```
Minclude <string.h>
Fdefine MAXPAROLA 30
#define MAXRIGA 80
   int treq[MAXPAROLA]; /* vettore di contatoni
delle frequenze delle lunghazza delle pitrole
   char riga[MAXRIGA] ;
lint i, inizio, lunghezza ;
```

# **Synchronization**

# **Task Programming in C++**

Stefano Quer
Dipartimento di Automatica e Informatica
Politecnico di Torino

## **License Information**

#### This work is licensed under the license









(cc) (i) (S) (=) CC BY-NC-ND 4.0

#### Attribution-NonCommercial-NoDerivatives 4.0 International

This license requires that reusers give credit to the creator. It allows reusers to copy and distribute the material in any medium or format in unadapted form and for noncommercial purposes only.

- BY: Credit must be given to you, the creator.
- S NC: Only noncommercial use of your work is permitted. Noncommercial means not primarily intended for or directed towards commercial advantage or monetary compensation.
- ND: No derivatives or adaptations of your work are permitted.

To view a copy of the license, visit: https://creativecommons.org/licenses/by-nc-nd/4.0/?ref=chooser-v1

### Introduction

- Multi-threading in C++ has two main limitations
  - The number of software threads may be higher than the number of hardware threads
    - Over-subscription occurs every time the number of software threads ready to start is higher than the number of hardware threads available in the system
    - Over-subscription implies system overhead and some performance penalty
    - With threads, the global load must be managed manually

➤ To solve this problem, C++11 introduced taskbased parallel programming

#### Introduction

# Multi-threading in C++ has two main limitations

2. Threads (the std::thread library) do not offer any direct way to return a value to the caller

- In POSIX
  - A simple strategy is return a value with pthread\_exit
  - More general strategies must be user-implemented (through global or local objects)
- In native C++ only the general strategy is available (you must manipulate objects explicitly)
- 3. Native C++ Threads: In native C++ (using std::thread), you don't have a direct way to return values. You need to manually manage shared objects or use other synchronization mechanisms to get data from threads.
- 4. Futures and Promises:

Returning Values from Threads: In traditional multi-threading (using std::thread), there is no built-in mechanism to return a value from a

2. POSIX Threads:

threads.

thread directly. This means you have to use workarounds like global or local objects to share data between threads.

POSIX threads (pthreads) offer a function pthread\_exit to return a value from a thread. However, for more complex

scenarios, you need to implement your own strategies to share data between

C++11 introduced futures and promises to address this limitation. These constructs provide a way to return values from threads in a more straightforward and manageable manner.



Promise: An object that allows you to set a value that will be available in the future. Future: An object that retrieves the value set by a promise once it becomes available.

To solve this problem, C++11 introduced futures
and promises

A promise object is created to hold the value that will be computed by the thread.
A future object is obtained from the promise, which will be used to retrieve the value.

A future object is obtained from the promise, which will be used to retrieve the value. A thread is started, and it sets the value in the promise.

The main thread waits for the future to get the value and then prints it.

A task is an entity that runs asynchronously, producing output data that will become available (and useful) at a later time.

- A task is an entity that runs asynchronously producing output data that will become available (and useful) at a later time
  - The operating system associate a thread to a task in an automatic way
  - Balancing tasks is automatic, through workstealing features
  - > Tasks have the possibility of handling return values

#### ❖ In C++

➤ Thread-based parallel programming relies on std::thread objects

std::thread is used to create a new thread that runs thread\_function. This is a low-level approach where you manually manage threads.

```
std::thread t(thread_function);
```

Task-based parallel programming relies on

**std::asynch** objects

Automatically manages task execution.

Returns a std::future object to retrieve the result.

Simplifies asynchronous programming by abstracting thread management

```
auto fut = std::asynch(thread_function);
```

std::async is a higher-level abstraction that automatically manages the creation and execution of tasks. It returns a std::future object, which can be used to retrieve the result of the asynchronous operation.

Comparison:

Thread-based programming with std::thread gives you more control but requires manual management of threads.

Task-based programming with std::async simplifies asynchronous execution and automatically handles thread management, making it easier to work with return values and manage tasks.

