

UNO. Anuncios.

Tengo problemas de sincronización de documentos entre computadoras, respondo mensajes o dudas sobre las notas del examen hasta llegue a la casa. Para usar la compu de la casa, que estoy seguro todo esta actualizado.

Revisión 3era tarea: lunes 24 de noviembre y martes 25 nov. Los citas para revisar el martes se rifarán.

Fecha para 2do Examen: Miercoles 3 de diciembre, a las 930 am o 1230 am. (3 horas). Expectativa de resultados del examen, el viernes 5 de diciembre. Ya cada estudiante sabrá cuando puntos necesita para pasar.

Ultima clase, es el 21 de noviembre, es presencial por 2 motivos: me hacen la evaluación del profesor y hay quiz.

Mañana jueves se avisará para que envíen msj privado para consultar la nota previa. Que incluirá las notas del quiz de hoy.

4ta tarea para quienes ocupen realizarla seria el martes 9 de diciembre.

Entrega de actas para este curso: 10 de diciembre.

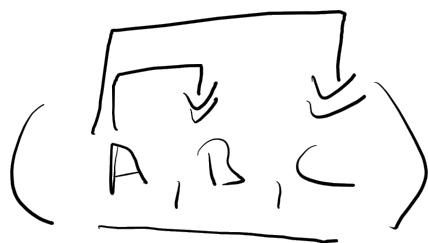
DOS. Revisión del quiz 2.

TRES. Seguimos con normalización.

4ta forma normal. El diseño esta en 3 NF y no existen 2 o mas DMV que son independientes entre si.

O sea, que se viola el 4NF (la anomalía) cuando hay 2 o mas DMV estan representadas y son independientes entre sí.

La pk es la unión de a, b y c



Y la DMV A=>>B es independiente de la DMV A=>>C, esto es que NO existe una narrativa coherente que defina una relación (frase) entre los atributos B y C, por lo tanto no deben estar en la misma tabla.

E(idEmpleado, IdIdioma, IdTitulo) PK compuesta de los 3 atributos.

No esta en 4NF, idEmpleado =>> idIdioma (es 1 a n, y hay narrativa, le puedo poner nombre al rol, "empleado domina n idiomas", idEmpleado =>> idTitulo (1 a n, "Empleado obtuvo n titulos"), PERO no hay una narrativa coherente entre IdIdioma Y IdTitulo

Para que si cumpla, hay que particionar.

E1 (IdEmpleado, IdIdioma, ...)}

E2 (IdEmpleado, idTitulo,)

Recordemos que un buen diseño conceptual, nunca habría combinado en una misma tabla las PK de Idioma y Titulo.

Otros ejemplos:

(idRestaurante, IdDistritoEntregas, idOpcionMenu)

Un Restaurante tiene muchas opciones de Menu (pizzeria La reina ofrece Pizza Suprema, Pizza Hawaiana, Pizza rustica, etc.), o sea hay una DMV

Un restaurante entrega en Paraiso, Oreamuno, Cartago, Tejar. Hay una DMV

No existe una narrativa entre distritos y opción de menú. O sea que las DMV son independientes. a Nunca le dirán "En Oreamuno no entregamo pizza suprema"

Otro ejemplo:

(idTexto, IdCurso, IdIdioma)

El idTexto es libro de texto de muchos cursos

El idTexto esta editado en varios idiomas

Existe narrativa entre Curso e Idioma?, si no, viola 4NF.

5NF. Un diseño esta en 5NF si está en 4NF y todas sus dependencias de reunión son triviales.

Una dr (dependencia de reunión) es trivial, si todas sus particiones son inducidas por la PK de la tabla original.

Definición de dr: es una partición vertical de un diseño, talque al hacer join de las particiones respecto de los atributos que se llaman igual, se obtiene la tabla original. Vamos a suponer que el particionamiento vertical hereda los mismos nombres de atributos de la tabla original.

Una dr, es la definición a la cual se recurre para documentar la anomalía que hace que no se cumpla 5NF, sin embargo la anomalía puede ser oscura, el profe solamente ha encontrado 2

ejemplos documentados en libros de texto, uno de ellos es el ejemplo arquetípico que vemos en esta clase (Proyectos, Partes y Proveedores).

