

NOM ALUMNE:

PAC 2: RECOPILACIÓ PREPARATÒRIA PER A L'EXAMEN

Presentació

Aquesta prova consta de 5 exercicis enfocats a determinar el grau de comprensió dels mòduls didàctics després del seu estudi:

- Model conceptual
- Model lògic
- Normalització
- Arbres B+
- Preguntes teòriques

Objectius i competències

Els objectius fonamentals d'aquesta activitat són:

- Tenir una visió global de les diferents fases del procés de disseny de bases de dades.
- Preparar la prova final de l'assignatura.

La realització d'aquesta PAC ha de permetre demostrar que s'han assolit les competències bàsiques de l'assignatura.



NOM ALUMNE:

Descripció de la PAC

EXERCICI 1: model conceptual

Ens disposem a dissenyar part d'una base de dades que permeti emmagatzemar informació referent a exposicions d'art.

Les exposicions tenen un títol, se celebren a una ciutat determinada i s'identifiquen per la data d'inici i de fi. A cada exposició hi ha diverses obres i una mateixa obra pot estar en diferents exposicions al llarg del temps. Cada obra a cada exposició pot tenir un preu de venda (no té perquè ser el mateix en tots els casos).

Les obres d'art s'identifiquen per un número de registre, tenen un nom i també es vol guardar l'any de la seva finalització. Només es volen guardar les obres que s'exposen. Cada obra és d'un artista i, cadascun d'ells pot haver fet més d'una obra.

Dels artistes es vol guardar el nom, la data de naixement, la de la seva defunció (si és el cas) i un codi que els identifica, a més de l'artista o artistes que li serveixen d'inspiració (si n'hi ha algun). Un artista pot ser pintor, escultor i/o d'altres disciplines.

Dels artistes escultors es vol conèixer la seva forma de treballar. Per això es vol guardar quins materials utilitzen i amb quines eines els treballen. Un escultor pot treballar un material amb diferents eines. Un escultor pot treballar amb una mateixa eina diversos materials. Una eina la poden utilitzar diversos escultors per a treballar el mateix material.

Les eines tenen un nom i s'identifiquen per un codi. Els materials tenen un nom, poden ser o no perillosos i s'identifiquen per un codi numèric.

Dels artistes pintors es vol guardar a quines ciutats tenen estudis on treballar.

Les ciutats s'identifiquen pel seu nom relatiu al país on es troben, és a dir, que pot haver-hi diverses poblacions amb el mateix nom en països diferents. Dels països es vol guardar el seu nom, que els identifica, i el percentatge d'impostos que graven les obres d'art.

