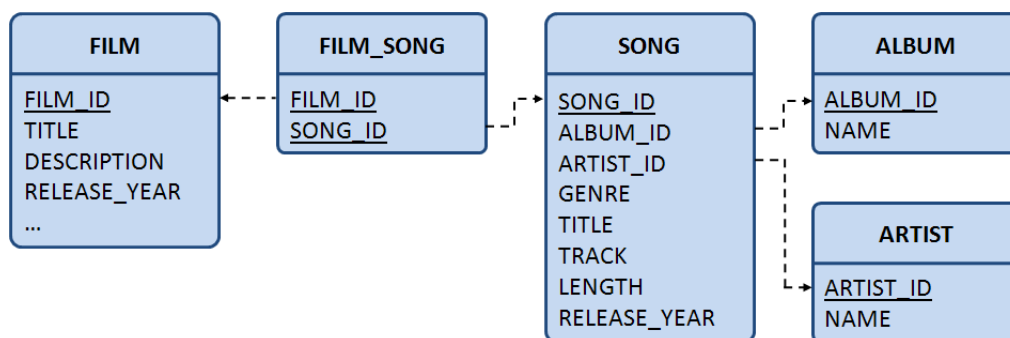

EXAMEN: Scriptingtalen

Prof. Dr. Peter Dawyndt
1^e Bachelor Informatica
groep 1

woensdag 20-08-2008, 8:30h
academiejaar 2007-2008
tweede zitting

Opgave 1

De DVD-verhuurketen heeft zijn Sakila databank uitgebreid, waardoor informatie kan worden opgezocht over soundtracks van films. Hiervoor werden vier extra tabellen aan de oorspronkelijke databank toegevoegd. Onderstaande figuur geeft een overzicht van de nieuwe tabellen, waarbij (samengestelde) primaire sleutels onderlijnd zijn en vreemde sleutels worden aangegeven met onderbroken pijlen.



De tabel **SONG** houdt een lijst bij van liedjes, waarbij aan elk liedje een uniek rangnummer toegekend wordt (kolom **SONG_ID**). De kolom **ALBUM_ID** is een vreemde sleutel die verwijst naar de gelijknamige kolom van de tabel **ALBUM**, en die aangeeft op welk album het liedje voorkomt. De kolom **ARTIST_ID** is een vreemde sleutel die verwijst naar de gelijknamige kolom van de tabel **ARTIST**, en die aangeeft welke artiest het nummer geschreven heeft. De tabel **ALBUM** (resp. **ARTIST**) bevat verder enkel de naam van het album (resp. de artiest). Voorts geeft de tabel **SONG** ook nog informatie weer over de titel van het nummer (**TITLE**), het genre (**GENRE**), de plaats van het nummer op het album (**TRACK**), de lengte van het nummer (**LENGTH**) en het jaar waarin het werd uitgegeven (**RELEASE_YEAR**). De tabel **FILM_SONG** geeft aan welk nummer in welke film gebruikt wordt, en bevat daarvoor twee vreemde sleutels die verwijzen naar de primaire sleutels van de tabellen **FILM** en **SONG**.

De uitgebreide Sakila databank werd ingeladen in een MySQL RDBMS en is bevragebaar via de phpMyAdmin webapplicatie (http://users.ugent.be/~dhollevo/scrita_sql/). Inloggen doe je in deze databank met je UGent account. Gevraagd wordt om SQL zoekopdrachten te formuleren die een antwoord bieden op onderstaande vragen. Zorg er telkens voor dat de kolommen van de resulterende tabel een zinvolle naam krijgen. Geef in je antwoordbestand ook aan hoeveel records de resultatentabel bevat.

