

## Comparative Programming Languages Prof. Alex Ufkes

Topic 10: Rust intro, typing, and control flow



## Notice!

# Obligatory copyright notice in the age of digital delivery and online classrooms:

The copyright to this original work is held by Alex Ufkes. Students registered in course C/CPS 506 can use this material for the purposes of this course but no other use is permitted, and there can be no sale or transfer or use of the work for any other purpose without explicit permission of Alex Ufkes.

## **Course Administration**



- Getting closer! Rust is our last language.
- Don't forget about the assignments!

# Moving on...

# ...to imperative.

Rust is an imperative language. However, we'll see many cool features that remind us of the functional languages we've seen.



## **Rust History**



- Grew out of a personal project by Mozilla employee Graydon Hoare in 2006
- Mozilla began sponsoring the project in 2009
- Officially announced in 2010
- Rust compiler successfully tested in 2011
- Pre-alpha version released in 2012
- Rust 1.0, the first stable release, arrived on May 15, 2015
- Youngest language we've seen so far
- Open source



#### **Systems Programming Language:**

- In contrast with *application* programming languages.
- System software includes things like operating systems, utility software, device drivers, compilers, linkers, etc.
- System languages tend to feature more direct access to physical hardware of a given machine.



#### Syntax:

- Similar to C/C++
- Blocks of code delimited by { }
- Familiar control structures supported (if, else, while, for, etc.)
- Supports pattern matching! (match)
- Need not use return, last expression creates return value
- Functions largely composed of expressions



#### Memory Safety:

- Rust is designed to be *memory safe*
- Null or dangling pointers are not permitted.

### "Null or dangling pointers are not permitted"

```
#include <stdio.h>
#include <stdlib.h>
int main(void)
£
   int *x = NULL;
    *x = 77;
   int *y = (int*) malloc(4 * sizeof(int));
   y[4] = 7; 🔫
    printf("%d \n", *x);
    printf("%d \n", y[4]);
    system("pause");
```

Alex Utkes, 2020, 202

- In C, we're allowed to <u>try</u> and access any memory we want.
- This code compiles!
  - It produces a run-time error when we try and index into pointer x.
- Overrunning array bounds does not
- necessarily give a run time error!
- Very unsafe use of memory.

### "Null or dangling pointers are not permitted"



#### Java is safer:

- This code *compiles*, but **always** throws an exception when we access outside array bounds.
- C/C++ only errors if going out of bounds accesses memory that your program doesn't have write permission for.
- Java still allows dangling references.
  - nums2 can be declared without instantiating its object.

© Alex Ufkes, 2020, 2022



#### Memory Safety:

- Rust is designed to be *memory safe*
- Null or dangling pointers are not permitted.
- What about linked lists? Null pointers are useful.
- Rust defines an *option* type, which can be used to test if a pointer has *Some* value or *None* 
  - What does this remind you of?



#### **Memory Management:**

- Rust does not do garbage collection
- Resource acquisition is initialization
- RAII Originated in C++
- Constructor used to acquire and initialize objects
- Resource *deallocation* is done by the destructor.
- No valid reference to object == no object.
- Not so in Java! Up to garbage collector.



#### **Types and Polymorphism:**

- Type system supports mechanism called "traits"
- Directly inspired by Haskell's type classes
- Supports type inference for variables declared with let keyword.
- Compile error if inference fails.
- Keyword **mut** for mutable variables.



#### **Pattern Matching:**

- Rust supports pattern matching!
- Pattern matching is considered a sticking point for people learning Rust.
- We already have experience with it



#### Strongly, statically typed

- Strong typing means limited implicit type conversions at compile time.
- C is happy to convert between numeric types without issue. Perhaps a compile warning in C++.
- Java raises compile error if there's a loss of precision (double to float for example).



