker som er stilt opp med en av langsidene ned, slik at begge de to foregående må velte før en brikke velter (på grunn av lavere tyngdepunkt i forhold til bredden på grunnflaten). Hvis vi da velter de to første brikkene i rekka, vil alle de andre følge etter av seg selv.

Vi kan vise induksjonskrittet for Fibonaccifølgen ved å vise to enkeltpåstander som ikke trenger så mye regning:

1. $$(\frac{1+\sqrt{5}}{2})^{n-1} + (\frac{1+\sqrt{5}}{2})^{n} = (\frac{1+\sqrt{5}}{2})^{n+1}$$

2. $$(\frac{1-\sqrt{5}}{2})^{n-1} + (\frac{1-\sqrt{5}}{2})^{n} = (\frac{1-\sqrt{5}}{2})^{n+1}$$

Disse følger av

1. $$1 + \frac{1+\sqrt{5}}{2} = (\frac{1+\sqrt{5}}{2})^{2}$$

2. $$1 + \frac{1-\sqrt{5}}{2} = (\frac{1-\sqrt{5}}{2})^{2}$$

og begge disse kan vises ved direkte utregning.

Eksempel 7.5 La $P_1, \ldots, P_k$ være utsagnsvariable, og la A være et sammensatt utsagn hvor vi begrenser oss til disse utsagnsvariablene og bindeordene $\neg$, $\wedge$ og $\vee$.
Ved rekursjon på A definerer vi A^* ved

- $P_i^* = \neg P_i$.
- $(\neg B)^* = \neg(B^*)$.
- $(B \vee C)^* = B^* \wedge C^*$.
- $(B \wedge C)^* = B^* \vee C^*$.

Vi vil vise at $A \Leftrightarrow \neg A^*$, og vi vil bruke induksjon.
Først må vi finne oss et passende tall i tilknytning til A som vi kan gjennomføre et induksjonsbevis for. Her vil vi velge *lengden* av A, det vil si antall symboler i A betraktet som et ord.
Da blir induksjonstarten at lengden av A er 1, og det svarer til $A = P_i$. I dette tilfellet er $\neg A^* = \neg\neg P_i \Leftrightarrow P_i = A$, så induksjonstarten er uproblematisk.
Induksjonskrittet er det litt værre med, hvis vi skal holde oss strengt til de former for induksjonskritt vi har sett på så langt. Vi får tre tilfeller

1. $A = \neg B$. I dette tilfellet er B et kortere utsagn enn A, og egentlig er B tre symboler kortere fordi A egentlig er på formen $(\neg B)$. Antagelsen i induksjonsprinsippet er at påstanden vi vil vise holder for utsagn hvor lengden er én mindre enn for den vi ser på. Vi vet imidlertid at hvis vi er midt inne i et induksjonsbevis, vil påstanden vår ikke bare holde for utsagn med ett symbol mindre, men underveis vil vi også ha vist den for

alle kortere utsagn. Vi kan derfor anta at egenskapen vår holder for B, det vil si at $B \Leftrightarrow \neg(B^*)$.
Da har vi

$$A = \neg B \Leftrightarrow \neg\neg B^* = \neg A^*.$$

2. $A = B \vee C$. I dette tilfellet er B og C kortere utsagn enn A, og som i tilfellet med negasjon vil vi anta at påstanden vår holder for B og C.
   Da holder også påstanden vår for A ved følgende utledning:

$$A = B \vee C \Leftrightarrow \neg B^* \vee \neg C^* \Leftrightarrow \neg(B^* \wedge C^*) = \neg A^*.$$

3. $A = B \wedge C$. Vi antar her at påstanden vår holder for B og C, og ser at da må påstanden vår holde for A.

Er så dette argumentet tilstrekkelig? Vi har fraveket formatet i induksjonsprinsippet, men vi har ikke fraveket intuisjonen bak. Tenk oss nå at vi i stedet for dominobrikker med tall bruker brikker med en eller to utløsningsmekanismer, og med utsagnsvariable eller sammensatte utsagn skrevet på seg. Vi kan utløse alle brikker med utsagnsvariable ved å trykke på en knapp. En brikke med et utsagn $\neg B$ på seg vil ha én utløsningsmekanisme og vil falle hvis brikken foran faller. Der setter vi selvfølgelig brikken med B på seg. En brikke med utsagn $B \vee C$ eller $B \wedge C$ på seg vil ha to utløsningsmekanismer, og vil falle hvis to brikker som står foran og ved siden av hverandre faller. Der setter vi selvfølgelig brikker for hhv. B og C. (Det er mulig at vi må bruke flere brikker for det samme utsagnet på forskjellige steder for fysisk å få plass.)
Poenget er at hvis trykker på knappen som får brikkene til utsagnsvariablene til å falle, så vil alle de andre brikkene falle etterhvert.
En litt avansert dominobrikkemodell tilsier altså at vi kan bevise en egenskap ved sammensatte utsagn hvor vi først viser egenskapen for utsagnsvariablene, og så viser at egenskapen bevares fra delutsagn til et utsagn. Vi skal komme tilbake til dette eksemplet i neste avsnitt.

I skolematematikken tar vi det for gitt at ethvert naturlig tall kan faktoriseres i primtall, og at dette kan skje på en og bare en måte. Egentlig er dette egenskaper ved tall som krever bevis, og vi skal se hvor induksjonsbevis kommer inn her.

Lemma 7.1 *Hvis $n \geq 2$ er et naturlig tall, finnes det primtall $p_1, \ldots, p_k$ slik at*

$$n = p_1 \cdot \ldots \cdot p_k.$$

Bevis.
Vi skal bruke et induksjonsbevis.
La n være gitt. Vi antar at påstanden holder for alle $m < n$.
Hvis n er et primtall, lar vi $p_1 = n$ og $k = 1$.
Hvis n ikke er et primtall, finnes det $m < n$ og $l < n$ slik at $n = k \cdot l$.
Ved induksjonsantagelsen finnes det primtall $p_1, \ldots, p_k$ og $q_1, \ldots, q_{k'}$ slik at

$$m = p_1 \cdot \ldots \cdot p_k$$

og

$$l = q_1 \cdot \ldots \cdot q_{k'}.$$

Da er

$$n = p_1 \cdot \ldots \cdot p_k \cdot q_1 \cdot \ldots \cdot q_{k'},$$

så n er et produkt av primtall.

Dette lemmaet sier oss ikke at oppspaltingen i primtall er entydig, bare at det finnes en faktorisering av ethvert tall. Poenget her er at vi har vi har bevist en egenskap $P(n)$ for alle n ved å bevise at for enhver k så vil $P(k)$ holde hvis vi antar at $P(m)$ holder for alle $m < k$. Det bildet vi vil bruke her er at hvis vi lar solen skinne på en rad snemenn, så vil en snemann holde seg så lenge det finnes en annen snemann lengere fremme som skygger for solen, mens en snemann som får direkte sollys vil smelte. Likevel vil snemennene smelte én for én, så vårt utvidede induksjonskritt svekker ikke konklusjonen.
Vi sammenfatter dette i

Induksjonsprinsippet II
La $P(n)$ være en påstand om et vilkårlig naturlig tall.
Anta at vi kan bevise for et vilkårlig tall k at

hvis $P(m)$ holder for alle $m < k$, så vil $P(k)$ holde.

Da kan vi konkludere med at $P(n)$ holder for alle naturlige tall n.

Vi skal gi en anvendelse til av induksjonsprinsipp II:

Lemma 7.2 *La n og m være to naturlige tall, og la a være største felles faktor i n og m.*
Da finnes det hele tall b og c slik at

$$a = b \cdot n + c \cdot m.$$

Bevis
Vi vil vise dette ved induksjon på $max\{n, m\}$, og beviset deles opp i tilfeller. Vi vil bruke induksjonsprinsipp II.

Så la n og m være gitt. Vi ser på to tilfeller, og vi antar at påstanden holder for alle tallpar u og v hvor $max\{u, v\} < max\{n, m\}$:

1. n er faktor i m eller m er faktor i n.
   Siden situasjonen er symmetrisk, kan vi anta at m er en faktor i n. Da er m største felles faktor i n og m, så $a = m = b \cdot n$ for en b. Setter vi $c = 0$ ser vi at lemmaet holder i dette tilfellet.

2. Ellers.
   Ved symmetri kan vi anta at $n > m$ (siden m er faktor i n hvis $m = n$).
   Da finnes det naturlige tall u og $v < m$ slik at $n = u \cdot m + v$, det vil si at $v = n - u \cdot m$.
   Vi har at den største felles faktoren i n og m er den samme som den største

felles faktoren a til m og v.
Ved induksjonsantagelsen finnes det da hele tall d og b slik at

$$a = d \cdot m + b \cdot v.$$

Setter vi inn for v i dette uttrykket får vi at

$$a = d \cdot m + b(n - u \cdot m),$$

som ved litt utregning gir at

$$a = b \cdot n + (d - u) \cdot m.$$

Ved å sette $c = d - u$ ser vi at påstanden holder.

Dette avslutter beviset.

Selv om lemmaene over først og fremst er tatt med for å illustrere bruken av induksjon, skal vi benytte anledningen til å fullføre beviset for entydig primtallsfaktorisering.
Fullføringen vil bestå i et lemma hvor beviset ikke illustrerer induksjon, og tilslutt et argument som igjen bruker induksjon.

