CMSC330: Rust Basics

Chris Kauffman

Last Updated: Tue Dec 5 12:10:09 AM EST 2023

Logistics

Reading: The Rust Programming Language

- Official tutorial guide from Rust foundation
- Chapters 1-10 should be sufficient for the course

Assignments

- Project 8: Rust Spell Checker Released Due Mon 11-Dec
- Quiz 4: Fri 08-Dec during Discussion

Goals

- ☑ Introduction to Rust
- □ Language Features / Relations
- Memory Ownership
- □ Datatypes and Traits
- Extras

Announcements: None

A Short History of Rust



Graydon Hoare according to Crunchbase



Rust Logo



- Started ~2006 as a side project by Graydon Hoare while working at Mozilla (makes of Firefox and other fine tools)
- Started after a software failure incapacitated an elevator at Hoare's apartment requiring him to climb 21 flights of stairs and inspiring a desire for a "safe" programming language
- Compiler originally written in OCaml, became self-hosting in 2010
- After Mozilla divested developers in 2020, Rust Foundation was created to manage language and community
- Stack Overflow Developer Survey named Rust the "most loved programming language" every year from 2016 to 2023 (Wikip)
- ► Hoare named Rust "...after a group of remarkably hardy fungi" (MIT Tech Review)

Exercise: Collatz Computation An Introductory Example

- collatz.rs prompts for an integer and computes the Collatz Sequence starting there
- ➤ The current number is updated to the next in the sequence via if cur is EVEN cur=cur/2; else cur=cur*3+1
- ► This process is repeated until it converges to 1 (mysteriously) or the maximum iteration count is reached
- The code demonstrates a variety of Python features and makes for a great crash course intro
- With a neighbor, study this code and identify the features you should look for in every programming language

Exercise: Collatz Computation An Introductory Example

```
1 // collatz.rs:
   use std::io:
   use std::str::FromStr;
   const VERBOSE : bool = true:
   fn collatz (start: i32, maxsteps: i32)
               -> (i32,i32)
10
     let mut cur = start;
     let mut step = 0:
11
     if VERBOSE {
       println!("start: {start}");
       println!("Step Current"):
       println!("{step:3} {cur:5}"):
16
17
     while cur != 1 && step < maxsteps {
       step += 1:
18
       if cur % 2 == 0{
         cur = cur / 2:
21
       elsef
         cur = cur*3 + 1;
24
25
       if VERBOSE {
         println!("{step:3} {cur:5}");
27
     (cur, step)
                            // return val
29
     // return (cur.step): // alt
31 }
```

```
32 fn main() { // entry point
     println!("Collatz start val: ");
33
     let mut instr = String::new();
34
     match io::stdin().read line(&mut instr) {
35
36
       Err(why) => panic!("Input failed: {}",why),
       0k(_) => {}
                                    // freak on error
37
     }:
                                    // proceed on an ok
38
     instr.pop():
                                    // remove trailing \n
39
     let start = match i32::from_str(instr.as_str()) {
40
       Err(whv) => panic!("Bad int 'finstr}': f}".whv).
41
       Ok(anint) => anint
                                    // freak on error
42
43
     1:
                                    // o/w return the int
     let (last,steps) = collatz(start, 500);
44
     println!("Reached {last} after {steps} iters");
46
```

Look for... Comments, Statements/Expressions, Variable Types, Assignment, Basic Input/Output, Function Declarations, Conditionals, Iteration, Aggregate Data, Library System

Answers: Collatz Computation An Introductory Example

Comments: // comment to end of line println!("Hi"); statement lines end with semicolons ✓ Variable Types: i32 integer, boolean, some types mentions others inferred \boxtimes Assignment: via let x = expr; or let mut x = expr; or x = 3*x+1; Basic Input/Output: println!() and ... oh boy... io::stdin.read_line() □ Function Declarations: fn funcname(param1: type1, param2: type2) -> RetType { Conditionals (if-else): if cond { ...} else {...}; also match{ } conditions ☐ Iteration (loops): clearly while cond {...}, also for iter {} Aggregate data (arrays, records, objects, etc): (rust, has, tuples) and Variant(types) ⊟ Library System: use std::io; is like import std.io

Compile and Run

```
>> rustc collatz.rs
>> file collata
collatz: ELF 64-bit LSB pie executable, x86-64,
version 1 (SYSV), dynamically linked, ...
>> collatz
Collatz start val:
start: 10
Step Current
      10
      5
     16
Reached 1 after 6 iters
>> ./collatz
Collatz start val:
apple
thread 'main' panicked at collatz.rs:66:17:
Bad int 'apple': invalid digit found in string
note: run with 'RUST BACKTRACE=1' environment
variable to display a backtrace
```

- rustc compiles code with a main() method to executables named after the file
- collatz fails on bad input like apple via the panic!() macro though code running an elevator may wish to take a different tack...

Rust's Prime Directive: Be Safe!

