

E00 (% af 2 timer)

Opgave 1 (20%, E00)

Givet en relation R med skemaet R(A, B, C, D, E) og følgende afhængigheder:

$\{A, B\} \twoheadrightarrow C$,

$D \rightarrow E$.

Spørgsmål 1

Kan det slutes at $\{A, B\} \rightarrow C$ også må gælde?

Kan det slutes at $D \twoheadrightarrow E$?

Giv i forbindelse med svaret en kort forklaring.

Spørgsmål 2

Givet at R indeholder 3 tupler (d,f,a,b,c), (d,f,b,d,e), (e,g,a,b,c).

Hvilke tupler må R da også indeholde?

Spørgsmål 3

Normaliser R til Boyce-Codd normalform.

Opgave 2 (50%, E00)

En database til registrering af produkter i en produktionsvirksomhed har følgende skema, hvor primærnøgler er understreget:

produkt(pnr, pnavn)

styklister(pnr, del-pnr)

lager(lnr, kvm, bynr)

opbevares(pnr, lnr, antal)

by(bynr, bynavn)

Til skemaet knyttes følgende kommentarer. Relationen **produkt** omfatter et produkts stamoplysninger, der her er produktets nummer og navn, hhv. **pnr**, og **pnavn**. **styklister** er en del-af relation, der udtrykker at produktet **del-pnr** er en del af produktet **pnr**. Relationen **lager** indeholder stamoplysninger om et lager, hvilket er et nummer, et areal i kvadratmeter og et bynummer, hhv. **lnr**, **kvm** og **bynr**. I **opbevares** registreres hvilke produkter, der opbevares på et lager; produktet **pnr** opbevares på lageret **lnr**. Endelig indeholder relationen **by** et nummer og et navn for en by i hhv. **bynr** og **bynavn**.

Spørgsmål 1

Skitser et ER-diagram, der kan give anledning til dette skema.

Spørgsmål 2

Formuler i SQL følgende forespørgsler:

a) I hvilke byer (angiv navn) oplagres produktet med pnavn "tryklodder"?

b) Angiv det samlede antal der produceres af produktet med pnavn "tryklodder" fordelt på bynavn, idet der ses bort fra byen med navnet "Herning".

Spørgsmål 3

Et produkts umiddelbare delprodukter er dem, der optræder sammen med produktet i stykliste.

Formuler endvidere i SQL følgende:

- a) I hvilke byer kan man producere produktet med nr. 17? Forudsætningen for at producere et produkt i en by er at alle produktets umiddelbare delprodukter opbevares på et lager i byen.
- b) I hvilke byer opbevares hverken produktet med nummer 17 eller dets umiddelbare delprodukter?
- c)

Spørgsmål 4

Virksomheden har gennemført en politik for produktnummerering, således at et produkt A, der er en del af et andet produkt B, altid har et lavere nummer end det andet produkt.

Skitser i SQL en del af databaseskemaet, der omfatter en begrænsning, som sikrer, at der ikke er "cykler" i styklisten (et produkt A må ikke være en del af sig selv og må ikke have delprodukter, der har A som delprodukt, osv.)

Spørgsmål 5

Virksomheden har endvidere en politik, der går ud på at ethvert produkt skal findes på mindst et lager.

Skitser hertil en trigger der, i forbindelse med oprettelsen af et nyt produkt med $pnr = X$, automatisk sikrer at dette opbevares med antal 1 på lageret med $lnr=0$ - altså opretter en tupel $(X, 0, 1)$ i opbevares.

