

i Front page

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: 24. mai 2019

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 (15 %)

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

Vi har studieprogram som har en unik studieprogramkode, studieprogramnavn og et antall semester. Studieprogram kan være av ulike typer: Årsstudier som går over to semester, treårige bachelorprogram som går over seks semester, toårige masterprogram som går over fire semester og femårige studieprogram som går over 10 semester. Innen et studieprogram nummereres semestrene fra 1 og oppover. Alle semester har i tillegg en kode som viser om det er høst- eller vårsemester.

Et semester i et program består av et antall emner som har en unik emnekode, emnenavn og omfang målt i studiepoeng. Emner kan inngå i ulike studieprogram. Et emne kan være et obligatorisk eller et valgfritt emne for et studieprogram. Et emne kan være obligatorisk for noen studieprogram og valgfritt emne for andre studieprogram. Emner har en semesterkode som viser om det undervises høst, vår eller begge. Emner trenger ikke å være i studieplanen til noen program.

Læringsutbyttebeskrivelser (LUB-er) beskriver det en person vet, kan og er i stand til å gjøre som et resultat av en læringsprosess. Læringsutbyttebeskrivelser utarbeides på to nivåer, for studieprogram som helhet og på emnenivå. Læringsutbyttebeskrivelser deles i de tre kategoriene kunnskaper, ferdigheter og generell kompetanse. For et studieprogram finnes det et antall overordnede læringsutbyttebeskrivelser. Tilsvarende finnes det et antall læringsutbyttebeskrivelser på emnenivå for alle emner. En læringsutbyttebeskrivelse, både overordnede og på emnenivå, består av en unik læringsutbytte-identifikator (LUB-ID), en beskrivende tekst og et felt som viser om den gjelder kunnskap, ferdighet eller generell kompetanse. En læringsutbyttebeskrivelse på overordnet nivå gjelder bare for ett studieprogram. Tilsvarende vil en læringsutbyttebeskrivelse på emnenivå gjelde for bare ett bestemt emne.

Vi ønsker å ha oversikt over hvordan læringsutbyttebeskrivelser på emnenivå bidrar til å oppfylle overordnede læringsutbyttebeskrivelser på studieprogramnivå. Vi skal derfor kunne registrere at et emne bidrar til et studieprogram gjennom at en eller flere av emne-LUB-ene bidrar til en eller flere studieprogram-LUB-er.

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

Denne oppgaven må løses på papir.

Maks poeng: 15

2 **Problem 2 (4 %)**

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

```
Course(CourseCode, CourseTitle, CourseCredit, CourseDecription)
Student(StudentNo, FirstName, LastName, Gender)
Exam(ID, ExamDay, ExamMonth, ExamYear, CourseCode, Duration)
    CourseCode is foreign key referencing Course, cannot take NULL value
ExamResults(ExamID, StudentNo, Grade)
    ExamID is foreign key referencing Exam, cannot take NULL value
    StudentNo is foreign key referencing Student, cannot take NULL value
```

Lag relasjonsalgebra som finner StudentNo, FirstName, Lastname, ExamYear og CourseCode for alle studenter som har fått karakteren «A» i emner som heter «Datamodellering og databasesystemer».

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

3 Problem 3 (4 %)

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

Course(CourseCode, CourseTitle, CourseCredit, CourseDescription)

Student(StudentNo, FirstName, LastName, Gender)

Exam(ID, ExamDay, ExamMonth, ExamYear, CourseCode, Duration)

CourseCode is foreign key referencing Course, cannot take NULL value

ExamResults (ExamID, StudentNo, Grade)

ExamID is foreign key referencing Exam, cannot take NULL value

StudentNo is foreign key referencing Student, cannot take NULL value

Skriv SQL som finner antall eksamensresultat i TDT4145 (CourseCode) for hvert av årene det har blitt arrangert eksamen i emnet. Resultatet skal sorteres på eksamensår i synkende rekkefølge.