Lemma 7.3 *La p være et primtall og anta at p er faktor i $n \cdot m$. Da er p en faktor i n eller en faktor i m (eller, siden dette er inklusiv eller, i begge).*

Bevis
Anta at p ikke er faktor i m. Vi skal vise at da må p være faktor i n.
Hvis p ikke er faktor i m, og siden p er et primtall, så er 1 den største felles faktoren i p og m. Ved forrige lemma finnes det da heltall b og c slik at

$$1 = b \cdot p + c \cdot m,$$

eller

$$b \cdot p = 1 - c \cdot m.$$

Siden vi har antatt at p er faktor i $n \cdot m$, vil vi også ha at p er faktor i $n \cdot m + n \cdot b \cdot p$.
Regner vi litt på dette får vi at

$$n \cdot m + n \cdot b \cdot p = n \cdot (m + 1 - c \cdot m) = n \cdot m \cdot (1 - c) + n.$$

Vi kan nå skrive

$$n = n \cdot m + n \cdot b \cdot p - n \cdot m \cdot (1 - c).$$

Siden p er faktor i alle leddene på høyre side, er også p faktor i n.
Dette viser lemmaet.

Teorem 7.1 *La $p_1, \ldots, p_n$ og $q_1, \ldots, q_m$ være primtall slik at*

$$p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_m.$$

Da inneholder de to listene de samme primtallene og hvert av primtallene forekommer i det samme antallet.

Dette viser at hvis vi ser bort fra rekkefølgen, er det bare en måte å primtallsfaktorisere et tall på.

Vi bruker induksjon på n.

Påstand
La $k \geq m$ og anta at p_n er faktor i $q_1 \cdot \ldots \cdot q_k$.
Da er $p_n = q_j$ for en $j \leq k$.

Bevis for påstanden
Vi bruker induksjon på k.
Hvis $k = 1$ må $p_n = q_1$.
Hvis $k > 1$, anta at påstanden holder for $k - 1$. Ved lemmaet over vil p_n enten være faktor i q_k og dermed være lik q_k, eller så er p_n en faktor i $q_1 \cdot \ldots \cdot q_{k-1}$.
I dette tilfellet kan vi bruke induksjonsantagelsen og får at $p_n = q_j$ for en $j < k$.
I begge tilfeller holder det vi vil vise, så påstanden er vist.

Fra påstanden får vi at $p_n = q_j$ for en $j \leq m$, og siden rekkefølgen på q-ene ikke spiller noen rolle kan vi anta at $p_n = q_m$.
Da har vi at

$$p_1 \cdot \ldots \cdot p_{n-1} = q_1 \cdot \ldots \cdot q_{m-1},$$

og ved induksjonsantagelsen i hovedbeviset får vi at disse faktoriseringene er like. Da må de opprinnelige faktoriseringene være like.
Dette avslutter beviset for entydig primtallsfaktorisering.

Oppgaver til avsnitt 7.1

Oppgave 7.1.1 Bruk induksjon på n til å vise at hvis a er et naturlig tall med n siffre, så får vi samme rest om vi deler a på 3 som om vi deler tverrsummen til a på 3.

Oppgave 7.1.2 Bruk induksjon på n til å vise at summen av de n første kvadrattallene blir

$$1 + 4 + \cdots + n^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}.$$

Oppgave 7.1.3 Vi har definert en funksjon $H : \mathbb{N}_0 \to \mathbb{N}_0$ ved rekursjon ved

- $H(0) = 2$.
- $H(1) = 3$.
- $H(n) = 3H(n-1) - 2H(n-2)$ når $n \geq 2$.

Bruk induksjon på n til å vise at $H(n) = 1 + 2^n$.

Oppgave 7.1.4 Vi har definert en funksjon G ved rekursjon ved

- $G(0) = 1$
- $G(1) = 4$

- $G(n) = 4G(n-1) - 4G(n-2)$ når $n \geq 2$.

Bruk induksjon til å vise at $G(n) = 2^n + n \cdot 2^n$ for alle $n \in \mathbb{N}_0$.

Oppgave 7.1.5 Vi minner om at $A \triangle B$ betegner den symmetriske differensen mellom mengdene A og B.
I denne oppgaven skal vi utvide definisjonen av symmetrisk differens til en operator på klassen av endelige, ikke-tomme mengder av mengder.

La $A_1, \ldots, A_n$ være mengder.
Vi definerer $\triangle\{A_1, \ldots, A_n\}$ ved rekursjon på n som følger

1. $\triangle\{A\} = A$
2. $\triangle\{A_1, \ldots, A_n, A_{n+1}\} = (\triangle\{A_1, \ldots, A_n\}) \triangle A_{n+1}$

Vis at $a \in \triangle\{A_1, \ldots, A_n\}$ hvis og bare hvis $\{i \mid a \in A_i\}$ har et odde antall elementer.
Forklar hvorfor dette betyr at parentessetting og rekkefølge ikke spiller noen rolle når vi ser på den symmetriske differensen mellom flere mengder.

Oppgave 7.1.6 Med *primtallsummen* til et tall n mener vi summen av alle primtallsfaktorene til n, hvor vi aksepterer gjentagelse.
Primtallsummen til 13 er 13, mens primtallsummen til 12 er $2 + 2 + 3 = 7$.
Bruk induksjon til å vise at primtallsummen til et tall n alltid er $\leq n$.
Hint: Hvis p er en primtallsfaktor i n, bruk induksjonsantagelsen for $\frac{n}{p}$.

7.2 Generell induksjon

Vi har jobbet oss gjennom eksempel 7.5 hvor vi definerte et sammensatt utsagn A^* ved rekursjon på oppbyggingen av det sammensatte utsagnet A, og så viste ved induksjon på lengden av utsagnet at $\neg A$ og A^* vil være ekvivalente.
Hvis vi går tilbake til det eksemplet, vil vi se at vi ikke brukte induksjonsantagelsen for alle mulige kortere utsagn, men for utsagn som er byggestener for A. Hvis $A = \neg B$, brukte vi induksjonsantagelsen for B, hvis $A = B \vee C$, brukte vi induksjonsantagelsen for B og for C, og hvis $A = B \wedge C$ brukte vi også induksjonsantagelsen for B og for C.

Dominoprinsippet for de naturlige tallene gir oss både at vi kan definere funksjoner ved rekursjon og at vi kan bevise påstander ved induksjon. I avsnittet om generell rekursjon påpekte vi at de naturlige tallene bare er ett av mange tilfeller hvor rekursive definisjoner gir mening. Ved akkurat den samme argumentasjonen kan vi se at vi kan bruke bevis ved induksjon når vi har en induktivt definert struktur som gir grunnlag for rekursivt definerte strukturer.
Tidligere i heftet har vi brukt induksjon over generelle strukturer uten å påpeke at det er det vi har gjort. Vi skal se på noen eksempler, og noen av disse eksemplene vil representere forbedrede bevis av påstander vi har kommet med tidligere.

Eksempel 7.6 Igjen skal vi se på utsagnslogikk, hvor vi har utsagnsvariablene $P_1, \ldots, P_k$, men hvor vi nå bruker alle bindeordene $\neg$, $\vee$, $\wedge$ $\rightarrow$ og $\leftrightarrow$.

Definisjon
Hvis A er et sammensatt utsagn, definerer vi $\Phi(A)$ ved rekursjon på oppbygningen av A ved

- $\Phi(P_i) = P_i$
- $\Phi(\neg B) = \neg\Phi(B)$ for alle B.
- $\Phi(B \wedge C) = \Phi(B) \wedge \Phi(C)$ for alle B og C.
- $\Phi(B \vee C) = \Phi(B) \vee \Phi(C)$ for alle B og C.
- $\Phi(B \rightarrow C) = \neg\Phi(B) \vee \Phi(C)$ for alle B og C.
- $\Phi(B \leftrightarrow C) = (\neg\Phi(B) \vee \Phi(C)) \wedge (\neg\Phi(C) \vee \Phi(B))$

Det vi har gjort her er å gi en rekursiv definisjonen av det utsagnet vi får når vi systematisk eliminerer alle forekomster av $\rightarrow$ og $\leftrightarrow$.
Det er selvfølgelig et viktig poeng at $\Phi(A)$ er utsagnslogisk ekvivalent med A for alle utsagn A. Dette er intuitivt opplagt, men bak intuisjonen ligger forståelsen av at dette holder fordi $\Phi(P_i)$ alltid er ekvivalent med P_i og fordi $\Phi(B) \Leftrightarrow B$ og $\Phi(C) \Leftrightarrow C$ vil medføre $\Phi(\neg B) \Leftrightarrow \neg B$, $\Phi(B \vee C) \Leftrightarrow B \vee C$, $\Phi(B \wedge C) \Leftrightarrow B \wedge C$, $\Phi(B \rightarrow C) \Leftrightarrow B \rightarrow C$ og $\Phi(B \leftrightarrow C) \Leftrightarrow B \leftrightarrow C$.