- Avoid memory bugs at all costs
- Provide mechanisms for error handling but force programs to contend with errors
- ► If failing, fail predictably

Memory Safe Languages

- ➤ You may be told that Rust is a memory safe language; this is true and often said to contrast it to C/C++ which may segfault as they give the power of unrestricted pointer operations
- ▶ You might respond to the speaker that there are a few other memory safe languages such as Java, Python, OCaml, Scheme, Racket, Clojure, Shell Script, Haskell, Fortran, Javascript, etc. and basically all other languages that don't have unrestricted pointer operations
- You might indicate to the speaker that perhaps they meant Rust is memory safe without using a Garbage Collector which is unusual
- You might end by mentioning that if being memory safe is good and lots of PLs have that quality, having a Garbage Collector might also be good and make a language more usable

Borrowing Ideas from C/C++

- ▶ **No Garbage Collector**: GC costs at runtime so avoid it
 - C/C++ handle this with manual management, e.g. malloc()/free() or new / delete
 - Rust follows a different model
- ➤ Aim for "zero cost abstractions": high-level looking code compiles down to very efficient machine instructions, no runtime penalties
- Use of the <T> syntax for generic / templated code
- Similar syntax for namespace navigation
 some::package::file::run_function(a,b)
- ► Shared with C++: allow operator overloading so that a+b can be overloaded for any types
- ► Shared with C++: variety of ways to pass parameter, as is, as refs, as mutable refs, etc.
- Shared with C++: add a LOT of stuff to the language; small, tightly integrated features are less fun than playing with everything and the kitchen sink

Borrowing Ideas from OCaml

- Default to immutable data as it is more easily shared and enables concurrency more readily
- Explicitly label data as mut to indicate mutability which comes with benefits and costs
- Strongly typed
- ➤ Some degree of type inference supported but certain places, particularly function signatures, require explicit typing
- Some degree of polymorphism supported though the model is closer to C++ Templates / Java Generics
- ► Support Pattern match-ing with rich variant/algebraic types, called enum in Rust; use these in the standard library

Borrowing Ideas from Python

- Provide a featurful array-like data structure (List in Python, Vector in Rust)
- ► Favor iterators in for loops and ensure that most standard container types support them for ease
- Like Python (and Java), makes use of code annotations such
 as #[test] to denote a function is a test case or
 #[derive(Debug)] to automatically derive some
 functionality for a data type associated with debugging

Borrowing Ideas from Java

- Uses method dispatch:
- Rust is NOT object-oriented, but then again Gosling would have removed class inheritance from Java given a second chance
- Java also features Interfaces which are a collection of methods implemented by a class
- Rust follows this model: data types impl collections of methods referred to as Traits allowing the data type to be used any place something with the given trait is present

Borrowing Ideas from Lisp

- ► Provide for powerful macro creation that enables manipulation of the syntax tree during compilation
- Macros like println!(...) aren't functions, rather they generate code in place which can make things more efficient and allow for compile time safety with convenience such as in println!("x: {x}")

Cautions and Disclosures

Cautions

- ▶ Rust is 17 years old with wide public attention for <10 years
- ▶ During that time it has undergone radical changes with much breakage to older code, a trend that is likely to continue
- ▶ It combines many features from many other languages
- ► It's only feature of real note is its memory model: no Garbage Collector, manage memory at compile time as possible

Disclaimer

I don't know Rust particularly well but I don't like what I see. In the name of safety, it makes the creation of linked data structures like lists and trees nearly impossible. I don't think this is a good tradeoff and would not select Rust for my projects at this time.

I will try hard not to left my negative view overly influence our discussion as there are still interesting things to learn.

But it doesn't have a ruddy REPL. WTF^M?

Ownership of Memory Locations

Ownership is Rust's most unique feature and has deep implications for the rest of the language. It enables Rust to make memory safety guarantees without needing a garbage collector, so its important to understand how ownership works.

The Rust Programming Language, Ch 4

Also from the book:

- Each value (memory block) 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 (de-allocated)

Ownership Relates to Scoping

- ▶ PLs provide bindings of names to values in memory
- ► Each PL has semantics about when names go out of scope and what becomes of the memory bound to it
- ▶ Below example shows a simple example in 3 related languages
- ▶ Rust is similar to others with Stack/Heap allocation BUT detects when values are no longer reachable and immediately de-allocates them
- ► This strategy has myriad consequences that we'll discuss

```
// C version
                              // Java version
                                                                // Rust Version
void print str(){
                              public static void printStr(){
                                                               fn printStr(){
  int i = 5:
                                int i = 5:
                                                                  let i = 5:
  char s[6] = "hello":
                                String s = "hello":
                                                                let s = "hello":
  char *h = malloc(6);
                                String h = new String("there"); let h = String::new("there");
  strcpy(h, "there");
 printf("%d %s %s\n",
                                System.out.printf("%d %s %s\n",
                                                                 println!("{i} {h} {s}")
        i,s,h);
                                                  i,s,h);
                              // i stack-allocated
// i stack-allocated
                                                                // i stack-allocated
// s stack-allocated
                              // s refs const
                                                                // s refs static str
// automatically de-allocated // automatically de-allocated
                                                               // automatically de-allocated
// h heap-allocated
                              // h heap-allocated
                                                               // h heap-allocated
// h out of scope: heap ref // h out of scope, subject
                                                               // h out of scope, immediately
// is lost, memory leak // to be GC'd later
                                                                // "dropped" to de-allocate it
```

