IMD0030
LINGUAGEM DE PROGRAMAÇÃO I

Aula 18 – Manipulação de Arquivos em C++

INSTITUTO METRÓPOLE DIGITAL
UFRN UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE

Objetivos desta aula

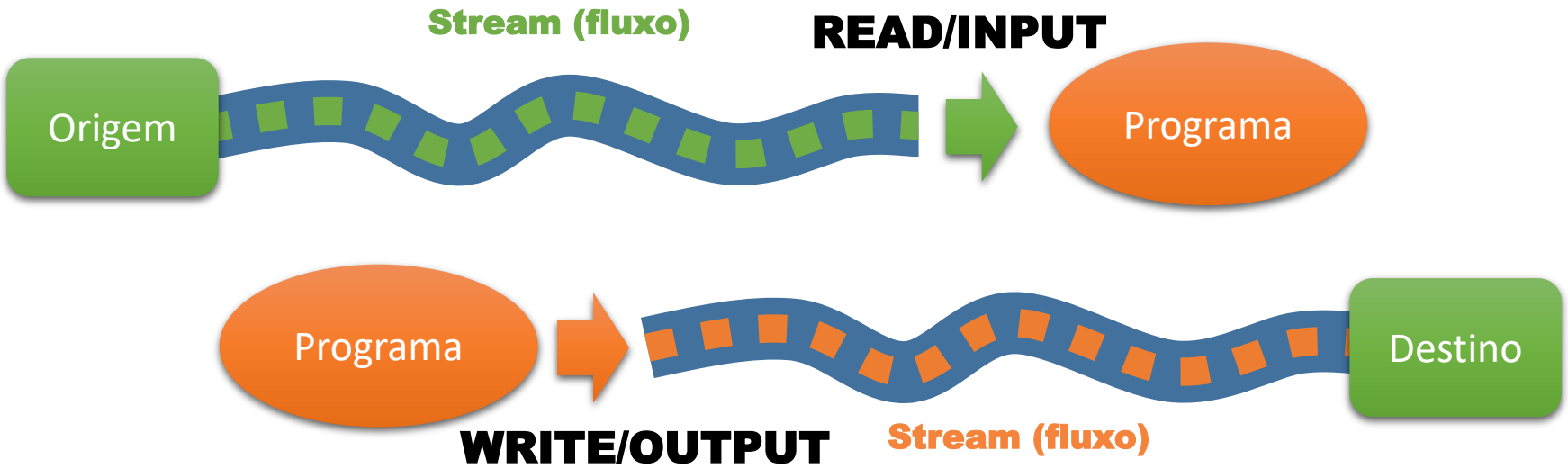
- Introduzir os conceitos e estruturas para a manipulação de arquivos em C++
- Para isso, estudaremos:
  - As principais bibliotecas e objetos do C++ que implementam operações a arquivos
  - As principais operações em arquivos no C++
  - Algumas operações de E/S em ***streams***
- Ao final da aula espera-se que o aluno seja capaz de:
  - Escrever programas em C++ que manipulem arquivos

Introdução

- As ações de entrada e saída de dados não fazem parte da linguagem C++.
- Como forma de uniformizar as primitivas através das quais um programa invoca estas ações de I/O (entrada e saída de dados), a linguagem C++ virtualiza todos os dispositivos envolvidos nestas ações como objetos ***streams***.
- C++ utiliza operações de E/S *type safe*
  - Cada operação de E/S é realizada de maneira sensível ao tipo de dado
- Extensibilidade
  - É possível especificar operações de E/S sobre tipos definidos pelo usuário da mesma forma como é feito para tipos padrão

O conceito de Stream no C++

- O conceito de ***Stream*** pode ser entendido como um fluxo de informação que pode entrar ou sair de um programa para uma fonte de informação que pode ser um arquivo, a memória, o teclado ou o vídeo
  - Desta forma, a escrita para qualquer um destes meios utiliza-se basicamente a mesmas classes e métodos, facilitando seu uso.



O conceito de Stream no C++

- Todos os dispositivos lógicos (streams) são semelhantes em comportamento, e bastante independentes dos dispositivos reais
- Distinguem-se dois tipos de streams
  - streams para texto
  - streams para palavras binárias
- Um stream, associa-se a um periférico realizando uma operação abertura (*open*), e desassocia-se dele com uma operação de fechamento (*close*).

