



Identifique todas as folhas. Exame sem consulta. Duração: 2h30.

Grupo 1 – Número: _____ **Nome:** _____

- 1a)** Apresente o diagrama Entidade-Associação (E-A) para um novo jogo multiutilizador inspirado no popular “Tetris” e caracterizado pelo seguinte domínio: (1,5v)

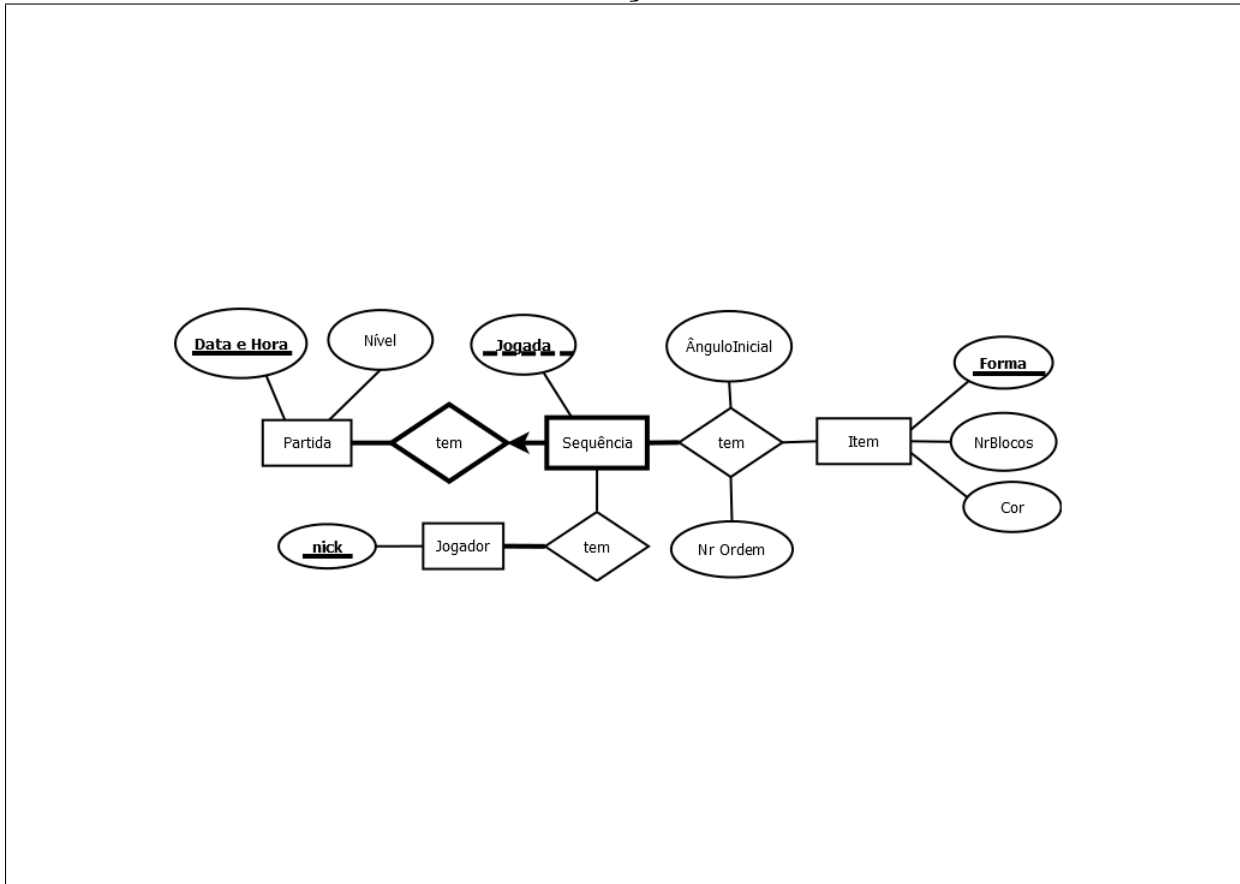
Ao longo de uma partida cada jogador, identificado por um nick, vai ter de empilhar uma sequência de itens. Os jogadores jogam em alternância vencendo aquele que empilhar mais itens. Os itens têm figuras associadas. Cada figura é identificada por uma letra que representa a sua forma (como “T”, “I”, “L” e “Z”), e tem uma cor (“verde”, vermelho, “amarelo” e “violeta”) e o número de blocos que a compõem (3, 4 ou 5).

Um item corresponde a uma figura que o jogador pode rodar com o objetivo de maximizar o nº de figuras empilhadas. Uma partida envolve o empilhamento da sua sequência de itens. Cada partida é ainda caracterizada pela data-hora de início e pode ser jogada por um ou vários jogadores.

