

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

05.567 19 01 19 EX

Enganxeu en aquest espai una etiqueta identificativa
amb el vostre codi personal
Examen

Aquest enunciat correspon també a les assignatures següents:

- 11.503 - Bases de dades

Fitxa tècnica de l'examen

- Comprova que el codi i el nom de l'assignatura corresponen a l'assignatura matriculada.
- Només has d'enganxar una etiqueta d'estudiant a l'espai corresponent d'aquest full.
- No es poden adjuntar fulls addicionals, ni realitzar l'examen en llapis o retolador gruixut.
- Temps total: **2 hores** Valor de cada pregunta: **indicat a cada exercici**
- En cas que els estudiants puguin consultar algun material durant l'examen, quins són?
no es pot fer servir material En cas de poder fer servir calculadora, de quin tipus? **CAP**
- Si hi hagi preguntes tipus test: Descompten les respostes errònies? **NO** Quant?
- Indicacions específiques per a la realització d'aquest examen:

Enunciats

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

EXERCICI 1 (3 punts)

Volem disposar d'una base de dades per enregistrar informació sobre pel·lícules, actors i directors. A continuació, es descriu cadascuna de les taules.

PERSON (PERSONA)

La taula *PERSON* conté les dades generals sobre les persones que apareixen a la BD. En concret, per a cada persona s'enregistra un nombre identificador (*id*) que és clau primària, el nom (*name*), la data de naixement (*birth*), la data de defunció (*death*), el gènere (*gender*), la nacionalitat (*nationality*) i l'adreça (*address*). Només els atributs data de defunció (*death*) i adreça (*address*) poden prendre valors nuls.

DIRECTOR (DIRECTOR)

La taula conté informació sobre els directors. En concret, per cada director, es guarda l'identificador (*id*) que és clau primària i clau forana de *PERSON*. L'any de debut (*debut_year*) i el nombre de premis que ha rebut durant la seva carrera (*num_awards*). Només l'atribut any de debut (*debut_year*) pot ser nul.

ACTOR (ACTOR)

Guarda informació sobre els actors. Per cada actor es guarda l'identificador (*id*) que és clau primària i clau forana de *PERSON* i l'any de debut (*debut_year*), que pot ser nul.

MOVIE (PEL·LÍCULA)

Informació sobre les pel·lícules. Per a cada pel·lícula s'enregistra el seu identificador (*id*) que és clau primària, el títol (*title*), el gènere (*genre*), l'any d'estrena (*year*), el país de producció (*country*), la duració en minuts (*duration*), el pressupost en euros (*budget*), el ingressos generats en euros (*incomes*), una referència a la pel·lícula anterior si la pel·lícula pertany a una saga (*id_movie_prev*) i l'identificador del director (*id_director*) que és clau forana a *DIRECTOR*.

Només l'atribut que fa referència a la pel·lícula anterior pot prendre valors nuls. L'atribut *duration* ha de ser positiu estricte i els atributs *incomes* i *budget* han de ser iguals o majors a zero.

L'atribut gènere només pot prendre els valors {'Action', 'Adventure', 'Comedy', 'Crime', 'Drama', 'Historical', 'Horror', 'Musical', 'Science-fiction', 'War', 'Western'}.

PARTICIPATION (PARTICIPACIÓ)

Informació sobre les participacions dels actors a les pel·lícules. S'emmagatzemen el valors {*id_actor*, *id_movie*} que són clau primària, on {*id_actor*} és clau forana de *ACTOR* i {*id_movie*} és

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

clau forana de *MOVIE*. El nom del personatge interpretat (*character*) i el rol (*role*) que només pot prendre els valors {'Leading', 'Supporting', 'Narrator', 'Cameo', 'Unclassified'}.

Cap atribut pot prendre valors nuls. Si s'esborra una pel·lícula de la BD, caldrà esborrar totes les participacions d'actors a la pel·lícula.

(1 punt) ES DEMANA, utilitzant expressions d'àlgebra relacional:

1) Obtenir les persones que són, alhora, actors i directors, però només si són de nacionalitat italiana i estan vius. Concretament volem l'identificador, el nom i la data de naixement d'aquestes persones.