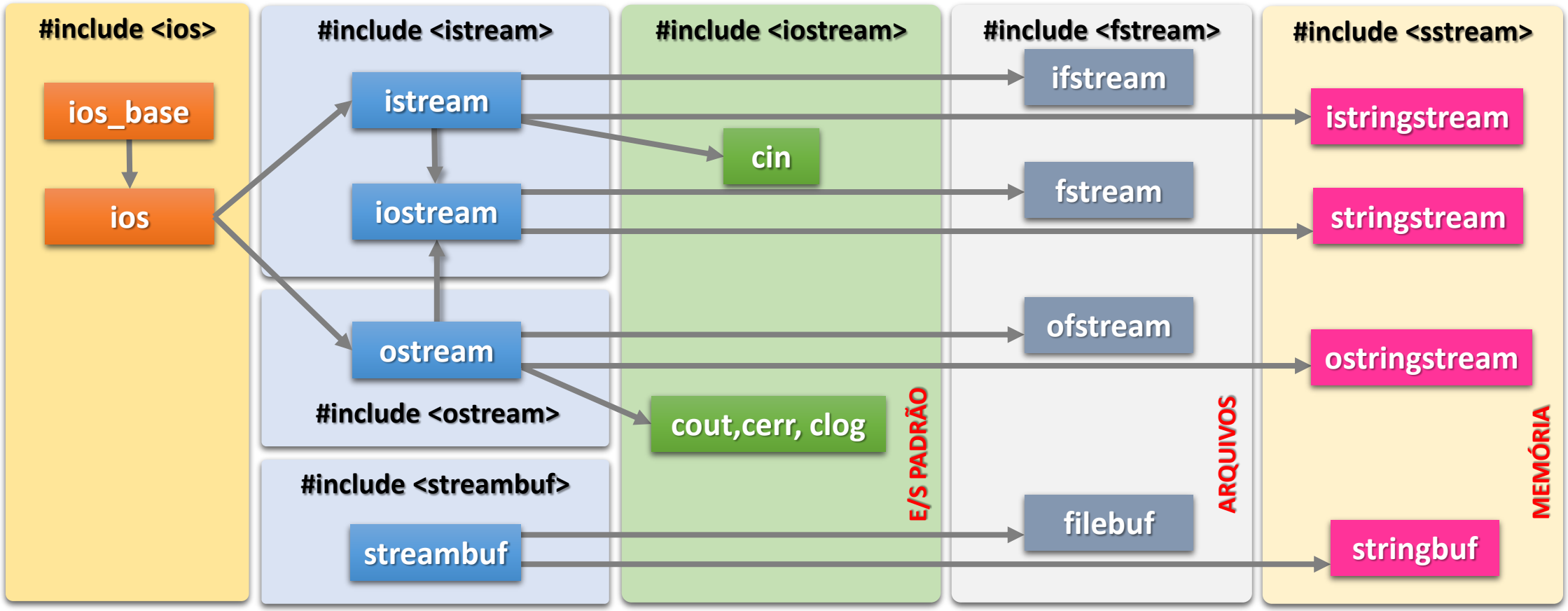
Buffer

- O buffer consiste em uma área de memória onde as informações são lidas ou escritas.
- O acesso ao buffer ocorre na velocidade da RAM enquanto os dispositivos de entrada e saída (teclado; disco; etc.) trabalham em uma velocidade de acesso muito menor.
- Quando parte de um arquivo é necessária, normalmente um volume maior de informações é lido e transferido para o buffer, quando uma próxima parte desta informação é necessária o seu conteúdo já se encontra no buffer.
- A utilização de um buffer reduz o número de acessos necessários a disco, por exemplo, e desta forma este tipo de I/O é mais eficiente.

E/S padrão

- Os serviços de E/S mais comuns do C++ (***cin***, ***cout*** e ***cerr***) são implementados através da biblioteca `iostream`
  - No C++, E/S pode ser feita tanto num arquivo quanto na memória (*streams*)
    - Um objeto ***stream*** (fluxo de caracteres) pode ser entendido como um lugar na memória reservado para o recebimento ou envio de *bytes*
  - Além disso, amplia o conceito de arquivo no sentido de considerar não somente os arquivos em disco mas também o teclado, o vídeo, a impressora e portas de comunicação (serial, USB, etc.)
    - Por exemplo, `cout` está associado ao vídeo (saída) e `cin` está associado ao teclado (entrada)
- O conceito de **herança** é largamente explorado no modelo de classes usado para E/S
  - Evita repetição de código em operações semelhantes
  - Permite a criação de classes e métodos polimórficos

