#### CS3211 Tutorial 7.5 & 8

#### Safety and Concurrency in Rust Simon – T5

Adapted From Sriram's Slides

### Why talk about a third language?

C++ (50% of the mod)

- Threads
- Synchronization
- Atomics
- Memory Ordering
- Debugging
- Lock-Free Programming

Powerful, but needs us to be really careful. Significant mental load?

The promise of performance, but programmer is responsible for safety

Go (25% of the mod)

- Lightweight Co-routines (Goroutines)
- Channels and Message Passing

Useful mental model, but garbage collection, copying in channels, etc, may impact performance.

# Rust (25% of the mod)

- Compile-time safety checking
- Safe futures / async
- Safe data parallelism
- Etc

Can we have a <u>fast</u>, <u>modern</u> language that reduces the programmer's <u>mental correctness</u> <u>burden</u> somewhat?

### What is the future of prog lang?

Chromium will support third-party Rust libraries



Google has announced that it will allow t browser project.

Chrome security team member Dana Jai decision.

Jansens says that Google is now actively

LINUX / OPEN SOURCE / RUST / SOFTWARE DEVELOPMENT

#### Rust in the Linux Kernel

Why it's all happening for the Rust programming language, how it made it into the Linux kernel, and where it will go from here.

Oct 5th, 2022 7:00am by Steven J. Vaughan-Nichols



**ENGINEERING & DEVELOPERS** 

# WHY DISCORD IS SWITCHING FROM GO TO RUST

lз

#### The Guarantees of Rust

#### Safe Rust ⇒ No Undefined Behavior, Yes to Memory Safety

- No use-after-free, double-free, yes to bounds checks, panics
- References are always valid and variables are initialized before use
- Data races are completely eliminated
- Etc etc

(\*not all can be done at compile time, also safe code calls unsafe code..)

#### But, in Real Rust, all of these are still possible!

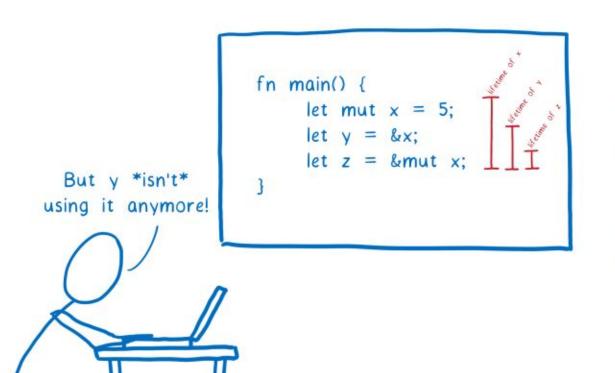
- Deadlock, livelock, etc
- Memory leaks
- Integer overflow...

#### Further reading:

https://doc.rust-lang.org/nomicon/meet-safe-and -unsafe.html

https://tiemoko.com/blog/blue-team-rust/

#### One more note: Rust is not simple!



Wait, you can't compile that!

z needs to be the only one borrowing x's value, but y is already using it.

# Let's learn (what we can) about Rust

What happens in this C++ program? [p]

```
#include <iostream>
     #include <string>
     void f(std::string* s) {
        free(s);
 9
     int main(int argc, char* argv[]) {
10
       std::string* s = new std::string("asdf");
11
12
       f(s);
       std::cout << *s;
13
```

#### What happens in this C++ program? [p]

 Undefined behavior: use-after-free (note: you could pick this up with ASan)

```
#include <iostream>
#include <string>

void f(std::string* s) {

free(s);

}

int main(int argc, char* argv[]) {

std::string* s = new std::string("asdf");

f(s);

std::cout << *s;

}</pre>
```

```
x86-64 qcc 12.2
                                          -fsanitize=address
Program returned: 1
 rogram stderr
   #0 0x7efc3e357898 (/opt/compiler-explorer/gcc-12.2.0/lib64/libasan.so.8+0xba898)
    #1 0x40236d in f(std::__cxx11::basic_string<char, std::char_traits<char>, std::al
   #2 0x402453 in main /app/example.cpp:12
   #3 0x7efc3dd37082 in __libc_start_main (/lib/x86_64-linux-qnu/libc.so.6+0x24082)
   #4 0x40229d in _start (/app/output.s+0x40229d)
 x603000000010 is located 0 bytes inside of 32-byte region [0x603000000010,0x60300000
   #0 0x7efc3e3596b8 in operator new(unsigned long) (/opt/compiler-explorer/gcc-12.2
   #1 0x4023f9 in main /app/example.cpp:11
   #2 0x7efc3dd37082 in __libc_start_main (/lib/x86_64-linux-gnu/libc.so.6+0x24082)
SUMMARY: AddressSanitizer: alloc-dealloc-mismatch (/opt/compiler-explorer/gcc-12.2.0/
==1==HINT: if you don't care about these errors you may set ASAN_OPTIONS=alloc_dealloc
=1==ABORTING
```

