

Anàlisi i disseny de bases de dades

Cicle formatiu superior

Desenvolupament d'aplicacions informàtiques

Isidre Guixà i Miranda

IES SEP Milà i Fontanals, d'Igualada

Setembre del 2008
© Isidre Guixà i Miranda
IES Milà i Fontanals
C/. Emili Vallès, 4
08700 - Igualada

En cas de suggeriment i/o detecció d'error podeu posar-vos en contacte via el correu electrònic iguixa@xtec.cat

Cap part d'aquesta publicació, incloent-hi el disseny general i de la coberta, no pot ser copiada, reproduïda, emmagatzemada o tramesa de cap manera ni per cap mitjà, tant si és elèctric, com químic, mecànic, òptic, d'enregistrament, de fotocòpia, o per altres mètodes, sense l'autorització prèvia per escrit dels titulars del copyright.

Índex

Índex.....	3
UD1. Introducció a les bases de dades	5
1. Base de dades i sistema gestor de bases de dades.....	5
2. Objectius del sistema gestor de bases de dades	7
2.1. Possibilitat de consultes complexes	7
2.2. Elasticitat a canvis	7
2.3. Evitar redundància de les dades	8
2.4. Assegurar la consistència de les dades	8
2.5. Prevenir accessos no autoritzats.....	13
2.6. Capacitat d'auditoria	14
2.7. Potents eines d'administració.....	14
3. Arquitectura d'un sistema gestor de bases de dades.....	14
4. Independència lògica i física de les dades	17
5. Llenguatges en sistemes gestors de bases de dades	18
6. Usuaris de bases de dades	20
7. Diccionari de dades	22
8. Arquitectura client-servidor.	23
9. Bases de dades distribuïdes i replicades.....	26
9.1. Possibilitats de distribuir i/o replicar les dades.....	28
9.2. Fragmentació.....	28
9.3. Avantatges i inconvenients	30
9.4. Gestió de bases de dades distribuïdes	31
UD2. Disseny de bases de dades i model entitat - relació	35
1. L'esquema conceptual i el model entitat - relació.....	35
1.1. Entitats	35
1.2. Relacions.....	38
1.3. Model estès.....	42
2. Disseny d'una base de dades	43
3. Models de sistemes gestors de bases de dades.....	48
Annex. Enunciats per exercitar el model entitat-relació.	50
1. L'empresa CFSDAI.....	50
2. La biblioteca.....	51
3. La base de dades atlètica	52
4. La multinacional.....	53

5. La facultat	54
6. L'hospital.....	55
7. L'empresa de programari.....	56
8. La caixa d'estalvis.....	57
9. La vinícola	58
10. La base de dades olímpica	59
11. La policia	60
12. Les eleccions.....	61
13. El cinema	62
14. La publicitat a TV	63
15. L'empresa de mercaderies	64
16. Les cases de colònies	65
17. La cadena de supermercats.....	66
18. L'ICS	68
19. Els cicles formatius a Catalunya.....	69
20. El turisme.....	70
21. L'hotel.....	71
22. Futbol professional.....	72
23. El restaurant	73
24. Monsters of Rock	74
25. Ciclisme	76

UD1. Introducció a les bases de dades

1. Base de dades i sistema gestor de bases de dades

En el present apartat intentarem definir els conceptes de base de dades i de sistema gestor de base de dades, establint la distinció entre aquests conceptes i els conceptes d'arxiu i de sistema gestor d'arxius¹.

Què és una base de dades? Vegem-ne algunes definicions.

Una **base de dades** és un conjunt de fitxers **interrelacionats**, amb **integració i compartició** de dades.

En aquesta definició apareixen tres conceptes importants. Anem a explicitar-los.

- Arxius **interrelacionats**, que indica que els registres d'un arxiu estan relacionats amb els registres dels altres arxius.

Les relacions entre els registres dels diferents arxius poden ser de tres tipus:

- Un a molts, com el cas que es produeix quan per un client hi pot haver moltes comandes de venda.
 - Molts a un, com el cas que es produeix quan moltes comandes de venda poden correspondre a un mateix client
 - Molts a molts, com el cas que es produeix quan un article es ven a molts clients i un client compra molts articles.
- **Integració** de dades, que indica que no hi ha d'haver repeticions ni redundàncies en el conjunt de dades de tots els arxius.

Això s'aconsegueix considerant els diversos conjunts de dades a gestionar (diferents arxius) com un únic conjunt.

- **Compartició** de dades, que indica que molts usuaris, de manera simultània, han de poder accedir a les dades, actualitzant-les fins i tot.

L'anterior definició presenta una base de dades com un conjunt d'arxius. Com a primera definició, no està malament, però no és la realitat. Una base de dades ha d'aparèixer com el resultat d'un anàlisi de la situació a informatitzar i en l'etapa d'anàlisi no es parla mai d'arxius, sinó de les entitats a gestionar, amb els seus atributs i els valors que poden prendre.

¹ **Arxius i sistema gestor d'arxius.**

La comprensió de les bases de dades i dels sistemes gestors de bases de dades és més fàcil si l'alumne ha treballat amb arxius i, per tant, amb sistemes gestors d'arxius. En el *CFS Desenvolupament d'aplicacions informàtiques* aquests conceptes es presenten amb deteniment en el crèdit *Programació estructurada i modular*.

Pel cas en que aquest crèdit es cursi paral·lelament a l'anterior, comentem que arxius i bases de dades tenen per objectiu l'emmagatzematge de les dades de manera que la seva gestió sigui simple i eficaç. Les tècniques informàtiques inicials per assolir aquest objectiu consistiren en la gestió dels arxius seqüencials, la qual va evolucionar cap la gestió d'arxius relatius, calculats i indexats. La continua evolució en l'emmagatzematge de les dades va portar a la gestió de les bases de dades, objectiu del present crèdit.

En determinats punts d'aquesta unitat didàctica es compara la gestió de bases de dades amb la gestió d'arxius, pressuposant que l'alumne és coneixedor de la gestió d'arxius. No cal preocupar-se si no es dona aquesta situació ja que en tal cas l'alumne estaria cursant simultàniament aquest crèdit i el crèdit *Programació estructurada i modular* que inclou l'estudi dels sistemes gestors d'arxius.

Així, seria millor presentar una nova definició de base de dades.

Una **base de dades** és un conjunt d'entitats **interrelacionades**, amb **integració i compartició** de coneixements.

Aquesta definició s'emmarca en el món dels coneixements. Però la informàtica tracta amb dades i, per tant, haurem de tractar la base de dades en el món de les representacions.

Ens cal, per tant, establir com representar les bases de dades. Tenim, d'entrada, dues possibilitats:

- Utilitzar tècniques d'arxius
- Utilitzar tècniques basades en estructures de dades, algunes possiblement conegudes si som coneixedors de la metodologia de la programació (arbres, objectes) i altres potser desconegudes (xarxes, taules -no confondre amb les taules com a tipus de dada en memòria interna-)

La decisió que es va prendre en el seu moment, va ser representar les bases de dades amb tècniques basades en estructures de dades, i així han aparegut diferents models de bases de dades:

- **Model jeràrquic**, basat en l'estructura de dada *arbre*, coneguda per nosaltres. No aprofundirem, però, en aquest model de base de dades.
- **Model en xarxa**, que es basa en l'estructura de dada *xarxa*, desconeguda en aquest cicle formatiu. Tampoc aprofundirem en aquest model de base de dades.
- **Model relacional**, que es basa en l'estructura de dada *taula* -no confondre amb el tipus de dada estàtica *taula*-, encara desconeguda per nosaltres, però que veurem en profunditat en properes unitats didàctiques. Aquest model serà el que posarem en pràctica en aquest crèdit.
- **Models orientats a objectes**, que tenen en compte els conceptes introduïts en l'anàlisi i disseny orientats a objectes².

Així, la definició més correcta de base de dades seria:

Una **base de dades** és un conjunt estructurat de dades amb **integració i compartició**.

I, com gestionarem les bases de dades? Doncs de manera similar a com es gestionen els arxius. Necessitarem d'un conjunt de programes que ens facilitin la seva gestió, de manera similar al conjunt de programes que formen el sistema gestor de fitxers i que faciliten la gestió dels arxius. En el món de les bases de dades, aquest conjunt de programes conformen el que s'anomena **sistema gestor de bases de dades**³. De manera semblant a la utilització de les sigles SGF per referenciar els sistemes gestors de fitxers, utilitzarem les sigles SGBD per referenciar els sistemes gestors de bases de dades.

² Orientació a Objectes

És una metodologia d'anàlisi i programació de nova implantació en comparació amb la metodologia estructurada i modular.

En el món informàtic s'ha demostrat que l'arrelament de les noves tecnologies es produeix bastant més enllà del moment en que comencen a sorgir. I això és el que està passant amb els models de bases de dades orientats a objectes. N'hi ha alguns prototipus, però poc o gens s'utilitzen en el món real.

En l'actualitat, el model de base de dades més arrelat és el relacional, pels que alguns fabricants han incorporat, a títol personal -sense seguir cap estàndard-, conceptes de l'orientació a objectes. Ens movem encara, però, en el model relacional.

³ Termes anglesos

Oi que un SGF permet gestionar diferents arxius? De la mateixa manera, un SGBD permet gestionar diferents bases de dades. No s'ha de confondre el concepte de SGBD (conjunt de programes que permeten gestionar bases de dades) amb el concepte de base de dades (conjunt estructurat de dades amb integració i compartició).

2. Objectius del sistema gestor de bases de dades

En aquest apartat establirem els principals objectius d'un sistema gestor de bases de dades.

2.1. Possibilitat de consultes complexes

Es disposa de mecanismes senzills per efectuar consultes complexes.

Així, imaginem-nos la necessitat de conèixer les principals dades (*codi, raó social, telèfon i NIF*) dels clients que han adquirit algun producte de la família "*Aliments*" al llarg d'un determinat exercici econòmic (per exemple, l'any 2000).

En una gestió d'arxius, possiblement tindríem un arxiu de clients (on hi hauria les dades dels clients), un arxiu de productes (on hi hauria les dades dels productes i entre elles el codi de la família a la que pertany cada producte), un arxiu de famílies de productes (on hi hauria la codificació de les famílies de productes) i un arxiu de vendes (on hi hauria la informació de les vendes de productes efectuades a cada client: quantitat, preu, data, descomptes,...).

Us imagineu la quantitat de feina, amb les tècniques que ens possibiliten els SGF, per tal d'aconseguir la informació demanada? Caldria efectuar un programa amb força codi. En un SGBD relacional (que encara no coneixem), es resoluria amb una instrucció semblant a:

```
select distinct cl.codi, cl.rao, cl.telefon, cl.nif
from client cl, productes pr, familia fm, vendes vn
where cl.codi = vn.codcli
  and pr.codi = vn.codpro
  and fm.codi = pr.codfam
  and upper(fm.desc) = "ALIMENTS"
  and vn.data >= "01-01-2000"
  and vn.data <= "31-12-2000";
```

Oi que val la pena? Compareu les anteriors línies amb el programa que hauríeu d'haver fet en un SGF per aconseguir la informació demanada.

Les anteriors línies corresponen a una instrucció desenvolupada amb el llenguatge SQL, que coneixerem al llarg d'aquest crèdit. Fixeu-vos que no ha calgut obrir ni tancar cap arxiu, ni efectuar cap recorregut ni situar-se en cap registre concret. Simplement hem demanat al SGBD que ens doni una certa informació explicitant on l'ha d'anar a cercar i ell fa la feina per nosaltres. Interessant, oi?

2.2. Elasticitat a canvis

Es disposa de la possibilitat d'efectuar canvis a la base de dades impossibles d'efectuar de manera immediata en SGF.

El terme anglès per un *sistema gestor de base de dades* és *Database Management System* i s'acostuma a abreujar amb DBMS. En canvi el terme per un *sistema gestor de fitxers* és *File Management System* i s'abreuja amb FMS.

Així, imagineu la necessitat d'afegir o suprimir atributs en una entitat o de modificar-ne el seu domini de valors. Imagineu també la necessitat d'afegir noves entitats i d'afegir o modificar relacions entre les entitats. Els SGBD permeten aquests canvis de manera senzilla i transparent a les aplicacions ja existents i no afectades de manera directa pels canvis efectuats.

Aquest objectiu dels SGBD té un nom molt més concret: **independència física i lògica de les dades**. És tal la importància d'aquest objectiu que li donarem, més endavant, un apartat per a tractar-lo amb la profunditat que es mereix.

2.3. Evitar redundància de les dades

Es facilita evitar:

- Repeticions de dades en diferents entitats, amb la implementació de les relacions entre entitats.

Per exemple, no ens caldrà tenir, en les *comandes de venda*, la raó social del client si hi tenim el codi de client, doncs amb aquesta informació es té la raó social de manera ràpida a partir de l'entitat *clients* en tenir relacionades les entitats *clients* i *comandes de venda*.

En SGF no estàvem obligats a tenir en l'arxiu *comandes de venda* la raó social del client, però moltes vegades s'hi tenia per tal de tenir accés de manera ràpida a aquesta informació en gestionar la comanda. Aquesta forma de treballar portava lloc a nombrosos problemes motivats per l'actualització de les dades repetides.

En SGBD no es té la necessitat de tenir la informació repetida ja que el SGBD ens sap facilitar de manera fàcil la informació relacionada existent en altres entitats.

- Emmagatzemar dades calculables

Per exemple, no ens caldrà tenir a l'entitat *clients*, l'import facturat i pendent de cobrar per cada client, ja que aquesta informació és calculable a partir dels cobraments pendents per les factures dels clients i aquesta informació resideix en altres entitats de la base de dades. Quan es necessiti saber aquesta informació, el SGBD ens permetrà calcular-la a l'instant.

D'aquesta manera s'estalvia espai i s'estalvia actualitzar la dada calculada quan hi ha modificacions en les dades a partir de la que es calcula.

A vegades, però, és recomanable renunciar als anteriors estalvis i mantenir la redundància de les dades calculables. Això es justifica quan:

- La consulta de les dades calculables s'efectua moltes vegades, ja que cada vegada cal efectuar els corresponents càlculs. Per exemple, el risc en curs d'un client.
- Els càlculs a efectuar són de cost considerable. Per exemple, les vendes resumides per mes a nivell de client.

Caldrà, per tant, decidir en temps d'anàlisi, el tipus de gestió a efectuar per cadascuna de les dades calculables.

2.4. Assegurar la consistència de les dades

El SGBD contempla mecanismes per a conservar correctes les dades dins la base de dades.

Hi ha diferents causes per les que les dades poden passar a ser incorrectes:

- Desastre en el suport físic on hi ha enregistrades les dades
- Errors de disseny en programes que no efectuen el tractament de dades de manera correcta.
- Actualització incompleta de dades per caiguda del sistema informàtic.
- Introducció de dades no vàlides a través de les eines que puguin existir per interactuar amb la base de dades.
- Actualitzacions concurrents que provoquen la pèrdua d'operacions d'actualització.

Com pot el SGBD fer front a les causes enumerades? Amb diferents mecanismes que explicitem a continuació.

2.4.1. Mecanismes de recuperació

Hi ha diferents situacions en les que pot ser necessari efectuar una recuperació de les dades. Per tant, també hi haurà diferents formes d'efectuar-la.

Es produeix una **recuperació en calent** quan després d'una caiguda del SGBD (per qualsevol motiu) han quedat operacions d'actualització a mig efectuar.

Les operacions d'actualització a mig efectuar poden correspondre a una única operació sobre la base de dades o a un conjunt d'instruccions agrupades sota una transacció.

La recuperació en calent consisteix en finalitzar o en desfer les operacions que havien quedat a mig efectuar. El SGBD decideix de manera automàtica, en tornar a estar operatiu, si finalitza les operacions o les desfà.

Per a poder dur a terme la recuperació en calent, el SGBD en execució normal, va enregistrant en un arxiu especial, anomenat **dietari** o **log**, la informació corresponent a les actualitzacions a efectuar sobre la base de dades. En finalitzar correctament les actualitzacions corresponents a una operació o transacció, el SGBD deixa una marca de verificació (**checkpoint**) en el diari i en un arxiu de rearrencada. D'aquesta manera, quan és necessària la recuperació en calent, el SGBD obté la marca de verificació més recent de l'arxiu de rearrencada i efectua un recorregut de l'arxiu dietari des d'aquest punt fins el final, per obtenir la informació de quines operacions ha de desfer o finalitzar.

Es produeix una **recuperació en fred** quan es detecta un desastre a la base de dades i s'ha de refer el seu contingut actual a partir de la darrera **còpia de seguretat** i dels arxius **dietaris** enregistrats des de la còpia.

Cal observar que per a poder efectuar aquesta recuperació és tan important disposar d'una còpia de seguretat recent com dels arxius dietaris. Per aquest motiu aquests arxius no es destrueixen.

Hi ha encara una tercera situació en la que pot ser necessari haver d'utilitzar els arxius dietaris: la recuperació semàntica.

Es produeix una **recuperació semàntica** quan es detecta que hi ha hagut, en un instant temporal passat, una actualització incorrecta segons el punt de vista semàntic i s'ha de refer dita actualització i tots els moviments posteriors fins el moment present.

Aquesta situació és la que es pot produir, per exemple, quan s'ha enregistrat un càlcul incorrecte que ha provocat uns posteriors moviments en cascada. Imaginem, per exemple, els problemes que pot ocasionar a un client d'una entitat bancària l'enregistrament d'un ingrés en metàl·lic efectuat en un caixer automàtic, en una data posterior de la realment efectuada (per error de l'operador que enregistra l'ingrés en el sistema). Evidentment, caldrà restaurar l'actualització a la data correcta i refer els moviments posteriors que s'haguessin desencadenat.

A vegades, aquesta recuperació es pot efectuar directament a través d'un històric de moviments enregistrats a la pròpia base de dades, però en altres ocasions és necessari restaurar el contingut de la base de dades a partir dels arxius dietaris. En tal situació, s'acostuma a partir de la base de dades actual i a desfer els moviments enregistrats en els dietaris (**recuperació cap enrera**), fins arribar al moment en que es va produir l'actualització semànticament incorrecta; s'enregistra l'actualització de manera correcta i s'inicia el procés de **recuperació cap endavant** a partir dels mateixos dietaris enregistrant els moviments pertinents a partir de la nova situació.

2.4.2. Definició de la integritat

Sota aquest mecanisme s'agrupen:

a) Regles d'integritat del model

Són les regles de funcionament que porta associat el model de base de dades que correspongui. Així, en el model jeràrquic, basat en l'estructura de dades *arbre*, es verifica, obligatòriament les regles d'integritat dels arbres:

- un node no arrel sempre té un i només un pare; l'arrel és l'únic node que no té pare
- si l'arbre és binari, cap node pot tenir més de dos fills
- si l'arbre és balancejat, s'ha de mantenir el factor d'equilibri

El SGBD controlarà sempre les regles d'integritat del model en el que es basa.

b) Regles d'integritat referencial

Són les regles que el SGBD obligarà a complir en funció de les relacions que s'estableixin entre les entitats de la base de dades.

Així, si es disposa d'una entitat *país* i d'una entitat *departament* (província) i s'estableix que entre *país* i *departament* hi ha una relació de *1 a molts*, de manera que tot *departament* ha de pertànyer obligatòriament a un *país* existent a l'entitat *país*, el SGBD controlarà aquest fet i no permetrà:

- donar d'alta un departament i no assignar-li país o assignar-li un país inexistent a la base de dades
- modificar el país assignat correctament a un departament per un país inexistent o deixar-lo sense país assignat

- eliminar un país quan hi hagi departaments que el tenen assignat
- modificar la clau primària d'un país que tingui departaments assignats a través d'aquest valor de clau

Alguns SGBD permeten definir el comportament a seguir davant la tercera situació presentada, que pot establir-se en no permetre l'eliminació o en efectuar una eliminació en cascada (esborrant país i departaments associats) o en efectuar l'eliminació i deixar els corresponents departaments sense assignar (per a poder fer això no hi hauria d'haver l'obligatorietat d'assignar cada departament a un país, és a dir, caldria l'existència d'un valor nul en aquesta assignació).

Alguns SGBD també permeten definir el comportament a seguir davant la quarta situació presentada, que pot establir-se en no permetre la modificació o en modificar la clau primària del país i el valor assignat als corresponents departaments, de manera que el lligam país-departament continuï mantenint-se.

c) Restriccions d'usuari

Són restriccions que es defineixen pels diferents atributs de les entitats, de manera que el SGBD les comprova en les operacions d'actualització. Poden ser:

- estàtiques, com per exemple, "10 < edat < 99"
- dinàmiques, com per exemple, "els preus mai baixaran"

2.4.3. Control de concurrència

Els SGBD han de controlar de forma efectiva els accessos concurrents a les dades. Del contrari es pot produir el problema conegut com a **operació perduda**, exemplificat en la següent situació.

Considerem diferents llocs de treball en un magatzem, ocupats per operaris que executen sortides de material. Considerem un article A del que en consten en existència 100 unitats. Suposem que cada operari ha d'efectuar una sortida de 20 unitats.

Imaginem-nos que els esdeveniments tenen lloc, en el temps, en el següent ordre:

Operari 1	Operari 2
Consulta l'existència de la base de dades i en queden 100 unitats	
	Consulta l'existència de la bases de dades i en queden 100 unitats
Efectua la sortida de 20 unitats, actualitzant l'existència a la lectura anterior (100) menys la sortida efectuada. És a dir, actualitza el registre a 80 unitats	
	Efectua la sortida de 20 unitats, actualitzant l'existència a la lectura anterior (100) menys la sortida efectuada. És a dir, actualitza el registre a 80 unitats

És clar que el valor final existent a la base de dades és incorrecte. S'ha perdut l'efecte d'una de les operacions d'actualització.

Per a solucionar aquest problema, el SGBD està programat per a controlar, mitjançant el **mecanisme de bloqueig**, les actualitzacions de dades a un determinat nivell (entitat, instància, atribut o, fins i tot, interval de valors) de manera que en

iniciar-se una actualització es bloqueja la dada al nivell que correspongui i ningú més pot accedir a la dada bloquejada fins que el propi SGBD l'allibera. Les sol·licituds d'actualització poden quedar en espera de l'alliberament o ser rebutjades, segons s'hagi definit l'actuació en el SGBD.

El nivell al què s'aplica el bloqueig acostuma a anomenar-se **grànul**. Així hi pot haver grànuls a nivell d'entitat, a nivell d'instància de l'entitat, a nivell d'atribut o, fins i tot, a nivell d'interval de valors per un atribut. Exemplificant-ho, pot donar-se el cas que es bloquegi tota l'entitat *clients* (el grànul és l'entitat) o un *client* concret (el grànul és la instància) o l'atribut *domicili* d'un client concret (el grànul és l'atribut) o l'atribut *codi postal* del client si està entre un determinat interval de valors (el grànul és un interval de valors de l'atribut).

L'elecció del nivell d'actualització a controlar pel SGBD és un tema polèmic. Si s'escull un nivell petit s'afavoreix un major paral·lelisme de les transaccions, però augmenta la quantitat de temps utilitzada en l'accés a les dades.

El mecanisme de bloqueigs soluciona el control concurrent, però provoca les conegudes **abraçades mortals**, consistents en una espera circular entre dues o més transaccions de manera que cadascuna sol·licita una actualització sobre el mateix element. Exemplifiquem-ho.

Transacció 1	SGBD	Transacció 2
Inicia actualització d'un element A		
	Bloqueja l'element A	
		Inicia actualització d'un element B
	Bloqueja l'element B	
Necessita accedir a l'element B, que està bloquejat		
	Deixa la transacció 1 en espera de desbloqueig de l'element B	
		Necessita accedir a l'element A, que està bloquejat
	Deixa la transacció 2 en espera de desbloqueig de l'element A	

Per resoldre aquest problema es poden seguir dues tècniques:

- Controlar que no succeeixi, cosa que no sempre és fàcil

Hi ha diferents mecanismes per a prevenir l'abraçada mortal (amb un d'ells n'hi ha prou). Suposem que la transacció T_1 té bloquejat el grànul i que la transacció T_2 intenta bloquejar el mateix grànul. Podem actuar amb una de les tres formes següents:

- "Wait-Die": Si T_2 és més vella, s'espera; en cas contrari, és cancel·la.
- "Kill-Wait": Si T_2 és més vella, T_1 és cancel·la; en cas contrari, T_2 s'espera.
- "Wound-Wait": Si T_2 és més vella, T_1 es cancel·la a no ser que hagi començat a confirmar actualitzacions; en cas contrari, T_2 s'espera.

