

## i Framside

Institutt for datateknologi og informatikk

**Eksamensoppgave i TDT4145 Datamodellering og databasesystemer**

**Eksamensdato: 17. august 2023**

**Eksamenstid (fra-til): 09:00 - 11:00**

**Hjelpemiddelkode/Tillatte hjelpemidler:**

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

**Faglig kontakt under eksamen:**

Svein Erik Bratsberg, Tlf.: 99539963

**Faglig kontakt møter i eksamenslokalet: JA**

### **ANNEN INFORMASJON:**

Skaff deg overblikk over oppgavesettet før du begynner på besvarelsen din.

Les oppgavene nøye, gjør dine egne antagelser og presiser i besvarelsen hvilke forutsetninger du har lagt til grunn i tolkning/avgrensing av oppgaven. Faglig kontaktperson skal kun kontaktes dersom det er direkte feil eller mangler i oppgavesettet. Henvend deg til en eksamensvakt hvis du ønsker å kontakte faglærer. Noter gjerne spørsmålet ditt på forhånd.

**Kommentaroppgave:** Dersom det er helt nødvendig å gi en melding til sensor har vi lagt til en "kommentaroppgave" (teller selvsagt ikke) til slutt der du kan forklare det du trenger. Vær så snill ikke å bruke denne oppgave til å legge inn generelle kommentarer på eksamen, det gjør du mye bedre i e-post til faglærer.

**InspiraScan:** I oppgave 1 og 2 er det lagt opp til å besvare på ark. Andre oppgaver skal besvares direkte i Inspira. Nederst i oppgaven finner du en sjusifret kode. Fyll inn denne koden øverst til venstre på arkene du ønsker å levere. Det anbefales å gjøre dette underveis i eksamen. Dersom du behøver tilgang til kodene etter at eksamenstiden har utløpt, må du klikke «Vis besvarelse».

**Vekting av oppgavene:** Vektingen av hver oppgave er angitt i overskriften til oppgaven.

Varslinger: Hvis det oppstår behov for å gi beskjeder til kandidatene underveis i eksamen (f.eks. ved feil i oppgavesettet), vil dette bli gjort via varslinger i Inspira. Et varsel vil dukke opp som en dialogboks på skjermen. Du kan finne igjen varselet ved å klikke på bjella øverst til høyre.

**Trekk fra/avbrutt eksamen:** Blir du syk under eksamen, eller av andre grunner ønsker å levere blankt/avbryte eksamen, gå til "hamburgermenyen" i øvre høyre hjørne og velg «Lever blankt». Dette kan ikke angres selv om prøven fremdeles er åpen.

**Tilgang til besvarelse:** Etter eksamen finner du besvarelsen din i arkivet i Inspira. Merk at det kan ta én virkedag før eventuelle håndtegninger vil være tilgjengelige i arkivet.

## 1 Oppgave 1: Datamodeller (35 %)

Bruktmarked AS skal åpne en ny tjeneste for nettbasert kjøp og salg av brukte varer, og trenger en datamodell for en database som skal kunne lagre nødvendige data for å understøtte tjenesten. Ta utgangspunkt i følgende beskrivelse av behovene og lag en EER-modell for denne situasjonen. Du kan bruke alle modell-elementer i pensum, men alle trenger ikke å være med i modellen.

Tjenesten baserer seg på at brukere som skal selge eller kjøpe en vare må registrere seg i systemet. Ved registrering tildeles bruker en unik brukerID. Brukeren må registrere navn, adresse, telefonnummer, postnummer og poststed. I tillegg legger systemet inn årstall for registreringen.

Når en bruker ønsker å selge en vare eller registrere et behov for å kjøpe en vare, oppretter brukeren en annonse. Det registreres når annonsen opprettes, annonsen får et entydig annonse-løpenummer, og brukeren må fylle inn en del opplysninger: Overskrift, beskrivelse, pris, tilstand (brukt eller nytt) og om det gjelder kjøp eller salg. Annonser kan være «påbegynt», «under godkjenning», «aktiv», «deaktivert», «solgt» eller «kjøpt».