Skriv ditt svar her...

Maks poeng: 4

4 Problem 4 (4 %)

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

Course(CourseCode, CourseTitle, CourseCredit, CourseDescription)

Student(StudentNo, FirstName, LastName, Gender)

Exam(ID, ExamDay, ExamMonth, ExamYear, CourseCode, Duration)

CourseCode is foreign key referencing Course, cannot take NULL value

ExamResults(ExamID, StudentNo, Grade)

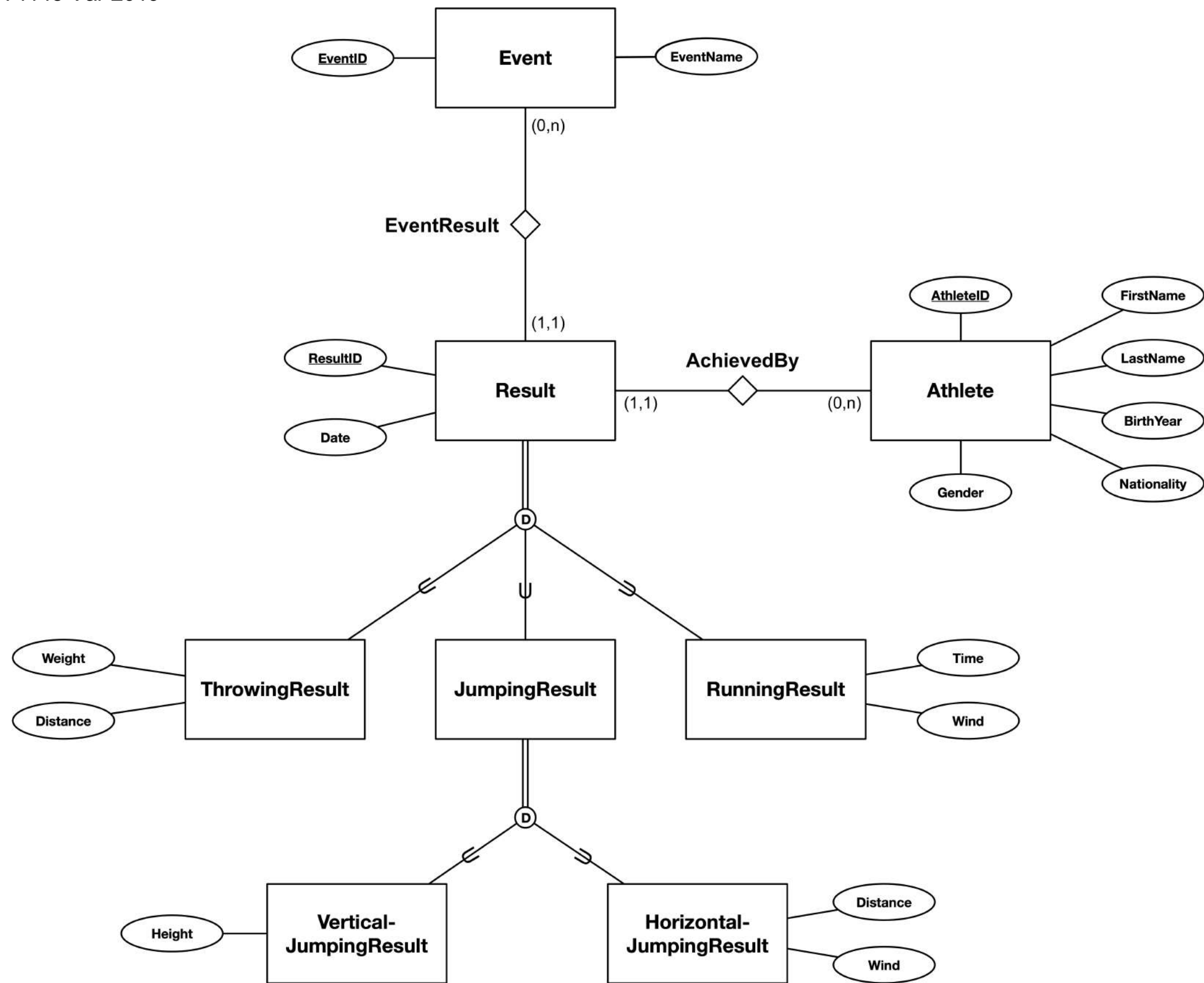
ExamID is foreign key referencing Exam, cannot take NULL value