Una transacció podria ser cancel·lada moltes vegades si no s'adopta alguna mesura per detectar tal situació i donar prioritat a dita transacció. L'*edat* de les transaccions es calcula a través d'un valor que té un comptador quan s'inicien.

- Detectar la seva aparició i trencar-la cancel·lant algunes transaccions (les víctimes).

Hi ha diferents maneres de detectar l'abraçada mortal. Una d'elles consisteix en mesurar l'interval de temps que dura una transacció, de manera que en superar uns determinats límits es pugui considerar que està en espera permanent.

2.5. Prevenir accessos no autoritzats

El SGBD ha de protegir les dades dels intents d'accés dels usuaris que no tinguin els privilegis corresponents. Aquesta prevenció s'efectua a nivell del SGBD i no és incompatible amb els mecanismes que pugui tenir el propi sistema informàtic (control d'accés a les instal·lacions, control d'entrada al sistema per part del sistema operatiu,)

Pel que fa referència al SGBD (tema que ens ocupa), els sistemes moderns permeten una o ambdues de les estratègies següents:

- **Control discrecional**, que permet definir privilegis d'accés a cada usuari (o grup d'usuaris) sobre cada objecte (o grup d'objectes).
- **Control obligatori**, en el que cada objecte (o grup d'objectes) és etiquetat amb un cert nivell de classificació i cada usuari (o grup d'usuaris) és etiquetat amb un cert nivell d'acreditació, de manera que cada objecte només pot ser accedit pels usuaris amb un determinat nivell d'acreditació.

Els esquemes obligatoris són més rígids que els esquemes discrecionals.

Per a que el SGBD pugui decidir quines restriccions ha d'aplicar ha de ser capaç de reconèixer l'origen de la sol·licitud. Per aquest motiu, els SGBD han de demanar a l'usuari la identificació en iniciar una sessió de treball en el SGBD. Normalment, aquesta identificació va acompanyada d'una contrasenya⁴.

Normalment, en iniciar una sessió de treball en un sistema informàtic, aquest, via sistema operatiu, exigeix la identificació de l'usuari. L'usuari, una vegada identificat, entra en el sistema informàtic i pot demanar d'iniciar una sessió de treball amb el SGBD. Aquest pot exigir altra vegada la identificació a l'usuari, la qual pot ser totalment diferent de l'exigida pel sistema operatiu per accedir al sistema informàtic. Alguns SGBD permeten utilitzar el mateix sistema d'accés que l'utilitzat per accedir al sistema informàtic, de manera que l'usuari no hagi d'anar identificant-se en les diferents capes que ha de travessar fins arribar al SGBD.

A més de totes aquestes tècniques, alguns SGBD permeten utilitzar la criptografia per a emmagatzemar les dades dins la base de dades. És a dir, les dades s'enregistren prèviament criptografiades (xifrades), de manera que en recuperar-les el SGBD ha de procedir a descriptografiar-les (desxifrar-les). D'aquesta manera es prevé el fet que s'accedeixi a les dades sense passar pel SGBD, és a dir, accedint directament al suport físic on estan enregistrades.

És clar que tota tècnica per prevenir els accessos no desitjats implica un esforç de temps i càlcul per part del SGBD. Per tant, caldrà establir els criteris d'accés en funció de les reals necessitats de control dins cada organització.

⁴ **Mecanismes d'identificació**

La forma més normal d'accedir al SGBD és via identificació escrita de l'usuari acompanyada d'una contrasenya. Aquests sistemes possiblement seran modificats en un futur, tal i com es veu en les pel·lícules actuals, passant a utilitzar altres mètodes basats en característiques físiques, com les empremtes dactilars, la veu, la pupil·la dels ulls,....

2.6. Capacitat d'auditoria

L'objectiu que es pretén amb l'auditoria és el de poder determinar quins usuaris van interactuar amb la base de dades i en quin moment.

Hi ha diferents nivells d'auditoria:

- Control dels usuaris que inicien sessió de treball amb una base de dades.

Aquest nivell d'auditoria precisa pocs recursos.

- Control de les modificacions efectuades en una base de dades (o en uns determinats objectes de la base de dades) per cada usuari.

Aquest nivell d'auditoria precisa més recursos que el nivell anterior i normalment s'activa per aquells objectes crítics.

- Control de les lectures efectuades en una base de dades per cada usuari.

Aquest nivell d'auditoria precisa molts més recursos que els nivells anteriors i normalment s'activa quan hi ha indicis de que s'està extraient informació confidencial (espionatge).

2.7. Potents eines d'administració

Per poder fer front a tots els objectius enumerats prèviament, el SGBD necessita d'un administrador (el presentarem posteriorment) i aquest necessita d'unes potents eines per fer front als requeriments que l'organització pretén obtenir del SGBD.

Algunes de les eines del SGBD poden formar part del propi SGBD. Altres poden ser programes d'aplicació que interactuen amb el SGBD.

3. Arquitectura d'un sistema gestor de bases de dades

En els apartats anteriors hem presentat els conceptes de *base de dades* i de *sistema gestor de bases de dades*, establint les diferències entre ells i els objectius que persegueix el darrer.

L'evolució dels SGF cap als SGBD ha estat motivada per noves prestacions que les organitzacions demanaven als SGF i que aquests no podien facilitar. En especial, una de les grans demandes era el fet que els arxius que fins el moment eren pensats per unes aplicacions concretes, poguessin passar a formar part d'una visió més global.

Amb el concepte de BD s'aconsegueix la integració de totes les dades (dispersades fins el moment en diferents arxius i fins i tot duplicades) en un sol objecte. Ara, però, apareix una necessitat. Cal que la BD es pugui veure a trossos simulant, d'alguna manera, els arxius als que tenien accés les diferents aplicacions en el SGF.

Aquesta necessitat és motivada pel fet que els usuaris d'un determinat departament de l'organització han de poder veure les dades des del seu punt de vista, des del seu enfocament. Exemplem-ho.

Imaginem-nos l'entitat *clients*. El departament comercial necessita gestionar les dades *codi, raó social, domicili fiscal, domicili de lliurament, venedor assignat, polítiques de descomptes i facturació*. En canvi, el departament comptable necessita gestionar les dades *codi, raó social, domicili fiscal, facturació, compte comptable, compte d'explotació i cobraments*. És clar que l'entitat *clients* és única i que englobarà la totalitat dels atributs descrits. Però l'organització exigeix que la visió que en tingui el departament comercial sigui diferent de la visió que en tingui el departament comptable, de manera semblant a les visions que tenien antigament en treballar amb SGF. I aquesta exigència és lògica ja que es vol centrar als usuaris en la gestió de les dades que els pertoca.

Entenc que ara mateix podeu tenir una mica d'embolic en el cap. Pot ser us esteu preguntant: quin problema hi ha en tenir tots els atributs en una única entitat si les aplicacions desenvolupades pels departaments comercial i comptable ja s'encarregaran de gestionar les dades que els pertocin? per què cal tenir diferents visions de la base de dades?

Hi ha diferents motius que donen resposta a la necessitat plantejada:

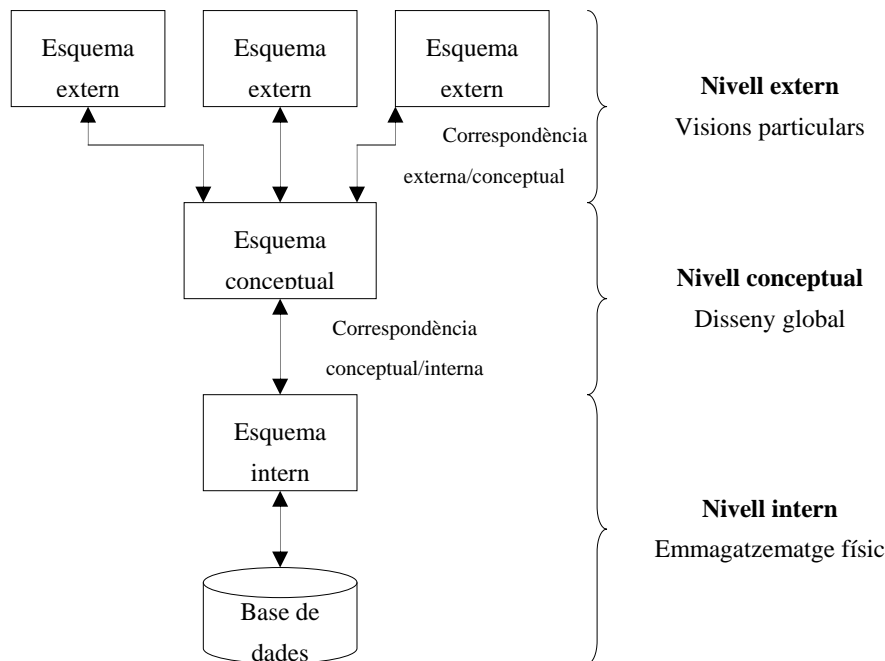
- El SGBD possibilita la realització de consultes complexes de manera senzilla. Recordeu l'exemple de l'apartat 2.1.? Un usuari que conegui la sintaxi del llenguatge (cosa no gens difícil com bé veurem més endavant), pot executar operacions per obtenir i modificar informació de la base de dades. Si l'usuari és del departament comercial hauria de veure els *clients* dels del punt de vista comercial, mentre que si l'usuari és del departament comptable els hauria de veure dels de punt de vista comptable. Cap d'ells n'ha de fer absolutament res amb les dades necessàries per l'altre departament.
- Si les aplicacions desenvolupades pels diferents departaments no es basen en l'entitat *clients* sinó en la particular visió de l'entitat que correspon al seu departament, qualsevol canvi en l'entitat *clients* que afecti a l'altre departament, li serà transparent, és a dir, no es veurà afectada doncs la particular visió de l'entitat continuarà invariable.

Així doncs, es vol treballar amb bases de dades per a integrar i compartir les dades però, a la vegada, es vol tenir diferents visions d'ella.

Com aconseguir aquesta dualitat?

La situació es va plantejar en els inicis dels SGBD i evidentment va sorgir als Estats Units d'Amèrica. Allí, l'*American National Standard Institute (ANSI)*, responsable de la definició de normes per la indústria del país, estructurat en diferents divisions, una de les quals és l'anomenada *X3* que s'ocupa de les normes relatives als ordinadors i a les màquines d'oficina, va crear el *System Planning And Requirements Committee (SPARC)*, del qual va sorgir l'informe que posa les bases per solucionar la situació plantejada i que proposa per les bases de dades una arquitectura de tres nivells que s'ha conegut com **arquitectura ANSI/X3/SPARC**.

L'esmentat informe planteja que l'arquitectura d'una base de dades ha de tenir tres nivells: **intern, conceptual i extern**, cadascun dels quals correspon a un punt de vista diferent de la base de dades: emmagatzematge físic de la BD, disseny global de la BD i visions particulars de la BD tal i com es veu a la figura:



Anem a veure amb una mica de profunditat cadascun dels tres nivells.

1) Nivell intern

Describeix l'estructura d'emmagatzematge físic que conté les dades, és a dir, descriu, entre altres elements, els arxius físics que contenen les dades (noms, tipus d'accés, ubicació,...), els registres físics d'aquests arxius (longitud, camps,...) i les rutes d'accés (índexs, encadenaments,...)

El nivell intern està descrit en el SGBD pel que s'anomena **esquema intern**. És el SGBD qui juntament amb el SO gestionarà, en temps d'execució, les entrades i sortides físiques, en base a la informació existent en l'esquema intern. En aquesta gestió el SGBD transforma dades de registre físic a registre lògic, operació anomenada de diferents formes: transformació conceptual/interna, correspondència conceptual/interna, mapeig conceptual/intern.

2) Nivell conceptual

Describeix el disseny global de tota la base de dades. En aquest nivell s'inclou la descripció de totes les entitats, amb els seus atributs i dominis de valors, així com les interrelacions entre les entitats, les regles d'integritat referencial, les restriccions, les autoritzacions,... En definitiva, representa la visió integrada de tots els usuaris. El nivell conceptual està descrit en el SGBD pel que s'anomena **esquema conceptual**.

3) Nivell extern

Describeix les diferents visions definides a la base de dades, cadascuna de les quals representa la percepció que de la base de dades en tenen els diferents departaments o grups d'usuaris o aplicacions.

Cadascuna de les visions de la base de dades definides en el nivell extern ha d'estar descrita en el que s'anomena **esquema extern**. Els anteriors nivells intern i conceptual estan descrits sempre en un únic esquema intern i conceptual. En canvi, en el nivell extern hi ha d'haver tants esquemes externes com visions diferents de la base de dades es vulgui mantenir.

Cada esquema extern correspon a la visió que es donarà als usuaris i aplicacions que hi tinguin accés. Les entitats aquí definides poden provenir directament d'entitats existents en l'esquema conceptual o ser el resultat d'un muntatge a partir de les entitats existents en l'esquema conceptual. Així, en el cas presentat a l'inici d'aquest apartat, en l'esquema extern corresponent al departament comercial hi hauria una entitat *clients* aconseguida com a subconjunt de la vertadera entitat *clients* existent en l'esquema conceptual. Les entitats de l'esquema extern poden ser muntades a partir de subconjunts, interseccions i unions d'altres entitats de l'esquema conceptual o del propi esquema extern. Els noms dels atributs també es poden redefinir i els dominis de valors es poden fer més restrictius que els existents en l'esquema conceptual.

És el SGBD qui gestionarà, en temps d'execució, les entrades i sortides sol·licitades pels usuaris en els esquemes externs, en base a la informació existent en el propi esquema. En aquesta gestió el SGBD transforma les sol·licituds de l'esquema extern a l'esquema conceptual, operació anomenada de diferents formes: transformació externa/conceptual, correspondència externa/conceptual, mapeig extern/conceptual.

4. Independència lògica i física de les dades⁵

En l'apartat destinat als objectius principals a assolir amb els SGBD apareixia l'*elasticitat a canvis* i es comentava que aquesta s'assolia gràcies a la independència física i lògica de les dades, fet molt important que es mereixia un apartat específic.

S'aconsegueix **independència física de les dades** quan l'accés a les dades no depèn del suport extern, ni del sistema operatiu, ni del maquinari, ni de l'estructura física en que estan distribuïdes les dades.

S'aconsegueix **independència lògica de les dades** quan el disseny lògic de les dades es pot modificar sense haver de modificar les aplicacions ja desenvolupades i que no estan directament implicades en els canvis efectuats.

Els SGBD asseguren els dos tipus d'independència.

La independència física s'assoleix amb la distinció entre nivells conceptual i intern, ja que qualsevol canvi en el suport físic, sistema operatiu, maquinari o estructura física afectarà a com es troben les dades en el suport físic i, per tant, es reflectiran en l'esquema intern, però l'accés a les dades per part dels usuaris i de les aplicacions no es veurà afectat ja que serà el SGBD qui en la transformació conceptual/interna reconduirà l'accés a la nova situació.

La independència lògica s'assoleix amb la distinció entre nivells externs i conceptual, ja que qualsevol canvi en el disseny de les dades afectarà l'esquema conceptual i, de retruc, algun(s) esquema(es) extern(s). Les aplicacions ja desenvolupades sobre un esquema extern que es manté intacte no han de patir cap reestructuració. Fins i tot, es pot donar el cas de modificacions en l'esquema conceptual que afectarien directament a un esquema extern però aquest no es veu afectat per què la transformació externa/conceptual ha anul·lat la modificació.

En efecte. Suposem que en l'esquema conceptual teníem les entitats:

⁵ **Independència física de les dades en sistemes gestors de fitxers**

En els sistemes gestors de fitxers hi ha independència física de les dades que s'assoleix amb la distinció entre arxiu extern (visió de l'arxiu des del sistema operatiu) i arxiu intern (visió de l'arxiu des dels programes), fet possibilitat per l'arquitectura de dos nivells existent en sistemes gestors d'arxius.

país (codi, nom, habitants)
província (codi, nom, codi_país)
hospital (codi, nom, num_llits)

Suposem que tenim dos esquemes externs:

Esquema extern 1:

país, mapejada a partir de *país* de l'esquema conceptual
província, mapejada a partir de *província* de l'esquema conceptual

Esquema extern 2:

hospital, mapejada a partir d'*hospital* de l'esquema conceptual

Suposem que s'arriba a la conclusió que l'atribut *habitants* de l'entitat *país* ha de passar a estar a nivell de l'entitat *província*, doncs interessa tenir aquest valor desglossat entre les diferents províncies que formen un país. Això implica una modificació en l'esquema conceptual que pot mantenir-se transparent pels dos esquemes externs, tal i com es veu a continuació.

Esquema conceptual:

país (codi, nom)
província (codi, nom, codi_país, habitants)
hospital (codi, nom, num_llits)

Esquema extern 1:

país, amb els atributs *codi* i *nom* mapejats de manera directa de *país* de l'esquema conceptual i l'atribut *habitants* mapejat com a suma dels valors de l'atribut *habitants* de *província* per les províncies assignades al corresponent país
província, amb els atributs *codi*, *nom* i *codi_país* de manera directa de província de l'esquema conceptual

Esquema extern 2:

hospital, mapejada a partir d'*hospital* de l'esquema conceptual

Observeu que les modificacions efectuades en l'esquema s'han pogut mantenir transparents en els dos esquemes externs. Això no acostumarà a passar en tots els esquemes externs, ja que si s'efectua alguna modificació de disseny en l'esquema conceptual és molt probable que algun esquema extern s'hagi de veure afectat. En el cas que hem presentat, és molt possible que la modificació efectuada es vulgui fer aparèixer en l'esquema extern 1. Si és així, les aplicacions desenvolupades sobre aquest esquema extern hauran de modificar-se en allò que es vegin afectades.

5. Llenguatges en sistemes gestors de bases de dades

Els SGBD han de proporcionar llenguatges per a poder descriure els diferents nivells de la base de dades i per a poder accedir a les dades i així efectuar consultes i actualitzacions. Aquests llenguatges es classifiquen en⁶:

⁶ Termes en anglès

En l'argot informàtic s'acostuma a utilitzar els termes:

- *Data Definition Language*, abreujat per DDL
- *Data Control Language*, abreujat per DCL
- *Query Language*, abreujat per QL
- *Data Manipulation Language*, abreujat per DML

a) **Llenguatges destinats a la definició de les dades (LDD)**, que permeten definir els objectes (entitats, atributs, dominis, regles d'integritat referencial, restriccions, rutes d'accés,...) en els diferents esquemes.

Cada SGBD disposa d'un o més llenguatges LDD específics. En uns casos existeix un únic llenguatge, amb sentències que permeten descriure els elements dels diferents esquemes i, en altres, hi ha un llenguatge diferent per cadascun dels nivells.

b) **Llenguatges destinats al control sobre les dades (LCD)**, que permeten concedir i retirar permisos sobre els diferents objectes de la base de dades.

En alguns SGBD no hi ha distinció entre LDD i LCD i es parla únicament de LDD per a les tasques de definició de dades i de concessió i revocació de privilegis.

c) **Llenguatges destinats a la consulta de les dades (LC)**, que permeten efectuar lectures o sortides lògiques en el nivell extern.

d) **Llenguatges destinats a la manipulació de les dades (LMD)**, que permeten actualitzar la base de dades (altes, baixes i modificacions).

En alguns SGBD no hi ha distinció entre LC i LMD i es parla únicament de LMD per a les consultes i actualitzacions.

Tots aquests llenguatges acostumen a tenir una sintaxi senzilla, semblant a les ordres de consola per un sistema operatiu, anomenada **sintaxi autosuficient**.

Els llenguatges LC i LMD poden presentar, però, una segona sintaxi, **sintaxi hostatjada**, consistent en un conjunt de sentències que són admeses dins d'un llenguatge, anomenat **llenguatge amfitrió**, de propòsit general.

Així ens podem trobar LC i LMD que es poden hostatjar en llenguatges de tercera generació com C, Cobol, Fortran,... i en llenguatges de quarta generació.

Per hostatjar en llenguatges de tercera generació compilats es necessari disposar d'un precompilador específic pel llenguatge, subministrat pel fabricant del SGBD, amb el que es precompila el codi font que conté les sentències LC i LMD i s'obté, en cas d'absència d'errors, un nou font on les sentències LC i LMD han estat tractades i encapsulades de manera que es pot procedir a la compilació normal per tal d'obtenir el codi objecte i passar a la fase d'enllaç per obtenir el codi executable.

Els llenguatges de quarta generació que permeten utilitzar sentències LC i LMD acostumen a ser llenguatges incorporats al propi SGBD proporcionats pel fabricant del SGBD. Aquest és el cas dels productes comercials que utilitzarem en el present crèdit. Els productes *Access '97* i *Personal Oracle 8* incorporen eines visuals de quarta generació per a desenvolupar pantalles (altes, baixes, consultes i modificacions), informes, gràfics,... que interactuen amb la base de dades.

Per finalitzar, distingir que podem trobar dos tipologies de llenguatges LC i LMD en funció de la manera que l'usuari haurà d'accedir a la base de dades:

- **implícits**, on l'usuari demana les dades a accedir però no diu la forma d'accedir-hi; també s'anomenen **no procedurals** ja que no s'indica el procediment per accedir a les dades,
- **explícits**, on l'usuari indica la forma d'accedir a les dades desitjades; també s'anomenen **procedurals** ja que s'indica el procediment a seguir per accedir a les dades.

6. Usuaris de bases de dades

Un SGBD permet emmagatzemar dades, les quals han de ser accedides pels usuaris de l'organització i, com és d'imaginar, no tots els usuaris tenen les mateixes necessitats ni les mateixes capacitats d'execució. Per aquest motiu es distingeixen diferents tipus d'usuaris.

1) Usuaris informàtics

És el grup més reduït d'usuaris de bases de dades. Han de tenir forts coneixements informàtics. Podem distingir-ne:

a) **Dissenyadors d'aplicacions** (analistes, dissenyadors i programadors), que són els qui desenvolupen les aplicacions informàtiques que interactuen amb bases de dades.

Per a desenvolupar la seva feina utilitzen els llenguatges LC i LMD hostatjats en llenguatges de tercera i quarta generació textuals o gràfics. En l'actualitat, però, es treballa quasi exclusivament en entorns gràfics.

Aquests usuaris utilitzen també, en les fases de disseny, codificació i verificació de les aplicacions, els llenguatges LC i LMD autosuficients per a tenir accés immediat a la base de dades i així comprovar els seus continguts.

Per últim, tot i que no és una funció que els pertoqui directament, ja que és tasca d'administració, també acostumen a conèixer els llenguatges LDD i LCD doncs en les fases de desenvolupament de les aplicacions hauran de ser, en ocasions, els que defineixin els diferents esquemes de la base de dades.

Així doncs, els programadors d'aplicacions són usuaris dels llenguatges LC i LMD autosuficients i hostatjats en 3GL i 4GL, procedurals i no procedurals. També són usuaris, en menor grau, dels llenguatges LDD i LCD.

b) **Administradors de bases de dades**, que com el seu nom indica, són els que s'encarreguen d'*administrar* la base de dades, tasca que pot ser desenvolupada per una o varies persones. I què s'entén per *administrar*?

Les responsabilitats de l'administrador de la base de dades (*Database Administrator - DBA*) són, entre d'altres:

- Creació i manteniment dels diferents esquemes (externs, conceptual i intern) de la base de dades.

Aquesta tasca inclou la definició de les entitats, amb els seus atributs, dominis de valors, regles d'integritat, restriccions i procediments de validació, junt amb les correspondències externa/conceptual i conceptual/interna. Per l'esquema intern s'inclou l'estratègia d'emmagatzematge a seguir pels SGBD junt amb els mètodes d'accés a utilitzar.

Per a poder fer front a aquesta responsabilitat, l'administrador utilitza el llenguatge LDD i les eines que el SGBD pugui incorporar. En situacions de manteniment (canvis en el disseny conceptual) li caldrà dissenyar i executar processos de reorganització de les entitats existents i, per aquest motiu, precisarà dels llenguatges LMD i LC autosuficients i, fins i tot, de les corresponents versions hostatjades en un llenguatge de tercera generació.

- Control de l'accés a les dades.

S'inclou en aquest apartat tota la gestió d'usuaris (altes, baixes i modificacions) amb els diferents privilegis d'accés (consulta i/o actualització) sobre els diferents objectes de la base de dades.

Hi ha SGBD que contemplen la possibilitat de treballar amb grups d'usuaris i grups de privilegis, de manera que faciliten la tasca repetitiva de concedir i/o revocar un determinat conjunt de privilegis a un conjunt d'usuaris.

Per a poder fer front a aquesta responsabilitat, l'administrador utilitza el llenguatge LCD i les eines que el SGBD pugui incorporar.

