

FUNDAMENTOS DE BASES DE DATOS

Examen Diciembre 2004

Presentar la resolución del examen:

- Con las hojas numeradas y escritas de un solo lado.
- Con la cantidad de hojas entregadas en la primer hoja.
- Con cédula de identidad y nombre en cada hoja.
- Escrita a lápiz y en forma prolija.

Parte 1. Modelo Entidad-Relación (25 puntos)

Ejercicio 1. (25 pts)

Una librería de una facultad desea implementar un sistema de ventas de apuntes por internet. De cada apunte se conoce un título, la fecha de edición y el costo en pesos. Además se conocen los autores del documento y una única asignatura de la cual proviene. Se asume que el título del apunte es único para cada asignatura. De cada autor, se conoce el nombre, su email (que es único), el instituto en donde trabaja y un resumen de su currículum.

De cada asignatura, se conoce un código que la identifica, un nombre, y un conjunto de temas que constituye el programa. De cada tema se conoce un nombre y un área temática. Hay que tener en cuenta que pueden haber temas con igual nombre en distintas áreas temáticas (por ejemplo el término inglés Bridge puede referirse a computación o ingeniería civil y son cosas diferentes). Los temas constituyen una jerarquía en la que hay temas subordinados a temas principales, en donde cada tema subordinado puede tener un único tema principal. Además cada tema puede ser visto en diferentes asignaturas.

Un apunte también tiene asociado un conjunto de temas de los cuales, alguno tiene que ser un tema de la asignatura asociada al apunte. Obviamente tanto los apuntes como las asignaturas tratan de por lo menos un tema.

Cada cliente debe registrarse antes de comprar y se le solicitan los siguientes datos:

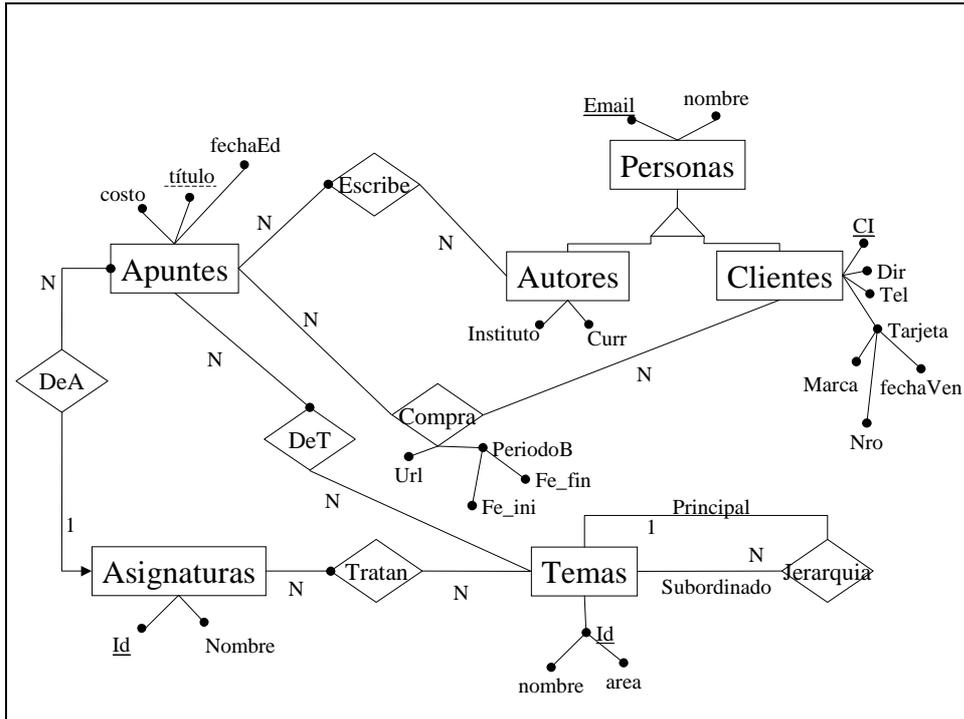
- nombre
- dirección
- email (que es único)
- Cedula de identidad
- Identificación de tarjeta de crédito formada por marca, número y fecha de vencimiento.
- teléfono

Para cada compra que se realiza se genera una url que depende del cliente y del apunte. Interesa registrar esa url y un "período de bajada" (fecha de inicio y fecha de fin) en que ese cliente puede acceder a ese artículo por esa url.

Hay que considerar que algunos autores pueden ser clientes.

