Sensorveiledning TDT4145 eksamen 14. mai 2022

Versjon 14. mai 2022

Læringsutbyttebeskrivelser for TDT4145

Kunnskaper:

- 1. Databasesystemer: generelle egenskaper og systemstruktur.
- 2. Datamodellering med vekt på entity-relationship-modeller.
- 3. Relasjonsdatabasemodellen for databasesystemer, databaseskjema og dataintegritet.
- 4. Spørrespråk: Relasjonsalgebra og SQL.
- 5. Designteori for relasjonsdatabaser.
- 6. Systemdesign og programmering mot databasesystemer.
- 7. Datalagring, filorganisering og indeksstrukturer.
- 8. Utføring av databasespørringer.
- 9. Transaksjoner, samtidighet og robusthet mot feil.

Ferdigheter:

- 1. Datamodellering med entity-relationship-modellen.
- 2. Realisering av relasjonsdatabaser.
- 3. Databaseorientert programmering: SQL, relasjonsalgebra og database-programmering i Java.
- 4. Vurdering og forbedring av relasjonsdatabaseskjema med utgangspunkt i normaliseringsteori.
- 5. Analyse og optimalisering av ytelsen til databasesystemer.

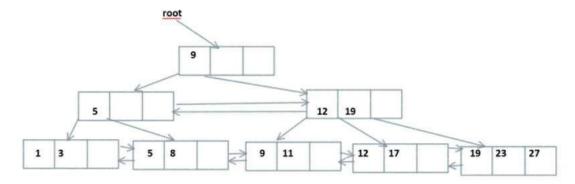
Generell kompetanse:

- 1. Kjennskap til anvendelser av databasesystemer og forståelse for nytte og begrensninger ved slike systemer.
- 2. Modellering av og analytisk tilnærming til datatekniske problemer

Poenggrenser brukt:

Dette er foreløpig poenggrensene brukt ved sensur.

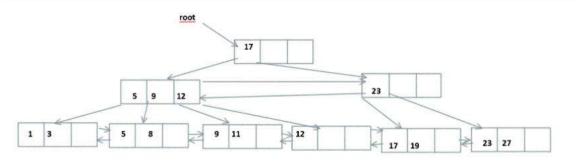
1. B-tre (10 %)



Følgende sekvenser av innsettinger er tilgjengelige. Hvilken av disse gir B+-treet illustrert i oppgaven? Det er plass til tre poster i hver blokk, og hver blokksplitt flytter over en post til den nye blokka til høyre.



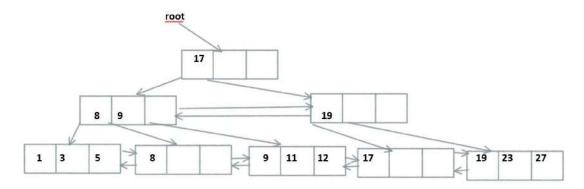
2. B+-tre (10 %)



Følgende sekvenser av innsettinger er tilgjengelige. Hvilken av disse gir B+-treet illustrert i oppgaven? Det er plass til tre poster i hver blokk, og hver blokksplitt flytter over en post til den nye blokka til høyre.



3. B+-tre (10 %)



Følgende sekvenser av innsettinger er tilgjengelige. Hvilken av disse gir B+-treet illustrert i oppgaven? Det er plass til tre poster i hver blokk, og hver <u>blokksplitt</u> flytter over en post til den nye blokka til høyre.