- Gestió dels mecanismes de recuperació, definint les estratègies oportunes en funció de l'organització i de la utilització que se'n fa de la base de dades.
- Execució del *tunning* o de la *posta a punt*

Aquest és un punt fonamental en l'administració de les bases de dades ja que consisteix en cercar el millor funcionament possible de la base de dades en funció de la utilització que se'n fa. És a dir, s'estudia el rendiment (*performance*).

Així, de forma periòdica, l'administrador ha d'analitzar, mitjançant estadístiques i eines que proporciona el propi SGBD, la utilització que s'efectua de la base de dades i la resposta que el SGBD està donant, per tal d'efectuar les correccions oportunes i, en conseqüència, millorar el funcionament.

La posta a punt s'ha de repetir periòdicament ja que la utilització de la base de dades pot canviar per les necessitats de funcionament de l'organització. També és possible que l'organització no hagi modificat el seu funcionament però que el volum de dades hagi tingut una important modificació de manera que l'estructura d'emmagatzematge i accés (nivell intern) precisi d'algun retoc.

Dels diferents punts anomenats, se'n dedueix que aquest l'administrador ha de conèixer els llenguatges LDD i LCD i les versions autosuficients de LMD i LC. En certs SGBD necessitarà també els llenguatges LMD i LC hostatjats en llenguatges de tercera generació.

2) Usuaris no informàtics

És el grup més ampli d'usuaris de bases de dades. No estan obligats a tenir coneixements informàtics i, si en tenen, són escassos. Podem distingir-ne:

a) **Usuaris paramètrics**, en els que la interacció amb la base de dades s'efectua en base a l'execució d'aplicacions dissenyades per atendre específicament les seves necessitats. L'usuari, com a molt, ha de donar resposta a les preguntes que l'aplicació li formula.

Així, per exemple, quan estem en un caixer automàtic d'una entitat bancària, som usuaris paramètrics, ja que responem a les opcions i preguntes que una aplicació informàtica ens efectua i és aquesta la que interactua directament amb la base de dades de l'entitat bancària.

Aquest usuari no té per què conèixer cap llenguatge de base de dades ni cap eina de generació de codi.

b) **Usuaris finals**, que poden efectuar consultes no previstes, diguem-ne espontànies, a la base de dades, generant petits programes.

Aquest és el cas, per exemple, del personal d'administració d'una empresa, que té al seu abast la possibilitat d'efectuar consultes i/o actualitzacions sobre la base de dades sense utilitzar cap de les aplicacions desenvolupades a l'efecte. Per a tenir aquesta possibilitat és menester dues condicions:

- disposar d'una eina que ho possibiliti (eines de quarta generació i llenguatges LC-LMD autosuficients i implícits)
- conèixer el contingut de l'esquema extern de la base de dades sobre el que té accés (entitats, atributs, dominis de valors, relacions entre les entitats, regles d'integritat, restriccions,...)
- tenir les adequades capacitats per a executar la consulta i/o actualització sense provocar cap desgavell

És molt important disposar en una organització d'usuaris d'aquesta tipologia ja que permet interactuar, en fets puntuals, amb la base de dades, sense necessitat de tenir una aplicació informàtica que ens ho faciliti.

7. Diccionari de dades

Les dades són part fonamental d'una base de dades de manera semblant als passatgers en un tren o avió. És a dir, de res ens servirà una base de dades que es mantingui buida. Concloem, per tant, en que les dades són part fonamental d'una base de dades.

Ara bé, el tren i l'avió no únicament són els passatgers. També hi ha la maquinària que els fa córrer i volar. En les bases de dades, la maquinària és el SGBD. Encara més, en el tren i a l'avió, hi ha uns conductors o pilots que condueixen la maquinària. En les bases de dades, el SGBD és gestionat per l'administrador.

La maquinària del tren i de l'avió segueix un conjunt de regles de funcionament, algunes mecàniques i la majoria, avui en dia, electròniques. Aquestes regles de funcionament estan enregistrades en els circuits electrònics i en les memòries disposades a aquestes finalitats. Els dissenyadors dels aparells (enginyers) són els responsables d'aquestes regles de funcionament.

En els SGBD passa quelcom semblant. El SGBD ha de regir-se per unes *regles* a l'hora de gestionar una base de dades. Aquestes *regles* han estat dissenyades pels analistes i comunicades al SGBD per l'administrador. Les *regles* a què ens referim són el conjunt de definicions dels esquemes (entitats, atributs, dominis, regles d'integritat, restriccions, procediments de validació,...), correspondències entre els esquemes, estratègies d'emmagatzematge físic i mètodes d'accés,... Tota aquesta informació ha d'estar emmagatzemada en algun lloc de manera que el SGBD hi tingui accés.

El **diccionari de dades** és l'espai on resideix el conjunt d'informació necessària per la gestió de la base de dades a càrrec del SGBD.

Cada BD té el seu diccionari de dades. Un SGBD tindrà accés a tants diccionaris de dades com BD gestioni. Cal observar que el diccionari de dades conté la informació que l'administrador de la BD ha hagut de comunicar al SGBD per la gestió de la BD. Recordem quina és aquesta informació a partir de les responsabilitats que tenia l'administrador de la BD:

- Definicions de les entitats, amb els seus atributs, dominis de valors, regles d'integritat, restriccions i procediments de validació.
- Definicions de les correspondències externa/conceptual i conceptual/interna.
- Definicions de l'estratègia d'emmagatzematge i dels mètodes d'accés a utilitzar.

- Definicions d'usuaris, grups d'usuaris, privilegis i grups de privilegis.

Observem que el diccionari de dades és el conjunt de dades necessàries per gestionar les dades de la base de dades. Per aquest motiu, en molts àmbits, s'acostuma a dir que és una **meta-base**⁷.

- **Qui té accés al diccionari de dades?**

La informació del diccionari usualment és introduïda per l'administrador de la BD i es genera a partir de les instruccions executades amb els llenguatges LDD i LCD i amb eines que aporta el mateix SGBD o que subministra, en paral·lel al SGBD, el fabricant del SGBD.

El propi diccionari contempla la possibilitat de permetre diferents tipus d'accés als usuaris que ell mateix té definits per l'accés a la base de dades. Així, es pot permetre als usuaris finals la consulta dels noms de les entitats i dels seus atributs per a facilitar la utilització dels llenguatges LMD i LC. També es pot permetre accés amb més privilegis als analistes i dissenyadors de la base de dades i d'aplicacions sobre la base de dades, ja que potser ells mateixos executaran alguns canvis en l'estructura de la base de dades utilitzant els llenguatges LDD i LCD i les eines existents.

- **Tipus de diccionaris de dades**

Segons la localització del diccionari de dades, classifiquem:

- *Diccionaris de dades independents o aïllats*, consistents en un conjunt d'arxius que es situen independentment de la base de dades de la qual en formen el diccionari i, fins i tot, es gestionen amb eines externes al propi SGBD. En l'actualitat se'n troben pocs.
- *Diccionaris de dades totalment integrats*, que són gestionats pel propi SGBD i formen part de la pròpia base de dades. És el tipus de diccionari emprat en la majoria de SGBD actuals.

8. Arquitectura client-servidor.

En la present unitat didàctica hem introduït un munt de conceptes involucrats en la gestió de bases de dades i que aniran apareixent directament o indirecta al llarg del present crèdit. El món de les bases de dades és, però, molt ampli i en continua expansió motivada per la creixent necessitat d'accedir a la informació i per l'evolució de les telecomunicacions. No hem tractat, fins el moment, els punts d'accés a les bases de dades.

Anomenem **client d'un SGBD** a qualsevol aplicació que permet establir un diàleg amb el SGBD per a accedir a la base de dades. El SGBD es constitueix en **servidor** que dona serveis als clients que ho precisin.

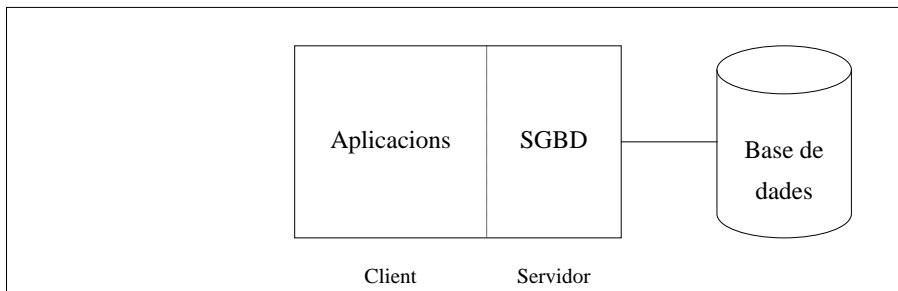
El propi SGBD incorpora clients per a facilitar les tasques de l'administrador i possibilitar la utilització dels llenguatges LMD i LC per part dels dissenyadors d'aplicacions i dels usuaris finals.

⁷ **Prefix meta**

El prefix meta indica posterioritat o més enllà de... o per damunt de... Per aquest motiu s'utilitza per a definir el diccionari de dades, ja que es tracta d'una base de dades utilitzada per gestionar la base de dades. És a dir, és una base de dades que està per damunt de la base de dades que veritablement es vol gestionar.

Així, el producte *Access XX* de la família *Microsoft Office* és un conjunt d'aplicacions clients del SGBD *Microsoft Jet*, tot i que s'acostuma a anomenar *SGBD Access XX*. El SGBD anomenat *Oracle* incorpora el veritable SGBD i moltes eines que en són clients, com *SQL*Plus*.

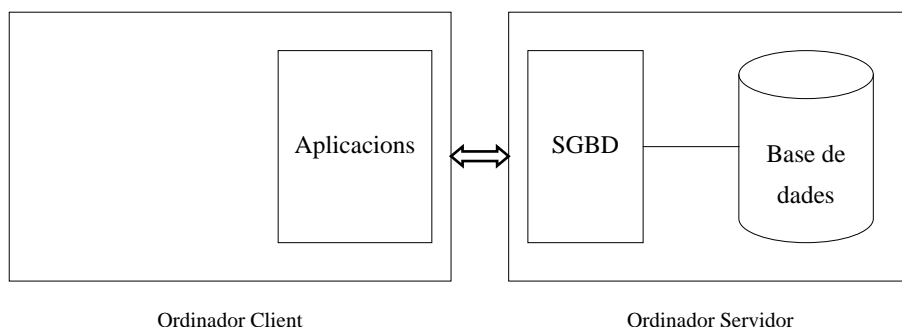
Podem veure el funcionament d'un SGBD i de les aplicacions que en requereixen els seus serveis, sota l'esquema de la dreta.



Aquesta arquitectura client-servidor a nivell de gestió de bases de dades es pot donar en una única màquina (SGBD i aplicacions s'executen en el mateix ordinador) o en diferents màquines. En aquest darrer cas, hi ha diferents possibilitats:

- Un ordinador actuant de servidor de dades (conté la base de dades i el SGBD) i diferents ordinadors actuant de clients (contenen les aplicacions que interactuen amb el SGBD).

Tot i que en l'esquema següent s'observa una única màquina client, la majoria de les vegades hi ha moltes màquines clients i una màquina servidora. En l'actualitat, el concepte d'arquitectura client/servidor s'identifica amb aquesta darrera situació, ja que en tenir separats els ordinadors que contenen les aplicacions -clients- de l'ordinador que conté les dades -servidor-, la distinció és clara i tothom reconeix una *arquitectura client-servidor*.



Cal remarcar, però, que el concepte d'arquitectura client-servidor no s'ha de vincular a la distinció entre màquines, sinó a la distinció entre aplicacions -clients- que utilitzen els serveis d'un SGBD -servidor- per accedir a la(es) base(s) de dades.

En l'esquema anterior s'observa una màquina client que conté aplicacions clients i una màquina servidora que conté el SGBD. Aquesta podria contenir, també, les aplicacions clients (com en l'esquema presentat inicialment) de manera que ella mateixa actués com a client i servidora. De fet, aquesta situació és molt probable, ja que l'administrador de la base de dades utilitza aplicacions clients per a desenvolupar la seva feina les quals segurament resideixen en la mateixa màquina on hi ha el SGBD.

- Variis ordinadors actuant de servidors de dades i diferents ordinadors actuant de clients (contenen les aplicacions que interactuen amb el SGBD).

Aquesta és la situació que es planteja quan la base de dades ha d'estar geogràficament distribuïda entre diferents punts. En tal situació es parla de bases de dades distribuïdes i/o replicades. Aprofundirem el concepte en el següent apartat d'aquesta mateixa unitat didàctica.

Així doncs, el concepte d'*arquitectura client-servidor*, àmpliament difós en la dècada dels noranta, distingeix les aplicacions i el(s) servidor(s) de dades. Com s'arriba a aquesta arquitectura?

Aquesta arquitectura va substituir l'arquitectura existent des dels anys setanta, consistent en grans ordinadors que actuaven de servidors de dades i que contenien les aplicacions a executar. Els usuaris executaven aquestes aplicacions des de terminals (conjunt format per pantalla i teclat connectats a l'ordinador central directament o per línies telefòniques punt a punt). Amb l'evolució, en els anys vuitanta, dels microordinadors, va iniciar-se la substitució dels *terminals no intel·ligents* per microordinadors que possibilitaven la mateixa gestió dels terminals (via programes d'emulació) i, a més, permetien a l'usuari l'execució d'altres tipus de programari de millors prestacions que les que oferien les aplicacions dels grans ordinadors, com per exemple, en el tema ofimàtic.

En generalitzar-se la utilització dels microordinadors en els sistemes informàtics i en assolir aquells unes prestacions cada cop més importants, els fabricants de programari arriben a la conclusió que poden utilitzar la seva potència per a executar les aplicacions clients del SGBD i així alliberar recursos en l'ordinador que conté el SGBD i la base de dades. En aquest moment el concepte d'*arquitectura client-servidor* més amunt detallada deixa d'utilitzar-se a nivell de programari (aplicacions clients del SGBD -aplicació servidora-) per passar a utilitzar-se a nivell de maquinari (màquines clients de màquina servidora). En el fons continua sent vàlida la utilització inicial del concepte, ja que les màquines clients ho són pel fet d'executar les aplicacions clients de l'aplicació servidora que resideix en la màquina servidora.

Amb aquest nou funcionament, les aplicacions clients s'instal·len, en part o en la seva totalitat, en les màquines clients (microordinadors), fet que provoca noves situacions:

- Cal instal·lar i mantenir cada aplicació client en cada màquina client des de la que es vulgui executar, enlloc de tenir-la instal·lada i mantinguda una única vegada en la màquina que conté el SGBD.
- L'usuari de microordinadors té accés a la configuració de la seva màquina i hi instal·la altres aplicacions de propòsit general i específic dins l'organització i, fins i tot, per propòsits personals, les quals poden interferir en el correcte funcionament de les aplicacions clients.

Les dues situacions provoquen l'augment de personal dedicat a mantenir en correcte funcionament el sistema informàtic. Per tot això, l'*arquitectura client-servidor* més amunt detallada i que anomenarem *de dos nivells*, ha evolucionat cap l'estai actual: *arquitectura client-servidor de tres nivells*.

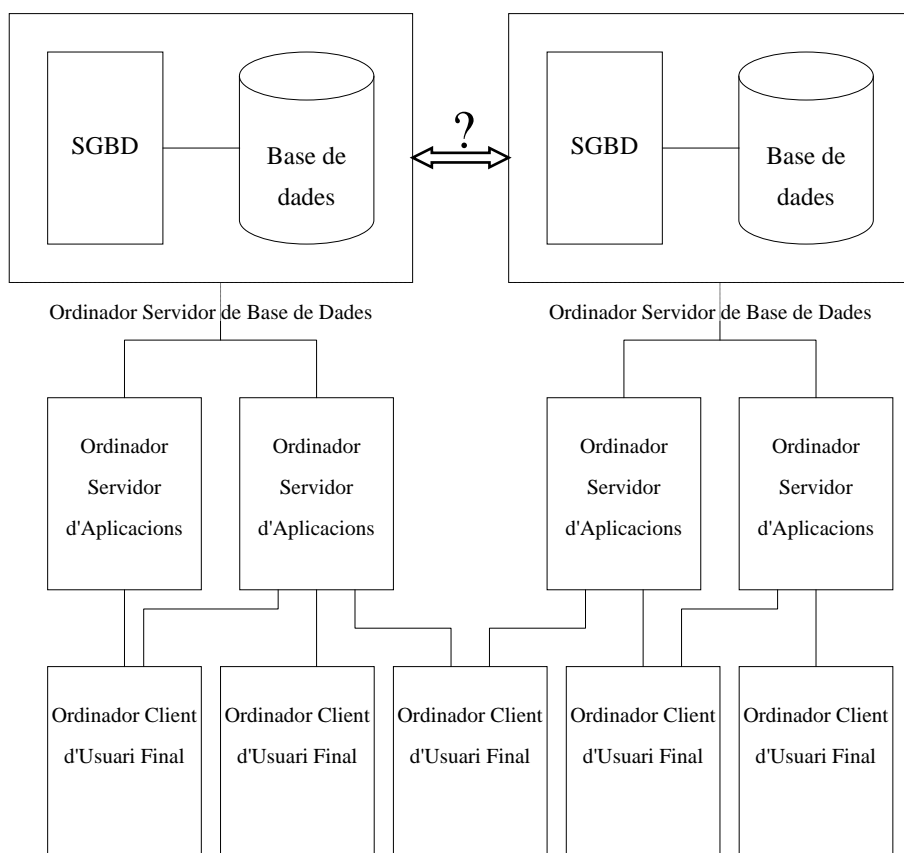
En aquesta nova arquitectura es manté el concepte de servidor(s) de base(s) de dades (una o varies màquines, segons es tracti de base de dades centralitzada o distribuïda o replicada, tal i com veurem en el proper apartat) però les aplicacions clients es distribueixen en dos tipus de màquines: màquines servidores d'aplicacions i màquines d'accés a les aplicacions.

Les màquines servidores d'aplicacions són ordinadors que contenen les aplicacions clients dels SGBD. Una determinada aplicació client pot estar en una o varies màquines servidores d'aplicacions. De manera anàloga al cas anterior, cal fer el manteniment de les diferents màquines servidores d'aplicacions, però aquestes són en número molt inferior al que ens trobàvem en l'altre cas i no hi ha els problemes referents a que els usuaris en modifiquin la configuració o instal·lin altre programari que interactui negativament amb les aplicacions.

I llavors, què tenen els usuaris? Des d'on executen les aplicacions? Doncs des de les seves màquines, amb molta o poca capacitat de procés, les quals utilitzen per accedir als servidors d'aplicacions. És a dir, d'alguna manera es torna al concepte que l'usuari accedeix a les aplicacions per medi de terminals poc intel·ligents. Avui en dia, amb l'explosió de l'accés a internet, interessa que els usuaris tinguin màquines amb qualsevol navegador instal·lat (segons preferències de l'usuari) el qual

permet l'accés -via internet o intranet- a les aplicacions web que es troben en execució en els corresponents servidors d'aplicacions.

En l'esquema següent es pot observar una típica situació d'aquesta arquitectura. En un primer nivell es disposa d'un parell de SGBD que poden estar enllaçats o no (d'aquí l'interrogant en l'enllaç entre elles). En un segon nivell es té quatre servidors d'aplicacions, cadascun dels quals té aplicacions clients d'un dels SGBD. Un servidor d'aplicacions podria tenir aplicacions clients de diferents SGBD situats en diferents servidors de bases de dades. En el tercer nivell estan els ordinadors dels usuaris, els quals accedeixen als servidors d'aplicacions que interessen. L'execució de les aplicacions la duen a terme els servidors d'aplicacions. Les màquines dels usuaris finals únicament funcionen com a interfície entre l'usuari i l'aplicació en execució. Aquest és el típic funcionament actual que l'usuari es troba en accedir via internet a pàgines web des de les que pot comprar, efectuar reserves,... Tot i que aquest funcionament ja existeix en el món internet, també es pot utilitzar en intranets. Una mateixa aplicació es pot estar executant en diferents servidors d'aplicacions. L'usuari final establirà connexió amb el servidor d'aplicacions pel que tingui un accés més ràpid i directe.



9. Bases de dades distribuïdes i replicades

Aquest concepte ja ha aparegut en l'anterior apartat. En moltes ocasions serà necessari i/o convenient tenir les dades d'una mateixa base de dades repartides i/o repetides en diferents punts. En tal situació, cada punt ha de tenir el seu SGBD per a gestionar les corresponents dades i els diferents SGBD han d'estar connectats i treballar de forma conjunta.

Una **base de dades distribuïda i/o replicada** és un conjunt de dades localitzades en diferents màquines, cadascuna de les quals pot participar en l'execució de les transaccions que gestionen dades d'un o varis conjunts.

Cadascuna de les màquines on hi ha part de la base de dades és un **node** i ha de tenir un SGBD per a la gestió de les seves dades. Els diferents nodes han de formar part d'una xarxa d'àrea local o d'una xarxa d'àrea estesa o d'una barreja d'ambdues situacions.

No es pot confondre el concepte **node** amb el concepte **punt d'accés** que s'utilitza per fer referència al punt des del que es sol·liciten transaccions sobre la base de dades i que, com molt bé podeu deduir, tan pot estar situat en el propi node com en estacions de treball remotes.

Quan es vol remarcar que una base de dades no és distribuïda ni replicada, sinó que tota ella es troba emmagatzemada en una màquina sota la supervisió d'un únic SGBD es pot dir que es tracta d'una base de dades **centralitzada**.

Notarem **distribució de dades** quan les dades es troben repartides en diferents nodes i notarem **rèplica de dades** quan les dades es troben repetides en diferents nodes, provocant redundància.

Una base de dades pot ser distribuïda i replicada a la vegada, és a dir, pot tenir dades repartides entre diferents nodes (distribució) i dades repetides en diferents nodes -alguns o tots- (rèpliques). Tot i que ja deveu estar intuïnt els problemes que pot suposar tenir una base de dades replicada pel fet de les actualitzacions de les dades repetides, la utilització de rèpliques pot estar justificada per aquelles dades que s'utilitzen molt en nodes diferents.

Les bases de dades distribuïdes i/o replicades poden tenir el mateix o diferent SGBD en els diferents nodes. Es necessita, però, un SGBD global (SGBDG) encarregat d'organitzar les comunicacions entre els diferents SGBD. La problemàtica del SGBDG varia molt segons els diferents SGBD parlin o no el mateix llenguatge. Per tant, l'interès radica no en el fet que els diferents nodes es regeixin pel mateix SGBD sinó en el fet que els diferents SGBD tinguin uns llenguatges semblants o, si més no, semblants en les parts on s'ha d'establir comunicació. Així doncs, distingirem:

- **SGBDG homogènis**, on els SGBD de tots els nodes disposen de llenguatges semblants.
- **SGBDG heterogènis**, on els nodes utilitzen SGBD amb llenguatges força diferents. En aquesta situació, la gestió del SGBDG és força complicada.

Degut a que la informació es troba distribuïda en diferents nodes i que des de cadascun d'ells es pot gestionar dades del seu node o dades d'altres nodes, apareixen els següents tipus de transaccions:

- **Transaccions locals**, que gestionen dades situades en el propi node.
- **Transaccions globals**, que gestionen dades situades en nodes diferents.

Per finalitzar aquesta breu introducció a la temàtica de les bases de dades distribuïdes i/o replicades, fem esment a com es produeix el naixement d'una tal base de dades, el qual pot ser fruit de:

- Disseny inicial que preveu la distribució i/o rèpliques de dades entre diferents nodes: **base de dades autònoma**.
- Agrupament de diferents bases de dades independents, existents amb anterioritat, majoritàriament centralitzades però amb possibilitat inicial de distribució i/o rèpliques en alguna d'elles, que s'han de continuar gestionant de manera relativament independent: **base de dades federal**.

La distinció entre bases de dades autonòmiques i federals és independent, en principi, del fet que el SGBDG sigui homogeni o heterogeni. La realitat és, però, que una base de dades federal té moltes possibilitats de tractar-se amb un SGBDG heterogeni, mentre que en una base de dades autonòmica hi ha més marge per escollir el SGBDG a utilitzar i per simplicitat sembla lògic escollir un SGBDG homogeni.