Eksempel 7.7 Vi kan fortsette eksemplet over, og anta at vi nå har begrenset oss til sammensatte utsagn hvor vi bare bruker bindeordene $\neg$, $\wedge$ og $\vee$.
Da vi studerte sammenhengen mellom sammensatte utsagn og strømkretser, så vi at vi kan lage en strømkrets for ethvert sammensatt utsagn hvor negasjonstegnet står helt inne ved utsagnsvariablene.
Vi sa at ethvert utsagn kan skrives om til et utsagn med denne egenskapen. Her vil vi gjenkjenne mønsteret fra forrige eksempel, vi definerer først resultatet $\Psi(A)$ av omskrivningen ved rekursjon på oppbyggingen av A, og deretter viser vi ved induksjon over oppbyggingen av A at A og $\Psi(A)$ er ekvivalente. Vi skal gjennomgå alle detaljene:

Definisjon

- La $\Psi(P_i) = P_i$
- La $\Psi(\neg P_i) = \neg P_i$
- La $\Psi(B \wedge C) = \Psi(B) \wedge \Psi(C)$
- La $\Psi(B \vee C) = \Psi(B) \vee \Psi(C)$
- La $\Psi(\neg\neg B) = \Psi(B)$
- La $\Psi(\neg(B \vee C)) = \Psi(\neg B) \wedge \Psi(\neg C)$
- La $\Psi(\neg(B \wedge C)) = \Psi(\neg B) \vee \Psi(\neg C)$

Dette er et sentralt eksempel, men det passer ikke helt inn i et formalisert mønster, ettersom vi i de to siste linjene bruker $\Psi(\neg B)$ og $\Psi(\neg C)$, mens $\neg B$ og $\neg C$ ikke er delutsagn av utsagnene $\neg(B \vee C)$ og $\neg(B \wedge C)$.
En måte å komme rundt dette problemet kan være å beskrive vår definisjon som en simultan definisjon av $\Psi(A)$ og $\Psi(\neg A)$ ved rekursjon på A. Hvis vi ser ser definisjonen vår på den måten, kan vi også vise følgende påstand ved induksjon på oppbyggingen av A:

Påstand
For hvert utsagn A vil $\Psi(A) \Rightarrow A$ og $\Psi(\neg A) \Rightarrow \neg A$.

Bevis

- Hvis $A = P_i$ har vi at $\Psi(A) = A$ og at $\Psi(\neg A) = \neg A$, og påstanden holder.
- Hvis $A = \neg B$ har vi fra induksjonsantagelsen at $\Psi(\neg A) = \Psi(\neg\neg B) = \Psi(B) \Leftrightarrow B \Leftrightarrow \neg A$.
  For å vise egenskapen for $\Psi(A)$ må vi dele beviset opp i ytterligere tilfeller ut fra hvordan B ser ut:
  - $B = \neg C$.
    Da er $\Psi(A) = \Psi(\neg\neg C) = \Psi(C) \Leftrightarrow C \Leftrightarrow A$
  - $B = C \wedge D$.
    Da er $\Psi(A) = \Psi(\neg(B \wedge C) = \Psi(\neg B) \vee \Psi(\neg C) \Leftrightarrow \neg B \vee \neg C \Leftrightarrow A$
  - $B = C \vee D$. Dette tilfellet håndteres som tilfellet over.
- Hvis $A = B \vee C$ har vi at $\Psi(A) = \Psi(B) \vee \Psi(C) \Leftrightarrow B \vee C = A$
  Videre har vi at $\Psi(\neg A) = \Psi(\neg B) \wedge \Psi(\neg C) \Leftrightarrow \neg B \wedge \neg C \Leftrightarrow \neg A$.
- Tilfellet $A = B \wedge C$ hånderes som tilfellet over.

Eksempel 7.8 Vi så på mengden av ord over et alfabet som et av eksemplene på induktivt definerte strukturer. Hvis vi ser på ord over alfabetet $\{0, 1\}$ som binære tall, har vi vist hvordan vi kan definere addisjon av binære tall ved rekursjon på de to tallene, vi definerte $+$ og $+'$, addisjon og addisjon hvor vi hadde tatt med oss en mente fra før. Egentlig bør vi da vise at denne konstruksjonen gjør det den skal ved induksjon på oppbyggingen av de to ordene.
For ikke å gi et for omfattende eksempel, skal vi anta at $+$ er definert. Vi definerer da en funksjon K ved

- $K(e) = e$
- $K(w0) = K(w)00$
- $K(w1) = K(w)00 + w01$

Det er sikkert ikke så lett å gjennomskue hva vi har definert her, så la oss se på $w = 101$, hvor 101 er binærtallet til 5, og la oss regne på det:

$$K(101) = K(10)00 + 1001 = K(1)0000 + 1001 = (00 + 01)0000 + 1001 = 011001.$$

Dette er et binært tall som svarer til 25. Et par eksempler til vil få oss til å tro at K står for 'kvadrat', noe som viser seg å stemme:

Påstand
Hvis w er en binærrepresentasjon av et tall n, hvor vi også betrakter e som en representasjon av 0, vil $K(w)$ være en binærrepresentasjon av n^2.

Bevis
Hvis $w = e$, er $K(w) = e$, så her har vi bestemt at påstanden skal holde.
Hvis $w = v0$ og v representerer n, vil w representere $2n$. Ved induksjonsantagelsen vil $K(v)$ representere n^2 og $K(w) = K(v)00$ vil representere $4n^2 = (2n)^2$. Påstanden holder altså også for w i dette tilfellet.
Hvis $w = v1$ og v representerer n, vil w representere $2n + 1$.
Da er $(2n + 1)^2 = 4n^2 + 4n + 1$.
Ved induksjonsantagelsen vil $K(v)$ representere n^2, så $K(v)00$ representerer $4n^2$. Samtidig vil $v01$ representere $4n+1$, og det betyr at $K(w)$ representerer $(2n+1)^2$ i dette tilfellet. Dermed holder også induksjonskrittet for $w = v1$.

Eksempel 7.9 Mengdelære er et nyttig verktøy i matematikken, og verd å læres av den grunn. Ved å bruke mengdelære og språket rundt mengdelæren, kan mange matematiske konstruksjoner og bevis gjøres mere presise og lettere tilgjengelige enn det som ellers ville vært mulig.
En annen grunn til å lære seg mengdelære er at den av mange oppfattes som matematikkens grunnlag, i prinsippet skal alle matematiske begreper kunne defineres i mengdelæren, og alle matematiske teoremer er egentlig å betrakte som teoremer i mengdelæren.
Vi har bare sporadisk gått inn på dette aspektet, men som leseren trolig vil huske har vi definert funksjoner og relasjoner som mengder av ordnede par, og lesere som har tatt utfordringer underveis vil huske at vi har definert et ordnet par som

$$(a, b) = \{\{a\}\{a, b\}\}.$$

Vi har også, som en utfordring, vist hvordan vi kan finne mengder som kan representere naturlige tall, ved å la $\emptyset$ representere 0 og la $x \cup \{x\}$ representere $n + 1$ hvis x representerer n.
Vi skal nå se på en klasse av mengder som vi vil kalle HF, *de hereditært endelige mengdene*. HF står for den engelske betegnelsen *hereditarily finite sets*.
Grunnen til at denne klassen er interessant, er at den i en viss forstand kan fungere som den globale datatypen. Dataobjekter er av natur endelige, og de er bygget opp induktivt ved bruk av ordnede sekvenser, mengdebyggeren og andre finitære konstruksjoner. Siden slike konstruksjoner kan formaliseres innen mengdelæren, kan vi godt si at alle dataobjekter kan representeres ved endelige mengder. Vi skal ikke utdype dette nærmere, men definere klassen HF induktivt og studere denne klassen. Mye av hensikten for oss er fortsatt å illustrere induksjon og rekursjon, men HF er av interesse ut over det.

Definisjon 7.1 Vi definerer HF som den minste klassen av mengder som oppfyller

- $\emptyset \in HF$
- Hvis $A \subseteq HF$ er endelig, så er $A \in HF$.

Vi skal nå definere en funksjon *numb* ved rekursjon på HF. *numb* er kortformen av *number*, og skal tilordne et naturlig tall til hver hereditært endelig mengde. *num* vil ha en invers, som vil være en opptelling av HF.
Hvis A er en endelig mengde, og $f : A \to \mathbb{Z}$, lar vi

$$\sum_{a \in A} f(a)$$

betegne summen av alle verdier $f(a)$ når a gjennomløper A.

Definisjon 7.2 La $A \in HF$. La

$$numb(A) = \sum_{a \in A} 2^{numb(a)}.$$

Hvordan kan vi påstå at dette er en grei definisjon? Det kan være to problemer:

1. Vi sa at dette er en rekursiv definisjon, og det betyr at vi antar at $numb(a)$ er definert når $a \in A \in HF$.
   Dette er uproblematisk siden vi må ha at $A \subseteq HF$ før vi kan få at $A \in HF$, så alle elementene i A er med i HF på et tidligere tidspunkt.

2. Vi har ikke gitt basis $\emptyset \in HF$ spesialbehandling.
   Vi kunne godt ha gjort det, og satt $numb(\emptyset) = 0$. Men hvis vi betrakter $\emptyset$ som en endelig delmengde av HF, får vi at $numb(\emptyset)$ er summen over den tomme mengden, og en tom sum vil vi normalt gi verdien 0.

