

## i TDT4145 spring 2018

Institutt for datateknologi og informatikk

### Eksamensoppgave i TDT4145 Datamodellering og databasesystemer

Faglig kontakt under eksamen:

Roger Midtstraum, mobil: 995 72 420

Svein Erik Bratsberg, mobil: 995 39 963

Eksamensdato: 30. mai 2018

Eksamenstid (fra-til): 09:00 - 13:00

Hjelpemiddelkode/Tillatte hjelpemidler:

D – Ingen trykte eller håndskrevne hjelpemidler tillatt.

Bestemt, enkel kalkulator tillatt.

Annen informasjon:

Merk! Studenter finner sensur i Studentweb. Har du spørsmål om din sensur må du kontakte instituttet ditt. Eksamenskontoret vil ikke kunne svare på slike spørsmål.

## 1 Problem 1 Data Models (19 %)

Lag en ER-modell (du kan bruke alle virkemidler som er med i pensum) for følgende forenklete beskrivelse av Sportsavisens database for eliteserien i fotball.

Eliteserien i fotball består av et antall lag. Hvert lag har en hjemmearena der lagets hjemmekamper spilles. Serien er organisert slik at hvert lag spiller to kamper mot alle de andre lagene; en hjemmekamp på egen arena og en bortekamp på motstanderens arena. Lagene har en unik lag-kode, lagnavn og stiftelsesår. Hver arena har et unikt arenanavn og maksimum tilskuerkapasitet. En arena er hjemmearena for ett eller to eliteserielag. Hver kamp i serien spilles på en bestemt arena og har et unikt kampnummer. For hver kamp registreres serierunde, kampdato, starttidspunkt, antall tilskuere, antall mål til hjemmelaget og antall mål til bortelaget. Hver kamp har oppnevnt fire dommere; en hoveddommer, to linjedommere og en dommerassistent. Hver dommer har en unik person-id, navn, fødselsår, mobilnummer og dommerstatus (kretsdommer, forbundsdommer eller internasjonal dommer). En dommer kan være medlem i ett av lagene i eliteserien. I så fall skal dette kunne registreres.

Lagene har en hovedtrener, kan ha en eller flere hjelpetrener og kan ha en målvakttrener. Trenerne registreres med en unik person-id, navn, fødselsår og mobilnummer. Lagene må ha minst 11 spillere som hver registreres med en unik person-id, navn, fødselsår og mobilnummer. Hver spiller har et fast trøyenummer og er enten utespiller eller målvakt. For hver kamp registreres kamplengden i minutter og hvilke spillere som deltar i spillet. Hver spiller får registrert spilletid (fra, til), antall mål, antall selvmål og om spilleren fikk gult eller rødt kort. Sportsavisen gir karakter for spillere som spilte lenge nok til å kunne vurderes. Karakterene er på en skala fra 1 til 10, der 1 er elendig og 10 er verdensklasse.

Dommere kan ikke delta som spiller eller trener for lagene i eliteserien. Det kan finnes spillere som også er trener for det laget de spiller på. Vi registrerer bare spillere og trenere som er tilknyttet et av lagene i eliteserien. Det er ikke anledning til å spille for flere lag i serien eller å være trener for flere lag. Vi velger å se bort fra at lag kan bytte trenere eller spillere i løpet av en sesong.

Gjør kort rede for eventuelle forutsetninger som du finner det nødvendig å gjøre.

Denne oppgaven må løses på papir.

Maks poeng: 19

## 2 Problem 2a (4 %)

Ta utgangspunkt i følgende relasjonsdatabase-skjema (primærnøkler er understreket):

**Book**(ID, BookTitle, PublishedYear, Price, PublisherID)

**Author**(ID, Firstname, Surname, Nationality, URL)

**Publisher**(ID, PublisherName, URL)

**Genre**(ID, GenreTitle, Description)

**BookAuthor**(BookID, AuthorID)

**BookGenre**(BookID, GenreID)

Lag et *ER-diagram* (du kan bruke alle virkemidler som er med i pensum) som i størst mulig grad samsvarer med relasjonsskjemaet. Gjør rede for eventuelle antagelser som du finner det nødvendig å gjøre.

Denne oppgaven må løses på papir.

Maks poeng: 4

## 3 Problem 2b (4 %)

Ta utgangspunkt i følgende relasjonsdatabase-skjema (primærnøkler er understreket):

```
Book(ID, BookTitle, PublishedYear, Price, PublisherID)
```

Author(ID, Firstname, Surname, Nationality, URL)

**Publisher(ID, PublisherName, URL)**

Genre(ID, GenreTitle, Description)

**BookAuthor (BookID, AuthorID)**

**BookGenre (BookID, GenreID)**

Lag en *SQL-spørring* som finner alle sjangre (genre) som har ordet *fantasy* i sjangerbeskrivelsen (Description). I resultatet skal du ta med både sjangeridentifikator og sjangertittel.