9.1. Possibilitats de distribuir i/o replicar les dades

En treballar amb bases de dades distribuïdes i/o replicades es pot tenir diferents tipologies segons sigui el repartiment que de les dades se'n fa en els diferents nodes. Les possibilitats existents són:

- **Amb multiplicació.** Totes les dades estan duplicades en cada node amb la finalitat de reduir els costos de comunicació i millorar el temps d'execució, ja que cada operació de consulta es pot executar en cada node. Presenta l'inconvenient de la quantitat d'espai necessari per tenir les dades duplicades, la redundància extrema entre diferents nodes i els problemes d'actualització de les dades en els diferents nodes. Aquesta tipologia no s'acostuma a utilitzar.
- **Amb partició.** Consisteix en que les dades es troben repartides entre diferents nodes sense cap tipus de rèplica. D'aquesta manera, qualsevol actualització sobre l'entitat s'ha d'executar únicament en el node que la conté. Presenta els inconvenients del cost de temps en efectuar operacions que afecten a dades de diferents nodes i de la caiguda total de la base de dades quan un node es queda sense servei.
- **Amb un node que conté tota la base de dades i altres nodes que en contenen parts replicades.** Una organització d'aquest estil es justifica per reduir al màxim els costos de comunicació en les transaccions globals. Presenta l'inconvenient dels processos d'actualització en les parts replicades. La base de dades pot quedar sense servei si cau el node que conté tota la base de dades.
- **Amb informació replicada en nodes seleccionats.** Consisteix en tenir la informació emmagatzemada, majoritàriament, en els nodes des dels que s'hi accedeix, és a dir, la informació accedida des de varis nodes estarà ubicada en tots ells (rèpliques), mentre que la informació accedida pràcticament des d'un únic node, estarà emmagatzemada únicament en ell. Aquest tipus d'organització continua tenint els inconvenients de l'actualització de les dades replicades i de l'accés a dades que es troben en diferents indrets.

9.2. Fragmentació

En l'apartat anterior hem estat presentant diferents possibilitats de combinar els conceptes de *repartir* i *replicar*. En totes elles hem parlat genèricament de *dades*, és a dir, *repartim i/o replicàvem dades*.

Ja sabem que en una base de dades hi representarem entitats amb els seus atributs (com en els arxius) i a més les integrarem i compartirem. Ara, però, estem repartint i/o replicant les dades entre diferents nodes. Ens apareixen les preguntes: *què entenem per dada repartida i/o replicada?*, *"dada" és igual a "entitat"?*, *"dada" és igual a "conjunt d'entitats"?*, *"dada" és igual a "part d'una entitat"?*,... La resposta no és única. Hi ha diferents possibilitats.

Anomenem **fragment** d'una entitat distribuïda i/o replicada a cada subconjunt d'informació que es distribueix i/o replica.

Hi ha diferents possibilitats d'escollir els fragments d'una entitat. Per exemplificar-les, considerem l'entitat *persones* amb les següents instàncies:

Dni	Nom	Data Naixement	Altura	Pes	Localitat	Comarca
450	Marius	09-12-1980	169	65	Igualada	Anoia
870	Empar	19-03-1936	165	70	Barcelona	Barcelonès
920	Manel	25-04-1967	168	74	Jorba	Anoia
225	Guifré	07-02-1995	100	16	Igualada	Anoia
342	Josep	09-12-1960	175	85	Olesa de Montserrat	Baix Llobregat
425	Regi	05-12-1992	105	18	Olesa de Montserrat	Baix Llobregat
972	Manela	24-11-1932	155	65	Igualada	Anoia
782	Elvira	09-01-1965	165	60	Barcelona	Barcelonès
645	Joan	07-01-1924	174	75	Barcelona	Barcelonès
123	Benet	03-08-1931	158	50	Igualada	Anoia
500	Maite	27-01-1958	165	48	Igualada	Anoia

1) **Fragmentació horitzontal**, on els fragments són conjunts d'instàncies de l'entitat.

Situant-nos en l'exemple, si suposem que cada comarca té un node de la base de dades, podríem arribar a la conclusió de fragmentar l'entitat *persones* de manera que cada comarca emmagatzemi les "seves" persones. Així, tindríem tres fragments:

Dni	Nom	Data Naixement	Altura	Pes	Localitat	Comarca
450	Marius	09-12-1980	169	65	Igualada	Anoia
920	Manel	25-04-1967	168	74	Jorba	Anoia
225	Guifré	07-02-1995	100	16	Igualada	Anoia
972	Manela	24-11-1932	155	65	Igualada	Anoia
123	Benet	03-08-1931	158	50	Igualada	Anoia
500	Maite	27-01-1958	165	48	Igualada	Anoia

Dni	Nom	Data Naixement	Altura	Pes	Localitat	Comarca
870	Empar	19-03-1936	165	70	Barcelona	Barcelonès
782	Elvis	09-01-1965	165	60	Barcelona	Barcelonès
645	Joan	07-01-1924	174	75	Barcelona	Barcelonès

Dni	Nom	Data Naixement	Altura	Pes	Localitat	Comarca
342	Josep	09-12-1960	175	85	Olesa de Montserrat	Baix Llobregat
425	Regi	05-12-1992	105	18	Olesa de Montserrat	Baix Llobregat

2) Fragmentació vertical, on els fragments són conjunts de valors d'atributs de l'entitat.

Situant-nos en l'exemple, podríem considerar que en un cert node hi ha les dades corresponents a la localització de les persones (dni, nom, localitat., comarca) i en un altre node hi ha dades mèdiques de les persones (dni, nom, data de naixement, pes, altura). Així, tindríem dos fragments:

Dni	Nom	Data N.	Alt	Pes	Dni	Nom	Localitat	Comarca
450	Marius	09-12-1980	169	65	450	Marius	Igualada	Anoia
870	Empar	19-03-1936	165	70	870	Empar	Barcelona	Barcelonès
920	Manel	25-04-1967	168	74	920	Manel	Jorba	Anoia
225	Guifré	07-02-1995	100	16	225	Guifré	Igualada	Anoia
342	Josep	09-12-1960	175	85	342	Josep	Olesa de Montserrat	Baix Llobregat
425	Regi	05-12-1992	105	18	425	Regi	Olesa de Montserrat	Baix Llobregat
972	Manela	24-11-1932	155	65	972	Manela	Igualada	Anoia
782	Elvis	09-01-1965	165	60	782	Elvis	Barcelona	Barcelonès
645	Joan	07-01-1924	174	75	645	Joan	Barcelona	Barcelonès
123	Benet	03-08-1931	158	50	123	Benet	Igualada	Anoia
500	Maite	27-01-1958	165	48	500	Maite	Igualada	Anoia

Evidentment hi ha d'haver algun atribut identificador que permeti enllaçar les parts separades de cada instància. En el nostre exemple suposem que és l'atribut *dni*. Observem, també, que un mateix atribut pot estar repetit en diferents fragments, com succeeix amb l'atribut *nom*.

3) Fragmentació mixta, que consisteix en considerar fragments verticals, algun dels quals poden, a la vegada, estar fragmentats horitzontalment.

Per finalitzar, notem que els tres tipus de fragmentació presentats són combinables amb les possibilitats de distribuir i replicar una base de dades.

9.3. Avantatges i inconvenients

En els apartats anteriors ja s'han anat intuït les avantatges i inconvenients de la gestió d'una base de dades distribuïda. Anem a resumir-les.

1) Entre les **avantatges**, cal considerar:

a) Reduir el cost en les comunicacions, ja que la majoria d'operacions es suposa que es realitzaran, si la base de dades està ben distribuïda, accedint a les dades del node local.

b) Certa autonomia en el node local, tot i que no es pot obviar que es forma part d'una organització superior.

c) Possibilitat de continuar el treball amb les dades emmagatzemades en el node local quan es produeix una caiguda d'algun altre node. En restablir-se el funcionament d'aquest, caldrà reintegrar-se a la xarxa amb el mínim de complicacions possible.

2) Entre els **inconvenients**, cal considerar:

a) Cost del programari, ja que el SGBDG ha de ser més potent

b) Actualització de les dades duplicades

c) Major possibilitat d'errors com a conseqüència d'una major concurrència des de diferents nodes a una mateixa dada.

d) Aconseguir la transparència d'ubicació de cara als diferents usuaris (aplicacions, dissenyadors i programadors d'aplicacions, usuaris finals).

e) Increment del temps de procés en les transaccions globals.

9.4. Gestió de bases de dades distribuïdes

En els següents apartats volem presentar una aproximació del funcionament intern que es produeix en el treball en bases de dades distribuïdes, des del moment en que l'usuari provoca una consulta i/o actualització a la base de dades (ja sigui directament via LMD-LC autosuficient o des d'una aplicació que utilitza LMD-LC hostatjat) fins el moment en que l'usuari rep la resposta proporcionada pel sistema gestor de base de dades.

9.4.1. Transaccions globals i locals

En un apartat anterior hem introduït la distinció entre transacció global i local, diferenciant-les en funció dels nodes als que han d'accedir per donar resposta a la sol·licitud de l'usuari.

Davant una transacció sobre una base de dades distribuïda (provocada per una sol·licitud d'usuari en un determinat punt d'accés), el SGBDG reconeix si es tracta d'una transacció global (que ha d'accedir a diferents nodes) o d'una transacció local (per un node en concret). En aquest darrer cas passarà la transacció al SGBD del node afectat per a que actuï en conseqüència i n'esperarà el corresponent resultat. Si es tracta, però, d'una transacció global, el SGBDG la descompon en subtransaccions de les que n'esperarà els corresponents resultats per tal de poder confeccionar el resultat final per la transacció inicial.

Cadascuna de les subtransaccions pot ser, a la vegada, global (encara intervenen diferents nodes) o local per un node en concret. El SGBDG torna a iniciar la gestió de les globals i traspasa les locals als corresponents SGBD. Els resultats corresponents a cada subtransacció es generen amb les dades a les que s'accedeix directament i amb els resultats d'altres subtransaccions que la pròpia subtransacció havia generat. D'aquesta manera s'arriba a confeccionar el resultat de la transacció inicial.

Si la transacció executa una actualització de dades existents en fragments replicats en altres nodes, es genera una subtransacció per actualitzar les rèpliques afectades.

En cas de transacció d'actualització que afecti a diferents nodes caldrà assegurar que l'actualització es produeixi totalment o s'anul·li en la seva totalitat, de manera que si s'avorta una subtransacció que actualitzava, també s'avortin la resta de

subtransaccions. Una excepció a aquest fet pot permetre's en les subtransaccions d'actualització a rèpliques, les quals es permet, en certes situacions que s'efectuïn en temps diferit⁸.

9.4.2. Arquitectura d'un SGBD distribuït

En un apartat anterior hem presentat l'arquitectura de tres nivells (intern, conceptual i extern) com l'arquitectura adoptada pels sistemes gestors de bases de dades. Hem vist, també, que ens proporcionava la independència física i lògica de les dades.

Com s'ha de modificar aquesta arquitectura en un SGBD distribuït?

La descripció del contingut de tota la base de dades distribuïda i l'especificació de quins fragments hi ha a cada node forma el que s'acostuma a anomenar **esquema global**, el qual haurà d'estar emmagatzemat en cada punt d'accés, ja que cal tenir-hi accés ràpid per poder descompondre cada transacció en subtransaccions.

L'arquitectura és sensiblement diferent segons es tracti d'una base de dades autonòmica o federal.

1) Base de dades autonòmica

En una base de dades autonòmica el disseny de la base de dades s'efectua preveient que les dades estaran distribuïdes i/o replicades en diferents punts d'emmagatzematge. Per tant, en la fase de disseny es defineixen:

- L'esquema conceptual, corresponent a tota la base de dades.
- L'esquema global, que conté l'esquema conceptual més els criteris de distribució i replicació.

Aquesta informació determina els esquemes conceptuals dels diferents nodes, els quals es podran generar de manera quasi automàtica.

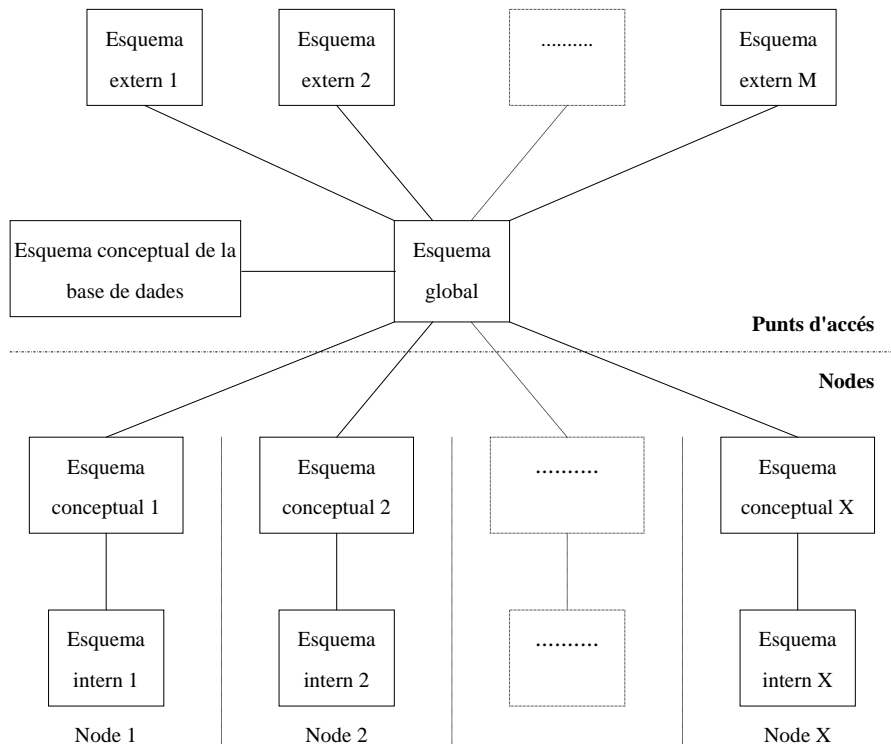
Els esquemes interns es defineixen per cada node en funció, també, dels criteris establerts en l'esquema global.

Els esquemes externs es defineixen a partir de l'esquema global, a diferència d'una base de dades centralitzada en la que es defineixen a partir de l'esquema conceptual.

En la figura següent es pot observar un cas de X nodes en una base de dades autonòmica, gestionada sota un esquema global que es correspon amb l'esquema conceptual de la base de dades. En tal situació, els M esquemes externs són dissenyats sobre l'esquema global.

⁸ **Rèpliques discordants**

En certes situacions de bases de dades distribuïdes es permet que fragments replicats no estiguin actualitzats. En dites situacions, hi ha subtransaccions pendents d'execució les quals s'executen en temps diferit.



Observem que els punts d'accés han de tenir la informació referent a l'esquema global i als esquemes externs que es pot accedir. Una transacció sol·licitada en un punt d'accés serà tractada dins l'esquema extern que correspongui i descomposta en el propi punt d'accés en diverses subtransaccions, cadascuna de les quals serà adreçada al node que pertorqui. Cada node crearà un esquema extern per la subtransacció rebuda, de manera que el SGBD del node actuarà com si es tractés d'una transacció local.

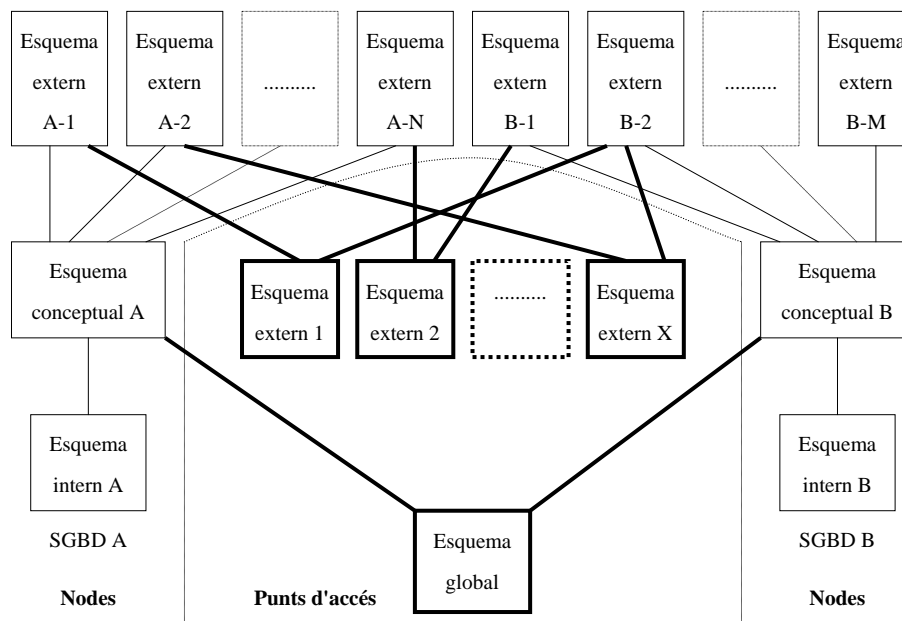
2) Base de dades federal

En una base de dades federal es parteix de diferents bases de dades ja existents. Per tant, cadascuna ja té els seus esquemes conceptual, intern i externs. En tal situació, l'arquitectura del SGBD distribuït és més complicada.

L'esquema global es construeix a partir dels esquemes conceptuals de les diferents bases de dades que es federen. Apareixen, segurament, conflictes relatius a sinònims i homònims, que es s'acostumen a resoldre afegint un prefix que indiqui la base de dades d'origen.

Els esquemes externs es podrien definir a partir dels esquemes externs existents a les diferents bases de dades que es federen.

Realitzar un dibuix que esquematitzi la situació plantejada no és tan simple com en el cas de la base de dades autonòmic. Mostrem, a la pàgina següent, un exemple, pel cas de una federació de dues bases de dades.



9.4.3. Components d'un SGBD distribuït

És de calaix que en un SGBD distribuït es necessita un gestor de comunicacions, és a dir, un sistema en xarxa, que garanteixi que tots els missatges arriben al seu destí en l'ordre en què s'han anat produint i sense duplicació. Considerant assolida aquesta situació, tenim com a components propis del SGBD:

- En els **punts d'accés**:
 - Els **catàlegs**, on s'emmagatzema l'esquema global i els esquemes externs vinculats al punt d'accés.
 - El **processador de transaccions font**, que comprova que només accedeixin a la xarxa els usuaris autoritzats i obtinguin la informació requerida, així com en quins punts es troben les dades demanades. A més, descompon cada transacció en subtransaccions tenint en compte la màxima reducció d'utilització de les línies de comunicacions, compila cada subtransacció totalment o parcial i li genera un pla d'execució.
 - El **gestor de transaccions**, que controla a nivell global el pla d'execució de cada transacció.
- En els **nodes**:
 - El SGBD que correspongui
 - El **gestor de subtransaccions**, que rep les transaccions, les acaba de compilar si és menester, i les executa. En finalitzar la seva execució, ho comunica al gestor de transaccions del punt d'accés que va enviar la transacció i envia els resultats obtinguts als nodes explicitats en el pla d'execució.

UD2. Disseny de bases de dades i model entitat - relació

1. L'esquema conceptual i el model entitat - relació

En l'anterior unitat didàctica hem presentat l'arquitectura de tres nivells com l'arquitectura adoptada per la majoria de SGBD. Es basava en la distinció dels esquemes conceptual, intern i externs.

Un bon disseny de base de dades s'inicia per l'esquema conceptual. Cal descriure les dades de totes les àrees d'activitat de l'organització amb les relacions que s'estableixen entre elles, independentment de la implementació física que se'n farà posteriorment i, que en cas d'utilitzar un SGBD, podrà seguir el model jeràrquic, el model en xarxa o el model relacional.

És molt important remarcar que el disseny de l'esquema conceptual que presentem en aquesta unitat didàctica és convenient efectuar-lo encara que es conegui que la posterior implementació s'efectuarà sobre un SGF.

Els textos destinats a l'*Anàlisi i Disseny d'Aplicacions Informàtiques i/o de Bases de Dades* remarquen la importància de descriure què cal fer abans d'especificar com fer-ho.

I és el que pretenem en aquesta unitat didàctica. Cal efectuar un eficaç disseny de l'esquema conceptual, es qual sorgirà com el producte d'una eficient anàlisi de l'organització i/o situació a informatitzar.

Què entenem per *eficaç disseny*? Volem aconseguir un **model** que ens defineixi l'esquema conceptual. Hi ha diferents maneres de modelar l'esquema conceptual. En aquesta unitat didàctica presentem un model gràfic per l'esquema conceptual: el **model entitat - relació (ER)**

El model ER va ser desenvolupat, en la dècada dels 70, per Sh. Chen⁹ i per aquest motiu és reconegut en molts textos com **model Chen**. Gira al voltant de dos conceptes fonamentals en el món de les representacions: **entitats i relacions**.

1.1. Entitats

Retrobem el concepte d'entitat i els components que el conformen: atribut, valor, domini, valor nul, clau, identificador,...

Una **entitat** és "quelcom" real, identificable i distingible.

Així, tota persona és una entitat, com també ho és un gos, una casa, un cotxe,... Podem agrupar les entitats d'un mateix tipus sota un nom comú i llavors es parla de l'*entitat persones*, *entitat gossos*, *entitat cases*, *entitat cotxes*. En tal cas, quan parlem d'un element concret d'una entitat acostumen a utilitzar el concepte d'**instància**.

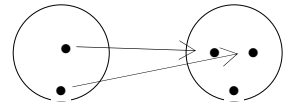
⁹ Referència bibliogràfica als treballs de Chen

P. P. - Sh. Chen (març del 1976). The Entity-Relationship Model - Toward a Unified View of Data. ACM Transactions on Database Systems. Vol. 1. N. 1. Pàgs. 9-36.

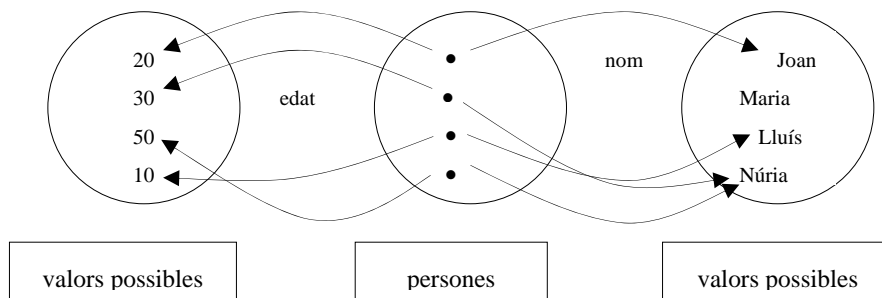
Un **atribut** és una aplicació d'una entitat (conjunt d'elements del mateix tipus) sobre un conjunt de valors.

Recordatori matemàtic

Una aplicació és una correspondència entre dos conjunts de manera que cada element del conjunt origen té una i només una imatge en el conjunt destí, tal i com s'observa en l'exemple del marge.



Considerem la situació de l'exemple:



Observeu l'entitat *persones*, que conté quatre instàncies (els punts negres). Cada persona és "quelcom" real, distingible i identificable. Observeu els dos atributs: *edat* i *nom*. En podríem considerar d'altres: *dni*, *direcció*, *pes*,...

L'atribut *pes* és una aplicació del conjunt de les persones sobre el conjunt de valors possibles com a pes. Convindreu que el conjunt de valors ha de ser, en aquest cas, un conjunt de nombres naturals.

L'atribut *nom* és una aplicació del conjunt de les persones sobre el conjunt de valors possibles com a nom, és a dir, el conjunt de noms de persones.

Observeu que hi pot haver noms que no siguin la imatge de cap persona (*Maria*) i altres que poden ser la imatge de diverses persones (*Núria*). El mateix pot succeir amb l'atribut *edat* tot i que en aquesta situació concreta no es produeix.

Anomenem **domini** d'un atribut al conjunt de valors possibles que pot prendre l'atribut, és a dir, que poden ser imatge d'alguna entitat per l'atribut considerat.

Així, en l'exemple anterior tenim clar que l'atribut *edat* té per domini el conjunt dels números naturals, ja que podem convenir que l'edat de les persones ha de ser numèrica, no negativa i sense decimals (això ja és més discutible...). Podem afinar més i considerar que l'edat de les persones pot tenir un límit superior, que podríem fixar en 200. Us sembla correcte?

Per altra banda, l'atribut *nom* té per domini el conjunt de noms possibles en les persones. Evidentment és el conjunt de cadenes constituït per les cadenes que tenen sentit com a nom de persones.

Hi ha diferents motius que causen l'existència d'un **valor nul**:

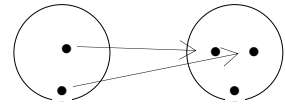
- Quan el valor de l'atribut d'una entitat és desconegut, com per exemple, l'*edat* d'una *persona*.
- Quan l'atribut no és aplicable sobre una instància concreta d'una entitat, com per exemple, *quants parts* ha tingut una persona quan la persona és de sexe masculí.

D'aquesta manera aconseguim mantenir el concepte d'atribut ja que en haver de ser una aplicació, tota instància ha de tenir un valor com a imatge per l'atribut. En les situacions anteriors es considera el *valor nul*.

Un atribut és **identificador** quan l'aplicació és injectiva.