SE PIDE: Esquema Entidad Relación completo, incluyendo restricciones de integridad.

SOLUCION:



RNE:

$$\forall a \in \text{Apuntes} . \exists t \in \text{Temas} . (<a,t> \in \text{DeT} \rightarrow (\exists s \in \text{Asignaturas} . (<a,s> \in \text{DeA} \wedge <t,s> \in \text{Tratan})))$$

Parte 2. Diseño Relacional (25 puntos)

Ejercicio 2 (8 pts)

Considere el esquema $R(A,B,C,D,E,G)$.

- Construya una instancia de R con 4 tuplas que cumpla con las restricciones de integridad del siguiente conjunto F : $F = \{AB \rightarrow CD, E \rightarrow D\}$
- Construya una instancia de R con 4 tuplas, con 4 valores distintos en el atributo E y que cumpla con las restricciones de integridad del siguiente conjunto F : $F = \{AB \rightarrow C|D, E \rightarrow D\}$

Sugerencia: Para los valores utilice el nombre del atributo en minúscula con un subíndice, asumiendo que subíndices distintos representan valores distintos para el mismo atributo.

SOL:

Solución:

a)

A	B	C	D	E
a ₁	b ₁	c ₁	d ₁	e ₁
a ₁	b ₁	c ₂	d ₁	e ₂
a ₁	b ₁	c ₁	d ₁	e ₂
a ₁	b ₁	c ₂	d ₁	e ₁

b)

A	B	C	D	E
a ₁	b ₁	c ₁	d ₁	e ₁
a ₁	b ₁	c ₂	d ₂	e ₂
a ₁	b ₁	c ₁	d ₂	e ₃
a ₁	b ₁	c ₂	d ₁	e ₄

Ejercicio 3 (9 pts)

Un servicio de salud quiere implementar su base de datos corporativa.

El servicio no posee hospitales propios sino que interna a sus pacientes en otros hospitales. De cada hospital se conoce un identificador, su nombre, su dirección, la ciudad en donde está y sus teléfonos. Cada hospital tiene un conjunto de salas de las que se conoce el nombre de la sala (CTI, Cuidados Intermedios, etc.) y la cantidad de camas. Tenga en cuenta que el nombre de la sala lo asigna cada hospital en forma independiente, y cuidando de no repetirlo en el hospital.

De cada afiliado se conoce su cédula de identidad, su nombre, dirección y teléfono. Además a cada afiliado se le da una tarjeta que identifica al grupo familiar por el cual se afilió. De esas tarjetas se conoce el número de la tarjeta y el plan de pago correspondiente, que es único para cada tarjeta. Todos los integrantes del grupo familiar, tienen la misma tarjeta, y cada afiliado posee una única tarjeta.

El servicio además tiene un conjunto de médicos que pueden ser afiliados al servicio o no. De los médicos se conoce el nombre, cedula de Identidad, teléfono de contacto, dirección del consultorio y la especialidad. Los médicos afiliados tienen un porcentaje de descuento que varía con cada médico.

Dado el siguiente esquema relacional para esta realidad:

HOSPITALES (id_hosp, nom_hosp, dir_hosp, ciu_hosp, tel_hosp)

SALAS (id_hosp, nom_sala, cant_camras)

AFILIADOS_NO_MEDICOS (ci_afil, nom_afil, dir_afil, tel_afil, num_tarj, plan_tarj)

MEDICOS_AFILIADOS (ci_med, porc_desc, num_tarj, plan_tarj)

MEDICOS (ci_med, nom_med, tel_med, dir_med, esp_med)

- Para cada relación dar: dependencias funcionales, dependencias multivaluadas, dependencias de inclusión y claves.
- Para cada relación decir en que forma normal se encuentra.

Solución

a)

HOSPITALES

Dfs: id_hosp -> nom_hosp, dir_hosp, ciu_hosp

Dmvs: id_hosp ->> tel_hosp

Dep. Inclusion: ninguna

Claves: (id_hosp, tel_hosp)

SALAS

Dfs: id_hosp, nom_sala -> cant_camras

Dmvs: ninguna

Dep. Inclusion: Salas.id_hosp \subseteq Hospitales.id_hosp
Claves: (id_hosp, nom_sala)

AFILIADOS_NO_MEDICOS

Dfs: ci_afil -> nom_afil, dir_afil, tel_afil, num_tarj
num_tarj -> plan_tarj
Dmvs: ninguna
Dep. Inclusion: ninguna
Claves: ci_afil