By using std::async, you can focus more on the logic of your tasks rather than the intricacies of thread management, which can lead to cleaner and more maintainable code.

Parameters for the

thread function

#include <future>
std::async, std::future, and related functionalities.
future<T>: This is a template class that represents a future value of type T. It is used to retrieve the result

#include <future>: This header file is necessary to use std::async, std::future, and related functionalities.

•future<T>: This is a template class that represents a future value of type T. It is used to retrieve the result of an asynchronous operation.

asynchrolicy function args. ): This function runs

async(policy, function, args...): This function runs function asynchronously, with args as its arguments, and returns a future<T>.

future<T> async(policy, function, args...);

<T> is the type of the future

Asynchronous policy

"Thread" function

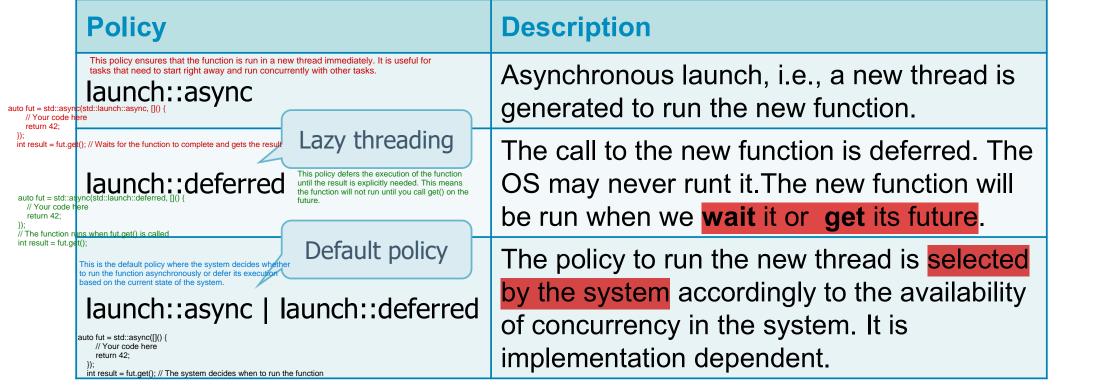
- <T>: The type of the future. This indicates the type of value that the asynchronous function will return.
- 2. policy: The asynchronous policy that dictates how the function should be executed.
- 3. function: The function to be executed asynchronously. 4. args...: The arguments to be passed to the function.

- Function async (namespace std)
  - ➤ Is an alternative to std::thread to execute functions in parallel
  - Has and extra parameter, i.e., the policy
  - Returns a future of type T

Alternative to std::thread: std::async is an easier and higher-level alternative to std::thread for executing functions in parallel. Extra Parameter (Policy): std::async takes an extra parameter called the policy, which determines how the function will be executed. Returns a Future: The function returns a future<T>, which can be used to retrieve the result of the asynchronous operation.

For now, ignore it

- The user may decide the running policy
  - > There are three different types of policies



# **Examples**

Running async tasks

```
auto f1 = std::async(std::launch::async, my f, 10);
// Thread function my f is run in a new thread
                                               For now, we do not
auto f2 = std::async(
                                              know what a future is
  std::launch::deferred, my f, 20);
// Thread function my f is not run until we get
// its results or wait for it
auto f3 = std::async(
   std::launch::async | std::launch::deferred,
   my f, 30);
// The system decides when running my f.
// Possibly, it never runs my f.
                Force task f2 to be associated
                 and run within a thread is
f2.wait()
// Invoke deferred function f2 (i.e., run it)
```

# **Example**

#### Running policy

```
auto f = std::async
                                                        (my f);
            // f.wait_for(0s): This checks the status of the future without waiting (0 seconds). It returns immediately with the status of the future
            if (f.wait for(0s)==std::future status::deferred)
                                                 It is deferred: Use wait to
                 f.wait();
                                                     force the execution.
                                                  It is asynch, it is already
             } else {
                                                            running.
                while (f.wait for(100ms)!=
                     std::future status::ready)
f.wait_for(100ms): This waits for 100
milliseconds and checks the status of
std::future status::ready: This status
indicates that the task is complete.
                              do something
The while loop continues to check the
status every 100 milliseconds.
```