Recordatori matemàtic

Una aplicació injectiva és una aplicació en la que cada element del conjunt destí és imatge d'un o cap elements del conjunt origen, com es pot observar en l'exemple del marge.



El concepte d'atribut **identificador** es correspon amb el fet que l'atribut no pugui tenir valors repetits. És un **concepte semàntic**, és a dir, no s'és atribut identificador en funció del cas particular que s'estigui tractant sinó del significat que l'atribut porta en sí mateix.

En el cas particular del darrer exemple, l'atribut *edat* es comportava com a *identificador* doncs no hi havia dues persones que tinguessin la mateixa edat. Ara bé, sabem que aquesta situació no és la que porta vinculada l'atribut *edat*, ja que de tots és conegut que hi ha moltes persones amb la mateixa edat. Per tant, l'atribut *edat* no és identificador. Seguint amb el mateix exemple, en la situació particular presentada s'observa que l'atribut *nom* no és identificador ja que *Núria* és el nom de dues persones.

És molt important poder tenir, en tota entitat, un -com a mínim- atribut identificador, ja que d'aquesta manera tenim l'entitat totalment identificada, valgui la redundància.

En el cas de l'entitat *persones*, quin identificador podem considerar? Cal tenir en compte l'entorn en el que s'està parlant. Hi ha la costum d'utilitzar el *dni* com a identificador de les persones. És correcte? Depèn! Si estem parlant de persones a nivell mundial, no és correcte, doncs el *dni* és un concepte de l'estat espanyol. Per altra banda, si pretenem gestionar persones en les que hi pot haver infants, no és correcte, doncs en l'actualitat els infants no estan obligats a tenir adjudicat un *dni*. I, per finalitzar, si anem a gestionar persones per encàrrec de qualsevol de les administracions de l'estat espanyol, el *dni* tampoc és correcte com a identificador, ja que hi ha hagut errors en l'adjudicació i hi ha valors repetits. En resum, cal fer un bon estudi sobre un atribut per a poder-lo prendre com a identificador.

I si no disposem de cap atribut que per sí sol sigui identificador? Intentarem, en tal cas, utilitzar el concepte més genèric de **clau**.

Una **clau** és un o varis atributs que conjuntament identifiquen, de manera unívoca, una entitat.

Com a exemple, suposem que volem identificar les ciutats amb el seu nom. De tots és conegut que això no és possible doncs hi pot haver ciutats amb el mateix nom en diferents països. Fins i tot dins el mateix país. Oblidem-nos d'aquesta darrera possibilitat. Suposem que dins un país, el nom de la ciutat és atribut identificador. Però hem de gestionar ciutats de diferents països. Quin identificador prendre? Podem considerar la clau formada per la parella *nom de país* i *nom de ciutat*, parella que identifica tota ciutat.

Fixeu-vos que el concepte de clau també engloba el concepte d'atribut identificador, ja que una clau pot estar formada per un únic atribut.

Cal distingir els següents tipus de claus:

- Clau **candidata**: Qualsevol de les claus possibles sobre una entitat.
- Clau **mínima**: Clau que no té atributs redundants, és a dir, que tot atribut és imprescindible per a que sigui una clau.
- Clau **redundant**: Clau que té atributs redundants, és a dir, que té algun atribut no imprescindible per a que sigui una clau.
- Clau **primària**: Clau **mínima** escollida per identificar l'entitat. Només n'hi ha una.
- Clau **alternativa**: Tota clau candidata mínima no primària.

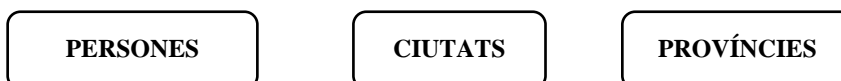
No cal dir que cap atribut d'una clau pot tenir definit el valor nul.

En presentar el model ER hem comentat que es tractava d'un model gràfic. Manca, per tant, donar la notació gràfica que s'utilitza per representar les entitats i els seus components.

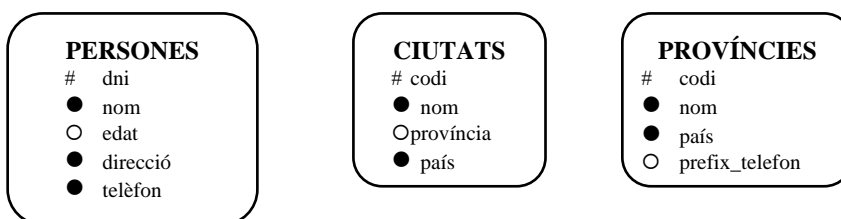
Una entitat es representa mitjançant un rectangle que conté, en el seu interior, el nom, en majúscules, de l'entitat que representa. Així, tindríem:



Alguns textos, però, utilitzen un caixa arrodonida:



Com molt bé hem vist, les entitats poden tenir molts atributs, els quals es poden incorporar (tots o alguns) en la representació gràfica. Al respecte també hi ha diverses maneres de fer. En alguns textos els atributs es posen en algun costat de la caixa que representa l'entitat. En altres textos es posen en l'interior de la caixa. Nosaltres optarem per aquesta darrera. Així:



Els atributs s'escriuen, normalment, en minúscules. Van precedits per un símbol que indica:

- # Atribut obligatori i que forma part de la clau primària
- Atribut obligatori, és a dir, no admet valors nuls.
- Atribut no obligatori, és a dir, admet valor nul.

En molts textos, els atributs que componen la clau primària van subratllats.

1.2. Relacions

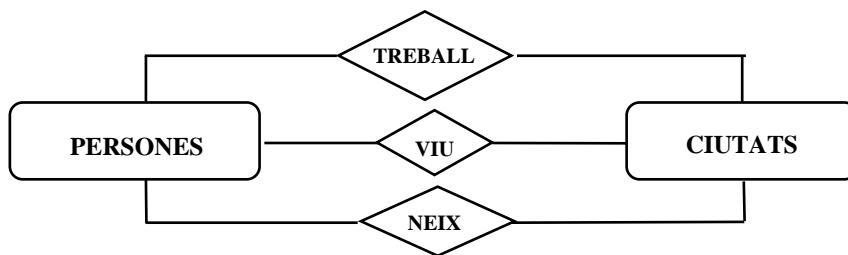
L'altre concepte bàsic en el model ER és el de relació entre entitats.

Anomenem **relació entre entitats** o **inter-relació** a qualsevol connexió existent entre dues o més entitats.

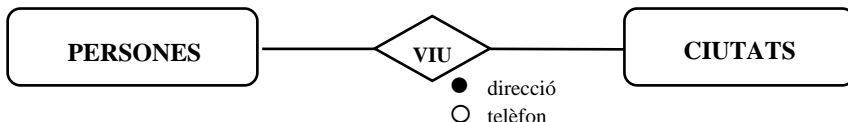
Entre diverses entitats es pot establir diverses relacions. Així, per exemple, considerem, les entitats *PERSONES* i *CIUTATS*, les quals poden estar relacionades sota diferents conceptes: *lloc de residència*, *lloc de treball*, *lloc de naixement*,... Cadascun d'aquests conceptes dona lloc a una relació diferent entre les entitats *PERSONES* i *CIUTATS*.

Una relació és, generalment, un **verb**. En l'exemple anterior observem que *Josep viu a Palau de Plegamans*, *Montserrat treballa a Cadaqués*, *Eva va néixer a Sitges*.

Les relacions entre les entitats s'acostumen a representar gràficament mitjançant un rombe que conté un nom que identifica la relació (probablement el **verb**) i línies que uneixen aquest rombe amb les entitats connectades per la relació. En l'exemple anterior:



Les relacions poden tenir atributs. Gràficament es situen al costat del corresponent rombe. Així, en la relació *VIU* podem considerar els atributs *direcció* i *telèfon*.



Els atributs de les relacions poden admetre l'existència d'un valor nul. Per aquest motiu cal indicar si són o no obligatoris. Ara bé, les relacions no tenen *clau*, ja que el que representen no existeix per sí mateix sinó que és fruit de l'existència de les entitats que connecta.

En funció del nombre d'entitats connectades per una relació, tenim:

- Relacions **reflexives**, quan es tracta de relacions que connecten una entitat amb sí mateixa.
- Relacions **binàries**, quan es tracta de relacions que connecten dues entitats.
- Relacions **ternàries**, quan es tracta de relacions que connecten tres entitats.

I, evidentment, podríem continuar anomenant les relacions segons el nombre d'entitats que hi intervenen.

Per finalitzar, una advertència. El model ER no contempla la connexió entre relacions ni entre relacions i entitats. És a dir, una relació només connecta entitats. Una relació mai pot ser part d'altres relacions.

1.2.1. Grau de les relacions

Una relació entitats es pot *llegir* des de tants punts de vista com entitats connectades per la relació. Exemplifiquem-ho.

- Suposem una relació entre dues entitats A i B. Podem llegir-la de dues maneres: d'A cap a B o de B cap a A.

Així, la relació binària entre PERSONES i CIUTATS definida per VIU es pot veure des de dos punts de vista:

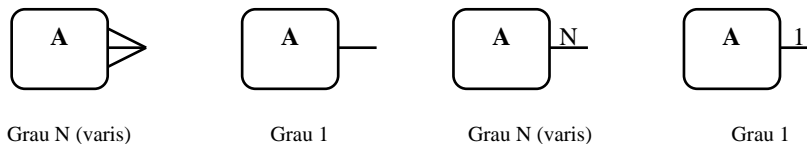
- Associació de CIUTATS cap a PERSONES, de manera que cada ciutat (instància de l'entitat CIUTATS) pot ser lloc on viuen diverses persones (instàncies de l'entitat PERSONES).
- Associació de PERSONES cap a CIUTATS, de manera que cada persona té una ciutat com a lloc on viure.
- Suposem una relació entre tres entitats A, B i C. Hi ha tres possibles lectures: d'A cap a parelles B-C, de B cap a parelles A-C i de C cap a parelles A-B.

Considerem, per exemple, les entitats CLIENTS, ARTICLES i COMANDES relacionades sota el concepte de comanda de venda. Tenim tres lectures:

- Hi pot haver varies comandes on un client hagi sol·licitat un article.
- Hi pot haver un o varis articles sol·licitats per un client en una comanda.
- Hi ha un únic client que ha demanat articles dins una comanda.

Anomenem **grau d'una relació entre entitats** a la relació numèrica que s'estableix entre les instàncies de les entitats connectades per la relació. S'utilitza únicament dues possibilitats: **1** o **N**

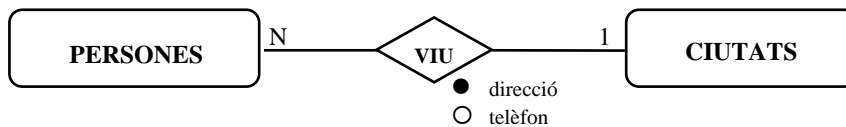
Aquest concepte es representa gràficament aprofitant la línia que connecta el rombe amb les diverses entitats i indicant-ho en l'extrem que llinda amb l'entitat. S'observen diverses possibilitats utilitzades en diversos textos



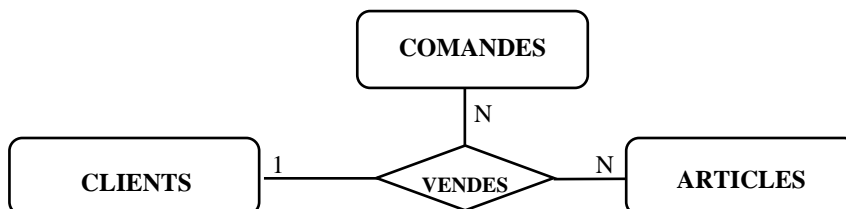
Hi ha textos en els que s'utilitza la notació de les línies i textos en els que s'utilitza la notació dels símbols. En aquest text acostumarem a utilitzar la representació dels símbols.

Seguint amb els anteriors exemples:

- La relació VIU és una relació binària N a 1 entre PERSONES i CIUTATS, ja que cada persona viu en 1 ciutat i en una ciutat hi poden viure N persones.

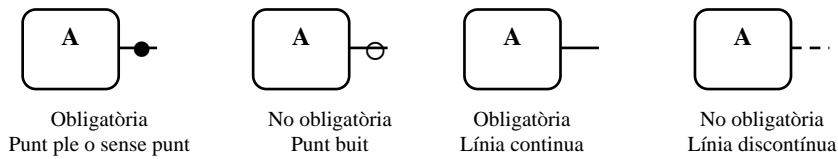


- La relació VENDES és, tal i com s'ha vist abans, una relació ternària 1:N:N entre CLIENTS, ARTICLES i COMANDES.



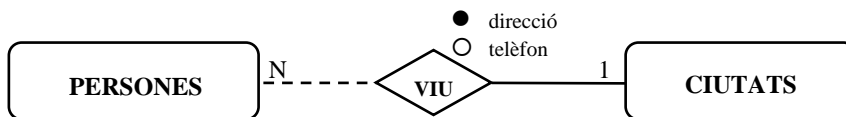
1.2.2. Obligatorietat de participació

Una relació s'estableix entre entitats, però el grau de participació de les instàncies de les entitats connectades pot ser obligatori o no obligatori. Aquest fet es representa gràficament de diverses formes:



Hi ha textos en els que s'utilitza la notació dels punts ple (o sense punt) i buit i textos en els que s'utilitza la notació de les línies contínua i discontinua. En aquest text acostumarem a utilitzar la representació de les línies.

En l'exemple que venim presentant:



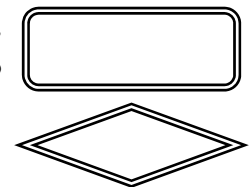
Observem que tota persona viu, obligatòriament, en una ciutat. En canvi, en una ciutat hi poden viure moltes persones, però també es pot donar el cas que en una ciutat de l'entitat *CIUTATS* no hi visqui cap de les persones que formen part de l'entitat *PERSONES*.

1.2.3. Entitats fortes i dèbils

Per finalitzar amb la presentació del model ER, manca presentar un cert tipus d'entitat que apareix en certes situacions.

Anomenarem **entitat feble** a una entitat tal que les seves instàncies no s'identifiquen per sí mateixes sinó que depenen d'una altra entitat, la qual s'anomena **entitat forta**.

Una entitat feble està vinculada a la corresponent entitat forta per una relació N a 1. Gràficament, l'entitat feble es representa mitjançant dues caixes concèntriques. La relació que la connecta amb l'entitat forta es representa mitjançant dos rombes concèntrics.

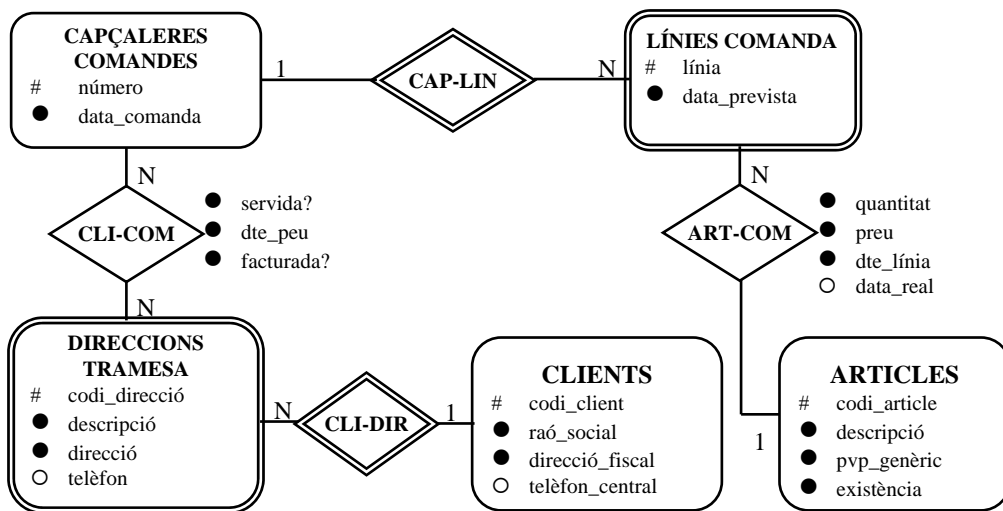


Vegem-ne alguns exemples.

- Suposem el cas de les línies d'una comanda de venda assignada a un client. La comanda de venda té un número o codi que la identifica. Existeix per sí mateixa. Però, i les línies de la comanda?

Suposem que enumerem les diferents línies d'una comanda, a partir del número 1. Evidentment, totes les comandes tenen la línia 1 i moltes tindran també la línia 2, la línia 3, ... És a dir, el número de línia no identifica una línia de comanda ja que hi pot haver moltes línies amb el mateix número. Ara bé, el número de línia sí que identifica una línia dins la comanda. Per aquest motiu, podem dir que les línies de comanda són una entitat feble de comandes.

- Suposem la situació que es produeix quan els clients de l'organització tenen diferents direccions de tramesa. És molt lògic, en tal situació, tenir les direccions codificades dins la nostra base de dades. Com codificar-les? Sembla lògic que una direcció no existeix per sí mateixa sinó que ha d'anar acompanyada del client corresponent. Per aquest motiu podem dir que les direccions són una entitat feble de clients.



1.3. Model estès

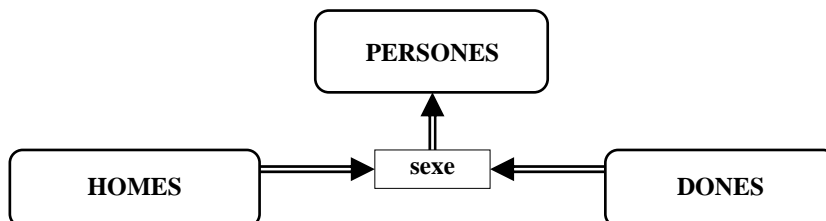
El model ER presentat per Chen es queda coix per modelar situacions com les dues que presentem a continuació. Per aquest motiu s'acostuma a estendre el model ER amb les dues representacions que es proposen i es parla del model EER. Molts autors incorporen directament les dues representacions següents en el model ER.

a) Representació d'una partició¹⁰ d'una entitat

Hi ha situacions en les que una entitat es subdivideix en subentitats disjunts de manera que tota instància de l'entitat pertany a una i només una de les subentitats.

Quan això succeeix, l'entitat inicial acostuma a tenir un atribut que identifica, per cada instància de l'entitat, l'única subentitat a la que pertany.

Un exemple d'aquesta situació és la que es produeix en l'entitat *PERSONES*, ja que tota persona, segons el seu *sexe* (atribut que genera la partició), pertany a una de les subentitats *DONES* o *HOMES*. Ho representarem com:



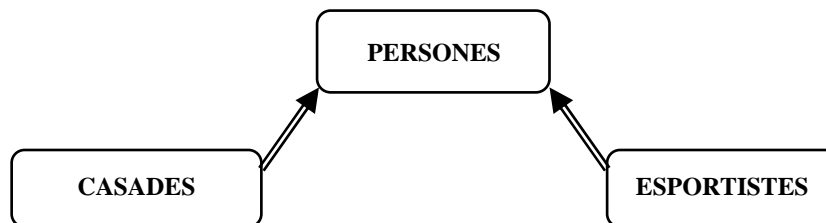
¹⁰ Concepte matemàtic de partició

Una partició d'un conjunt és una col·lecció de subconjunts disjunts de manera que la seva unió forma el conjunt inicial.

b) Representació de subconjunts (disjunts o no) d'una entitat.

Hi ha situacions en les que una entitat es subdivideix en subentitats que no tenen per què ser disjunctes o que, tot i ser-ho, no tota instància pertany a una de les subentitats.

Un exemple d'aquesta situació és la que es produeix en l'entitat *PERSONES* i els seus subconjunts *CASADES* i *ESPORTISTES*. Evidentment hi pot haver persones que pertanyin als dos subconjunts, persones que només pertanyin a un d'ells i persones que no pertanyin a cap d'ells. Ho representarem com:



Les dues situacions afegides en el model EER respecte el model ER s'utilitzen quan per les instàncies de l'entitat mare cal distingir atributs que només són aplicables en les subentitats. En tal cas, els atributs comuns es consideren a l'entitat mare i els atributs particulars a les subentitats. Cal fer nota que les subentitats mai tenen una clau primària pròpia, sinó que hereten la clau primària de l'entitat mare.

2. Disseny d'una base de dades

Tal i com hem comentat en la introducció d'aquesta unitat didàctica, no entrarem en els mètodes a seguir per aconseguir la informació necessària per efectuar el disseny de la base de dades. Suposarem que hi ha una persona que ens presenta un resum de la situació per la que cal dissenyar la base de dades. Suposem que se'ns presenta el següent escrit.

Volem informatitzar el circuit de vendes d'una empresa que ven uns determinats productes.

Una venda s'inicia quan un client efectua una comanda la qual és atesa pels empleats del departament comercial. Totes les vendes s'efectuen a clients de l'empresa i, per tant, els clients estan codificats. De cada client és obligat conèixer la seva raó social i és convenient conèixer la seva direcció, telèfon, fax i el nom i càrrec de la persona de contacte. Ara bé, un client pot efectuar la comanda i no ser ell el destinatari de la mercaderia, tot i que serà ell el responsable del facturat i del corresponent pagament.

Les comandes acceptades pels empleats del departament comercial reben, de manera automàtica, un número correlatiu que las identifica dins els circuits d'informació. Interessa deixar constància de la data en què s'ha efectuat la comanda així com l'empleat que l'ha acceptat i que en serà l'encarregat del posterior seguiment.

Les comandes acceptades estan formades per un o més articles dels que comercialitza l'empresa. Els productes comercialitzats tenen un codi identificatiu designat per la pròpia empresa i, obligatòriament, una descripció. Poden ser fabricats per la pròpia empresa o subministrats per diferents proveïdors. Interessa, com a molt, conèixer el darrer proveïdor que ha subministrat un determinat producte. Així mateix, l'empresa té definides unes determinades categories de productes, de manera que cada article pot pertànyer, com a molt, a una d'aquestes categories. Cada categoria té, obligatòriament, un nom i disposa d'un espai per a efectuar-ne una breu explicació, així com la possibilitat d'incloure un logotip -per cada categoria- que pugui acompanyar els corresponents articles i així fer-los més ràpidament identificables.

L'empresa assigna, a cada empleat, un codi i, possiblement, un càrrec i un tractament (Sra., Sr., Srta., D.,...). Interessa també poder mantenir les dates de naixement i de contractació, així com la seva direcció, el telèfon personal i, a ser possible, la seva foto. Evidentment, cada empleat disposa d'una extensió telefònica dins l'empresa, que ha de ser consultable. Seria molt interessant poder disposar d'un espai per incloure diverses observacions per cada empleat.

L'empresa també té codificats els seus proveïdors de manera semblant als clients. És a dir, de cada proveïdor és obligat conèixer la seva raó social i és convenient conèixer la seva direcció, telèfon, fax i el nom i càrrec de la persona de contacte. La majoria dels proveïdors de l'empresa disposen, en l'actualitat, de pàgines WEB, de manera que es vol mantenir la direcció de la seva pàgina principal en el WEB.

Dels productes sol·licitats en una comanda interessa conèixer el preu unitari que s'estableix, la quantitat que se'n demana i el descompte que s'efectuarà, si és el cas. En acceptar una comanda, l'empleat encarregat li assigna una data prevista de lliurament que cal intentar complir. Es vol tenir constància d'aquest fet i es demana poder emmagatzemar la data en què s'efectua l'envio de la mercaderia així com l'agència de transport que l'efectua. En aquest punt cal saber que l'empresa treballa amb unes determinades agències de transport, les quals estan codificades i de les que cal saber el nom i, si és possible, el telèfon de contacte.

Tota comanda es serveix en la seva totalitat o no es serveix, és a dir, no hi ha la possibilitat de fer lliuraments parcials de mercaderia. A més, tota comanda pot portar un càrrec afegit causat pel total d'impostos a aplicar.

El departament comercial està estructurat jeràrquicament, de manera que cada empleat pot tenir un cap que ha de ser, a la vegada, empleat de l'empresa. Aquesta estructura jeràrquica té, evidentment, un(s) cap(s) superior(s) que no depenen de ningú. Interessa mantenir, per cada empleat, el seu cap, de manera que donat un cap es pugui saber tots els empleats que té al seu càrrec.

Dels productes que l'empresa comercialitza interessa tenir, també, la presentació (és a dir, el contingut d'una unitat de compra - venda) i l'existència actual (en unitats de compra - venda) així com la quantitat demanada a proveïdor i que està pendent de recepció. Es vol tenir, també, també, la quantitat mínima a comprar als proveïdors. A la vegada, un article -tingui o no existència- ha de poder-se marcar com article suspès per a no permetre la seva compra - venda fins que sigui desmarcat.