Ownership Basics

To get a start on Ownership, examine ownership_basics.rs which has a series of 4 examples

- 1. i32 integers as params
- 2. Broken String ownership
- Working String ownership with cloning
- Working String ownership with references

These will start to give a sense of the rules the Rust compiler enforces on ownership >> rustc ownership_basics.rs

>> ./ownership_basics

2 plus 3 is 5

3 plus 2 is 5

two plus three is two_three
two plus three is three_two

Ownership Basics 1 / 4: Copy-able Values

- ► Some types like i32 (32-bit signed integers) copy their bits when assigned or passed as parameters¹
- ▶ Identical semantics to C / Java / OCaml
- Copying means everyone owns distinct copies

¹Rust denotes this "copyable" quality with its Copy Trait which is implemented by i32 the type of integers. We'll look at Traits and supporting them in the near future though likely not Copy.

Ownership Basics 2 / 4: Problems

- Strings in Rust are heap-allocated, passed as pointers to the heap location just as in C / Java / OCaml
- Only one owner of String can exist and ownership can change hands
- This breaks the code below.

```
1 // ownership_basics: string append Version 1 (broken)
2 fn show_append() {
3 let s = String::from("two"); // allocate strings
4 let t = String::from("three");
5 let st = append2(s,t); // append2() assumes ownership of s and t
6 let ts = append2(t,s); // ownership lost and compiler forbids re-use
   println!("{s} plus {t} is {st}");
    println!("{s} plus {t} is {ts}");
8
9 }
10 fn append2(x: String, y: String) -> String{
11 let mut z = String::new();
12 z.push_str(&x);
13 z.push_str(&y);
14 return z;
15 } // x,y now dropped and de-allocated
```

Ownership Basics 2.5 / 4: Compiler Errors

rustc ownership_basics.rs generates some loud errors

These are frequent in Rust development, requires learning to follow the compiler and "borrow checker" rules

Ownership Basics 3 / 4: Cloning

A fix for the ownership problems is to Clone the Strings using the clone() method from its Clone Trait, identical in name and semantics to Java's idea of Clone.

```
1 // ownership basics: string append Version 2 (works via cloning)
 2 fn show append() {
     let s = String::from("two"):
 3
     let t = String::from("three");
 4
    let st = append2(s.clone(),t.clone()); // append2() gets its own copies
 5
6
     let ts = append2(t.clone(),s.clone()); // of s and t
7
     println!("{s} plus {t} is {st}");
     println!("{s} plus {t} is {ts}");
8
9 }
10 fn append2(x: String, y: String) -> String {
     let mut z = String::new():
11
     z.push str(&x):
12
   z.push str(" ");
13
14
     z.push str(&v):
15
     return z;
16 }
```

Cloning works but is dissatisfying as data must be duplicated every time a function is called. Rust provides a more efficient alternative.

Ownership Basics 4 / 4: References

A more efficient fix for the ownership problem is to adjust the function parameters and call site to use a **Reference** with notation

- &myvar at a call site or in an assignment generates a reference to myvar
- param:&String for a function parameter type of param is a String reference

```
1 // ownership basics: string append Version 3 (works via refs)
2 fn show_append() {
3 let s = String::from("two");
4 let t = String::from("three");
5 let st = append2(&s,&t); // append2() gets references to s/t
6 let ts = append2(&t,&s); // show append() retains ownership
  println!("{s} plus {t} is {st}");
7
    println!("{s} plus {t} is {ts}"):
8
9 }
10 fn append2(x: &String, y: &String) -> String {
    let mut z = String::new(): // params x.v are refs to String
11
  z.push str(&x);
12
  z.push_str("_");
13
   z.push str(&y);
14
15
   return z;
16 }
```

Working with refs is as essential to Rust as pointers are to C

References

Reference in Rust use the & syntax as in

fn myfunc(param: &some_type) {} // param has ref type

- Refs are like pointers in they give access to the data pointed at
- They differ from full pointers in that they do not give the ability to end the life of a memory block which is restricted to the owner of the block
- Refs allow borrowing of memory blocks