1. Bereken het minimum, maximum en gemiddeld aantal liedjes per artiest.
2. Geef de titels van de films die een soundtrack hebben die in hetzelfde jaar werd uitgebracht als het jaar waarin ook de film werd uitgebracht.
3. Geef de naam en het aantal soundtracks van de artiest die het meeste soundtracks heeft geschreven die gebruikt worden in films uit de databank van de verhuurketen.
4. Geef de namen van elk paar artiesten die liedjes hebben geleverd aan minstens twee gemeenschappelijke albums. Zorg ervoor dat binnen elk paar de namen alfabetisch gerangschikt zijn, en sorteer voorts de lijst van paren eerst op de eerste naam en dan op de tweede naam.

Opmerking: Het is niet toegelaten om gebruik te maken van MySQL-specifieke uitbreidingen die niet voldoen aan de SQL standaard. Dit geldt bijvoorbeeld voor het sleutelwoord `LIMIT`.

Opgave 2

Het bestand `doolhof.zip` bevat een directorystructuur, waarvan elke subdirectory één of meerdere bestanden kan bevatten. Bestandsnamen hebben een naamcomponent en een extensie die beide uit cijfers en letters zijn opgebouwd, en van elkaar worden gescheiden door een punt. Los van wat de extensie zou kunnen doen vermoeden (`.pdf`, `.doc`, ...) zijn alle bestanden platte tekst, die elk een aantal getallen bevatten die op afzonderlijke regels zijn geplaatst. Gevraagd wordt

1. Schrijf een `bash` shell script `som.sh` dat de som van de regels uit de bestanden die overeenkomen met een bepaalde beschrijving uitschrijft naar standaard uitvoer. Aan het script moeten vier argumenten worden doorgegeven

```
som.sh dir naam extensie regel
```

Het eerste argument (`dir`) bevat de wortel van de directorystructuur die moet doorzocht worden. Het tweede argument (`naam`) bevat een substring die moet voorkomen in de naamcomponent van het bestand om het in rekening te brengen. Het derde argument (`extensie`) bevat de extensie van de bestanden die in rekening moeten gebracht worden. Het vierde argument (`regel`) bevat het regelnummer van de term die in de som moet worden opgenomen. Bestanden die minder regels hebben dan wat door het regelnummer wordt aangegeven moeten niet in rekening worden gebracht. Zo moet het commando

```
som.sh doolfhof abc pdf 4
```

de som uitschrijven van elke vierde regel van de bestanden onder de directory `doolhof` met een naam die de substring `abc` bevat en die de extensie `.xml` hebben.

2. Schrijf een `bash` shell script `product.sh` dat het product van een aantal sommen van regels uit de bestanden die overeenkomen met een bepaalde beschrijving uitschrijft naar standaard uitvoer. Aan het script wordt als eerste argument een ZIP-bestand doorgegeven met de bestanden die moeten doorzocht worden. Voorts kunnen aan het script ook nog een willekeurig aantal (`naam extensie regel`)-tripletten worden doorgegeven, die dezelfde betekenis hebben als bij het shell script `som.sh`. Indien er geen tripletten worden meegegeven, dan wordt de waarde 1 teruggegeven. Zo moet het commando

```
product.sh doolfhof.zip abc pdf 4 def js 1
```

het product uitschrijven van

- de som van de getallen op elke vierde regel van de bestanden met substring `abc` in de naam en extensie `.pdf`
- de som van de getallen op elke eerste regel van de bestanden met substring `def` in de naam en extensie `.js`

Dit shell script moet natuurlijk gebruik maken van het shell script `som.sh`. Het script moet bovendien zelf het ZIP-bestand uitpakken alvorens de berekeningen uit te voeren, en zorgen dat er daarvan na het uitvoeren van het script geen sporen meer achterblijven.

De scripts `som.sh` en `product.sh` moeten de nodige controles uitvoeren op de aanwezigheid en de geldigheid van de argumenten die ze verwachten.

Ter info: De uitkomst van de voorbeeldsom bij de beschrijving van `som.sh` is 27921, en de uitkomst van het voorbeeldproduct bij de beschrijving van `product.sh` is 535189728.