Anem a proposar unes línies de treball per assolir el model ER.

1) Lectura inicial del text per obtenir una primera visió de la situació a modelar.

En aquesta primera lectura anem anotant els conceptes susceptibles de ser entitats, relacions i atributs. Notem que el text no deixa de ser un relat efectuat per algú que coneix l'organització, i que no té per què estar ben estructurat. És la nostra responsabilitat treure'n el suc.

Els conceptes que n'obtenim de la primera lectura, ordenats segons van apareixent, els tenim a la taula:

Paràgraf	Concepte	Susceptible de ser...
1r.	circuit de vendes	???
	empresa	entitat
	productes	entitat
2n.	venda	entitat
	client	entitat
	comanda	entitat
	client efectua comanda	relació entre client i comanda

	empleat accepta comanda	relació entre empleat i comanda
	empleats	entitat
	departament comercial	entitat
	codi, raó social, telèfon, fax, nom i càrrec de la persona de contacte dels clients	atributs de l'entitat client
	destinatari de la mercaderia	atributs de comanda
3r.	codi i data de comanda	atributs de comanda
4t.	productes de la comanda	relació entre productes i comanda
	codi i descripció dels productes	atributs de producte
	article fabricat o subministrat?	atribut de producte
	proveïdors	entitat
	darrer proveïdor que subministra un producte	relació entre proveïdor i productes
	categories de productes	entitat
	producte pertany a categoria	relació entre productes i categories
	nom, descripció i logotip	atributs de categories
5è.	codi, càrrec i tractament, dates de naixement i contractació, direcció, telèfon personal, foto, extensió telefònica i observacions	atributs dels empleats
6è.	codi, raó social, telèfon, fax, nom i càrrec de la persona de contacte i pàgina principal dins el WEB dels proveïdors	atributs dels proveïdors
7è.	preu unitari, quantitat i descompte	atributs de la relació entre productes i comanda
	data prevista de lliurament	atribut de comanda
	agències de transport	entitat
	agència lliura comanda	relació entre agències i comanda
	data d'envio	atributs de la relació entre agències i comanda
	codi, nom i telèfon de contacte	atributs d'agències de transport
8è	situació de la comanda (servida o pendent)	atributs de comanda
	càrrec afegit	atribut de comanda
9è.	cap de cada empleat	relació entre empleats
10è.	presentació i existències actual, pendent de recepció i mínima dels productes	atributs de productes
	situació de l'article (suspès o no)	atribut de productes

2) Anàlisi dels conceptes anotats

En aquest moment cal analitzar amb deteniment tots els conceptes anotats de manera que es pugui reafirmar les anotacions preses en la primera lectura. Repetirem la lectura tantes vegades com sigui necessari retocant les anotacions.

Analitzant les anotacions que hem efectuat, trobem quatre punts obscurs, corresponents als conceptes *circuit de vendes*, *empresa*, *venda* i *departament comercial*. En efectuar la primera lectura hem considerat aquests termes com susceptibles de ser entitats. Però en l'anàlisi que ara estem efectuant comprovem que no hi ha cap atribut per aquestes entitats i això no és possible. Analitzem-los.

- El concepte *circuit de vendes* l'hem anotat però ja no teníem clar, des d'un inici, de què es tractava. L'hem anotat per què ens semblava un concepte fort, un concepte de pes, tot i que després veiem que té pes dins el text però no és res real, tangible, palpable (com ha de ser una entitat) ni tampoc és fruit d'una relació entre dues o més entitats. En realitat és el concepte que resumeix tota la situació a modelar. Descartem, per tant, aquest concepte.
- El concepte *empresa* ens ha fet caure en un parany que acostuma a existir en dissenyar models. L'empresa és, evidentment, una cosa real, tangible i palpable. Per tant, té tots els números per a ser una entitat. Observem, però, que el model que volem assolir és per a gestionar una única empresa. Per aquest motiu, no ha de tractar-se com una entitat. Caldria considerar l'entitat *empreses* en cas que la nostra base de dades hagués de donar suport a diferents empreses de manera simultània. Això no vol pas dir que no es pugui utilitzar el mateix model per a la gestió de diferents empreses (imaginem-nos el cas d'estar dissenyant un model per a un estàndard). Descartem, per tant, aquesta possible entitat.

- El concepte *venda* ens ha còpsat de la mateixa manera que el concepte *circuit de venda*. En l'anàlisi que estem efectuant observem que no apareix cap atribut per a *venda* i ens adonem que el concepte *venda* equival al concepte de *comanda* que sí té atributs. Descartem, per tant, aquesta possible entitat.
- El concepte *departament comercial* ens ha enganyat de manera semblant al concepte *empresa*. El descartem.

3) Estructuració dels conceptes anotats.

En aquest punt reestructurem la taula de manera que apareguin entitats i relacions amb els seus atributs. És possible que el nostre sentit comú ens porti a afegir algun ens calgui afegir algun atribut no deduït de les anotacions inicials.

Entitats	Atributs
productes	codi, descripció, article fabricat o subministrat?, presentació, existència actual, existència pendent de recepció, existència mínima, situació de l'article (suspès o no)
clients	codi, raó social, telèfon, fax, nom i càrrec de la persona de contacte
comandes	codi, data de comanda, data prevista de lliurament, situació de la comanda (servida o pendent), càrrec afegit, destinatari de la mercaderia
empleats	codi, cognoms i nom, càrrec, tractament, dates de naixement i contractació, direcció, telèfon personal, foto, extensió telefònica i observacions
proveïdors	codi, raó social, telèfon, fax, nom i càrrec de la persona de contacte i pàgina principal dins el WEB dels proveïdors
categories	codi, nom, descripció i logotip
agències transport	codi, nom i telèfon de contacte

Hem afegit, a més dels atributs que ja havíem anotat en l'anterior taula, el *codi* de les *categories* i els *cognoms* i *nom* dels *empleats*.

Relacions	Entitats relacionades	Atributs
client que efectua comanda	comandes - clients	
empleat que accepta comanda	comandes - empleats	
productes de la comanda	productes - comandes	preu unitari, quantitat i descompte
darrer proveïdor que subministra un producte	proveïdors - productes	
producte pertany a categoria	productes - categories	
cap de cada empleat	empleats - empleats	
agencia lliura comanda	comandes - agències	data d'envio

4) Execució del disseny de la base de dades amb el model ER.

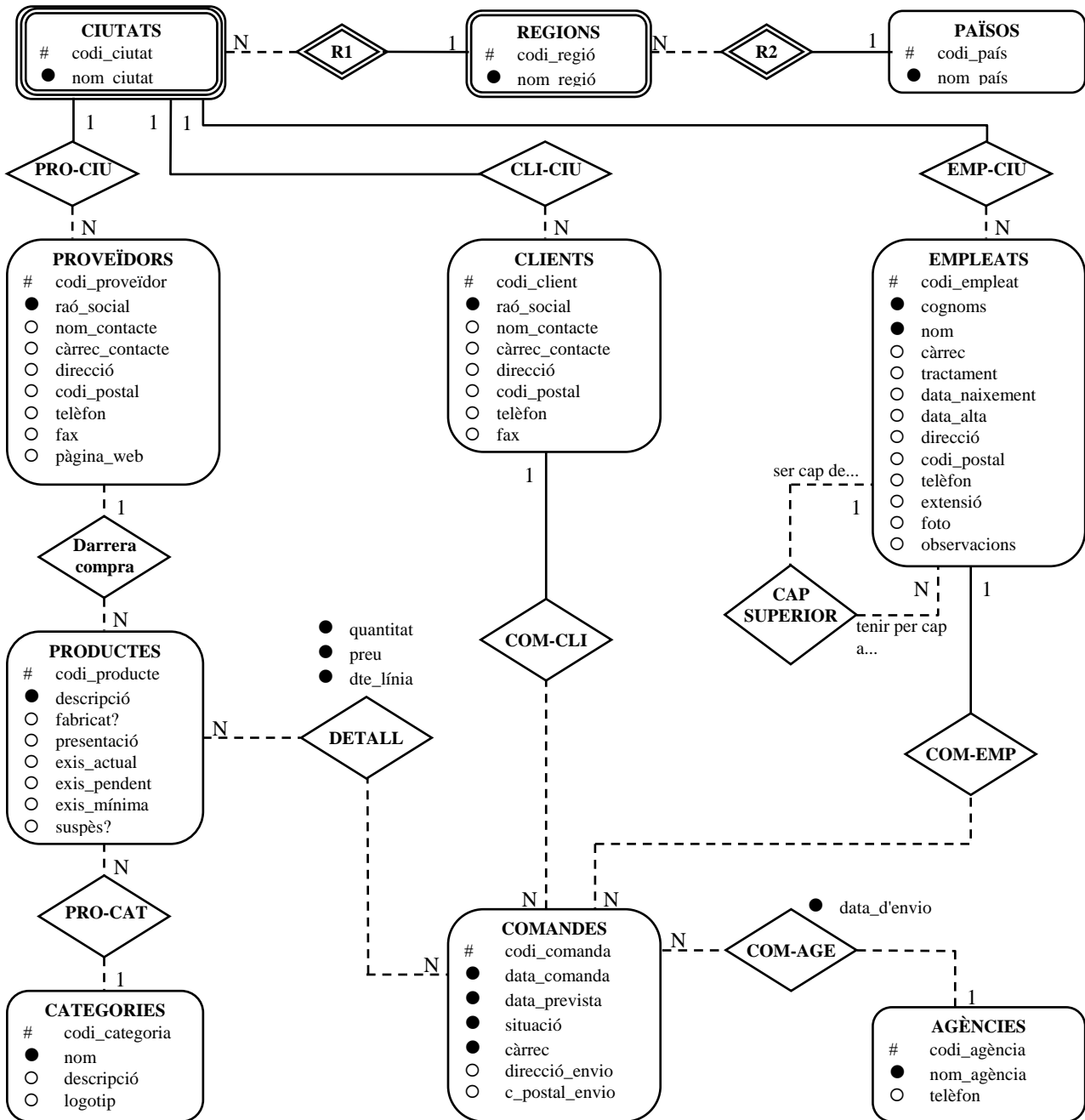
En aquest darrer apartat es tracta d'efectuar el disseny de la base de dades amb el model ER utilitzant les anotacions efectuades i decisions relatives a les claus primàries de les entitats, l'obligatorietat u opcionalitat dels atributs, el grau de les relacions i la obligatorietat u opcionalitat de les entitats en les relacions. És molt possible que per prendre aquestes darreres decisions sigui menester efectuar més lectures del text.

Hi ha una decisió a prendre referent a les dades que cal incloure en una *direcció*. Observem que aquest concepte apareix de manera explícita en *clients*, *proveïdors* i *empleats* i de manera implícita en el *destinatari* de la comanda.

Què entenem per *direcció*? Per decidir-ho cal saber quin és l'entorn geogràfic de la nostra empresa: comarcal, regional, estatal, continental, mundial,... Bé, considerem que la nostra empresa té àmbit mundial i, per tant, la direcció (per totes les entitats) ha de tenir un tractament que incorpori des de la direcció del carrer fins el país, tractant el codi postal i unes possibles regions o departaments o províncies dins el país.

Considerarem el *país*, la *regió* i la *ciutat* com a entitats a mantenir codificades dins la nostra base de dades. Al contrari, la *direcció del carrer* i el *codi postal* els considerarem com a informacions no codificables. Identificarem les ciutats com a elements d'una regió concreta d'un país i les regions com a elements d'un país en concret. És a dir, com que en el món hi pot haver diferents ciutats de nom *Barcelona*, identificarem aquesta dins d'una regió de nom *Catalunya* la qual està identificada dins un estat de nom *Espanya*.

Així doncs, obtenim el disseny següent:



- Tenim la relació reflexiva *CAP DE* sobre l'entitat *EMPLEATS* de grau 1:N ja que *un empleat pot ser cap de varis (o ningun) empleats* i *un empleat pot tenir per cap a un (o ningun) empleat*. La comprensió d'aquesta relació a partir del rombe que defineix l'aplicació pot ser complicada i, per aquest motiu, hem anotat *ser cap de...* i *tenir per cap a...* en les línies que uneixen el rombe amb l'entitat.

- L'entitat *CIUTATS* és feble de l'entitat *REGIONS* i aquesta és feble de l'entitat *PAÏSOS*. Per aquest motiu, l'entitat *REGIONS* es representa dins dues caixes concèntriques i l'entitat *CIUTATS* es representa dins tres caixes concèntriques.

3. Models de sistemes gestors de bases de dades

En l'anterior unitat didàctica havíem comentat l'existència de diferents models de base de dades en funció de les estructures de dades en què es sustentaven. Així teníem:

- Model jeràrquic, basat en l'estructura de dada *arbre*.
- Model en xarxa, que es basa en l'estructura de dada *xarxa*.
- Model relacional, que es basa en l'estructura de dada *taula* -no confondre amb el tipus de dada estàtica *taula*-.
- Models orientats a objectes, basats en l'anàlisi i disseny orientats a objectes.

Els dos primers -jeràrquic i en xarxa- varen dominar el mercat en els anys setanta i vuitanta i avui en dia continuen funcionant amb un bon rendiment en les organitzacions que en aquell moment els van adquirir.

El model relacional és el model base pels sistemes gestors de bases de dades actuals. Per aquest motiu hi ha un crèdit dedicat íntegrament al seu estudi.

Els models orientats a objectes estan en fase experimental i encara veig llunyà el seu arrelament com a base pels sistemes gestors de bases de dades. Els actuals fabricants de sistemes gestors de bases de dades relacionals incorporen en els seus productes alguns conceptes de l'orientació a objectes, els quals es poden o no utilitzar. Penso, i és una opinió totalment personal, que el futur passa per continuar utilitzant el model relacional incorporant conceptes de l'orientació a objectes, més que en una nova revolució dels sistemes gestors de bases de dades que provoqui l'aparició de sistemes gestors de bases de dades basats totalment en l'orientació a objectes. Però això ja serà una altra guerra!

Una darrera observació referent als models relacionals amb incorporacions de conceptes no relacionals: cada fabricant de SGBDR segueix el seu propi criteri a l'hora d'evolucionar el seu producte. És a dir, no hi ha, fins el moment, cap criteri estàndard semblant a la normativa ANSI/X3/SPARC esmentada en l'anterior unitat didàctica.

L'anterior observació és un punt molt important a tenir en compte pels analistes i dissenyadors d'aplicacions informàtiques que accedeixen a SGBDR via LMD i LC hostatjats. Si es pretén que l'aplicació informàtica sigui independent del SGBDR al que accedeixi, cal dissenyar les bases de dades gestionades pel SGBDR utilitzant el model relacional estàndard, sense emprar els conceptes no estàndards que el fabricant hagi incorporat. D'aquesta manera és molt probable que no s'utilitzi tota la potència que podria donar-nos cada SGBDR però s'assoleix la independència de les aplicacions respecte el SGBDR.

Exemplificació del darrer paràgraf

Suposem que desenvolupem una aplicació informàtica per la gestió acadèmica dels Cicles Formatius, destinada als centres docents que imparteixen aquest tipus d'estudi.

És clar que la gestió a efectuar serà la mateixa tan si el centre docent imparteix un únic Cicle Formatiu com si n'imparteix vint, però estarem d'acord en què el volum de dades no és el mateix i que la diferència s'engrandeix a mida que passen els cursos acadèmics i s'ha d'anar mantenint les dades dels cursos anteriors per poder efectuar certificats, matricular alumnes que els queden alguns crèdits pendents,...

Per tant, l'aplicació serà la mateixa, però potser caldrà diferenciar el SGBD que emmagatzemi les dades en cada situació. Així, per un centre docent amb poca gestió de Cicles Formatius n'hi haurà prou amb un SGBD simple, com *MS-Access*. En canvi, per un centre docent amb molta gestió de Cicles Formatius caldrà un SGBD potent, com *Oracle*.

Hem esmentat dos productes molt coneguts, però en el mercat n'hi ha d'altres: *FoxPro*, *SQL-Server*, *Informix*, *MultiBase*,... Cal escollir el que s'adapti millor a les necessitats de gestió i a les possibilitats econòmiques d'adquisició i posterior manteniment.

Per últim, referent a l'estudi dels sistemes gestors de bases de dades relacionals en el CFS *Desenvolupament d'aplicacions informàtiques*, comentar que no s'estudia els models jeràrquic i en xarxa per considerar-los obsolets tot i que hi hagi encara sistemes informàtics basats en aquests models. Cal pensar que el futur titulat en aquest cicle coneixerà en profunditat la gestió d'arxius indexats, la gestió de dades dinàmiques, el model entitat - relació i el model relacional, de manera que si casualment ha de treballar amb SGBD jeràrquics o en xarxa, disposarà de les capacitats necessàries per assolir el seu funcionament en base a la multitud de bibliografia existent.

Annex. Enunciats per exercitar el model entitat-relació.

1. L'empresa CFSDAI

Els alumnes del CFS Desenvolupament d'Aplicacions Informàtiques que s'han anat despenjant degut a que han pogut tastar que la programació informàtica no és el seu futur, han decidit muntar una empresa que es dediqui a la fabricació i venda de material d'oficina. Com que tenen uns companys que encara continuen en el CF, els hi encarreguen la informatització del departament de vendes de la nova empresa.

El departament de vendes de qualsevol empresa tracta bàsicament amb els dos tipus de dades següents: clients i productes. La nova empresa no és un cas especial i, per tant, interessa tenir constància dels clients que té l'empresa i dels productes que es comercialitzen. Dels clients interessa conèixer la raó social, el cognom i nom de la persona de contacte, la direcció, el codi postal, ciutat i província. A més, l'empresa assigna un codi de client que l'identificarà. Dels productes que comercialitza, que també estan codificats, interessa una breu descripció i el cost unitari.

Aquesta empresa ha decidit que per poder donar a conèixer els seus productes necessita tenir una plantilla de venedors que es desplacin a la recerca de clients i comandes. Dels venedors, també codificats, interessa tenir constància del seu nom i cognom, la data d'alta, el salari, la comissió que li pertoca i la ciutat i província de residència. La majoria dels venedors tenen un supervisor que a la vegada es també un venedor.

Les comandes que es van produint passen a formar part de la cartera de l'empresa amb un codi que es va assignant a mesura que es reben. Interessa tenir constància de la data de comanda, el client i el venedor corresponents, els productes que la integren i la quantitat de cadascun d'ells. Interessa saber també si ja s'ha enviat cadascun dels productes. A tal efecte cal considerar que una comanda pot estar formada per diferents productes. Pot ser que diferents productes es serveixin en diferents dates. No és possible que diferents quantitat d'un mateix producte d'una comanda es serveixi en diferents dates. Interessa també saber si ja s'ha facturat. Només es facturarà una comanda (una única factura) quan ja s'hagin enviat tots els productes que la formaven.

El departament comercial d'aquesta empresa també ha de tenir constància de l'estoc de productes que hi ha en els diferents magatzems que disposa. Dels magatzems., que estan codificats, només interessa saber la ciutat en la que es troben.

Per finalitzar, recordar que en mobles d'oficina és molt possible que es puguin vendre productes que formi part d'altres productes (per exemple, una cadira pot vendre's sola i també pot formar part d'una oficina sencera). Evidentment cal tenir constància de quin producte pot formar part d'un altre producte.

2. La biblioteca

Una biblioteca privada de gran volum, constituïda per socis lectors que abonen una quota mensual, desitja informatitzar el seu funcionament i dissenyar una base de dades adequada a les seves necessitats.

El títol de cada obra de la citada biblioteca és únic, és a dir, obres diferents tenen títols diferents. En general, una obra ha estat escrita per un o més autors en un ordre que figura en les primeres pàgines de l'obra, que pot constar de varis volums o toms. A més, degut a l'elevat número d'assidus lectors, es disposa de varis exemplars de les obres més populars, encara que a vegades no de tots els volums o toms.

Particularment en obres de caràcter científic, per allò de "les ciències avancen que és una barbaritat", s'han anat incorporant noves edicions de determinades obres, que, lògicament interessen molt més, encara que no falta cap historiador interessat en estudiar l'evolució que les successives edicions representen.

Alguns socis sol·liciten les obres especificant el seu títol exacte, però altres només recorden un autor o autors, que no sempre és el primer o primers en la llista que figura en les primeres pàgines de l'obra en qüestió. Hi ha també socis interessats en certs temes, dels que desitgen saber quines obres existeixen a la biblioteca. Degut al gran volum de llibres i a l'especialització d'alguns socis, es desitja poder accedir, també, per subtemes a varis nivells, per exemple, si algú demana llibres sobre "informàtica", li apareixeran com subtemes "programació", "sistemes operatius", "bases de dades", "informàtica aplicada a medicina", etc. i si demana llibres de "medicina", li apareixeran els subtemes "anatomia", "patologia", "informàtica aplicada a la medicina", etc.

En adquirir una obra, la biblioteca assigna a cada exemplar un número de referència, amb el que es facilita la seva localització en el dipòsit de llibres.

Cada soci té dret a retirar en préstec els llibres que desitgi, per un període de quinze dies prorrogables a altres quinze sempre que els exemplars estiguin disponibles o no sol·licitats per altres socis. Si en sol·licitar una obra en préstec aquesta no està disponible, el soci pot situar-se a la cua d'espera per ser avisat en el moment en que la seva sol·licitud pugui ser atesa (és a dir, quan un altre soci li retorni l'obra). Els socis poden també consultar el llibre o llibres que desitgin en una sala de lectura magníficament condicionada prop el dipòsit de llibres.

3. La base de dades atlètica

La federació estatal d'atletisme ha pensat establir una base de dades que li permeti un millor seguiment de les activitats de les diferents federacions autonòmiques, així com dels clubs associats i dels atletes que desenvolupen la seva activitat esportiva com membres d'un club determinat. A més, és propòsit de la federació central mantenir informació al dia respecte les millors marques del atletes en cadascuna de les activitats esportives a les que es dediquen.

És també d'interès registrar les reunions atlètiques que tenen lloc durant l'any entre federacions autonòmiques. Un encontre d'aquest tipus involucra a dues o més federacions (hi ha trobades triangulars, quadrangulars, etc. àdhuc algunes -com el campionat d'Espanya "indoor" o "outdoor"- en el que participen -en principi- totes les federacions autonòmiques i/o autonòmiques). Es desitja registrar quines federacions autonòmiques han participat en cada encontre, així com on i quan es van celebrar i la marca de cada atleta participant en un encontre determinat hauria d'estar present també a la base de dades, de manera que en qualsevol moment fos possible, per exemple, consultar la història de marques d'un atleta donat, des del principi de temporada fins el moment d'efectuar la consulta. S'ha de tenir en compte que un atleta pot participar en un mateix encontre competint en més d'una especialitat, cas en que es desitja tenir dades sobre les marques de cadascuna.

Es celebren també competicions entre clubs (enlloc de federacions autonòmiques), de les quals es desitja també mantenir informació en la base de dades. No és precís registrar el detall de les marques que realitza cada atleta (només si algú millora la millor marca de l'any fins el moment, es posaria al dia la dada corresponent). El que interessa tenir al dia són els resultats dels encontres entre clubs (que poden, en general, ser també entre més de dos clubs), és a dir, punts obtinguts per cada club en cada encontre en què ha participat i lloc ocupat a la classificació final de l'encontre.

4. La multinacional

Una empresa té diverses fàbriques cadascuna de les quals produeix diversos productes finals específics de cada fàbrica.

A l'empresa treballen molts empleats, assignats cadascun a una fàbrica. Algun d'ells formen part de la directiva de la fàbrica corresponent i res impedeix que un empleat pugui ésser contractat per una altra fàbrica de la mateixa empresa, pugui acomiadar-se o ser acomiadat o pugui retornar a una fàbrica on ja havia treballat. Es desitja tenir una història completa de la contractació de personal per part de cada fàbrica, amb les dades habituals d'un treballador, si ocupen un lloc directiu o no i una llista de les seves especialitats, si es que les tenen (que ve donada per títols acadèmics, diplomes de curssets o notable experiència professional), i qualificades aquestes especialitats quan existeixin, de 1 a 5 segons apreciacions dels directius.

Cada fàbrica compra a diversos proveïdors les matèries primeres que degudament elaborades es convertiran en els productes finals, abans esmentats. Alguns d'aquests productes encara que amb nom i presentació diferents, són equivalents a efectes dels clients (penseu, per exemple, què succeeix amb alguns productes farmacèutics). Serà convenient saber la composició elemental de cada producte final elaborat.

Els clients formulen les seves comandes de diferents productes a l'esmentada empresa i aquesta, que disposa de més magatzems que fàbriques per una millor distribució dels seus productes, té assignat a cada client el magatzem més proper al seu domicili, d'on procura, si l'estoc és suficient, servir totes les seves comandes. Es vol conservar un històric de clients i comandes servides, així com dels proveïdors i dels seus subministraments.

