MAT1030 – Forelesning 20

Kombinatorikk

Dag Normann - 7. april 2010

(Sist oppdatert: 2010-04-07 12:43)

Fortsettelse fra 06.04.2010

Rekursjon og induksjon

- I går så vi på formelle språk generelt, og på de formelle språkene av utsagnslogiske formler og av korrekte parentesuttrykk spesielt.
- Eksemplene vi har sett på bruk av rekursjon på oppbyggingen av ord eller formler er så sentrale at de kan bli brukt som grunnlag for oppgaver gitt til eksamen eller til Oblig. 2.
- Det neste eksemplet har ikke den samme betydningen i informatikk, og kan derfor best sees på som et eksempel for innøvelse av forståelse og ferdigheter.
- I eksemplet beskriver vi "den universelle datatype", en universell mengde hvor vi kan finne alle *virkelige* datatyper som delmengder.

Eksempel.

- Vi definerer de hereditært endelige mengdene HF som den minste mengden slik at
  - $\emptyset \in \mathsf{HF}$
  - Hvis $X = \{a_1, \ldots, a_n\}$ er en endelig delmengde av HF, så er $X \in \mathsf{HF}$.
- Vi definerer $f : \mathsf{HF} \to \mathbb{N}_0$ ved rekursjon over HF ved
  - $f(\emptyset) = 0$
  - $f(\{a_1, \ldots, a_n\}) = 2^{f(a_1)} + \cdots + 2^{f(a_n)}$.

  Her må vi anta at mengden er beskrevet uten gjentagelse, det vil si at alle a_i'ene er forskjellige.

Eksempel (Fortsatt).

- Ved induksjon over HF ser vi at $f(X) \in \mathbb{N}_0$ for alle $X \in \mathsf{HF}$.

  Hvis $X = \emptyset$ ser vi det direkte, og hvis $X = \{a_1, \ldots, a_n\}$ kan vi bruke induksjonsantagelsen som sier at $f(a_i) \in \mathbb{N}_0$ for alle i.
- Hvis $k \in \mathbb{N}_0$ vil det fins en og bare en $X \in \mathsf{HF}$ slik at $f(X) = k$.
- Her bruker vi induksjon på k.

  For $n = 0$ ser vi at det bare er $X = \emptyset$ som kan gi $f(X) = 0$ siden $2^{f(a_i)} > 0$ for alle mulige $a_i \in X$.

For $k > 0$, fins det en og bare en måte å skrive k som en sum av forskjellige 2'erpotenser på, og hver eksponent k_i vil komme fra en og bare en mengde a_i.

Det betyr at vi kan finne X slik at $f(X) = k$ fra k.

Oppgave.

La HF og f være definert som i eksemplet over, og la $S : \mathrm{HF} \to \mathrm{HF}$ være definert ved

$$S(X) = X \cup \{X\}.$$

a) Vis at $X \in \mathrm{HF} \Rightarrow S(X) \in \mathrm{HF}$.

b) La $X, Y \in \mathrm{HF}$.
Vis at $X \in Y \Rightarrow f(X) < f(Y)$.

c) Forklar hvorfor det ikke fins noen $X \in \mathrm{HF}$ slik at $X \in X$

d) La $N \subseteq \mathrm{HF}$ være den minste mengden slik at $\emptyset \in N$ og slik at hvis $X \in N$ så vil $S(X) \in N$.
Drøft sammenhengen mellom N og $\mathbb{N}_0$.

- Hvis vår induktive definisjon av mengden av utsagnslogiske formler skulle vært gitt i en lærebok i informatikk på et avansert nivå, eller i en forskningsartikkel, ville den sett ut som noe slikt:

  Utsagnsvariabel v

  $v ::= p \mid q \mid r$

  Formel A

  $A ::= v \mid \neg A \mid (A \wedge A) \mid (A \vee A)$

- De to første linjene skal leses som følger:

  I denne definisjonen lar vi v betegne en vilkårlig utsagnsvariabel.

  v kan stå for p, q eller r.

- De to neste linjene skal leses som følger:

  I denne definisjonen lar vi A betegne en vilkårlig formel.

  En formel kan enten være en utsagnsvariabel, fremkommet fra en annen formel ved å skrive $\neg$ foran, eller fremkommet fra to andre formler ved å skrive $\wedge$ eller $\vee$ mellom, og parenteser rundt.