#### Output processo

Show output from: Build

| 🖺 | 🖆 🛓 | 🖉 | 🖏

1>----- Build started: Project: Tester, Configuration: Debug Win32 -----

1> Source.cpp

1>d:\googledrive\teaching - humber\atmn 253\visual studio projects\tester\source.cpp(7): warning C4244: 'initializing' : conversion from 'double' to 'int', possible loss of data
1> Tester.vcxproj -> D:\GoogleDrive\Teaching - Humber\ATMN 253\Visual Studio Projects\Tester\Debug\Tester.exe

======= Build: 1 succeeded, 0 failed, 0 up-to-date, 0 skipped =========

p(7): warning C4244: 'initializing' : conversion from 'double' to 'int', possible loss of data

cts\Tescer \pepug\rescer.exe







#### No "Undefined Behavior"

- Null pointer dereferencing
  - $\circ$   $\;$  Attempt to dereference address 0  $\;$



🔳 quincy

0



#### No "Undefined Behavior"

- Null pointer dereferencing
  - $\circ$   $\;$  Attempt to dereference address 0  $\;$
- Use of variable before it's initialized
  - $\circ$   $\,$  In C, we get whatever was in
    - memory before that.
  - $\circ$  Only globals auto-initialize to 0



#### **No "Undefined Behavior"**

- Null pointer dereferencing
  - $\circ$   $\;$  Attempt to dereference address 0  $\;$
- Use of variable before it's initialized
  - In C, we get whatever was in memory before that.
  - $\circ$  Only globals auto-initialize to 0
- Array index out of bounds
  - May or may not cause runtime error (in C), depends who owns memory

<pre>#include <stdio.h> #include <stdlib.h></stdlib.h></stdio.h></pre>	Ν
int x;	
int main(void) {	
int y[5];	
y[6000] = 8;	
}	

#### **Microsoft Visual Studio**







#### **No "Undefined Behavior"**

• Signed integer overflow & optimization

X+1 > X

- If overflow is undefined, compiler can just optimize this to simply **true**.
- Dangerous if X can overflow!
- Forcing compiler to consider overflow means we lose certain optimizations.

## **Rust Non-Goals**



- We do not employ any particularly cutting-edge technologies.
   Old, established techniques are better.
- We do not prize expressiveness, minimalism or elegance above other goals. These are desirable but subordinate goals.
- We do not intend to cover the complete feature-set of C++, or any other language. Rust should provide majority-case features.
- We do not intend to be 100% static, 100% safe, 100% reflective, or too dogmatic in any other sense. Trade-offs exist.
- We do not demand that Rust run on "every possible platform". It must eventually work without unnecessary compromises on widely-used hardware and software platforms.



## **Installing Rust**

#### https://www.rust-lang.org/en-US/index.html

Install



Documentation

Community

Contribute

**Rust** is a systems programming language that runs blazingly fast, prevents segfaults, and guarantees thread safety.



See who's using Rust, and read more about Rust in production.

© Alex Ufkes, 2020, 2022 Featuring





## **Installing Rust**



To install Rust, download and run rustup-init.exe

then follow the onscreen instructions.

If you're a Windows Subsystem for Linux user run the following in your terminal, then follow the onscreen instructions to install Rust.

curl https://sh.rustup.rs -sSf | sh

Rust 1.26.0

May 10, 2018

## **Editing Rust Code**

#### Any text editor will do, but I like VSCode:

💐 m	ain.rs - RustCode - Visual Studio Code		— I	×
<u>F</u> ile <u>E</u>	dit <u>S</u> election <u>V</u> iew <u>Go</u> <u>D</u> ebug <u>T</u> asks <u>H</u> elp			
ര	EXPLORER	🖲 main.rs 🛛 🗙	ශ්	
	▲ OPEN EDITORS	1 fn main() {		
Ω	main.rs	<pre>2 println!("Hello, world!");</pre>		
~	▲ RUSTCODE	3 }		
Ŷ	® main.rs	5		
0				
æ				
S				
		Visual Studio Code:		
		Supports Rust syntax colori	ng	
			0	
		Useful for other languages		

© Alex Ufkes, 2020, 2022

## **Compiling Rust Code**

#### Command Line - rustc





#### Much of the syntax is reminiscent of C/C++

# fn main() { println!("Hello, world!"); }

Like C, C++, Java, Haskell, and many others, main() defines the entry point for executing a Rust program.

# fn main() { println!("Hello, world!"); }

#### println vs println!

- The ! indicates we're calling a macro.
- A standard function call doesn't include !