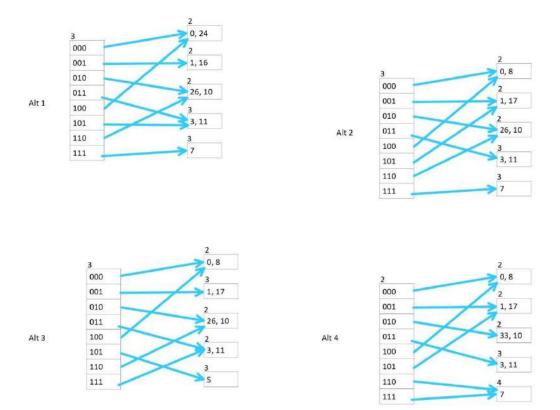
Velg ett alternativ:

X17, 19, 8, 9, 1, 3, 12, 5, 11, 23, 27

O 1, 3, 5, 8, 9, 11, 12, 17, 19, 23, 27

O 5, 12, 17, 1, 3, 19, 23, 27, 8, 9, 11

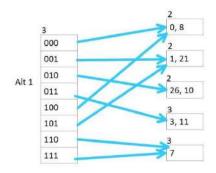
4. Extendible hashing (6%)

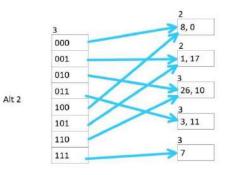


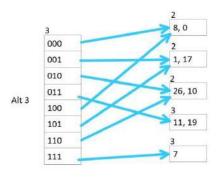
Hvilke av figurene er lovlige strukturer i extendible hashing? Det kan være flere riktige svar. Hvis figuren er liten, kan du trykke CTRL-+ i browseren.

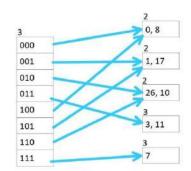


5. Extendible hashing (6 %)







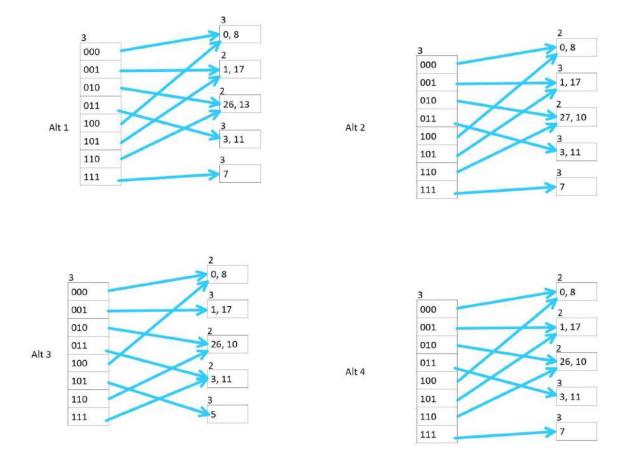


Hvilke av figurene er lovlige strukturer i extendible hashing? Det kan være flere riktige svar. Hvis figuren er liten, kan du trykke CTRL-+ i browseren.

Alt 4



6. Extendible hashing (6 %)



Hvilke av figurene er lovlige strukturer i extendible hashing? Det kan være flere riktige svar. Hvis figuren er liten, kan du trykke CTRL-+ i browseren.



7. Misc (6 %)

Hva er riktig for ARIES?

Velg ett eller flere alternativer on logging write-ahead logging force, steal no-force, steal no-force, no-steal force, no-steal

I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

8. Misc (6 %)

Hvilke alternativ er sanne?



I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

9. Misc (6 %)

Hvilke alternativ er sanne?

Velg ett eller flere alternativer □ DPT forteller hvor UNDO kan starte DPT forteller hvor REDO kan starte ransaksjonstabellen forteller hvem som er vinnere og tapere av transaksjonene etter analysen □ REDO starter ved eldste "levende" loggpost, dvs. eldste loggpost blant tapere. UNDO slutter ved eldste "levende" loggpost, dvs. eldste loggpost blant tapere. □ PageLSN blir alltid oppdatert for alle loggposter ved REDO

I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

10. 2PL-utføring (5 %)

Databasesystemet får inn følgende sekvens av operasjoner. Vi innfører tofaselåsing (rigorous). I hvilken rekkefølge committer transaksjonene?