Exercise: Motivating References from C

What is wrong with the following C program

```
1 void use arr(int *arr, int len){
     printf("arr: [");
2
                                            17 int main(int argc, char *argv[]){
     for(int i=0; i<len; i++){</pre>
                                                 int len = 5;
                                            18
       printf("%d ",arr[i]);
                                                 int *a = malloc(sizeof(int) * len);
                                            19
5
                                                 for(int i=0; i<len; i++){</pre>
                                            20
     printf("]\n");
6
                                            21
                                                   a[i] = (i+1)*10;
7
                                            22
8
                                            23
   void use up arr(int *arr, int len){
                                                 use arr(a,len);
                                            24
     printf("arr: [");
10
                                                 use up arr(a,len);
                                            25
     for(int i=0; i<len; i++){
11
                                            26
                                                 use arr(a,len);
       printf("%d ",arr[i]);
12
                                            27
13
                                            28
                                                 return 0:
14
     printf("]\n");
                                            29 }
     free(arr):
15
16 }
```

Answers: Motivating References from C

```
// c mem problems.c:
void use up arr(int *arr, int len){
  printf("arr: [");
  for(int i=0; i<len; i++){
    printf("%d ",arr[i]);
  printf("]\n");
  free(arr):
int main(int argc, char *argv[]){
  int len = 5:
  int *a = malloc(sizeof(int) * len);
  for(int i=0; i<len; i++){</pre>
    a[i] = (i+1)*10:
  use arr(a.len):
                        // okav
  use up arr(a,len); // free'd
  use_arr(a,len);
                        // not okav
  return 0;
```

- A classic use after free error
- 2nd call to use_arr() accesses a free()'d block
- ► Java / OCaml don't allow user free()'s, GC does this

Exercise: How Rust "Fixes" the C Mistakes

```
1 // rust_owner_problems.rs:
  fn use arr(arr: Vec<i32>){
     print!("arr: ["): // arr owned
     for x in arr{
       print!("{x} ");
     println!("]");
                         // arr dropped
9
   fn use up arr(arr: Vec<i32>){
     print!("arr: [");
                        // arr owned
12
    for x in arr{
       print!("{x} ");
13
14
     println!("]");
16
                         // arr dropped
17
   fn main(){
     let len = 5:
     let mut a = vec![]:
    for i in 0..len {
       a.push((i+1)*10):
23
24
     use arr(a); // ownership lost
     use_up_arr(a); // compiler error
25
     use_arr(a);
26
27 }
```

Side Notes

- ► Check out vec! [] macro to create Vector
- push(x) to add on to a vector
- Iteration over a range via start..stop

Ownership

- Rust fixes the C problem by passing ownership of a memory block to functions by default
- Once passed into the first function, a is lost and cannot be used

What's the fix for this in Rust

Answers: How Rust "Fixes" the C Mistakes

```
1 // rust owner borrow.rs:
  fn use arr(arr: &Vec<i32>){
     print!("arr: ["); // arr borrowed
    for x in arr{
       print!("{x} "):
     println!("]");
8
                        // arr not dropped
   fn use up arr(arr: &Vec<i32>){
    print!("arr: ["): // arr borrowed
11
    for x in arr{
       print!("{x} "):
     println!("]");
16
                        // arr not dropped
17
   fn main(){
    let len = 5:
    let mut a = vec![];
    for i in 0..len {
       a.push((i+1)*10);
23
24
     use arr(&a):
                   // ownership retained
     use up arr(&a): // ownership retained
25
     use arr(&a);
26
                       // ownership retained
27 F
                        // a is now dropped
```

Fixes

```
27 | use_arr(a); // ownership lost
| value moved here
28 | use_up_arr(a); // compiler error
| value used here after move
```

note: consider changing this parameter type in function `use_arr` to borrow instead if owning the value isn't necessary

- Adjust parameter to be reference types
- Adjust calls pass references to functions
- References do not affect ownership nor cause drops (deallocation)

Mutable References

- Synatx &mut x may be used in place of & to indicate reference may be mutated
- ► Multi-threaded programs are restricted to use only 1 mutable reference at a time OR as many immutable refs as desired

```
1 // mut ref.rs:
2 fn add some(vec: &mut Vec<i32>){
  for i in 1..=3 {
      vec.push(i);
                                   // alters param vector
7 fn main(){
    let mut v = vec![10,11];
  add_some(&mut v);
                                   // pass with ability to mutate
  add_some(&mut v);
                                   // and again
10
   println!("{:?}",v);
                                   // use hand debug print formating
11
12 }
  >> rustc mut ref.rs
  >> ./mut ref
   [10, 11, 1, 2, 3, 1, 2, 3]
```

Vectors

Vectors are Rust's goto data structure, an extensible array in the same vein as Python Lists / Java ArrayList

```
1 // vec demo.rs:
    let mut v : Vec<i32> = Vec::new(); // required type annotation for new()
2
    v.push(10); v.push(20); v.push(30); // extend vector
3
4
    let mut v2 = vec![10,20,30];
                                         // vec! macro is commonly used
    v2[1] = 40:
                                          // standard [] indexing
    println!("v2[1]: {}", v2[1]);
6
7
8
    // Oh so many ways to iterate
    for x in &v2 {
9
                                          // implicit slice
       print!("{x} ");
10
11
12
    for x in &v2[..] {
                                          // explicit slice
       print!("{x} "):
13
14
    for x in v2.iter() {
                                          // explicit iterator
15
       print!("{x} ");
16
17
    for i in 0..v2.len(){
                                          // traditional via range
18
19
       print!("v2[{i}]: {} ",v2[i]);
20
    for (i,x) in v2.iter().enumerate() { // iterator + index
21
22
       print!("v2[{i}]: {x} ");
23
    // Vectors are Generic / Polymorphic; type annotation below is optional
24
    let vs1 : Vec<&str> = vec!["katz:","all","your","bass","..."];
25
                         = vec!["are:"."belong"."to"."us"]:
    let vs2
26
27 ...
```

Slices in Rust

19 ...