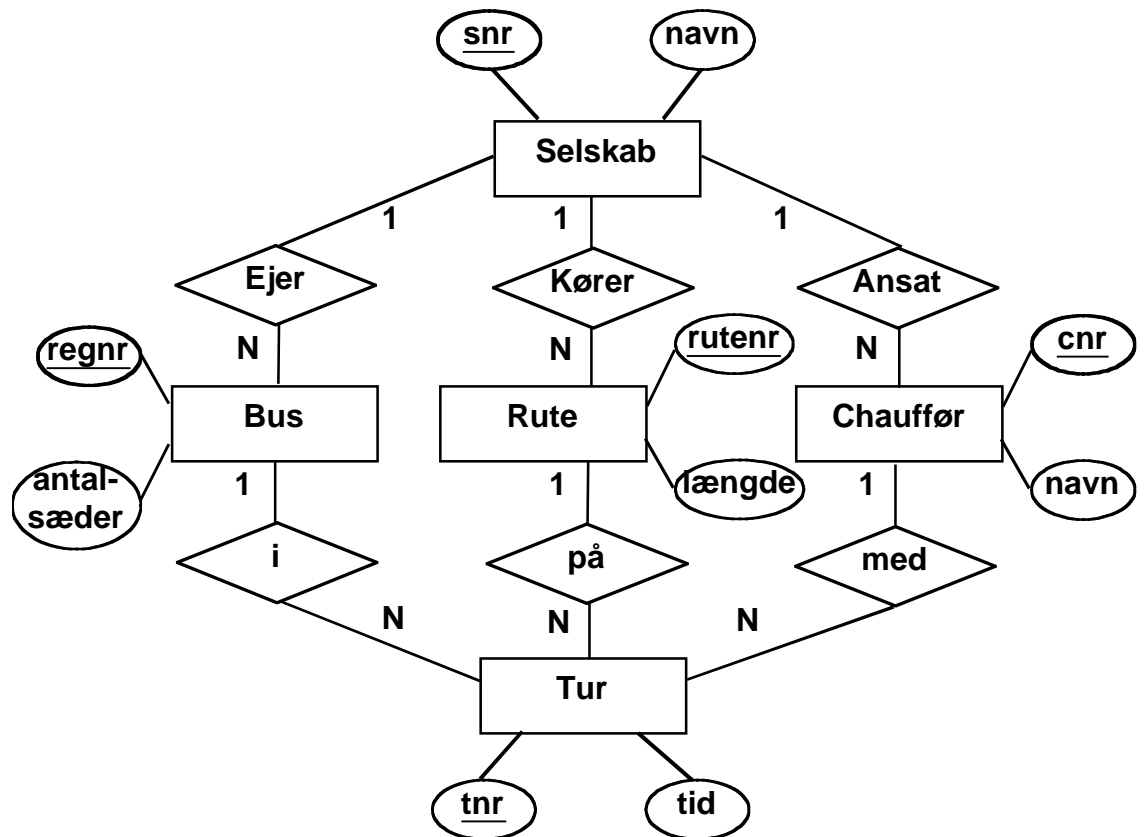
Opgave 3 (30%, E00)

Til administration af den del af det offentlige trafiknet, der vedrører busdrift, skal der konstrueres et system med en database, der registrerer følgende.

Om alle busdrift-**selskaber** registreres et entydigt nummer **snr** samt navn **navn**. For de involverede **busser** registreres bussens registreringsnummer **regnr** og antallet af sæder **antal-sæder**. De **ruter** der køres på registreres med et entydigt **rutenr** og længde for ruten angive i kilometer **længde**. Involverede **chauffører** registreres med et entydigt nummer **cnr** samt med navn **navn**. Endelig registreres samtlige **ture** der køres på en dag med et entydigt nummer **tnr** samt med tidspunkt for påbegyndelse af turen **tid** (der er altså ikke tale om en komplet køreplan – kun tures starttidspunkt er registreret og der findes kun en slags dag – altså ingen specielle lørdage eller søndage, f.eks.)

Selskab er forbundet med **bus**, **rute** og **chauffør** i sammenhænge hhv. **ejer**, **kører** og **ansat**. Disse registrerer for et selskab hvilke busser det ejer, hvilke ruter det kører og hvilke chauffører det har ansat. **Tur** er forbundet med **bus**, **rute** og **chauffør** i sammenhænge hhv. **i**, **på** og **med**. Disse registrerer om en tur at den gemmenføres "i" den givne bus, "på" den givne rute og "med" den givne chauffør.

Et design for denne database er illustreret i nedenstående ER-diagram noterede skema.



Spørgsmål 1

Oversæt skemaet angivet herover til et relationelt databaseskema. Noter evt. blot skemaet ved prædikatnotation (såsom "selskab(snr, navn)"). Husk at angive primærnøgler.

Spørgsmål 2

Formuler i relationel algebra følgende forespørgsel til en database med skemaet fra spørgsmål 1.

- a) På hvilke ruter (angiv rutenr) kører en chauffør med navn "Knud Hansen"?
- b) Hvilke selskaber har ikke en chauffør med navn "Knud Hansen", der kører på en rute med rutenr 17?