- Vi kan bruke den samme effektive notasjonen for å definere mengden av korrekte parentesuttrykk:

  Parentesuttrykk P

  $P ::= e \mid (P) \mid PP$

  som uttrykker at vi får parentesuttrykkene ved å starte med det tomme ordet e og deretter enten sette parenteser rundt et korrekt uttrykk eller sette to korrekte uttrykk sammmen.

- Som tidligere nevnt, liker informatikere og logikere å la de naturlige tallene starte med 0, altså å arbeide med $\mathbb{N}_0$ i stedet for $\mathbb{N}$.
- Følgende definisjon er påtruffet i informatikklitteratur:

  $n : \mathsf{NAT}$

  $n ::= 0 \mid S(n)$.
- Dette skal leses som at vi definerer en datatype NAT ved å la n stå for et vilkårlig objekt, og vi finner objektene i NAT, enten som *symbolet* 0 eller som en *symbolsekvens* $S(w)$ hvor w er en symbolsekvens vi allerede vet er av type NAT.

Oppsummering

Kapittel 1–3

- algoritmer
- pseudokoder
- kontrollstrukturer
- representasjon av tall (hele og reelle tall)
- tallsystemer

Kapittel 4

- logisk holdbare argumenter
- utsagnslogikk og predikatlogikk
- utsagnslogiske bindeord og kvantorer
- parenteser
- sannhetsverditabeller, tautologier, kontradiksjoner
- logisk konsekvens, logiske lover
- bevisteknikker

Kapittel 5

- mengdelære
- notasjon
- mengdealgebra (union, snitt, komplement, mengdedifferens)
- Venndiagrammer
- kardinaltall, potensmengder, ordnede par
- binære relasjoner og egenskaper ved disse
- ekvivalensrelasjoner og partielle ordninger

Kapittel 6

- funksjoner
- terminologi, verdiområde, definisjonsområde
- surjektive og injektive funksjoner
- sammensetning og invers av funksjoner

Kapittel 7

- rekursjon og induksjon
- rekursive funksjoner
- induksjonsbevis
- rekurrenslikninger
- induktivt definerte mengder, f.eks.
  - mengden av ord over et alfabet
  - mengden av utsagnslogiske formler
  - mengden av parentesuttrykk
- generell rekursjon og induksjon

Kapittel 9: Kombinatorikk

Kombinatorikk

- Kombinatorikk er studiet av *opptellinger*, *kombinasjoner* og *permutasjoner*.
- Vi finner svar på spørsmål "Hvor mange måter ...?" uten å telle.
- Viktig del av f.eks. kompleksitetsanalyse av algoritmer.
  - Hvor mye *tid* bruker en algoritme?
  - Hvor mye *plass* bruker en algoritme?
- Grunnleggende, nyttig og fascinerende matematikk som dere må beherske.
- Vi bruker cirka to dobbeltimer på kapittel 9.

Eksempel.

- Spørsmål:

  Når de syv rette lottotallene på en kupong med 34 tall er trukket ut, hvor mange forskjellige rekker har seks rette?
- Svar:

  Det er syv forskjellige måter å plukke ut seks rette fra de syv rette.

  Det er 27 måter å velge ut det ene tallet som er feil.

  Det gir $7 \cdot 27 = 189$ forskjellige rekker med seks rette.

- Kombinatorikk inngår som et vesentlig element i sannsynlighetsteori.
- Kombinatorikk inngår også når man skal vurdere hvor lang tid et program trenger for å nå i mål og når man skal vurdere hvor stor lagringsplass man må sette av for at et program eller en programpakke skal få det nødvendige arbeidsrommet.
- Lagring av data i forskjellige registere kan illustreres ved lagring av kuler i bokser.
- Et naturlig spørsmål vil da være hvor mange forskjellige måter dette kan gjøres på.
- I dette eksemplet skal vi anta at vi har tretten kuler og fire bokser.

- Hvis vi spør om på hvor mange måter vi kan fordele 13 kuler på fire forskjellige bokser, er det to mulige presiseringer.

Eksempel (Tilfelle 1).

- Alle kulene er forskjellige.
- Da har vi 13 kuler, og vi har fire muligheter for plassering av hver kule.
- Det gir
$$4^{13}$$
mulige fordelinger.

Eksempel (Tilfelle 2).

- Kulene er like, mens boksene fortsatt er forskjellige, kall dem A, B, C og D.
- La
$$xxxxxxxxxxxxxxxx$$
være en mengde med 16 elementer (en mindre enn antall kuler pluss antall bokser).
- Det fins
$$\binom{16}{3} = \frac{16!}{3! \cdot 13!} = 960$$
måter å omgjøre tre x til X på.