Lemma 7.4 *$numb : HF \to \mathbb{N}_0$ er både injektiv og surjektiv.*

Bevis
For å vise at *numb* er injektiv, må vi vise at hvis A og B er i HF og $numb(A) = numb(B)$, så er $A = B$. Dette viser vi ved induksjon over A og B:
Anta som en induksjonsantagelse at for alle $a \in A \cup B$ og for alle $b \in A \cup B$, hvis $numb(a) = numb(b)$, så vil $a = b$.
Anta videre at $numb(A) = numb(B)$. Det betyr at $\sum_{a \in A} 2^{numb(a)} = \sum_{b \in B} 2^{numb(b)}$.
To summer av forskjellige 2-er potenser vil være like hvis og bare hvis det er de samme leddene som inngår.
Det betyr at $\{numb(a) \mid a \in A\} = \{numb(b) \mid b \in B\}$.
Ved induksjonsantagelsen vil vi ha at vi for alle $a \in A$ og alle $b \in$ har at $numb(a) = numb(b) \Rightarrow a = b$. Det betyr at $A = B$, og vi er fremme. For å vise at *numb* er surjektiv, må vi bruke induksjon over $\mathbb{N}_0$, det vil si at vi må vise at det for alle $n \in \mathbb{N}_0$ finnes en $A \in HF$ slik at $numb(A) = n$.
For $n = 0$ lar vi $A = \emptyset$.
La $n > 0$, og anta at påstanden holder for alle $m < n$.
Siden ethvert positivt heltall kan skrives som en sum av 2-er potenser på en og

bare en måte, finnes det én endelig mengde $C \subseteq \mathbb{N}_0$ slik at $n = \sum_{m \in C} 2^m$, og hvis $m \in C$ vil vi ha at $m < n$.
Da kan vi bruke induksjonsantagelsen og det at *numb* er injektiv, og får at det finnes en endelig mengde $A \subseteq HF$ slik at $C = \{numb(a) \mid a \in A\}$. Da er $n = numb(A)$.

Dette viser at i en viss forstand er $\mathbb{N}_0$ og HF like store, og vi kunne tilsynelatende like gjerne brukt $\mathbb{N}_0$ som vår globale datatype, alt kunne jo oversettes via *numb* i alle fall. I oppgave 7.2.1 skal vi se hvorfor dette ikke er så opplagt likevel.

Oppgaver til avsnitt 7.2

Oppgave 7.2.1 Vi har hevdet at vi kan la $\{\{a\}\{a,b\}\}$ være den mengdeteoretiske tolkningen av det ordnede paret av (a,b).

a) La $OP(n,m) = 2^{2^n} + 2^{2^n+2^m}$
Vis at hvis $n = numb(a)$ og $m = numb(b)$, så er $OP(n,m) = numb(\{\{a\},\{a,b\}\})$.
Vis at $OP(2,3) = 4112$.

b) Vis at hvis $OP(n,m) = OP(k,l)$, så er $n = k$ og $m = l$.

c) Representerer OP en fornuftig måte å representere ordnede par på? Gi en drøfting av dette spørsmålet.

Oppgave 7.2.2 Vi definerer mengden av *kjedebrøker* ved induksjon.

- 0 er en kjedebrøk.
- Hvis K er en kjedebrøk og $n \in \mathbb{N}$, vil

$$\frac{1}{n+K}$$

  også være en kjedebrøk.

Kjedebrøken vil være selve uttrykket, og vi vil derfor snakke om kjedebrøk og dens *numeriske verdi*. Det er meningen at du skal bruke induksjon til å løse oppgavene under.

a) La K være en kjedebrøk med numerisk verdi q. Vis at $q \in \mathbb{Q}$.

b) la K og q være som over.
Vis at $0 \leq q \leq 1$.

c) La q være et rasjonalt tall slik at $0 \leq q \leq 1$.
Vis at da finnes det en kjedebrøk K med numerisk verdi q.
Hint: Skriv q som en brøk og bruk induksjon på nevneren til q.

Oppgave 7.2.3 Parenteser er viktige i algebra, vi bruker dem til å gruppere deler av uttrykk slik at vi vet hvilke deluttrykk som skal multipliseres med hverandre ol. I utsagnslogikken brukte vi parenteser for å kunne tolke hvordan bindeordene skal brukes. I programmering er også bruk av parenteser viktig. I den sammenhengen bruker man forskjellige typer parenteser, ofte kamuflert som par

$$\text{begin}\{---\}\ \text{delprogram}\ \text{end}\{---\}$$

eller som andre former for markering av en begynnelse og en matchende slutt. Det er derfor viktig å kunne avgjøre om parenteser er brukt på en riktig måte. I denne oppgaven skal vi definere hva vi mener med *korrekte parentesuttrykk*, hvor vi ser bort fra all mellomtekst. Vi skal arbeide med ord i alfabetet $\{(,)\}$, og vi (det vil si dere) skal vise en karakterisering av mengden av korrekte parentesuttrykk som gjør det enklere å sjekke maskinelt (eller manuelt under korrekturlesning) om et parentesuttrykk er lovlig eller ikke.
Den interesserte leser kan utvide vår definisjon til å omfatte flere former for parenteser, for eksempel hakeparenteser og klammeparenteser, og finne en tilsvarende karakterisering av mengden av korrekte slike uttrykk.

1. Det tomme ordet er et korrekt parentesuttrykk.
2. Hvis u og v er korrekte parentesuttrykk, er (u) og uv også korrekte parentesuttrykk.

a) Vis at hvis w er et korrekt parentesuttrykk, så vil w ha like mange venstre- som høyreparenteser.

b) Vis at hvis w er et korrekt parentesuttrykk, og vi deler w opp i to delord

$$w = uv$$

så vil u ha minst like mange venstre- som høyreparenteser.

c) [u] Vis at hvis w er et ord i alfabetet $\{(,)\}$ og konklusjonene i a) og b) holder for w, så er w et korrekt parentesuttrykk.
Hint: Bruk induksjon på lengden av w. Vis at en av tre muligheter vil gjelde

1. $w = e$ (det tomme ordet)
2. $w = (w_1)$ hvor w_1 oppfyller a) og b).
3. $w = w_1w_2$ hvor ingen av w_1 og w_2 er tomme og både w_1 og w_2 oppfyller a) og b).

Strategien for å avgjøre om vi har satt parentesene på lovlig måte er da å telle parenteser fra venstre mot høyre, vi teller oppover for hver venstreparentes (og nedover for hver høyreparentes). Hvis vi ender ut med 0 og aldri er innom negative tall, er parentessettingen i tråd med reglene for god parentesføring.

Oppgave 7.2.4 (u) La $*$ være et objekt, uten at det betyr noe hva skalgs objekt det er.
Vi definerer familien av *binære trær* induktivt ved

1. $*$ er et binært tre.
2. Hvis T_0 og T_1 er binære trær, er (T_0, T_1) et binært tre.

Bildet skal være at vi har et forgreningspunkt nede ved roten, hvor T_0 er fortsettelsen langs den venstre grenen og T_1 er fortsettelsen langs den høyre grenen. Ytterst sitter bladene, representert ved $*$.

Hvis $\sigma = (a_1, \ldots, a_n)$ er en ordnet sekvens av 0-er og 1-ere og T er et binært tre, definerer vi relasjonen

$$\text{`}\sigma \text{ er en gren i } T\text{'}$$

ved

1. e er en gren i $*$.
2. $(0, a_2, \ldots, a_n)$ er en gren i (T_0, T_1) hvis $(a_2, \ldots, a_n)$ er en gren i T_0.
3. $(1, a_2, \ldots, a_n)$ er en gren i (T_0, T_1) hvis $(a_2, \ldots, a_n)$ er en gren i T_1.

La $\{a_n\}_{n\in\mathbb{N}}$ være en uendelig følge av 0-er og 1-ere og la T være et binært tre. Vis at det finnes en og bare en n slik at $(a_1, \ldots, a_n)$ er en gren i T.
Hint: Bruk induksjon på oppbyggingen av T. Det kan være lurt å vise en sterkere egenskap, nemlig at det for alle $m \geq 1$ finnes en og bare en $n \geq m$ slik at $(a_m, \cdots, a_n)$ er en gren i T.

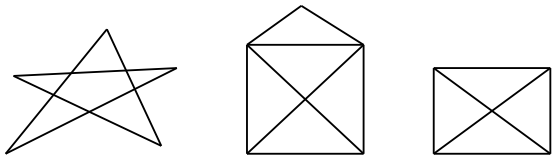
7.3 Utfordringer

7.3.1 Eulerstier og Eulersykler

Definisjon 7.3 En *graf* vil i dette avsnittet bestå av en endelig mengde punkter (grafens *hjørner*) og en eller flere linjer som forbinder par av punkter (grafens *kanter*).
En *sti* i en graf vil være en sekvens $P_1, K_1P_2 \cdots P_{n-1}K_{n-1}P_n$ hvor hver P_i er et hjørne og hver K_i er en kant mellom P_i og P_{i+1}.
En sti kalles en *Eulersti* hvis alle kantene i grafen benyttes en og bare en gang. En Eulersti kalles en Eulersykel hvis den starter og slutter i det samme hjørnet.

