Universidade de São Paulo Instituto de Ciências Matemáticas e de Computação Bacharelado em Ciências de Computação SCC0215 - Organização de Arquivos – Turma B

Twitter

Frederico de Azevedo Marques	8936926
Henrique Martins Loschiavo	8936972
Lucas Kassouf Crocomo	8937420
Roberto Pommella Alegro	8936756

Sumário

Seção 1. Estrutura interna de dados	2
Seção 2. Estrutura de funções internas	3
Funções implementadas	3
insertTweet	3
requestAllTweets	4
requestTweet	5
findTweetByUser	6
removeTweet	7
printTweet	9
Seção 3. Interface com o usuário	10
Tela Inicial	10
Tela de Inserção	10
Tela de Requisição	11
Tela de Remoção	11
Tela de Ajuda	12
Tela de Créditos	12
Seção 4. Bateria de testes	13
Executando teste	13
Bateria de testes	13

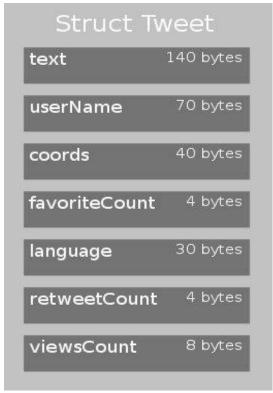
Seção 1: Estrutura de dados interna

A estrutura definida para um tweet foi de:

- Uma String de no máximo de 140 caracteres para o texto do tweet.
- Uma String de no máximo 70 caracteres para o nome do usuário.
- Uma String de no máximo 40 caracteres para as coordenadas geográficas do tweet.
- Uma String de no máximo 30 caracteres para o idioma do tweet
- Um inteiro longo para a contagem de visualizações do tweet
- Um inteiro para a contagem de favoritos do tweet
- Um inteiro para a contagem de retweets.

O valor do texto foi escolhido como 140, pois o tamanho máximo no twitter é de 140 caracteres.

Os outros valores foram escolhidos de forma arbitraria, levando em consideração apenas ter o tamanho suficiente para acomodar as informações de cada campo.



Estrutura dos registros

Seção 2: Estrutura de funções internas

Funções Implementadas

insertTweet

A função *insertTweet* insere um tweet no arquivo.

A função recebe por parâmetro o nome do arquivo no qual deve ser adicionado o tweet, o texto do tweet, o usuário, as coordenadas geográficas, o idioma, o número de visualizações, de favoritos e de retweets.

Para a inserção, é levado em consideração que um registro que foi removido logicamente é um espaço válido para salvar o novo registro.

O registro é inserido no arquivo usando a função *fwrite* da biblioteca *stdio.h*, padrão do sistema.

Ao inserir em um arquivo em branco temos:

```
--> xxd test.dat
00000000: 5465 7874 6f00 0000 0000 0000 0000 0000
                      Texto.....
00000080: 0000 0000 0000 0000 0000 4e6f 6d65
00000090: 2044 6520 5573 75c3 a172 696f 0000 0000
                       De Usu..rio....
000000d0: 0000 436f 6f72 6465 6e61 6461 7320 4765
                       ..Coordenadas Ge
000000e0: 6f67 72c3 a166 6963 6173 0000 0000 0000
                       ogr..ficas.....
000000f0: 0000 0000 0000 0000 0000 4964 696f 6d61
00000110: 0000 0000 0000 0000 0100 0000 0100 0000
00000120: 0100 0000 0000 0000
```

Podemos observar no arquivo que cada espaço de bytes corresponde a um campo na estrutura do tweet.

