

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i : IN 212 — Databaseteori
Eksamensdag : Mandag 24. august 1998
Tid for eksamen : 09.00 - 15.00
Oppgavesettet er på : 4 sider
Vedlegg : Ingen
Tillatte hjelpemidler : Alle trykte og skrevne

Les hele oppgaven nøye og kontroller at oppgavesettet er komplett før du begynner å besvare det. Alle oppgavene bør besvares, så disponer tiden godt.

Oppgave 1 (35 %)

Nedenfor følger et begrepsmessige skjema for en relasjonsdatabase for et medlemsregister i en landsomfattende forening med flere lokallag.

Skjema 1:

MEDLEM (LokalLagKode, MedlemsNr, MedlemsKat, Navn, FødselsDato, Vei, PostNr, PostSted, Kontingent, Betalt)

(LokalLagKode, MedlemsNr) er primærnøkkel i MEDLEM

Kontingent har domene heltall. *Betalt* er boolsk (domene Boolean). *FødselsDato* har domene dato. Alle de andre attributtene har domene tekst.

Kontingenten er avhengig av hvilken kategori (*MedlemsKat*) medlemmet tilhører (hovedmedlem, husstandsmedlem, ungdomsmedlem, pensjonistmedlem etc.) . Hvert lokallag bestemmer selv hvilke medlemskategorier man vil operere med og kontingentsatsen for hver av dem. *Betalt* angir om kontingenten er betalt for inneværende år.

Oppgave 1 A Hva er den høyeste normalformen til relasjonen MEDLEM? Begrunn svaret.

Oppgave 1 B Dekomponer denne relasjonen over på et antall relasjoner som alle er på Boyce-Codd-normalform (BCNF)?

Oppgave 1 C MEDLEM er ikke på BCNF. Begrunn hvorfor det likevel kan hevdes at en database som omfatter bare denne ene relasjonen, er godt strukturert?

Oppgave 1 fortsetter på neste side

Et alternativt skjema for relasjonsdatabasen er skjema 2:

LOKALLAG (LokalLagKode, MedlemsKat, Kontingent)

MEDLEM (LokalLagKode, MedlemsNr, MedlemsKat, Navn,
FødselsDato, Vei, PostNr, PostSted, Betalt)

Oppgave 1 D Basert på opplysningene gitt tidligere i oppgaven, hvilke kandidat- og fremmednøkler er det rimelig å anta gjelder i skjema 2?

I oppgavene nedenfor skal du ta utgangspunkt i skjema 2.

Oppgave 1 E Formuler i både SQL, relasjonsalgebra og tuppelkalkyle et spørsmål som returnerer en liste med alle opplysninger (unntatt kontingentsatsen) om alle medlemmer.

Oppgave 1 F Formuler i både SQL, relasjonsalgebra og domenekalkyle et spørsmål som returnerer en liste over alle medlemmer og kontingentsatsen for hver enkelt av dem.

Oppgave 1 G Formuler i SQL en oppdatering som angir at et gitt medlem har betalt kontingenten.

Oppgave 1 H Bruk SQL til å lage en liste over alle lokallag, og hvor mange medlemmer som er registrert i hver enkelt av dem, samt det totale antall medlemmer i hele foreningen. Listen skal være sortert på antall medlemmer i avtagende rekkefølge. Lokallag uten medlemmer skal utelates.

Slutt på oppgave 1

Oppgavesettet fortsetter på neste side

Oppgave 2 (15 %)

Oppgave 2 A Beskriv på norsk hva følgende SQL-spørsmål vil returnere:

```
SELECT MedlemsNr, Navn, LokalLagKode  
FROM LOKALLAG L, MEDLEM M  
WHERE Betalt =TRUE AND  
L.LokalLagKode = M.LokalLagKode
```

Oppgave 2 B Tegn et kanonisk (initielt) spørretre for SQL-spørsmålet gitt i oppgave 2A.