StudentNo is foreign key referencing Student, cannot take NULL value

Skriv SQL som finner CourseCode og CourseTitle for alle emner der det er arrangert minst en eksamen og ingen student ennå har fått karakteren «A».

Skriv ditt svar her...

Maks poeng: 4



Oversett denne ER-modellen til et tilsvarende relasjonsdatabase-skjema (modell). Gi en begrunnelse for hvorfor din løsning bør foretrekkes fremfor alternative løsninger.

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

Skriv ditt svar her...

Format | B | I | U | x₂ | x² | I_x | [Icons: Copy, Paste, Undo, Redo, Bulleted List, Numbered List, Link, Table, Highlight, Sum, Spell Check, Clear]

Words: 0

Maks poeng: 8

7 Problem 7 (2 %)

Tabeller (relasjoner) vil alltid ha minst en *nøkkel*. Forklar hvorfor det er slik.

Skriv ditt svar her...

Maks poeng: 2

8 Problem 8 (2 %)

Forklar hva som skiller begrepene *nøkkel* og *supernøkkel*.

Skriv ditt svar her...

Maks poeng: 2

11 **Problem 11 (3 %)**

I hvilke situasjoner er det hensiktsmessig å bruke *kategorier* (eng: category) i en EER-modell?

Skriv ditt svar her...