Byte offset	te offset Valor			
00 - 8B	Texto (Verde)	140		
8C - D1	Nome do Usuário (Azul)	70		
D2 - F9	Coordenadas (Vermelho)	40		
FA - 117	Idioma (Amarelo)	30		
118 -11B	Favoritos(Aqua)	4		
11C - 11F	Retweets(Aqua)	4		
120 - 127	Visualizações (Aqua)	8		

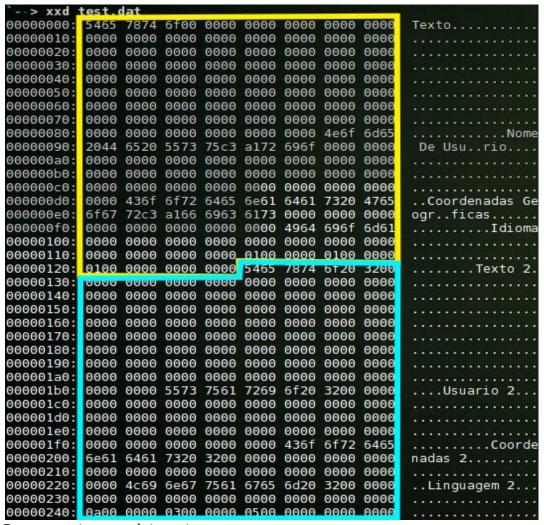
requestAllTweets

A função *requestAllTweets* consulta o arquivo e retorna todos os tweets que forem encontrados, se houver algum, e retorna também o número de tweets que foram encontrados.

A implementação percorre por todos os tweets do arquivo e verifica se ele existe ou foi removido logicamente, avaliando se os valores de viewCount, retweetCount e favoriteCount não são negativos.

A função recebe por parâmetro o nome do arquivo do qual deve ler os tweets e um ponteiro para onde colocar o número de tweets encontrados.

Ao buscar em um arquivo com 2 registros, temos:



Busca arquivo com dois registros

Note que cada tweet está selecionado. Cada seleção corresponde a um *fread* no arquivo.

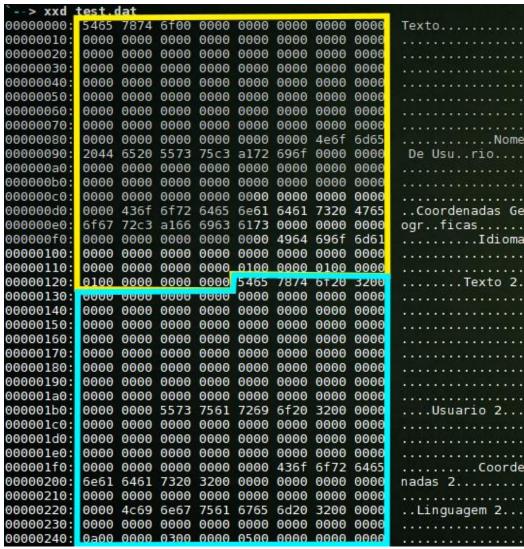
requestTweet

A função *requestTweet* consulta o arquivo e retorna o tweet correspondente ao *RRN* dado, se esse existir.

É levado em consideração que o registro com o *RRN* dado pode ter sido removido logicamente. Nesse caso a função considera que o *RRN* é inválido.

A função recebe por parâmetro o nome do arquivo do qual deve ler os tweets e o *RRN*, de onde deve pegar o registro.

Ao buscar em um arquivo com 2 registros, temos:



Busca em arquivo com dois registros

A *requestTweet* converte o *RRN* recebido em byte offset, através do produto do *RRN* por 0x128 (O tamanho de um tweet em bytes). Em seguida o ponteiro da stream é posicionado nesse offset, através do comando *fseek*, e a leitura do registro é feita por um *fread*.

