

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



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

Faglig kontakt under eksamen: Svein Erik Bratsberg og Roger Midtstraum

Tlf.: 99539963 (Bratsberg) og 99572420 (Midtstraum)

Eksamensdato: 10. august 2011

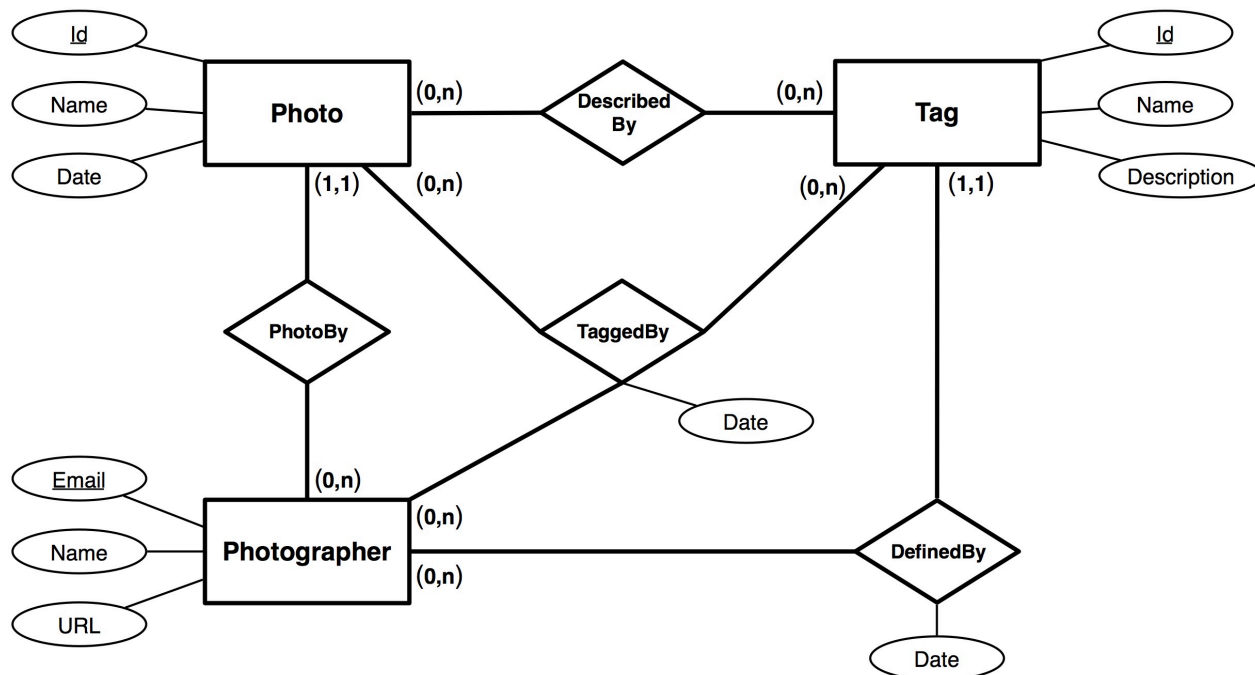
Eksamenstid: 09.00-13.00

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

Språkform: Bokmål

Sensurdato: 31. august 2011

Oppgave 1 – ER- og relasjonsmodell – 10 %



Ta utgangspunkt i ER-diagrammet over. Lag et relasjonsskjema som i størst mulig grad samsvarer med denne ER-modellen. For hver tabell i relasjonsskjemaet skal du gi opp attributter, primærnøkler, fremmednøkler og evt. andre restriksjoner. Gjør rede for eventuelle antagelser som du finner det nødvendig å gjøre.

Oppgave 2 – ER-modeller – 10 %

Ved spesialisering av entitetsklasser kan vi ha *restriksjoner* (Engelsk: constraints) av typen total/delvis (Engelsk: total/partial) og av typen overlappende/disjunkt (Engelsk: overlapping/disjoint). Dette gir fire klasser av spesialiseringer. Gi et eksempel på en spesialisering av hver klasse og forklar for hvert eksempel hvilke forutsetninger som må gjelde i miniverdenen for at eksempelet skal være gyldig.

Oppgave 3 – Relasjonsalgebra og SQL – 20 %

Ta utgangspunkt i følgende relasjonsdatabase (primærnøkler er understreket):

Student(Studentnr, Studentnavn, Adresse, Postnr, Poststed, Kjønn, Fødselsår)

Emne(Emnenr, Emnenavn, Studiepoeng)

Eksamen(Emnenr, Eksamensdato, OppgaveURL)

Sensur(Studentnr, Emnenr, Eksamensdato, Poengsum, Karakter)

Forutsetninger(Emnenr, Byggerpaaemnenr)

I Forutsetninger-tabellen representerer Emnenr et emne som forutsetter et emne med Byggerpaaemnenr.