Consideremos este diseño Empleado(idEmp K, nombre, fechaNacimiento, idPuesto)

Si hacemos una partición de Empleado, tal que al hacer join de la misma sobre los atributos que se llaman igual, obtenemos la tabla original, entonces tengo una dr.

dr1 (e1[idEmp K, Nombre], e2[idEmp K, FechaNacimiento], e3[idEmp K, IdPuesto]) ✓

```
select *  
from e1  
inner join e1.IdEmp=E2.Nombre ....
```



Esto es sin sentido, no se vinculan tablas sobre atributos que representan diferentes conceptos

```
select E1.IdEmp, e1.Nombre, e2.FechaNac, e3.IdPuesto  
from e1  
inner join e1.IdEmp=E2.IdEmp  
inner join e1.IdEmp= e2.IdEmp
```

Haciendo join de las 3 tablas, vinculando a través de atributos que se llaman igual (o sea de misma naturaleza), obtengo la tabla original, entonces la dr es válida.

Una partición es inducida por la PK, si la PK está presente en todas las particiones. ✓

dr2: (e1[idEmpleado, Nombre, FechaNac], e2[idEmpleado, IdPuesto]) ✓

dr3: (e1[idEmpleado, IdPuesto, Nombre], e2[idEmpleado, FechaNac]) ✓

dr4: (e1[idEmpleado, Nombre, FechaNac], e2[Nombre, IdPuesto]) ✗

en dr4, la única posibilidad de hacer join es vinculando nombre en ambas tablas, y ya que el nombre no es único, si tengo 2 nombres iguales, producirá 2 instancias de más al hacer join, por lo tanto la tabla resultante ya no es igual a la original. Se producen tuplas espurias, aquellas que no están en la tabla original.

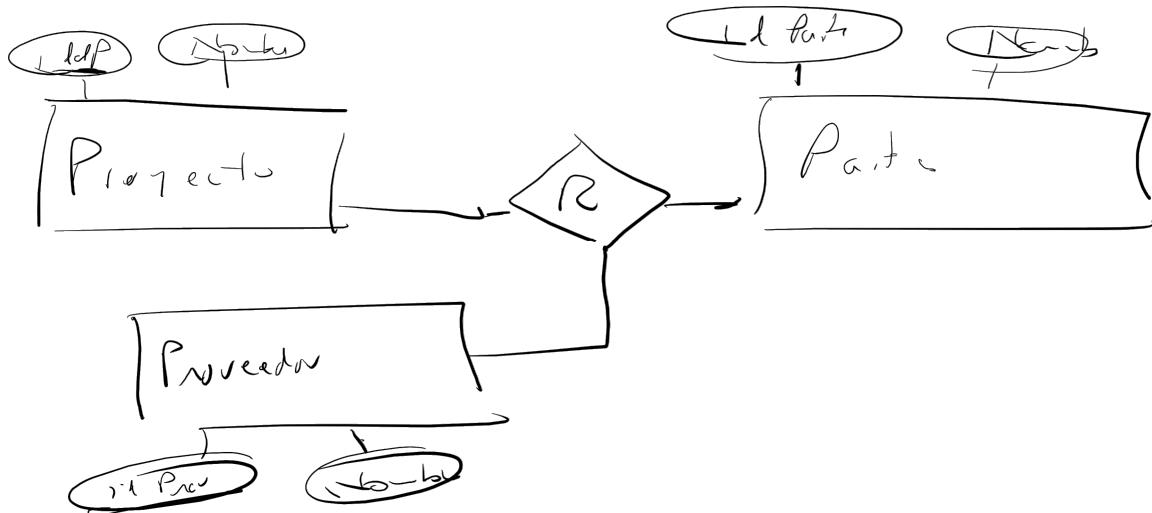
Si enumeramos todas las dr de Empleado, veremos que todas están inducidas por la PK, entonces si cumple 5NF.

En el caso de tabla Empleado, es autoevidente que está en 5NF, hay dificultades cuando tenemos tablas muy vinculadas, por ejemplo la tabla líneas (de la factura que veremos más abajo)

Aunque las dr nos ayudan a determinar si la tabla no está en 5NF, es más importante deducir lógicamente la razón del error en el diseño de tablas, o sea entender el pq el diseño no es una representación de la realidad.