Translated to Rust: what happens? [np]

```
3 fn f(x: String) {
    // Think C++'s free(x) for a start
    drop(x)
6 }
7
8 fn main() {
    let x: String = String::from("asdf");
10    f(x);
11    print!("{}", x)
12 }
```

#### Translated to Rust: what happens? [np]

https://play.rust-lang.org/?version
 =stable&mode=debug&edition=202
 1&gist=336f34516c2edf45a31694facf
 cf955b

```
fn f(x: String) {
    // Think C++'s free(x) for a start
    drop(x)
}

fn main() {
    let x: String = String::from("asdf");
}
```

What about this? [p]

```
3 fn f(x: String) {
4
5 }
6
7 fn main() {
8 let x: String = String::from("asdf");
9 f(x);
10 print!("{}", x)
11 }
```

- What about this? [p]
  - Same issue!
  - Output
    But why??

https://play.rust-lang.org/?version=stable&mode=debug&editon=2021&gist=6ecc6467b2aafdf94405647a201bf570

```
3 fn f(x: String) {
4
5 }
6
7 fn main() {
10 8 let x: String = String::from("asdf");
9 f(x);
```

- What about this? [p]
  - Same issue!
  - Too hard to decide validity of ownership on case-by-case basis, the strict rules might lead to some "valid + safe" code to be rejected => no false positives

```
fn f(x: String) {
7 - fn main() {
        let x: String = String::from("asdf");
        f(x);
        print!("{}", x)
10
 3 fn f(x: String) {
       if (based_on_user_input_at_runtime) {
           drop(x)
```

What about this? [p]

#### What about this? [p]

- Totally valid!
- Ownership of string passed to f
- f owns the string, so
   ownership can be passed
   back through return value
  - String is "move"d
- Still only one explicit owner
- https://play.rust-lang.org/?versio n=stable&mode=debug&edition= 2021&gist=2701de5060e894e20 cc960a57a5bb82a

```
3 fn f(x: String) -> String {
         x // same as return x
 8 fn main() {
         let x: String = String::from("asdf");
         let y = f(x);
10
         print!("{}", y)
      Compiling playground v0.0.1 (/playground)
       Finished dev [unoptimized + debuginfo] target(s) in 1.46s
        Running `target/debug/playground`
   asdf
```

#### What did we just learn?

Concept of "Ownership"

Governs compiler-checked automatic memory management \*\*without GC\*\*

```
3 fn f(x: String) {
4     // Think C++'s free(x) for a start
5     drop(x)
6 }
7
8 fn main() {
9     let x: String = String::from("asdf");
10     f(x);
11     print!("{}", x)
12 }
```

- Each value in Rust has an owner.
- There can only be one owner at a time.
- When the owner goes out of scope, the value will be dropped.

- Working with owners:borrowing
- &T: Temporary
   immutable
   reference to the
   owned object there can be
   many at a time!

```
3 fn f(x: &Vec<u8>) {
4    print!("Borrowed version of x in f: {:?}\n", x)
5 }
6
7 fn main() {
8    let x: Vec<u8> = vec![1, 2, 3, 4];
9    print!("x from start of main: {:?}\n", x);
10    f(&x);
11    print!("x from end of main: {:?}\n", x);
12 }
```

```
x from start of main: [1, 2, 3, 4]
Borrowed version of x in f: [1, 2, 3, 4]
x from end of main: [1, 2, 3, 4]
```

Working with owners: immutableborrowing

```
3 fn f(x: &Vec<u8>) {
4     print!("Borrowed version of x in f: {:?}\n", x)
5 }
6
7 fn main() {
8     let x: Vec<u8> = vec![1, 2, 3, 4];
9     print!("x from start of main: {:?}\n", x);
10     f(&x);
11     print!("x from end of main: {:?}\n", x);
12 }
```

&T: Temporary
 immutable
 reference to the
 owned object there can be
 many at a time!

```
x from start of main: [1, 2, 3, 4]
Borrowed version of x in f: [1, 2, 3, 4]
x from end of main: [1, 2, 3, 4]
```

Why print! macro doesn't own the arguments: https://stackoverflow.com/questions/30450399/does-println-borrow-or-own-t he-variable

#### &mut T:

Temporary
mutable
reference to the
owned object:
there can only be
one at a time!