Format

B


I


U


x_2


x^2


I_x




















Ω





Σ

ABC

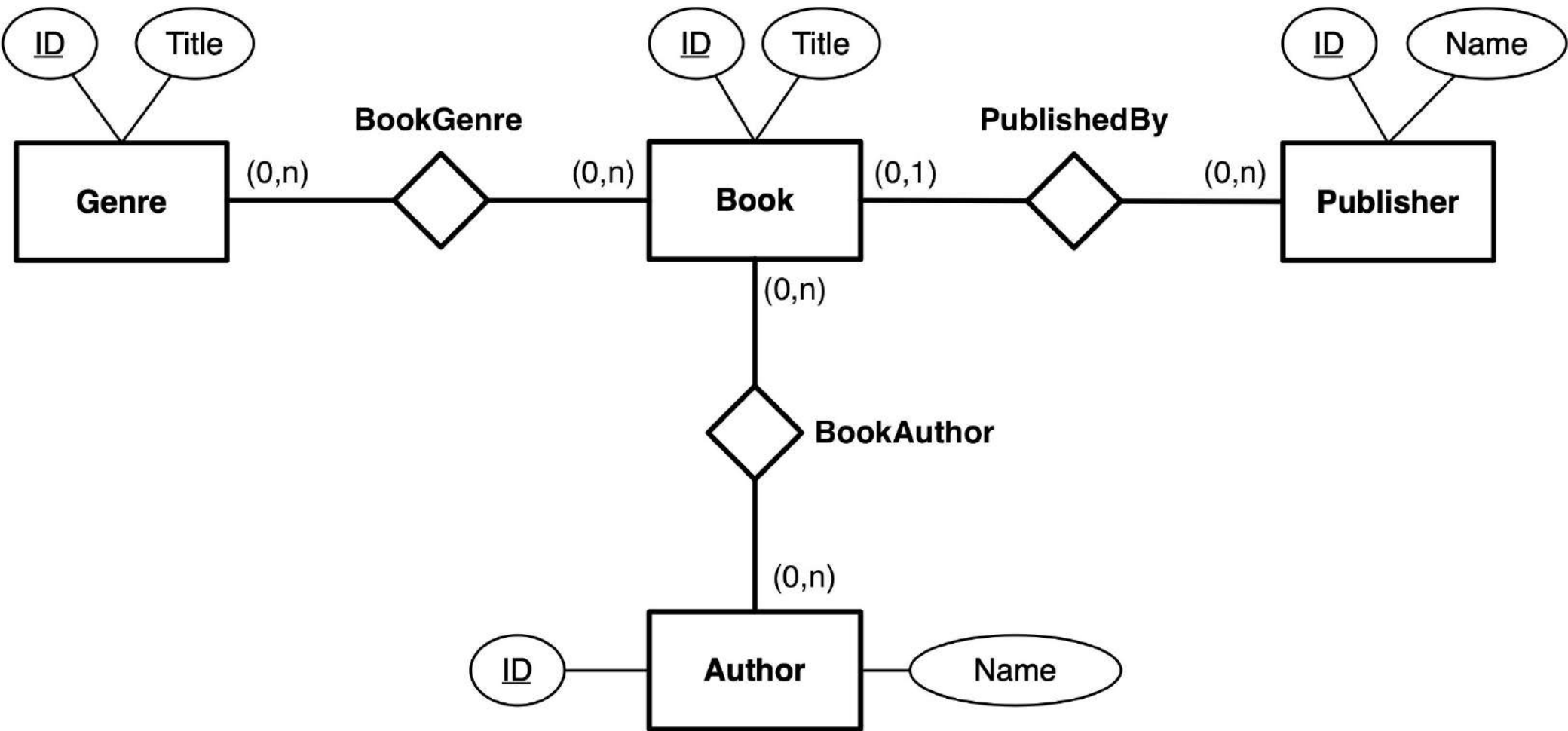


Words: 0

Maks poeng: 3

12 Problem 12 (4 %)

Ta utgangspunkt i følgende modell for informasjon om bøker:



Tabellene under viser tabellforekomster for en relasjonsdatabase som lagrer data i overenstemmelse med ER-modellen.

Book			Author		Publisher	
BookID	Title	PublisherID	AuthorID	Name	PublisherID	Name
1	Sult	1	1	Knut Hamsun	1	Gyldendal
2	Victoria	1	2	Hans Rosling	2	Cappelen
3	Factfulness	2	3	Daniel Kahneman	3	Penguin
4	Thinking, Fast and Slow	3			4	Oktober

Genre		BookGenre		BookAuthor	
GenreID	Title	BookID	GenreID	BookID	AuthorID
1	Poetry	1	2	1	1
2	Novel	2	2	2	1
3	Nonfiction	3	3	3	2
		4	3	4	3

Tegn et forekomstdiagram («instance diagram») for ER-modellen med utgangspunkt i dataene fra den oppgitte relasjonsdatabasen.

Denne oppgaven må løses på papir.

Maks poeng: 4

13 **Problem 13 (5 %)**

Vi har en *extendible hashing*-datastruktur. Når vi startet hadde den *global dybde* 2. Akkurat nå har den *global dybde* 5. Hvilke verdier kan *lokal dybde* ha i denne extendible hashing-datstrukturen? Forklar svaret ditt.

Skriv ditt svar her...

Format

B

I

U

x_2

x^2

I_x

Ω

Σ

ABC

Words: 0

Maks poeng: 5

15

Problem 15 (2.5 %)

Vi har en tabell

Student (pno, studno, lastname, firstname, email)

Tabellen er lagret som heapfil. Vi har en statisk hashindeks på **pno**. I tillegg har vi et unclustered B+-tree på **lastname**. Hver post i hashindeksen er på formatet **(pno, RecordID)** og hver post på løvnivå i B+-treet er på formatet **(lastname, RecordID, RecordID, ...)**, dvs. et variabelt antall RecordID per lastname.