Oppgave 2 C Gjennomfør en transformasjon av treet du tegnet under oppgave 2B, til et optimalisert spørretre. Redegjør også for de enkelte stegene i transformasjonen – enten i form av tegninger eller tekstlig beskrivelse.

Oppgave 2 D Beskriv kort og poengtert egenskaper ved de to spørre-trærne du tegnet under 2B og det optimaliserte treet i 2C.

Oppgave 3 (10 %)

Oppgave 3 A Lag et objekt-orientert alternativ til skjema 2 fra oppgave 1. Skriv skjemaet enten i syntaksen til O₂ eller ObjectStore. Du kan sløyfe metoder/signaturer.

Oppgave 3 B Beskriv fordeler og ulemper ved det objekt-orienterte alternativet du lagde under oppgave 3 A av skjema 2, i forhold til relasjons-versjonen.

Oppgave 4 (10 %)

Siden lokallagene er fordelt over hele landet, er det nærliggende å bruke en form for distribuert database der hver lokal database inneholder f.eks. lokallagets medlemmer.

Oppgave 4 Ta utgangspunkt i skjema 2 fra oppgave 1 og beskriv med ord (norsk) hvordan du mener de enkelte relasjoner i databasen mest hensiktsmessig bør fragmenteres, allokeres og evt. replikeres. Beskriv fragmenteringen ved hjelp av relasjonsalgebra. Hvis du mener det er nødvendig å gjøre forutsetninger om bruks- mønsteret, så gjør rede for disse forutsetningene.

Slutt på oppgave 4

Oppgavesettet fortsetter på neste side

Oppgave 5 (15 %)

Gitt relasjonen R med skjema $R(R) = \{A,B,C,D,E\}$.

La $\mathcal{F} = \{CD \emptyset B, AE \emptyset BE, D \emptyset E\}$ være mengden av funksjonelle avhengigheter over R.

Oppgave 5 A Finn en kandidatnøkkel for R.

Oppgave 5 B Beregn \mathcal{F}^+_E (mengden av de elementære FD-ene i \mathcal{F}^+).

Oppgave 5 C R er kun på 1NF. Bruk algoritmen i avsnitt 7.3.4 i kompendiet til å dekomponere R tapsfritt til en mengde relasjoner som alle er på BCNF.

Oppgave 5 D Vis at dekomposisjonen du kom fram til i oppgave 5C ikke er FD-bevarende med hensyn på \mathcal{F} .

Oppgave 5 E Finnes det noen tapsfri og FD-bevarende dekomposisjon av R der alle skjemaene i dekomposisjonen er på BCNF? Gi en overbevisende begrunnelse for svaret.

Oppgave 6 (15%)

La R være en relasjon med skjema $R(R) = \{A,B,C,D\}$.

Nedenfor følger 4 slutningsregler for FD/MVD-er. For hver slutningsregel skal du bevise eller motbevise at regelen holder på R. (En slutningsregel motbevises enklest ved å angi et moteksempel, dvs. en instans av R hvor antagelsene holder, men hvor konklusjonen ikke holder.)

Du skal forutsette at det ikke gjelder andre integritetsregler for R enn de som følger fra antagelsene i den slutningsregelen du skal bevise/motbevise.

- (i) Hvis $A \twoheadrightarrow B$ og $B \twoheadrightarrow C$, så har vi at $A \twoheadrightarrow C$
- (ii) Hvis $A \twoheadrightarrow B$ og $A \emptyset C$, så har vi at $A \twoheadrightarrow D$
- (iii) Hvis $A \twoheadrightarrow BC$ og $B \twoheadrightarrow CD$, så har vi at $A \twoheadrightarrow CD$
- (iv) Hvis $A \twoheadrightarrow BC$ og $A \twoheadrightarrow CD$, så har vi at $A \twoheadrightarrow BD$

*Slutt på oppgavesettet
Lykke til!*