**Skriv ditt svar her...**

Maks poeng: 4

#### 4 Problem 2c (4 %)

Ta utgangspunkt i følgende relasjonsdatabase-skjema (primærnøkler er understreket):

**Book(ID, BookTitle, PublishedYear, Price, PublisherID)**

Author(ID, Firstname, Surname, Nationality, URL)

**Publisher(ID, PublisherName, URL)**

Genre(ID, GenreTitle, Description)

BookAuthor (BookID, AuthorID)

**BookGenre (BookID, GenreID)**

Lag en *SQL-spørring* som finner antall bøker som er gitt ut på hvert forlag (publisher). I svaret skal du ta med både identifikator og navn for hvert forlag.

**Skriv ditt svar her...**

Maks poeng: 4

**5 Problem 2d (4 %)**

Ta utgangspunkt i følgende relasjonsdatabase-skjema (primærnøkler er understreket):

**Book(ID, BookTitle, PublishedYear, Price, PublisherID)**

**Author(ID, Firstname, Surname, Nationality, URL)**

**Publisher(ID, PublisherName, URL)**

```
Genre(ID, GenreTitle, Description)
```

BookAuthor (BookID, AuthorID)

**BookGenre (BookID, GenreID)**

Lag en *SQL-setning* som legger inn den danske forfatteren Kim Leine i forfattertabellen. Han har forfatterid 100 og har ingen kjent URL.

**Skriv ditt svar her...**

Maks poeng: 4

6 **Problem 2e (4 %)**

Ta utgangspunkt i følgende relasjonsdatabase-skjema (primærnøkler er understreket):

Book(ID, BookTitle, PublishedYear, Price, PublisherID)

**Author (ID, Firstname, Surname, Nationality, URL)**

Publisher(ID, PublisherName, URL)

```
Genre(ID, GenreTitle, Description)
```

BookAuthor (BookID, AuthorID)

**BookGenre (BookID, GenreID)**

Lag en spørring i *relasjonsalgebra* som finner alle forfattere (ID, Firstname, Surname og Nationality) som har skrevet en eller flere bøker sammen med Linn Ullmann. Pass på at Linn Ullmann selv ikke blir med i resultatet. Du kan gå ut fra at det bare finnes en forfatter som heter Linn Ullmann.

Relasjonsalgebra kan formuleres som tekst eller graf. Hvis du behersker begge notasjonene foretrekker vi at du svarer med graf, men du blir ikke trukket for å svare med tekst.

Denne oppgaven må løses på papir.

Maks poeng: 4

## 7 Problem 3a (4 %)

I friidrettsøvelsen lengdehopp får hver utøver et antall forsøk (attempt) på å hoppe lengst mulig. Hver utøver er identifisert ved et unikt startnummer. For hvert forsøk blir det registrert hvor mye medvind (wind) hopperen hadde og hvor langt det ble hoppet. Under er vist et eksempel på en tabellforekomst for en lengdehopp-konkurranse.

StartNumber	Attempt	Length	Wind
1	1	3,10	1,5
2	1	3,10	2,0
1	2	3,20	2,5
2	2	3,30	2,5

Ut fra denne tabellforekomsten er det mulig at Wind  $\rightarrow$  Attempt kan være en restriksjon for tabellen. Forklar hvorfor det er mulig. Er det rimelig å anta at denne funksjonelle avhengigheten gjelder i den aktuelle miniverdenen? Formuler med ord hvilken restriksjon som i så fall ville gjelde.

**8 Problem 3b (4 %)**

<https://ntnu.inspera.no/admin#/author/test/30303247>

9 Problem 3c (4 %)

<https://ntnu.inspera.no/admin#/author/test/30303247>



**10 Oppgave 3d (4 %)**

<https://ntnu.inspera.no/admin#/author/test/30303247>

**Skriv ditt svar her...**

Maks poeng: 4

### 11 Problem 4 (5 %)

Studenter er identifisert med student-ID og har et navn. Emner er identifisert med emne-ID og har et emnenavn. En student kan ta eksamen i ett eller flere emner. For hver eksamen som en student tar, lagres karakter og dato, slik at vi kan skille mellom flere forsøk i samme emne. En student kan være registrert i et emne, i så fall skal registreringsdato kunne registreres. Det kan finnes emner uten registrerte studenter og det trenger ikke å finnes studenter som har tatt eksamen i et emne. Et emne har et øvingsopplegg for hvert år som emnet undervises. Øvingsopplegg identifiseres ved en unik øvingsopplegg-id. Dersom en student har gjennomført et øvingsopplegg, skal dette kunne registreres og det skal være mulig å registrere status (godkjent / ikke godkjent).

I figuren under har vi vist resultatet av et forsøk på å lage en datamodell for denne miniverdenen. I denne modellen er det minst fem feil. Finn feilene og forklar for hver feil hva som er problemet.



Maks poeng: 5

## 11/19

Anta at det er plass til tre nøkler i hver blokk, og at det er plass til tre nøkler og fire pekere i hver blokk med nivå  $> 0$ . Vis tilstanden til B+-treet hver gang du skal til å splitte en blokk. Vis også tilstanden til B+-treet til slutt.

Denne oppgaven må løses på ark.