Hierarquia de entrada/saída em C++



Entrada e saída em C++

- Operações de E/S:
  - **Entrada** (***Input***): Um *stream* flui de um dispositivo de entrada para a memória principal
  - **Saída** (***Output***): Um *stream* flui da memória principal para um dispositivo de saída
- E/S de baixo nível (*low-level I/O*)
  - Sem formato definido
  - Opera sobre *bytes* individuais
  - Boa para manipular grande volume de dados com alta velocidade
- E/S de alto nível (*high-level I/O*)
  - Formatada
  - Opera sobre conjunto de *bytes* agrupados em unidades com significado (inteiro, character, etc.)
  - Boa para todo tipo de E/S, exceto processamento de arquivos muito volumosos

E/S de *streams* em C++

- `ios`:
  - `istream` e `ostream` são classes herdadas de `ios`
  - `iostream` herda características das classes `istream` e `ostream`
    - `#include <iostream>` permite utilizer os *streams* de E/S `cin`, `cout`, `cerr` e `clog`
- << (*left-shift operator*)
  - Operador sobrecarregado que funciona como operador de inserção em *stream*
- >> (*right-shift operator*)
  - Operador sobrecarregado que funciona como operador de extração de *stream*
- Ambos os operadores << e >> são usados com os *streams* já conhecidos `cin`, `cout`, `cerr`, `clog`, assim como com objetos do tipo *stream* definidos pelo usuário

E/S de *streams* em C++

- `istream`: *input streams*
  - `cin >> varX`
    - `cin` sabe qual o tipo do dado a ser associado a `varX` (baseado no tipo definido para `varX`)
- `ostream`: *output streams*
  - `cout << varX`
    - `cout` sabe qual o tipo do dado a ser associado a `varX` (baseado no tipo definido para `varX`)
  - `cerr << varX`
    - Não armazena em *buffer* (*unbuffered*), ou seja, imprime o valor de `varX` imediatamente

Manipuladores

- Existem manipuladores para *streams*
  - de entrada, para alterar o formato das extrações de *stream*
  - de saída, para alterar o formato das inserções em *stream*
- A área de ação para cada manipulador começa com o aparecimento do manipulador e acaba depois de cancelado por outro manipulador

Manipuladores

Manipulador	*In*	*Out*	Definição
`endl`		√	Mudar de linha e *flush* do ostream
`ends`		√	Inserir '\0' para terminar *string*
`flush`		√	Esvaziar (*flush*) o *buffer* do ostream
`dec`	√	√	Conversão para base decimal
`hex`	√	√	Conversão para base hexadecimal
`oct`	√	√	Conversão para base octal
`ws`	√		Eliminar caracteres separadores
`setbase(int b)`	√	√	Fixar a base de conversão em *b*
`resetiosflags(long b)`	√	√	Desativar *bit*-vector *flags* de acordo com *b*
`setiosflags(long b)`	√	√	Ativar *bit*-vector *flags* de acordo com *b*
`setfill(int f)`		√	Definir o caracter de preenchimento de espaços do campo com (`char`)*f*
`setprecision(int n)`		√	Situar em *n* dígitos a precisão de um número em ponto-flutuante
`setw(int n)`	√	√	Colocar em *n* caracteres a largura do campo

Consulte a lista completa de manipuladores em: http://en.cppreference.com/w/cpp/io/manip

Acessando arquivos em C++

- Em **C++**, pode-se definidos objetos associados a arquivos e passar a interagir com esses objetos com os mesmos operadores, métodos e manipuladores que se utilizam para `cin` e `cout`
- Entre os vários objetos que podem ser criados para ter acesso a arquivos, destacam-se:
  - `ifstream` - quando queremos abrir um arquivo para leitura
  - `ofstream` - quando queremos abrir um arquivo para escrita
  - `fstream` - quando se deseja que o arquivo possa ser lido e escrito ao mesmo tempo
- Para poder utilizar E/S em arquivos, utilizar a biblioteca `fstream`
  - `#include <fstream>`

Manipulando streams

- Um objeto `ifstream` pode ser construído da seguinte forma:
  - `ifstream arqDados;`
    - o objeto `arqDados` é criado mas nenhum arquivo é aberto
  - Para abrir é preciso usar o método `open()`: `arqDados.open("dados.dat");`
- Um objeto `ifstream` também pode ser construído da seguinte forma:
  - `ifstream arqDados("dados.dat");`
    - o objecto `arqDados` é criado e o arquivo `dados.dat` é aberto em modo de texto para leitura