Suponga un modelo que incluye los proyectos de una empresa constructora (construcción de casa, puentes, edificios), las partes que necesita cada proyecto (cemento, clavos de zinc, varillas, zinc, bloques, reglas de madera, vidrios ..etc) y los proveedores a quienes se les compra las partes.

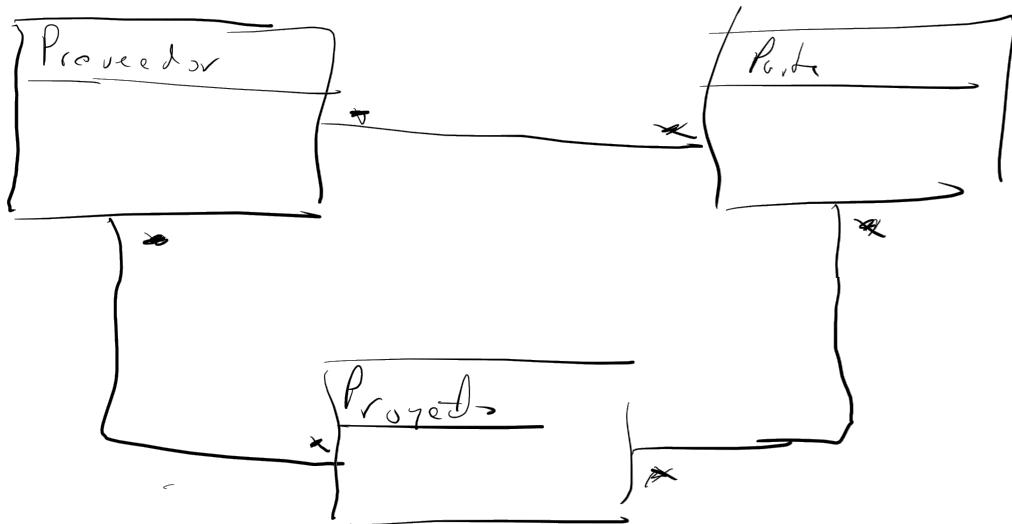
Existe una narrativa entre todos los Proyectos y la partes: La casa necesita laminas de zinc para protegerse de la lluvia). También existe una narrativa entre Proyecto y proveedores: El proyecto casa en San Vito compra partes en la Ferretería Central de San Vito (los proyectos de construcción se abastecen localmente de manera preferible) y también existe una narrativa entre Proveedor y partes. Ferretería EPA vende mangueras para proyecto Casa en San Vito, VICESA vende vidrios a casa en San Vito). Ya que existe una narrativa en los 3 elementos, hay tentación de modelarlo así, lo cual es permitido en Entidad-Relacion:



Cual es la narrativa de R: Proyecto P1 necesita Partes Pa1 y Pa2 que vende el proveedor Prov5, el problema es que como R es una relación sin limitaciones (de n a n entre 3 CE), no impide que existan tuplas espurias que representen asociaciones entre CE invalidas. Si VICESA (vidrios centroamericanos SA) solo provee vidrios, embargo no se impide que haya una relación entre Proyecto P1, Parte Tornillos y VICESA, porque R permite cualquier combinación. La narrativa es irrestricta, por lo tanto permite tuplas espurias, que no tengan sentido.