En bruker kan registrere interesse for en annonse, slik at det blir lett å finne igjen interessante annonser. Vi registrerer dato for innlegging av interesse. Dersom en vare blir solgt gjennom systemet, registreres det hvilken bruker som gjør kjøpet, og når det skjer. Både kjøper og selger har anledning til å gi en vurdering av prosessen, ved å registrere en til seks stjerner og eventuelt legge inn en kort kommentar i tillegg. Dersom noen ønsker å tilby en ønsket vare, blir selve transaksjonen håndtert av de involverte, uten spesiell støtte fra systemet.

Brukere kan kommunisere med utgangspunkt i en annonse. Dette gjøres ved å sende meldinger som blir registrert i systemet. En melding har et sekvensnummer som er unikt for den aktuelle annonsen, dato og tid for sending av meldingen, og en meldingstekst. Kommunikasjon er kun mellom eier av annonsen og interesserte kjøpere/selgere, med en privat meldingstråd for hver kombinasjon av annonseeier og annen bruker.

En bruker kan laste opp bilder som kan brukes for å illustrere egne annonser. Bruker velger visningsrekkefølge (fra en og oppover) for bildene innen annonsen. Samme bilde kan brukes for flere annonser, gjerne med ulik bildetekst og ulikt visningsnummer. Bruker kan administrere egne bilder. Bildene registreres med et unikt bildenummer som tildeles av systemet, en obligatorisk tittel og en beskrivende tekst. Det skal være mulig å finne annonser gjennom å «browse» gjennom annonser eller gjennom søk i annonse-titler og -beskrivelser. For å gjøre det enklere å finne interessante annonser, har systemet et antall hovedkategorier (for eksempel bil) med underkategorier (for eksempel Toyota). Når en bruker legger inn en annonse, må vedkommende velge den mest relevante hovedkategorien og en underkategori av denne hvis det er relevant. Hovedkategorier har en unik ID og et beskrivende navn. Det samme gjelder underkategorier, som må være knyttet til en, men ikke flere hovedkategorier.

Systemet har også en oversikt over fylker, kommuner og områder i kommuner. Fylker har fylkesnummer som er unikt, og fylkesnavn. Kommuner har kommunnummer som er unikt, og kommunenavn. Områder har områdenavn og et områdenummer som er unikt innenfor den kommunen området ligger i. En bruker må ved registrering legge inn det området der vedkommende holder til. På den måten kan brukere velge annonser fra andre brukere i sitt nærområde, for å unngå å betale frakt for sending av varer.

For å kunne analysere trender, selge «brukerprofiler» og tilby brukerne skreddersydde forslag på relevante annonser, skal systemet registrere hver gang en bruker går inn på en annonse og hvor lang tid det tar før brukeren klikker videre.

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

Selve datamodellen må tegnes på papir. Forutsetninger og andre kommentarer kan skrives i boksen under.

**Skriv ditt svar her**

Format

**B**


*I*


U


$x_2$


$x^2$


$I_x$
































Words: 0

Maks poeng: 35

## 2 Oppgave 2: B+-tre (15 %)

Tegn opp B+-treet som finnes etter innsetting av følgende sekvens av nøkler:

2, 3, 5, 1, 7, 6, 4, 13, 17, 9

Hver blokk har plass til 3 nøkler og hver blokk på nivå  $> 0$  har plass til 4 pekere og 3 nøkler.

Vis tilstanden til B+-treet hver gang du skal til å splitte en blokk, og vis treet til slutt

Denne oppgaven skal leveres på ark.

**Skriv ditt svar her**

Format | **B** | *I* | U |  $\times_2$  |  $\times^2$  |  $\mathcal{I}_x$  | | | | | | |  $\Omega$  | | |  $\Sigma$  |

Words: 0

Maks poeng: 15

### 3 Oppgave 3: Aksessmetoder (3 %)

Anta vi har en tabell **Birdwatcher(bwid, name, address, age, ...)** med 8000 rader. Attributtet bwid er primærnøkkel for tabellen. Tabellen er lagret i et clustered B+-tre med bwid som søkenøkkel, der løvnodene (level=0) i B+-treet har 550 blokker. Det er tre nivåer med blokker i B+-treet. Hvor mange blokkaksesser får vi med

SQL-setningen “**SELECT name, age FROM Birdwatcher where bwid=2001;**”?

