

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



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

Faglig kontakt under eksamen: Svein Erik Bratsberg

Tlf.: 99539963

Eksamensdato: 13. august 2008

Eksamenstid: 15.00-19.00

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

Språkform: Bokmål

Sensurdato: 3. september 2008

Oppgave 1 – Datamodellering – 20 %

Vi ønsker å designe en database for å kunne gjøre prissammenligning for flyreiser. Databaseen skal inneholde priser for flyreiser i 6 måneders tid framover. En flyreise går mellom to steder, men kan ha mellomlandinger. Databaseen skal også inneholde alle mellomlandinger. Flyreisen har et gitt flightnummer og starter på en gitt dato/tidspunkt og slutter på en gitt dato/tidspunkt. Det kan være flere flyplasser for et sted. Vi ønsker å vite hvilke flyplasser reisen gjelder, også for mellomlandinger, men prissammenligningen går mellom forskjellige flyreiser mellom de samme stedene. Hver flyreise er utført av et gitt flyselskap. Databaseen skal kunne oppdateres hver dag med nye priser. Hver flyplass har en unik flyplasskode. Det skal også registreres avstanden i kilometer mellom en flyplass og stedet flyplassen er registrert på. Det kan være flere priser for samme flyreise, avhengig av hvilken billetttype det gjelder, og om det er for voksne eller barn.

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

- Den billigste flyreisen mellom to steder.
- Oversikt over alle flyreiser mellom to steder på en gitt dato.
- Oversikt over alle flyreiser som et flyselskap har.

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, karakter, sensorID)
Sensor (sensorID, sensornavn, sensoradr)
Emne (emnekode, emnenavn, faglærerpersnr)
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".

- Finne adressen til studenter som har fått karakteren A i emnet 'Objektorientert systemutvikling'.
- Finne studentnummeret til alle studenter som har tatt eksamen i både emnet 'Systemutvikling' og emnet 'Objektorientert systemutvikling'.

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

- c) Finn navn på sensorer som har samme adresse som studenter de har vært sensor for.
- d) Lag en liste av karakter, antall studenter for eksamen i emnet med emnekode TDT4145?
- e) Finn hvilken karakter det er gitt flest av. NB: Det kan være flere karakterer som oppfyller kriteriet.
- f) Finn sensorID til sensorer som har gitt flest F (uavhengig av emner). NB: Det kan være flere sensorer som oppfyller kriteriet.

Oppgave 3 – Lagring og systemkatalog – 20%

- a) Beskriv to forskjellige formater for poster (records), dvs. et format for fastlengde poster og et format for variabelle lengde poster.
- b) Beskriv to forskjellige formater for blokker (pages), dvs. et format hvor det er fastlengde poster og et format hvor det er variabelle lengde poster.
- c) Beskriv to måter å implementere heapfil på.
- d) Hvilken informasjon finnes i systemkatalogen/SQL dictionary?

Oppgave 4 – Queryevaluering – 5%

Nevn to grunner for at en queryoptimalisator som regel kun jobber med såkalte "left-deep plans"?

Oppgave 5 – Logging og Recovery – 15%

- a) Forklar hensikten med de forskjellige feltene i en update-loggpost: **prevLSN, transID, type, pageId, lenght, offset, before-image, after-image**.
- b) Forklar begrepet "Compensating Log Record".
- c) Beskriv hva som skjer under de forskjellige fasene av recovery etter systemkrasj.

Oppgave 6 – Normalisering – 20 %

Gitt følgende tabell for å registrere resultater fra lokale langrenn:

Navn	Renn	Tid	Klubb	Arena	Plassering	Klasse
Jon Bang	1	24.58	Strindheim	Granåsen	1	M2
Jon Bang	2	36.12	Strindheim	Nilsbyen	1	M2
Per Dal	2	37.45	Byåsen IL	Nilsbyen	2	M2
Jon Bang	3	39.15	Strindheim	Granåsen	1	M2
Gry Sand	2	27.34	Nidelv	Nilsbyen	2	K3
Eva Nord	2	26.59	Byåsen IL	Nilsbyen	1	K3

- Hvilke funksjonelle avhengigheter (eng: functional dependencies) gjelder for tabellen beskrevet over? Gjør rede for de forutsetningene du eventuelt bygger på.
- Tabellen har en rekke uheldige egenskaper (anomalier). Gjør rede for disse og gi eksempler på problemer som kan oppstå.
- 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).