Here the future f is ready

```
If it is async or deferred;
    it may never run
```

Wait for 0 seconds, i.e., do not wait, check the status

Check status every 100 msecons

If it is not ready, do something in parallel The slide demonstrates how to use std::async to

To wait for 0s or 100ms use std::literals

#### **Futures**

- An async object will eventually hold the return value of the thread function in a future
  - ➤ A future is an object that can represent a value generated by some provider
  - Function <future>::get applied to a valid future
    - Blocks the thread until the object is ready
    - Returns the object (return with "return") once it is ready

<T> is the type of the future

future<T> async(policy, function, args...);

# **Example**

```
#include <future>
                                     Check if num is
                                         prime
bool is prime (int n) {
  if (num <= 1) return false;</pre>
  if (num <= 3) return true;</pre>
  return false;
                                             Run a new function
                                                in a thread
int main () {
  std::future<bool> fut = std::async(
     std::launch::async, is prime, 117);
  // ... do other work ...
  bool ret = fut.get();
                                           Wait for function is_prime to
  cout << ret;</pre>
                                           return and make the Boolean
                                                value available
  return 0;
```

# **Example**

```
#include <future>
#include <iostream>
                              Run a new function
                                   thread
auto fut = std::async
  std::launch::async,
                                     Work with lambda
  [](){
                                        expressions
    std::vector<int> v;
    for (int i=0; i<100; i++)
       v.push back(i);
    return v;
                                Wait for the future to be ready
                                   and get the return value
auto ret = fut.get();
for (auto e: ret)
  std::cout << e << std::endl;</pre>
```

### **Shared futures**

# In C++ there are two types of future

Unique Future (std::future<T>): Only one instance referring to the event. Cannot be copied. Used when only one thread needs to retrieve the result

Single Instance: There is only one instance of a std::future referring to the event. This means that once a std::future object is created, it cannot be copied or shared with other std::future objects.

Ownership: The std::future object has unique ownership of the result. Once the result is retrieved using get(), the future becomes invalid.

Unique future, i.e., std::future<T>

There is only one instance referring to the event

Shared future, i.e., std::shared\_future<T>

Usage: Typically used when only one thread needs to wait for and retrieve the result of an asynchronous

std::future<int> fut = std::async(std::launch::async, []() { return 42; }); int result = fut.get(); // Retrieve the result, fut is now

Shared Future (std::shared\_future<T>): Definition: A shared future is similar to a unique future but can be copied and shared among multiple instances

Characteristics:

Copyable: A std::shared future object can be copied. allowing multiple instances to refer to the same event. Multiple Instances: Multiple std::shared\_future objects can refer to the same asynchronous result. Simultaneous Readiness: All instances of std::shared future will become ready at the same time, and the result can be retrieved from any of the

Signaling Multiple Threads: std::shared future can be used to signal multiple threads simultaneously, similar to how std::condition variable::notify all works.

- A shared\_future object behaves like a future object, except that it can be copied
- Multiple instances may refer to the same event
- All intances will become ready at the same time and can be retrieved
- May be used to signal multiple threads // Multiple threads can now use shared\_fut to get the result std::thread t1/(shared fut)) { simultaneously, similarly to std::condition variable::notify all

std::shared\_future<int> shared\_fut = std::async(std::launch::async, []() { return 42; }).share(); std::thread t1([shared\_fut]() { int result = shared\_fut.get() std::cout << "Thread 1: " << result << std::endl; std::thread t2([shared\_fut]() { int result = shared\_fut\_get() std::cout << "Thread 2: " << result << std::endl; t1.join(); t2.join();

Shared Future (std::shared\_future<T>): Can be copied and shared among multiple instances Multiple instances can refer to the same event. All instances become ready at the same time. Useful for signaling multiple threads simultaneously.