Viss flere transaksjoner blir vekt opp etter en låsventing, blir de vekt opp i den rekkefølgen de la seg til å vente, dvs. en kø.

r1(A); w2(B); r1(B); r3(A); w3(A); c1; c2; c3;

```
Velg ett alternativ:

O T1; T2; T3;
O T1; T3; T2;
O Vranglås mellom T1 og T3.

T2; T1; T3;
O T2; T3; T1;
O T3; T1; T2;
O T3; T2; T1;
```

11. 2PL-utføring (5 %)

Databasesystemet får inn følgende sekvens av operasjoner. Vi innfører tofaselåsing (rigorous). I hvilken rekkefølge committer transaksjonene?

Viss flere transaksjoner blir vekt opp etter en låsventing, blir de vekt opp i den rekkefølgen de la seg til å vente, dvs. en kø.

r3(A); w2(B); r3(B); w1(A); w3(B); c1; c2; c3;

```
Velg ett alternativ:

O T1; T2; T3;

O T1; T3; T2;

O Vranglås mellom T1 og T3

O T2; T1; T3;

▼T2; T3; T1;

O T3; T1; T2;

O T3; T2; T1;
```

12. 2PL-utføring (5 %)

Databasesystemet får inn følgende sekvens av operasjoner. Vi innfører tofaselåsing (rigorous). I hvilken rekkefølge committer transaksjonene?

Viss flere transaksjoner blir vekt opp etter en låsventing, blir de vekt opp i den rekkefølgen de la seg til å vente, dvs. en kø.

r1(B); r3(B); w1(A); w2(A); w3(A); c1; c2; c3;

13. Konfliktserialiserbarhet (6 %)

Hvilke historier er konfliktserialiserbare? Det kan være flere.

```
Velg ett eller flere alternativer

□ r1(X); r2(X); w1(X); w2(Y); r3(Z); w2(X);

xr2(Y); r1(Y); r3(Y); w1(X); w2(Z); w2(X);

xr3(X); w3(X); r1(Y); r2(Y); r3(Y); w2(Y);

□ w1(A); w2(B); r1(B); r3(A); w2(A); r1(A);

xw2(B); w1(A); r3(B); w3(B); r3(A); w2(C);

□ r1(C); w2(B); r1(B); r3(B); w1(B); w3(B);
```

I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

14. Konfliktserialiserbarhet (6 %)

Hvilke historier er konfliktserialiserbare? Det kan være flere.

```
Velg ett eller flere alternativer

□ r1(X); w1(X); r2(X); r3(Y); w2(Y); w3(Y);

w2(X); r1(X); w1(X); r3(Y); w3(X); r2(Y);

xr1(Z); r2(X); w3(Z); r1(X); w2(X); w3(X);

□ r1(A); r2(A); w3(B); r3(A); w3(A); w2(A);

□ r2(B); w3(B); r1(A); w2(A); r3(Z); w1(Z);

xr1(A); w2(C); r3(C); w1(A); r2(A); w2(A);
```

15. Konfliktserialiserbarhet (6 %)

Hvilke historier er konfliktserialiserbare? Det kan være flere.

Velg ett eller flere alternativer □ w1(X); r2(X); r1(Y); w3(X); w2(X); w1(Y); r1(Y); w2(Y); r1(X); w3(X); r3(Y); r1(Z); w2(Y); r3(Y); r1(Z); w2(Z); w3(Z); w2(X); □ r1(A); w2(B); r2(A); w3(A); r3(Z) w1(A); □ r3(Z); w2(Y); r1(Z); w1(Y); w2(Z); w2(X); w3(A); r2(A); r1(B); w1(B); w3(C); w2(C);

I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

16. Gjenopprettbarhet (6 %)

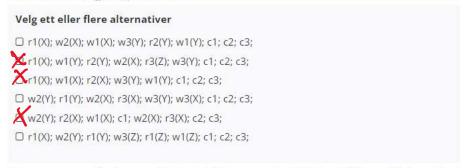
Hvilke historier er gjenopprettbare?