Com a informação registada na base de dados, um jogador poderá consultar todo o seu histórico, que consiste na relação das partidas jogadas, o nível de dificuldade com que foi jogada cada partida, e os itens da sequência que foram empilhados por cada jogador nessa partida.

Cada item da sequência é caracterizado por um número de ordem, pela figura e pelo o ângulo inicial de apresentação, que pode ser “180”, “-90” ou “90” graus (por exemplo, a figura “I” deitada terá um ângulo de 90°). As sequências são geradas no início de cada partida. A dificuldade da partida pode variar entre de 1 e 9, determinando a frequência com que são apresentados os itens.

Solução:

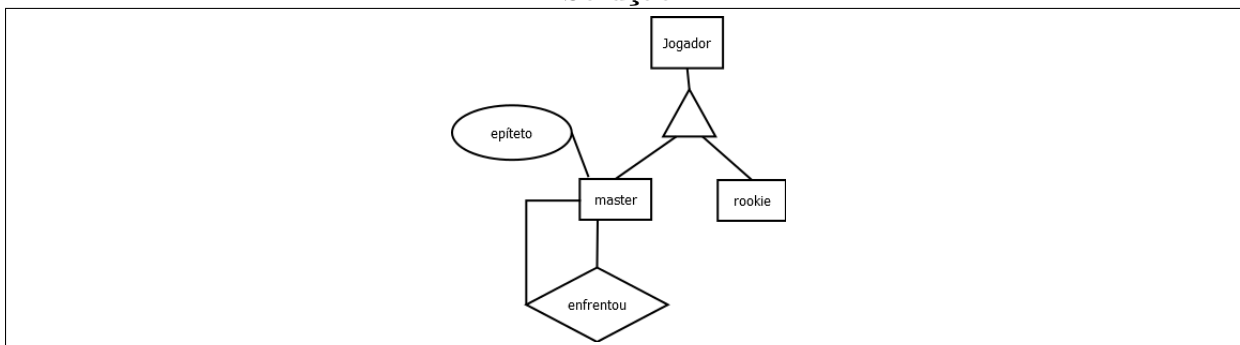


1b) Ao domínio da alínea anterior acrescente as seguintes afirmações:

Os jogadores são classificados em duas categorias, “rookie” e “master” consoante os resultados obtidos, sendo o escalão da classificação determinado pela média das 3 melhores pontuações obtidas em todas as partidas jogadas. Além disso, os jogadores “master” podem juntar um epíteto ao nick e mantêm-se para estes uma relação dos outros “masters” que já enfrentaram.

Represente o modelo correspondente E-A. Não é necessário repetir os elementos que se mantêm iguais à alínea anterior, mas deve indicar as alterações se as houver. (1v)

Solução:



- 1c) Converta o diagrama resultante para modelo relacional. Use a seguinte notação para indicar as chaves primárias (sublinhado) e estrangeiras (FK): (1,5v)

$relacao_A(\underline{atributo_1}, atributo_2, \dots)$ $atributo_2 : FK(relacao_B)$

Solução:

$partida(\underline{datahora}, nivel)$
 $sequencia(\underline{jogada}, \underline{datahora})$
 $datahora : FK(Partida)$
 $item(\underline{forma}, nrblocos, cor)$
 $sequenciaTemItem(\underline{jogada}, \underline{datahora}, \underline{forma}, AnguloInicial, NrdeOrdem)$
 $jogada, datahora : FK(sequencia)$
 $forma : FK(item)$
 $jogador(\underline{nick})$
 $jogadorTemSequencia(\underline{nick}, jogada, datahora)$
 $jogada, datahora : FK(sequencia)$
 $nick : FK(jogador)$
 $master(\underline{nick}, epiteto)$
 $nick : FK(jogador)$
 $enfrentou(\underline{nick1}, \underline{nick2})$
 $nick1 : FK(master)$
 $nick2 : FK(master)$

Grupo 2 – Número: _____ Nome: _____

Considere uma base de dados em que se guardam dados sobre um jogo multiutilizador.

<u>datahora</u>	<u>ordem</u>	nblocos
... 12:31:00	123	1
... 12:31:00	124	2
...

