

# UNIVERSITETET I OSLO

## Det matematisk-naturvitenskapelige fakultet

Eksamen i:	INF1300 — Introduksjon til databaser
Eksamensdag:	28. november 2016
Tid for eksamen:	09.00 – 15.00
Oppgavesettet er på:	7 sider
Vedlegg:	ORM 2 Graphical Notation, 6 sider
Tillatte hjelpemidler:	Halpin & Morgan: <i>Information Modeling and Relational Databases</i> . 2 A4-ark (4 sider) håndskrevne notater.

*Kontroller at oppgavesettet er komplett før du begynner å besvare spørsmålene.*

### **Del A - Idrett og dopingkontroll (60%)**

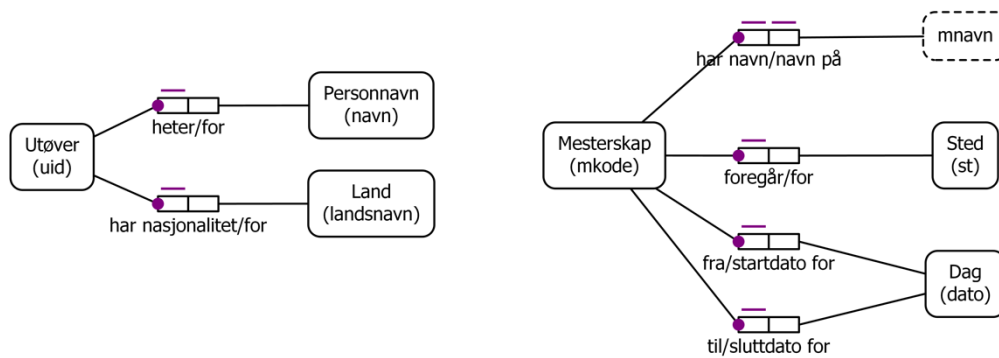
***For å bestå eksamen må del A oppnå karakteren E eller bedre.***

*Generelle retningslinjer til del A:*

- *Les gjennom hele del A før du begynner å modellere.*
- *Ha gjerne med kommentarer i modellen. Om du mener noe er uklart i oppgaveteksten, kan du spesifisere med kommentarer hvilke antakelser du har lagt til grunn.*
- *Du kan med fordel dele opp ORM-modellen over flere sider. Skriv på besvarelsen hvilke sider som gjelder hvilke oppgaver.*
- *Tegn og skriv tydelig.*

Alle toppidrettsutøvere (i det videre bare kalt "utøvere") kan når som helst bli innkalt til dopingkontroll der det tas en urinprøve eller blodprøve. Slike kontroller blir ikke varslet på forhånd. Når en utøver blir varslet om kontroll, vil han eller hun være under konstant oppsyn av en dopingkontrollør helt til prøven er tatt.

Informasjon om idrettsmesterskap og utøvere er reflektert i relasjonsdatabaseskjemaet i del B i oppgavesettet. I figur 1 viser vi deler av den tilsvarende ORM-modellen slik at dere ikke behøver å definere disse begrepene selv (og slik at dere kan bruke dem i den videre modelleringen).



Figur 1: ORM-modeller av utøvere og mesterskap.

Vi skal se på tre forskjellige komponenter som inngår i informasjonssystemet til Antidoping Norge.

### Oppgave 1 - ADAMS: Oppholdsstedsrapportering (15%)

Alle utøverne må fire ganger i året innrapportere detaljer om hvilke mesterskap de skal delta i, og hvor de til enhver tid befinner seg. Dette er for at utøverne når som helst skal kunne bli innkalt til dopingkontroll. Informasjonssystemet som mottar og lagrer informasjon om hvor utøverne befinner seg, heter ADAMS (Anti-Doping Administration and Management System). Under er en sterkt forenklet beskrivelse av de opplysningene dette systemet inneholder.

Informasjonen i ADAMS organiseres som fire årlige kvartalsoversikter for hver utøver - dvs. én oversikt for 1. kvartal (januar, februar, mars), én for 2. kvartal (april, mai, juni), osv. Hver kvartalsoversikt inneholder noen overordnede opplysninger, som blant annet hvilke mesterskap utøveren skal delta i. Resten av opplysningene gjelder for kortere tidsrom. En av disse er den såkalte "garantitiden". Dette er et trefftidspunkt på én time mellom kl. 05:00 og 23:00. For enkelhets skyld organiseres denne og liknende opplysninger i form av dagsoversikter for hver utøver, én for hver dag i kvartalet.