MEDICOS_AFILIADOS

Dfs: ci_med -> porc_desc, num_tarj
num_tarj -> plan_tarj
Dmvs: ninguna
Dep. Inclusion: Medicos_Afiliados.ci_med \subseteq Medicos.ci_med
Claves: ci_med

MEDICOS

Dfs: ci_med -> nom_med, tel_med, dir_med, esp_med
Dmvs: ninguna
Dep. Inclusion: ninguna
Claves: ci_med

b)

HOSPITALES – 1NF (dependencia parcial: id_hosp -> nom_hosp, dir_hosp, ciu_hosp)

SALAS – 4NF

AFILIADOS_NO_MEDICOS – 2NF (dependencia transitiva: num_tarj -> plan_tarj)

MEDICOS_AFILIADOS – 2NF (dependencia transitiva: num_tarj -> plan_tarj)

MEDICOS – 4NF

Ejercicio 4 (8 pts)

Dado un esquema de relación R.

- Qué significa que ρ tiene Join sin Pérdida?
- Dados el esquema R(A,B,C,D,E), la descomposición $\rho = \{ R1(ABC), R2(ACE), R3(CD) \}$ y las dependencias $F = \{ A \rightarrow B, C \rightarrow ADE, E \rightarrow A \}$, ρ tiene join sin pérdida?
- Indique en qué forma normal está ρ .
- Lleve R directamente a 3NF con respecto a F, con join sin pérdida y sin pérdida de dependencias. En cada momento indique que paso del algoritmo está aplicando.

Solucion ejercicio 4

- a) la propiedad de JSP significa que al aplicar la operación de JOIN sobre la descomposición no existe pérdida de la calidad de la información, o sea no se generan tuplas espúreas.

b)

$R(A,B,C,D,E)$

$F = \{A \rightarrow B, C \rightarrow ED, E \rightarrow A\}$

$\rho = \{R1(ABC), R2(ACE), R3(CD)\}$

Test de JSP

	A	B	C	D	E
R1	a1	a2	a3	b14	b15
R2	a1	b22	a3	b24	a5
R3	b31	b32	a3	a4	b35

Por $A \rightarrow B$

	A	B	C	D	E
R1	a1	a2	a3	b14	b15
R2	a1	a2	a3	b24	a5
R3	b31	b32	a3	a4	b35

Por $C \rightarrow ED$

	A	B	C	D	E
R1	a1	a2	a3	a4	a5
R2	a1	a2	a3	a4	a5
R3	b31	b32	a3	a4	a5

La descomposición ρ tiene JSP.

c) Forma normal de ρ

Proyecto dependencias sobre cada uno de los subesquemas, busco claves en cada uno de ellos y a partir de esto determino la forma normal de la descomposición ρ ,

$R1(ABC)$ $IIR1 = \{A \rightarrow B, C \rightarrow A\}$ C clave única

La dependencia $A \rightarrow B$ viola 3NF, por lo tanto R1 está en 2NF

$R2(ACE)$ $IIR2 = \{C \rightarrow AE, E \rightarrow A\}$ C clave única

La dependencia $A \rightarrow B$ viola 3NF, por lo tanto R2 está en 2NF

$R3(CD)$ $IIR3 = \{C \rightarrow D\}$ C clave única R3 está en 4NF

La descomposición está en 2NF

d)

Para llevar a 3NF con JSP y sin pérdida de dependencias busco un cubrimiento minimal de F. Como F ya es minimal no es necesario.

Hallo todas las claves de F.

Como C no aparece a la dcha. en ninguna df está en toda clave. Pruebo si C basta para formar una

clave.

$C^+ = \{C, E, D, A, B\}$ entonces C es clave, y además única porque cualquier otra clave debe contener a C y por lo tanto sería superclave.

Armo la descomposición:

$\rho_1 = \{R_1(AB), R_2(CED), R_3(EA)\}$ está en 3NF y conserva dfs.

Parte 3. Consultas (25 puntos)

Ejercicio 5 (25 pts)

La empresa *ArtePlus* se encarga de traer al país espectáculos artísticos (teatro, ballet, recitales, desfiles, etc.) y organizar su exhibición. Para ello alquila salas adecuadas para exhibir el espectáculo, en diferentes regiones del país. La empresa maneja la información relativa a los espectáculos, salas, funciones y reservas en una base de datos con las siguientes tablas:

ESPECTACULOS (codE, nombreE, tipo, procedencia, director, critica)

En esta tabla se almacena la información relativa a los espectáculos. De cada uno se conoce su código que lo identifica (codE), su nombre (nombreE), el tipo de espectáculo (tipo), país de procedencia (procedencia), director (director) y una puntuación asignada por la crítica (critica). Los tipos de espectáculos pueden ser teatro, ballet, recitales, etc. Esos son todos los tipos de espectáculos existentes.