<u>nick</u>	nome	categoria
zé	José Ribeiro	master
quim	Joaquim Ribeiro	rookie
manel	Manuel Ribeiro	rookie

<u>datahora</u>	<u>nick</u>	nível
... 12:31:01	zé	7
... 12:31:01	quim	2
... 17:25:00	quim	2

Apresente uma resolução em álgebra relacional para cada uma das seguintes questões:

2a) Quais os nicks e nomes completos dos jogadores que participaram em partidas nos últimos 10 dias? (1v)

Solução:

$$\pi_{nick, nome} \sigma_{now() - datahora > 10} (jogador \bowtie partida)$$

2b) Qual o jogador que jogou mais blocos em todas as partidas? (1,5v)

Solução:

$$\begin{aligned} & \rho(r_1(nick, total),_{nick} \mathcal{G}_{sum(nblocos)}(partida \bowtie sequencia)) \\ & \rho(r_2(total), \mathcal{G}_{max(nblocos)}(r_1)) \\ & \pi_{nick}(r_1 \bowtie r_2) \end{aligned}$$

2c) Quais os jogadores que jogaram em todas as partidas? (1,5v)

Solução:

$$\pi_{nick}(partida / (\pi_{datahora}(partida)))$$

Grupo 3 – Número: _____ **Nome:** _____

Considerando a mesma base de dados do exercício anterior, escreva uma consulta em SQL para responder a cada uma das seguintes questões:

3a) Qual o nome dos jogadores que nunca jogaram nenhuma partida? (1,5v)

Solução:

```
select nome
from jogador
where nick not in (select nick from partida)
```

3b) Para cada nível, quais os nicks dos jogadores que obtiveram maior pontuação (numa partida desse nível)? Considere que a pontuação é obtida multiplicado o número de blocos pelo nível de dificuldade. (1,5v)

Solução:

```
select nick, datahora, nivel, sum(nivel*nblocos)
from (partida natural join sequencia) as t
group by nick, datahora, nivel
having sum(nivel*nblocos) >= all (
select sum(nivel*nblocos)
from partida natural join sequencia
where nivel = t.nivel
group by nick, datahora
)
```

3c) Quais os nicks dos jogadores que jogaram em todas as partidas de nível dificuldade superior a 5? (1v)

Solução:

```
select nick
from jogador j
where nick not exists ((select datahora
from partida
where nivel > 5)
except
(select datahora
from partida p
where p.nick = j.nick)
```

Grupo 4 – Número: _____ Nome: _____

4a) Considere os dois seguintes escalonamentos:

- i. T1:R(X), T2:R(X), T1:W(X), T2:W(X)
- ii. T1:R(X), T2:W(X), T1:W(X), T2:Abort, T1:Commit

Para cada um, indique se é serializável, *conflict-serializable* (serializável em conflitos ou com conflito serializável), recuperável, evita *rollbacks* em cadeia. Justifique. (1,5v)

Solução:

- | |
|---|
| <ul style="list-style-type: none">i. Serializável: não porque não é equivalente a nenhum escalonamento em série; com conflito serializável: não, porque o grafo de precedências tem ciclos; recuperável: sim porque não existem leituras de T1/T2 depois de escritas de T2/T1; evita rollbacks em cadeia: sim, pela mesma razão.
ii. Serializável: sim, porque é equivalente ao escalonamento em série T2, T1; com conflito serializável em conflitos: não porque o grafo de precedências continua a ter ciclos, ou de outro modo não conseguimos trocar o T1:W(X) com o T2:W(X); recuperável: sim pela mesma razão de i; evita rollbacks em cascata: sim, pela mesma razão. |
|---|

4b) Considere uma base de dados com objectos X e Y e assuma que existem duas transacções T1 e T2 a executar-se. A transacção T1 lê os objectos X e Y e depois escreve o objecto X. A transacção T2 lê os objectos X e Y e depois escreve os objectos X e Y.

Dê um exemplo de escalonamento com acções das transacções T1 e T2 sobre os objectos X e Y que resulte num conflito leitura-escrita. Indique se a utilização do protocolo *2PL* evitaria esse conflito. (1v)

Solução:

<p>T2:R(X), T2:R(Y), T2:W(X), T1:R(X)</p> <p>No protocolo Strict 2PL, T1 não poderia obter um lock partilhado sobre X porque T2 possuiria um lock exclusivo sobre X. Assim, T1 teria que esperar que T2 acabasse. Se fosse apenas 2PL, poderia não garantir, porque os locks de uma transacção podem libertar-se enquanto a outra se executa.</p>
--

4c) Considere a execução de transacções correspondente ao seguinte ficheiro de *log* (em que cada registo está representado na sua forma compacta) e considere um sistema de recuperação que obedece ao algoritmo ARIES:

LSN	LOG	prevLSN
00	update: T1 writes P2	
10	update: T1 writes P1	
20	update: T2 writes P5	
30	update: T3 writes P3	
40	T3 commit	
50	update: T2 writes P5	
60	update: T2 writes P3	
70	T2 abort	

- i. Preencha a figura acima de modo a mostrar os valores dos campos prevLSN.
- ii. Descreva quais as acções que são executadas e indique quais os registos que são acrescentados ao ficheiro de *log* durante o processo de cancelamento de T2 (que tem lugar depois de T2 ter sido abortada), incluindo os valores dos campos prevLSN e undoNextLSN.