**Deloppgave 1.1 Kvartalsoversikt (5%):** Hver kvartalsoversikt gjelder én utøver, ett år og ett kvartal i dette året. Kvartalsoversikten inneholder opplysninger om postadresse og navnene på de mesterskapene utøveren skal delta i. Modeller dette i ORM.

**Deloppgave 1.2 Dagsoversikt (5%):** En dagsoversikt har alltid informasjon om garantitiden og overnattingsadressen til utøveren for denne dagen (overnattingsadressen gjelder natt til neste dag). Dessuten kan dagsoversikten ha informasjon om treningstid og treningsadresse (dette er ikke aktuelt under mesterskap, derfor er det ikke alle dagsoversiktene som inneholder slik informasjon). Modeller dette i ORM.

**Deloppgave 1.3 Dagsoversikter og kvartalsoversikter (5%):** Hvis vi har en gitt dato og en kvartalsoversikt for en utøver, så skal det bare være én mulig dagsoversikt for denne datoen. Hvis dette ikke allerede er ivarettatt av ORM-modellen din, så gjør endringer eller tilføyelser slik at ORM-modellen reflekterer dette.

## **Oppgave 2 - NAPMU: Biologiske pass (30%)**

For å styrke dopingarbeidet benytter Antidoping Norge biologiske pass. De nordiske landene har et felles informasjonssystem for dette, NAPMU (Nordic Athlete Passport Management Unit). Hver utøver har sitt personlige biologiske pass. Det biologiske passet er basert på analyseresultatene fra dopingprøvene, og viser blant annet hvordan blodverdiene endrer seg over tid.

Helt konkret inneholder det biologiske passet analyseresultater fra alle urinprøvene og blodprøvene. I tillegg kan det blant annet inneholde en blodprofil (hematologisk profil) som blir utarbeidet på grunnlag av blodprøvene, og en hormonprofil (steroidprofil) som blir utarbeidet på grunnlag av urinprøvene. Én prøve kan gi flere analyseresultater fordi den kan testes for flere ting. For eksempel kan én blodprøve testes for både alkohol, veksthormon og CERA (et middel som får nyrene til å produsere mer EPO-hormoner).

Det biologiske passet kan benyttes til å planlegge og gjennomføre ytterligere analyser, enten på allerede innhentede prøver, eller til å planlegge videre målrettet testing ("intelligent testing"). Siden prøvene blir oppbevart i lengre tid, kan det utføres nye analyser på gamle prøver. Derfor er det viktig at systemet ikke bare har informasjon om når en prøve ble tatt, men også når hver enkelt analyse ble gjennomført, og av hvilket laboratorium. Endelig kan det biologiske passet benyttes som bevis på brudd på dopingreglene.

**Deloppgave 2.1 (10%)** Modeller prøver og prøvetyper i ORM, dvs. det som skal til for å beskrive at en prøve enten er en blodprøve eller en urinprøve, hvilken utøver prøven gjaldt, og når prøven ble tatt.

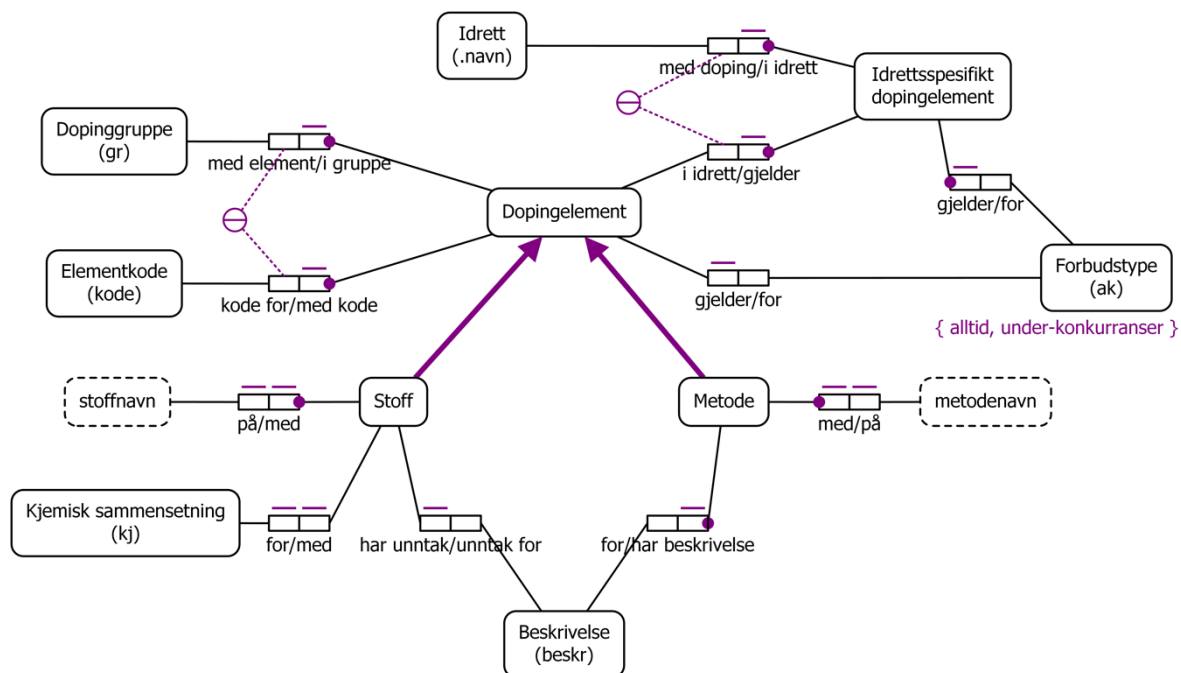
**Deloppgave 2.2 (15%)** Modeller analyser og analyseresultater i ORM.