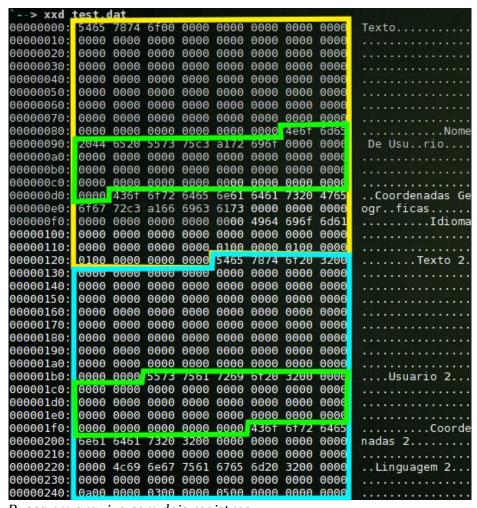
findTweetByUser

A função *findTweetByUser* consulta o arquivo e retorna todos os tweets de um determinado usuário, se houver algum, e retorna também o número de tweets encontrados.

A implementação percorre por todos os tweets do arquivo e verifica se ele existe ou foi removido logicamente, avaliando se os valores de viewCount, retweetCount e favoriteCount não são negativos.

A função recebe por parâmetro o nome do arquivo do qual deve ler os tweets e o nome do usuário para busca.

Ao buscar em um arquivo com 2 registros, temos:



Busca em arquivo com dois registros

Note que agora, além dos tweets, os bytes correspondentes ao campo userName foram destacados em verde. A função compara esses bytes com a string recebida por parâmetro usando a função *strcmp* oferecida pela biblioteca *string.h.*

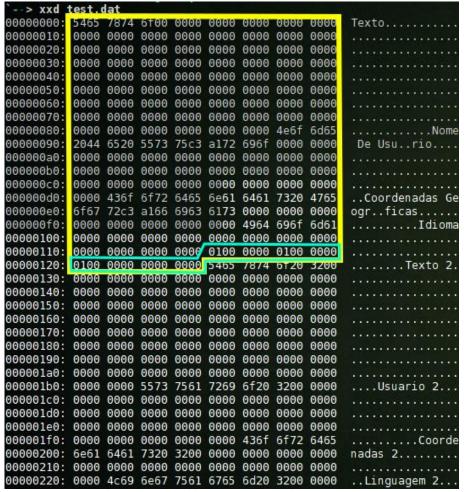
removeTweet

A função *removeTweet* remove um tweet logicamente, colocando os valores negativos no número de visualizações, de favoritos e de retweets.

A função recebe por parâmetro o nome do arquivo do qual deve remover o tweet e o *RRN* do tweet a remover.

Após realizar um *requestTweet* com o *RRN* recebido por parâmetro, os campos viewCount, retweetCount e favoriteCount são mudados para -1. Ao fim a função sobrescreve o registro com o mesmo *RRN*, atualizando assim os campos alterados.

Antes de remover o primeiro registro (RRN = 0), temos:



Pré- remoção

Note que os últimos bytes, destacados em azul, correspondem aos campos viewCount, retweetCount e favoriteCount.

Após a remoção, temos:

`> xxd	test.	dat							
00000000:	5465	7874	6100	0000	0000	0000	0000	0000	Texto
00000010:	0000	0000	0000	0000	0000	0000	0000	0000	
00000020:	0000	0000	0000	0000	0000	0000	0000	0000	to the second se
00000030:	0000	0000	0000	0000	0000	0000	0000	0000	
00000040:	0000	0000		0000		0000	0000	0000	
00000050:	0000	0000	0000	0000	0000	0000	0000	0000	
00000060:	0000	0000	0000	0000	0000	0000	0000	0000	
00000070:	0000	0000	0000	0000	0000	0000	0000	0000	
00000080:	0000	0000	0000	0000	0000	0000	4e6f	6d65	Nome
00000090:	2044	6520	5573	75c3	a172	696f	0000	0000	De Usurio
000000a0:	0000	0000	0000	0000	0000	0000	0000	9000	
000000b0:	0000	0000	0000	0000	0000	0000	0000	0006	
000000c0:	0000	0000	0000	0000	0000	0000	0000	9000	
000000d0:	0000	436f	6f72	6465	6e61	6461	7320	4765	Coordenadas Ge
000000e0:	6f67	72c3	a166	6963	6173	0000	0000	0000	ogrficas
000000f0:	0000	0000	0000	0000	0000	4964	696f	6d61	Idioma
00000100:	0000		0000						*****
00000110	0000	9999	9999	0000	ffff	ffff	ffff	fff	
00000110:									The state of the s
			ffff	ffff	5465	/8/4	0120	320U	Texto 2.
	ffff	ffff 0000	ffff 0000	ffff	5465 0000		A STATE OF THE STATE OF	The second second	The state of the s
00000120:	ffff	ffff 0000	ffff 0000	ffff	5465 0000	/8/4	0120	320U	Texto 2.
00000120: 00000130:	ffff 0000 0000 0000	ffff 0000 0000	ffff 0000	ffff 0000 0000	5465 0000	7874 0000 0000	6120 0000 0000 0000	3200 0000	Texto 2.
00000120: 00000130: 00000140:	ffff 0000 0000	ffff 0000 0000 0000	ffff 0000 0000	ffff 0000 0000 0000	5465 0000 0000	7874 0000 0000 0000	0000 0000	3200 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150:	ffff 0000 0000 0000	ffff 0000 0000 0000	ffff 0000 0000 0000	ffff 0000 0000 0000	9000 0000 0000 0000	7874 0000 0000 0000	6120 0000 0000 0000	3200 0000 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150: 00000160:	ffff 0000 0000 0000 0000	ffff 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000	9000 9000 9000 9000 9000	7874 0000 0000 0000 0000	0000 0000 0000 0000 0000	0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150: 00000160:	ffff 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000	9465 9000 9000 9000 9000 9000	7874 0000 0000 0000 0000 0000	0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150: 00000160: 00000170: 00000180: 00000190:	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	9463 9099 9099 9099 9099 9099 9099	7874 0000 0000 0000 0000 0000 0000 0000	8120 0000 0000 0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150: 00000160: 00000170: 00000180: 00000190: 000001b0:	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	0000 0000 0000 0000 0000 0000 0000 0000 7269	7874 0000 0000 0000 0000 0000 0000 0000	0120 0000 0000 0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150: 00000170: 00000180: 00000190: 000001b0: 000001c0:	6fff 0000 0000 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	0000 0000 0000 0000 0000 0000 0000 0000 7269 0000	7874 0000 0000 0000 0000 0000 0000 0000	6120 0000 0000 0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000140: 00000150: 00000170: 00000180: 00000190: 000001b0: 000001c0:	6fff 0000 0000 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000	0000 0000 0000 0000 0000 0000 0000 0000 7269 0000 0000	7874 0000 0000 0000 0000 0000 0000 0000	6120 0000 0000 0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000140: 00000150: 00000160: 00000170: 00000180: 00000100: 000001c0: 000001d0: 000001e0:	6fff 0000 0000 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 0000	0000 0000 0000 0000 0000 0000 0000 7269 0000 0000	7874 0000 0000 0000 0000 0000 0000 0000	8120 0000 0000 0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000140: 00000150: 00000160: 00000170: 00000180: 00000100: 000001c0: 000001d0: 000001f0:	6fff 0000 0000 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 0000	0000 0000 0000 0000 0000 0000 0000 7269 0000 0000 0000 0000	7874 0000 0000 0000 0000 0000 0000 0000	6120 0000 0000 0000 0000 0000 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000140: 00000150: 00000160: 00000170: 00000180: 00000100: 000001c0: 000001d0: 000001f0: 000001f0: 000001f0:	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 0000 0000 3200	5465 0000 0000 0000 0000 0000 0000 7269 0000 0000 0000 0000 0000	7874 0000 0000 0000 0000 0000 0000 6f20 0000 000	6120 0000 0000 0000 0000 0000 0000 3200 0000 0000 0000 6f72 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000150: 00000150: 00000170: 00000180: 00000100: 000001c0: 000001c0: 000001f0: 000001f0: 000001f0: 00000200: 00000210:	6fff 0000 0000 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 0000 0000 3200 0000	5465 0000 0000 0000 0000 0000 0000 7269 0000 0000 0000 0000 0000 0000	7874 0000 0000 0000 0000 0000 0000 6f20 0000 000	8120 0000 0000 0000 0000 0000 0000 3200 0000 0000 6f72 0000 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000150: 00000160: 00000170: 00000180: 00000100: 000001c0: 000001c0: 000001f0: 00000200: 00000220:	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 0000 3200 0000 7561	5465 0000 0000 0000 0000 0000 0000 7269 0000 0000 0000 0000 0000 0000 6765	7874 0000 0000 0000 0000 0000 0000 6f20 0000 436f 0000 0000 6d20	6120 0000 0000 0000 0000 0000 0000 3200 0000 6f72 0000 0000 3200	3200 0000 0000 0000 0000 0000 0000 0000	Usuario 2Coorde
00000120: 00000130: 00000150: 00000160: 00000170: 00000180: 00000100: 00000100: 00000100: 00000160: 00000160: 00000160: 00000160: 00000160: 00000160: 00000160: 00000100: 00000100:	6fff 0000 0000 0000 0000 0000 0000 0000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 7561 0000 7561	5465 0000 0000 0000 0000 0000 0000 7269 0000 0000 0000 0000 0000 0000 6765 0000	7874 0000 0000 0000 0000 0000 0000 6f20 0000 436f 0000 6d20 0000	8120 0000 0000 0000 0000 0000 0000 3200 0000 6f72 0000 0000 3200 0000 3200 0000	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.
00000120: 00000130: 00000150: 00000160: 00000170: 00000180: 00000100: 000001c0: 000001c0: 000001f0: 000001f0: 00000200: 00000220:	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 000	ffff 0000 0000 0000 0000 0000 0000 5573 0000 0000	ffff 0000 0000 0000 0000 0000 0000 7561 0000 0000 7561 0000 7561	5465 0000 0000 0000 0000 0000 0000 7269 0000 0000 0000 0000 0000 0000 6765 0000	7874 0000 0000 0000 0000 0000 0000 6f20 0000 436f 0000 6d20 0000	6120 0000 0000 0000 0000 0000 0000 3200 0000 6f72 0000 0000 3200	3200 0000 0000 0000 0000 0000 0000 0000	Texto 2.