## Variables

- By default, Rust variables are immutable
- Once initialized, can't change.
- Like final or const in other languages
- Declare using let keyword:

© Alex Ufkes, 2020, 2022

## as a C/C++ style placeholder

## Variables



## **Mutable Variables**

#### Use **mut** keyword:

Command Prompt

- fn main() {
   let mut x = 7;
   x = 5;
   println!("value: {}", x);<sup>2</sup>
  - = note: #[warn(unused\_assignments)] on by default

C:\\_RustCode>rustc main.rs

let mut x = 7;

warning: value assigned to `x` is never read

- We get a warning, and it's sensible.
- We change the value of x before the initial value is ever read.
- Pointless.

C:\\_RustCode>main value: 5

--> main.rs:2:9

## **Constant/Global Variables**

#### Rust still has them:

- Use const instead of let
- Always immutable
- Can be declared in global scope, unlike let
- Must indicate data type (u32)
- More on types coming up.
# **Constant/Global Variables**

Can be declared in global scope, unlike let



Variables with the same name?

In Java, variables can have the same name so long as their scope does not overlap:

Variables with the same name?

C++ is less strict. Scopes can overlap, but they can't be identical:

## Variables with the same name?



## Variables with the same name?



- What we're doing here is like re-binding in Haskell or Elixir.
- This doesn't work with mutable variables.
- Think of this mathematically We're simply saying let x = something else.

# Shadowing VS mut

Why not just use shadowing? Why do we need **mut**?



# Shadowing VS mut

Why not just use shadowing? Why do we need **mut**?



# Shadowing VS mut

Why not just use shadowing? Why do we need **mut**?

- With **mut**, we're *mutating* a variable in memory.
- Storing a different value in the same variable.
- The name still refers to the same place, thus the type must stay the same.
- With shadowing, we're getting a new variable in memory each time.
- We're changing what a given name is referring to.
- We're not changing the existing value.

# Data Types

Two subsets: Scalar and Compound

**Reminder:** Rust is statically typed. Must know all variable types at compile time.

#### Scalar types represent a single value:

• Rust has four: integers, floating-point, Booleans, characters.

## **Compound types group multiple values:**

• Two primitive compound types: tuples and arrays.

## Scalar Types: Integers

Length	Signed	Unsigned
8-bit	i8	u8
16-bit	i16	u16
32-bit	i32	u32
64-bit	i64	u64
arch	isize	usize

•	Signed integers are stored
	using 2s comp

- Arch will be 32 bits on a 32 bit system, 64 bits on a 64 bit system.
- When not specified, Rust defaults to i32

# Specify Type?

Rust has type inference, but we can be explicit:

🐵 main.rs	×	
1	fn main() {	Command Prompt
-		
2	let x: u8 = 3;	C:\_RustCode>rustc main.rs
3	let y: i64 = 5;	C:\_RustCode>main
4	<pre>let z: isize = 999;</pre>	x: 3
5	<pre>println!("x: {}", x);</pre>	z: 999
6	<pre>println!("y: {}", y);</pre>	C:\ RustCode>
7	<pre>println!("z: {}", z);</pre>	
8	}	
9		

## **Integer Literals**

In addition to just writing the value...

Number literals	Example
Decimal	98_222
Hex	Oxff
Octal	0077
Binary	0b1111_0000
Byte (u8 only)	b'A'

Bytes can be character literals

© Alex Ufkes, 2020, 2022

## Notice the \_

- This is a handy visual sugar
- Hard to count the zeroes in 1000000000.
   What number is this?
- Easy to see 1\_000\_000\_000 is one billion.

## Scalar Types: Floating Point

- Two kinds 32 and 64 bit (float and double, single and double precision)
- Represented using standard IEEE-754

 $\bigcirc$ 

	Command Prompt
🐵 main.rs 🗙	
1	C:\_RustCode>rustc main.rs
Default let x: f32 = 1.0/3.0;	C·\ RustCode\main
<pre>let y: f64 = 1.0/3.0;</pre>	x: 0.33333334
<pre>4 println!("x: {}", x);</pre>	y: 0.33333333333333333
5 println!("y: {}", y);	
6 <b>}</b>	C:\_RustCode>_
7 Hex Ufkes, 2020, 2022	4

