

# UNIVERSITETET I OSLO

## Det matematisk-naturvitenskapelige fakultet

Eksamen i :                INF3100/INF4100 — Databasesystemer  
Eksamensdag :            Onsdag 8. juni 2005  
Tid for eksamen :        14.30 – 17.30  
Oppgavesettet er på : 5 sider (inkl. vedlegg)  
Vedlegg :                 side nr. 5  
Tillatte hjelpemidler : Kalkulator og ordbøker

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

**Det er fem oppgaver**

**Alle har samme vekt, så beregn ca 35 minutter på hver**

### **Oppgave 1 Relasjonsalgebra**

Med basis i databasestrukturen beskrevet i vedlegg 1 på side 5, skal du besvare følgende oppgaver med relasjonsalgebra:

- Oppgave 1 a**            En avdelings kunder er de kundene som har minst ett prosjekt hvis prosjektleder er ansatt i avdelingen.  
Finn navn og adresse på kundene til avdelingen ledet av Berit Blom.
- Oppgave 1 b**            Finn navn og adresse på de kundene som ikke har noen ufakturerte, godkjente timer.
- Oppgave 1 c**            Dersom relasjonene er implementert som bager (og ikke mengder), hvilke endringer må du da gjøre i din besvarelse av oppgave 1 a?

## Oppgave 2 SQL

Med basis i databasestrukturen beskrevet i vedlegg 1 på side 5 og som du allerede har brukt i oppgave 1, skal du besvare følgende oppgaver med SQL:

- Oppgave 2 a** Blant de prosjektledere som leder prosjekter for kunden "HelseSør" og som har stilling "konsulent", skal det lages en sortert liste av de av disse konsulentene (navn) med timepris hvor denne timeprisen er mer enn 1400 kr. per time. Listen skal være sortert stigende på konsulentenes timepris.
- Oppgave 2 b** En avdelings kunder er de kundene som har minst ett prosjekt som er ledet av en ansatt på avdelingen. Lag en liste av navn og adresse på kundene til avdelingen hvor "Geir Hansen" er ansatt.
- Oppgave 2 c** Lag en liste med som for en gitt kunde, år og uke gir fakturaopplysninger for alle uker det ennå ikke er sendt faktura for. Listen skal ha én linje per prosjekt, og hver linje skal inneholde p-r, p-navn, ans-nr, summen av timer og det tilsvarende beløpet.

## Oppgave 3 Transaksjoner

Anta vi har følgende eksekveringsplan S:

-----

$T_1$ :	$w_1(A)$	$r_1(B)$	$w_1(B)$
$T_2$ :	$r_2(A)$	$w_2(B)$	
$T_3$ :		$r_3(B)$	$w_3(C)$
$T_4$ :		$r_4(C)$	$w_4(B)$ $w_4(C)$

-----

- Oppgave 3 a** Gi, gjerne med dine egne ord, en kort definisjon av presedensgrafen til en eksekveringsplan; og lag presedensgrafen for S.
- Oppgave 3b** Er S konfliktserialiserbar? Begrunn svaret.
- Oppgave 3c** Gi tilsvarende en kortfattet definisjon av polygrafen til en eksekveringsplan; og lag polygrafen for S.
- Oppgave 3d** Er S view-serialiserbar? Begrunn svaret.
- Oppgave 3e** Forklar hva det betyr hvis du finner en sykel i den ene, men ikke i den andre grafen.

## Oppgave 4 Parsering, spørreplan og optimering

I denne oppgaven skal du vise din forståelse av hvordan en spørring parseres, en logisk spørreplan lages og hvordan spørreplanen kan optimeres. Vi skal bruke kun følgende spørring i denne oppgaven, som henter ut navn, etternavn og kunde ID av alle kvinnelige kunder som har fullkasko bilforsikring i 2005:

```
SELECT KUNDE.kundeID, KUNDE.navn, KUNDE.etternavn
FROM KUNDE, POLISE, FORSIKRING
WHERE KUNDE.kundeID = POLISE.kundeID AND
      POLISE.dato LIKE '2005%' AND
      POLISE.produktID = FORSIKRING.produktID AND
      FORSIKRING.forsikringstype = 'BIL' AND
      FORSIKRING.dekning = 'FULLKASKO' AND
      KUNDE.kjønn = 'K'
```

- Oppgave 4a** Bruk den enkle grammatikken i vedlegg 2 til å lage et parserings-tre for spørringen ovenfor.
- Oppgave 4b** Konverter parserings-treet i oppgave 4a ovenfor til en logisk spørreplan i relasjonsalgebra (dvs. tegn uttrykkstreet). NB! Denne oppgaven skal løses uten optimering. Optimering hører til neste oppgave!
- Oppgave 4c** Optimering:
- i) Hvilke regler benyttes ofte (gir oftest stor ytelsesgevinst) for optimering av logiske spørreplaner?
  - ii) Optimer den logiske spørreplanen i deloppgave 4b ovenfor (tegn det nye uttrykkstreet).

## Oppgave 5 Semi-strukturerte data

**Oppgave 5a** Forklar hva semi-strukturerte data er (dvs. hva det er som karakteriserer semi-strukturerte data).

**Oppgave 5b** Forklar hva den semi-strukturerte datamodellen kan brukes til og hvorfor.

----- o 0 o -----

Ragnar Normann, Naci Akkøk, Jon Grov, Gerhard Skagestein og Arne Maus

**Vedlegg 1 — Relasjonsdatabaseskjema for en prosjekt- og timelistedatabase**

Primærnøkler er markert med **fete** typer, kandidatnøkler er i *kursiv*

Fremmednøkler fremgår av attributtnavnene

Ansatt (**ans-nr**, navn, adresse, telefon, stilling, avdeling)  
 Stilling (**s-kode**, *tittel*, timepris)  
 Avdeling (**avd-nr**, *avdnavn*, avdleder)  
 Prosjekt (**p-nr**, p-navn, p-leder, faktura-år, faktura-uke, kunde)  
 Kunde (**k-nr**, *navn*, adresse, telefon)  
 Timeliste (**ans-nr**, **år**, **uke**, godkjent)  
 Timelistelinje (**ans-nr**, **år**, **uke**, **dag**, **p-nr**, timer)

Hver post i Timelistelinje angir hvor mange timer en ansatt har arbeidet på et prosjekt på en gitt dag.

Faktura-år og faktura-uke angir, for hvert prosjekt, den siste uken det er sendt faktura for.

**Vedlegg 2 — Grammatikk for parsing av spørsmål**

```

<query>      ::= <SFW>

<SFW>       ::= SELECT <selList>
                FROM <fromList>
                WHERE <condition>

<selList>   ::= <attribute>, <selList> |
                <attribute>

<fromList>  ::= <relation>, <fromList> |
                <relation>

<condition> ::= <condition> AND <condition> |
                <attribute> = <attribute> |
                <attribute> = <pattern> |
                <attribute> LIKE <pattern>
  
```

Elementære syntaktiske kategorier som <attribute>, <relation> og <pattern> har ingen regler, men oversettes henholdsvis med navnet på attributtet, navnet på relasjonen eller en streng i anførselstegn.