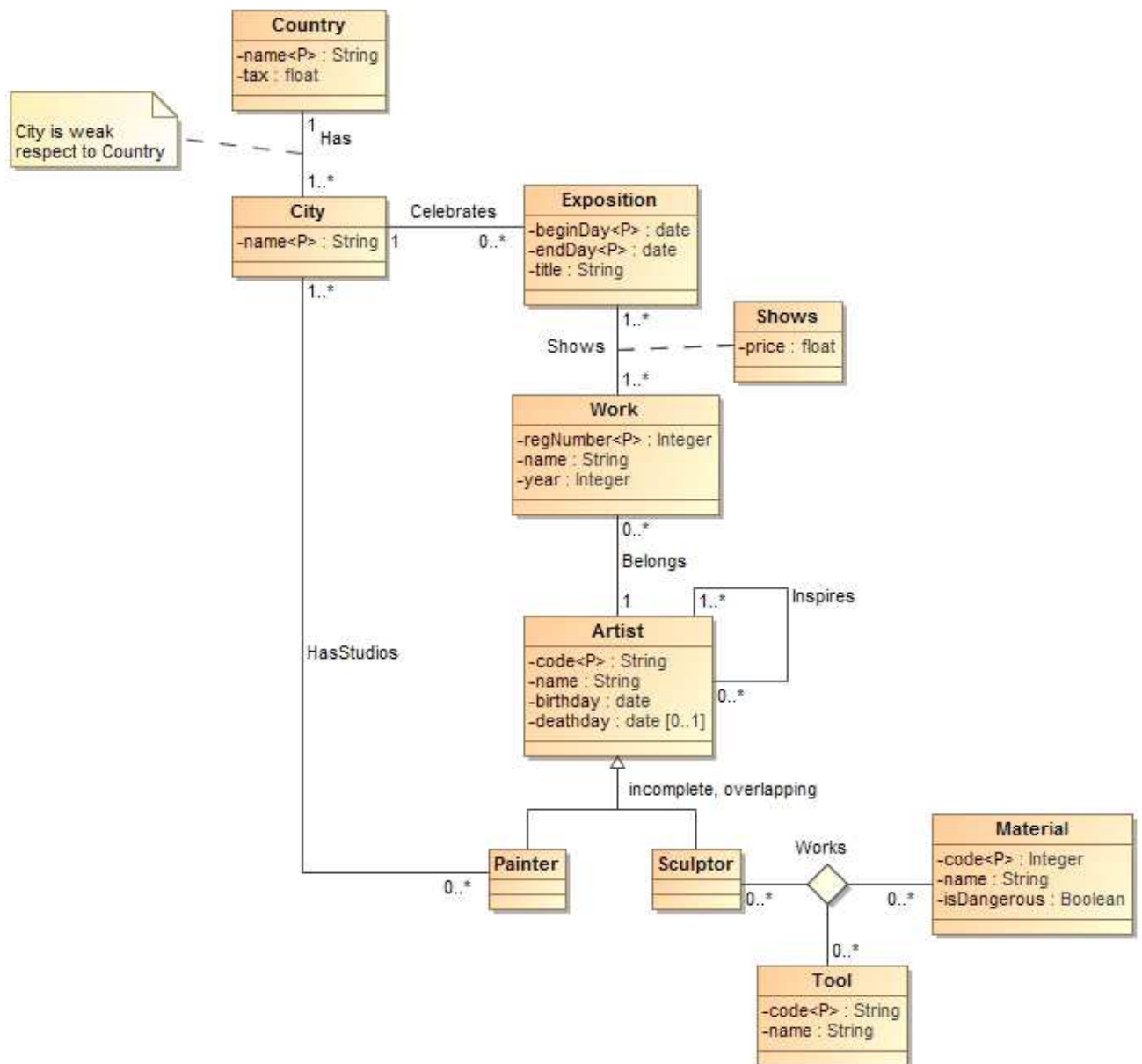
Es demana: realitzar el disseny conceptual, mitjançant un diagrama de classes UML, que reculli tota la semàntica de l'enunciat. Cal indicar els atributs dels tipus d'entitats, les cardinalitats dels tipus de relacions, així com els requisits que no han

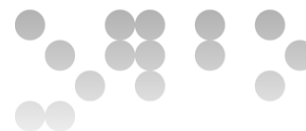


NOM ALUMNE:

quedat reflectits en l'esquema proposat. Si s'ha realitzat alguna suposició semàntica addicional també s'ha d'indicar.

SOLUCIÓ:

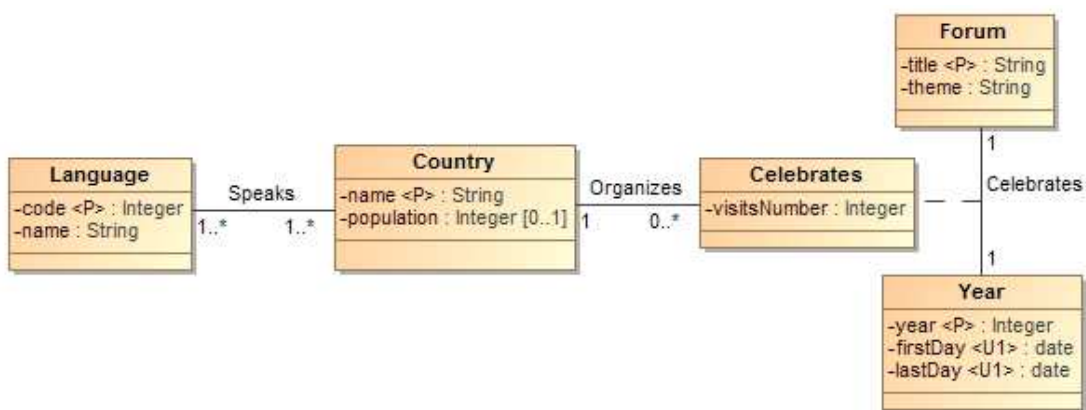




NOM ALUMNE:

EXERCICI 2: model lògic

Donat el model conceptual següent:



Es demana: obtenir el model lògic relacional corresponent, tot indicant si, a causa de la transformació, apareixen atributs que poden prendre valor NULL.

SOLUCIÓ:

Language (code, **name**)

Country (name, population)

Speaks (country, language)

{country} is foreign key to Country

{language} is foreign key to Language

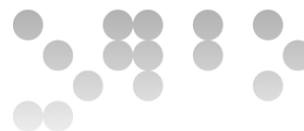
Forum (title, **theme**, **year**, **firstDay**, **lastDay**, **visitsNumber**, **country**)

{year} is alternative key

{firstDay, lastDay} is alternative key

{country} is foreign key to Country

Nota: els atributs que no són clau i no poden ser null estan en negreta.



NOM ALUMNE:

EXERCICI 3: normalització

Partim d'una base de dades que emmagatzema dades sobre les festes que organitza cada any l'ajuntament d'una ciutat.

L'ajuntament organitza diverses festes cada any. Cada festa s'identifica per un codi alfanumèric més el seu any de celebració. També volem guardar el recinte on se celebra la festa. Els recintes s'identifiquen per un codi numèric i en volem guardar el nom i la capacitat màxima (l'aforament). Les emissores de ràdio FM de la ciutat patrocinen les festes. Les emissores s'identifiquen per la seva freqüència i en volem guardar també el nom. Una emissora pot patrocinar una o més festes, en un mateix any o diversos anys, i cada festa concreta està patrocinada per com a mínim una emissora.

Per acabar, també volem guardar el nombre total de persones que finalment ha assistit a cada festa, així com el nombre de persones que han assistit convidats per cada emissora patrocinadora.

Com a part del model relacional corresponent, tenim la relació següent (atributs que formen la clau primària subratllats):

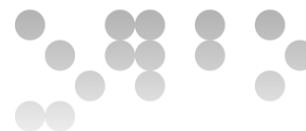
Events (partyCode, partyYear, **stationName**, venueCode, **venueName**, **venueCapacity**, stationFrequency, totalAttendees, stationAttendees)

Es demana:

- Indiqueu si la relació està en FNBC, tot raonant la resposta.
- Si no està en FNBC, descriuiu quin problema es produeix i doneu un model relacional equivalent normalitzat a FNBC. Assenyaleu les claus primàries i foranes de les relacions obtingudes. Indiqueu si hi ha alguna restricció que no es pugui reflectir al model relacional.

SOLUCIÓ:

- La relació presentada no està en FNBC perquè no totes les dades depenen de tot el conjunt de la clau primària. Concretament, el nom de l'emissora de ràdio (*stationName*) només depèn de l'emissora (*stationFrequency*); el codi i el nom del recinte (*venueCode* i *venueName*) només depenen de la festa (*partyCode* + *partyYear*); i el nombre total de persones que han assistit a la

**NOM ALUMNE:**

festa (*totalAttendees*) depèn també només de la festa (*partyCode* + *partyYear*). El nombre d'assistents convidats per l'emissora de ràdio (*stationAttendees*) sí que depèn de tota la clau primària, ja que aquest nombre vindrà donat tant per la festa en sí mateixa com per cada emissora patrocinadora.