SALAS (codS, nombreS, direccion, telefono, ciudad, capacidad, categoría)

En esta tabla se almacena la información relativa a las salas. De cada una se conoce un código que la identifica (codS), su nombre (nombreS), su dirección (direccion), un teléfono de contacto (telefono), su ciudad (ciudad), la cantidad máxima de personas que puede albergar (capacidad) y una puntuación del 1 al 5 que define su categoría (categoría).

FUNCIONES (codE, codS, fecha, horaCom, horaFin)

En esta tabla se maneja la información relativa a la exhibición de los espectáculos en las salas (funciones). De cada función se registra el espectáculo (codE), la sala (codS), la fecha (fecha), la hora de comienzo (horaCom) y la hora de finalización (horaFin). Cada sala exhibe a lo sumo un espectáculo por día. Un espectáculo puede exhibirse en varias salas y en varios días.

RESERVAS (codS, fecha, sector, fila, numero, estado)

En esta tabla se mantiene la información relativa a las reservas de asientos para asistir a un espectáculo. Una sala está dividida en sectores, por ejemplo platea alta, platea baja, palcos, etc., y los sectores se dividen en filas y asientos numerados.

De cada reserva se registra la sala (codS), la fecha (fecha), el sector (sector), fila (fila), número de asiento (numero) y el estado del asiento (estado). El estado puede ser vendido o reservado.

PRECIOS (codE, codS, sector, precio)

En esta tabla se mantiene la información relativa a los precios de los sectores de asientos para asistir a los espectáculos. Se registra el espectáculo (codE), la sala (codS), el sector (sector) y el precio. Un espectáculo puede tener precios distintos en distintas salas y sectores. Un sector de una sala puede valer distintos precios para distintos espectáculos.

NOTAS:

- No existen tablas vacías.
- $\Pi_{\text{codE}}(\text{FUNCIONES}) \subseteq \Pi_{\text{codE}}(\text{ESPECTACULOS})$
- $\Pi_{\text{codE}}(\text{PRECIOS}) \subseteq \Pi_{\text{codE}}(\text{ESPECTACULOS})$
- $\Pi_{\text{codS}}(\text{SERVICIOS}) \subseteq \Pi_{\text{codS}}(\text{SALAS})$
- $\Pi_{\text{codS}}(\text{FUNCIONES}) \subseteq \Pi_{\text{codS}}(\text{SALAS})$
- $\Pi_{\text{codS}}(\text{RESERVAS}) \subseteq \Pi_{\text{codS}}(\text{SALAS})$

- $\Pi_{\text{codS}}(\text{PRECIOS}) \subseteq \Pi_{\text{codS}}(\text{SALAS})$

SE PIDE:

Resolver las siguientes consultas en Álgebra Relacional:

- Obtener el código de sala y la fecha de las funciones del espectáculo “Una noche en Viena” que no tienen asientos reservados.
- Obtener los nombres de las salas donde se han exhibido espectáculos de todos los tipos.

Resolver las siguientes consultas en Cálculo Relacional:

- Obtener los nombres de las salas ubicadas en ciudades, tales que en esa ciudad se ha exhibido algún espectáculo dirigido por Federico García Vigil.
- Obtener los nombres de los espectáculos que vendieron las entradas más caras.

Resolver las siguientes consultas en SQL sin utilizar vistas

- Obtener las fechas en las que se vendieron todas las entradas en la sala “Zitarrosa” (se llenó toda la capacidad de la sala).
- Obtener el código y nombre de los espectáculos argentinos que se exhibieron en salas de Montevideo y la categoría promedio de dichas salas.