Spørgsmål 3

Formuler endvidere i SQL følgende.

- a) Hvormange siddepladser er der sammenlagt i busserne på de ture, der køres af chaufføren med navn "Knud Hansen"?
- b) Formuler en forespørgsel, der resulterer i en liste over chauffører med navn og antal kilometer, der køres

dagligt, idet der på listen kun medtages chauffører der kører mere end 200 km. dagligt.

F00 (% af 2 timer)

Opgave 1 (20%, F00)

Antag at vi har en relation R med fire attributter A, B, C og D. Under punkterne 1-3 nedenfor er angivet tre alternative sæt af funktionelle afhængigheder, der vedrører R.

Angiv for hver af disse,

- (i) nøgle(r) for R og
- (ii) hvilke af de angivne funktionelle afhængigheder, der bryder BCNF, samt
- (iii) en BCNF normalisering af R fremkommet ved at dekomponere R svarende til de angivne afhængigheder.

OBS: Der er tale om tre alternative sæt af afhængigheder, der således skal behandles isoleret.

1. C D, C A, B C
2. B C, D A
3. A B, BC D, A C

Opgave 2(10%, F00)

Antag at vi, for en relation S med skema S(A,B,C), har en forekomst med følgende fire tupler: (1,2,3),(4,2,3),(5,3,3),(5,3,4). Hvilke af afhængighederne 1-6 herunder kan man slutte ikke kan gælde for S (m.a.o. hvilke afhængigheder brydes af den givne forekomst).

1. A -> B
2. A ->> B
3. BC -> A
4. BC ->> A
5. B -> C
6. B ->> C

Opgave 3 (40%, F00)

En database til håndtering af et flyselskabs flyvninger er designet med følgende skema:

Lufthavn(Lnr, By)

Flytype(Tnr, Navn, Antal-sæder, Hastighed)

Kan_lande(Lnr, Tnr)

Fly(Flynr, Tnr)

Rute(Rnr, Afstand, Fra-Lnr, Til-Lnr)

Flyvning(Flyvningnr, Dag, Tid, Flynr, Rnr)

I skemaet er nøgler understreget, og der er indført særlige attributter til entydig identifikation (Lnr for lufthavnsnummer, Tnr for Flytypenummer, m.fl).

Lufthavn registrerer lufthavne med navn på byen, hvor de hører hjemme. I Flytype kan registreres flytyper med typebetegnelse (Navn), antal sæder, samt den hastighed de skal flyve med (det antages her kun at benyttes én hastighed).

For flytyper er der tillige med Kan_lande registreret, hvilke lufthavne de kan lande i.

Konkrete fly er registreret i Fly med angivelse af flytype (Tnr). En rute, der registreres i Rute, er blot en kombination af en afgangslufthavn (Fra-Lnr) og en ankomstlufthavn (Til-Lnr) med angivelse af afstand i kilometer imellem de to lufthavne.

En flyvning, der registreres i Flyvning, foregår på en bestemt ugedag (Dag), til et fastlagt tidspunkt, med et bestemt fly og på en bestemt rute.

Spørgsmål 1

Skitser et ER-diagram der kan give anledning til det angivne relationelle databaseskema.

Spørgsmål 2

Formuler følgende forespørgsler, til en database med det givne database-skema, i SQL.

d) Hvilke ruter (rnr), med en afstand (Afstand) på mindre end 500 km, går til lufthavnen Lnr = "L21".

e) Hvilke ruter (rnr) har bestemmelsessted (by for lufthavn der flyves til) "København", der udgår fra lufthavne i en afstand af mindre end 500 km. fra "København".

f) Angiv dag (Dag), afgangstidspunkt (Tid) og flyvetid for samtlige flyvninger på ruten med rutenummer Rnr="R35". Flyvetiden beregnes som Afstand/Hastighed.