- Vectors support slices, a borrowed portion of a data structure
- Allows for efficient borrowing of portions of Vectors / DS's
- Syntax for slice types is... interesting

Type	Rust Parlance	Elems	C Equiv
&[i32]	slice of i32	i32	array of int
&[&str]	slice of &str	&str	array of char[]
&str	string slice (bleck!)	char	plain char[]

```
1 // vec_demo.rs: SLICES: borrowed portions of vectors
    let v100: Vec < i32 > = (0..100).collect(); // range to vector
    println!("v100[50]: {}",v100[50]):
    let v20 40: &[i32] = &v100[20..40]; // SLICE of vector w/ explicit
    for x in v20_40 {
                                        // type annotation
       print!("{x} ");
7
8
    let v50\ 70 = &v100[50..70]:
                                     // SLICE omitting type annotation
9
    for x in v50 70 {
10
      print!("{x} "):
11
12
13
    let vs =
                                         // vector of primitive str
       vec!["katz:","all","your","bass","are","belong","to","us"];
14
15
    let sl vs : &[&str] = &vs[1..4]; // SLICE of vector, type annotation optional
    for x in sl vs {
                                        // iterate over slice
16
       print!("{x} ");
17
18
```

str and String 1/2

The str type, also called a string slice, is the most primitive string type. It is usually seen in its borrowed form, &str. It is also the type of string literals, &'static str. ... A &str is made up of two components: a pointer to some bytes, and a length.

Rust Docs for str

The String type is the most common string type that has ownership over the contents of the string. It has a close relationship with its borrowed counterpart, the primitive str.

Rust Docs for String

str and String 2 / 2

```
1 // string vs str.rs:
 2 fn main(){
 3 let a = "hello world":
                                            // primitive str not meant to grow in size
 4
    let b = String::from("hello world"):
                                            // standard String buffer of characters
5
6
     let mut c = "goodbye mut":
                                            // both can be made mutable
7
     let mut d = String::from("goobye mut");
8
     for (i.ch) in c.chars().enumerate() {
                                            // iterate over characters in a str
9
10
       println!("c[{i}]: {ch}");
                                            // .iter() pops of chars while
                                            // .enumerate() gives index, char pairs
11
12
     for (i.ch) in d.chars().enumerate() {
                                            // likewise for characters in a String
13
       println!("d[{i}]: {ch}");
14
15
16
     let cs1 : &str = &c[2..11];
                                           // slice of &str is &str
     let ds1 : &str = &d[2..11]:
                                            // slice of String is &str
17
18
     let clen = c.len():
                                             // length methods for both str and String
19
20
     let dlen = d.len():
21
     // c.push_str(" again"); // no ability to grow a str
22
     d.push str(" again"):
                                  // methods for String
23
24
     // c.replace_range(0..1, "And "); // no supported method as str isn't meant to grow
25
26
     d.replace range(0..0, "And "); // grow
27 }
```

Ownership and Data Structures

- vec_owner.rs provides a demo of several uses for how ownership changes wrt to the Vector data structure
- In some cases, ownership of data transfers to the Vector: memory block will be de-allocated when the Vector is de-allocated
- ► In other cases, Vector only contains a reference to data owned elsewhere
- Some versions and lines are commented out as they are rejected by the compiler: being able to explain why these won't compile is good practice for...
 - Writing your own code
 - Exam Questions which might ask for an explanation

Exercise: Ownership within Data Structures

Data structures like Vectors can be composed of

- Owned data: de-allocated when DS is dropped
- Borrowed data: data persists when DS is dropped

```
1 // vec ownership.rs:
  /////// A ///////
    let s = String::from("abc"); // string is born
3
    let mut v = vec![]:
                     // vec is born
   v.push(&s);
                         // string ref from vec
    println!("v[0]: {}",v[0]);  // okay as s is in scope
7
8
    /////// B ///////
9
    let s = String::from("abc"); // string is born
10
   let mut v = vec![]; // vec is born
11
   v.push(s);  // vec now owns string
// println!("s: {s}");  // ERROR: s lost ownership
   v.push(s);
12
13
14
```

What is the inferred type of v in A / B examples?