$R1 := \text{ACTOR}[id]$

$R2 := \text{DIRECTOR}[id]$

$R3 := R1 \cap R2$

$R4 := R3 * \text{PERSON}$

$R5 := R4(\text{nationality} = \text{'italian'} \wedge \text{death is null})$

$R6 := R5[id, \text{name}, \text{birth}]$

(1 punt) ES DEMANA, utilitzant expressions d'SQL:

2) Trobar el nombre total de pel·lícules que de cada gènere s'han fet durant l'any 2018. En concret, es vol el gènere de la pel·lícula, el nombre total de pel·lícules que s'han fet de cada gènere, la duració mitjana de les pel·lícules de cada gènere, i el total d'ingressos que han recaptat les pel·lícules de cada gènere. El resultat es vol ordenat de forma ascendent pel gènere de la pel·lícula.

```
SELECT genre, count(*), avg(duration), sum(incomes)
FROM MOVIE
WHERE year=2018
GROUP BY genre
ORDER BY genre;
```

(1 punt) ES DEMANA, donada la seqüència següent d'expressions d'àlgebra relacional:

$R1 := \text{MOVIE}(\text{genre} = \text{'Science-fiction'})$

$R2 := R1[id=id_movie] \text{PARTICIPATION}$

$R3 := R2[id_actor]$

$R4 := \text{MOVIE}(\text{genre} = \text{'War'})$

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

R5:= R4[id=id_movie]PARTICIPATION

R6:= R5[id_actor]

R7:= R3-R6

R8:= R7[id_actor=id]ACTOR

R9:= R8*PERSON

R10:= R9[id, name, debut_year]

3.1) **(0.5 punts)** Explicar en llenguatge natural què fa la seqüència d'operacions prèvies (en definitiva, es tracta de donar l'enunciat de la consulta que es planteja).

Obtenir els actors que han participat en alguna pel·lícula de ciència ficció (*Science-fiction*), però mai en pel·lícules bèl·liques (*War*). Concretament volem l'identificador, el nom, i l'any de debut d'aquests actors.

3.2) **(0.5 punts)** Proposar una consulta equivalent en llenguatge SQL.

```
SELECT pe.id, pe.name, a.debut_year
FROM PERSON p JOIN ACTOR a ON p.id=a.id
      JOIN PARTICIPATION pa ON a.id=pa.id_actor
      JOIN MOVIE m ON m.id ON pa.id_movie=m.id
WHERE m.genre='Science-fiction' AND pa.id_actor NOT IN
      (SELECT pa1.id_actor
      FROM PARTICIPATION pa1 JOIN MOVIE m1 ON pa1.id_movie=m1.id
      WHERE m1.genre='War');
```

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

EXERCICI 2 (2.5 punts)

Sobre la base de dades de pel·lícules del primer exercici, hem afegit la columna nombre total de pel·lícules amb beneficis del director (*total_profits_movies*) a la taula *DIRECTOR*. Aquesta columna indica el nombre total de pel·lícules dirigides pel director tals que els seus ingressos (*incomes*) han estat superiors a les despeses (*budget*).

- Volem mantenir el càlcul del nombre total de pel·lícules amb beneficis (*total_profits_movies*) del director actualitzat. En concret, volem que aquesta columna sempre reflecteixi el nombre total de les pel·lícules on ha estat director i han tingut beneficis. Podem suposar que els usuaris o programes mai no actualitzaran directament la columna *total_profits_movies* de la taula *DIRECTOR*, i que en el moment d'inserir un nou director, el valor de la columna *total_profits_movies* serà zero.

Digueu quins elements d'SQL faríeu servir per implementar en PostgreSQL les regles anteriors.

En cas que trieu disparadors, cal indicar l'operació (*INSERT*, *DELETE*, *UPDATE*) que activaria l'execució de cada disparador i les taules implicades. En cas d'operacions d'*UPDATE* cal especificar les columnes rellevants.

Adicionalment, indiqueu el tipus de disparador (*FOR EACH ROW/FOR EACH STATEMENT* i si ha de ser *BEFORE/AFTER*), així com una breu descripció de les comprovacions que caldria realitzar.

