

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i INF3100 — Databasesystemer

Eksamensdag: 8. juni 2011

Tid for eksamen: 9.00 – 13.00

Oppgavesettet er på 4 sider.

Vedlegg: Ingen

Tillatte hjelpemidler: Kalkulator og ordbok

Kontroller at oppgavesettet er komplett før
du begynner å besvare spørsmålene.

Les oppgavene nøye, og lykke til!

Folkeregisteret inneholder informasjon om alle innbyggere i Norge og hvor de bor. Vi skal se på et lite (og ikke nødvendigvis optimalt) utsnitt av disse opplysningene.

Alle personer har et entydig 11-sifret fødselsnummer **fnr**, f.eks. 14105833355. En persons navn består av **fornavn**, **etternavn** og eventuelt **mellomnavn**. For eksempel har Johan Herman Lie Vogt fornavnet 'Johan Herman', mellomnavnet 'Lie' og etternavnet 'Vogt', mens Marius Sophus Lie har fornavn 'Marius Sophus' og etternavn 'Lie' (men intet mellomnavn). Når noen flytter, må de sende inn en flyttemelding til folkeregisteret. Flyttemeldingen inneholder **flyttedato** og gammel og ny adresse (henholdsvis **fraadr** og **tiladr**) samt hvilke personer flyttemeldingen omfatter. Når en flyttemelding registreres i folkeregisterets database, tilordnes den en identifikator **mid**.

Denne informasjonen er samlet i følgende to relasjoner der primærnøkklene er understreket:

Person(fnr, fornavn, mellomnavn, etternavn)

Flyttemelding(mid, fnr, flyttedato, fraadr, tiladr)

En person kan ikke flytte to ganger på samme dag. Hver person skal bare sende inn én flyttemelding pr. flytting. Det er imidlertid tilstrekkelig å sende inn en felles flyttemelding når det er flere som flytter samtidig fra en adresse til en annen.

(Fortsettes på side 2.)

For at databasen skal ha de rette egenskapene, må derfor *Flyttmelding* oppfylle de to funksjonelle avhengighetene (FDene)

$mid \rightarrow \text{flyttedato, fraadr, tiladr}$
 $\text{fnr, flyttedato} \rightarrow mid$

I tillegg skal *fraadr* og *tiladr* i en og samme flyttmelding alltid være forskjellige, men dette må ivaretas av en annen type integritetsregel.

Oppgave 1 FDer og normalformer (30%)

- (i) (5%) Finn kandidatnøkklene i *Flyttmelding*.
- (ii) (10%) Hvilken normalform er *Flyttmelding* på? Begrunn svaret.
- (iii) (5%) Dekomponer *Flyttmelding* tapsfritt til Boyce-Codd (BCNF). Avgjør om dekomposisjonen er FD-bevarende.

En alternativ dekomposisjon av *Flyttmelding* er

$G1(\text{mid, fraadr, tiladr})$
 $G2(\text{mid, fnr, flyttedato})$

- (iv) (5%) Vis at denne dekomposisjonen er tapsfri.
- (v) (5%) Hvilke fordeler og ulemper har ($G1$, $G2$) i forhold til den dekomposisjonen du kom frem til i oppgave 1 (iii)?

Oppgave 2 SQL (30%)

Ta utgangspunkt i de opprinnelige relasjonene *Person* og *Flyttmelding*. Besvar følgende spørsmål ved hjelp av SQL.

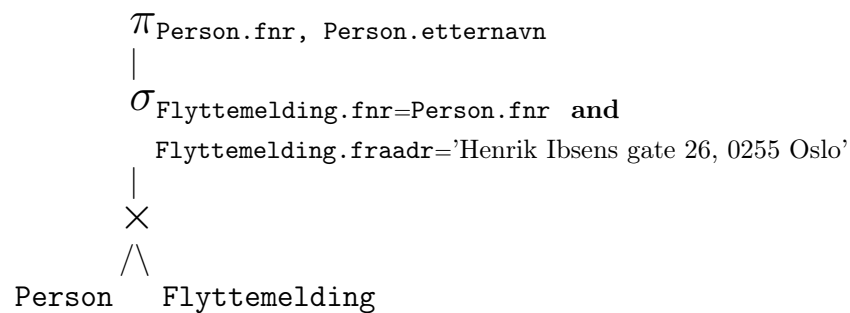
- (i) (5%) Definer relasjonen *Flyttmelding* med SQL. Husk primærnøkkel, eventuelle andre kandidatnøkler og fremmednøkler. Ta også eventuelt med de integritetsreglene som kan håndteres med CHECK. Du kan anta at *mid* og *fnr* er heltall, *flyttedato* har datatypen DATE og at *fraadr* og *tiladr* har datatypen VARCHAR(100).
- (ii) (10%) Sjekk om *Flyttmelding* oppfyller FDen $\text{mid} \rightarrow \text{flyttedato, fraadr, tiladr}$ ved å skrive et SQL-uttrykk som returnerer de *mid*-ene som eventuelt bryter FDen.

(Fortsettes på side 3.)

- (iii) (5%) Finn nåværende adresse til alle som heter Marius Sophus Lie (og som har flyttet minst én gang). Skriv ut fødselsnummer og adresse, sortert på fødselsnummer.
- (iv) (10%) Sjekk at alle fra- og til-adresser i *Flyttmelding* er konsistente, dvs. at hvis en person har flyttet mer enn én gang, så er *fraadr* lik innholdet i *tiladr* i personens forrige flyttmelding. For hver person hvor dette ikke stemmer, skriv ut personens *fnr* samt *mid* for de to flyttmeldingene det gjelder.

Oppgave 3 Relasjonsalgebra og spørreplaner (25%)

Betrakt følgende uttrykkstre i relasjonsalgebra:



- (i) (5%) Forklar med ord hva som er resultatmengden til uttrykket (dvs. hvilket spørsmål uttrykket finner svaret på).
- (ii) (10%) Uttrykkstreet utgjør en logisk spørreplan. Optimer den logiske spørreplanen (tegn det nye uttrykkstreet).
- (iii) (10%) Den optimerte logiske spørreplanen fra oppgave 3 (ii) er utgangspunkt for valg av fysisk spørreplan (konkrete algoritmer). Anta at de to relasjonene *Person* og *Flyttmelding* bare har indekser på primærnøkklene. Forklar om og i tilfelle hvordan den fysiske spørreplanen kan nyttiggjøre seg indeksene. (Hvis du ikke har løst oppgave 3 (ii), kan du besvare spørsmålet med hensyn på det opprinnelige uttrykkstreet.)

(Fortsettes på side 4.)

Oppgave 4 Transaksjonsprotokoller (15%)

- (i) (10%) Forklar kort prinsippet bak transaksjonsprotokollen Snapshot Isolation (SI).
- (ii) (5%) Hva vil det si at SI ikke garanterer serialiserbarhet? Hvilke konsekvenser kan dette ha for resultatet av en transaksjonsutførelse? Hvorfor er SI likevel så mye benyttet i kommersielle databasesystemer?

Ellen Munthe-Kaas