Institut for Matematik og Datalogi
Syddansk Universitet, Odense

3. maj 2000
KSL

DM26 – EKSAMEN
Obligatorisk opgave, del 2

Da dette er en eksamensopgave, er der en række begrænsninger på, hvad der tilladt; lige som til en skriftlig eksamen. Sidst i denne opgave præciseres disse krav nærmere.

Opgaven er stillet onsdag, d. 3. maj 2000 kl. 16:00 og skal afleveres senest tirsdag, d. 23. maj 2000 kl. 12:00 på IMADA's sekretariat.

Denne opgave må løses i grupper på 2 personer.

Opgaven handler om ER-modellering, evaluering af relationelle operatorer og optimering af forespørgsler.

Opgave 1

Spørgsmål 1

I denne delopgave får man nogle oplysninger om datasammenhænge i et parlament. Man skal udfra disse oplysninger producere ét ER-diagram, der indfanger alle de givne oplysninger. Er der oplysninger, man ikke mener at kunne udtrykke, skal der gives en forklaring.

Efter konstruktionen af ER-diagrammet skal der laves en oversættelse til SQL via CREATE TABLE definitioner. Man skal forsøge at indfange de informationer om entydighed, totalitet, mm., der indgår i diagrammet. Der skal redegøres for eventuelle valg, der er foretaget under processen.

Herunder følger oplysningerne.

- Ministerier får en entydig ministerium-id. Desuden skal navnet på ministeriet samt årsbudgetet (et heltal) registreres.
- Parlamentsmedlemmer registreres ved deres CPR-nummer med oplysning om navn og partitilknytning. Desuden registreres for hvert medlem et mindre antal ministerier, hvis arbejdsområder det pågældende medlem ved specielt meget om.
- Udvalg har en entydig identifikation og et navn. Et udvalg sorterer altid under ét ministerium.
- Parlamentsmedlemmer er med i et antal udvalg. Datoen for indtræden i hvert enkelt udvalg registreres.
- Hvert ministerium har en minister, som altid er et af parlamentsmedlemmerne. Datoen for ministerens overtagelse af et ministerium registreres.
- Medarbejdere i ministerierne registreres ved deres CPR-nummer sammen med oplysninger om navn, hvilken overenskomst, de er ansat efter, samt anciennitet og udløbsdato for ansættelsen. Hvis ansættelsen er permanent, sættes udløbsdatoen til den dato, hvor medarbejderen fylder 67 år. Medarbejderne i ministerierne er ikke parlamentsmedlemmer.

- Alle medarbejderne er ansat ved ét ministerium.
- Nogle medarbejdere har lederbeføjelser, og der registreres en reference til, hvor deres underskriftprøve er arkiveret.
- Nogle medarbejdere er departementchefer. For disse gemmes en længere beskrivelse af deres præcise ansvar.
- Alle departementchefer refererer til deres egen minister. Desuden kan de have et antal andre ministre, de har pligt til at orientere om væsentlige beslutninger.
- Alt arbejde (ud over faste administrative opgaver, som vi ikke vil beskæftige os med her) organiseres som projekter, der har en identifikation, titel og deadline.
- Det faktiske arbejde (ud over ledelsen) i projekterne laves i arbejdsgrupper. Hvert projekt kan have mange sådanne arbejdsgrupper. Arbejdsgrupperne nummeres 1, 2, 3, osv. separat for hvert projekt. Dvs. at alle projekter, der overhovedet har arbejdsgrupper på et givet tidspunkt, har en arbejdsgruppe nummer 1. For de enkelte grupper gemmes navnet på arbejdsgruppen, samt deadline. Arbejdsgrupperne har deltagelse af et antal medarbejdere og parlamentsmedlemmer.

Spørgsmål 2

Skriv en “assertion”, der sikrer, at hver enkelt minister er minister for et ministerium, hvis arbejdsområde de ved specielt meget om.

Opgave 2

I denne delopgave bliver vi indenfor samme begrebsverden som i opgave 1, men de relationer, der bruges her, antyder intet om, hvordan design og implementation skal foretages i opgave 1. Denne opgave er helt uafhængig af den foregående.

Spørgsmål 1

Vi antager, at vi har en relation `Ansat` med attributter `CPR`, `name`, `overens`, `anc` og `dato`. `CPR` er nøgle for relationen. Attributterne fylder henholdsvis 11, 25, 54, 2 og 8 bytes. Relationen fylder 7500 sider, hvor en side er 1024 bytes.

Vi betragter følgende forespørgsel, der finder navnene på de ansatte, der er ansat i henhold til overenskomsten med magistrenes fagforening og som har 7 års anciennitet.

```
SELECT A.name
FROM Ansat A
WHERE A.overens = "Magistrenes" AND A.anc = 7
```

Vi regner med, at ca. 30% af tuplerne har værdien `Magistrenes` i feltet `overens`, ca. 5% har værdien `7` i feltet `anc`, og ca. 3% opfylder begge betingelser.

Der er 15 buffer-sider til rådighed for evalueringen.

Det antages, at der anvendes en directory filimplementation *uden* mulighed for fra en side at komme til den næste. Dvs. at man godt kan scanne hele filen, men hvis man ønsker tilgang til en delsekvens af siderne gennem et B^+-træ-index, så er det nødvendigt også at læse alle de tilsvarende index-blade (disse *er* hægtet sammen).

Angiv bedste plan, samt en detaljeret udregning af prisen for den, for hver af nedenstående seks uafhængige muligheder.