Velg ett eller flere alternativer Xr1(X); r2(X); w1(Y); r2(Y); r3(Y); w3(Y); c1; c2; c3; □ w2(X); r1(X); r3(X); w2(Y); r3(Y); w1(X); c1; c2; c3; Xw1(X); w1(Y); r2(Y); w3(Y); w1(Z); c1; c2; c3; □ r2(X); w1(X); w2(Y); r1(Y); r3(Y); c1; r3(X); c2; c3; Xr1(Z); r2(Z); w1(Y); r3(Y); w3(Y): c1; w3(Z); c2; c3; □ w3(Z); r1(X); r1(Z); w2(Y); r3(Y); w3(Y); c1; c2; c3;

I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

17. Gjenopprettbarhet (6 %)

Hvilke historier er gjenopprettbare?



18. Misc (6 %)

For en tabell Student (<u>studnr</u>, navn, epost) har vi ei heapfil med 2000 blokker som lagrer tabellen, og et unclustered B+-tre for nøkkelen studnr. B+-treet har 500 blokker på løvnivå og høyde 3. Verdiene av epost er ikke unike.

Hvilke alternativ er sanne?

Velg ett eller flere alternativer □ "SELECT * FROM Student" gir 2003 aksesser til blokker "SELECT navn FROM Student WHERE studnr=10001" gir 4 aksesser til blokker. □ "SELECT studnr FROM Student ORDER BY studnr" gir 502 aksesser til blokker. □ "SELECT studnr FROM Student WHERE navn='Hans Hansen'" gir 4 aksesser til blokker. □ "SELECT epost, navn FROM Student WHERE epost='ole@stud.ntnu.no'" gir 2000 aksesser til blokker. □ "INSERT INTO Student VALUES (10002, 'Ole Olesen', 'ole@stud.ntnu.no) gir 503 aksesser til blokker og tre skrivinger.

I denne oppgaven får du poeng for hvert riktige svar og trekk for hvert feil svar. Minimum 0 poeng på hele oppgaven.

Her er det et poeng at epost ikke er unik, eller ville det siste svaret vært 1000 i gjennomsnitt. Det er en litt tvilsom antagelse, men står oppgitt i oppgaven. F.eks. hvis man registrerer 'ukjent' når man ikke har epost, så er den ikke unik.

19. Misc (6 %)

Vi har en tabell Ansatt(<u>aid</u>, navn, epost) som er lagret i et clustered B+-tre hvor primærnøkkel i tabellen og søkenøkkel i B+-treet er aid. B+-treet har 2000 blokker på løvnivå og har høyde 3.

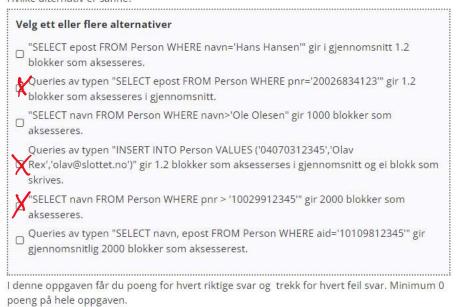
Hvilke alternativ er sanne?



20. Misc (6 %)

Vi har ei clustered hashfil hvor gjennomsnittlige antall aksesser er 1,2 blokker når du akssseser på søkenøkkel. Tabellen Person(pnr. navn, epost) er lagra i den clustered hashfila med pnr som søkenøkkel. Det er 2000 blokker totalt i hashfila.

Hvilke alternativ er sanne?



Her var det en feil i det siste queriet. De skulle vært pnr, ikke aid. Det ble sendt ut melding til alle studenter om dette, selv om det gjelder kun 1/3 av studentene.

21. Recovery (ARIES) (4 %)

LSN	PrevLSN	TransID	Type	PageID
101			End_ckpt	
102	Null	T1	Update	A
103	Null	T2	Update	В
104	102	T1	Commit	
105	103	T2	Update B	
106	105	T2	Update	С
107	Null	T3	Update C	
108	106	T2	Update	Α

Anta loggen over etter et krasj med ARIES. Etter analysen har DPT følgende innhold: (A, 98), (B, 103), (C, 106) og datablokkene har følgende PageLSN (A, 98), (B, 97), (C, 107). Hvilke loggposter blir det gjort REDO for?