 If there is a valid &mut T, then no &T allowed simultaneously

```
3 fn f(x: \&mut Vec< u8>) {
        x[1] = 244;
        print! ("Borrowed version of x in f: \{:?\}\n", x)
   fn main() {
        let mut x: Vec<u8> = vec![1, 2, 3, 4];
        print!("x from start of main: {:?}\n", x);
        f(\&mut x);
11
12
        print!("x from end of main: {:?}\n", x);
```

```
x from start of main: [1, 2, 3, 4]
Borrowed version of x in f: [1, 244, 3, 4]
x from end of main: [1, 244, 3, 4]
```

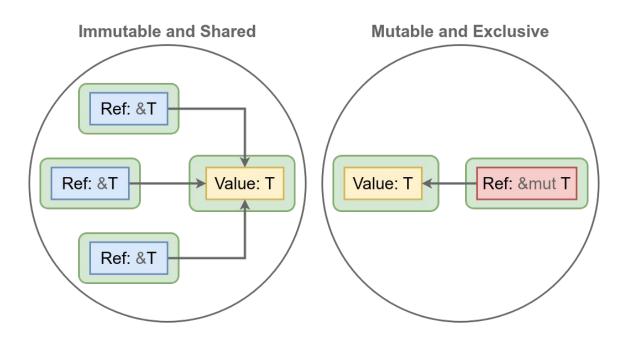
#### Extra Challenge

```
let mut x = \text{vec}![1,2,3];
let first = x[0];
let mut second = x[1];
second += 1;
```

Is this allowed?

# What did we just learn?

Concept of "Borrowing"



(Exclusivity guaranteed at compile time)

#### Lifetimes

- This rust code doesn't compile!
- Problem: what is the appropriate lifetime of the "borrowed" return value? Will it "live as long" as our inputs?

```
1 fn longest(x: &str, y: &str) -> &str {
        if x.len() > y.len() {
        } else {
   fn main() {
        let string1 = String::from("abcd");
        let string2 = "xyz";
13
        let result = longest(string1.as_str(), string2);
        println!("The longest string is {}", result);
14
15
```

https://github.com/dtolnay/rust-faq #why-arent-function-signatures-infe rred

#### Lifetimes

```
1 fn longest<'a>(x: &'a str, y: &'a str) -> &'a str {
2    if x.len() > y.len() {
3         x
4    } else {
5         y
6    }
7 }
```

- We use lifetimes to tell the compiler here: the output str will live as long as the inputs (which have the same lifetime)
- Now the compiler can go and verify this (as part of its type system)!

#### Lifetimes

https://play.rust-lang.org/?version=stable&m ode=debug&edition=2021&gist=0cb691feed7 b15c04e1e5a6e90d41cc7

```
1 fn print_longest_return_first_arg<'a, 'b>(x: &'a str, y: &'b str) -> &'a str {
2     if x.len() > y.len() {
3         print!("X is longer!\n")
4     } else {
5         print!("Y is longer!\n")
6     }
7     x
8 }
```

- An example of multiple lifetimes that the compiler can verify!
- We only return the first argument, so we can specify that the return value follows the lifetime of the first argument

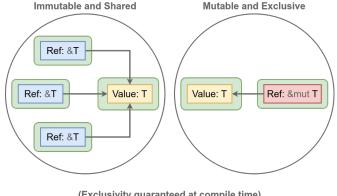
#### What did we just learn?

 Lifetimes: proving to the compiler that we always have valid references in our program!

```
fn longest<'a>(x: &'a str, y: &'a str) -> &'a str {
    if x.len() > y.len() {
                           Return values checked against
                             annotations in signature
fn example_with_helper() {
                                      Lookup lifetime annotations in type
    let string1 = "abcd";
                                       signature to determine how long
    let string2 = "xyz";
                                           reference should live for
     let result
                   longest(&string1, &string2);
    println!("The longest string is {}", result);
```