Manipulando streams

- Um objeto `ofstream` pode ser construído da seguinte forma:
  - `ofstream arqDados;`
    - O objeto `arqDados` é criado mas nenhum arquivo é aberto
  - Para abrir é preciso usar o método `open()`: `arqDados.open("dados.dat" , ios::binary);`
- Um objeto `ofstream` também pode ser construído da seguinte forma:
  - `ofstream arqDados("dados.dat", ios::binary);`
    - O objeto `arqDados` é criado e o arquivo `dados.dat` é aberto em modo binário para leitura
    - O parâmetro `ios::binary` indica que o arquivo deve ser aberto em modo binário

Manipulando streams

- Um objeto `fstream` (`ifstream` + `ofstream`) pode ser criado da seguinte forma:
  - `fstream arqDados("dados.dat", ios::in | ios::out | ios::binary);`
  - Neste exemplo, é criado o objeto `arqDados` para leitura/escrita em modo binário, associado ao arquivo `dados.dat`
- O modo de abertura padrão do fstream é `ios::in | ios::out`
  - Caso o arquivo não exista, a operação de abertura irá falhar, pois a porção `ios::in` irá falhar pelo fato de o arquivo não existir
  - Sempre que for necessário utilizar um `fstream` para leitura e escrita, é necessário **garantir que o arquivo existe**
    - Ou utilize um stream de escrita: `fstream(arq, ios::out)` ou `ofstream(arq)`

Manipulando streams

Importante lembrar:

- Um arquivo aberto por um objeto `ofstream` não necessita que definamos o modo de arquivo `ios::out`, pois este modo já é definido para este tipo de objeto por definição
- O mesmo ocorre com o modo `ios::in` e os objetos `ifstream`
- Além disso, por definição, um arquivo aberto por um objeto `ofstream` irá truncar os dados já existentes no arquivo, sobrescrevendo os dados novos nos antigos
  - Porém, o modo `ios::app` permite anexar dados ao final de um arquivo

Verificando a abertura do arquivo

- A verificação da abertura efetiva de um arquivo deve ser sempre realizada antes de efetuar qualquer operação de E/S sobre o mesmo
- Exemplos de verificação:

```
ifstream arqDados("dados.dat");

if(arqDados.bad()) {
    cerr << "o arquivo nao foi aberto" << endl;
    exit(1);
}

if(!arqDados) {
    cerr << "o arquivo nao foi aberto" << endl;
    exit(1);
}
if(arqDados.is_open() == 0) {
    cerr << "o arquivo nao foi aberto" << endl;
    exit(1);
}
```

Verificando o fim do arquivo

- O método `eof()` permite saber se foi atingido o fim do arquivo
- Exemplo:

```
ifstream arqDados("dados.dat");
while(!arqDados.eof()) {
    // lê arquivo
    // ...
}
```

Leitura de arquivos em C++

- A leitura de ***strings*** e **arquivos** é efetuada de forma semelhante, pois as suas classes derivam de uma classe de E/S comum (`ios`)
- Exemplos de leitura:

```
char str[10];
istrstream in_str("String com os dados a serem lidos.");
ifstream in_file("dados.dat");
in_str >> str;              // extrai “o tipo de dado” definido por str do arquivo
in_str
in_file >> str;             // extrai “o tipo de dado” definido por str do arquivo
in_file
in_str.read(str, 5);        // lê 5 caracteres do string in_str
in_file.read(str, 5);       // lê 5 caracteres do arquivo in_file
in_str.getline(str, 80);    // lê uma linha de no máximo 80 caracteres do string in_str
in_file.getline(str, 80); // lê uma linha de no máximo 80 caracteres do arquivo
in_file
```

Leitura de arquivos em C++

- Uma forma popular de leitura de arquivos no C++ utiliza o operador de extração
- Exemplo:

```
ifstream arqDados("dados.dat");
while(arqDados >> valor) {
    /* O operador de extração retornará 0 (false) quando encontrar EOF e o
       laço de repetição chegará ao fim */
}
```

Escrita de arquivos em C++

- Assim como na leitura, a escrita de ***strings*** e **arquivos** é efetuada de forma semelhante, pois as suas classes derivam de uma classe de E/S comum (`ios`)
- Exemplos de escrita:

```
char str1[] = "LP1";
strstream out_str;
fstream out_file("data.dat", ios::out);
out_str << str1;              // Escreve o conteúdo da string str1 no string out_str
out_file << str1;             // Escreve o conteúdo da string str1 no string out_file
out_str.write(str1, 3);       // Escreve, no máximo, 3 caracteres no string out_str
out_file.write(str1, 3);      // Escreve, no máximo, 3 caracteres no string out_file
```