Antall studenter i tabellen er 50 000. Heapfilen inneholder 1000 blokker. B+-treet inneholder 500 blokker på løvnivå, og har 3 nivåer. Hashindeksen inneholder 300 blokker med 60 overløpsblokker og gjennomsnittlig antall blokker per aksess av indeksen er 1.2.

Gi et overslag på hvor mange blokker som aksesseres ved følgende SQL-setning. Gi en begrunnelse for ditt svar.

SELECT * FROM Student WHERE pno=12121212345;

Skriv ditt svar her...

Format

B

I

U

x₂

x²

I_x

↶

↷

↺

☰

☷

Ω

Σ

ABC

Words: 0

Maks poeng: 2.5

16

Problem 16 (2.5 %)

Vi har en tabell

Student (pno, studno, lastname, firstname, email)

Tabellen er lagret som heapfil. Vi har en statisk hashindeks på **pno**. I tillegg har vi et unclustered B+-tree på **lastname**. Hver post i hashindeksen er på formatet **(pno, RecordID)** og hver post på løvnivå i B+-treet er på formatet **(lastname, RecordID, RecordID, ...)**, dvs. et variabelt antall RecordID per lastname.

Antall studenter i tabellen er 50 000. Heapfilen inneholder 1000 blokker. B+-treet inneholder 500 blokker på løvnivå, og har 3 nivåer. Hashindeksen inneholder 300 blokker med 60 overløpsblokker og gjennomsnittlig antall blokker per aksess av indeksen er 1.2.

Gi et overslag på hvor mange blokker som aksesseres ved følgende SQL-setning. Gi en begrunnelse for ditt svar.

SELECT * FROM Student WHERE lastname='Hansen';

Skriv ditt svar her...

Format

B

I

U

x₂

x²

I_x

↶

↷

↺

☰

☷

Ω

▢

Σ

ABC

✖

Words: 0

Maks poeng: 2.5

Vi har følgende sekvens av operasjoner som kommer inn til databasen. Vis hvordan denne sekvensen blir utført gitt at vi bruker tofaselåsing av typen rigorous og vranglåsoppdagelse (deadlock detection).

Skriv ditt svar her...

Words: 0

Maks poeng: 5

Vi har følgende sekvens av operasjoner som kommer inn til databasen. Vis hvordan denne sekvensen blir utført gitt at vi bruker tofaselåsing av typen rigorous og vranglåsoppdagelse (deadlock detection).

Skriv ditt svar her...

Maks poeng: 5

21 **Problem 21 (5 %)**

Anta ARIES-recovery.
Etter en krasj fant vi følgende loggposter:

	LSN	PrevLSN	TransId	Type	Pageld
10		0	T1	update	A
11		0	T2	update	B
12		10	T1	commit	
13				begin_checkpoint	
14				end_checkpoint	
15		0	T3	update	B
16		11	T2	update	C
17		16	T2	commit	

Følgende transaksjonstabell ble funnet i loggpost med LSN 14:

TransId	LastLSN	Status
T1	12	commit
T2	11	active

Følgende Dirty Page Table (DPT) ble funnet i loggpost med LSN 14:

Pageld	RecLSN
A	10
B	11

Vis hvordan DPT og transaksjonstabellen ser ut etter at analysen er ferdig.

Skriv ditt svar her...

Format

B

I

U

\times_2

\times^2

*I*_x

ABC

Words: 0

Maks poeng: 5

22 **Problem 22 (5 %)**

	LSN	PrevLSN	TransId	Type	PageId
10	0		T1	update	A
11	0		T2	update	B
12	10		T1	commit	
13				begin_checkpoint	
14				end_checkpoint	
15	0		T3	update	B
16	11		T2	update	C
17	16		T2	commit	

	TransId	LastLSN	Status
T1	12		commit
T2	11		active

	Pageld	RecLSN
A	10	
B	11	

	PageId	PageLSN
A	7	
B	15	
C	8	