## TLDR - First pass of interesting parts of Rust

- Ownership
- Borrowing (Mutable & Immutable)
- Lifetimes



(Exclusivity guaranteed at compile time)

```
3 fn f(x: String) {
     drop(x)
  fn main() {
      let x: String = String::from("asdf");
      f(x);
      print!("{}", x)
```

```
fn longest<'a>(x: &'a str, y: &'a str) \rightarrow &'a str {
    if x.len() > v.len() {
                          Return values checked against
                              annotations in signature
fn example with helper() {
                                       Lookup lifetime annotations in type
    let string1 = "abcd";
                                       signature to determine how long
    let string2 = "xyz";
                                           reference should live for
    let result = longest(&string1, &string2);
    println!("The longest string is {}", result);
```

# **Tutorial: (1) Concurrent Counter**

 Recall: Our data-race / undefined behavior C++ code to add to a counter from multiple threads

```
int counter;
int main() {
  std::thread t0{[]() { ++counter; }};
  std::thread t1{[]() { ++counter; }};
  t0.join();
  t1.join();
  std::cout << counter << std::endl;
  return 0;
```

 Direct translation to Rust has <u>compile</u> errors! Let's try to fix...

```
use std::thread;
fn main() {
  let mut counter = 0;
  let t0 = thread::spawn(|| { counter += 1; });
  let t1 = thread::spawn(|| { counter += 1; });
  t0.join();
  t1.join();
  println!("{}", counter);
```

```
error[E0373]: closure may outlive the current function, but it borrows
 --> src/main.rs:7:26
     let t1 = thread::spawn(|| { counter += 1; });
                                  ----- `counter` is borrowed here
                             may outlive borrowed value `counter`
note: function requires argument type to outlive `'static`
 --> src/main.rs:7:12
     let t1 = thread::spawn(|| { counter += 1; });
help: to force the closure to take ownership of `counter` (and any ot
     let t1 = thread::spawn(move || { counter += 1; });
error[E0502]: cannot borrow `counter` as immutable because it is also
  --> src/main.rs:12:18
      let t0 = thread::spawn(|| { counter += 1; });
                                   first borrow occurs due to use of
                              mutable borrow occurs here
                argument requires that 'counter' is borrowed for '!sta
```

https://play.rust-lang.org/?version=stable &mode=release&edition=2021&gist=19d0 9528931f6431500585c66cd49661 29

- Let's listen to the compiler
- Output of this program? [p]

```
fn main() {
     let mut counter = 0;
5
    let t0 = thread::spawn(move || { counter += 1; });
     let t1 = thread::spawn(move | | { counter += 1; });
8
     t0.join();
     t1.join();
     println!("{}", counter);
```

https://play.rust-lang.org/?version=stable&mode=release&edition=2021&gist=738271fb046792be74c3aaf10a0c8008

- Let's listen to the compiler (but not the right warning :))
- Output of this program? [p] (Outputs 0!)
- Each thread tries to move counter into its scope (to own it)
  - Counter is Copyable so each thread just makes a private copy of counter!!

```
3 fn main() {
      let mut counter = 0;
 5
     let t0 = thread::spawn(move | { counter += 1; });
     let t1 = thread::spawn(move || { counter += 1; });
     t0.join();
     t1.join();
10
     println!("{}", counter);
13
```

- Problem 1: counter is shared and mutable not allowed!
- Problem 2: compiler thinks threads might outlive counter!
  - Possible dangling reference to counter?

```
use std::thread;
   fn main() {
      let mut counter = 0;
      let t0 = thread::spawn(|| { counter += 1; });
      let t1 = thread::spawn(|| { counter += 1; });
      t0.join();
10
      t1.join();
11
      println!("{}", counter);
12
13
```

- Fixing problem 1: put our counter in a Mutex!
- Problem?

```
use std::thread;
  use std::sync::{Mutex};
3
  fn main() {
    let counter: Mutex<i32> = Mutex::new(0)
    let t0 = thread::spawn(|| *counter.lock().unwrap()
    let t1 = thread::spawn(|| *counter.lock().unwrap() += 1);
8
    t0.join();
    t1.join();
    println!("{}", *counter.lock().unwrap());
```

Problem 2 appears: compiler is not convinced that our threads finish before the end of main(), when counter is dropped
 https://play.rust-lang.org/?version=stable&mode=release&edition=2021&gist=b22193c1e9298500fd432a055ace0497

```
use std::thread;
  use std::sync::{Mutex};
3
   fn main() {
    let counter: Mutex<i32> = Mutex::new(0);
    let t0 = thread::spawn(|| *counter.lock().unwrap() += 1);
     let t1 = thread::spawn(|| *counter.lock().unwrap() += 1);
8
    t0.join();
    t1.join();
    println!("{}", *counter.lock().unwrap());
```

https://play.rust-lang.org/?version=s table&mode=release&edition=2021 &gist=fe97826e618ddbdea7f4c8c6 3b33aaa3