Vi skal finne ut av når det finnes en Eulersti og når det finnes en Eulersykel. Det ene kravet som opplagt må være tilfredstilt er at grafen må være *sammenhengende*, det vil si at alle par av punkter kan forbindes med en sti. Vi skal se

på noen eksempler, for å finne et par andre nødvendige betingelser:



I eksemplet til venstre har vi en Eulersykel. I eksemplet i midten har vi en Eulersti, men ikke en Eulersykel. Grunnen til at vi ikke har en sykel er at det kommer et odde antall kanter fra to av hjørnene. Hvis et hjørne hverken er begynnelse eller slutt på en sti, må stien benytte seg av et partall antall kanter ut fra det aktuelle hjørnet, stien kommer både inn langs en kant og går ut langs en annen et visst antall ganger. Dette argumentet forteller oss at vi må starte og slutte stien i de to nederste hjørnene, men litt prøving og feiling vil vise at da er det mulig å tegne stien.
I eksemplet til høyre har vi ikke engang en Eulersti. Grunnen til det er at vi har fire hjørner som det går tre kanter ut fra. For hvert av disse hjørnene må en sti enten begynne eller slutte der. Siden vi bare kan begynne ett sted og slutte ett annet sted, er dette umulig.

Vi skal bruke induksjon på antall kanter i en graf, og vise Eulers teorem. Merk at siden hver kant vil ha to ender, er summen av antall kanter ut fra hjørnene et partall, nemlig det dobbelte av antall kanter i grafen.

Teorem 7.2 *La G være en endelig sammenhengende graf.*

a) Hvis antall kanter ut fra hvert hjørne alltid er et partall, finnes det en Eulersykel.

b) Hvis det finnes høyst to hjørner med et odde antall kanter ut fra seg, finnes det en Eulersti.

Bevis
Vi bruker induksjon på antall kanter. Det kan lønne seg å illustrere de enkelte delene av beviset med figurer.
Hvis det bare er en kant har vi to muligheter:

1. Vi har ett punkt P og en kant K fra P til P. Da utgjør denne kanten en sykel.

2. Vi har to punkter P og Q og en kant mellom dem. Da utgjør denne kanten en Eulersti.

Anta nå at vi har $n+1$ kanter, og at teoremet holder for grafer med $\leq n$ kanter.
Vi viser a) og b) hver for seg.
Det blir en del spesialtilfeller:

a) Hvis vi tar bort en kant K har vi to muligheter:

   1. K går fra et punkt P til seg selv.
      Da oppfyller den reduserte grafen fortsatt betingelsen i a), så den grafen har en Eulersykel. Legger vi til kanten K til den sykelen, får vi en Eulersykel for den gitte grafen.
   2. K går fra et punkt P til et annet punkt Q.
      Da reduserer vi antall kanter ut fra P med en og antall kanter ut fra Q med en. Da er betingelsen i b) oppfylt, og det finnes en Eulersti fra P til Q i den reduserte grafen. Legger vi til kanten K får vi en sykel i den opprinnelige grafen.

   For ingen av disse tilfellene er det noen fare for at den reduserte grafen ikke skal bli sammenhengende. I tilfelle 2. ville det resultert i at vi hadde to grafer hvor det totale antall kanter inn mot hjørner var oddetall, og det er umulig.

b) Vi velger nå et hjørne P som det går odde antall hjørner ut fra, og vi fjerner en av kantene K som går ut fra P til et punkt Q. Vi deler argumentet opp i fire tilfeller:

   1. Q er det andre punktet med odde antall kanter, og den reduserte grafen er sammenhengende.
      Da har den reduserte grafen en Eulersykel fra Q til Q, og hekter vi på K får vi en Eulersti fra P til Q.
   2. Q er det andre punktet med odde antall kanter og den nye grafen er ikke sammenhengende.
      Da består den nye grafen av to separate deler G_1 med Q i og G_2 med P i.
      Både G_1 og G_2 vil oppfylle betingelsen i a), så de har Eulersykler. Vi lager en Eulersti fra P til Q ved først å følge en Eulersykel i G_2 fra P til P, så bruke K og tilsist følge en Eulersykel i G_1 fra Q til Q.
   3. Q er ikke det andre punktet med odde antall kanter og den reduserte grafen er sammenhengende.
      Da vil den reduserte grafen oppfylle betingelsen i b), så det finnes en Eulersti i den reduserte grafen som starter i Q. Legger vi til K får vi en Eulersykel for den opprinnelige grafen som starter i P.
   4. Q er ikke det andre hjørnet med et odde antall kanter, men den reduserte grafen er ikke sammenhengende.
      Som over består den reduserte grafen av to separate deler, en graf G_1 hvor Q er et hjørne og en graf G_2 hvor P er et hjørne.
      Nå vil G_2 oppfylle betingelse b), mens G_1 vil oppfylle betingelse a).

Vi limer sammen sykler, stier og K til en Eulersti for den opprinnelige grafen på tilsvarende måte som før.

Dette avslutter beviset.

7.3.2 Kubens fordobling og tredeling av vinkler

Anta at vi har gitt et rettvinklet aksekors med x-akse og y-akse hvor vi har avsatt en enhetslengde langs den ene aksen. Anta at vi ut over dette har til disposisjon en passer og en linjal uten merker på, det vil si de klassiske hjelpemidlene i konstruksjonsoppgaver. Vi skal se på et par klassiske problemer omkring konstruksjon: Kubens fordobling og vinkelens tredeling.
Sidekanten i en kubus med volum 2 vil være $\sqrt[3]{2}$, og hvis vi kan vise at det er umulig å konstruere et linjestykke med lengde $\sqrt[3]{2}$ ved hjelp av passer og linjal, har vi bevist at konstruksjonsoppgaven om kubens fordobling er uløselig.
Alle som har lært litt om bruk av passer og linjal, vet at det er mulig å konstruere en vinkel på 30°. Hvis det i tillegg hadde vært mulig å dele en vinkel i tre like store deler ved passer og linjal, ville vi kunnet konstruere en vinkel på 10°. Tenker vi kontrapositivt er det altså tilstrekkelig å vise at det er umulig å konstruere en vinkel på 10° for å vise at problemet med tredelingen av vinkelen er uløselig. Siden vi kan tegne enhetssirkelen ved hjelp av passeren, er det nødvendig og tilstrekkelig å vise at vi ikke kan kan konstruere et linjestykke med lengde $\sin(10°)$ for å vise at vi ikke kan konstruere 10°.

Vi skal gi en induktiv definisjon av en punktmengde G på den reelle tallinjen. Vi vil kalle G *mengden av geometriske punkter*, og vi skal vise at det er nøyaktig linjestykker med lengder som ligger i G det er mulig å konstruere ved hjelp av passer og linjal. Punktene i G vil vi finne ved gjentatte ganger å løse annengradslikninger. Den siste spikeren i kista for de to problemene våre er da at det er umulig å løse tredjegradslikninger ved hele tiden å løse annengradslikninger, vi kan ikke få en kubikkrot ved bare å trekke kvadratrøtter.

Definisjon 7.4 Vi definerer klassen av geometriske punktmengder som den minste klassen av delmengder av $\mathbb{R}$ som oppfyller

- $\mathbb{Q}$ er en geometrisk punktmengde.
- Hvis $K \subseteq \mathbb{R}$ er en geometrisk punktmengde, og $c \in K$ er positiv, så er

$$K_1 = \{a + b\sqrt{c} \mid a \in K \wedge b \in K\}$$

  også en geometrisk punktmengde.

Vi lar G være mengden av alle reelle tall som ligger i minst en geometrisk punktmengde, og vi vil kalle elemeentene i G for *geometriske punkter*

Vi kan gi en direkte definisjon av G:

Lemma 7.5 *Mengden G av geometriske punkter kan gis en alternativ definisjon ved:*

- *Hvis $a \in \mathbb{Q}$ er $a \in G$*
- *Hvis $a \in G$ og $b \in G$ er $a + b$ og ab i G*
- *Hvis $a \in G$ er positiv, er $\sqrt{a} \in G$.*

Bevis
Beviset overlates til leseren som Oppgave 7.4.2.

Lemma 7.6 *Hvis K er en geometrisk punktmengde, er K lukket under de fire regningsartene, det vil si at hvis a og b er i K vil $a + b \in K$, $a - b \in K$, $ab \in K$ og hvis i tillegg $b \neq 0$ vil vi ha at $\frac{a}{b} \in K$.*

Bevis
Vi beviser dette ved induksjon på oppbyggingen av geometriske punktmengder. Induksjonstarten er at $\mathbb{Q}$ er lukket under de fire regningsartene, mens induksjonskrittet ble gitt som oppgave 4 i annet innleveringsett.
Det eneste krevende punktet er å vise at hvis $K_1 = \{a + b\sqrt{c} \mid a \in K \wedge b \in K\}$, så er $\frac{1}{a+b\sqrt{c}} \in K_1$ når $a + b\sqrt{c}$ er det.
Vi multipliserer oppe og nede med $a - b\sqrt{c}$, og bruker at K er lukket under de fire regningsartene.

En viktig grunn til å kalle punktene i G for geometriske punkter, er at det i en viss forstand er disse punktene vi kan konstruere ved hjelp av passer og linjal.