**Deloppgave 2.3 (5%)** Modeller biologiske pass i ORM.

## **Oppgave 3 - WADA-listen: Forbudte stoffer og metoder (15%)**

Den såkalte WADA-listen (i dagligtale "dopinglisten") inneholder informasjon om forbudte stoffer. De forbudte stoffene er inndelt i grupper etter blant annet virkemåte. Hvert forbudt stoff tilhører én slik gruppe. Hvert stoff har et entydig navn og eventuelt en entydig kjemisk beskrivelse. Noen stoffer er totalforbudt, mens andre bare er forbudt hvis de benyttes i forbindelse med en konkurranse. Noen stoffer er forbudt i enkelte idretter, men tillatt i andre. For noen stoffer finnes det unntaksbestemmelser, som at et stoff er tillatt hvis dosen ikke overstiger en gitt mengde.

Videre inneholder WADA-listen informasjon om forbudte metoder, som bloddoping, gendoping og annet. Også for de forbudte metodene er noen metoder totalforbudt, mens andre bare er forbudt hvis de brukes i forbindelse med en konkurranse. Også her er det slik at noen metoder kan være forbudt i enkelte idretter, men tillatt i andre.

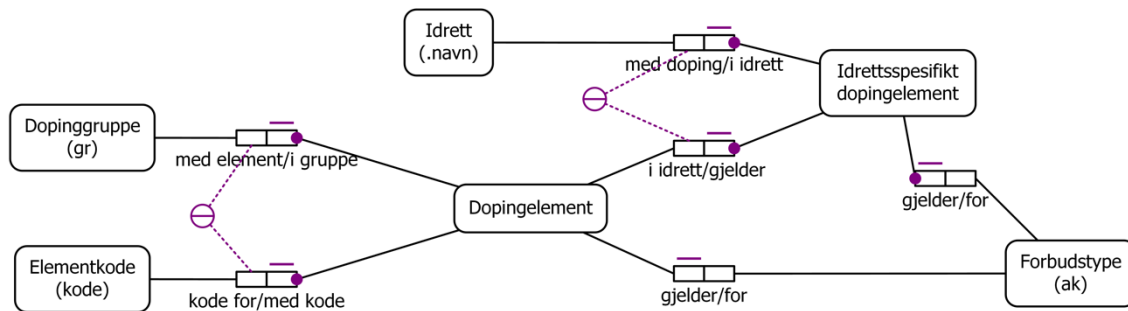


Figur 2: ORM-modell av WADA-listen.

**Deloppgave 3.1 (5%)** I figur 2 finner du en ORM-modell av WADA-listen. For hver av påstandene under, angi om påstanden er sann eller gal i modellen.

- a** Et dopingelement er enten et stoff eller en metode.
- b** Et stoff kan ikke tilhøre mer enn én dopinggruppe.
- c** En dopinggruppe inneholder enten bare stoffer eller bare metoder.
- d** Til hvert dopingelement er det nøyaktig én forbudstype (dvs. om elementet alltid er forbudt, eller om det bare er forbudt under konkurranser).

**Deloppgave 3.2 (10%)** I denne deloppgaven skal vi se på realisering (gruppering) av ORM-modeller som relasjonsdatabaseskjema. Du skal ikke angi datatyper for attributtene og ikke bruke SQL. For hver relasjon, angi relasjonens navn og navnet til hvert attributt. Undertrykkede relasjoner skal ikke være med i listen.



Figur 3: Utsnitt av ORM-modellen i figur 2.

**a** Figur 3 er et utsnitt av modellen i figur 2. Realiser ORM-modellen i figur 3 som et relasjonsdatabaseskjema. Understrek primærnøkklene.