22. Recovery (ARIES) (4 %)

LSN	PrevLSN	TransID	Туре	PageID
101			End_ckpt	
102	Null	T1	Update	A
103	Null	T2	Update	В
104	102	T1	Commit	
105	103	T2	Update	В
106	105	T2	Update	С
107	Null	T3	Update	C
108	106	T2	Update	А

Anta loggen over etter et krasj med ARIES. Etter analysen har DPT følgende innhold: (A, 108), (B, 98), (C, 106) og datablokkene har følgende PageLSN (A, 102), (B, 103), (C, 97). Hvilke loggposter blir det gjort REDO for?



23. Recovery (ARIES) (4 %)

LSN	PrevLSN	TransID	Туре	PageID
101			End_ckpt	
102	Null	T1	Update	А
103	Null	T2	Update	В
104	102	T1	Commit	
105	103	T2	Update	В
106	105	T2	Update	С
107	Null	T3	Update	С
108	106	T2	Update	A

Anta loggen over etter et krasj med ARIES. Etter analysen har DPT følgende innhold: (A, 102), (B, 103), (C, 96) og datablokkene har følgende PageLSN (A, 98), (B, 105), (C, 96). Hvilke loggposter blir det gjort REDO for?



24. Join (5 %)

To tabeller Klasse og Student skal joines ved en nested loop join. Bufferet har 6 plasser til blokker, tabellen Klasse har 12 blokker og tabellen Student har 1000 blokker. Hvor mange lesinger av blokker skjer ved joinen?

Velg ett alternativ:			
O 1012			
O 2024			
O 3036			
O 2012			
3012			

25. Join (5 %)

To tabeller Klasse og Student skal joines ved en nested loop join. Bufferet har 8 plasser til blokker, tabellen Klasse har 12 blokker og tabellen Student har 1400 blokker. Hvor mange lesinger av blokker skjer ved joinen?

Velg ett alternativ:		
O 1512		
O 2824		
2812		
O 4524		
O 1520		

26. ER-modellering (34 %)

Lag en ER-modell (du kan bruke alle virkemidler som er med i pensum, også spesialisering og kategorier) ut fra følgende situasjonsbeskrivelse («miniverden»):

Eiendomsselskapet Bo-i-ro forvalter en rekke bygninger som hver har en unik bygnings-id, en gateadresse, et postnummer og et poststed. For hver bygning registrerer man antall etasjer og samlet areal. Bygningene er delt inn i rom. Et rom har et romnummer som er unikt innenfor bygningen, et areal og ligger i en av bygningens etasjer. For å overvåke tilstanden i bygningene, kan man måle ulike miljøparametre, for eksempel lufttemperatur, relativ luftfuktighet og CO2. Bo-i-ro holder oversikt over de miljøparameterne som kan være aktuelle, der hver parametertype har en unik parameter-id, et navn, en måleenhet og en kort beskrivelse. For alle aktuelle miljøparametre, er det slik at en måling består av en enkelt tallverdi.

For å gjøre målinger av miljøparametre bruker Bo-i-ro sensorer. Hver sensor er av en bestemt sensormodell, for eksempel «SmartLab CO2 reader». En sensormodell har en unik sensormodell-id og et produktnavn. Hver sensormodell kan måle en eller flere miljøparametre, og dette holder man rede på. Man kan ha mange sensorer av samme sensormodell. Hver sensor har et sensornummer som er unikt innenfor Bo-i-ro. For hver sensor registrerer man år og måned for innkjøp. En sensor som er tatt i bruk, er plassert i et bestemt rom i en av selskapets bygninger. Når man monterer en sensor, registrerer man en kort beskrivelse av plasseringen i rommet, for eksempel «midt i taket». Det kan være mange sensorer (også av samme sensormodell) i et rom. Det er ikke et krav at alle rom må ha sensorer.