- Fixing problem 2: wrap the mutex in a shared-pointer equivalent ("Atomically Reference Counted") type
- The counter will be dropped when both references die so the compiler is satisfied!

```
4 fn main() {
     let counter = Arc::new(Mutex::new(0));
     let counter1 = counter.clone();
    let counter2 = counter.clone();
     let t0 = thread::spawn(move || { *counter1.lock().unwrap() += 1 });
     let t1 = thread::spawn(move || { *counter2.lock().unwrap() += 1 });
     t0.join();
     t1.join();
     println!("{}", *counter.lock().unwrap());
```

## Why does this question matter?

- Rust will literally prevent you from writing non-thread-safe code!
- The compiler is not some genius oracle
- But knows that if you follow a set of (more restrictive that necessary) rules, your program will be safe

## Safely writing code that isn't thread-safe

An under-appreciated Rust feature

# **Tutorial: (2) Scoped Threads**

Or: do we really need Arc?

https://play.rust-lang.org/?version=stable&mode=release&edition=2021&gist=79d8080575170cb2972a25bae9b69ef5

- Compiler not smart enough! (w.r.t scope of counter vs threads)
- We can help it by using an automatically scoped thread (must join before the end of the scope) - think std::jthread from C++

```
fn main() {
    let counter = Mutex::new(0);
    thread::scope(|s| {
        s.spawn(|| *counter.lock().unwrap() += 1);
        s.spawn(|| *counter.lock().unwrap() += 1);
8
    });
    println!("{}", *counter.lock().unwrap());
```

# **Tutorial: (3) Interior Mutability**

### 3. Interior Mutability

https://play.rust-lang.org/?version=stable&mode=release&edition=2021&gist=79d8080575170cb2972a25bae9b69ef5

- Doesn't it seem crazy that we can do this?
- Two areas of code can <u>mutably</u> change a non-mutable reference!

```
fn main()
 let counter = Mutex::new(0);
 thread::scope(|s|
      s.spawn(|| *counter.lock().unwrap()
      s.spawn(|| *counter.lock().unwrap()
 });
  println!("{}", *counter.lock().unwrap());
```

### What is a Mutex type in Rust?

- An actual mutex to protect in data
- 2. And an UnsafeCell that contains our data

```
pub struct Mutex<T: ?Sized> {
    inner: sys::Mutex,
    poison: poison::Flag,
    data: UnsafeCell<T>,
}
```

```
pub struct UnsafeCell<T: ?Sized> {
    value: T,
}
```

### What is a Mutex type in Rust?

- When we deference the mutex after locking...
- The Mutex library code uses unsafe Rust to get a mutable reference from the UnsafeCell!
  - o It's OK: we know that a locked mutex allows only 1 thread to access the data
  - But the compiler isn't a human and we need to disable some checks with unsafe

#### Unsafe...

- As long as the library is written correctly, we can still have guarantees in our program
- Task: can we use atomics to avoid Mutexes?



#### 3.1 Atomics version

https://play.rust-lang.org/?version=stable&mode=release&edition=2021&gist=c478db74d562120e42f8e5b04138ccff

- Almost identical to C++!
- Quick aside: why is Relaxed OK here? [np]

#### **Atomics**

```
use std::sync::atomic::{AtomicI32, Ordering};
   use std::thread;
   fn main() {
        let counter = AtomicI32::new(0);
        thread::scope(|s| {
            s.spawn(||
                counter.fetch_add(1, Ordering::Relaxed);
10
            });
            s.spawn(|
                counter.fetch_add(1, Ordering::Relaxed);
13
            });
        });
14
15
        println!("{}", counter.load(Ordering::Relaxed));
16
```

Rust pretty blatantly just inherits the memory model for atomics from C++20.