Spørgsmål 3

Forklar hvad svaret udgør på forespørgslerne (udtrykt i hhv. SQL og relationel algebra) :

a)

```
select fra.by,til.by
from rute, Lufthavn fra, Lufthavn til
where rute.fra-lnr=fra.lnr and rute.til-lnr=til.lnr
union
select fra.by,til.by
from rute r1, rute r2, Lufthavn fra, Lufthavn til
where r1.fra-lnr=fra.lnr and r1.til-lnr=r2.fra-lnr and r2.til-
lnr=til.lnr
union
select fra.by,til.by
from rute r1, rute r2, rute r3, Lufthavn fra, Lufthavn til
where r1.fra-lnr=fra.lnr and r1.til-lnr=r2.fra-lnr and r2.til-
lnr=r3.fra-lnr and r3.til-lnr=til.lnr
```

b)

$$\pi_{\text{navn}}(\text{Flytype} \bowtie \text{Fly} \bowtie \text{Flyvning} \bowtie \sigma_{\text{By}='Stokholm'}(\text{Rute} \bowtie_{\text{Fra-Lnr=Lnr}} \text{Lufthavn}))$$

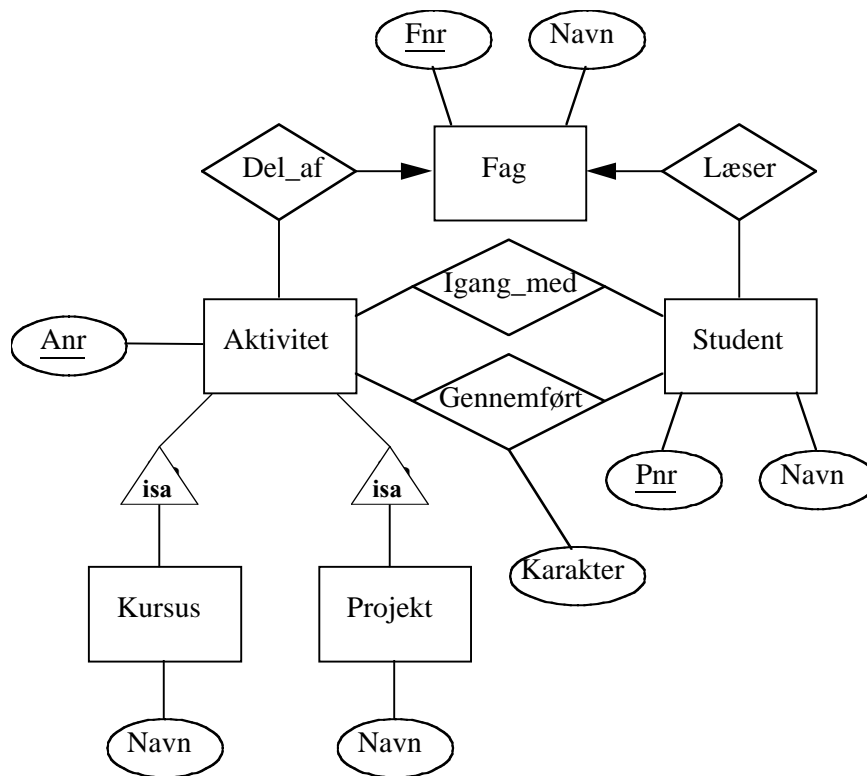
c)

$$\pi_B(\rho_{R(A,B,C,D)}(\text{Flytype}) - \pi_{A,B,C,D}(\rho_{R(A,B,C,D)}(\text{Flytype}))) \bowtie_{D=H}(\rho_{S(E,F,G,H)}(\text{Flytype})))$$

Opgave 4 (30%, F00)

Nedenstående ER-diagram repræsenterer et design af en database til registrering på et uddannelses-sted, af:

- hvad de forskellige fag omfatter
- hvilke studerende, der er aktive på stedet samt
- for de studerende, hvilke aktiviteter de er i gang med og har gennemført.



Spørgsmål 1.

Lav en oversættelse det skitserede skema til et relationelt database skema. Nøgler skal medtages.

Spørgsmål 2

Forklar hvad der naturligt – ud fra beskrivelse og diagram ovenfor – må være begrænsninger på databasen, svarende til referenceintegritet.

Formuler tillige disse begrænsninger ved brug af relationel algebra, idet der refereres til skemaet fra spørgsmål 1.