Solución:

Álgebra Relacional:

- Obtener el código de sala y la fecha de las funciones del espectáculo “Una noche en Viena” que no tienen asientos reservados.

$$A = \Pi_{\text{codS, fecha}}(\text{FUNCIONES}) - \Pi_{\text{codS, fecha}}(\sigma_{\text{estado}=\text{"reservado"}}(\text{RESERVAS}))$$

$$\text{Sol} = \Pi_{\text{codS, fecha}}(\sigma_{\text{nombreE}=\text{"Una noche en Viena"}}(A * \text{FUNCIONES} * \text{ESPECTACULOS}))$$

- Obtener los nombres de las salas donde se han exhibido espectáculos de todos los tipos.

$$A = \Pi_{\text{tipo}}(\text{ESPECTACULOS})$$

$$B = \Pi_{\text{codS, tipo}}(\text{FUNCIONES} * \text{ESPECTACULOS})$$

$$C = B \% A$$

$$\text{Sol} = \Pi_{\text{nombreS}}(C * \text{SALAS})$$

Obs: La división deben hacerla por codS y no por nombreS

Cálculo Relacional:

- Obtener los nombres de las salas ubicadas en ciudades, tales que en esa ciudad se ha exhibido algún espectáculo dirigido por Federico García Vigil.

$$\{ t.\text{nombreS} / \text{SALAS}(t) \wedge$$

$$(\exists s) (\text{SALAS}(s) \wedge s.\text{ciudad} = t.\text{ciudad} \wedge$$

$$(\exists f) (\text{FUNCIONES}(f) \wedge f.\text{codS} = s.\text{codS} \wedge$$

$$(\exists e) (\text{ESPECTACULOS}(e) \wedge e.\text{codE} = f.\text{codE} \wedge$$

$$e.\text{director} = \text{"Federico García Vigil"}$$

$$)$$

$$)$$

$$)$$

}

- d. Obtener los nombres de los espectáculos que vendieron las entradas más caras.

```
{ e.nombreE / ESPECTACULOS (e) ^
  (∃p) (PRECIOS(p) ^ p.codE = e.codE ^
    ¬ (∃q) (PRECIOS(q) ^ q.precio > p.precio)
  )
}
```

SQL

- e. Obtener las fechas en las que se vendieron todas las entradas en la sala "Zitarrosa" (se llenó toda la capacidad de la sala).

```
SELECT F.fecha
FROM SALAS S, FUNCIONES F
WHERE S.codS = F.codS
  AND S.nombreS = "Zitarrosa"
  AND S.capacidad = (
    SELECT count(*)
    FROM RESERVAS R
    WHERE R.codS = S.codS
      AND R.fecha = F.fecha
      AND R.estado = "vendido"
  )
```

- f. Obtener el código y nombre de los espectáculos argentinos que se exhibieron en salas de Montevideo y la categoría promedio de dichas salas.

```
SELECT E.codE, E.nombreE, avg(S.categoria)
FROM ESPECTACULOS E, SALAS S, FUNCIONES F
WHERE E.codE = F.codE
  AND S.codS = F.codS
  AND E.procedencia = "Argentina"
  AND S.ciudad = "Montevideo"
GROUP BY E.codE, E.nombreE
```

Parte 4 Optimización y Concurrencia (25 puntos)

Ejercicio 6 (12 pts).

Dado el siguiente esquema:

```
EMPLEADO ( Ci, Nombre, Dirección , Sexo, Salario, NumDepto )
DEPARTAMENTO ( NumDepto, NomDepto, Ci_Jefe )
ΠNumDepto (EMPLEADO) ⊆ ΠNumDepto (DEPARTAMENTO)
```

Comparar el costo de 2 planes físicos (correspondientes a 2 planes lógicos diferentes, no necesariamente optimizados) para la siguiente consulta:

```
SELECT *
FROM EMPLEADO E, DEPARTAMENTO D
WHERE E.NumDepto = D.NumDepto AND E.Salario = 40000
```

Especificar plan lógico y plan físico en cada caso. No olvidar considerar los costos de grabar para los

resultados intermedios. Considere que se tienen 3 buffers disponibles.

Datos:

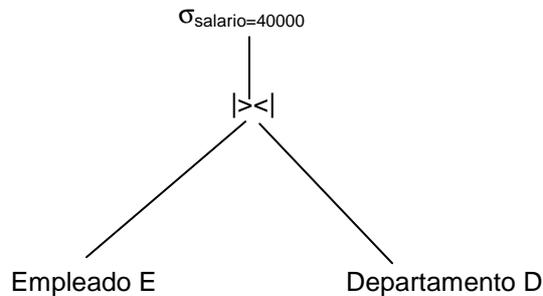
Tabla	Columna	Valores distintos
Departamento	NumDepto	50
Empleado	NumDepto	50
Empleado	Salario	500