# **Example**

#### Unique future

```
int sum(int a, int b) {
   std::this_thread::sleep_for(std::chrono::seconds(2));
   return a + b;
int main() {
                                            How is it different for deferred? Deferred Execution: The function sum runs only when fut.get() is called
                                            Blocking: fut.get() will block until the function completes and the result is available
   std::future<int> fut =
      std::async(std::launch::async, sum, 10, 20);
                                               Wait and then get the future
   int result1 = fut.get();
   std::cout << "Result: " << result1 << endl;</pre>
                                                                           Result: 30
   int result2 = fut.get();
   std::cout << "Result: " << result2 << endl;</pre>
   return 0;
                               terminate called after throwing an instance of 'std::future error'
                                       what(): std::future error: No associated state
                                                 Aborted (core dumped)
```

# **Example**

#### Shared future

```
int sum(int a, int b) {
  std::this_thread::sleep_for(std::chrono::seconds(2));
  return a + b;
int main() {
  std::shared future<int> fut =
    std::async(std::launch::async, sum, 10, 20);
                                  Wait and then get the future
  int result1 = fut.get();
  std::cout << "Result: " << result1 << endl;</pre>
                                                      Result: 30
  int result2 = fut.get();
  std::cout << "Result: " << result2 << endl;</pre>
                                                      Result: 30
  return 0;
```

- The most common situation where you encounter a future is with a call to an async
  - An async returns a future
  - ➤ A future represents a value that you do not yet have but will have eventually
- At the lowest level, a future comes from an associated promise
- 1. Common Use Case:
- "The most common situation where you encounter a future is with a call to an async": When you use std::async, it returns a std::future object.
- "An async returns a future": std::async creates an asynchronous task and returns a future that will hold the result of that task.
- "A future represents a value that you do not yet have but will have eventually": A future is a placeholder for a result that will be available at some point in the future.
- Lowest Level: Future from Promise: "At the lowest level, a future comes from an associated promise": A future is often created from a promise.
- "A promise is an object that can store a value to be retrieved by a future object". A promise is used to set a value that a future will eventually hold.
- "When the value is set, it will be made available through its corresponding future": Once the promise sets the value, the future can retrieve it

- A promise is an object that can store a value to be retrieved by a future object
  - A promise is an object that you will eventually set
  - When the value is set, it will be made available through its corresponding future

Think of promise and future as creating a single-use channel for data": A promise and future pair can be thought of as a one-time communication channel.

# Think of promise and future as creating a single-use channel for data "Creates the channel": The promise creates the channel. "Writes data in the channel": The promise sets the value that will be communicated through the channel.

- > A promise
  - Creates the channel
  - Writes data in the channel

Promise Name

```
std::promise<type> pn;
pn.set_value(...);
```

#### > A future

- Connects to the other end of the channel
- Waits and reads the data once it has been written

Connects to the other end of the channel": The future is connected to the promise and will receive the value set by the romise.

"Waits and reads the data once it has been written": The future waits for the promise to set the value and then ret<u>rieves it.</u>

Future Name

```
auto fn = pn.get_future();
fn.get();
```

- The principal steps are
  - > The main thread
    - Defines a promise
    - Associate a future to the promise
  - > The working thread
    - Receive the promise
    - Executes the function and fulfills the promise

```
pn.set_value(...);
```

std::promise<type> pn;

auto fn = pn.get future();

> The main thread retrieves the result

```
fn.get();
```



If a promise is destroyed without setting a value (using set\_value or set\_exception), the associated future will store an exception.

When get is called on the future, it will throw an exception if the promise was destroyed without setting a value.

If we destroy the promise without setting a value, an exception is stored in the object

- Function get will return
- > The object associated to a promise is usually stored in the heap as it cannot be stored
  - In the setter of the promise, as the setter can die
  - In the getter of the future, as we futures can be shares among several getters

The object associated to a promise is usually stored in the heap as it cannot be stored":

The value or exception associated with a promise is typically stored on the heap because the promise and future may outlive the scope in which they were created.