Answers: Ownership within Data Structures

Data structures like Vectors can be composed of

- Owned data: de-allocated when DS is dropped
- ▶ Borrowed data: data persists when DS is dropped

What is the inferred type of v in A / B examples?

```
1 // vec_ownership.rs:
2 ////// A //////
   let s = String::from("abc"); // string is born
    let mut v = vec![]; // vec is born
4
   v.push(&s);
                         // string ref from vec
    6
    println!("v[0]: {}",v[0]);  // okay as s is in scope
7
8
    // v: Vec<&String> "Vector of String Refs"
9
10
    /////// B ///////
11
    let s = String::from("abc"); // string is born
12
    let mut v = vec![]:
                         // vec is born
13
                           // vec now owns string
14
   v.push(s);
    // println!("s: {s}"); // ERROR: s lost ownership
15
16
   println!("v[0]: {}",v[0]); // okay as v owns string
17
    // v: Vec<String> "Vector of Strings", owns strings
18
```

Exercise: Issues in DS Ownership

- Examples C-F of code involve a Vector owning a String
- ▶ Will each example compile or will rustc find ownership issues?

This is tricky but is good practice for exam questions

```
1 // vec ownership.rs:
                                              25 // vec ownership.rs:
                                              26 ////// E ///////
 2 ////// C ///////
 3 fn make_vec() -> Vec<String>{
                                              27 fn make vec() -> Vec<&'static String>{
 4 let s = String::from("abc");
                                              28 let s = String::from("abc");
 5 let mut v = vec![]:
                                              29 let mut v = vec![]:
 6 v.push(s);
                                              30 v.push(&s);
    return v;
                                              31
                                                 return v;
 8 }
                                              32 }
 9 fn use make vec() {
                                              33 fn use make vec() {
     let v = make vec();
                                              34 let v = make vec();
10
     println!("v[0]: {}".v[0]):
                                                  println!("v[0]: {}",v[0]):
11
                                              35
12 }
                                              36 }
13
                                              37
14 ////// D ///////
                                              38 ////// F //////
15 fn make vec() -> Vec<&String>{
                                              39 fn vec ref(s: &String) -> Vec<&String>{
   let s = String::from("abc");
                                                  let mut v = vec! []:
16
                                              40
17 let mut v = vec![];
                                                 v.push(s);
                                              41
18
   v.push(&s);
                                              42
                                                  return v;
    return v:
                                              43 }
19
                                              44 fn use_vec_ref(){
20 }
21 fn use_make_vec() {
                                                 let s = String::from("abc");
22
     let v = make vec():
                                              46  let v = vec ref(&s):
23
     println!("v[0]: {}",v[0]);
                                              47
                                                  println!("v[0]: {}",v[0]);
24 }
                                              48 }
```

Answers: Issues in DS Ownership 1 / 2

```
1 // vec ownership.rs:
2 ////// C ////// OKAY
3 fn make vec() -> Vec<String>{ // CORRECT: return vec of String
4 let s = String::from("abc"); // string is born
5 let mut v = vec![]:
                                 // vec is born
6 v.push(s);
                                 // string moves into vec, vec owns it
7
   return v;
                                  // safe to return
8 }
9 fn use make vec() {
   let v = make vec();
10
    println!("v[0]: {}",v[0]);
11
12 }
13
14 ////// D ////// ERROR
15 fn make vec() -> Vec<&String>{ // ERROR: unable to resolve types
16 let s = String::from("abc"): // string is born within function
  let mut v = vec![];
                                 // vec is born
18 v.push(&s);
                                  // ref to string from vec, string still owned by s
                                  // ERROR: returning from func would de-allocate s
19 return v:
20 }
                                  // but refs remain to it in v
21 fn use make vec() {
    let v = make_vec();
22
    println!("v[0]: {}",v[0]);
24 }
```

Answers: Issues in DS Ownership 2 / 2

```
25 // vec ownership.rs:
26 ////// E ////// ERROR
27 fn make vec() -> Vec<&'static String>{
28 let s = String::from("abc"); // for (D), compiler suggests adding &'static
29 let mut v = vec![]:
                                 // which is a "lifetime" for the string. This
30 v.push(&s);
                                 // does not help in this case as the fundamental
31
   return v;
                                 // error is that s's String needs to be owned by
32 }
                                  // the escaping vector
33 fn use make vec() {
   let v = make vec();
34
    println!("v[0]: {}",v[0]);
35
36 }
37
  ////// F ////// OKAY
39 fn vec ref(s: &String) -> Vec<&String>{ // string owned from elsewhere
40
  let mut v = vec![]:
                                         // vec is born
  v.push(s);
                                         // vec refers to string, doesn't own it
41
42
   return v;
                                         // safe to return as string isn't owned by v
43 }
44 fn use vec ref(){
45
  let s = String::from("abc"); // string is born
  let v = vec ref(&s):
46
47
    println!("v[0]: {}",v[0]);
48 }
```