## **Numeric Operations**

🖲 main.rs	×	
1	<pre>fn main() {</pre>	Command Prompt – 🗆 × X: 0.333333334
2	let r1 = 2 + 3	y: 10000
3	let r2 = 3/4;	C:\_RustCode>rustc main.rs
4	let r3 = 2 % 3	> main.rs:4:16
5	println!("r1:	4 let r3 = 2 % 3.0;
6	println!("r2:	<pre>^ no implementation for `{intege p} % {float}`</pre>
7	println!("r3:	
8	}	<pre>= help: the trait `std::ops::Rem&lt;{float}&gt;` is not implemented for `{integer}`</pre>
9		

## **Numeric Operations**

📾 main rc	~	
		Command Prompt
1	<pre>fn main() {</pre>	C:\ RustCode>rustc main.rs
2	let r1 = 2 + 3 * 6;	
3	let r2 = 3/4;	C:\_RustCode>main r1: 20
4	let r3 = 2 % 3;	r2: 0
5	println!("r1: {}", r1);	r3: 2
6	println!("r2: {}", r2);	C:\_RustCode>
7	println!("r3: {}", r3);	
8	}	
9		





© Alex Ufkes, 2020, 2022

53



Division may truncate, good reason to avoid implicit conversion...

```
Command Prompt
                                                                              \times
C:\_RustCode>rustc main.rs
 error [E0277]: cannot add a float to an integer --> main.rs:2:16
        let r1 = 3 + 4.0;
                   ^ no implementation for `{integer} + {float}`
  = help: the trait `std::ops::Add<{floatl>` is not implemented for `{integer}`
                                      main.rs
                                               ×
 error: aborting due to previous erro
                                              fn main() {
For more information about this erro
                                           2
                                                     let r1 = 3 + 4.0;
                                                     println!("r1: {}", r1);
                                           3
C:\ RustCode>_
Alex Lifkes 2020 2022
```

55

# Why?!

- Adding **float** to **int** means converting the integer to a floating-point type, then adding.
- CPU doesn't add different types.
- Float and int arithmetic is done using different instructions, in different locations on CPU.
- It's possible to introduce errors in precision!
- An integer in binary is *exactly precise*.
- The same value represented as a floating point may lose significant digits.
- Most languages don't even warn about this Rust doesn't allow it at all.

```
public class MethodTester
```

```
public static void main(String[] args)
```

```
int a = 2111111111;
System.out.println(a);
```

```
float b = a;
a = (int) b;
```

System.out.println(a);

4 BlueJ: Terminal Window - HelloWorld

Options

2111111111 2111111168



© Alex Ufkes, 2020, 2022

Can only ent

## Scalar Types: Boolean

## true, false. Easy:



## Scalar Types: Characters

## Rust supports Unicode:



# Compound Types: Tuples



## De-structuring!

🔊 main rc	~	
		Command Prompt
1	<pre>fn main()</pre>	$(\cdot)$ BustCada busts main no
2	{	C: \_RustCode>rustC main.rs
3	<pre>let tup = (42, 3.141592, '!');</pre>	C:\_RustCode>main
4	let $(x, y, z) = tup;$	42, 3.141592, !
5		C:\_RustCode>_
6	println!("{}, {}, {}", x, y, z);	
7	}	
8		

Can also access directly:



## Out of bounds:



## Can we fool it?





## Out of bounds:



# **Array of Tuples**

## Same rules as Haskell:



# **Array of Tuples**

Same rules as Haskell: Tuple types must be the same



## Types & Literals: Summary

#### **4 Scalar types:**

Integer - u8, u16, u32, u64, usize, i8, i16, <u>i32</u>, i64, isize
Floating Point - f32, <u>f64</u>
Boolean - bool (true, false)
Character - Unicode: 'Z', 'a', '&', '\u{00C5}', etc

## **2** Compound types:

Tuple – heterogeneous Arrays – homogeneous

Rust supports other data structures such as strings and vectors. These are not base types, but very useful.