**Future** 



Storage



Promise

Promise and future. One thread (lambda)

# **Example**

```
Thread function
int f(int x) { return x + 1; }
                                                                  Define the future from
        Define the promise
                                                                       the promise
   This is the producer-write end
                                                                   This is the consumer-
                                                                         read end
std::promise<int> promise;
auto future = promise.get future();
                                                        Get the promise in
                                                          the capture list
// Launch f asynchronously
std::thread thread([&promise]
                                                        (int x)
       promise.set value(f(x));
                                                      The thread function is a lambda that captures the promise by reference (&promise) and takes an integer x as ar
                                                         the lambda, promise.set value(f(x)); is called. This sets the value of the promise to the result of f(x). In this case, f(5) is called, which
                             Set the value (to be
                      communicated) into the promise
                                                                                 Get the value
```

std::cout << future.get() << std::endl;</pre>

#### Promise and future. Two threads (lambdas)

# **Example**

This example demonstrates how to use std::promise and std::future with two separate threads: one for setting the promise and another for getting the future. Here's a step-by-step explanation:

Define the promise and the future from the promise

```
auto promise = std::promise<std::string>(); A std::promise object is created to hold a std::string. This is the producer end, which will set a value in the future.
auto future = promise.get future();
auto producer = std::thread([&]
                                                                                                     Run the thread
    promise.set value("Hello World");
                                                                                                  setting the promise
              Inside the lambda, promise.set_value("Hello World"); is called. This sets the value of the promise to the string "Hello World"
                                   Another thread is launched. The thread function is a lambda that captures the future by reference (&).
auto consumer = std::thread([&]
                                                                                                     Run the thread
     std::cout << future.get();</pre>
                                                                                                   getting the future
        set, it retrieves the value (which is "Hello World") and prints it.
producer.join();
                                         producer.join();: This ensures that the producer thread completes execution before the program continues.
                                         consumer join();: This ensures that the consumer thread completes execution before the program continues.
consumer.join();
```

Promise and future. One thread (function)

# **Example**

```
One-way communication:
#include <future>
                                               The thread set the promise and get
                                                         the future
using namespace std;
```

```
void factorial (const int &N, promise<int>& pr) {
  int res = 1:
  for (int i=N; i> 1; i--)
  res *=i;
 pr.set value(res);
int main () {
 promise<int> p;
  future<int> f = p.get future();
  thread t = thread(factorial, 4, ref(p));
  // here we have the data
  int x = f.get();
  t.join();
```

Define the promise and the future from the promise

the reason we use std::move to pass the promise to the thread function is because std::promise objects are not copyable, only movable. This means you cannot make a copy of a promise dbject, but you can ise it ensures that only one specific thread has ownership and therefore can fulfill the promise.

std::ref is used to create a reference\_wrapper, which is a copyable and assignable object that emulates a reference. It's typically used when you need to pass references to functions that expect copies, such as std::thread. However, in the context of std::promise, using std::ref would not be appropriate. This is because std::promise is meant to be moved, not copied or passed by reference.

ref generates an object of type promise<int> to hold a reference to p

One async task. Two way sync

# **Example**

```
Two-ways communication:
                                          The caller set the promise and get
#include <future>
                                                  the future
using namespace std;
int factorial (std::future<int>& f ) {
  int res = 1;
  int N = f.get();
                                             The future must be passed by
                                               reference, since it doesn't
  for ( int i=N; i> 1; i-- )
                                               support copy semantics
  res *=i;
  return res;
                                       Define the promise and the
                                         future from the promise
int main () {
  std::promise<int> p;
  std::future<int> f = p.get future();
  std::future<int> fu =
    async(std::launch::async,factorial,std::ref(f));
  p.set value(4);
  int x = fu.get();
```

Two async tasks.
Two way sync

# Example

```
void func1 (promise<int> p) {
  int res = 18;
 p.set value(res);
int func2 (future<int> f) {
  int res=f.get();
  return res;
                        The move semantics is
int main ()
                         achieved by std::move
  promise<int> p;
  future<int> f = p.get future();
  future<void> fu1 = async(func1, move(p));
  future<int> fu2 = async(func2, move(f));
  int x = fu2.qet();
  return 0;
```

# **Example**

#### Shared future

```
usign namespace std;
int factorial (shared future<int> f) {
  int res = 1;
  int N = f.qet();
  for (int i=N; i> 1; i--) res *=i;
  return res;
int main () {
  promise<int> p;
  future<int> f = p.get future();
  shared future<int> sf = f.share();
  future<int> fu1 = async(std::launch::async, factorial, sf);
  future<int> fu2 = async(std::launch::async,factorial, sf);
  future<int> fu3 = async(std::launch::async, factorial, sf);
  p.set value(4);
  int r1=fu1.get(); int r2=fu2.get(); int r3=fu3.get();
  return 0;
```