**b** Realiser som et relasjonsdatabaseskjema det av ORM-modellen i figur 2 som ikke er dekket av skjemaet i **a**. Understrek primærnøkklene. Hvis det er kandidatnøkler, markerer du dem med én strek og primærnøkkelen med to.

## Del B Skøytemesterskap (hurtigløp) (40%)

Hurtigløp på skøyter kan sammenlignes med baneløp i friidrett, der det løpes en gitt distanse på en oval bane. Kvinner og menn konkurrerer hver for seg. De distansene det konkurreres i, er delvis de samme for kvinner og menn, og delvis forskjellige. For eksempel konkurrerer både kvinner og menn på distansen 500 meter, mens bare menn konkurrerer på distansen 10 000 meter. Et mesterskap består av flere disipliner. En disiplin kan identifiseres entydig ved kjønn og distanse. Eksempler på noen disipliner er (k,500), (m,500) og (m,10000).

I resten av dette oppgavesettet skal vi se på en liten database som tar for seg skøytemesterskap. (For de av dere som kjenner denne sporten godt: Vær oppmerksom på at databasen ikke kan brukes til å representere resultatene fra ekte skøytemesterskap fordi vi har gjort en del forenklinger.)

Utøver(uid, navn, land)

Mesterskap(mkode, mnavn, sted, fradato, tildato)

Resultatliste(mkode, kjønn, distanse, uid, [tid], [poeng], [plassering])

- Utøvere er identifisert ved uid. Alle utøvere må ha en nasjonalitet (land).
- I Mesterskap er mesterskapskoden (mkode) primærnøkkel og navnet på mesterskapet en kandidatnøkkel.
- I Resultatliste er distanse lengden på løpet i meter, mens tid er lagret i sekunder. For eksempel svarer 395,67 til tiden 6 minutter og 35,67 sekunder. Poeng(p) beregnes utifra tid(t) og distanse(d) slik:  $p = 500 * t / d$ .

Her er noen eksempler på mulige tupler i Mesterskap og Resultatliste:

#### Mesterskap

mkode	mnavn	sted	fradato	tildato
svm16	Sprint-VM 2016	Seoul	27.2.2016	28.2.2016
avm16	Allround-VM 2016	Berlin	5.3.2016	6.3.2016

#### Resultatliste

mkode	kjønn	distanse	uid	tid	poeng	plassering
avm16	k	5000	45677	431,95	43,195	8
avm16	m	500	34630	34,60	34,600	2
avm16	k	1500	45677	107,70	35,900	3
avm16	m	5000	54002	395,67	39,567	13

Poengberegningen gjør at vi i tabellen Resultatliste blant annet har den funksjonelle avhengigheten (FDen)

distanse, tid → poeng

#### Oppgave 4 - normalformer (5%)

Gi et begrunnet svar på hvilken normalform Resultatliste er på.

#### Oppgave 5 - SQL (35%)

**Deloppgave 5.1 (5%)** Bruk SQL til å definere tabellen Resultatliste.

I de følgende deloppgavene skal du skrive select-setninger som definerer/skriver ut tabellen det spørres etter. Det er lov å bruke views.

#### Deloppgave 5.2 (5%)

Lag en tabell med navn og nasjonalitet på utøvere som deltok i mesterskapet med mkode avm16. Tabellen skal ordnes slik at alle utøvere fra samme nasjon kommer etter hverandre.

#### Deloppgave 5.3 (5%)

Lag en tabell over antall utøvere pr. land i Enkeltdistanse-VM 2015 (mkode 'evm15'). Ordne listen etter fallende antall utøvere.

#### Deloppgave 5.4 (5%)

Finn navn på utøvere som deltok i allround-VM 2016 (avm16) og hadde en samlet poengsum på mellom 148 og 152 poeng.

**Deloppgave 5.5 (5%)**

Hvert allround-VM har en kode på formen 'avm16'. Allround-VM i 1967 har for eksempel kode 'avm67'. Lag en tabell over antall utøvere for hver disiplin i hvert enkelt allround-VM. Et tuppel fra tabellen skal ha fire verdier:

(mesterskapsnavn, kjønn, distanse, antallutøvere)

Tuplene skal ordnes etter mesterskapsnavn.

**Deloppgave 5.6 (5%)**

Finn navn på alle utøverne som fikk bedre tid enn alle norske utøvere på 500-meter for menn (m,500) i allround-VM 2016 (avm16).

**Deloppgave 5.7 (5%)**

Finn navn på utøvere som aldri har gått 500-meteren på dårligere tid enn 40,00 sekunder.

---

*Oppgavesett slutt. Lykke til!*

*Ellen Munthe-Kaas og Stein Michael Storleer*