Fechando o arquivo

- Como todo recurso em **C++**, os objetos associados a arquivos devem ser liberados após sua utilização, ou seja, quando já não forem mais necessários
  - No caso de arquivos em **C++**, os objetos que implementam os *streams* já se encarregam disso, através de métodos destrutores, logo **não é necessário fechar um *stream* manualmente**
  - Não há problema em liberar um *stream* manualmente, **mas não é o "estilo C++"** de fazê-lo
- Para liberar um objeto associado a um arquivo, deve-se utilizar o método `close()`
  - Exemplo:

```
ifstream arqDados("dados.dat");
// Utiliza o arquivo
// ...
arqDados.close();
```

std::getline

- O método std::**getline()** extrai caracteres de um stream até que o final de linha ou um delimitador especificado seja encontrado e armazena numa variável std::string passada como parâmetro.
- Assinaturas:

```
template< class CharT, class Traits, class Allocator >
std::basic_istream<CharT,Traits>& getline( std::basic_istream<CharT,Traits>&& input,
                                           std::basic_string<CharT,Traits,Allocator>& str,
                                           CharT delim );
```

  - Parâmetros:
    - ***input*** ***- o stream de entrada de onde devem ser lidos os dados***
    - ***str*** ***- a variável string que será usada para armazenar o valor lido***
    - ***delim*** ***- o character a ser usado como delimitador (seu valor padrão é '\n')***

Posicionando o ponteiro do arquivo

- O ponteiro de arquivo indica a posição no arquivo onde será feita a próxima leitura ou escrita
- No C++, há um conjunto de métodos que podem ser usados para movimentar o ponteiro do arquivo
  - `seekg(pos)`
    - Movimenta a posição atual de leitura para `pos`
    - Ex: `in_file.seekg(0); // retorna para o início do arquivo`
  - `seekp(pos)`
    - Movimenta a posição atual de escrita
  - `tellg()`
    - Retorna a posição atual de leitura (em *bytes*), a partir do início do arquivo
  - `tellp()`
    - Retorna a posição atual de escrita (em *bytes*), a partir do início do arquivo

Lendo e gravando para a memória

- As classes `sstream` interagem com a memória usando a mesma sintaxe usada na manipulação de arquivos em disco
- Exemplo:

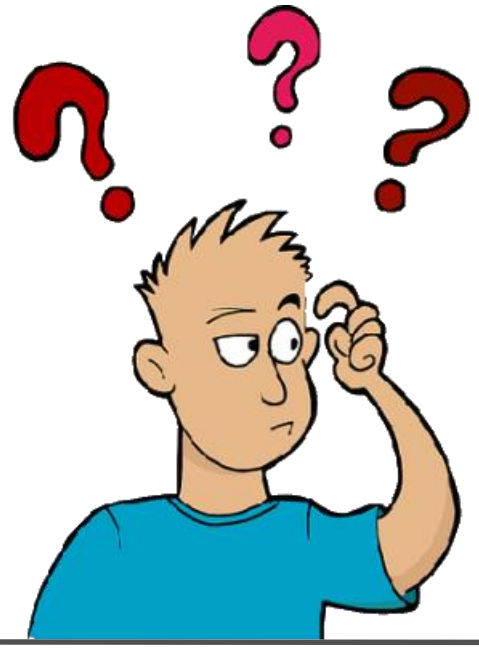
```
#include <iostream>
using std::cout;
using std::endl;

#include <string>
using std::string;

#include <sstream>
using
std::ostringstream;
```

```
int main() {
    ostringstream oss;
    oss << "Testando a escrita em memoria\n";
    oss << 123 << '\n';
    string s = oss.str();
    cout << s << endl;
    return 0;
}
```

Alguma Questão?



Exercício

Escreva um programa em C++ que leia um arquivo de texto no formato CSV (valores separados por vírgulas) referente às notas dos alunos de uma turma. O programa deverá computar a média de cada um e a situação de aprovação (aprovação com média maior ou igual a 7.0) ou reprovação e imprimir essas informações tanto na saída padrão (tela) quanto em um outro arquivo de texto. No arquivo de entrada, cada linha deve conter o nome do aluno seguido de três notas. A média deverá ser impressa com apenas uma casa decimal.

Exemplo de entrada:

```
Antonio Silva;10.0;9.0;8.0
Maria Joaquina;10.0;10.0;10.0
Carla Sousa;9.0;5.0;0.0
```

Exemplo de saída:

```
Antonio Silva 9.0 Aprovado
Maria Joaquina 10.0 Aprovado
Carla Sousa 4.7 Reprovado
```