**Velg ett alternativ:**

- ☐ 550
- ☐ 1
- ☐ 3
- ☐ 8000

---

Maks poeng: 3

### 4 Oppgave 4: Aksessmetoder (3%)

Anta vi har følgende nøkler som skal settes inn i en statisk hash-struktur:

16, 17 , 10, 11, 12, 13, 14, 15 og 18.

Hash-strukturen er tom fra før og har 4 blokker, der hver blokk kan inneholde to nøkler. Vi bruker separat lenket overløp hvis en blokk blir full. Bruk hashfunksjonen  $h(k) = k \bmod 4$  til å sette inn nøklene i rekkefølgen gitt over.

Hvor mange blokkaksesser får vi gjennomsnittlig her på aksess med en tilfeldig nøkkel over? Rund av svaret til to desimaler.

**Velg ett alternativ:**

- ☐ 1.22
- ☐ 1.33
- ☐ 1.11
- ☐ 1.00

---

Maks poeng: 3

## 5 Oppgave 5: Aksessmetoder (3 %)

Anta vi har en tabell **Birdwatcher(bwid, ...)** med 8000 rader. Attributtet bwid er primærnøkkel for tabellen. Tabellen er lagret i en clustered, statisk hash-struktur med bwid som søkenøkkel, der alle fugletitterne får plass i 450 blokker. Gjennomsnittlig aksesseres 1.25 blokker per søk på bwid. Hvor mange blokkaksesser får vi med SQL-setningen **"SELECT \* FROM Birdwatcher;"**?

Velg ett alternativ:

- ☐ 562
- ☐ 10000
- ☐ 8000
- ☐ 450

---

Maks poeng: 3

## 6 Oppgave 6: SNAPSHOT ISOLATION (6 %)

Hvilke påstander er riktige?

Velg ett eller flere alternativer

- ☐ SNAPSHOT ISOLATION tillater at to skrivetransaksjoner skriver samme dataelement samtidig
- ☐ SNAPSHOT ISOLATION fører til at lesetransaksjoner ofte må abortere
- ☐ SNAPSHOT ISOLATION tillater mer samtidighet mellom transaksjoner enn SERIALIZABLE
- ☐ SNAPSHOT ISOLATION tillater at to transaksjoner leser samme dataelement
- ☐ Ingen av de andre alternativene er riktige
- ☐ SNAPSHOT ISOLATION er den metoden som gir mest isolasjon mellom transaksjoner
- ☐ SNAPSHOT ISOLATION lar transaksjoner lese data som er låst av skrivetransaksjoner

---

Maks poeng: 6

## 7 Oppgave 7: Gjenopprettbarhet (3 %)

Hva er recovery-egenskapen til følgende historie:  $w1(A)$ ;  $r2(A)$ ;  $w1(B)$ ;  $w2(B)$ ;  $c1$ ;  $c2$ ;

**Velg ett alternativ:**

- ☐ Strict
- ☐ Gjenopprettbar (recoverable)
- ☐ ACA
- ☐ Ikke gjenopprettbar (non-recoverable)

---

Maks poeng: 3

## 8 Oppgave 8: Gjenopprettbarhet (3 %)

Hva er recovery-egenskapen til følgende historie:  $w1(A)$ ;  $w1(B)$ ;  $w2(A)$ ;  $c1$ ;  $r2(B)$ ;  $r3(B)$ ;  $c2$ ;  $c3$ ;

**Velg ett alternativ:**

- ☐ ACA
- ☐ Gjenopprettbar (recoverable)
- ☐ Ikke gjenopprettbar (non-recoverable)
- ☐ Strict

---

Maks poeng: 3

## 9 Oppgave 9: Gjenopprettbarhet (3 %)

Hva er recovery-egenskapen til følgende historie:  $r1(A); w2(A); w1(A); r3(A); c1; c2; c3;$

**Velg ett alternativ:**

- ☐ Strict
- ☐ Ikke gjenopprettbar (non-recoverable)
- ☐ ACA
- ☐ Gjenopprettbar (recoverable)

---

Maks poeng: 3

## 10 Oppgave 10: Konfliktserialiserbarhet (6 %)

Hvilke historier (schedules) er konfliktserialiserbare?