- b) El problema que es produeix és que hi ha redundància de dades. D'una banda s'estarà repetint el nom de l'estació per cadascuna de les festes que hagi patrocinat; d'una altra banda s'estarà repetint el codi, el nom i la capacitat del recinte per cada festa que s'hi hagi organitzat i emissora que l'hagi patrocinat; i d'una altra banda també s'estarà repetint el nombre total d'assistents per cada festa i emissora patrocinadora.

Per solucionar aquest problema, podem oferir el següent model relacional normalitzat equivalent:

RadioStation (frequency, name)

Venue (code, name, capacity)

Party (code, year, venue, totalAttendees)

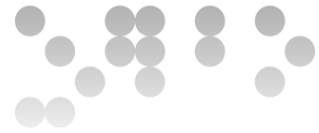
{venue} is foreign key of Venue (code)

Events (partyCode, partyYear, stationFrequency, stationAttendees)

{partyCode, partyYear} is foreign key of Party (code, year)

{stationFrequency} is foreign key of RadioStation (frequency)

Hi ha un parell de restriccions que no es poden reflectir al model relacional. D'una banda, el total d'assistents a una festa que han estat convidats per emissores de ràdio hauria de ser igual o menor al nombre d'assistents total d'aquella festa (a no ser que desconeguem aquesta darrera xifra). I, al mateix temps, el nombre d'assistents total d'aquella festa hauria de ser menor o igual a l'aforament del recinte a on s'ha celebrat (en aquest cas sí que l'hauríem de conèixer perquè no pot ser nul).

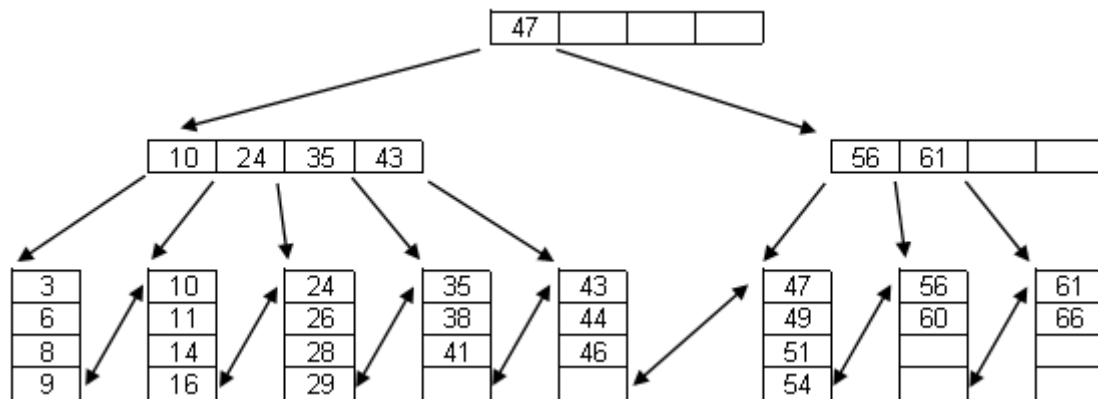


NOM ALUMNE:

EXERCICI 4: arbres B+

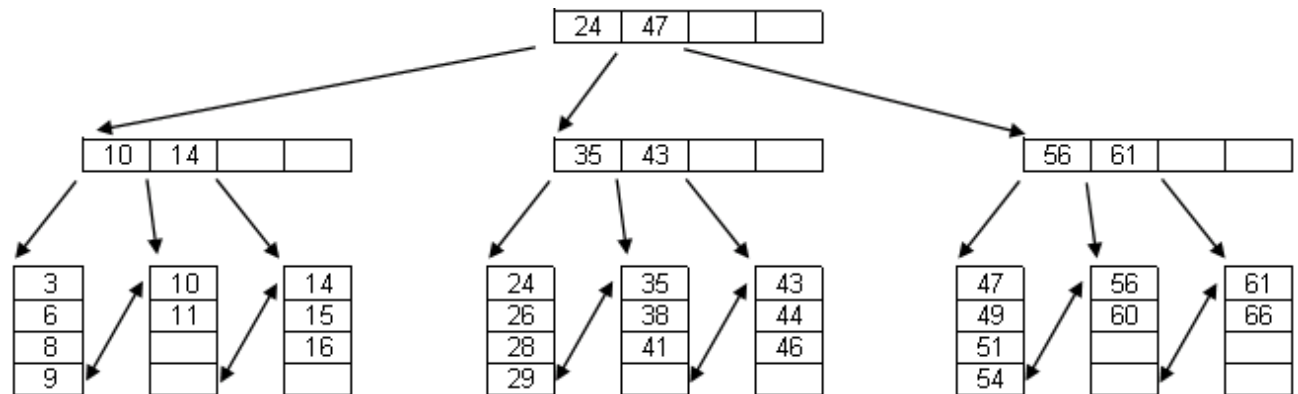
APARTAT A

A la següent figura tenim representat un arbre B+ d'ordre 2:



Mostreu com quedaria l'arbre després d'inserir el valor 15.

SOLUCIÓ:



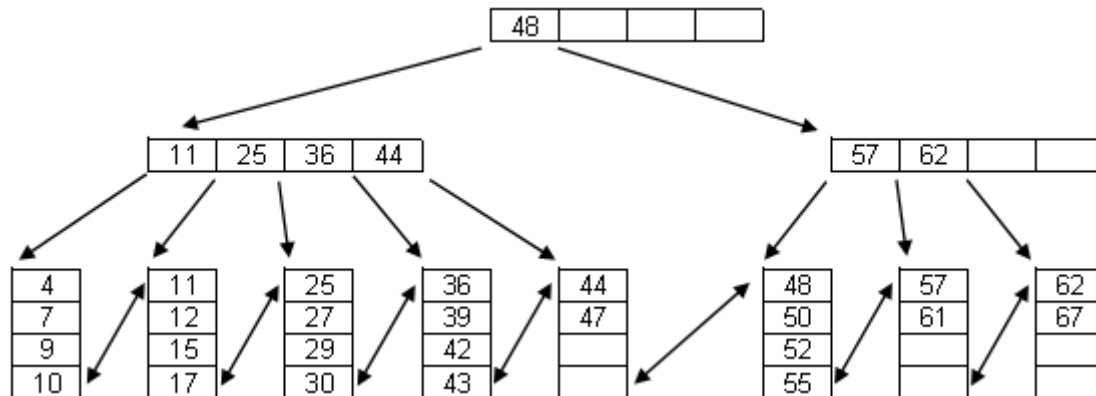
Com que la fulla on s'hauria d'anotar el 15 està plena, s'haurà de crear una de nova i dividir el seu contingut deixant els dos primers valors a la fulla de l'esquerra. Al crear-se una nova fulla s'haurà d'anotar al node pare el nou valor, però a l'estar ple, també s'haurà de dividir (els dos primers valors quedaran a l'esquerra) i es pujarà el valor de la nova branca al node arrel



NOM ALUMNE:

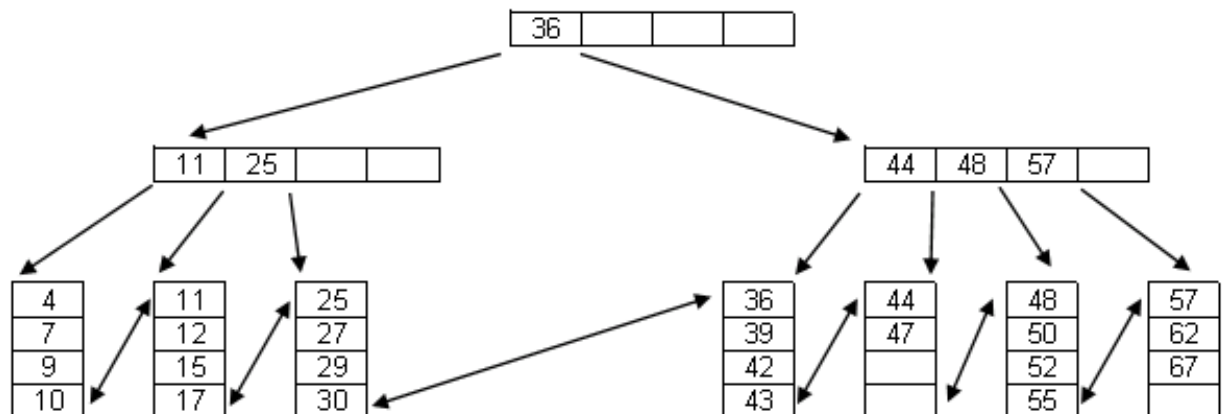
APARTAT B

A la següent figura tenim representat un arbre B+ d'ordre 2:

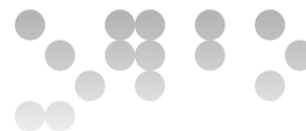


Mostreu com quedaria l'arbre després d'eliminar el valor 61.

SOLUCIÓ:



En eliminar-se el valor 61 queda la fulla amb un únic valor. Com que el germà de la dreta té dos valors es fusionaran. El pare però quedaria amb un únic valor, pel que també s'haurà de redistribuir l'arbre en el nivell superior.



NOM ALUMNE:

EXERCICI 5: preguntes teòriques

APARTAT A

Volem guardar informació sobre els empleats d'una empresa: a quin departament treballen i a quina oficina ho fan, així com algunes dades bàsiques sobre els empleats, les oficines i els departaments. Hem de tenir en compte que no tindrem dades històriques, sinó només el departament/oficina on treballa un empleat concret a l'actualitat i, a més, un empleat només pot treballar per un departament i en una oficina determinada al mateix temps. També haurem de tenir en compte que podem tenir guardats molts milers de treballadors. Tenim les següents taules:

```
CREATE TABLE Office (
    numOffice NUMBER(2,0) CONSTRAINT PK_Office PRIMARY KEY,
    capacity NUMBER(2,0) CONSTRAINT NN_officeCapacity NOT NULL
);

CREATE TABLE Employee (
    idEmployee NUMBER(6,0) CONSTRAINT PK_Employee PRIMARY KEY,
    name VARCHAR2(40 CHAR) CONSTRAINT NN_nameEmployee NOT NULL,
    surname VARCHAR2(40 CHAR) CONSTRAINT NN_surnameEmployee NOT NULL,
    birthday DATE
);

CREATE TABLE Department (
    name VARCHAR2(20 CHAR) CONSTRAINT PK_Department PRIMARY KEY,
    manager NUMBER(6,0) CONSTRAINT FK_Manager REFERENCES Employee (idEmployee)
);

CREATE TABLE EmployeeDepartment (
    idEmployee NUMBER(6,0) CONSTRAINT FK_Employee REFERENCES Employee
(idEmployee),
    department VARCHAR2(20 CHAR) CONSTRAINT FK_Department REFERENCES
Department (name),
    office NUMBER(2,0) CONSTRAINT FK_Office REFERENCES Office (numOffice),
    deskNumber NUMBER(2,0) CONSTRAINT NN_deskNumber NOT NULL
);
```

Com podeu veure, la taula *EmployeeDepartment* no té clau primària. Digueu la vostra opinió sobre cadascuna de les següents propostes de clau primària per *EmployeeDepartment*. Indiqueu quins avantatges/inconvenients hi veieu a cadascuna. Indiqueu també quina de les claus primàries seria l'adequada i si faria falta afegir alguna restricció addicional per assegurar tots els requeriments. En tot cas, justifiqueu les vostres respostes.