Lemma 7.7 *La $a \in G$. Da kan vi konstruere et linjestykke med lengde $|a|$.*

Bevis
Vi bruker induksjon basert på den alternative definisjonen av G gitt i Lemma **??**.
Det er lett å se at vi kan konstruere linjestykker med lengde a for alle positive rasjonale tall a.
Hvis vi kan konstruere et linjestykke med lengde a og et annet med lengde b, kan vi legge linjestykkene etterhverandre og få et linjestykke av lengde $a + b$.
For å konstruere et linjestykke av lengde ab bruker vi en standard metode: Tegn en vinkel ut fra et punkt P. Avsett ett punkt A i avstand a fra P langs det ene vinkelbenet. og to punkter C og B i avstand hhv. 1 og b fra P langs det andre. Trekk linjen fra C til A og parallellforskyv den til en linje gjennom B. La D være skjæringspunktet mellom denne linjen og det første vinkelbenet. Da vil ab være avstanden fra P til D.
For å konstruere kvadratrøtter utnytter vi Pythagoras. Konstruksjonsmåten vil være avhengig av hvor stor a er. Eksempelvis, hvis $a > 1$ konstruerer vi en rettvinklet trekant hvor lengden på hypotenusen er $\frac{a+1}{2}$ og lengden av den ene kateten er $\frac{a-1}{2}$. Da er $\sqrt{a}$ lik lengden av den andre kateten.
Husk at vi ikke trenger en felles konstruksjonsmåte som gjelder for alle tall a, bare at det for hvert tall a finnes en måte å konstruere $\sqrt{a}$ på, hvor metoden kan være avhengig av a.
Dette avslutter det skissemessige beviset.

Omvendingen gjelder også

Lemma 7.8 *Anta at vi kan konstruere punktet (x, y) ved hjelp av passer og linjal. Da er x og y geometriske punkter.*

Bevis
Vi konstruerer punkter ved å finne skjæringspunktet mellom linjer, mellom sirkler eller mellom en sirkel og en linje. Poenget er at når vi trekker en linje, så trekker vi den mellom to punkter vi allerede har konstruert. Når vi slår en sirkel, har den et sentrum som allerede er konstruert, og radius er avstanden mellom to punkter vi allerede har konstruert.
Hvis (a, b) og (u, v) er geometriske punkter, er avstanden mellom punktene

$$\sqrt{(a-u)^2 + (b-v)^2}$$

som også er et geometrisk punkt.
For å finne skjæringspunktet mellom to linjer, løser vi to førstegradslikninger med to ukjente, og er alle koeffisientene i G vil løsningene ligge i G.
For å finne skjæringspunktet mellom en linje og en sirkel, må vi løse to likninger med to ukjente hvor den ene er av første grad og den andre av annen grad. Hvis vi løser førstegradslikningen med hensyn på y og så setter løsningen inn for y i 2. gradslikningen, får vi en 2. gradslikning i x. Hvis alle koeffisientene er i G, er løsningen av denne 2. gradslikningen også i G, og vi finner at skjæringspunktene mellom linjen og sirkelen er i G.
For å finne skjæringspunktet mellom to sirkler må vi løse to 2. gradslikninger med to ukjente. La oss se på sirklene med sentrum i hhv. (a, b) og (c, d) og med radier hhv. r og s. Vi antar at a, b, c, d, r og s er i G.
Likningene for sirklene er

$$(x-a)^2 + (y-b)^2 = r^2$$

og

$$(x-c)^2 + (y-d)^2 = s^2.$$

Regner vi litt på dette får vi skrevet likningene om til

$$x^2 - 2ax + a^2 + y^2 - 2by + b^2 = r^2$$

og

$$x^2 - 2cx + c^2 + y^2 - 2dy + d^2 = s^2.$$

Tar vi differensen mellom den nederste og den øverste likningen, får vi

$$(2a - 2c)x + (2b - 2d)y = s^2 - r^2$$

som er en førstegradslikning med koeffisienter i G. Da har vi redusert oppgaven til å finne skjæringspunktene mellom en linje og en sirkel, og disse kan vi finne i G.
Dette avslutter det skissemessige beviset av lemmaet.

Vi skal avslutte dette avsnittet med å vise at $\sqrt[3]{2}$ ikke er i G. Det betyr at dette tallet ikke kan konstrueres med passer og linjal.

Lemma 7.9 *La K være en geometrisk punktmengde. Da er $\sqrt[3]{2} \notin K$.*

Bevis
Vi viser dette ved induksjon på definisjonen av geometriske punktmengder.
Hvis $K = \mathbb{Q}$, er påstanden kjent.
Anta at $\sqrt[3]{2} \notin K$, men at det finnes $a, b, c \in K$ slik at

$$\sqrt[3]{2} = a + b\sqrt{c}.$$

Da kan ikke $\sqrt{c}$ være i K og vi må ha at $b \neq 0$. Da har vi videre, ved å opphøye begge sider i 3, at

$$2 = a^3 + 3a^2b\sqrt{c} + 3ab^2c + b^3c\sqrt{c}.$$

Hvis denne likningen kan løses mht. $\sqrt{c}$ vil vi få at $\sqrt{c} \in K$, så det er umulig. Men det må bety at $3a^2b + b^3c = 0$, eller at $b^2c = -3a^2$. Setter vi $-3a^2$ for b^2c inn i den opprinnelige likningen får vi at leddene med $\sqrt{c}$ blir borte, og at

$$2 = a^3 - 9a^3 = -8a^3,$$

og dette gir at $\sqrt[3]{2} = -2a \in K$, noe som er mot forutsetningen.
Dette beviser induksjonskrittet.

Vi har nå gått igjennom hovedelementene i beviset for

Teorem 7.3 *Det er umulig å konstruere et linjestykke med lengde $\sqrt[3]{2}$ ved hjelp av passer og linjal.*

Ved å bruke at vi kjenner en formel for $\sin(3v)$ uttrykt i $\sin(v)$, og at vi vet at $\sin(30°) = \frac{1}{2}$, kan vi finne en tredjegradslikning med $sin(10°)$ som løsning. Det er mulig å vise at denne likningen ikke har noen løsning i noen geometrisk punktmengde. Vi bruker induksjon på oppbygningen av den geometriske punktmengden. Dette kan den ivrige leseren forsøke seg på selv. Et problem vil være å vise at den aktuelle likningen ikke har rasjonale løsninger.

7.4 Blandede oppgaver

Oppgave 7.4.1 (u) I beviset for teorem 3.1 brukte vi en påstand som vi i bemerkningen under erkjente kunne trenge et mer stringent bevis. I oppgave 6.4.1 ga vi en mer stringent definisjon av $UV(\phi)$. Bruk denne definisjonen og generell induksjon til å bevise påstanden vi tok for gitt i beviset av teorem 3.1

Oppgave 7.4.2 Definer mengden H som den minste mengden av punkter som oppfyller

- Hvis $a \in \mathbb{Q}$ er $a \in H$
- Hvis $a \in H$ og $b \in H$ er $a + b$ og ab i H
- Hvis $a \in H$ er positiv, er $\sqrt{a} \in H$.

Vis at $H = G$, hvor G er mengden av geometriske punkter.
Hint: Vis ved induksjon på oppbyggingen av geometriske punktmengder K at $K \subseteq H$ for alle slike K.
Vis omvendt at det for alle endelige delmengder E av H, så finnes det en geometrisk punktmengde K slik at $E \subseteq K$. Det lønner seg å bruke induksjon på det totale antall ganger vi har trukket kvadratrøtter for å konstruere elementene i E.

Kapittel 8

Noen løsningsforslag/fasitsvar

Etter ønske fra kursdeltagerne suppleres heftet med fasit for noen av oppgavene. Der det er aktuelt, gir vi også mer utfyllende forslag til hvordan oppgaven kan løses.
Teksten i dette kapitlet bør ikke oppfattes som komplette løsningsforslag. Det finnes nemlig ingen "eneste" måte å løse de fleste oppgaver på rent tekstlig, måten man formulerer en løsning på kan godt være individuell.
Når fasit eller løsningsforslag ikke er gitt for en oppgave kan det skyldes en eller flere av følgende:

1. Det er for arbeidskrevende å skrive løsningen i tekstbehandlingsystemet/har ikke prioritert tid til å gi løsningen.
2. Oppgaven er en 'oppdag verden selv'-oppgave hvor det å gi løsningen ødelegger den pedagogiske hensikten forfatteren hadde med å gi den.
3. Oppgaven er en utfordringsoppgave hvor det å gi løsningen ødelegger utfordringsaspektet.

Kapittel 1

1.1.1
$A = \{-1, 0, 1\}$, $B = \{0, 1\}$, $C = A$, $D = [-1, 1]$, $E = A$.
Vi har $A = C = D$ og de to andre er forskjellige fra A og fra hverandre.

1.1.2
B er ekte inneholdt i $A = C = E$ som er ekte inneholdt i D.
Da er B ekte inneholdt i D.

1.1.3
$A \subseteq B$, $A \subseteq C$, $A \subseteq D$, $B \subseteq D$ og $C \subseteq D$.

Alle disse inklusjonene er ekte, og ingen andre inklusjoner holder.

1.1.4
$36 \in A$, $21 \notin B$, $27 \in B$, $6 \in D$, $s \notin C$, $9 \notin C$ og $27 \notin A$. Eksempelvis skyldes den siste påstanden at 27 ikke er delelig med 18.

1.1.5
$120 \in A$ er feil fordi 120 ikke er et kvadrattall.
$121 \in A$ er rett fordi $121 = 11 \cdot 11$ og $121 > 0$.
$144 \notin A$ er feil fordi 144 er et positivt kvadrattall.
$17 \notin A$ er rett fordi 17 ikke er noe kvadrattall.

1.1.8
$B \subseteq A$ betyr at hver gang vi lar $x \in B$, så vil $x \in A$. Vi kan aldri finne $x \in \emptyset$, så hver gang vi gjør det (dvs. aldri) har vi også funnet en $x \in A$.