Defining New Data Types: struct and enum

- Basic data type declaration are laid out like in OCaml
 - struct : like an OCaml record (or C struct)
 - enum : like an OCaml Algebraic (Variant) types
- ► There are few surprises, just slightly different syntax
- Next few slides demo basic aspects
- ► Then we'll proceed to Rust's version of methods via impl-ementations

struct: Creating New Types with Fields

```
1 // struct enum demo.rs:
2 struct Omelet {
                               // a new data type with fields
                                  // floating point field
 3 cook_time: f32,
 4 is cooked: bool.
                                  // boolean field
 5
     ingredients: String,
                                  // string field
 6
7
   fn use omelet(){
                                  // demonstrates use of Omelet type
     println!("===USE OMELET====");
9
     let om = Omelet {
                                  // create an immutable Omelet
10
       cook_time: 2.5,
11
12
   is cooked: false.
13
       ingredients: String::from("bacon cheddar")
     }:
                                  // print out a field
14
15
     println!("om.ingredients: {}",om.ingredients);
16
     let mut mom = Omelet {
                                  // create a fully mutable Omelet
17
18
       cook time: 0.0,
       is_cooked: false,
19
20
       ingredients: String::from("spinach swiss")
21
     };
     mom.cook time += 4.999:
                                  // alter fields
22
     mom.is cooked = true:
23
     mom.ingredients.push str(" mushroom");
24
     println!("mom.ingredients: {}",mom.ingredients);
25
26
     println!("mom.cook time: {:.2}",mom.cook time);
27 }
```

enum: New Data Type with Variants

```
1 enum Breakfast {
                                   // like OCaml's algebraic types, 4 variants of type
                                   // Variant with no additional data
 2
     None,
 3 Meager(String).
                                  // Variant with String data
 4 Hearty (Omelet),
                                  // Variant with Omelet data
    Misc(u32,String)
                                   // Variant with pair of u32 and String
 5
6 }
7 use Breakfast::*;
                                   // allow bare names of variants like Meager()
8
   fn breakfast count(br: &Breakfast) -> u32 {
10
     return match br {
                                  // match on the variants of Breakfast
       None
                      => 0.
11
12
       Misc (count, ) => *count, // read about * on your own time
13
                      => 1
                                  // all other variants
14
     }:
15
  }
16 fn use breakfast(){
     let dog br = Meager(String::from("kibble"));
17
     let md br = Misc(4,String::from("cups oatmeal"));
18
     let ck_br = Hearty(Omelet{cook_time: 5.00, is_cooked: true,
19
20
                                ingredients: String::from("ham swiss")});
21
     let br vec = vec! [dog br, md br, ck br, None]; // all of Breakfast type
22
     for (i,x) in br_vec.iter().enumerate() { // iterate over array handling
23
       let desc = match x {
                                               // each variant in a match
                   => "none", Meager(_) => "meager",
24
         Hearty( ) => "hearty", Misc(_,_) => "misc",
25
       }:
26
       let count = breakfast_count(&x);
27
28
       println!("{i}: {desc}, count: {count}");
29
30
```

Implementing "Methods" with impl Blocks

- Rust is NOT object-oriented but can feel that way due to its data syntax and types which feature the "dot" notation used in OO languages
- Rust favors an approach similar to C++: define data type (struct) in one place, define associated functions in another place via impl construct
- Rust supports various syntatic "sugar" around impl such as "method" invocation
- Several impl can exist for a struct allowing associated functions to be defined across several files (though this is atypical)

```
struct Omelet {
                    // a new data type
  . . . ;
impl Omelet {    // "methods" for Omelets
  // construct an Omelet
  fn new(ingr: &str) -> Omelet {
    . . .
 // cook an omelet
 fn cook(&mut self, time: f32){
impl Omelet{    // MORE "methods" for Omelets
  // add an ingredient
  fn add ingredient(&mut self, ingr: &str){
    . . .
  // check for overcooked
 fn is_overcooked(&self) -> bool {
    return self.cook time > 8.0:
  // non-"method" function
 fn denver_ingredients() -> &'static str {
    return "ham cheese peppers";
```

impl_demo.rs Highlights 1 / 2

. . .

```
// impl demo.rs:
struct Omelet { cook time: f32, is cooked: bool, ingredients: String}
impl Omelet {
                           // "methods" for Omelet struct
 return Omelet{cook time: 0.0,
            is cooked: false,
            ingredients: String::from(ingr)};
 fn cook(&mut self, time: f32){      // cook an omelet
  if self.cook_time >= 5.0 { // which is assigned by the compiler
                         // to &mut Omelet based on the context
    self.is_cooked = true;
fn main(){
 let ingr = Omelet::denver ingredients(); // invoke function in Omelet namespace
 Omelet::cook(&mut denver, 0.25); // direct invocation of cook() function
 denver.cook(0.25):
                        // "dot" invocation of cook() function
 denver.add ingredient(" mushrooms"); // "dot" invocation of function
 denver.cook(0.5);
 }
```

Aside: "methods" and Methods

Methods are similar to functions: we declare them with the fn keyword and a name, they can have parameters and a return value, and they contain some code thats run when the method is called from somewhere else. Unlike functions, methods are defined within the context of a struct (or an enum or a trait object, which we cover in Chapter 6 and Chapter 17, respectively), and their first parameter is always self, which represents the instance of the struct the method is being called on.