Eksempel (Tilfelle 2, fortsatt).

- Eksempel
$$xxxXxxxxXxxXxxxx$$
- I dette tilfellet plasserer vi tre kuler i A (foran første X), fire kuler i B (mellom første og andre X), to kuler i C (mellom andre og tredje X) og fire kuler i D (bak siste X).
- Alle plasseringer av de tre X'ene gir oss en fordeling av kulene på de fire boksene.

Eksempel (Tilfelle 2, fortsatt).

- Omvendt vil en fordeling av 13 kuler på boksene A, B, C og D gi oss en plassering av X'ene.
- Har vi to kuler i A, to i B, fem i C og fire i D svarer det til
$$xxXxxXxxxxxXxxxx.$$

- Dette eksemplet er hva vi kaller generisk.
- Det betyr at vi kan lese en generell setning ut av eksemplet:

Teorem.
Det finnes

$$\binom{n+m-1}{m-1} = \frac{(n+m-1)!}{(m-1)! \cdot n!}$$

måter å fordele n like objekter på m ulike beholdere på.

Merk.
- Vi kunne ha formulert problemet om antall fordelinger også i det tilfellet hvor det ikke er forskjell på boksene.
- Det krever imidlertid at vi går ut over læreboka i en retning som ikke er prioritert, så det skal vi la ligge.
- Vi kunne også ha sett på tilfellet der vi har seks hvite og syv røde kuler.
- Dette er en utfordring dere bør kunne håndtere selv når vi er ferdige med kapittel 9.

- Vi skal komme tilbake til opptelling av mulige fordelinger i forskjellige situasjoner.
- Først skal vi imidlertid se på en sammenheng mange kjenner fra sannsynlighetsteorien.
- I læreboka går den under betegnelsen Inklusjons- og eksklusjonsprinsippet.

Inklusjons- og eksklusjonsprinsippet

Eksempel.
- På en skole er det 177 elever som driver aktiv idrett.

  103 av elevene er aktive om sommeren og 85 av elevene er aktive om vinteren.

  Det betyr at det er 188 aktiviteter fordelt på 177 elever.

  For at dette skal stemme, må 11 av elevene drive både sommer- og vinteridrett, siden det er 11 flere aktiviteter enn det er elever.
- Hvis vi lar S være mengden av elever som driver sommeridrett og V være mengden av elever som driver vinteridrett, ser vi at
  - $|S| = 103$
  - $|V| = 85$
  - $|S \cup V| = 177$
  - $|S \cap V| = 11$

Eksempel (Fortsatt).

- Det siste tallet regnet vi ut på grunnlag av de tre første.
- Vi ser at $|S| + |V| = |S \cup V| + |S \cap V|$ fordi det at det er flere aktiviteter enn elever skyldes at noen driver to aktiviteter, og differensen mellom antall aktiviteter og antall elever må være nøyaktig antallet på de som driver både sommer- og vinteridrett.

Teorem (Inklusjons- og eksklusjonsprinsippet).

La A og B være to endelige mengder.

Da er $|A \cup B| = |A| + |B| - |A \cap B|$.

Bevis.

Hvis vi først teller opp elementene i A og deretter elementene i B, har vi talt elementene i $A \cap B$ to ganger.

For å få antall elementer i $A \cup B$ må vi derfor trekke fra det vi har talt for mye, nemlig antallet i $A \cap B$.

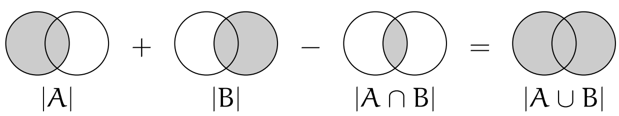
Merk.

- Det er en nær sammenheng mellom *inklusjons- og eksklusjonsprinsippet* og en tilsvarende lov om sannsynlighet:

$$P(A) + P(B) = P(A \cup B) + P(A \cap B)$$

  når A og B er uavhengige hendelser og P måler en sannsynlighet.
- Begge lovene illustreres greit med et Venn-diagram, hvor man ser at hvis vi skraverer sirkelskiven som markerer A i en retning og sirkelskiven som markerer B i en annen retning, er det akkurat feltet som markerer $A \cap B$ vi skraverer to ganger.

- Inklusjons- og eksklusjonsprinsippet:



Eksempel.

- Anta at vi har fått følgende oppgave:

  Av 231 studenter var det 174 som greide oppgave 1 og 175 som greide oppgave 2.

