

# Andre sett obligatoriske oppgaver i INF3100 V2014

Oppgavesettet skal i utgangspunktet løses av grupper på to og to studenter. Vi godkjenner også individuelle besvarelser, men oppfordrer dere altså til heller å finne en å samarbeide med. Vi godkjenner ikke grupper med mer enn to studenter. Vi vil ved behov foreta stikkprøver der vi plukker ut enkeltpersoner som må gjennomgå sin besvarelse med oss.

Gjennomføring og innlevering av oppgaven skal skje i henhold til gjeldende retningslinjer ved Institutt for informatikk, se

<http://www.mn.uio.no/ifi/studier/admin/obliger/>

*Enhver innlevering av besvarelse på en obligatorisk oppgave tas som en bekref-  
telse på at retningslinjene er lest og forstått.*

**Innleveringsfrist: Fredag 2. mai kl. 23.59.**

*Fristen er absolutt, og det blir ikke gitt utsettelse. Alle spørsmålene må besvares  
for å få godkjent besvarelsen.*

## Oppgave 1–5 Eksamen 2013

Løs oppgave 1 (FDer og MVDer), oppgave 2 (SQL), oppgave 3 (relasjonsalgebra), oppgave 4 (transaksjonsprotokoller) og oppgave 5 (distribuerte transaksjoner) fra eksamen i INF3100 våren 2013. (Scenario: Romreservasjoner.)

## Oppgave 6 Implementasjon

Folkeregisteret inneholder informasjon om alle innbyggere i Norge og hvor de bor. Alle personer har et entydig 11-sifret fødselsnummer **fnr**. Når noen

flytter, må de sende inn en flyttemelding til folkeregisteret. Flyttemeldingen kan omfatte mer enn én person, f.eks. hvis en hel familie flytter. Hver flyttemelding er identifisert ved et løpenummer *mid*. Forøvrig inneholder den informasjon om flyttedato, gammel og ny adresse. Anta at denne informasjonen er organisert i følgende relasjoner der primærnøklerne er understreket:

```
Person(fnr, fornavn, etternavn)
Flyttemelding(mid, flyttedato, fraadr, tiladr)
FlyttetPerson(mid, fnr)
```

Betrakt følgende SQL-spørring som finner alle flyttemeldingene (i form av løpenummer og flyttedato) til Jo Å:

```
SELECT Flyttemelding.mid, Flyttemelding.flyttedato
FROM   Person, Flyttemelding, FlyttetPerson
WHERE  Person.fnr = FlyttetPerson.fnr AND
       Flyttemelding.mid = FlyttetPerson.mid AND
       Person.fornavn = 'Jo' AND
       Person.etternavn = 'Å';
```

## 6a Parsering

- (i) Bruk den enkle grammatikken under til å lage et parseringstre for spørringen over.

Elementære syntaktiske kategorier som `<attribute>`, `<relation>` og `<pattern>` oversettes med henholdsvis navnet på attributtet, navnet på relasjonen og en streng i enkle anførselstejn.

```
<query> ::= <SFW>
```

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

```
<selList> ::= <attribute>
```

```
<selList> ::= <attribute>, <selList>
```

```
<fromList> ::= <relation>
```

```
<fromList> ::= <relation>, <fromList>
```

```
<condition> ::= <condition> AND <condition>
```

```
<condition> ::= <attribute> = <attribute>
```

```

<condition> ::= <attribute> = <pattern>
<condition> ::= <attribute> LIKE <pattern>
<condition> ::= <attribute> IS NULL

```

## 6b Logisk spørreplan

- (i) Konverter parsingstreet i oppgave 5a til en logisk spørreplan i relasjonsalgebra (tegn uttrykkstreet).
- (ii) Optimer den logiske spørreplanen i punkt (i) hvis den ikke alt er på optimal form (tegn det nye uttrykkstreet).

## 6c Datalagring

Anta at vi har en disk med følgende spesifikasjoner for lagring av våre data:

Diskplater:	10 (med 2 overflater hver)
Spor:	10.000 pr. overflate
Antall sektorer pr. spor:	1000 (en ikke-sonet disk)
Byte pr. sektor:	512
Byte pr. "gap":	64
Gjennomsnittlig søketid:	3,7 ms
Spor-spor søk:	0,1 ms
Rotasjonshastighet:	15.000 RPM

Databasen inneholder følgende antall tupler hva gjelder de to relasjonene **Person** og **Flyttmelding**:

<b>Person:</b>	5.000.000
<b>Flyttmelding:</b>	10.000.000

I tillegg gjelder følgende informasjon om diverse størrelser:

- Hver blokk har en "header" (hode) på 20 byte.
- Hver "record" (post) har et hode på 10 byte.
- Hvert attributt har følgende størrelse i antall bytes:

Person		Flyttemelding	
fnr:	4	mid:	4
fornavn:	64	flyttedato:	4
etternavn:	64	fraadr:	128
		tiladr:	128

- (i) Hva er diskens utnyttbare kapasitet?
- (ii) Hvilke faktorer inngår i å aksessere en blokk på disken, og hva er gjennomsnittlig aksestid for en vilkårlig 4 Kbyte (4096 byte) blokk?
- (iii) Hvor stor plass trenger disse relasjonene på disken i tilfellet “unspanned” lagring (dvs. hvis ingen enkelt post er delt over flere blokker)?
- (iv) Hvor lang tid tar det å lese hele relasjonen **Person** uavbrutt hvis vi antar vilkårlig plassering av data i diskblokker på disken?
- (v) Hvis pekere er på 8 byte, hvor mange blokker må aksesserer og hva er gjennomsnittlig aksestid totalt for å finne ett tuppel (én post) i **Person** for en gitt verdi av attributtet **etternavn** hvis **person** har tett indeks på **etternavn** og indeksen er realisert ved et  $B^+$ -tre?

## Oppgave 7 Logging

Det semantiske innholdet av transaksjonen  $T_3$  i oppgave 4 består av følgende operasjoner der  $x$ ,  $y$  og  $z$  er lokale arbeidsvariable for  $T_3$  og derfor ikke skal logges:

$$T_3 : x := a; y := b; z := c; a := x + y; c := z + 1;$$

- (i) Beskriv postene i undo/redo-loggen for transaksjonen  $T_3$  når vi initielt har verdiene  $a = 150$ ,  $b = 20$  og  $c = 6$ .
- (ii) Når skal de forskjellige typene loggposter skrives til disk?

## Slutt på obligatorisk oppgave 2