Per a la regla de negoci cal definir disparadors sobre les taules i esdeveniments següents:

	<i>INSERT</i>	<i>UPDATE</i>	<i>DELETE</i>
<i>MOVIE</i>	<p>S'afegirà una nova pel·lícula (el valor de la columna <i>total_profits_movies</i> s'incrementa en una unitat si els <i>incomes</i> són major que el <i>budget</i> pel director relacionat).</p> <p><i>FOR EACH ROW</i> (cal accedir a valors de la fila a inserir mitjançant <i>NEW</i>).</p> <p><i>AFTER</i> així garantim que no s'ha violat cap regla de negoci.</p>	<p>En cas d'<i>UPDATE</i> hem de tenir en compte la modificació de les columnes següents:</p> <p><i>id_director</i>: Si canviem aquesta columna, hem de tornar a calcular el valor de la columna <i>total_profits_movies</i> per a l'antic director i per al nou.</p> <p><i>incomes</i> o <i>budget</i>: Si canviem una d'aquestes columnes, hem de tornar a calcular el valor de la columna <i>total_profits_movies</i> si la diferència entre <i>incomes</i> i <i>budget</i> antic i nou varia (abans donava beneficis i ara no, o a la inversa). Segons el cas, sumarem o restarem una unitat a <i>total_profits_movies</i>.</p>	<p>En cas d'esborrar una fila de la taula, s'haurà de tornar a calcular el valor de la columna <i>total_profits_movies</i> del director relacionat, però només si la pel·lícula tenia beneficis.</p> <p><i>FOR EACH ROW</i> (cal accedir a valors de la fila a esborrar mitjançant <i>OLD</i> per mirar si tenia beneficis la pel·lícula i restar una unitat a <i>total_profits_movies</i> del director relacionat, si fos el cas).</p> <p><i>AFTER</i> així garantim que no s'ha violat cap regla de negoci.</p>

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

		<p><i>FOR EACH ROW</i> ja que les sentències d'<i>UPDATE</i> poden actuar sobre un conjunt de files (accedirem als valors de la fila mitjançant <i>NEW</i> i <i>OLD</i>).</p> <p><i>AFTER</i> així garantim que no s'ha violat cap regla de negoci.</p>	
--	--	---	--

EXERCICI 3 (2.5 punts)

Sigui un SGBD **sense cap mecanisme de control de concurrència**, i suposem que es produeix l'horari que mostrem a continuació (on R=lectura, RU=lectura amb intenció d'actualització, W=escriptura; les accions s'han numerat per facilitar la seva referència):

Acc#	T1	T2	T3
10	R(F)		
20		RU(B)	
30		W(B)	
40			RU(C)
50			W(C)
60			R(D)
70	RU(A)		
80	W(A)		
90		RU(D)	
100		W(D)	
110			R(E)
120		R(B)	
130		R(D)	
140	R(B)		
150			R(D)
160			COMMIT
170	COMMIT		

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

180		ABORT	
-----	--	-------	--

1) (1 punt) Indiqueu la interferència o interferències que s'hi produeixen: tipus, grànuls i transaccions implicades, així com el nivell mínim d'aïllament d'SQL amb el que haurien de treballar les transaccions per tal d'evitar simultàniament totes les interferències trobades.

Acc#	T1	T2	T3
10	R(F)		
20		RU(B)	
30		W(B)	
40			RU(C)
50			W(C)
60			RD
70	RU(A)		
80	W(A)		
90		RU(D)	
100		W(D)	
110			R(E)
120		R(B)	
130		R(D)	
140	R(B)		
150			RD
160			COMMIT
170	COMMIT		
180		ABORT	

Lectura no repetible entre T2 i T3 pel grànul D.

Lectura no confirmada entre T1 i T2 pel grànul B.

El nivell d'aïllament mínim d'SQL per evitar la interferència de lectura no repetible és REPEATABLE READ i per evitar la lectura no confirmada és READ COMMITTED.

Per tant, per evitar totes dues interferències simultàniament ens hem de quedar amb el nivell més

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

restrictiu, és a dir, REPEATABLE READ.

2) **(0.5 punts)** L'horari proposat, és recuperable? Quins horaris serials donarien els mateixos resultats que l'horari proposat?