(1,5v)

Solução:

A figura estendida é:

LSN	prevLSN	undoNextLSN (de um CLR que corresponde a este registo de update)
00	-	-
10	00	00
20	-	-
30	-	-
40	30	- (não é um registo de update)
50	20	20
60	50	50
70	60	- (não é um registo de update)

As acções executadas durante o *rollback* de T2 são as seguintes:

- Repôr P3 com os valores antigos tal como guardados no LSN 60.
- Repôr P5 com os valores antigos tal como guardado no LSN 50.
- Repôr P5 com os valores antigos tal como guardado no LSN 20.

O resto do ficheiro de log é como se segue:

LSN	prevLSN	transID	type	pageID	undonextLSN
80	70	T2	CLR	P3	50
90	80	T2	CLR	P5	20
100	90	T2	CLR	P5	-
110	100	T2	END	-	-

Grupo 5 – Número: _____ Nome: _____

5a) Preencha a tabela abaixo com as dependências funcionais observadas nas 3 relações do Grupo 2 (considerando o domínio descrito no Grupo 1) e com a forma normal em que se encontra cada uma dessas relações.

Relação	Dependências Funcionais	Forma Normal
<i>sequencia</i>		
<i>jogador</i>		
<i>partida</i>		

(1v)

Solução:

Relação	Dependências Funcionais	Forma Normal
<i>sequencia</i>	$datahora, ordem \rightarrow nblocos$	BCNF
<i>jogador</i>	$nick \rightarrow nome, categoria$	BCNF
<i>partida</i>	$datahora, nick \rightarrow nivel$	BCNF

Houve porém várias outras interpretações também consideradas, nomeadamente:

- considerar não haver chaves definidas à partida, identificar todas as DF a partir dos dados, assumindo existência de dependência sempre que os dados não permitem concluir a não existência de dependência funcional, e a partir daí as chaves e a FN.
- aceitar as chaves, mas considerar que *nick* não faz parte da chave de *partida*
- haver mais DF inferidas, como $nick \rightarrow nivel$ em *partida*, da qual resultaria que a relação poderia estar na 1NF, 2NF ou 3NF consoante *nick* faz ou não parte da chave.

5b) Para a interrogação da questão 3a) *Qual o nome dos jogadores que nunca jogaram nenhuma partida?*, indique quais os índices necessários para que o resultado pudesse ser obtido exclusivamente por acesso a esses índices (sem acesso ao ficheiro de dados). Diga também, justificando, se, obtendo-se o resultado por acesso exclusivo a esses índices, eles poderiam ser *esparsos*. (1v)

Solução:

Um índice composto em jogador(nick, nome) e outro em partida(nick). Não poderiam ser esparsos desde logo porque a computação obrigaria a encontrar todos os valores de *nick* nas duas tabelas, logo cada uma das entradas da chave de pesquisa de cada índice teria de corresponder uma entrada na tabela correspondente. Caso contrário, seria necessário aceder aos ficheiros com as tabelas.

- 5c) Em processamento analítico de informação é frequente recorrer ao modelo multidimensional para representação dos dados. Explique a sua organização e indique porque razão é preferido à representação com o habitual modelo relacional. (1v)

Solução:

O MDD permite visualizar intuitivamente e computar rapidamente medidas que resultam de agregações sucessivas de dados em intervalos variáveis em várias dimensões. Estas agregações podem ser précomputadas o que possibilita a exploração interactiva de dados analíticos. Essa précomputação pode estar armazenada num array multi-dimensional ou num esquema relacional

- 5d) Diga quais são as diferenças entre um sistema de recuperação de informação e um sistema de Gestão de Bases de Dados, no que respeita à formulação das interrogações e resultados apresentados. (1v)

Solução:

- Linguagem de interrogação de dados baseada em lógica de 1ª ordem (como SQL) vs. lista de termos;
- Pesquisa exacta onde a ordem dos resultados apresentados que satisfazem a interrogação é irrelevante vs. probabilística por ordem de relevância
- Apresentação dos dados que satisfazem a interrogação vs. documentos cujo conteúdo satisfaz necessidade de informação expressa na interrogação.

Antes de entregar, separe e verifique que identificou todas as folhas.