Opgave 3

Je neemt het jezelf kwalijk dat je steeds de verjaardagen van je beste vrienden blijft vergeten. Je wil daar iets aan doen, en beslist als goeie informaticus om je vrienden vanaf nu automatisch een mailtje te laten sturen op hun verjaardag. In je home directory plaats je een bestand `.verjaardag.data` met daarin de gegevens van je vrienden. Elke regel bevat de voornaam, de familienaam, de geboortedatum (formaat DD/MM/JJJJ) en het emailadres van een vriend, en het feit of die vriend al dan niet moet verwittigd worden van verjaardagen (zie verder). De verschillende informatievelden worden daarbij gescheiden door een komma. Commentaarregels starten met een hekje (`#`). Een voorbeeld van zo een bestand zou dus kunnen zijn

```
# lijst van verjaardagen
#
# formaat:
# voornaam,familienaam,geboortedatum (DD/MM/JJJJ),emailadres,verwittigen (J|N)
Donald,Knuth,10/01/1938,Donald.Knuth@UGent.be,J
Linus,Torvalds,28/12/1969,Linus.Torvalds@UGent.be,N
Bill,Gates,28/10/1955,Bill.Gates@UGent.be,J
Steve,Jobs,24/02/1955,Steve.Jobs@UGent.be,J
```

Gevraagd wordt

1. Je hebt een standaardtekst opgemaakt, die je gepersonaliseerd wil doorsturen naar elke vriend op zijn of haar verjaardag. De standaardtekst zit opgeslagen in een tekstbestand in je home directory onder de naam `.verjaardag.mail`, en bevat de volgende inhoud

```
Beste __voornaam__,

Van harte gefeliciteerd met je __leeftijd__e verjaardag.

Met vriendelijke groeten,

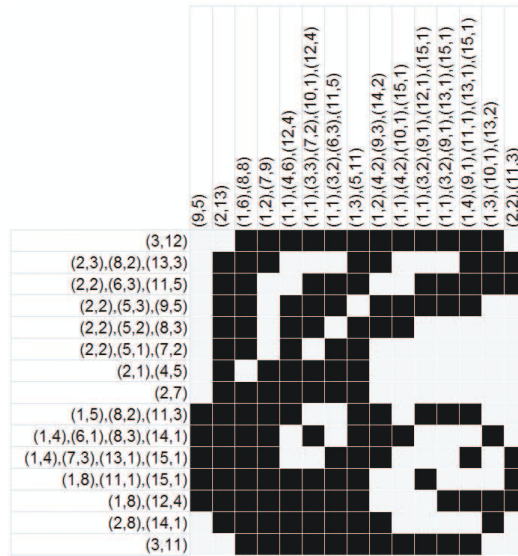
Ikke ;)
```

Schrijf een `bash` shell script `verjaardag_feliciteer.sh` dat een gepersonaliseerde email verstuurt naar alle vrienden uit het bestand `.verjaardag.data` die vandaag jarig zijn. Maak hierbij gebruik van `sed` voor het personaliseren van de standaardtekst uit `.verjaardag.mail`, waarbij de voornaam van de jarige moet worden ingevuld op de plaats van `__voornaam__`, en zijn of haar leeftijd op de plaats van `__leeftijd__`.

2. Door de stiptheid waarmee je je vrienden feliciteert met hun verjaardag dwing je al gauw heel wat respect af. Je beslist om een er nog een schepje bovenop te doen, en je vrienden aan de verjaardagen van anderen te helpen herinneren. Schrijf hiervoor een `awk` script `verjaardag_herinner.awk` dat een herinneringsemail stuurt naar alle vrienden die hebben aangegeven dat ze aan verjaardagen herinnerd willen worden. Dit wordt aangegeven door een J in de laatste kolom van het bestand `.verjaardag.data`. Iedere geadresseerde ontvangt dus een email met een lijst van namen (voor- en familienaam) en leeftijden van alle personen die vandaag jarig zijn, exclusief de naam en leeftijd van de geadresseerde zelf. Zorg ervoor dat er geen lege lijsten worden doorgemaild.
3. Geef aan hoe je ervoor kan zorgen dat op een Unix besturingssysteem de twee voorgaande scripts `verjaardag_feliciteer.sh` en `verjaardag_herinner.awk` elke dag automatisch uitgevoerd worden om zes uur 's morgens. Op die manier wek je bovendien ook nog eens de indruk dat je iedere dag reeds vroeg in de weer bent.