E99 (% af 4 timer)

Opgave 3 (25%)

Til administration af busdrift er konstrueret en database med følgende relationelle skema (nøgler er understreget):

selskab(snr, navn)

bus(regnr, antal-sæde, snr)

rute(rutenr, længde, snr)

chauffør(cnr, navn, snr)

tur(tid, regnr, rutenr, cnr)

Om alle busdriftselskaber er i **selskab** registreret et nummer samt et navn i hhv **snr** og **navn**. I bus er for hver bus registreret et registreringsnummer, et antal sæder og et selskabsnummer for selskabet, som bussen tilhører (i hhv. **regnr**, **antal-sæde** og **snr**). En **rute** er registreret ved et nummer (**rutenr**), en længde i kilometer (**længde**) og et selskabsnummer (**snr**) for selskabet, som har ansvaret for ruten, og en **chauffør** er registreret ved et nummer (**cnr**), et navn (**navn**) samt et nummer for selskabet (**snr**), der har ansat chaufføren. Endelig registreres med **tur**, for hver tur en bus (**regnr**) kører på, en rute (**rutenr**) med en chauffør (**cnr**) og den tid der tager at køre turen (**tid**).

Spørgsmål 1

Skitser et ER-diagram, der kan give anledning til ovenstående relationelle databaseskema.

Spørgsmål 2.

Skitser ændringer i både ER-diagram og relationelt skema, der tager højde for at der indføres et entydigt nummer **tnr** på **tur**. I det følgende kan der ses bort fra disse ændringer.

Spørgsmål 3

Formuler i relationel algebra følgende forespørgsler til en database med skemaet fra spørgsmål 1.

- På hvilke ruter (angiv rutenr) kører chaufføren med navn "Knud Hansen"?
- Hvilke selskaber (angiv navn) har har enten busser med flere end 60 sæder eller ruter, der er længere end 200 km.
- Hvilke selskaber (angiv navn) har ingen chauffør, der hedder "Knud Hansen", men har enten busser med flere end 60 sæder eller ruter, der er længere end 200 km.

Spørgsmål 4

Angiv med SQL navn for selskab og det totale antal sæder i samtlige selskabets busser opdelt på selskaber, idet der kun medtages selskaber med et gennemsnitligt antal sæder større end 10.

Spørgsmål 5

Med en naturlig join i udtrykket
selskab*bus

kan de to indgående relationer samles til een. (* er her naturlig join operatoren i relationel algebra)

Angiv skemaet for resultatrelationen (der f.eks. kan kaldes sb) med nøgle og forklar endvidere hvorfor denne ikke er på Boyce-Codd normalform.

Opgave 4 (10%)

Spørgsmål 1

Betragt mængden Γ af fem propositionelle formler opskrevet på klausulform:

$$\Gamma = \{Q \vee \neg P, R \vee \neg P, S \vee \neg P, P, \neg Q \vee \neg S \vee \neg R \vee T, \neg Q\}$$

Vis ved resolution at Γ er inkonsistent.

Spørgsmål 2

Benyt modbevis ved resolution til at vise at med

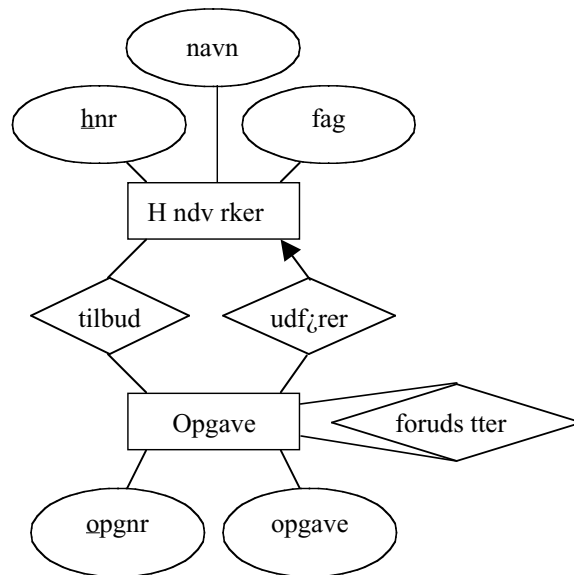
$$\Gamma' = \{Q \vee \neg P, R \vee \neg P, S \vee \neg P, P, \neg Q \vee \neg S \vee \neg R \vee T\}$$

har vi at $\Gamma' \vdash T$, altså at formlen T er beviselig ud fra formlerne Γ' .