NOM ALUMNE:

- a) `CONSTRAINT PK_EmployeeDepartment PRIMARY KEY (idEmployee, department)`
- b) `CONSTRAINT PK_EmployeeDepartment PRIMARY KEY (idEmployee, office)`
- c) `CONSTRAINT PK_EmployeeDepartment PRIMARY KEY (idEmployee)`
- d) `CONSTRAINT PK_EmployeeDepartment PRIMARY KEY (idEmployee, department, office)`

SOLUCIÓ:

a) Aquesta clau primària només permetria una fila per empleat / departament i, per tant, una única oficina, fins aquí és correcte. Ara bé, d'una banda ens trobaríem amb el problema que es podria introduir un valor *NULL* al camp *office*, cosa que no s'hauria de permetre ja que en aquest cas no podríem saber a quina oficina treballa l'empleat. Per tant, s'hauria d'afegir una restricció *NOT NULL* al camp *office*. I d'altra banda seria possible que un mateix empleat existís a la taula en més d'un departament diferent (i per tant fins i tot a oficines diferents), cosa que no s'ha de permetre segons els requeriments.

b) Ens trobem en el mateix cas que a l'apartat a), només permetria una fila per empleat / oficina i, per tant, un únic departament. Però d'una banda permetria introduir un valor *NULL* al camp *department*, amb la qual cosa perdríem informació. Per tant, s'hauria d'afegir una restricció *NOT NULL* al camp *department*. I d'altra banda seria possible que un empleat existís a la taula en més d'una oficina diferent (i per tant fins i tot a departaments diferents), cosa que tampoc s'ha de permetre segons els requeriments.

c) Aquesta clau primària seria vàlida ja que només permet tenir un empleat determinat una sola vegada a la taula, amb la qual cosa en principi només permetria una oficina i un departament per empleat. Però al mateix temps combina part dels inconvenients que hem trobat als apartats a) i b). Concretament, tant el camp *department* com el camp *office* podrien contenir alguns valors *NULL*, cosa que no s'hauria de permetre ja que si no podríem perdre informació. Per tant, triaríem aquesta clau primària i a més hauríem d'afegir dues *CONSTRAINT* per evitar que els camps *department* i *office* puguin contenir valors *NULL*. També té l'avantatge davant tota la resta que, en tractar-se d'una clau sintètica de només un

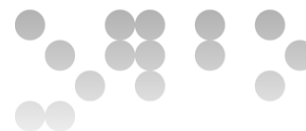


NOM ALUMNE:

camp, podem tenir valors realment petits a l'índex relacionat, cosa que agilitzaria molt les cerques.

d) Aquesta clau primària en un principi pot semblar adequada ja que assegura que una combinació empleat / departament / oficina només pot existir una vegada a la taula i que ni empleat ni departament ni oficina poden contenir valors *NULL*. En canvi, res impedeix que a la taula hi aparegui un mateix empleat més d'una vegada, en combinacions departament / oficina diferents.

Tenint en compte aquestes consideracions, triaríem com a millor opció la clau c), per tres motius principals. Per un costat, en introduir dades, és bastant fàcil recordar que hem de posar algun valor als camps *department* i *office* (a més que afegiríem *CONSTRAINT* per evitar valors *NULL*). Per un altre costat, tot i que poséssim valors *NULL* en aquests dos camps, almenys estem complint la lògica que un empleat només podrà estar a la taula una vegada. Amb això, seria bastant fàcil consultar quins empleats tenen algun valor *NULL* en aquells camps per arreglar-los, al contrari que amb la clau primària de l'apartat d), ja que si trobéssim dues (o fins i tot més) files per un mateix empleat no sabríem de cert quina és la correcta. Un tercer motiu important és que amb l'opció c) es reduiria dràsticament la mida de la clau primària i, per tant, de l'índex relacionat. Amb això ens estalviem molt d'espai d'emmagatzematge i millorem el rendiment de l'índex, fet important sabent que podem tenir molts milers de treballadors a la base de dades.



NOM ALUMNE:

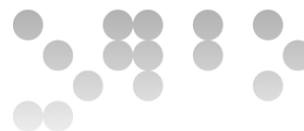
APARTAT B

Explica en què influeix a la BD cadascuna de les següents accions:

- a) Eliminar un índex que sabem que no és necessari.
- b) Actualitzar les estadístiques d'una taula.
- c) Disminuir el nombre de columnes de què es compona un índex.