Maks poeng: 10

**13 Problem 6a (2.5 %)**

Anta følgende tabell er lagret i et Clustered B+-tre med den sammensatte nøkkelen (exno, studno) som søkenøkkel:

Exercise (exno, studno, datedelivered, approval\_status)

Det er 1000 blokker på løvnivå (nivå=0) i B+-treet og det er tre nivå med blokker i treet.

Hvor mange blokker aksesseres ved utføring av det følgende query. Begrunn svaret ditt.

```
SELECT datedelivered, approval_status
FROM Exercise
WHERE exno=2 AND studno=123456;
```

**Skriv ditt svar her...**

Maks poeng: 2.5

## 14 Problem 6b (2.5 %)

Anta følgende tabell er lagret i et Clustered B+-tre med den sammensatte nøkkelen (exno, studno) som søkenøkkel:

Exercise (exno, studno, datedelivered, approval\_status)

Det er 1000 blokker på løv nivå ( $nivå=0$ ) i B+-treet og det er tre nivå med blokker i treet.

Hvor mange blokker aksesseres ved utføring av det følgende query. Begrunn svaret ditt.

```
SELECT *
FROM Exercise
WHERE studno=123456 AND approval_status="Disapproved";
```

**Skriv ditt svar her...**

[illegible]

Maks poeng: 2.5

15 **Problem 6c (2.5 %)**

Anta følgende tabell er lagret i et Clustered B+-tre med den sammensatte nøkkelen (exno, studno) som søkenøkkel:

Exercise (exno, studno, datedelivered, approval\_status)

Det er 1000 blokker på løvnivå (nivå=0) i B+-treet og det er tre nivå med blokker i treet.

Hvor mange blokker aksesseres ved utføring av det følgende query. Begrunn svaret ditt.

```
SELECT exno, studno, approval_status
FROM Exercise
ORDER BY exno ASC, studno ASC;
```

**Skriv ditt svar her...**

Maks poeng: 2.5

16 **Problem 6d (2.5 %)**

Anta følgende tabell er lagret i et Clustered B+-tre med den sammensatte nøkkelen (exno, studno) som søkenøkkel:

Exercise (exno, studno, datedelivered, approval\_status)

Det er 1000 blokker på løvnivå (nivå=0) i B+-treet og det er tre nivå med blokker i treet.

Hvor mange blokker aksesseres ved utføring av det følgende query. Begrunn svaret ditt.

```
SELECT exno, count(*)
FROM Exercise
WHERE approval_status="Approved"
GROUP BY exno;
```

**Skriv ditt svar her...**

Maks poeng: 2.5

**17 Problem 7 (5 %)**

Avgjør recoveryegenskapene (ikke gjenopprettbar, gjenopprettbar, ACA og strikt) ved de følgende historiene:

S1: r2(Z); w2(X); w2(Y); c2; w1(X); r1(X); c1; r3(X); r3(Z); r3(Y); c3;

S2: r1(X); r2(X) w2(X); w2(Y); c2; w1(X); r3(Z); w1(Y); c1; r3(Y); r3(X) c3;

S3: r2(X); w2(X); r1(X); r2(Y); w1(Y); c1; r2(Z); w2(Z); c2;

Begrunn svarene dine.

**Skriv ditt svar her...**

Maks poeng: 5

## 18 Problem 8 (5 %)

Anta følgende historie som det skal settes låser for:

S3: r2(X); w2(X); r1(X); r2(Y); w1(Y); c1; r2(Z); w2(Z); c2;

Vi antar tofaselåsing (2PL) av typen rigorous. Operasjonene kommer inn i databasen i den rekkefølgen som er angitt, men pga. låsene kan historien skje i en annen rekkefølge. Innfør både lese- og skrive-låser. Hvis en transaksjon må vente på en lås, kan en annen transaksjon senere i historien likevel prøve å fortsette:

Vis hvordan historien utføres når du innfører read\_lock, write\_lock og unlock som operasjoner i historien.

Tegn utførelsen på et ark.

Maks poeng: 5

**19 Problem 9a (5 %)**



LSN	PrevLSN	TransactionID	Operation	PageID
237			End_ckpt	
238	NULL	T1	Update	A
239	NULL	T2	Update	B
240	239	T2	Commit	
241	238	T1	Update	A
242	NULL	T3	Update	B

Hvordan ser DPT og TT ut etter at analysen i Recovery er ferdig? DPT har feltene PageId og RecoveryLSN. TT har feltene TransactionID, LastLSN og TransactionStatus.

17/19

**20 Problem 9b (5 %)**

Anta følgende logg blir funnet etter en krasj:

LSN	PrevLSN	TransactionID	Operation	PageID
237			End_ckpt	
238	NULL	T1	Update	A
239	NULL	T2	Update	B
240	239	T2	Commit	
241	238	T1	Update	A
242	NULL	T3	Update	B

Anta at Dirty Page Tabel (DPT) og transaksjonstabellen (TT) i sjekkpunktloggposten med LSN 237 (End\_ckpt) er tomme.

Hva skjer under undofasen av recovery? Hvilke transaksjoner rulles tilbake og hvilke nye loggposter lages?

**Skriv ditt svar her...**

Maks poeng: 5