F99 (% af 4 timer)

Opgave 3 (25%)

En database til at støtte en bygherre er skitseret i følgende ER-diagram:



Med et skema som det viste skal kunne håndteres et byggeprojekt, der involverer håndværkere og opgaver. Tilbud gives af håndværkere på opgaver, og til en opgave udvælges, på baggrund af de givne tilbud, en håndværker der skal udføre den. En opgave kan forudsætte at andre opgaver er udført (f.eks. fundamentet skal støbes før murerne kan mures).

Spørgsmål 1

Oversæt ovenstående ER-skema til et relationelt databaseskema. Angiv heri nøgler og sørg for at det relationelle skema kan udtrykke det samme som ER-skemaet.

Spørgsmål 2

Formuler til en database, med skemaet fra spørgsmål 1, følgende forespørgsler i SQL:

A: Hvilke opgaver findes der flere tilbud på?

B: Angiv de fag, hvor der findes håndværkere der kan starte byggeriet (dvs. håndværkere der skal udføre opgaver, der ikke forudsætter andre opgaver).

Spørgsmål 3

Modificer ER-skemaet således at højst to tilbud kan afgives på den samme opgave.

Skitser tillige en oversættelse til et relationelt skema af denne ændring.

Spørgsmål 4

Formuler i SQL en forespørgsel, der som svar angiver forudsatte opgaver med to attributter (opgave, forudsat-opgave). Forespørgslen skal udtrække forudsatte opgaver i to niveauer (hvis tag forudsætter mur og mur forudsætter fundament så skal ikke kun (tag,mur) og (mur, fundament) være i svaret, men også (tag, fundament)). Forklar hvordan en sådan forespørgsel kan formuleres til flere niveauer af forudsatte opgaver.

Spørgsmål 5

Antag at vi som (delvist overlappende) supplement til ovenstående har en registrering af håndværkere med deres fag og de opgaver de kan udføre i følgende skema:

Kan-udføre(hnr, fag, opgave)

og at der gælder følgende multiværdi-afhængighed: hnr \rightarrow fag.

Normaliser Kan-udføre til 4' de normalform.

Opgave 4 (10%)

Betragt mængden Γ af tre propositionelle formler opskrevet på klausulform (idet \neg er negation og \vee er logisk eller):