**Velg ett eller flere alternativer**

- ☐  $w3(A); r1(A); w1(B); r2(B); w2(C); r3(C);$
- ☐  $r1(A); w1(B); r2(B); r3(A); r2(A); w2(B);$
- ☐  $r1(A); r2(A); w1(B); w2(B); r1(B);$
- ☐  $r1(A); r2(A); r1(B); r3(C); w1(A); w2(C);$
- ☐ Ingen av de andre alternativene er konfliktserialiserbare
- ☐  $r2(C); r1(C); w2(C); r3(B); r2(B); w2(B);$
- ☐  $r3(C); r1(C); w3(C); r2(B); r3(B); w1(C);$

---

Maks poeng: 6



## 11 Oppgave 11: ARIES (10 %)

Anta at vi bruker ARIES-recovery og la A, B, og C være dataelementer. Etter en krasj ble følgende logg funnet med en loggpost i hver rad av tabellen:

LSN	Last_lsn	Transaction	OpType	Page_id	Other_info
101	0	T1	Update	A	...
102	0	T2	Update	B	...
103	101	T1	Commit		...
104			Begin_ckpt		
105			End_ckpt		
106	0	T3	Update	A	...
107	102	T2	Update	C	...
108	107	T2	Commit		...

Anta T1, T2 og T3 er alle transaksjonene som finnes og at transaksjonstabellen som finnes i loggpost 105 ser slik ut:

Transaction	Last_lsn	Status
T1	103	Commit
T2	102	In progress

Dirty Page Table som finnes i loggpost 105 ser slik ut:

Page_id	Rec_lsn
A	101
B	102

Hvordan ser transaksjonstabellen og Dirty Page Table ut etter analysefasen av recovery?

**Skriv ditt svar her**

Format ▼ | **B** *I* U  $\times_2$   $\times^2$  |  $\int_x$  | | | |  $\Omega$  |

$\Sigma$  |

Words: 0

Maks poeng: 10

## 12 Oppgave 12: ARIES (10 %)

Anta at vi bruker ARIES-recovery og la A og C være dataelementer. Etter en krasj ble følgende logg funnet med en loggpost i hver rad av tabellen:

LSN	Last_lsn	Transaction	OpType	Page_id	Other_info
103	101	T1	Commit		...
104			Begin_ckpt		
105			End_ckpt		
106	0	T3	Update	A	...
107	0	T2	Update	C	...
108	107	T2	Commit		...

Loggposten 105 inneholder en tom DPT.

Under redofasen trenger datablokk A redo av loggpost 106, mens datablokk C trenger ikke redo. Hvilke verdier for PageLSNene kan datablokk A og C ha i det redofasen starter?

**Skriv ditt svar her**

Format
|
**B**
*I*
U
 $x_2$ 
 $x^2$ 
 $I_x$ 
|


|


|


|

Σ
|

Words: 0

Maks poeng: 10

### 13 Oppgave 13: Kommentarer (0 %)

Denne "oppgaven" er en mulighet for å informere om *omstendigheter* som du tenker er **helt nødvendige** å kommunisere til sensor, for at din besvarelse skal bli riktig vurdert. Dette kan for eksempel gjelde antakelser som det var tvingende nødvendig å gjøre.

Du skal *ikke* bruke dette feltet til å gi *generelle kommentarer* til eksamen, det kan gjøres i Piazza eller i e-post til faglærer.

**Skriv ditt svar her**

Format

**B**


*I*


U


$x_2$


$x^2$


$I_x$



























$\Sigma$



Words: 0

Maks poeng: 0