1.2.1

1. Aer mengden av positive kubikktall.
2. B er mengden av løsninger av den gitte tredjegradslikningen. En annen måte å skrive den på er $(z+1)^3 = 0$, så B er mengden med -1 som eneste element.
3. C er den tomme mengden, fordi likningen $a^2 + a + 1 = 0$ ikke har heltallsløsninger.
4. D er mengden av punkter på linjen i planet som går gjennom punktene $(1, 0)$ og $(0, 1)$.
5. E er mengden av punkter på parabelen gitt ved likningen $y = x^2$.

1.2.2

1. Sann fordi $21+2 = 23$ som er et primtall.
2. Sann fordi alle tall på formen n^4 spesielt er kvadrattall.
3. Sann fordi $\emptyset$ er en delmengde av alle mengder, og dermed element i alle potensmengder.
4. Usann fordi $4^2 < 24$, mens $4^3 > 24$.

1.2.3
$\mathcal{P}(\emptyset) = \{\emptyset\}$ og har altså $\emptyset$ som element. $\emptyset$ har ingen elementer, så disse to mengdene er forskjellige.

1.2.4
Hvis $C \in \mathcal{P}(A)$ er $C \subseteq A$, så $C \subseteq B$ og $C \in \mathcal{P}$.
Hvis $\mathcal{P}(A) \subseteq \mathcal{P}(B)$ er $A \in \mathcal{P}(B)$. Det betyr at $A \subseteq B$.

1.3.1
Den symmetriske differensen av A og B består av de objektene som er element i én av mengdene, men ikke i begge.

1.3.2

1. $A \cap B = \{3, 9\}$, $A \cup B = \{3, 5, 8, 9, 6, 11\}$ og $A \setminus B = \{5, 8\}$.

2. $A \cap B = \{x \in \mathbb{R} \mid -1 < x < 1\}$, $A \cup B = \mathbb{R}$ og $A \setminus B = \{x \in \mathbb{R} \mid x \geq 0\}$.

3. $A \cap B = \emptyset$, $A \cup B = \{2, 3, 4, 5, \ldots\}$ og $A \setminus B = A$.

4. $A \cap B$ er første kvadrant i planet inklusive origo og punktene på de positive delene av aksene
   $A \cup B$ er hele planet med unntak av punktene som ligger ekte innenfor tredje kvadrant.
   $A \setminus B$ er venstre del av øvre halvplan, hvor punktene langs den negative delen av x-aksen får være med, mens ingen punkter på den ikke-negative delen av Y-aksen får være med.

1.3.3

1. C.

2. D.

3. D.

1.3.4
JA!

1.3.5
Siden $A \cap B \subseteq A$ vil $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(A)$.
Av tilsvarende grunn vil $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(B)$, så $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(A) \cap \mathcal{P}(B)$. Omvendt, hvis $C \in \mathcal{P}(A) \cap \mathcal{P}(B)$, vil $C \subseteq A$ og $C \subseteq B$, så $C \subseteq A \cap B$. Det betyr at $C \in \mathcal{P}(A \cap B)$.

1.3.6
Det er bare når $A \subseteq B$ eller $B \subseteq A$, for ellers vil ikke $A \cup B \in \mathcal{P}(A) \cup \mathcal{P}(B)$.

1.4.3

1. Feil.

2. Feil.

3. Riktig.

4. Feil.

Kapittel 2

2.1.1
1. er et spørsmål, 2. er et utsagn, 3. er et fromt ønske, men ikke noe utsagn, 4. er et utsagn som er usant, 5. er et utsagn, 6. er et sammensatt utsagn og 7. er et utsagn som ikke er helt spesifisert, ettersom verdien på x kan variere, og dermed varierer sannhetsverdien til utsagnet.

2.1.2

1. er situasjonsavhengig siden vi på den ene siden kan snakke om de tre brødrene fra eventyret mens vi på den andre siden for eksempel kan snakke om tre spillere på samme fotball-lag. Derimot er 2. et utsagn som er sant. 3. er situasjonsavhengig mens 4. er sann uansett. 5. er et utsagn som ikke er sant, 6. er sant 7. er usant, 8. er sant og 9. er sant. Ingen av disse er situasjonsavhengige.

2.2.2

Det finnes to måter å gi en sannhetsverdi til P_1 på. For hver av disse finnes det to måter å gi en sannhetsverdi til P_2 på. Det gir tilsammen fire måter. For hver av disse fire måtene finnes det to måter å gi en sannhetsverdi til P_3 på. Hvis $n > 3$ kan vi fortsette tankerekken og finner at det er 2^n måter å tildele sannhetsverdier til alle utsagnsvariablene på.

Når det gjelder neste del, tenker vi på samme måten, vi har 2^n foskjellige fordelinger av sannhetsverdier, og vi kan på fritt grunnlag tildele en av to sannhetsverdier til hver av dem. Til sammen blir det 2^{2^n} mulighteter.

To sammensatte utsagn som gir oss den samme sannhetsverdifunksjonen er ekvivalente, og er derfor ikke essensielt forskjellige. Siden $2^{2^3} = 256$ betyr det at vi bare kan beskrive 256 forskjellige sannhetsverdifunksjoner ved hjelp av tre utsagnsvariable.

2.3.1

Vi antar at R_i står for "du kan få 'rett nr i i opplistingen' ". Da kan vi uttrykke menyen ved

$$(R_1 \vee R_2) \wedge \neg(R_1 \wedge R_2) \wedge (R_3 \vee R_4 \vee R_5) \wedge \neg(R_3 \wedge R_3) \wedge \neg(R_3 \wedge R_5) \wedge \neg(R_4 \wedge R_5) \wedge (R_6 \vee R_7 \vee R_8) \wedge \neg(R_6 \wedge R_7) \wedge \neg(R_6 \wedge R_8) \wedge \neg(R_7 \wedge R_8) \wedge (R_9 \vee R10) \wedge \neg(R_9 \wedge R_{10})$$

2.3.2

1. A : Det er midt på dagen
   B : Det er overskyet
   C : Jeg kan ikke lese boka mi utendørs
   Svar $A \wedge B \rightarrow C$

2. A : Maleren kommer i morgen
   B : Du har ryddet rommet ditt
   C : Du mister lommepengene dine for en måned
   Svar $A \wedge (\neg B \rightarrow C)$

3. A : Jeg får med nok penger
   B : Jeg skal kjøpe meg en jakke
   C : Jeg skal kjøpe meg et par sko
   Svar $A \rightarrow (B \vee C)$

   Her kan 'eller' tolkes på to måter. Vi har valgt den inklusive tolkningen i løsningen, men den andre kan være vel så sannsynlig.

4. Denne oppgaven er litt søkt, for det er veldig lite av meningsinnholdet i setningen vi kan få med oss når vi tolker den rent utsagnslogisk:

   A : Tro kan flytte fjell

   B : TNT egner seg bedre enn tro til å flytte fjell

   Svar $A \wedge B$

5. A : Du er snill og hjelper meg

   B : Du får penger til kino

   C : Du skal gå til sengs uten mat

   Svar $(A \to B) \wedge (\neg A \to C)$

6. A : Solen står opp i vest

   B : Jeg havnet på flyet til Brisbane

   C : Jeg havnet på flyet til Vancouver

   Svar $A \to B \wedge \neg C$

2.3.4

1. OK.
2. $(\neg A \vee B) \wedge C \leftrightarrow \neg A \vee \neg B$ eller $\neg A \vee (B \wedge C) \leftrightarrow \neg A \vee \neg B$.
3. OK.
4. $(A \to \neg B \vee C) \to A$ eller $A \to (\neg B \vee C \to A)$.
5. OK.

2.3.5

Begge utsagnene får verdien $\top$ nøyaktig når én av variablene eller alle sammen har verdien $\top$.
Siden $A \leftrightarrow B$ og $B \leftrightarrow A$ også uttrykker det samme betyr det at vi kan bytte rekkefølge og flytte på parenteser så mye vi vil uten å endre betydningen av et utsagn med $\leftrightarrow$ som eneste bindeord. Vi kan komme fra et hvilket som helst slikt uttrykk til et hvilket som helst annet med de samme utsagnsvariablene i de samme antall ved å bytte rekkefølge og skyve på parenteser. Derfor betyr to slike uttrykk alltid det samme.

2.4.1

a) svarer til tautologien

$$(A \wedge B \to C) \to (A \to C) \vee (B \to C).$$

b) svarer til tautologien

$$(A \to C) \wedge (B \to D) \to (A \to D) \vee (B \to C).$$

c) svarer til tautologien

$$(A \leftrightarrow B) \vee (A \leftrightarrow \neg B).$$

Kapittel 3

3.1.1

a) $(x \in A \vee (x \in B \wedge \neg(x \in A)) \wedge (\neg(x \in A) \wedge x \in B)$ Obs! Parentesfeil i oppgaven.

b) $\neg(\neg(x \in A \vee \neg(x \in B)) \wedge \neg(x \in B \vee x \in C) \wedge x \in A$

c) $\neg(\neg(x \in C) \wedge \neg(x \in A)) \vee (x \in B \wedge x \in C)$

Kapittel 4

4.1.4
$(1, 1)$ og $(1^2, (-1)^2)$ er like.
$(4, 1)$ og $(2^2, 1)$ er like.
De andre er forskjellige fra disse og fra hverandre.