Der er et clustered B^+-træ på:

1. `anc`
2. `name`
3. ⟨`anc, name`⟩
4. ⟨`anc, overens`⟩
5. ⟨`anc, overens, name`⟩
6. ⟨`name, anc, overens`⟩

Spørgsmål 2

Vi har nu to relationer.

Den ene relation hedder `Parlamentsmedlem`, og den har attributterne `pCPR`, `pnavn`, `parti` og `budget`. De fylder henholdsvis 11, 30, 30 og 7 bytes.

Den anden relation hedder `Ansat`, og den har attributterne `aCPR`, `anavn`, `mpCPR` og `overens`. De fylder henholdsvis 11, 30, 11 og 20 bytes.

En ansat er nu tilknyttet ét bestemt medlem af parlamentet.

En side er igen 1024 bytes, og `Parlamentsmedlem` fylder 500 sider, mens `Ansat` fylder 1000 sider.

Vi ser nu på følgende forespørgsel, der finder navnene på de ansatte med overenskomst under Handel og Kontor, som er tilknyttet et medlem med et budget på under 40.000 kr.

```
SELECT A.aname
FROM Parlamentsmedlem P, Ansat A
WHERE     P.pCPR = A.mpCPR
      AND A.overens = "Handel og Kontor"
      AND P.budget < 40000
```

Ca. 10% af de ansatte er under Handel og Kontor, 20% af parlamentsmedlemmerne har budgetter under 40000, og 5% af de ansatte under Handel og Kontor er tilknyttet et parlamentsmedlem med et budget på under 40000.

Angiv en detaljeret udregning af prisen for følgende planer. Der er 15 buffer-sider til rådighed.

1. Relationerne joines sammen vha. "block nested loop", og resten (projektion og selektion) klares "on-the-fly".
2. Lige sådan, men der joines vha. "sort-merge".
3. Selektion skubbes mest muligt nedad, hvorefter relationerne joines sammen vha. "block nested loop", og projektion klares efterfølgende "on-the-fly".
4. Lige sådan, men der joines vha. "sort-merge".
5. Både selektion og projektion skubbes mest muligt nedad, så joines relationerne sammen vha. "block nested loop", hvorefter den endelige projektion klares "on-the-fly".
6. Lige sådan, men der joines vha. "sort-merge".

Generelle krav i forbindelse med eksamensprojekter

Status

En obligatorisk opgave er en eksamensform. Alle sædvanlige regler for afvikling af eksaminer gælder derfor også for obligatoriske opgaver. Den eneste væsentlige forskel fra en almindelig eksamen er det manglende opsyn. Det ikke tilladt på nogen form at modtage eller afgive information i forbindelse med opgaven med mindre det forudgående eksplicit er blevet tilladt af den ansvarlige lærer. Diskussion, afskrift eller elektronisk kopiering af dele af opgaverne vil blive betragtet og behandlet som eksamenssnyd. Man er forpligtiget til at beskytte sine filer mod at andre kan læse dem (f.eks. vha. `chmod`). Hvis du ikke ved, hvordan man gør, så spørg.

Er du i tvivl om, hvad du må, så tænkt på, om det ville være i orden til en normal eksamen.

Hvis eksamensprojektet laves i grupper har alle gruppens medlemmer på alle måder fuldt ansvar for alt, hvad der afleveres.

Besvarelsen

Besvarelsen skal være velstruktureret og opdelt logisk i afsnit. Der skal skrives præcist og klart.

Der skal læses grundig korrektur på besvarelsen. Der forventes et sprog, herunder bl.a. stavning og tegnsætning, af en rimelig kvalitet. Gør opmærksom på, hvis der er særlige forhold, der er skyld i, at ens besvarelse ikke kan blive sprogligt fejlfri.

Besvarelsen skal udfærdiges på tekstbehandlingsanlæg og udskrives på en højkvalitetsprinter. Udskriften skal være tydelig. Sider, der er skrevet ud med for lav toner, slidt farvebånd eller andet, betragtes som ikke-afleveret.

På forsiden skal der være fuldt navn, fødselsdato og brugernavn (login).

Programmer

Programmer skal være velstrukturerede, letlæselige og kommenterede. Kræves programudskrift vedlagt besvarelsen, gælder der som hovedregel samme regler for denne som for udskrift af besvarelsen.

Hvis der både kræves elektronisk aflevering af programmer samt aflevering af programudskrift, skal udskriften og de elektronisk afleverede programmer være identiske.

Aflevering

Der gives altid betragteligt mere tid til opgaverne end nødvendigt. Til gengæld accepteres for sent afleverede opgaver ikke; heller ikke bare 5 minutter!

Man bør lave og aflevere opgaven mindst 48 timer før afleveringsfristen. At systemet eller alle printere eksempelvis er nede de sidste 24 timer før afleveringsfristen betragtes ikke nødvendigvis som en legal undskyldning.

Der gælder samme regler for en evt. elektronisk aflevering.

Kvittering for aflevering af
Obligatorisk opgave, del 2, DM26, foråret 2000

Udfyldes af de(n) studerende inden afleveringen:

Navne: ....

Fødselsdatoer: ....

Brugernavne: ....

Udfyldes af sekretariatet eller forelæseren:

Modtaget den .... kl. .... af ....
(dato) (klokkeslet) (initialer)

Forside til
Obligatorisk opgave, del 2, DM26, foråret 2000

Navne: ....

Fødselsdatoer: ....

Brugernavne: ....