$$\Gamma = \{ \neg P \vee \neg Q \vee \neg R \vee S, P, R \}$$

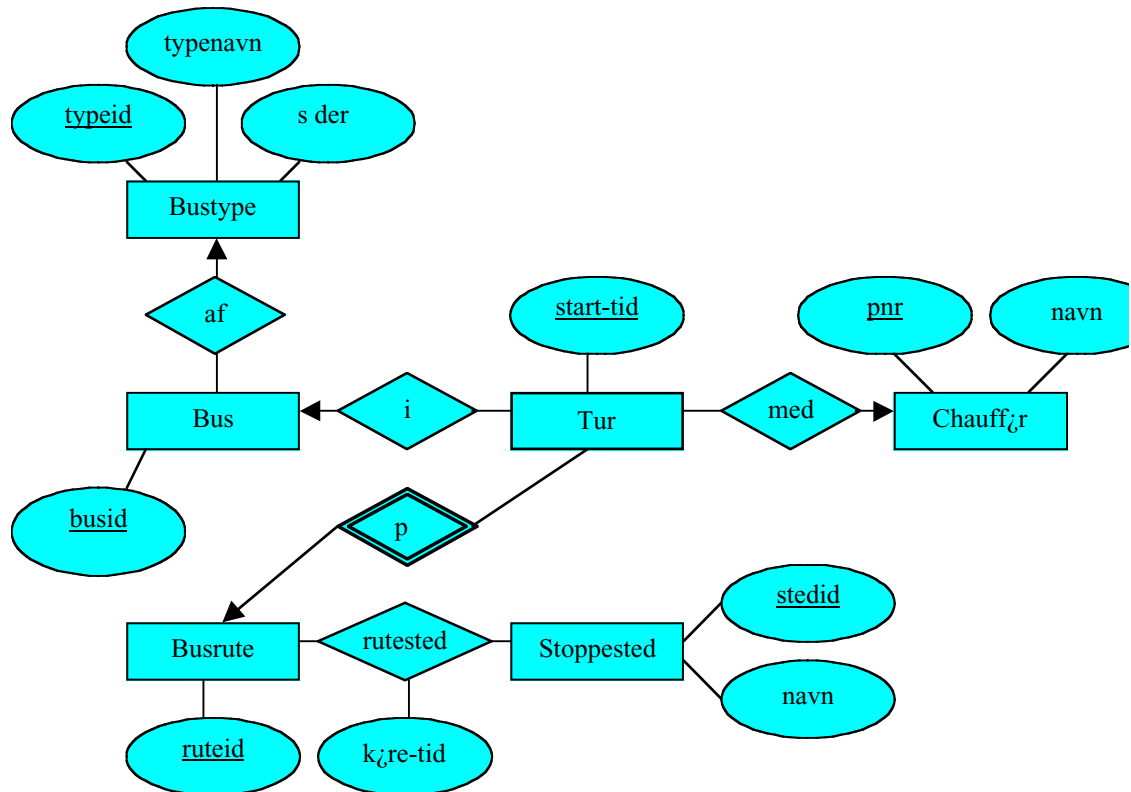
Spørgsmål 1

Tilføj een formel til Γ således at $\neg Q$ bliver beviselig og bevis $\neg Q$ ved modbevis ved resolution.

E98 (% af 4 timer)

Opgave 4 (15%)

En database til et busselskab, der kan understøtte administrationen af køre- og arbejdsplaner er skitseret i følgende ER-diagram:



Til skemaet skal bemærkes følgende:

- en tur er noget, der foregår på en busrute, i en bus med en chauffør og som starter til et bestemt tidspunkt (start-tid)
- køreplaner er registreret i databasen sådan, at der til hvert sted på ruten er tilknyttet en køre-tid, dvs. den tid det tager på en tur at nå fra første stoppested til det aktuelle (første sted har derfor køretid 0).

Spørgsmål 1

Oversæt ovenstående ER-skema til et relationelt databaseskema. Angiv heri nøgler (men se bort fra begrænsninger i øvrigt) og sørg for at det relationelle skema kan udtrykke det samme som ER-skemaet.

Spørgsmål 2

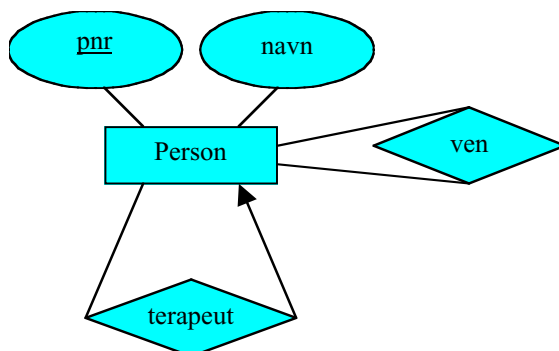
Modifier ER-skemaet således at "rutedsted" bliver en entitetstype i stedet for en sammenhæng ("relationship"), mens en oversættelse som i spørgsmål 1 giver samme resultat (også hvad angår primærnøgler).

Spørgsmål 3

Modifier ER-skemaet således at Tur bliver en sammenhæng i stedet for en entitetstype, mens en oversættelse igen giver samme resultat som i spørgsmål 1 (også hvad angår primærnøgler).

Opgave 5 (15%)

I nedenstående ER-diagram skitseres en database, der kan omfatte oplysninger om personer med nr. og navn samt personers "venne- og terapeut-relationer".



Spørgsmål 1

Hvordan vil et relationelt databaseskema, der svarer hertil, med to relationsskemaer *person* og *ven*, komme til at se ud?

Hvilke integritets-begrænsninger bør medtages fordi de er udtrykt i ER-skemaet?

Spørgsmål 2

Formuler til databasen forespørgslerne:

- Hvem er ven med deres terapeut?
- Hvilke terapeuter har venner?

Det kan antages at, hvis der står i databasen, at A har en ven B, så står der også i databasen, at B har en ven A.

Spørgsmål 3

Hvilke terapeuter har flere venner, der er terapeuter?

Dan som svar en liste over terapeuter, der har mindst to venner, der er terapeuter.