# **Conclusions**

# The task-based approach

- Makes the OS in charge of the parallelism
- ➤ Makes the return value of a thread/task accessible
- > Run threads with a smart policy
  - CPU load balancing
    - The C++ library can run the function without spawning a thread
  - Avoid the raising of std::system\_error in case of thread number reached the system limit
- Allows futures to catch exceptions thrown by the function
  - With std::thread() the program terminates

## Conclusions

# Thread-based approach

- Is used to execute tasks that do not terminate till the end of the application
  - A thread entry point function is like a second, concurrent main
- > It is a more general concurrency model
  - Can be used for thread-based design patterns
- > Allows us to access to the pthread native handle
  - Makes the programmer in charge of the parallelism
  - Useful for advanced management (priority, affinity, scheduling policies, etc.)

#### **Exercise**

Resorting only to asynchrnous tasks, write a C++ program to perform the matrix multiplication

$$\triangleright$$
 C = A x B

#### Constraints

- ➤ The number of columns in A must equal the number of rows in B
- ➤ Use C++ containers
  - Implement the matrices as vector of vectors
    - std::vector<std::vector<int>> a, b;
- ➤ To compute a sum of products, it is possible to use the function **std::inner\_product**

```
Utility function:
#include ...
                                      Generate a random matrix
void generateRandomMatrix
  vector<vector<int>>& m, int nRow, int nCol) {
  for (int i = 0; i < nRow; i++) {
    vector<int> row;
    for (int j = 0; j < nCol; j++)
      row.push back(rand()%3);
    m.push back(row);
                                    Utility function:
                                    Display a matrix
void printMatrix(const std::vector<std::vector<int>>& m) {
  for (auto & row : m) {
    for (int element : row)
      cout << element << " ";</pre>
    cout << endl;</pre>
```

#### Function threads

```
int computeSumOfProducts (
  const std::vector<int>& v1, const std::vector<int>& v2) {
  return std::inner_product (
    v1.begin(), v1.end(), v2.begin(), 0
  );
}
```

Return result

#### Until C++11

Computes the sum of products on the range [v1.begin, v1.end] and the range beginning at v2.begin, initializing the accumulator at 0

#### Main thread

```
int main() {
                                                             Matrices
  int nRowA = 2, nColA = 3, nRowB = 3, nColB = 2;
  vector<std::vector<int>> a, b;
  vector<std::vector<std::future<int>>> futures;
                                                           Matrix of
  ofstream outputFile;
                                                            futures
  cout << "Insert size of matrix A" << endl;</pre>
  cin >> nRowA >> nColA;
  cout << "Insert size of matrix B" << endl;</pre>
  cin >> nRowB >> nColB;
                                                         Read the size
                                                          of A and B
  if(nColA != nRowB) {
    cout << "Wrong size colA != rowB" << endl;</pre>
    return -1;
```