5. La facultat

Cal dissenyar una base de dades per a gestionar la docència d'un curs acadèmic en una facultat universitària per tal de controlar qüestions com ara les notes, horaris, matriculacions, professors, alumnes, etc.

La facultat està estructurada en departaments, cadascun dels quals imparteix la docència de diverses assignatures. No es dona el cas que una mateixa assignatura sigui impartida per més d'un departament. De les assignatures interessa saber les sigles, nom, crèdits, nivell, etc. També es vol saber quines assignatures són prerrequisit¹¹ d'altres.

Cada assignatura té un o més grups (per exemple, grups 10, 20, 30 i 40 de l'assignatura ADBD, grups 10, 20 i 30 de l'assignatura SGBD, etc.) De cada grup interessa saber les aules i "horaris" corresponents al curs actual (ententent per "horari" el dia de la setmana i l'hora d'inici). Cal tenir en compte que un mateix grup pot donar-se en aules diferents sempre que sigui en "horari" diferent. De les aules es vol conèixer quina és la seva capacitat.

Dels professors interessa saber la seva categoria laboral i la seva dedicació (a temps parcial o a temps complet). Dels alumnes interessa conèixer el número de matrícula i l'any d'ingrés a al facultat. Tant dels professors com dels alumnes, que podríem considerar sota una nova entitat que en podríem dir "persones" interessa el DNI, el nom, l'adreça i el telèfon.

Tots els professors estan adscrits a un i només a un departament. En general estan encarregats d'algun o alguns grups, però és possible que hi hagi professors que no estiguin encarregats de cap grup el curs actual (per exemple, per a dedicar-se a tasques de recerca o a tasques administratives). És possible també que un grup tingui més d'un professor encarregat (per què per exemple, diferents professors es responsabilitzen de diferents parts de l'assignatura). Es vol constància de tots els fets anteriorment esmentats, inclòs quina és la part de que cada professor es responsabilitza.

A l'inici del curs, els alumnes es matriculen en diversos grups i, naturalment, es vol tenir informació d'aquest fet. La base de dades hauria de guardar també informació de les notes que els alumnes han obtingut en les diferents convocatòries d'examen (febrer, juny i setembre) de cada assignatura.

La biblioteca d'aquesta facultat disposa d'una gran quantitat de textos (llibres, revistes, apunts, etc.) dels que es voldria tenir informació. Aquests textos s'identifiquen per un codi que assigna la pròpia biblioteca. En els diferents grups de cada assignatura es recomanen un o més d'aquests textos inclòs un cert ordre de preferència o prioritat. Sovint es dona el cas que un mateix text és recomanat a més d'un grup (especialment si són grups d'una mateixa assignatura). Vol enregistrar-se quin són els textos recomanats i amb quina prioritat a cadascun dels grups.

Com ja es dedueix dels paràgrafs anteriors, només es desitja guardar informació del curs actual i no la informació històrica de cursos anteriors.

¹¹ **Assignatures prerrequisit i corequisit**

Donada una assignatura A, les seves assignatures prerrequisit són aquelles assignatures que cal tenir cursades amb avaluació positiva abans de poder cursar l'assignatura A i les assignatures corequisit són aquelles assignatures que cal tenir cursades amb avaluació positiva abans de cursar l'assignatura A o que cal cursar simultàniament amb l'assignatura A.

6. L'hospital

Un hospital desitja establir una base de dades per millorar el control de les seves operacions (no només les quirúrgiques).

A l'actualitat hi ha dificultats per localitzar les fitxes dels pacients quan es necessiten. Moltes vegades, donat que un pacient pot tenir varies fitxes (per exemple perquè ha estat visitat en diverses ocasions per diferents metges), no es localitza tota la informació disponible d'un pacient determinat, amb el perill en el tractament d'aquest pacient.

L'hospital consta de diversos pavellons, corresponents a diverses especialitats (ginecologia, sistema digestiu, etc.) Els pacients ingressats ho són d'un pavelló determinat, en les habitacions a l'efecte reservades en aquells. Els metges, a part d'atendre els malalts que els corresponen per la seva especialitat (corresponents pavellons), poden atendre els pacients d'altres pavellons que necessitin atenció puntual, és a dir, algú amb una malaltia d'estómac pot ser atès per un ginecòleg si es presenta alguna complicació. Així, en general, en un moment determinat, varis metges poden atendre un mateix pacient.

Interessa també controlar les anàlisis que es fan als pacients, en els diferents laboratoris de l'hospital. És convenient poder consultar els resultats de l'anàlisi que un pacient té pendents de rebre d'un determinat laboratori, per exemple. En casos greus, pot sol·licitar-se la mateixa anàlisi a diferents laboratoris per tenir més seguretat en els resultats.

7. L'empresa de programari

Una empresa de desenvolupament de programari desitja una base de dades per mantenir informació sobre els seus treballadors tècnics, amb la finalitat d'assignar personal a projectes de manera eficient. En particular, es pensa que la següent informació hauria d'estar present a la base de dades:

- Dades usuals de cada treballador (nom, direcció, etc.)
- Informació sobre els estudis acadèmics realitzats pels treballadors amb una breu descripció (per exemple, una descripció de la carrera de la facultat si algú dels treballadors fos llicenciat en informàtica) i indicant la data en que cada treballador es va graduar dels diferents estudis.
- En connexió amb diferents especialitats laborals (per exemple "programador en C", "analista", "dissenyador de bases de dades") es desitja especificar, en una escala d'1 a 5 el denominat "grau de capacitat de cada treballador" per treballar en l'especialitat o especialitats per les que es considera capacitat, independentment dels estudis cursats (és a dir, en base a experiència laboral prèvia).
- El grau de possible idoneïtat per treballar en una determinada especialitat que es suposa tenir pel fet d'haver cursat determinats estudis, sense necessitat d'experiència laboral prèvia.
- El districte geogràfic que per diversos motius (residència, etc.) s'associa a cada treballador, indicant que prioritàriament el treballador hauria de ser assignat a projectes a desenvolupar en el corresponent districte.

8. La caixa d'estalvis

Es tracta de dissenyar una base de dades per a la gestió del personal d'una caixa d'estalvis implantada amb moltes agències a tot un país.

Els empleats estan identificats per un codi d'empleat. Serà important conèixer la seva ciutat de residència així com saber a quines ciutats són ubicades les diferents agències de l'entitat. Aquestes agències estan identificades per un codi numèric de quatre xifres. Un empleat, en un moment donat, treballa en una agència però pot ser traslladat a una altra, fins i tot pot tornar a treballar en una agència on ja hi havia treballat anteriorment. Interessa conèixer l'historial del pas de cada empleat en les diferents agències.

Els empleats poden tenir titulacions acadèmiques i es vol tenir constància d'aquest fet, concretament l'any i el centre acadèmic que els hi ha concedit. Els empleats tenen una determinada categoria laboral (auxiliar, oficial 2a, oficial 1a,...) a la que van accedint, sempre per millorar (sense degradacions) i que només perden si marxen definitivament de l'entitat. Es vol conèixer les diferents categories que ha assolit cada empleat. Considerarem que una vegada un empleat ha deixat definitivament l'entitat, no hi tornarà a treballar mai més. Segons un conveni de l'empresa, a cada categoria li correspon un determinat sou base, un preu d'hora extra i un coeficient a aportar a la seguretat social.

Alguns empleats estan afiliats a una (única) central sindical. S'ha arribat al pacte de descomptar en la nòmina mensual la quota sindical pels afiliats a cada central. Aquesta quota és única per a tots els afiliats a una determinada central sindical.

Els empleats de la caixa gaudeixen del dret d'excedència per a un període de temps determinat que pot ser exercit en diverses ocasions en la seva vida laboral i per diverses causes. També gaudeixen de la possibilitat de demanar diferents tipus pre-establerts de préstecs (per exemple, per matrimoni, per adquisició d'habitatge, per estudis, etc.) que poden ser o no concedits. En principi no hi ha cap limitació en demanar diferents tipus de préstecs, fins i tot simultàniament, i es vol constància de la concessió o no dels préstecs. Cada préstec té establertes diferents condicions de les quals, en particular, interessa el tipus d'interès i el període de vigència.

9. La vinícola

S'ha convocat un concurs per a dissenyar una base de dades que servirà per a portar el control de la producció i distribució de vins d'un país.

En el país en qüestió hi ha una colla de "firmes comercials" que són, en realitat, els fabricants i els embotelladors dels vins. Aquestes firmes atorguen "concessions en exclusiva" per les diferents zones del territori del país a diferents empreses distribuïdores. Una empresa distribuïdora pot treballar per a varies firmes comercials. Aquestes concessions es firmen per a un període de temps determinat, prorrogable sempre que es convingui i són vàlides per a vendre totes les marques de vi que la firma comercial fabrica. De les diferents zones del país interessa saber la seva extensió i població a efectes d'estratègia de mercat.

De cada marca de vi hi ha diverses varietats o tipus, com el blanc, el rosat, l'escumós, etc.) i la majoria d'aquestes marques tenen denominació d'origen (per exemple Alella, Priorat, Rioja, Penedès,...) però algunes no en tenen (ja que estan confeccionades amb ingredients més aviat químics i de poca qualitat). També es vol tenir constància de les varietats de vins equivalents, particularment de cara al seu preu.

Les empreses distribuïdores tenen uns magatzems que faciliten la distribució de les botelles als diferents clients (com restaurants, bars, botigues de queviures, etc.) encara que no sempre aquests magatzems estan situats en les zones de distribució que una empresa distribuïdora hagi pogut aconseguir. Interessa saber l'estoc actual respecte les diferents varietats o tipus.

Les empreses de distribució serveixen comandes als diferents clients, identificant la comanda per un número de comanda, intern a cada empresa distribuïdora. En la comanda hi consten, entre altres, la quantitat de botelles demanades per cada varietat.

Les firmes comercials, les empreses distribuïdores i els clients poden identificar-se pel seu codi d'identificació fiscal.

10. La base de dades olímpica

Es tracta de dissenyar la base de dades necessària per a que les proves olímpiques d'atletisme es puguin desenvolupar sense problemes, proporcionant als informadors dels diferents medis de comunicació les dades precises per a que puguin dur a terme la seva missió de la millor manera possible.

Així, s'haurà de proporcionar diàriament una llista de les proves a realitzar, indicant a quina especialitat corresponen (100 metres, salt d'altura, per exemple) i a quin nivell de competició (sèries, semifinals, final, etc.) fent constar quins passaran a la següent ronda (per exemple, els dos primers de cada sèrie i dos per marques o el que correspongui) i a quina hora està prevista la celebració de cada prova.

Es pretén també tenir informació sobre qui està previst que participi a cada prova, indicant en quin ordre (o en quin carrer). Cal tenir present que un mateix atleta pot participar en més d'una especialitat. Es desitja també tenir accés als records del món i olímpic oficials en el moment de celebrar-se les proves corresponents.

En atenció als països participants, es vol també proporcionar una informació esportiva el més completa possible respecte als diferents participants (edat, millor marca en què participa, etc.) així com el rècord nacional de cada país en cada prova.

Els resultats de les diferents proves s'han d'incloure també a la base de dades, de manera que en qualsevol moment es pugui consultar, per exemple, la classificació definitiva dels participants en una prova ja celebrada. En aquest sentit, es desitja tenir una especial atenció als atletes que aconsegueixen medalles i diplomes, deixant constància de les mateixes a la base de dades de manera que sigui possible conèixer no només qui va obtenir or, plata,... sinó també comptabilitzar les medalles i diplomes dels diferents tipus per països.

11. La policia

La policia de tràfic d'un determinat país, dividit en diferents demarcacions de tràfic, vol informatitzar el seu funcionament per tal de tenir un millor control i també poder obtenir certes dades estadístiques, especialment pel que fa a multes i accidents.

De cada demarcació, globalment, només es volen tenir les dades actuals referents al seu nombre d'habitants, nombre de vehicles i nombre de conductors amb carnet, però en canvi, es vol disposar en tot moment de les dades històriques referents a multes, accidents, vehicles i als diversos propietaris i matrícules que cada vehicle hagi pogut tenir.

Tant les multes com els accidents es produeixen en una demarcació determinada i una consulta que es preveu freqüent serà la de conèixer les multes o accidents que s'ha produït en una data determinada o en un interval de dates.

Multes i accidents són identificats per un número, però només dins de cada demarcació. D'un accident interessa saber l'hora i el lloc on s'ha produït així com el(s) vehicle(s) involucrat(s) i també el(s) ciutadà(ns) afectat(s). D'aquest(s) darrer(s) volem saber si era(n) o no conductor(s) en l'accident i també el dany físic rebut, que es qualificarà d'1 a 5 segons la gravetat, indicant la mort amb la qualificació d'1. Dels vehicles anotarem l'estat en què han quedat, valorat d'1 a 5, indicant la inexistència de destrosses amb la qualificació d'1.

En un determinat moment, un vehicle té un únic propietari que pot anar canviant en successives vendes. Interessa, en els diferents moments de cada compra, el seu quilometratge (zero si és nou) i el seu estat de conservació (valorat d'A a E en sentit decreixent).

Les multes es posen a un vehicle i al seu propietari i especifiquen el lloc, el tipus d'infracció (excés de velocitat, conducció temerària,...) i l'import. La base de dades ha de poder incloure el nom del conductor infractor que pot no coincidir amb el propietari del vehicle.

Un vehicle ve identificat pel seu número de xassís (penseu que ha pogut tenir diferents matrícules).

12. Les eleccions

Donada la gran quantitat d'eleccions generals que ha de supervisar un estat, el ministeri de l'interior ha decidit constituir una base de dades pel control de les diferents eleccions.

En primer lloc, es voldria tenir un arxiu històric dels polítics del país, que contingui informació de les seves afiliacions als diferents partits al llarg de la seva trajectòria política, fins i tot en moments no electorals. Un polític pot ser independent, és a dir, no pertànyer a cap partit. Simultàniament, un polític no pot pertànyer a més d'un partit.

La base de dades ha de poder tractar les coalicions de dos o més partits que de manera conjunta decideixen presentar-se a una convocatòria electoral.

Cada elecció té una data de convocatòria (que l'identifica) que és la del dia en que es fan les votacions. El país està dividit en autonomies, que a la vegada poden tenir una o més circumscripcions electorals. El número d'escons que corresponen a cada circumscripció és fix en cada convocatòria (depèn del padró d'habitants de la circumscripció). Interessa enregistrar el resultat, en número d'escons, de les diferents convocatòries aconseguit per cada partit o coalició en cada circumscripció.

Els polítics dels partits que es presenten a les convocatòries d'eleccions s'agrupen en forma de candidatures. Els partits o coalicions presenten aquestes candidatures a totes o només a algunes de les circumscripcions electorals. Un partit o coalició pot presentar polítics independents a les seves llistes.

Voldrien, també, tenir constància dels partits i/o coalicions "afins" sigui per la ideologia o sigui per col·laboracions en els poders executiu o legislatiu. Es suposa que només volen tenir constància de l'afinitat vigent o actual.

13. El cinema

Un cinèfil afeccionat a la informàtica vol crear una base de dades que reculli informació diversa sobre el món cinematogràfic des dels orígens del cinema fins a l'actualitat, al·l el contingut que es descriu a continuació.

Lògicament vol tenir classificades moltes pel·lícules que vindran identificades per un codi. De cadascuna en vol el nom, l'any d'estrena, el pressupost, el director, etc. A més, de cada pel·lícula vol conèixer també quins actors hi van intervenir així com el paper que hi representaven (actor principal, secundari, etc.) i el possible premi que van rebre per la seva interpretació.

Les pel·lícules són d'un tema determinat. Es ben sabut que hi ha actors especialitzats en un tema encara que un actor és capaç d'interpretar varis temes amb diferent habilitat.

Com que el nostre cinèfil és una mica curiós, vol emmagatzemar també dades personals dels actors, que ha anat recollint en llegir diferents revistes del món artístic. Per exemple, quins actors són en certa manera substitutius d'altres amb un grau de possible substitució que pot anar d'1 a 10. També quins actors són incompatibles, és a dir, que mai han treballat ni treballaran junts en una mateixa pel·lícula o escena.

Els actors estan contractats en un moment donat per una companyia, però poden canviar si tenen una oferta millor. Fins i tot poden retornar a una companyia en la que ja havien treballat. Les companyies produeixen pel·lícules però cap pel·lícula és coproduïda per dues o més companyies.

Com que el nostre amic fa molt de turisme vol saber per a cada ciutat, quines companyies hi tenen representació i a quina adreça. Evidentment, les companyies solen tenir representació a quasi totes les ciutats importants. Al mateix temps, vol també informació de quines pel·lícules s'estan rodant a cada ciutat i en quin moment, tenint en compte que una pel·lícula es pot rodar a varies ciutats i també en una mateixa ciutat en diferents fases del rodatge.

14. La publicitat a TV

Es proposa dissenyar una base de dades per millorar el control dels espots publicitaris per televisió i de tot el seu entorn, com són els canals de TV, franges horàries, agències de publicitat, tipus de productes que s'hi anuncien, etc.

En el país hi ha diverses entitats de televisió. Volem reflexar que unes són de titularitat pública (per exemple CCRTV - *Corporació Catalana de Ràdio i Televisió*-, RTVE -*Radio Television Española*-) i d'altres de titularitat privada. Aquestes entitats disposen d'un o més canals d'emissió (per exemple, CCRTV disposa de TV3 i Canal 33, RTVE disposa de TVE1 i TVE2, etc.) que venen identificats sempre per un nom.

Suposarem que tots els espots publicitaris són identificats per un codi assignat per una suposada *Oficina de Medis de Comunicació*. Cal considerar l'existència dels espots equivalents com podrien ser aquells que amb les mateixes imatges tenen diferent idioma o per altres causes que no interessa precisar.

Cada spot fa referència a un (situació normal) o més (situació excepcional) tipus de productes (per exemple, el cas de marques de rentadores i de detergents que anuncien els seus productes en un mateix spot). Fins i tot s'ha de pensar que hi pot haver tipus de productes independentment de si hi ha o no espots que s'hi refereixin. Els espots són sempre propietat d'una única firma comercial que es qui els paga, fins i tot, en el cas d'anunciar més d'un tipus de producte. Els espots són filmats normalment per una agència de publicitat encara que en alguns casos és la mateixa firma comercial que amb mitjans propis els produeix sense intervenció de cap agència publicitària.

De les agències de publicitat interessa conèixer el nom del director artístic. De les firmes comercials interessa conèixer el nom del cap de màrqueting. Però tan d'unes com de les altres, que podem considerar sota una nova entitat anomenada empreses, interessa saber el codi d'identificació fiscal, el nom, l'adreça i el telèfon.

Les agències de publicitat disposen de directors cinematogràfics per filmar els espots i interessa conèixer la història de contractació per part de les diferents agències d'aquests directors. Un director treballa, en un moment donat, per una sola agència. Les agències volen conèixer, per cada spot, quin director l'ha dirigit. Dels espots produïts directament per les firmes comercials no interessa ni es vol constància del seu director.

Tal com s'ha indicat a l'inici, interessa també conèixer l'emissió dels espots en els diversos canals televisius. A efectes d'audiència, les 24 hores del dia no s'entenen com a hores sinó com a franges horàries, com poden ser *matí, migdia, vespre, nit*,... perfectament diferenciades i unificades per l'*Oficina de Medis de Comunicació*. Cada franja horària té un preu vigent (únic preu que interessa mantenir a la base de dades). Per simplificar el cas, farem la hipòtesi que el preu no depèn del dia de la setmana. Es vol constància del número de vegades que cada spot s'ha emès en els diferents canals, la seva data d'emissió i la corresponent franja horària.

Finalment hi ha un aspecte legal molt important a considerar. Existeixen alguns tipus de productes, realment molt pocs, que no està permès d'anunciar en certes franges horàries. L'incompliment serà penalitzat amb multes de gravetat qualificada d'1 a 3. Cal tenir constància, a la base de dades, de les prohibicions esmentades.

15. L'empresa de mercaderies

Cal dissenyar una base de dades per a una empresa dedicada al transport internacional de mercaderies amb extensió a tot l'àmbit europeu. El dissenyador després d'estudiar l'empresa ha pres nota dels següents aspectes importants.

L'empresa disposa d'un determinat nombre no massa gran de delegacions i d'una gran quantitat d'oficines repartides unes i altres per tota la geografia europea. Cada delegació té la responsabilitat d'un bon nombre d'oficines de la qual depenen. Tan les delegacions com les oficines s'identifiquen per un nom. En una determinada ciutat poden haver-hi vàries oficines, però mai més d'una delegació. S'ha de suposar que no hi ha ciutats amb nom repetit (al menys en l'àmbit d'aquest cas).

El personal de l'empresa es pot diferenciar en dos grans grups:

- Administratius, dels que interessa el nivell d'estudis a codificar per una lletra
- Conductors, dels que interessa l'any d'obtenció del permís de conduir i el tipus de permís.

De tot el personal de l'empresa es vol saber el codi d'empleat (que l'identifica), el nom, l'adreça, l'any de naixement i l'estat civil.

Els administratius es consideren assignats a una determinada oficina, mentre que els conductors es consideren assignats a una determinada delegació. Evidentment, uns i altres poden anar canviant d'oficina o delegació respectivament i es vol constància històrica d'aquests fets.

L'objectiu de l'empresa és aconseguir el màxim nombre de clients i efectuar els viatges pertinents per tal de transportar les mercaderies segons la petició d'aquests clients, els quals estan identificats per un codi i pels que es vol conèixer el nom, telèfon i persona de contacte, etc.

Evidentment, l'empresa per a dur a terme la seva activitat, disposa de molts camions identificats per un codi de camió. Dels camions es vol saber també la matrícula, marca, tara, etc.

Els viatges els organitza sempre una delegació i s'identifiquen per un codi intern per cada delegació. Per cadascun dels viatges cal conèixer:

- Camió utilitzat (cada viatge utilitza un únic camió).
- Conductor(s) utilitzat(s) en el viatge. Es vol saber l'import de les dietes pagades a cada conductor (que pot no ser la mateixa pels diferents conductors d'un mateix viatge).
- Número de paquets carregats i paquets descarregats en cada ciutat i per cadascun dels clients. En un mateix viatge es pot deixar i/o recollir paquets a diferents ciutats per a un mateix client.
- Recorregut del viatge, és a dir, dia i hora en que el camió va parant o simplement passant per les diferents ciutats per carregar o descarregar. Suposarem que un viatge no passa ni s'atura mai dues vegades en una mateixa ciutat.

16. Les cases de colònies

El departament d'infància i joventut d'un determinat país està interessat en informatitzar en alguns aspectes la gestió de les cases de colònies o esplai. Aquestes cases, que s'identifiquen per un nom, estan situades en diferents comarques catalanes encara que no uniformement repartides (de manera que n'hi ha moltes a platja i a muntanya però poques o cap en zones d'interior).

En aquestes cases s'organitzen "tandes" que són sempre del dia 1 al 15 o del dia 16 a fi de mes (durant tots els mesos de l'any i no només a l'estiu). Les tandes les identificarem per la seva data inicial (any-mes-dia). En les cases es realitzen diferents activitats (com natació, esquí, rem, pintura, fotografia, música, etc.) Segons la situació i instal·lacions de cada casa, aquestes activitats tindran un nivell de qualitat més alt o més baix que qualificarem numèricament.

Als nens que assisteixen a alguna casa de colònies se'ls lliura un carnet numerat, atorgat pel departament d'infància i joventut, la primera vegada que assisteixen a una tanda i que ja servirà per sempre. Aquest número serà l'identificador del nen o jove, del que també es vol conèixer el nom, cognoms, adreça, sexe i data de naixement. Per alguns efectes cal tenir constància dels germans entre tots els que estan enregistrats a la base de dades. A part del carnet en el moment de cada inscripció en una tanda d'una casa se'ls dona un número d'ordre d'inscripció.

El personal de les cases de colònies del que sempre interessarà el seu dni, nom, cognoms, adreça i telèfon, està agrupat en dos tipus de persones:

- L'administrador, que és únic per cada casa durant un any natural, pel que es vol tenir constància del sou que ha cobrat cada any.
- Els monitors, que es comprometen a assistir del primer al darrer dia d'una tanda en una casa, no canviant mai de casa dins la mateixa tanda, però podent comprometre's a diverses tandes i cases en la seva vida.

Dels administradors es vol conèixer el seu número de la seguretat social. Dels monitors la seva data de naixement. És possible que un administrador hagi estat monitor en la seva joventut.