  Alle studentene greide minst en oppgave.

  Hvor mange studenter greide begge oppgavene?

Eksempel (Fortsatt).

- Løsning:

  La A være mengden av studenter som greide oppgave 1 og B mengden av studenter som greide oppgave 2.

  Da er $|A \cup B| = 231$, $|A| = 174$ og $|B| = 175$.

  Inklusjons- og eksklusjonsprinsippet sier oss at

  $$231 = 174 + 175 - |A \cap B|.$$

  Det gir oss at

  $$|A \cap B| = 174 + 175 - 231 = 118.$$

  Det var 118 studenter som greide begge oppgavene.

Eksempel.

La oss se på følgende oppgave:

- Medlemmene i et idrettslag blir bedt om å registrere seg på nettet med navn, adresse og enten e-postadresse eller mobilnummer.

  Ledelsen ønsker å automatisere utsendelsen av informasjon, uten å bruke to informasjonskanaler til samme medlem, men av de 728 medlemmene er det 94 som har oppgitt både e-postadresse og mobilnummer.

  Når vi vet at 562 medlemmer oppga e-postadresse, hvor mange oppga da mobilnummer?

Eksempel (Fortsatt).

- Løsning:

  La A være mengden av medlemmer som oppga e-postadresse og B være mengden av de som oppga mobilnummer.

  Da sier inklusjons- og eksklusjonsprinsippet at

  $$728 = 562 + |B| - 94$$

så

$$|B| = 728 + 94 - 562 = 260.$$

Det var 260 medlemmer som oppga mobilnummer.

Multiplikasjonsprinsippet

- Det neste prinsippet for beregning av antall muligheter vi skal se på er multiplikasjonsprinsippet.
- Multiplikasjonsprinsippet sier at hvis vi skal treffe en serie uavhengige valg, vil det totale antall muligheter være produktet av antall muligheter ved hvert valg.
- Igjen fins det en klar parallell i sannsynlighetsteori, hvor vi fins sannsynligheten for at en serie uavhengige hendelser finner sted ved å ta produktet av sannsynlighetene for enkelthendelsene.
- Vi skal illustrere dette prinsippet ved et par eksempler.

Eksempel.
- Et norsk registreringsnummer for bil består av to store bokstaver og fem siffre.
- Vi bruker ikke bokstavene G, I, O, Q, Æ, Ø eller Å, fordi de enten ikke forekommer utenlands, eller fordi de kan forveksles med tall eller andre bokstaver.
- Da står vi igjen med $22 \cdot 22 = 484$ bokstavkombinasjoner.

Eksempel (Fortsatt).
- Første siffer i nummeret må være et av de ni tallene 1, 2, 3, 4, 5, 6, 7, 8, 9, mens de fire andre sifrene kan hentes fra alle de ti tallsymbolene.
- Det gir tilsammen

$$22 \cdot 22 \cdot 9 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 43.760.000$$

mulige registreringsnummere på norske biler.

Eksempel (Fortsatt).
- Svenskene bruker tre bokstaver og tre tall, og hadde de begrenset seg til de bokstavene vi bruker i Norge, ville de kunne registrert færre biler.

Oppgave.

Hvor mange bokstaver må svenskene tillate for å kunne registrere like mange biler (eller fler) enn det nordmennene kan?

Eksempel.

- Amerikanske lærebokforfattere lever i den tro at amerikanske collegestudenter lever en ikke ubetydelig del av livet sitt med å spise sammen.
- Derfor er følgende oppgave typisk for amerikanske lærebøker i diskret matematikk.
- En sandwich-bar tilbyr:
  1. Fire typer brød: Fint, mellomgrovt, grovt og glutenfritt.
  2. Tre typer smøring: Smør, majones og sennep.
  3. Seks typer hovedpålegg: Kalkun, roastbeef, skinke, tunfisk, skalldyr og soyaprotein.
  4. Fire typer tilbehør: Stekt bacon, salat, agurk og tomat.
  5. Tre valg på dressing, Thousen Islands, tomatdressing og hvitløksdressing.
- Hvor mange forskjellige sandwicher er det mulig å komponere?

Eksempel (Fortsatt).

Selv om ikke alle sammensetningene vil være like vellykkede rent smaksmessig, er det ingen føringer på hvilke valg som kan kombineres.

Da finner vi det totale antall muligheter ved å bruke multiplikasjonsprinsippet.

- Vi har da

$$4 \cdot 3 \cdot 6 \cdot 4 \cdot 3 = 864$$

forskjellige sammensetninger.