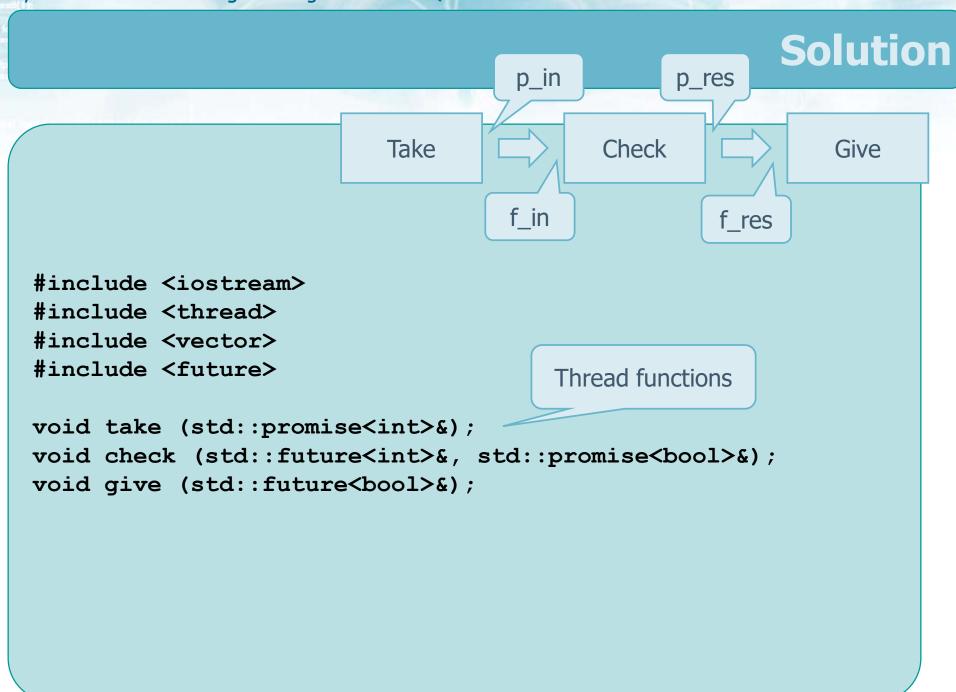
```
Generate matrix A
generateRandomMatrix(a, nRowA, nColA);
// generate the transpose of B (columns -> rows)
// so that we can easily access the columns of B
// for the multiplication
                                                Generate matrix B
generateRandomMatrix(b, nColB, nRowB);
// Generate the futures to compute the products
for (int i = 0; i < nRowA; i++) {
  vector<std::future<int>> futureRow;
                                                  Create a future for
  for (int j = 0; j < nColB; j++) {
                                                    each element
    future<int> f = std::async(
      std::launch::async | std::launch::deferred,
      computeSumOfProducts, a[i], b[j]);
    // Futures are not copyable so we have to use move
    // to store them in a vector
                                                Insert futures in the
    futureRow.push back(std::move(f));
                                                 matrix of futures
  futures.push back(std::move(futureRow));
```

```
cout << "A" << endl;</pre>
                                         Display input matrices
printMatrix(a);
cout << "B" << std::endl;</pre>
printMatrix(b);
outputFile.open("./output-matrix.txt");
for (auto & row : futures) {
                                                  Wait futures to be ready
  for (std::future<int> & f : row) {
                                                    and display result
     outputFile << f.get() << " ";</pre>
  outputFile << std::endl;</pre>
outputFile.close();
return 0;
```

#### **Exercise**

Exam 5 July 2021

- Write a C++ program with three tasks
  - > Thread **take** reads a number from command line
  - ➤ Thread **check** checks whether the number is prime
  - Thread give displays the answer to standard output
- Thread communication should be made using promises and futures
  - > All functions are acyclic



```
Solution
                                   p_in
                                              p_res
                                         Check
                         Take
                                                          Give
int main(){
                                  f_in
                                                 f_res
  std::promise<int> p in;
  std::future<int> f in = p in.get future();
 std::promise<bool> p res;
 std::future<bool> f_res = p_res.get_future();
  std::thread t1(take, std::ref(p in));
  std::thread t2(check, std::ref(f in), std::ref(p res));
  std::thread t3(give, std::ref(f res));
  t1.join();
  t2.join();
  t3.join();
  return 0;
```

#### Reading thread

```
void take (std::promise<int> &p in) {
  int in;
  std::cout << "Insert a number" << std::endl;</pre>
  std::cin >> inp;
                             Set promise "in"
  p in.set value (in);
             Writing thread
void give (std::future<bool>& f res) {
  bool answer = f res.get();
                                                Get future "ref"
  std::string s0 (" ");
  if(!answer)
    s0=" NOT";
  std::cout << "Number is" << s0 << " prime";</pre>
```

#### Computation thread

```
void check
  std::future<int> &f_in, std::promise<bool>& p res)
    int n = f in.get();
                                 Get future "in"
    bool prime=true;
    if (n \le 1) {
      prime = false;
    // Check from 2 to n-1
    for (int j=2; j<n; j++) {
      if (n % j == 0) {
        prime = false;
        break;
                                   Set promise
    p_res.set_value(prime);
                                      "res"
```