# Strings



# FUNCTIONS

## **Functions**

## We've seen main()



- Returns nothing, accepts no arguments.
- Convention for naming functions is snake\_case.
- Words separated by underscores.
# **Functions**



### **Parameters**

```
main.rs
       ×
       fn main()
           print_val (5);
           print_two_vals (5, 3.14);
       }
      fn print_val (n: i32)
           println!("{}", n);
  10
      }
                                                  5
      fn print_two_vals (n1: i32, n2: f64)
  12
  13
       ł
           println!("{}, {}", n1, n2);
  15
  Ufkes, 2020, 2022
```

### identifier: type

- Parameters separated by commas.
- Indicating type is *mandatory*
- Nothing too unusual here

#### Command Prompt

- C:\\_RustCode>rustc main.rs
- C:\\_RustCode>main
- 5, 3.14
- C:\\_RustCode>\_

# **Careful Now...**

```
8 main.rs
      ×
                                                   Command Prompt
                                                                                           \times
      fn main()
                                                  C:\_RustCode>rustc main.rs
                                                   error[E0308]: mismatched types
           print_val (5);
                                                   --> main.rs:4:24
           print two vals (5, 3);
                                                           print_two_vals (5, 3);
                                                                               ^ expected f64,
      fn print_val (n: i32)
                                                  found integral variable
           println!("{}", n);
                                                    = note: expected type `f64`
                                                                found type `{integer}`
  10
                                                  error: aborting due to previous error
      fn print_two_vals (n1: i32, n2: f64)
  12
                                                  For more information about this error, try
          println!("{}, {}", n1, n2);
                                                   rustc --explain E0308`.
    Jfkes, 2020, 2022
```

Rust is *primarily* expression based, but still has statements.

#### Two types of statements:

- Declaration statements return nothing
- Expression statements return empty tuple ()

let x = 6; // This is a declaration statement

The above does not return a value. We can't do the following:

let y = (let x = 6);

Rust is *primarily* expression based, but still has statements.

#### Two types of statements:

- Declaration statements return nothing
- Expression statements return empty tuple ()
- 5 + 2; // This is an expression statement

The above expression is evaluated, but the result is ignored (not saved).

- 5 + 2 is an **expression**. It evaluates to 7.
- y = 5+2; is an **expression statement**. It returns (), but the result of the nested expression 5+2 is saved to y

let 
$$y = (let x = 6);$$



```
C:\_RustCode>rustc main.rs
error: expected expression, found statement (`let`)
       let x = (let y = 6);
                 ^^^ expected expression
  = note: variable declaration using `let` is a statem
error: aborting due to previous error
```





- Variable **y** gets re-assigned.
- The *expression statement* (y=8) returns an empty tuple in Rust.
- Can't assign an empty tuple to a variable declared to hold **i32**!



### Here:

- Value of x will be 3
- Value of y will be () empty tuple

x + 6



// This is an expression

// This is an expression statement
// containing an expression

Declaration statement

Creating a new scope block?

We can do this in Java and C/C++, though again it isn't so common:



#### Scope blocks like this are expressions in Rust:



#### There's a few things going on here:

- We're trying to bind a value to y.
- Thus, the block { } should evaluate to something.
- Notice there's no semicolon after z + 1
- z + 1 is an expression.
- Adding a semi-colon would make it an *expression statement*.
- Thus, the block { } would return ( ).
- Probably not what we want.

Scope blocks like this are expressions in Rust:

expression!



#### Scope blocks like this are expressions in Rust:



# **Return Value**

Think of functions the same way.

The last line should be an expression – no semi-colon.



© Alex Ufkes, 2020, 2022

-> type

DuctCodes

- Explicitly indicate return type
- Result of expression gets returned

C:4. (	Command Prompt —	
c:`	\_RustCode>rustc main.rs	
C:` 13	\_RustCode>main	
	87	

# **Return Value**

Add semicolon? It becomes expression statement, returns (), type mismatch:



# Control Flow



TRAFFIC

# if/else

🖲 main.rs	•		
1	fn main()	• As with other imperative	languages, the else is optional.
2	{	• Recall that this is not the	case with Haskell!
3	<pre>let num = 3;</pre>	• We were required to have	e a complete if-then-else
4		·	
5	if num > 5 {		Command Prompt
6	printin!("Gr	eater than 5!");	
7	}		C:\_RustCode>rustc main.rs
8	else {		
9	println!("No	ot that thing I just said");	C:\_RustCode>main
10	}		Not that thing I just said
11			
12	}		C:\_RustCode>
13			

# **Boolean Conditions?**

#### Mandatory.

#### In C/C++ (and Elixir, with caveats):

- Non-zero values are "truthy".
- Only 0/nil considered false.

```
if (3.141592)
    cout << "Valid!" << endl;</pre>
```

### In Java (and Haskell, Rust):

• Conditions must be Boolean

Converting non-Boolean to Boolean requires implicit conversion, which, as we've seen, Rust does not do.