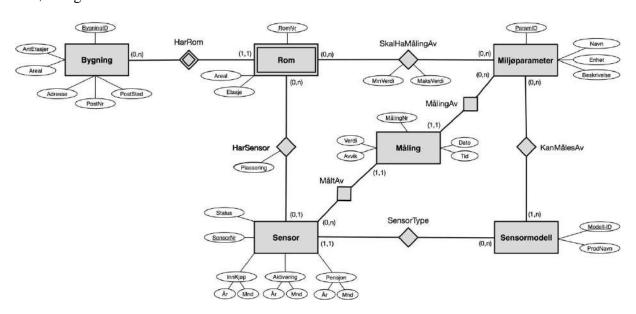
Det skal være mulig å registrere miljøparametre uten at man (ennå) har registrert en sensormodell som kan gjøre slike målinger. Det skal også være mulig å registrere en sensormodell uten at man har en sensor av denne modellen. For å gjøre modelleringen enklere, kan du anta at en sensor aldri blir flyttet når den først har blitt tatt i bruk. En sensor har status «ubrukt» frem til den tas i bruk. Da settes status til «aktiv» og man registrerer år og måned. Når en sensor tas ut av bruk, får den statusen «pensjonert» og man registrerer år og måned når dette skjer.

For et rom registreres de miljøparametrene som skal måles for rommet, og det registreres en minimumsverdi og en maksimumsverdi for målinger av hver type. Målinger utenfor disse (normal-)intervallene skal registreres som avvik. For alle sensorer som er i aktiv bruk, registreres de målingene som gjøres. Målinger har et unikt måling-nummer, en verdi for den aktuelle parametertypen, dato og tid. Dersom målingen er utenfor normalområdet for rommet, registreres det at målingen representerer et avvik. Hvis en sensor måler flere ulike miljøparametre, for eksempel både temperatur og luftfuktighet, skal dette registreres som to eller flere separate målinger.

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

Under er det vist utkast til løsning for de tre ulike modelleringsoppgavene. Det skal legges vekt på at de ulike modell-virkemidlene brukes på riktig måte. God (overordnet) «struktur» i datamodellen tillegges større vekt enn mer ubetydelige feil og mangler. Det finnes en del alternative modelleringsvalg og alternative forutsetninger som kan være like riktige som de som er vist i løsningsskissene.

Dersom det gjøres hensiktsmessige forutsetninger, skal disse legges til grunn ved vurderingen av løsningen.



27. ER-modellering (34 %)

Lag en ER-modell (du kan bruke alle virkemidler som er med i pensum, også spesialisering og kategorier) ut fra følgende situasjonsbeskrivelse («miniverden»):

Politiet ønsker å lage en database over fartsmålinger i trafikken. En fartsmåling bestemmer hastigheten til et kjøretøy. En fartsmåling har et unikt løpenummer. I tillegg til løpenummer og hastighet registreres dato og tid for målingen, og det er mulig å legge inn et notat som dokumenterer relevante omstendigheter. Fartsmålinger gjøres teknisk av såkalte «fotobokser»

som er fastmontert ved veien, eller ved hjelp av mobile lasermålere som betjenes av ansatte i politiet.

Automatisk fartskontroll ved hjelp av fotobokser er enten punktmåling eller strekningsmåling. Ved punktmåling måles farten idet kjøretøyet passerer fotoboksen. Dersom kjøretøyet holder høyere hastighet enn fartsgrensen på stedet, tas det bilde av kjøretøyet. Dette bildet lagres som en del av fartsmålingen for å identifisere kjøretøy og fører. Ved strekningsmåling måler man gjennomsnittsfarten mellom to fotobokser. Dette gjøres ved at man tar bilde av kjøretøyet og registrerer passeringstid når kjøretøyet passerer den første fotoboksen. Når kjøretøyet etter hvert passerer den andre fotoboksen, registreres passeringstid og det tas et nytt bilde av kjøretøyet. Ut fra passeringstidene og avstanden mellom de to fotoboksene, beregnes gjennomsnittshastigheten på strekningen. Dersom gjennomsnittsfarten er høyere enn fartsgrensen på strekningen, lagres passeringstidene, de to bildene og beregnet hastighet.