SOLUCIÓ:

- a) Amb aquesta acció, recuperariem l'espai d'emmagatzematge que ocupa aquest índex i el tornariem a tenir disponible. A més, agilitzaria la inserció, esborrat i modificació de registres a la taula relacionada, ja que no s'hauria de mantenir aquest índex no necessari.
- b) Actualitzant les estadístiques d'una taula aconseguim que els índexs d'aquella taula funcionin d'una manera més eficient, ja que l'SGBD té una informació més acurada sobre la informació que tenim a la BD.
- c) Com a l'apartat a), aquesta acció alliberaria part de l'espai que ocupava l'índex i agilitzaria en un cert grau la inserció, esborrat i modificació de registres a la taula relacionada. D'altra banda, però, podria influir positivament o negativament al rendiment de l'índex en funció de les consultes que volguéssim fer sobre la taula. Per exemple, si estem eliminant de l'índex un camp que realment participi sovint a les consultes que fem sobre la taula, el rendiment d'aquestes consultes disminuiria. En canvi, si els camps que eliminem de l'índex no estaven relacionats amb les consultes que fem a la taula, el rendiment augmentaria.



NOM ALUMNE:

Recursos

Per solucionar aquesta PAC és necessari utilitzar els continguts dels mòduls 1 a 5 del material docent.

Criteris de valoració

El pes dels exercicis en la nota total de la PAC és el següent:

- Exercici 1: 25%
- Exercici 2: 20%
- Exercici 3: 15%
- Exercici 4: 20%
- Exercici 5: 20%

Aquesta PAC té un pes del 50% en la nota d'avaluació continuada.

Format i data de lliurament

El format del fitxer ha de ser Word o OpenOffice, i s'ha de lliurar una versió del mateix fitxer en PDF. Feu enviaments independents per a la versió doc/odt i el fitxer PDF.

El nom del fitxer tindrà el format següent:

PAC2_Cognom1_Cognom2_Nom.extensió

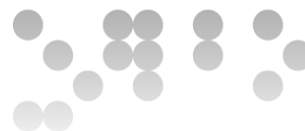
Els cognoms s'escriuran sense accents. Per exemple, un estudiant que es digui Alfredo García Melgar posaria el següent nom a l'arxiu:

PAC2_Garcia_Melgar_Alfredo.doc (u .odt) i .pdf

IMPORTANT: El nom i cognoms de l'estudiant també han d'aparèixer a la portada del document amb la solució.

És responsabilitat de l'estudiant assegurar-se que els documents lliurats s'han pujat correctament i corresponen a l'activitat que cal presentar.

La data límit per lliurar la PAC 2 és dimecres dia 3 de juny.

**NOM ALUMNE:**

Nota: Propietat intel·lectual

Sovint és inevitable, en produir una obra, fer ús de recursos creats per terceres persones. És per tant comprensible fer-ho en el marc d'una pràctica dels estudis del Grau d'Informàtica, sempre que això es documenti clarament i no suposi plagi en la pràctica.

Per tant, en presentar una pràctica que faci ús de recursos aliens, s'ha de presentar juntament amb ella un document en què es detallin tots ells, especificant el nom de cada recurs, el seu autor, el lloc on es va obtenir i el seu estatus legal: si l'obra està protegida pel copyright o s'acull a alguna altra llicència d'ús (CreativeCommons, llicència GNU, GPL...). L'estudiant haurà d'assegurar-se que la llicència que sigui no impedeix específicament seu ús en el marc de la pràctica. En cas de no trobar la informació corresponent haurà d'assumir que l'obra està protegida pel copyright.

Hauran, a més, adjuntar els fitxers originals quan les obres utilitzades siguin digitals, i el seu codi font si correspon.

Un altre punt a considerar és que qualsevol pràctica que faci ús de recursos protegits pel copyright no podrà en cap cas publicar-se en altres mitjans, a no ser que els propietaris dels drets intel·lectuals donin la seva autorització explícita.