Dels monitors es vol saber la tasca que han realitzat cada dia entenent el dia com a relatiu dins la tanda, és a dir, el primer dia, el segon dia,... en les tandes on han participat.

17. La cadena de supermercats

Una cadena de supermercats, estesa per tota la geografia estatal, ens ha sol·licitat el desenvolupament de les aplicacions necessàries per a gestionar, des de la seva seu central, la informació referent a compres, vendes, estoc, personal, clients,...

Aquesta cadena té un bon nombre de centres, cadascun dels quals té un nom pel que és conegut dins l'entorn geogràfic, però es volen tenir codificats de manera que no hi hagi lloc a ambigüitats entre diferents ciutats. Per cada centre es vol tenir NIF, direcció, població, codi postal i telèfon.

Per aquesta cadena hi passen un bon nombre de treballadors, els quals són contractats per un centre. Interessa tenir constància de tots els contractes que hi ha hagut en els diferents centres, amb una qualificació d'0 a 5 pel treballador, la qual haurà estat donada pel director del centre on està contractat. Un treballador pot tornar a treballar en un centre tantes vegades com vulgui (o, millor dit, tantes vegades com el contractin). Interessa tenir el nom del treballador i adreça, població, codi postal i telèfon més actuals. Els treballadors venen identificats per un codi de treballador. Pels treballadors espanyols ens interessa tenir el dni, i pels estrangers, el número de passaport i la seva nacionalitat (codificada, evidentment). Així mateix ens interessa saber quins treballadors estan actualment contractats en algun centre.

Els centres d'aquesta cadena són molt grans i estan subdividits en departaments. Cada centre assigna un nom als diferents departaments en què està estructurat i un codi que els identifica dins del centre. Cada centre t, un director general. Interessa tenir-ne constància, a l'igual que del director de cada departament i les dates des de que en són directors.

És evident que aquesta cadena de supermercats té proveïdors que subministren els articles que posen a la venda. Aquests proveïdors estan codificats i interessa tenir constància de NIF, raó social, direcció, població i telèfon. La seu central decideix els articles dels diferents proveïdors que els diferents centres podran vendre. A tal efecte codifiquen tots els productes i els assignen una descripció i un PVP. Interessa guardar l'únic proveïdor que li subministra el producte, la referència del producte pel proveïdor, el múltiple d'unitats que cal comprar i el preu de cost de cada unitat.

Les comandes dels diferents productes als proveïdors s'efectuen a nivell de departament de cada centre. Cada departament assigna un número a la comanda. Interessa tenir constància de les comandes existents, la data i els productes que la formen. Quan es fa una comanda, el departament sol·licita al proveïdor una data màxima de servei. Interessa tenir-ne constància, a l'igual que la data de rebuda (per penalitzar econòmicament al proveïdor si és posterior a la data màxima de servei). Per simplificar considerem que un proveïdor serveix tota la comanda sencera o no la serveix.

La base de dades també ha de permetre gestionar l'estoc de productes a nivell de departaments. A tal efecte, cada departament fixa, per cada producte, un mínim d'unitats que caldria tenir sempre en estoc. Interessa tenir-ne constància, a l'igual que la quantitat real emmagatzemada.

Si important és la gestió de comandes, no menys important és la gestió de vendes. Cada departament efectua múltiples vendes, les quals va numerant consecutivament. Interessa tenir constància del detall de la venda i del tipus de pagament, que pot ser en efectiu, amb targeta de crèdit o amb la targeta de client dels propis supermercats. En el darrer cas interessa tenir constància de la targeta a la què s'ha carregat l'import de la venda. Aquestes vendes han de quedar detallades per la posteritat.

Aquesta cadena de supermercats té establert unes franges horàries de treball. Els empleats poden anar canviant de franja. Interessa saber la franja en què està treballant cada empleat i la data en què va començar. És clar que el departament de personal preveu amb temps el canvi de franja i, per tant, interessa que la base de dades pugui emmagatzemar no només la franja de treball actual sinó també les properes que estiguin definides, amb la data d'inici de cadascuna d'elles. Les franges horàries poden tenir un plus per nocturnitat i/o festivitat.

I com no, també es vol emmagatzemar les dades dels clients que disposen de tarja de client dels magatzems. Els clients tenen un codi que els identifica. Cal saber-ne també el dni, població, telèfon i codi postal. Tot client té una targeta al seu nom, però també pot donar permís a altres persones (normalment de la seva família) per a tenir una targeta que carregui a la seva compte bancària. Cal enregistrar totes les targetes associades al client, el número de compte bancària on es carrega l'import de totes les targetes associades i el nom i dni de la persona titular de cada targeta.

18. L'ICS

Com molt bé sabeu, l'Institut Català del Sol fa diverses actuacions en els diferents municipis de Catalunya referents a proporcionar noves àrees urbanitzables.

En primer lloc cal dir l'ICS anomena POLÍGONS a les àrees que urbanitza. Tot polígon té un codi, un nom i una extensió. Cal tenir molt clara la distinció entre polígons residencials, destinats a habitatges, i polígons industrials, destinats a indústries. Ens interessa tenir constància sobre els municipis als que pertanyen els diferents polígons de l'ICS. Els municipis s'identifiquen per un codi i ens interessa tenir constància del seu nom, nombre d'habitants i la comarca a la que pertanyen.

En els polígons industrials s'hi instal·len indústries. L'ICS vol conèixer les indústries que utilitzen els polígons industrials. A tal efecte els hi assigna un codi a cadascuna i vol emmagatzemar la raó social i el domicili, telèfon i fax de la seva raó social. Cal tenir en compte que una mateixa indústria pot estar instal·lada en diferents polígons industrials. Es vol saber a quins polígons hi té alguna nau i quantes naus té a cada polígon.

Els polígons residencials estan subdividits en parcel·les cadascuna de les quals està destinada a un habitatge d'un determinat tipus. Els tipus d'habitatges (unifamiliars, blocs de pisos,...) estan codificats. Les parcel·les s'identifiquen dins els seu polígon per un codi i tenen un carrer i un número. Es vol tenir també el màxim número de plantes que està permès d'edificar.

Les parcel·les, mitjançant concurs, s'adjudiquen a una única persona o empresa constructora. Tant si es tracta d'una persona com d'una empresa, l'ICS l'identifica pel NIF. Per les constructores, l'ICS vol guardar la raó social i el domicili, població, telèfon i fax de la raó social. Per les persones, l'ICS vol enregistrar nom, domicili, població i telèfon. Cal tenir present que una persona o empresa pot tenir assignades vàries parcel·les.

19. Els cicles formatius a Catalunya

Volem crear una base de dades que reculli la informació sobre els cicles formatius (CF) que es cursen a Catalunya.

Cal tenir informació sobre els CF que el Departament d'Ensenyament de la Generalitat ha elaborat. Els CF estan agrupats en famílies professionals. Cada família té un codi que la identifica, un nom i un responsable dins el Departament d'Ensenyament. Els CF venen identificats per un codi i interessa saber quin és el seu nom i el seu nivell (mig o superior), la família a la que pertanyen i el número de DOGC en el que es va publicar el corresponent disseny curricular.

A Catalunya hi ha diferents centres que imparteixen CF. Cada centre té un codi que l'identifica i un nom. Dels centres interessa saber si la seva titularitat pertany a la Generalitat, a l'Ajuntament del corresponent municipi, si són privats docents o privats no docents (com pot ser el cas d'empreses automovilístiques que imparteixen CF de la família professional d'automoció). Evidentment cal tenir constància de l'adreça, telèfon, codi postal i municipi dels centres.

Interessa també poder esbrinar els CF que cada centre ha impartit fins l'actualitat amb els corresponents cursos i el número de DOGC en que es dona la corresponent autorització i el curs en que es va començar. Pot donar-se la situació que un CF hagi deixat d'impartir-se en un centre durant un període de temps per causes que no venen al cas, situació en la que cal poder saber el darrer curs en què es va impartir i el número de DOGC en que es comunica el cessament.

Volem saber també, per cada curs i centre i CF el número d'alumnes matriculats a cada curs, el número de baixes i el número d'alumnes que han superat el curs (tots els crèdits del curs superats). Cal tenir en compte que un CF s'estructura en un o dos cursos acadèmics, tot i que alguns centres poden tenir autoritzacions especials (tipus d'alumnat que obliga a estructurar el CF en més cursos per possibilitar l'assistència).

En els CF hi ha un crèdit anomenat *Formació en Centres de Treball* en el que l'alumne ha de desenvolupar pràctiques reals en una empresa del sector. Interessa saber totes les empreses amb el seu nom, adreça, telèfon, població i codi postal que han proporcionat alguna vegada pràctiques. A la vegada també es vol saber quants alumnes de cada centre han fet pràctiques a cada empresa per un curs acadèmic i un CF.

20. El turisme

Com que estem en un país molt turístic s'ha pensat en la conveniència de dissenyar i construir una base de dades que proporcioni informació relativa a hotels, empleats, agències de viatge, etc.

Els hotels tenen noms que no es repeteixen a una mateixa ciutat, però sí que poden repetir-se a diferents ciutats. Dels hotels interessa conèixer d'entre altres informacions, la categoria (nombre d'estrelles), adreça i telèfon. Els hotels que, independentment de la seva categoria tinguin semblances per alguna raó (que no ens interessa) en direm hotels "equivalents".

A efectes d'aquesta base de dades, només considerarem els hotels que en podríem dir "turístics", que classificarem en dos tipus ben diferenciats: hotels de platja i hotels de muntanya. Dels hotels de platja interessa saber si disposa o no de platja privada i si hi ha o no lloguer d'embarcacions. Dels hotels de muntanya interessa saber si té o no piscina i a més quina(es) activitat(s) organitzen, com "rafting", "mountain bike", sortides a cavall, etc. Segons les instal·lacions de l'hotel, cadascuna d'aquestes activitats tindrà un nivell de qualitat adequat. La majoria d'hotels (però no tots) pertanyen a cadenes hoteleres, que en són les propietàries. De les cadenes hoteleres, que s'identifiquen pel seu nif, s'ha de saber el nom i l'adreça fiscal.

Un important col·lectiu a considerar a la base de dades són els empleats identificats pel seu dni. Només tindrem informació dels que són contractats per les cadenes hoteleres i caldrà tenir constància de l'historial de la contractació d'aquests empleats. Aquest empleats tenen molta mobilitat i, per això, es desitja conèixer per a cada empleat el seu historial laboral en els diversos hotels (de la base de dades) on ha estat (o està treballant). Hem de destacar que un empleat no només pot treballar en els diferents hotels de la cadena que l'ha contractat sinó que pot treballar en altres hotels (per exemple) en condicions de cedit.

Volem tenir informació sobre el nom i categoria d'algunes agències de viatges i de les seves delegacions (interessa l'adreça i el número de fax) en algunes ciutats. Les agències les identificarem pel seu nom i les delegacions per un número propi assignat per cada agència.

Finalment voldrem informació del nombre de turistes que les agències de viatge assignen als diferents hotels en cadascun dels mesos de l'any.

21. L'hotel

Un hotel de la Riviera Maya, molt ben valorat, disposa de 2 classes d'habitació, les bàsiques i les luxe, cada una d'elles ubicada en una de les seves 3 plantes identificades per un número únic dins la planta. De totes les habitacions ens interessa saber si tenen terrassa o balcó, llit normal o doble, m2, i per les habitacions de luxe (les úniques que tenen jacuzzi), si el tenen a l'habitació a la terrassa, i en qualsevol cas cc3 de capacitat.

El personal de l'hotel es distribueix en alimentació / neteja / organització. De tots els treballadors interessa saber el nom, dni (que els identifica), direcció, telèfon. Cada membre de neteja té assignades unes determinades habitacions, de les que se'n responsabilitza. Tota habitació té un únic responsable (de neteja) en un període determinat de dates.

Cada planta de l'hotel té un responsable (de la secció d'organització), a qui els clients es dirigiran per efectuar preguntes o queixes (que fins al moment són inexistents).

Els clients ingressen a l'hotel en una determinada data, i per un període fixat, i resideixen sempre en una habitació.

És cortesia de l'hotel, oferir sopars especials als clients: Per aquest objectiu, en les instal·lacions tenen 4 restaurants temàtics (Mediterrani, Brasileny, Japonès, Francès), cada un amb els seus plats típics, i els clients elegeixen a quins restaurants volen sopar. No cada dia se sopa en restaurant temàtic, alguns dies poden sopar en restaurant de bufet lliure o a la carta amb plats típics mexicans. Interessa tenir constància d'aquest fet, del dia en que cada client ha sopat a cada restaurant, i del restaurant temàtic on hi ha (de forma definitiva) el personal d'alimentació

Finalment alguns clients han conegut l'hotel a partir de referències d'altres clients que havien passat anteriorment, interessa conèixer aquesta relació entre clients.

22. Futbol professional

1. La base de dades ha de gestionar les dades relatives al futbol professional. Interessa aquella informació sobre: enfrontaments, persones, resultats, ..
2. Cada any hi ha una competició futbolística entre els diferents equips. Hi ha dues voltes, en les quals cada equip juga contra tots els altres (un partit a casa i un altre a fora)
3. Per cada equip: els jugadors, dels quals en volem saber nom, dni (que els identifica), data de naixement i a quina posició volen jugar; el/s entrenador/s i altres persones o empreses vinculades (transport de jugadors, subministrament d'equip, ...)
4. Cada setmana, (una jornada), tots els partits s'enfronten, excepte, que siguin imparells, aleshores, un equip descansa. Volem saber quin dia i a quina hora hi ha hagut cada enfrontament.
5. També volem saber a quina posició ha jugat cada jugador, i si era titular o suplent, així com els minuts jugats. Volem tenir informació dels gols marcats: quin jugador l'ha marcat, a quin minut i si ha estat de penal.
6. També és d'interès conèixer les faltes que s'han comès, i per cada falta, si s'ha vist targeta (color) i si algun jugador s'ha vist lesionat.
7. Finalment, per cada jugador lesionat, volem saber, la data de baixa, la causa i la possible data d'incorporació.
8. Cal que la base de dades ens permeti saber pels gols marcats de penal, el jugador que ha causat la falta i, que en conseqüència ha generat la ocasió.

23. El restaurant

Un important restaurant d'Igualada (cdm), ens ha demanat que li dissenyem una base de dades que posteriorment utilitzaran algunes aplicacions. Les dades d'aquesta BD, es troben implícites en les següents situacions:

El restaurant es divideix en zones (taules, taules íntimes, terrassa ...), que s'identifiquen per un nom breu que ajudi a ubicar-les. Cada una d'aquestes zones té diverses taules, identificades per un número (únic dins de la zona). Interessa conèixer la capacitat màxima i mínima de cada taula.

El restaurant només serveix dinars i sopars i si hi ha reserva prèvia. Per tal de reservar taula, cal indicar el client, dia i hora de la reserva i taula on volen menjar. (Per simplificar, suposem que cada taula només es servirà un sopar per nit o dinar per migdia)

El restaurant ofereix diversos plats (cada un d'ells està pensat per ser servit com a entrant, primer, segon o postres), té un nom, un ingredient principal, i diversos ingredients secundaris. Cal guardar el preu de cada plat, i les proteïnes i greixos de cada ingredient. Cal guardar quins plats s'han servit diàriament per taula. Existeixen plats alternatius a cada plat ofert, per si algun client demana un plat que s'ha acabat.

Aquest any, han introduït la possibilitat del "Sopar-Regal o Dinar-Regal", de manera que es fa una reserva per un amic / familiar, el qual menjarà i no pagarà, i el reservant passarà a posteriori per pagar la compta (tot, menys una paga i senyal, lliure, que haurà entregat el reservant a l'hora de fer la reserva).

Finalment, destacar que els cambrers estan destinats a una taula (que pot variar diàriament), i serà el cambrer l'únic responsable de fer l'estància del client el més agradable possible.

De totes les persones (clients, cuiners i cambrers), interessa saber el nom, i un codi únic (intern pel programa per identificar-lo). Opcionalment, els clients poden donar adreça i/o mail, per estar informat de qualsevol novetat del restaurant. Finalment, només el personal (cambrers i cuiners) donaran dades econòmiques (compta corrent) per cobrar la seva nòmina.

24. Monsters of Rock

Després de l'èxit fulgurant de Roberto Chiquilicuatre a Eurovision (any 2008), el nostre amic Paco ha decidit dedicar la seva vida a la música. Abans de fer els seus primers passos, vol aprendre dels grans mestres, i s'ha proposat construir una base de dades per a seguir els seus grans ídols musicals.

El Paco ha revisat tots els discs i ha fet una llista de tots els grups que li agraden, de cadascun en té el nom, l'any de formació del grup, i els integrants originals (els músics que van crear la banda). També a numerat els grups per tenir-los sempre identificables. Com que és un currante, en Paco s'ha fet també una llista de tots els músics que coneix de totes les bandes. Per cada músic en sap el nom, cognoms, nom artístic, data de naixement i data de defunció (bé, aquest últim no sempre, es clar). També ha buscat la ciutat on van néixer cada músic. Com que la base de dades podrà tenir músics d'arreu del món caldrà saber també el país d'origen.

Tot i que els grups musicals els funden una sèrie de persones, els integrants dels grups solen anar canviant, alguns marxen a d'altres grups, d'altres formen el seu propi grup en solitari, i d'altres moren ofegats en els seus vòmits.... Sigui com sigui, es vol tenir constància de les persones que integren els grups, i dels canvis en els seus integrants. En Paco s'ha adonat que en alguns casos hi ha músics que han tornat a la seva banda original després de passar per d'altres formacions.

També seria interessant tenir enregistrada informació més detallada dels components dels grups, els músics poden ser instrumentistes o vocalistes, tot i que de vegades ens trobem que hi ha monstres que fan els dos papers (toquen la guitarra i canten, o en un grup són cantants i en un altre bateries !) Voldrem saber els instruments que és capaç de tocar cada instrumentista, i amb quin nivell (geni, mestre, pro, aficionat). En el cas dels vocalistes només en sabrem els anys d'experiència cantant.

Quan un grup contracta a un músic en agradaria saber quin és l'instrument principal que tocarà al grup i si cantarà (si o no). Per exemple, Ladislao Fiero és cantant i toca els teclats, la guitarra i el baix. En el seu primer grup , "*La marmota envenada*", ell tocava el baix i canta. Un any després va canviar de grup a "*Animales muertos*" i allí toca la guitarra.

Els grups treuen al mercat àlbums, que tenen un nom únic, i dels que en Paco vol guardar-ne l'any de publicació i la duració en minuts. De cada àlbum en sabrem totes les cançons, que s'identifiquen dins del disc amb un número d'ordre. De cada cançó se'n guardarà el títol i la duració. A banda de les cançons, els grups molt sovint confessen que estan influïts per d'altres grups que els han precedit. Per exemple, el grup "*Deshollando al cochino*" s'han vist molt influïts per *Nickelback*, *Whitesnake* i *Iron Maiden*.

Una de les coses més importants per la base de dades és que tingui un registre de les gires que munta cada grup. Habitualment les gires s'identifiquen amb el seu nom (amb noms com "*A Matter of Life and Death World Tour*") i en sabrem la data d'inici i de fi. A la gira es fan molts concerts per diferents ciutats, i és típic dels grups grans que es facin més d'un concert a la mateixa ciutat en dies consecutius abans de prosseguir la gira a un altre ciutat. Volem tenir tota aquesta informació a la base de dades.

Respecte a les gires, sovint s'organitzen per a promocionar un àlbum del grup, si aquest és cas, voldrem tenir forma de saber l'àlbum que dona origen a la gira. Així mateix, quan els grups grans van de gira, és habitual endur-se d'altres grups com a teloners.

A cada concert el grup porta una llista numerada de cançons que es tocaran, i de vegades si el públic es posa pesat poden tornar a tocar algun tema d'èxit. Es voldrà tenir constància del repertori tocat a cada concert de la gira.

Per acabar, en Paco també voldria saber quines cançons del disc dels àlbums s'han gravat en estudi i quines en directe. Si són en directe, haurem de poder saber a quin concert es van enregistrar.

25. Ciclisme

El director de la federació catalana de ciclisme vol fer una base de dades on guardar tota la informació relativa a curses ciclistes que es corren cada any. A tal efecte ens les dades més significatives per a dissenyar la base de dades.

En el món del ciclisme existeixen un nombre reduït d'equips ciclistes registrats a la UCI (Unió de Ciclistes Internacional), cadascun amb el seu nom i patrocinador. A efectes pràctics la UCI assigna un codi a cada equip per identificar-los. Els equips contracten ciclistes federats per a córrer les curses d'una temporada., cal dir que només volem emmagatzemar els equips vigents dels corredors, no ens cal saber on corrien anteriorment els ciclistes, només el contracte actual. A més, s'enregistraran les dades personals de cada corredor (NIF; nom, cognom, data de naixement, i adreça).i dades fisiològiques pel control antidopatge (capacitat toràcica, pulsacions, índex de globulina)

Els equips són liderats per un únic director esportiu, del qual en tenim les dades personals (NIF, nom, etc.). Es dona sovint el cas els de directors esportius han estat ciclistes anteriorment, però no sempre és així.

Generalment una setmana abans de la celebració d'una competició, els equips inscriuen a alguns (o tots) els seus corredors. Les competicions s'identifiquen amb un codi i tindran un nom i una data de celebració. És necessari saber també el nombre total de kilòmetres que té el traçat. Les curses de ciclisme solen tenir dues estructures ben diferenciades, per un costat tenim les curses clàssiques que són d'un dia (la clàssica de San Sebastián, el mundial...), i les curses per etapes (o voltes, com ara el *Tour de França*) que es perllonguen durant varies jornades. Les clàssiques s'organitzen a una ciutat concreta, que necessitem enregistrar (penseu que les curses són a nivell europeu). Pel que fa a les curses per etapes, l'organització prepara les jornades o etapes, assignant-li un número d'ordre (etapa 1, 2, etc.). Cada jornada té un kilometratge., un temps de tall previst per l'organització, i s'inicien i acaben a una població (que pot ser o no la mateixa).

En el transcurs d'una competició es duu un registre del temps acumulat, posició parcial i puntuacions (veure punt següent) de cada corredor inscrit. Aquestes dades es van actualitzant després de cada jornada, desant només el valor actual.

A les carreres per etapes, com ara el *Tour de França* o la *Volta* a més de premiar al corredor que arriba el primer (el de millor temps) es tenen en compte dues categories més de premis: el premi de muntanya i el premi de la regularitat. Cada corredor pot acumular punts de muntanya i/o punts de regularitat al llarg de la competició, i qui en té més a cadascuna de les categories guanya un o altre premi. L'organització desitja tenir constància dels punts de muntanya i de regularitat que tenen els ciclistes a la cursa.

Els punts de regularitat s'obtenen a les "metes volants" o "sprints". Les jornades poden tenir (opcionalment) varies metes intermèdies o "metes volants" a les que s'associa una puntuació determinada per a la competició de regularitat. Les metes volants s'identifiquen indicant la jornada a la que pertanyen i un número d'ordre (1,2,...) dins de la jornada. També es guarda el punt quilomètric on es situen. El sistema haurà de guardar l'ordre de pas dels corredors inscrits per la meta volant, i els punts de regularitat assignats. El primer que passa per la meta volant s'emporta tots els punts, i cada corredor posterior s'emporta un punt menys que l'anterior. Per exemple, la tercera etapa del *Tour* del 2006 hi havia 2 metes volants als quilòmetres 123 i 230. La primera meta tenia una puntuació de 3 punts i la segona de 6 punts.

Els punts de muntanya es guanyen a unes metes especials situades als ports de muntanya. A la base de dades es tindrà una llista completa de tots els ports de muntanya que existeixen a Europa, dels que en sabrem un codi, el país, el nom i la categoria (ESP, 1^a, 2^a o 3^a) (p.ex. FR023, França, l'Alpe d'Huez, ESP). Quan es decideix el recorregut de cada etapa, cal saber per quins ports es passarà i en quin punt quilomètric es farà. Com que de vegades els traçats de l'etapa són circulars es pot passar varies vegades per un mateix port ! També cal tenir en compte que alguns ports de muntanya del traçat de l'etapa poden tenir

assignats una puntuació pel premi de muntanya (p.ex. 6 punts), i el sistema ho ha de tenir emmagatzemat. Penseu que un mateix port de muntanya pot ser usat per moltes curses diferents.

Al llarg de cada etapa s'enregistra l'ordre d'arribada a cadascun dels ports de muntanya de l'etapa, i en el cas d'aquells ports amb premi, cal saber també la puntuació aconseguida pel corredor. (en general el primer s'emporta tota la puntuació del port, el segon un punt menys, el tercer 2 punts menys, etc.)