- a) Lag et *ER-diagram* som i størst mulig grad samsvarer med relasjonsskjemaet. Gjør rede for eventuelle antagelser du finner det nødvendig å gjøre.
- b) Lag en spørring i *relasjonsalgebra* som finner alle studenter (Studentnr, Studentnavn, Eksamensdato og Karakter) som har tatt eksamen i emnet med Emnenr TDT4145.
- c) Lag en spørring i *SQL* som finner alle emner (Emnenr og Emnenavn) som er en direkte forutsetning for emnet med emnenr TDT4145 (NB! Du skal *ikke* nøste ut eventuelle transitive avhengigheter som for eksempel at TDT4145 forutsetter TDT4100 som forutsetter TDT4105).
- d) Lag en *SQL-spørring* som finner laveste, høyeste og gjennomsnittlig poengsum for studentene som fikk karakteren C ved eksamenen i TDT4145 som ble holdt 1. juni 2011.
- e) Ta utgangspunkt i eksamenen i TDT4145 som ble holdt 1. juni 2011. Sett karakteren til 'A' for alle studenter som fikk mer enn 87 poeng ved denne eksamenen.
- f) Lag en *SQL-spørring* som finner alle studenter (Studentnr, Studentnavn) som har gjort sin beste eksamensprestasjon (høyeste poengsum) ved en eksamen i TDT4145. NB! Det er greit at studenten har eksamensprestasjoner i andre emner som er like gode.

Oppgave 4 – Normaliseringsteori – 20 %

- a) La X være en mengde attributter. Forklar hva som menes med *tillukningen* (Engelsk: closure) til mengden med attributter, X^+ .

Ta utgangspunkt i tabellen R(a b c d) der følgende funksjonelle avhengigheter gjelder,

$F = \{ a \rightarrow b, bc \rightarrow d \}$.

- b) Beregn følgende tillukninger: a^+ og ac^+ .
- c) Forklar begrepet kandidatnøkkel. Finn alle kandidatnøkler i R.
- d) Anta at R oppfyller første normalform (1NF). Hva er den høyeste normalformen som oppfylles av R? Svaret må begrunnes.
- e) (10 %) Ta utgangspunkt i tabellen

Hund(Registreringsnr, Rase, Hundegruppe, Fødselsår, Kjønn) med avhengighetene

$F = \{ \text{Registreringsnr} \rightarrow \text{Rase}, \text{Registreringsnr} \rightarrow \text{Fødselsår}, \text{Registreringsnr} \rightarrow \text{Kjønn},$
 $\text{Rase} \rightarrow \text{Hundegruppe} \}$

- i) Anta at Hund oppfyller første normalform. Forklar hvorfor tabellen *ikke* oppfyller Boyce-Codd-normalform (BCNF) og avgjør hva som er den høyeste normalformen som oppfylles av tabellen.
- ii) Finn en oppdeling av tabellen der alle del-tabellene er på BCNF.
- iii) Oppfyller den foreslåtte oppdelingen (ii) kravene om tapsløst-join, attributtbevaring og bevaring av funksjonelle avhengigheter? Svaret må begrunnes.
- iv) Diskuter fordeler og ulemper med den foreslåtte oppdelingen i forhold til å beholde den opprinnelige tabellen.

Oppgave 5 – Lagring og indekser – 25 %

Gitt at det er definert en tabell på følgende måte i SQL:

```
CREATE TABLE Ansatt (  
    anr INT NOT NULL PRIMARY KEY,  
    fornavn VARCHAR(20),  
    etternavn VARCHAR(30),  
    epost CHAR(8)UNIQUE  
);
```

- a) Forklar en måte (postformat) å lagre *ansattposter* (records) slik at den utnytter plassen i blokker best mulig. Vis hvordan denne ansattposten ser ut ved å lage en figur som viser hvert felt i posten:

```
INSERT INTO Ansatt VALUES (110, 'Per', 'Hansen', 'perh');
```

- b) De aller fleste aksesser av ansattposter i denne databasen er via *epost*. Foreslå en lagringsmåte for denne tabellen hvor oppslag via *epost* er spesielt effektivt.
- c) Forklar hvorfor en clustered hashindeks på *fornavn* er en dårlig ide for denne tabellen
- d) Forklar en måte å sørge for at feltet *anr* er unikt ved hjelp av indekser. Begrunn svaret.
- e) Anta du har en blokkstørrelse på 8192 bytes og at det finnes 5000 ansattposter. Hvor mange blokker vil din foreslåtte lagringstruktur i b) og d) til sammen utgjøre? Forklar hvilke antakelser du gjør.

Oppgave 6 – Transaksjoner – 15 %

- a) (10%) Beskriv de fire isolasjonsnivåene for transaksjoner i SQL og forklar hvordan de kan implementeres ved hjelp av låser.
- b) (5%) Når en transaksjon committer må effekten av dens endringer gjøres permanente (Engelsk: durable). Forklar hvorfor NO-FORCE er mer fordelaktig enn FORCE.