En fotoboks er registrert i systemet med et unikt fotoboksnummer og GPS-koordinater (breddegrad og lengdegrad) for boksens plassering. Dersom to fotobokser utgjør et par i forbindelse med strekningsmåling, skal dette registreres og man lagrer også avstanden mellom de to fotoboksene. Systemet skal holde oversikt over alle fartsgrenser (30, 40, 50, ...) som er i bruk på norske veier. En fartsgrense registreres med en unik fartsgrense-id og en høyeste tillatt hastighet i km/t. For hver fotoboks skal det registreres hvilken fartsgrense som gjelder på stedet.

Fartsmåling ved hjelp av mobil lasermåler gjøres ved at en ansatt i politiet bruker lasermåleren til å måle hastigheten på et kjøretøy som passerer. Dersom kjøretøyet holder høyere fart enn fartsgrensen, blir kjøretøyet stoppet og man sikrer seg registreringsnummer på kjøretøyet og førerens identitet. Når fartsmålingen gjøres ved mobil lasermåler, må man registrere GPS-koordinater (breddegrad og lengdegrad) for stedet der målingen gjennomføres.

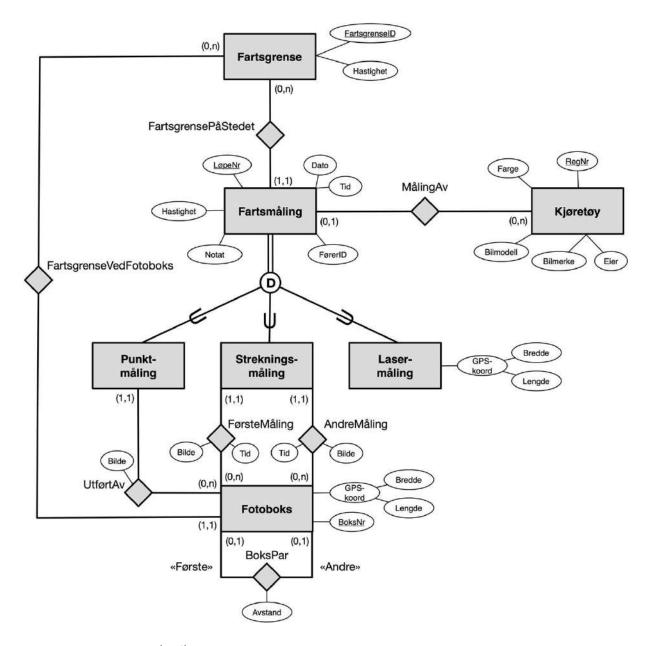
Systemet skal bare lagre fartsmålinger der kjøretøyet har holdt høyere hastighet enn fartsgrensen. For alle fartsmålinger skal det registreres hvilken fartsgrense som gjelder på stedet for fartsmålingen.

Bildene fra fotoboksene brukes til å identifisere kjøretøy og fører av kjøretøyet. Man har oversikt over alle registrerte kjøretøy som er lagret med unikt registreringsnummer, bilmerke, bilmodell, farge og fødselsnummer for eier. For hver fartsmåling blir det etter hvert registrert hvilket kjøretøy som er målt, bortsett fra i tilfeller der det er umulig å få klarhet i dette. Dersom man klarer å finne ut hvem som kjørte kjøretøyet da fartsmålingen ble foretatt, blir fødselsnummer for bilfører registrert for fartsmålingen.

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

Under er det vist utkast til løsning for de tre ulike modelleringsoppgavene. Det skal legges vekt på at de ulike modell-virkemidlene brukes på riktig måte. God (overordnet) «struktur» i datamodellen tillegges større vekt enn mer ubetydelige feil og mangler. Det finnes en del alternative modelleringsvalg og alternative forutsetninger som kan være like riktige som de som er vist i løsningsskissene.

Dersom det gjøres hensiktsmessige forutsetninger, skal disse legges til grunn ved vurderingen av løsningen.



28. 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.