® main.rs	×		
1	<pre>fn main()</pre>		
2 3 1 4 5	<pre>{     let temp = 33;     if temp &lt; 0 {</pre>	<ul> <li>As</li> <li>We one</li> </ul>	we'd expect. e use { } even though there's only e statement per branch
6 7 8 9	<pre>println!("Frozen"); } else if temp &lt; 100 {     println!("Liquid");</pre>	<ul> <li>Thi</li> <li>Wł</li> <li>wh</li> </ul>	s is required. ny? Rust treats these as blocks ose last line can be an expression.
10	}		
11	else {		
12	<pre>println!("Boiling");</pre>		
13	}		
14	}		
© Ålex	Ufkes, 2020, 2022		

```
main.rs
       ×
      fn main()
       {
           let temp = 33;
           let state = if temp < 0 { "Frozen" }</pre>
                         else if temp < 100 { "Liquid" }</pre>
                         else { "Boiling" };
           println!("Water is {}!", state);
  10 }
```

```
*
fn main()
{
    let temp = 33;
    let state = if temp < 0 { "Frozen" }
        else if temp < 100 { "Liquid" }
        else { "Boiling" };
    println!("Water is {}!", state);
}</pre>
```

- let state = {...}; is a statement
- {...} is an expression that will evaluate to a string. if == expression!
- "Frozen", "Liquid", or "Boiling"
- Each option is in a scope block { }
- The value of a scope block is the last expression
- Leaving the ; off makes these strings expressions.

```
Command Prompt
×
fn main()
                                                   C:\_RustCode>rustc main.rs
    let temp = 33;
                                                   C:\_RustCode>main
                                                   Water is Liquid!
    let state = if temp < 0 { "Frozen" }</pre>
                else if temp < 100 { "Liquid" }</pre>
                                                   C:\_RustCode>
                else { "Boiling" };
    println!("Water is {}!", state);
```

# **Problem?**



**Remember:** Strong, static typing. No implicit conversion!



# Looping



# Conditional Looping: while



# **Conditional Looping: for**

#### Just like a for loop in Python:

<pre>x fn main() {     let nums = [1, 2, 3, 4, 5, 6]</pre>	<ul> <li>Invoke iter() method of array nu</li> <li>elem takes the value of each element in the array.</li> <li>Safe! Never go out of bounds.</li> </ul>	Jms
<pre>for elem in nums.iter() {     print!("{} ", elem); } println!(); }</pre>	<pre> Command Prompt - □ C:\_RustCode&gt;rustc main.rs C:\_RustCode&gt;main 1 2 3 4 5 6 7 8 C:\_RustCode&gt;</pre>	) ×
© Alex Ufkes, 2020, 2022		100

# **Conditional Looping: for**

#### Use .. to create a range

```
×
fn main()
{
    let nums = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10];
    for i in (0..10).rev()
                                          Create a Range containing 0 to 9
                                        •
                                           Top of range not included!
        print!("{} ", nums[i]);
                                           Just like range() in Python
                                        ullet
    print!("\nLIFTOFF!\n");
}
```

# A loop is a loop is a loop



# Wait, what?

```
fn main()
    let nums = [1, 2, 3, 4, 5, 6, 7,
    let mut i = 9;
    loop
        if i < 0 { break; }</pre>
        print!("{} ", nums[i]);
        i -= 1;
                                       Cit.
    print!("\nLIFTOFF!\n");
```

© Alex Ufkes, 2020, 2022

• We didn't specify the type of **i**, but shouldn't it default to **i32**?

- Rust infers type, **i32** should be default.
- HOWEVER!
- Rust doesn't allow signed integers to be used as array indexes!
- It inferred the type as unsigned! Thus
- checking less than zero is pointless.

```
C:\_RustCode>rustc main.rs
warning: comparison is useless due to type limits
--> main.rs:8:12
```

### Rust doesn't allow signed integers to be used as array indexes!



### Need to adjust our logic a bit...



# **Fantastic Rust Reference:**

# https://doc.rust-lang.org/book/second-edition/

