

**Norges teknisk-naturvitenskapelige universitet  
Institutt for datateknikk og informasjonsvitenskap**



**EKSAMENSOPPGAVE I FAG TDT4145 – DATAMODELLERING OG  
DATABASESYSTEMER**

**Faglig kontakt under eksamen: Svein Erik Bratsberg**

**Tlf.: 50382**

**Eksamensdato: 4. juni 2008**

**Eksamenstid: 09.00-13.00**

**Tillatte hjelpemiddel: D: Ingen trykte eller håndskrevne hjelpemiddel tillatt. Bestemt, enkel kalkulator tillatt.**

**Språkform: Bokmål**

**Sensurdato: 25. juni 2008**

## Oppgave 1 – Datamodellering – 20 %

GodJobb er et firma som tilbyr utleie av forskjellige typer personell til alle fagområder. Dette er både ufaglærte og faglærte medarbeidere, samt ledere og spesialister. GodJobb ønsker å designe en database for å lagre informasjon om personell, kunder og oppdrag. For hver person registrert i databasen skal vi vite navn, adresse, fødselsdato, kjønn og telefonnummer. Vi ønsker også å registrere historien av oppdrag og kompetanse for hver person. For hvert oppdrag ønsker vi å vite stilling, firmanavn, adresse, lønn og oppdragsperiode. Vi ønsker også å vite hvilke utdannelser og kurs personen har, dvs. navn på kurs/sertifisering, grad, ant. år utdannelsen har normert, utdanningsinstitusjon og utdannelsesår. Kompetansen til personell er også klassifisert i forskjellige områder, for eksempel bank og finans, IT og telekommunikasjon, og lignende. En person kan gjerne være kompetent på flere områder. Hver person har en personlig konsulent i GodJobb som veileder dem.

**Lag en ER -modell for en database som dekker kravene ovenfor.** Husk å angi alle kardinalitetsrestriksjoner og nøkler. Databasen skal blant annet kunne brukes for å finne følgende (du trenger ikke å skrive disse spørringene):

- Adressen til en person med et bestemt telefonnummer.
- Oversikt over alle oppdrag fra et firma.
- Oversikt over alle oppdrag utført av en person.
- Alle personer som har mer enn 4 års høyere utdanning.

Forklar kort eventuelle forutsetninger du finner det nødvendig å gjøre.

## Oppgave 2 – Relasjonsalgebra og SQL – 20 %

I denne oppgaven skal du ta utgangspunkt i en database som inneholder informasjon om resultat fra alle eksamener ved NTNU våren 2008. Anta følgende relasjonsskjema (primærnøkler er understreket, attributter med samme navn som en annen tabell sin primærnøkkel er fremmednøkler):

```

Student (studnr, studnavn, studadr)
Eksamen (studnr, emnekode, poengsum, karakter, sensorID)
Sensor (sensorID, sensorNavn, sensoradr)
Emne (emnekode, emnenavn, faglærerpersnr)
Ansvarlig (faglærerpersnr, emnekode)
Faglærer (faglærerpersnr, faglærernavn, email)

```

For de følgende spørringene skal du gi svaret i *relasjonsalgebra*. Er du usikker på hvordan symbolene for operatorene ser ut kan du skrive dem med ord, f.eks. "JOIN".

- a) Finn navn på sensorer for emnet 'Objektorientert systemutvikling'?

- b) Finn navnet til alle studenter som har tatt eksamen med faglærer 'Per Hansen' som ansvarlig?

For hver av spørringene nedenfor skal du gi svaret i *SQL*.

- c) Finn navn på sensorer for emnet 'Systemutvikling'?
- d) Lag en liste av studentnavn, karakter sortert etter karakteren de har fått i emnet TDT4145?
- e) Finn antall studenter som har avlagt eksamen i TDT4145 og fått karakteren A?
- f) Finn studentnummer til studenter som har høyest poengsum ved eksamen uansett fag? NB: Det kan være flere studenter.

### Oppgave 3 – Lagring og indekser – 20%

- a) Anta det er plass til tre poster i hver blokk i et B-tre, men at det er plass til 4 pekere i en indeksblokk. Vis hvordan B-treet ser ut etter at du har satt inn poster med følgende nøkler i den gitte rekkefølgen under:  
1, 9, 2, 10, 11, 6, 7, 15, 28, 33, 4, 8

- b) Anta B-treet slik du svarte på oppgave a) og at postene er lagret i løvnodene i B-treet. Hvis det tar 10 millisekunder å lese inn en blokk fra disk, hvor lang tid tar det følgende søket:

```
SELECT * FROM TABELL WHERE key=28;
```

- c) Hvor lang tid tar det følgende søket:

```
SELECT * FROM TABELL;
```

- d) Hvor lang tid tar det å utføre det følgende søket:

```
SELECT * FROM TABELL WHERE key > 11 AND key < 28;
```

- e) Anta at de samme postene er lagret i en heapfil med plass til tre poster per blokk. Det er ingen indeks. Anta den sammenlenkede implementasjonen av heapfil. Hvor lang tid tar det følgende søket:

```
SELECT * FROM TABELL WHERE key=28;
```

- f) Hvor lang tid tar det å utføre det følgende søket:

```
SELECT * FROM TABELL;
```

- g) Hvor lang tid tar det å utføre det følgende søket:

```
SELECT * FROM TABELL WHERE key > 11 AND key < 28;
```

## Oppgave 4 – Queryevaluering – 10%

Anta følgende tabell:

Reserves(sid,bid,day,rname)  
50 bytes poster, 100 poster per blokk, 1500 blokker i tabellen.

Anta følgende SQL-spørsmål:

```
SELECT *  
FROM Reserves  
WHERE rname < 'C%';
```

Anta 10% av postene tilfredstiller betingelsen.

Hvor mange I/O-operasjoner får vi ved utføring av spørsmålet når vi bruker:

- Tabellskan.
- Clustered B-tre-index på rname.
- Unclustered B-tre-index på rname.

## Oppgave 5 – Transaksjoner – 15%

Forklar de tre samtidighetsproblemene med flettet utføring av databaseoperasjoner:

- "Reading uncommitted data" (WR conflicts)
- "Unrepeatable reads" (RW conflicts)
- "Overwriting uncommitted data" (WW conflicts)

## Oppgave 6 – Normalisering – 15%

Gitt følgende tabell og funksjonelle avhengigheter:

R(A,B,C,D,E,F,G) A->DG, AB->C, B->E, CG->F

- Finn alle kandidatnøkler for R.
- Hva er den høyeste normalformen som oppfylles av tabellen? Svaret må begrunnes.
- Vis hvordan tabellen bør normaliseres slik at alle de resulterende tabellene oppfyller kravene til Boyce-Codd normalform (BCNF).