L'horari proposat no és recuperable, ja que, per exemple, T3 a l'acció 150 llegeix el grànul D, prèviament escrit per T2 a l'acció 100, i T3 acaba la seva execució confirmant abans que T1.

No hi ha cap horari serial equivalent ja que l'horari proposat no és serialable degut a que hi ha dues interferències.

3) **(1 punt)** Apliqueu (indicant com quedaria l'horari) el protocol de reserves en dues fases (PR2F) estricte a l'horari anterior. Quins horaris serials donarien els mateixos resultats que l'horari resultant?

Acc#	T1	T2	T3
10	LOCK(F,S)		
20	R(F)		
30		LOCK(B,X)	
40		RU(B)	
50		W(B)	
60			LOCK(C,X)
70			RU(C)
80			W(C)
90			LOCK(D,S)
100			R(D)
110	LOCK(A,X)		
120	RU(A)		
130	W(A)		
140		LOCK(D,X)	
150			LOCK(E,S)
160			R(E)
170	LOCK(B,S)		
180			R(D)
190			COMMIT UNLOCK(C) UNLOCK(D)

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

			UNLOCK(E)
200		RU(D)	
210		W(D)	
220		R(B)	
230		R(D)	
240		ABORT UNLOCK(B) UNLOCK(D)	
250	R(B)		
260	COMMIT UNLOCK(F) UNLOCK(A) UNLOCK(B)		

T2 bloqueja la seva execució fins poder aconseguir la reserva sobre el grànul D i T1 bloqueja la seva execució fins poder aconseguir la reserva sobre el grànul B. L'estona que cada transacció està bloquejada es marca en gris a la solució.

L'horari que obtenim no té cap interferència i és equivalent a l'horari T3; T1. No s'inclou T2 perquè ha cancel·lat la seva execució, i en conseqüència, oficialment mai ha existit.

L'horari també és recuperable. El protocol de reserves en dues fases estricte garanteix horaris serialitzables i recuperables.

EXERCICI 4 (2 punts)

3) (1 punt) Explica què fan les següents línies de codi usat per resoldre el tercer lliurament de la pràctica. **Extensió màxima mitja pàgina.**

```
...
fileContents = fileUtilities.readFileFromClasspath("exercisel.data");
...
List<String> row : fileContents)
...
String updateSql = "UPDATE ACTOR " + "SET    num_movies= num_movies + 1
"+"WHERE id=?" ;
updatePreparedStatement = conn.prepareStatement(updateSql);
...
updatePreparedStatement.clearParameters();
setUpdateParameters(updatePreparedStatement, row);
updatePreparedStatement.executeUpdate();
System.out.println("Executing Update Director Number of films");
...
```

Per una línia del fitxer `exercisel.data` es realitza un increment d'una unitat del nombre de

Examen 2018/19-1

Assignatura	Codi	Data	Hora inici
Ús de bases de dades	05.567	19/01/2019	09:00

pel·lícules de l'actor amb qui coincideix *id* a la taula *ACTOR* de la base de dades.

2) **(1 punt)** Què és una base de dades geogràfica i quina és la seva utilitat principal? En el context d'aquest tipus de base de dades, què és la representació vectorial? **Extensió màxima mitja pàgina.**

Una base de dades geogràfica és un tipus de base de dades que permet representar tant informació geogràfica com informació alfanumèrica (aquest tipus d'informació és la que clàssicament es guarda en una base de dades com les que hem estudiat a l'assignatura). Per informació geogràfica s'entén la representació digital d'entitats, objectes o fenòmens que s'esdevenen a la superfície de la Terra o a prop d'aquesta.

Per a emmagatzemar informació geogràfica, la base de dades ha de disposar de tipus de dades que permetin representar-la. Una forma de representar la informació geogràfica és mitjançant gràfics vectorials, que permeten la representació d'objectes puntuals o de dimensió zero (objectes que es troben en un punt concret com, per exemple, un arbre), objectes lineals o de dimensió 1 (objectes que es distribueixen de forma lineal pel territori com, per exemple, una carretera), i objectes d'àrea o de dimensió 2 (objectes que ocupen l'àrea d'un territori com, per exemple, una casa).