Pós-remoção

Note que os bytes destacados foram atualizados para o valor -1 (0xFFFFFFFF), o que corresponde a remoção lógica.

printTweet

A função *printTweet* retorna uma representação de um tweet em uma string, no formato:

Text: <Texto> User: @<Usuário>

Coordinates: <Coordenadas>

Language: <Idioma>
Views: <Visualizações>
Favorites: <Favoritos>
Retweets: <Retweets>

RRN: <RRN>

A função recebe por parâmetro o nome de um arquivo no qual o tweet existe e um ponteiro para o mesmo tweet

Para calcular o *RRN* do tweet, a função percorre o arquivo de dados, buscando por um registro correspondente ao tweet recebido por parâmetro.

Seção 3. Interface com o usuário:

Tela Inicial

```
.-(~/documents/USP/Programação/C/Cristina/Tweet)
`--> ./main
please input the name of a file to work with. it will be created if it does not exist:
test.dat
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~>
```

Tela Inicial

Tela de Inserção

```
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> insert
Please type a user name:
Minerva
Type your geographic coordinates:
Sao Carlos - Sao Paulo
Specify the language of your text:
English
Type your text:
Hello, World!
Type how many favorites your tweet has:
999
Type how many views your tweet has:
1000
Type how many retweets your tweet has:
Tweet inserted successfully!
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~>
```