Tabla	Cant. tuplas	Factor de bloqueo
Departamento	50	10
Empleado	10000	5
Departamento \bowtie Empleado		3

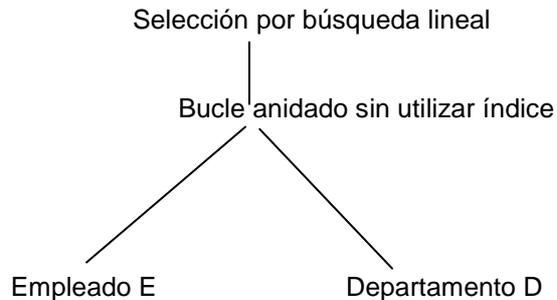
Indice	Tabla/Atributo	Tipo	Cant. niveles
Emp_NumDepto	Empleado/NumDepto	Primario	1
Emp_Sal	Empleado/Salario	Secundario	1

Solución:

Plan lógico 1:



Plan físico 1:



Costo:

$$C(\text{join}) = b_E + b_E * b_D = 2000 + 2000 * 5 = 12000$$

$$C(\text{grabar}) = b_{\text{join}} = |\text{join}| / \text{factor de bloqueo} = 10000 / 3 = 3334$$

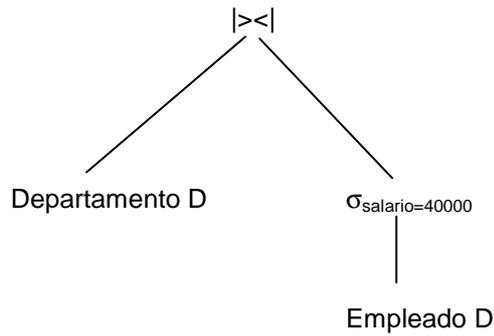
$$C(\text{selección}) = b_{\text{join}} = 3334$$

Aclaración:

$|\text{join}| = |E| = 10000$, porque NumDepto es clave en Departamento y clave foránea en Empleado

Ctotal = 18668

Plan lógico 2:



Plan físico 2:



Costo:

$$C(\text{seleccion}) = x + |\text{selección}| = 1 + 20 = 21$$

$$C(\text{grabar}) = b_{\text{seleccion}} = |\text{seleccion}| / \text{factor de bloqueo} = 20 / 5 = 4$$

$$C(\text{join}) = b_D + b_D * b_{\text{seleccion}} = 5 + 5 * 4 = 25$$

Aclaración:

$$|\text{seleccion}| = |E| / V(\text{salario}, E) = 10000 / 500 = 20$$

Ctotal = 50

Vemos que el plan 2 es mucho menos costoso que el plan 1.

Ejercicio 7. (13 pts.)

Decir si son verdaderas o falsas las siguientes afirmaciones, justificando cuando son verdaderas y dando un contraejemplo cuando son falsas.

- Toda historia que Evita Abortos en Cascada es Serializable.
- Toda historia Estricta es Recuperable.
- Toda historia cuyas transacciones siguen el protocolo 2PL-Riguroso es Estricta.
- Toda historia que no sigue el protocolo 2PL-Básico no es serializable.

Solución:

- Falso

Contraejemplo:

H: r1(X) r2(X) w1(X) w1(Y) c1 r2(Y) w2(Y) c2

H cumple EAC y no serializable.

b) Verdadero

Def. de Recuperable:

Una historia es recuperable cuando se cumple que si T1 lee de T2, T2 hace commit antes que T1.

En las historias estrictas todas las lecturas y escrituras se hacen en valores que fueron confirmados ("commiteados"). Por lo tanto si una historia es estricta cumple la definición de recuperable, ya que cuando se lee de una transacción esta ya hizo commit.

c) Verdadero.

Porque en cada transacción los *unlock*, tanto de los *read-lock* como de los *write-lock*, se hacen después del *commit*. Esto garantiza que una transacción solo va a poder trabajar sobre un ítem escrito por otra después que esta otra haya hecho commit.

d) Falso

Contraejemplo:

H: r1(X) r1(X) r1(X) r2(X) u2(X) w1(X) w1(X) u(X) w2(Y) w2(Y) u2(Y) r2(Y) r1(Y) u1(Y) c1 c2

Sacando los bloqueos (para que se vea mas claramente):

r1(X) r2(X) w1(X) w2(Y) r1(Y) c1 c2

H es serializable y no cumple 2PL-Básico.