F98 (% af 4 timer)

Opgave 4 (20%)

I forbindelse med arrangement af en konference er der behov for en database. Behovet er analyseret og resultatet foreligger i form af følgende relationelle databaseskema.

person(pnr, pnavn)

artikel(anr, atitel, præsenteres-af-pnr, holdes-i-snr)

session(snr, stitel, ordstyrer-pnr)

forfatter(pnr, anr)

referee(pnr, anr)

tilmelding(pnr, knr)

deltager-kategori(knr, pris)

De understregede attributter er primærnøgler. Skemaet er kommenteret herunder. Det skal bemærkes at konferencen er opdelt i "sessioner", dvs. samlinger af (typisk 3-4) artikel-præsentationer, der hver styres af en "ordstyrer" og at artikler i forbindelse med optagelse på konferencen gennemgår en kontrol-proces, hvor to eller flere personer vurderer hver artikel. En person der deltager i denne proces kaldes en "referee". For at kunne beregne forskellige priser for deltagelse i konferencen er der indført såkaldte "deltager-kategorier".

person: en person, der har noget med konferencen at gøre

pnr: personnummer

pnavn: personens navn

artikel: en artikel, der er indsendt til præsentation på konferencen

anr: artikelnummer

atitel: artiklens titel

præsenteres-af-pnr: pnr for personen, der præsenterer artiklen

holdes-i-snr: snr for sessionen hvori artiklen præsenteres

session: en samling af (typisk 3-4) artikel-præsentationer på konferencen

snr: sessionens nummer

stitel: sessionens titel

ordstyrer-pnr: pnr for ordstyreren på sessionen

forfatter: en forfatter til en artikel

pnr: pnr for den person, der er forfatter

anr: anr for den artikel personen er forfatter

referee: en person der vurderer en artikel

pnr: pnr for den person der vurderer artiklen

anr: anr for den artikel personen vurderer

tilmelding: en tilmelding til deltagelse i konferencen

pnr: pnr for den person der er tilmeldt

knr: knr for den deltager-kategori personen er tilmeldt under

deltager-kategori: en opdeling af tilmeldinger

knr: kategoriens nummer

pris: prisen for deltagelse under den pågældende kategori

Spørgsmål 1

Skitsér et database-skema i form af et ER-diagram, der ved oversættelse kan give anledning til det relationelle skema skitseret ovenfor. Sørg for at få alle attributter med og noter også nøgler for entiteter.

Spørgsmål 2

Formuler følgende som forespørgsler i SQL:

- a) Hvilke sessioner har ordstyrere, der deltager gratis.
- b) Hvilke artikler præsenteres af "substitutter". En substitut er en person, der præsenterer en artikel som vedkommende ikke er forfatter til.
- c) Hvilke personer mangler at tilmelde sig (personer der enten skal præsenterer en artikel eller være ordstyrer for en session)

Spørgsmål 3

Hvilke spørgsmål (på almindeligt dansk) er der formuleret med følgende SQL-udtryk:

- a)

```
select pnavn
from person
where pnr in
(select ordstyrer-pnr from session where snr=3)
union
(select pnr from referee, artikel
where referee.anr=artikel.anr and artikel.holdes-i-snr=3)
union
(select pnr from forfatter, artikel
where forfatter.anr=artikel.anr and artikel.holdes-i-snr=3)
union
(select præsenteres-af-pnr from artikel where holdes-i-snr=3)
```
- b)

```
select gruppe, count(pris), avg(pris)
from
(select 'o' as gruppe,ordstyrer-pnr as pnr from session
union
select 'r' as gruppe, pnr from referee
union
select 'f' as gruppe, pnr from forfatter
union
select 'p' as gruppe,præsenteres-af-pnr as pnr from artikel
) as alle, tilmelding, deltager-kategori
where alle.pnr=tilmelding.pnr and tilmelding.knr=deltager-
kategori.knr
group by gruppe
```

Opgave 5 (10%)

For en relation med skemaet

R(A, B, C, D, E)

er følgende funktionelle afhængigheder givet:

BC -> D

A \rightarrow B
D \rightarrow E

Spørgsmål 1

Følgende er IKKE en tilladt forekomst af relationen R. Forklar hvorfor.

R	A	B	C	D	E
	a	b	c	d	e
	e	a	b	c	d
	f	b	c	e	d
	e	a	b	c	e

(I de følgende to spørgsmål ses naturligvis bort fra denne forekomst.)

Spørgsmål 2

Angiv en nøgle for relationen R .

Spørgsmål 3

Normaliser R til BCNF (Boyce-Codd normalform). Angiv nøgler i de resulterende relationsskemaer.