Opgave 4

Een *bitmap* is een bestandsformaat dat gebruikt wordt voor het opslaan van digitale zwartwitafbeeldingen. De term zelf is afgeleid van de datastructuur die voor de gegevensopslag gebruikt wordt, namelijk een tweedimensionale array (ook een rooster of een map genaamd) van bits. Een alternatieve manier om digitale zwartwitafbeeldingen te coderen bestaat erin om voor elke rij of voor elke kolom de beginpositie en de lengte van elke reeks opeenvolgende 1-bits op te geven. Al deze verschillende voorstellingsvormen worden samen geïllustreerd in onderstaande figuur. Daarin is bijvoorbeeld te zien dat de eerste rij bestaat uit een reeks van twaalf opeenvolgende 1-bits die start op positie drie, en dat de derde kolom bestaat uit een reeks van zes opeenvolgende 1-bits die start op de eerste positie en een reeks van acht opeenvolgende 1-bits die start op de achtste positie.



Gevraagd wordt om een subprocedure `BepaalBitmap` te schrijven in VBA, die de bitmap en de kolomrepresentatie bepaalt voor een gegeven rijrepresentatie van een digitale afbeelding. De subprocedure moet de rijrepresentatie uitlezen uit de eerste kolom van het actieve werkblad in Excel, en moet daarbij zelf bepalen uit hoeveel rijen en kolommen de bitmap bestaat. De bits van een rij worden beschreven door een reeks koppels, waarbij elk koppel een reeks van opeenvolgende 1-bits voorstelt. Het eerste element van het koppel geeft de beginpositie van de reeks aan, en het tweede element de lengte van de reeks. Elk koppel wordt ingesloten tussen ronde haakjes, en koppels worden van elkaar gescheiden door komma's. De beschrijving van de eerste rij van de digitale afbeelding is terug te vinden op de tweede rij van het Excel werkblad, en de rijrepresentatie eindigt op de eerste lege cel in de eerste kolom van het werkblad (eerste rij niet in acht genomen). Je mag ervan uitgaan dat in elke kolom minstens één 1-bit voorkomt.

Bij de weergave van de bitmap van een digitale afbeelding moet je ervoor zorgen dat cellen van het Excel werkblad die corresponderen met 1-bits een zwarte achtergrondkleur krijgen, en dat cellen die corresponderen met 0-bits een lichtgrijze achtergrondkleur krijgen. Verder moet de subprocedure de breedte van alle kolommen die voor de bitmapvoorstelling gebruikt worden instellen op een waarde van 1,8. De kolomrepresentatie van de digitale afbeelding moet worden uitgeschreven op de eerste rij van het Excel werkblad. Voor de beschrijving ervan moet een analoge notatie gebruikt worden als voor de rijrepresentatie. De subprocedure moet er ook voor zorgen dat de tekst van de kolomrepresentatie verticaal wordt weergegeven, en dat die verticaal onderaan en horizontaal gecentreerd wordt uitgelijnd zoals in bovenstaande figuur. Zorg ervoor dat het scherm pas wordt bijgewerkt nadat de bitmap en de kolomrepresentatie volledig zijn weergegeven, zodat het scherm tussendoor niet voortdurend staat te flikkeren.

Het Excel document `bitmap.xls` bevat enkele voorbeelden van digitale afbeeldingen die worden weergegeven in rijrepresentatie. Je kan dit document gebruiken om je subprocedure te testen.