4.1.5
Første og tredje sekvens er like.
Andre og fjerde sekvens er like.
Femte og sjette sekvens er like.

4.1.6
Ved å trekke andre likning fra den første får vi at

$$b - c = d - f$$

Ved å legge denne likningen sammen med den siste, får vi at $b = d$, og derfra er det bare å nøste opp.

4.2.1
$A \times B$ har 8 elementer, og en mengde med 8 elementer har 256 delmengder.

4.2.2
A er argumentområdet til R, og B er verdiområdet.
B er argumentområdet til S og C er verdiområdet.
$S^{-1} = \{(6, 3), (6, 4), (5, 3)\}$ og $R^{-1} = \{(3, 1), (4, 2)\}$.
$S \circ R = \{(1, 5), (1, 6), (2, 6)\}$
Verdiområdet til R^{-1} er A og argumentområdet til S^{-1} er C. Siden disse ikke er like, kan vi ikke definere sammensetningen.
Likheten til slutt er ingen tilfeldighet, dette holder for alle relasjoner hvor sammensetningen er definert.

4.2.3
$(n, k) \in (<)^2$ hvis det finnes en m slik at $(n, m) \in <$ og $(m, k) \in <$, det vil si at hvis det finnes en m slik at $n < m$ og $m < k$. wi har eksempelvis at $2 < 3$ men ikke at $2(<)^2 3$ fordi det ikke finnes noe tall imellom.
$\leq^2 = \leq$.

4.2.4

a): Vi har $nRmn$ og $mnR^{-1}m$ for alle tall n og m, så $nR^{-1} \circ Rm$ for alle tall n og m.
b): Ved å gå veien om 1 ser vi at $R \circ R^{-1}$ også består av alle tallpar.

4.3.1
a): f er injektiv men ikke surjektiv.
b): f er både injektiv og surjektiv med $g(x) = \ln(x)$ som den inverse.
c): f er både injektiv og surjektiv med seg selv som invers.
d): f er injektiv men ikke surjektiv.
e): f er surjektiv, men ikke injektiv.

4.3.6
Bruk vanlig algebra til å vise at

$$P(n-1, m+1) = P(n, m) + 1$$

og at

$$P(n+1, 0) = P(0, n) + 1.$$

Dette viser at P 'teller opp' mengden av tallpar på en systematisk måte.

4.4.1
Sørg for at antall fyrstikker som er igjen etter ditt trekk er delelig med fire. Har du oppnådd dette på et tidspunkt i spillet, kan du fortsette med at det er slik, og når det er $0 \cdot 4$ antall fyrstikker igjen har du vunnet.

4.5.1
Hvis aRb og bRa vil $a = b$ siden R er antisymmetrisk. Men det er i konflikt med at R også er irrefleksiv.

4.5.3
a): Hvis $a - b = 17n$ og $c - d = 17m$, vil $(a + c) - (b + d) = 17(n + m)$.
b): Punkt a) sier at det ikke spiller noen rolle hvilke punkter vi plukker ut av to ekvivalensklasser. Legger vi dem sammen får vi ekvivalente punkter, så vi havner i den samme ekvivalensklassen.
c): Vi har også at om aRb og cRd så vil $acRbd$.
d): $[4] \cdot [11] = [44] = [10]$ så her er det feiltrykk i oppgaveteksten for de som leser utgaven av 10. februar.

Kapittel 5

5.1.1
Bevis for at $A \subseteq B$ og $B \subseteq A$ medfører at $A = B$.

Bevis
To mengder er like hvis de har de samme elmentene. Hvis alle elementer i A også er elementer i B og omvendt, må mengdene ha de samme elementene.

Bevis for at $A \subseteq B$ og $B \subseteq C$ medfører at $A \subseteq C$:

Bevis
La $x \in A$. Siden $A \subseteq B$ vil vi også ha at $x \in B$. Siden $B \subseteq C$ har vi videre at

$x \in C$. Siden $x \in A$ var vilkårlig ser vi at $A \subseteq C$.

Når det gjelder så enkle observasjone som i disse tilfellene her, kan det ofte være vanskelig å finne et passende detaljerthetsnivå for et bevis. Disse løsningsforslagene er bare forslag, det ene hovedsaklig i dagligtale og det andre basert på å uttrykke definisjonen av inklusjon via symboler.

5.1.2
Bruk at 4 og 6 er partall og at summen av partall alltid er partall. Det er ikke nødvendig å dele beviset opp i tilfeller her.

5.1.4
Feilen ligger i beviset for at R^T er total. Det at aSb for en total utvidelse S av R, betyr ikke at dette holder for alle totale utvidelser av R, da kan a og b være ordnet den andre veien.

5.2.3
Feilen ligger i påstanden om at $x = \sqrt{x^2}$, som bare holder for $x \geq 0$.

5.2.4
Beviset er i det alt vesentlige riktig. Hvis $m = p_1 \cdot \ldots \cdot p_k + 1$ får vi alltid rest 1 om vi deler m på p_i. Siden listen er antatt å inneholde alle primtallene, betyr det at m ikke har noen primtallsfaktorer. Siden alle tall > 1 kan faktoriseres i primtall, gir dette en motsigelse.

Kapittel 6

6.1.2
$C(0) = 1$, $C(1) = 2$ og $C(n) = C(n-1) + 4(n-1)$ når $n > 1$. $C(7) = 94$ mens $2^7 = 128$, så der bryter det sammen.

6.1.2
$y^0 = 1$ og $y^{S(x)} = y \cdot y^x$.

6.1.3
a): Hvis det ikke er noen fyrstikker igjen, kan spillet spilles på en måte, nemlig ved ikke å gjøre noe, og hvis det er én fyrstikk igjen kan spillet spilles på én måte, nemlig ved å ta bort fyrstikken.
Hvis det er mer enn én fyrstikk igjen, kan første spiller velge mellom å ta en eller to fyrstikker. I der første tilfellet er det $F(n-1)$ måter å fortsette spillet på, og i det andre tilfellet er det $F(n-2$ måter å fortsette spillet på.
Vi har gjenfunnet Fibonacci-tallene i dette enkle spillet.
b): $G(0) = 1$, $G(1) = 1$ og $G(2) = 2$. Hvis $n > 2$ har vi at $G(n) = G(n-1) + G(n-2) + G(n-3)$.

Kapittel 7

7.1.1
Induksjonstart: Hvis $n = 1$ så er a sin egen tverrsum, så vi får samme rest om vi deler a eller tverrsummen på tre.

Induksjonskritt: Anta at påstanden holder for tall med k siffre, og anta at a skrives med $k+1$ siffre, med første siffer c La b være tallet som skrives med resten av siffrene i a. Da er $a = c \cdot 10^k + b$. Siden $10^k - 1$ kan deles med 3, får vi samme rest om vi deler a med 3 som når vi deler $c + b$, med 3, og i følge induksjonsantagelsen, samme rest som om vi deler c + tverrsummen(b) på 3. Men da er det tverrsummen til a vi deler på 3.
Dette viser induksjonskrittet.

7.1.3
$1 + 2^0 = 2$ så formelen holder for $n = 0$.
$1 + 2^1 = 3$ så formelen holder for $n = 1$.
La $n \geq 2$ og anta at $H(n-2) = 1 + 2^{n-2}$ og at $H(n-1) = 2^{n-1}$.
Da er

$$H(n) = 3H(n-1) - 2H(n-2) = 3(1 + 2^{n-1}) - 2(1 + 2^{n-2})$$

$$= 1 + 6 \cdot 2^{n-2} - 2 \cdot n^{n-2} = 1 + 2^n.$$

7.1.6
Primtallsummen til 1 er 0 og primtallsummen til 2 er 2.
La n være gitt og anta at egenskapen holder for alle tall $m < n$.
Hvis n er et primtall, er n sin egen primtallsum.
Ellers må $n \geq 4$. La p være den minste primtallsfaktoren i n. La $m = \frac{n}{p}$.
Da er $p \leq m \leq \frac{n}{2}$, og primtallsummen til n er summen av p og primtallsummen til m, som samlet er mindre eller lik n.
dette viser induksjonskrittet og påstanden er vist.

7.2.2
Her vil vi bare gi noen hint til hvordan oppgaven kan løses.

a) Husk at $\mathbb{Q}$ betegner de rasjonale tallene.
$0 \in \mathbb{Q}$ Dette gir induksjonstarten.
Hvis $K \in \mathbb{Q}$ og $n \in \mathbb{N}$, vil $\frac{1}{n+K} \in \mathbb{Q}$.
Dette gir induksjonskrittet.

b) $0 \leq 0 \leq 1$.
$0 \leq K \leq 1 \wedge n \in \mathbb{N} \Rightarrow 0 \leq \frac{1}{1+K} \leq 1$.

c) Hvis $q = 0$ eller $q = 1$ kan q skrives som en kjedebrøk, så anta at $0 < q < 1$ og la $q = \frac{n}{m}$ hvor $n < m$. La $m = an + b$ hvor $b < n$.
Da kan vi skrive q som en brudden brøk

$$q = \frac{n}{m} = \frac{1}{\frac{m}{n}} = \frac{1}{a + \frac{b}{n}},$$

og vi kan bruke en passende induksjonsantagelse på $\frac{b}{n}$.

Resten er en fortsatt utfordring.

Register