Tela de inserção

Tela de Requisição

```
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> request rrn 0
Text: Hello, World!
User: Minerva
Coordinates: Sao Carlos - Sao Paulo
Language: English
Views: 1000
Favorites: 999
Retweets: 2
RRN: 0
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~>
```

Tela de requisição

Tela de Remoção

```
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> remove 0
The Tweet 0 was successfully removed
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> request all
Could not find requested Tweet(s). Maybe it does not exist
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~>
```

Tela de remoção

Tela de Ajuda

Tela de Ajuda

Tela de Créditos

```
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> credits
This program was made by:

* Frederico de Azevedo Marques - N USP: 8936926

* Henrique Martins Loschiavo - N USP: 8936972

* Lucas Kassouf Crocomo - N USP: 8937420

* Roberto Pommella Alegro - N USP: 8936756

Using the specification available at:
http://wiki.icmc.usp.br/images/6/63/SCC0215012015projeto01TurmaBa.pdf
and the knowledge obtained from the classes of professor Cristina Dutra de Aguiar Ciferri

~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~>
```

Tela de Créditos

Seção 4. Bateria de testes:

Executando teste

Criar um novo arquivo de entrada (in), contendo a entrada da bateria de testes.

```
./main novoArquivo.dat < in > out
```

Verificar se o arquivo out corresponde com a saída esperada.

Bateria de testes

Entrada:

insert

User1

Coordenadas

Português

Texto

10

80

5

insert

User2

Coordenadas2

Portugês

Texto2

-1

80

23

insert

User3

Coordenadas3

Português

Texto3

4

35

2

request all

remove 0

request rrn 0

request rrn 1

insert

User1

Coordenadas4

Português

Texto4

80

291

34

```
request all
request user User1
Saída esperada:
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Please type a us
er name:
Type your geographic coordinates:
Specify the language of your text:
Type your text:
Type how many favorites your tweet has:
Type how many views your tweet has:
Type how many retweets your tweet has:
Tweet inserted successfully!
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Please type a us
er name:
Type your geographic coordinates:
Specify the language of your text:
Type your text:
Type how many favorites your tweet has:
Type how many views your tweet has:
Type how many retweets your tweet has:
Could not insert the tweet. Is the filesystem read only?
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Please type a us
er name:
Type your geographic coordinates:
Specify the language of your text:
Type your text:
Type how many favorites your tweet has:
Type how many views your tweet has:
Type how many retweets your tweet has:
Tweet inserted successfully!
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Text: Texto
User: @User1
Coordinates: Coordenadas
Language: Português
Views: 80
Favorites: 10
Retweets: 5
RRN: 0
Text: Texto3
User: @User3
Coordinates: Coordenadas3
Language: Português
Views: 35
Favorites: 4
Retweets: 2
```

```
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> The Tweet 0 was
successfully removed
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Could not find r
equested Tweet(s). Maybe it does not exist
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Text: Texto3
User: @User3
Coordinates: Coordenadas3
Language: Português
Views: 35
Favorites: 4
Retweets: 2
RRN: 1
-----
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Please type a us
er name:
Type your geographic coordinates:
Specify the language of your text:
Type your text:
Type how many favorites your tweet has:
Type how many views your tweet has:
Type how many retweets your tweet has:
Tweet inserted successfully!
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Text: Texto4
User: @User1
Coordinates: Coordenadas4
Language: Português
Views: 291
Favorites: 80
Retweets: 34
RRN: 0
_____
Text: Texto3
User: @User3
Coordinates: Coordenadas3
Language: Português
Views: 35
Favorites: 4
Retweets: 2
RRN: 1
_____
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~> Text: Texto4
User: @User1
Coordinates: Coordenadas4
Language: Português
Views: 291
Favorites: 80
Retweets: 34
RRN: 0
~(/home/roberto/Copy/USP/Programação/C/Cristina/Tweet)~>
```