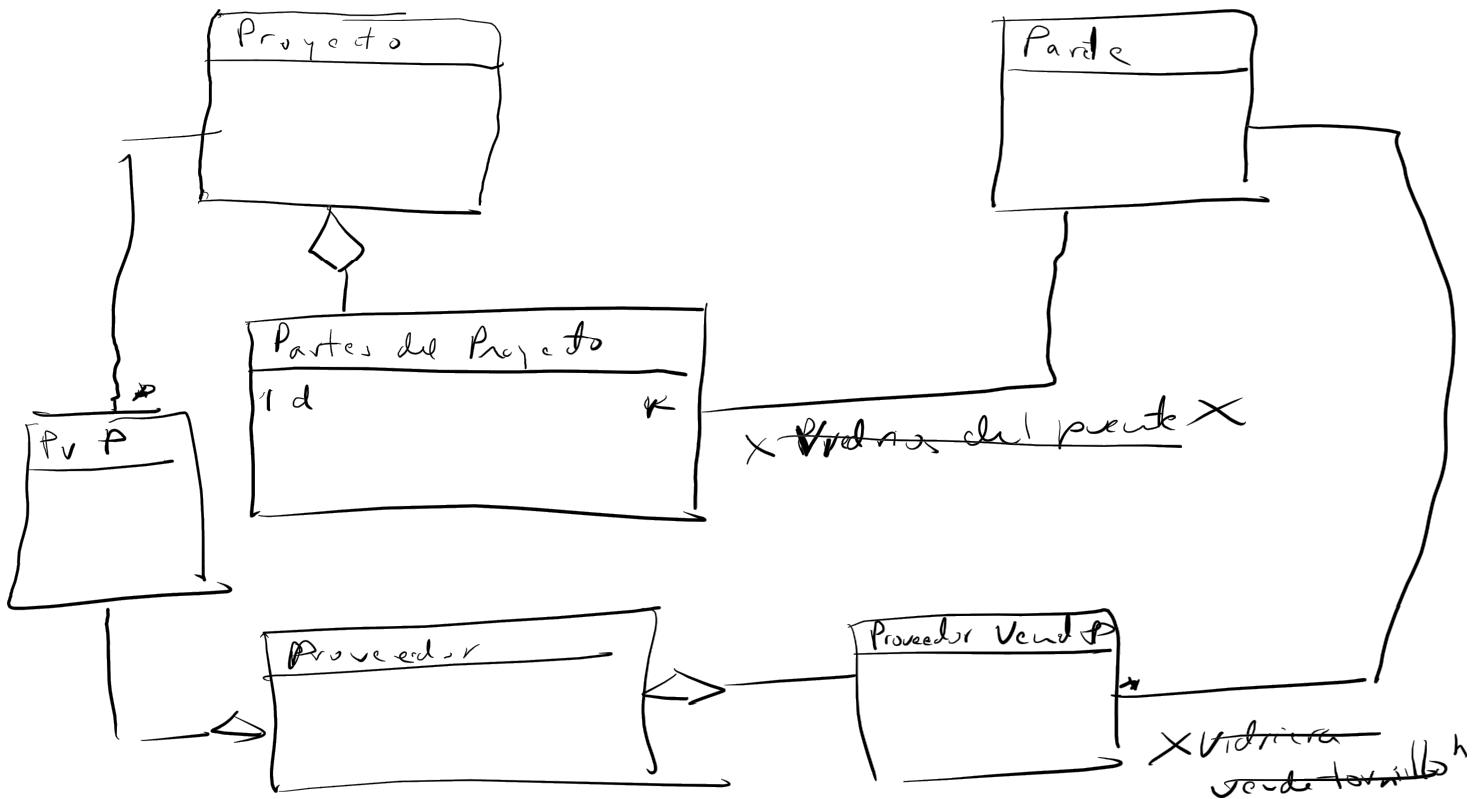
En UML, ya que no hay relaciones de 2 o más CE, no podríamos hacer una asociación de 3 CE, y las relaciones N a N se convierten en agregaciones y composiciones.

Remodelando en UML, son 3 relaciones n a n, de las que surgen 3 agregaciones.



