

**Institutt for datateknologi og informatikk**

**Eksamensoppgave i TDT4150 Avanserte databasesystemer**

**Eksamensdato: 2022–05–18**

**Eksamenstid (fra-til): 15:00–19:00**

**Hjelpemiddelkode/Tillatte hjelpemidler: A / Alle hjelpemidler tillatt**

**Faglig kontakt under eksamen: Norvald H. Ryeng**

**Tlf.: 971 74 980**

**Teknisk hjelp under eksamen: NTNU Orakel**

**Tlf: 73 59 16 00**

Får du tekniske problemer underveis i eksamen, må du ta kontakt for teknisk hjelp snarest mulig, og senest innen eksamenstida løper ut/prøven stenger. Kommer du ikke gjennom umiddelbart, hold linja til du får svar.

## **ANNEN INFORMASJON**

**Ikke ha Inspira åpen i flere faner, eller vær pålogget på flere enheter, samtidig,** da dette kan medføre feil med lagring/levering av besvarelsen din.

**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 kan kontaktes dersom du mener det er feil eller mangler i oppgavesettet.

**Juks/plagiat:** Eksamen skal være et individuelt, selvstendig arbeid. Det er tillatt å bruke hjelpemidler, men vær obs på at du må følge eventuelle anvisningen om kildehenvisninger under. Under eksamen er det ikke tillatt å kommunisere med andre personer om oppgaven eller å distribuere utkast til svar. Slik kommunikasjon er å anse som juks.

Alle besvarelser blir kontrollert for plagiat. [Du kan lese mer om juks og plagiering på eksamen her.](#)

**Kildehenvisninger:** Det forventes ikke at du trenger sitater i besvarelsen. Hvis du siterer, husk å oppgi referanser. Du kan referere til nummererte dokumenter på pensumlisten som «artikkel X».

**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 i Inspira. Du kan finne

igjen varselet ved å klikke på bjella øverst i høyre hjørne på skjermen. Det vil i tillegg bli sendt SMS til alle kandidater for å sikre at ingen går glipp av viktig informasjon. Ha mobiltelefonen din tilgjengelig.

**Vekting av oppgavene:** Alle oppgavene teller like mye.

## BESVARE OG LEVERE

**Besvare i Inspira:** Hvis oppgavesettet inneholder oppgaver som *ikke* er av typen filoplasting, skal de besvares direkte i Inspira. I Inspira lagres svarene dine automatisk hvert 15. sekund.

NB! Klipp og lim fra andre programmer frarådes, da dette kan medføre at formatering og elementer (bilder, tabeller etc.) vil kunne gå tapt.

**Automatisk innlevering:** Besvarelsen din leveres automatisk når eksamenstida er ute og prøven stenger, forutsatt at minst én oppgave er besvart. Dette skjer selv om du ikke har klikket «Lever og gå tilbake til Dashboard» på siste side i oppgavesettet. Du kan gjenåpne og redigere besvarelsen din så lenge prøven er åpen. Dersom ingen oppgaver er besvart ved prøveslutt, blir ikke besvarelsen din levert. Dette vil anses som “ikke møtt” til eksamen.

**Trekk/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:** Du finner besvarelsen din i Arkiv etter at sluttida for eksamen er passert.

# 1 Problem 1

Hva er hovedoppgavene til en spørringsoptimalisator (query optimizer)?

**Skriv ditt svar her**

Format

**B**


*I*


U


$x_2$


$x^2$


$I_x$
































Words: 0

Maks poeng: 10

## 2 Problem 2

Gi et eksempel på en SQL-spørring som skrives om til å bruke semijoin før optimalisering (oppgi både SQL-spørringen og den algebraiske formen etter omskriving). Hvordan påvirker semijoin-omskrivning arbeidet som gjøres av spørringsoptimalisatoren (the query optimizer)?

**Skriv ditt svar her**

Format

**B**

*I*

U

$x_2$

$x^2$

$I_x$

$\Omega$

$\Sigma$

Words: 0

Maks poeng: 10

### 3 Problem 3

Anta et DBMS med egenskaper tilsvarende System R og et skjema og datasett med indekser på alle kolonner. Dette systemet utfører den følgende spørringen:

```
SELECT *  
FROM Hotels h  
WHERE h.city = 'Nassau' AND NOT EXISTS (  
    SELECT *  
    FROM Hotels h1  
    WHERE h1.city = 'Nassau' AND  
        h1.distance <= h.distance AND  
        h1.price <= h.price AND  
        (h1.distance < h.distance OR  
        h1.price < h.price));
```

Oppgi alle sargbare predikater (sargable predicates) i spørringen og forklar for hvert predikat hvorfor det er sargbart (sargable).

**Skriv ditt svar her**

Format

**B**

*I*

U

$x_2$

$x^2$

$I_x$

Words: 0

Maks poeng: 10

## 4 Problem 4

I et DBMS som ligner på System R har vi disse selektivitetsfaktorene (selectivity factors):

column1 = value:  $F = 1 / 10$

(pred1) AND (pred2):  $F = F(\text{pred1}) * F(\text{pred2})$

Hvilke antagelser ligger bak disse formlene? Gi et eksempel hvor disse antagelsene ikke holder og forklar hvorfor.

**Skriv ditt svar her**

Format ▾ | **B** *I* U  $x_2$   $x^2$  |  $I_x$  | | | |  $\Omega$  |  $\Sigma$  |

Words: 0

Maks poeng: 10

## 5 Problem 5

Forklar hvordan System R velger mellom eksplisitt sortering og bruk av en indeks til å lese data i sortert rekkefølge.

**Skriv ditt svar her**

Format

**B**


*I*


U


$x_2$


$x^2$


$I_x$
































Words: 0












Maks poeng: 10


## 6 Problem 6

Forklar hvordan score guided hash rank join avviker fra den basale hash rank join-algoritmen og hvordan det påvirker ytelsen.

**Skriv ditt svar her**

Format ▾

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



Words: 0

---

Maks poeng: 10



## 7 Problem 7

Gitt et LSM-tre  $C_0$ ,  $C_1$ ,  $C_2$ , forklar hvorfor  $C_2 > C_1 > C_0$  og hvorfor det er bedre å ha flere komponenter enn å la eksisterende komponenter alltid vokse større og større når data legges til.

**Skriv ditt svar her**

Format

**B**


*I*


U


$x_2$


$x^2$


$I_x$
































Words: 0












Maks poeng: 10


## 8 Problem 8

Gi et eksempel hvor et lazy update anywhere-system får inkonsistens (inconsistency). Forklar hvordan/hvorfor inkonsistensen oppstår.

**Skriv ditt svar her**

Format ▾

**B** *I* U  $x_2$   $x^2$   $\frac{1}{x}$            



Words: 0












Maks poeng: 10


## 9 Problem 9

Hva gjør operatoren DistributedUnion i Spanner, og hvorfor forsøker optimalisatoren (the optimizer) å dytte den oppover i spørringsplanen (the query plan)?

**Skriv ditt svar her**

Format ▾

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



Words: 0

Maks poeng: 10

## 10 Problem 10

Anta to R-trær over det samme datasettet, ett med mye overlapp (node overlap) og ett med lite overlapp. Hvordan påvirker denne forskjellen eksekveringen av punktspøringer (point queries)?

**Skriv ditt svar her**

Format

**B**


*I*


U


$x_2$


$x^2$


$I_x$



























$\Sigma$



Words: 0

Maks poeng: 10