# **Tutorial: (4) Static Items**

### https://play.rust-lang.org/?version=stable&mode=release&edition=2021&gist=5ed58ec8a4b5c11cbc6b3d7e5de65fcb

#### 4 Static version

 No need scoped threads now

Compiler can prove that
 COUNTER will last for the whole program!

```
puse std::thread;
use std::sync::atomic::{AtomicI32, Ordering};

static COUNTER: AtomicI32 = AtomicI32::new(0);

fn main() {
    let t0 = thread::spawn(|| { COUNTER.fetch_add(1, Ordering::Relaxed); });
    let t1 = thread::spawn(|| { COUNTER.fetch_add(1, Ordering::Relaxed); });

t0.join().unwrap();
t1.join().unwrap();
println!("{}", COUNTER.load(Ordering::Relaxed));
}
```

# Tutorial: (5) Rayon

https://play.rust-lang.org/?version=stable&mod e=release&edition=2021&gist=82d3cb263bd0b a2e59f5db8608bac675

### 5. Rayon

Imagine this trivially parallelizable program

```
fn magic_sum(from: u128, to: u128) -> u128 {
    (from..to).filter(|i| i % 7 == i % 5).sum()
fn main() {
   let (from, to) = {
        // Comment out the line below if you are using the Rust Playground
        let mut args = std::env::args();
        // Use the line below instead if you are using the Rust Playground
        args.next(); // skip argv[0]
        (args.next().unwrap(), args.next().unwrap())
   };
    println!("{}", magic_sum(from.parse().unwrap(), to.parse().unwrap()));
```

https://play.rust-lang.org/?version=stable&mod e=release&edition=2021&gist=54a7f51d6e474 074db215fef69f44b60

### 5. Rayon

Safe parallelism in Rust is as simple as this!

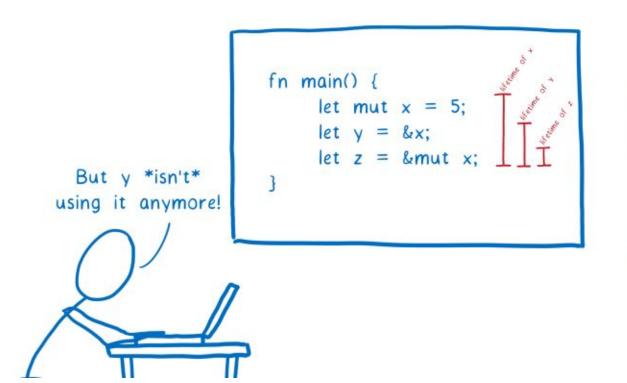
```
use rayon::prelude::*;
   fn magic_sum(from: u128, to: u128) -> u128 {
        (from..to].into_par_iter().filter(|i| i % 7 == i % 5).sum()
   fn main() {
       let (from, to) = {
            // Comment out the line below if you are using the Rust Playground
10
            let mut args = ["", "0", "100000000"].iter();
13
           args.next(); // skip argv[0]
14
            (args.next().unwrap(), args.next().unwrap())
15
       };
16
       println!("{}", magic_sum(from.parse().unwrap(), to.parse().unwrap()));
```

### Summary

- Rust is a language with many useful guarantees
  - No UB, memory safety, type safety...
- More rules to be followed if you want this reward
  - Ownership
  - Borrowing
  - Lifetimes, etc
- Really see the use when doing concurrent work!

# Extra Quiz

# Spot the bug here !!!



Wait, you can't compile that!

z needs to be the only one borrowing x's value, but y is already using it.

### Interesting complexities with scopes

Problem?

https://play.rust-lang.org/?version=stable&mode=debug &edition=2021&gist=fe901c4e48353a6f2ca450b9fe1042 7a

```
fn main() {
       let arr = vec![String::new(); 10];
       thread::scope(|s| {
6
           for i in 0..10 {
             s.spawn(|| println!("{}", &arr[i]));
       });
```

### Interesting complexities with scopes

Problem?

https://play.rust-lang.org/?version=stable&mode=debug &edition=2021&gist=fe901c4e48353a6f2ca450b9fe1042 7a

- The thread needs the value i but it might be dropped before the thread(s) start! (end of for loop)
  - Cannot borrow via reference! Must move into thread!

```
4 fn main() {
5    let arr = vec![String::new(); 10];
6    thread::scope(|s| {
7         for i in 0..10 {
8              s.spawn(|| println!("{}", &arr[i]));
9         }
10    });
11 }
```

### Interesting complexities with scopes

- We can **borrow** one thing (**arr**)
- https://play.rust-lang.org/?version=stable&mode=debug&edition =2021&gist=0057fccdfb504a00f183afbe2b91c435

- And **move** the other (i)!
- move statement is very coarse, so we have to handle it ourselves..

```
3 fn main() {
      let arr = vec![String::new(); 10];
      thread::scope(|s| {
           let borrowed_arr = &arr; // We can borrow this, and move i
           for i in 0..10 {
               s.spawn(move || println!("{}", borrowed_arr[i]));
```