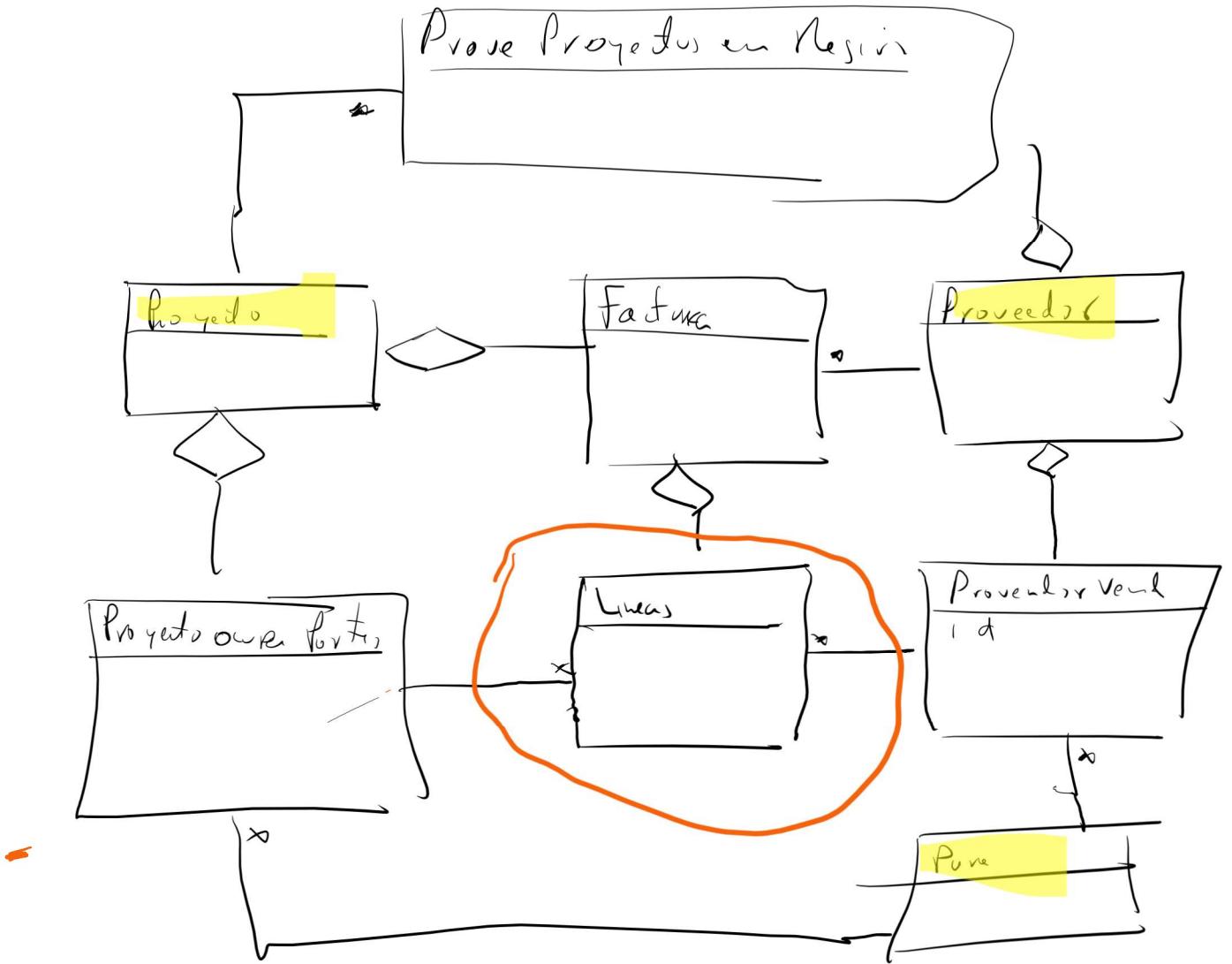
Sin embargo, es necesario de alguna forma representar una asociación entre los 3CE, talque Proyecto compra a Proveedor ciertas partes, pero solo le puede comprar aquellas parte que el proveedor provee. De esa manera evitamos que se modele la compra de una parte a un

proveedor y ese proveedor no provee la parte, por ejemplo: "Casa de Pedro compra tornillo a Vidriera", aunque si se debe poderse representar "Casa de Pedro compra tornillo a La casa del tornillo".



Este modelo protege de una tupla sin sentido que refleja ~~Ferreteria San Vito vende a Proyecto Casa en Los Chiles~~, por lo lógico es que la Ferreteria vende a clientes en su área regional.

Que relacione los 3 CE?, debemos encontrar una narrativa que asocie los 3 CE. Candidatos de esa narrativa es una Factura (del Proyecto, al Proveedor, comprando partes), otra narrativa podria ser "acarrea"...



Moraleja, no hacemos relaciones de 3 CE, representando muchos a muchos entre ellas, de manera directa. Solo representamos relaciones binarias (1 CE vs 1 CE), si son n a n, se convierten en agregaciones. La relación ternaria, no se hace de manera directa, siempre a través de relaciones binarias entre las agregaciones, por ejemplo: Lineas vs Factura (Factura a su vez asocia Proyecto y Proveedor), Lineas vs ProveedorVendeParte) y Lineas vs ProyectoOcupaPartes. Todas las asociaciones deben tener una narrativa coherente.

Importante: es poder justificar en términos de dr, que un diseño este en 5NF.

Como justificar que un diseño NO en 5NF?. NO lo vamos a preguntar, ni lo vamos a justificar teóricamente, para no oscurecer más.