- The Rust Programming Language, Ch 5.3
- Rust "methods" are mostly just functions
- Proper Methods are a family of functions with overriding usually through sub-classing and dynamic dispatch to select one of several possible functions at runtime based on the specific type of the data being used though alternatives exist
- Rust does not have sub-classing / overriding
- ► It does have a dynamic dispatch facility that is still evolving and requires use of smart pointers like Box and Trait Objects



Traits: Data Supporting Common Operations

- ► Common idea: data with operations XYZ can be used here (in this function, in this algorithm, in this data structure, etc)
- Common XYZ examples are (1) ability to compare data to sort sorting it and (2) ability to write/read a representation of data to files for saving/loading
- Pure OO-Solution accomplishes this with a class hierarchy, abstract parent classes with concrete child classes extending parents and overriding necessary methods, has downsides acknowledged by folks who know OO well
- ▶ Alternative: denote XYZ is present without inheritance
 - ▶ interface in Java, Type Class in Haskell, Protocol in Clojure
 - Trait in Rust
- ► A Trait is a list of functions that must be present that data can implement to be compatible with other functionalities
- ▶ A function or data structure type can be annotated to accept only parameters that implement Traits
- ► As Java uses Classes a lot, Rust uses Traits alot!

(A) New Trait / New Datatype Implementations

A reasonably common case: new Datatype and new Trait

```
// trait impl.rs:
struct Omelet {
                              // new datatype Omelet
 cook time: f32, is_cooked: bool, ingredients: String,
impl Omelet {
                // "methods" for Omelet
 fn new(ingr: &str) -> Omelet { ... }
 fn cook(&mut self, time: f32){ ... }
trait Updateable {
                              // a new trait Updateable
 fn update(&mut self);
impl Updateable for Omelet {
                              // (A) new Trait Updateable, new type Omelet
 fn update(&mut self){
   self.cook(0.25):
                               // update an Omelet by cooking it a bit
```

(B) Existing Trait / New Datatype Implementations

Common allows New datatypes to use existing Rust functionality

```
impl Iterator for Omelet{
                               // (B) existing Trait Iterator, new datatype Omelet
  type Item = ();
                               // iterator returns unit, denoted ()
  fn next(&mut self) -> Option<Self::Item> { // required "method" for Iterator
    if self.is cooked {
                                            // when cooked, stop iterating
     return None:
    else {
      self.cook(0.5):
                              // iterate by cooking a little and returning
     return Some(());
                               // some of unit
 }
impl Display for Omelet (// (B) existing Trait Display, new datatype Omelet
  fn fmt(&self, f: &mut Formatter<'_>) -> Result {
    return write!(f, "Omelet{{ cook_time: {:.2}, is_cooked: {}, ingredients: {}}}",
                 self.cook time, self.is cooked, self.ingredients);
                               // allows Omelet to be println!()'d
fn main() {
  for in &mut omelet {
                                         // iterate on omelet due to Iterator
    println!("Iterating on omelet"):
    // println!("omelet: {}",&omelet); // ERROR: mut ref to omelet already owned above
  println!("Finished omelet: {}".omelet): // print formatted omelet due to Display
```

(C) New Trait / Existing Datatype Implementations

Somewhat common: teach an old data type a new trick

```
trait Updateable {
    fn update(&mut self);
}
impl Updateable for i32 {
    fn update(&mut self){
       *self = *self + 1;
    }
}
impl Updateable for String {
    fn update(&mut self){
       *self = *self + 1;
    }
}
impl Updateable for String {
    fn update(&mut self){
       self.push('_');
    }
}
// (C) new Trait Updateable, existing type String
fn update(&mut self){
       self.push('_');
}
```

(D) Existing Trait / Existing Datatype Implementations

ERROR: prevented by Rust for reasons that are not explained

- Suggested work around is to define a new type that mirrors existing type and use it instead
- Somewhat dissatisfying but so it goes

```
type Item = char;
                         // ERROR: this case is not allowed by rust
 fn next(&mut self) -> Option<Self::Item> {
   return self.pop(); // return Some(last_char) or None
>> rustc trait impl.rs && trait impl
error [E0117]: only traits defined in the current crate can be implemented for types define
 --> trait_impl.rs:83:1
                           // (D) exiting Trait Iterator, existing type i32
83 I
    impl Iterator for String {
    --<del>-</del>-----
                   'String' is not defined in the current crate
    impl doesn't use only types from inside the current crate
  = note: define and implement a trait or new type instead
```

error: aborting due to previous error

Rust Trait Implementation Summary

Rust Traits

	Trait	Туре	Ok?	Notes
Α	New	New	Yes	Can implement MyTrait for MyData
В	Exists	New	Yes	Can implement Iterator for MyData
C	New	Exists	Yes	Can implement MyTrait for i32
D	Exists	Exists	Sub	Cannot implement Iterator for i32
				Can create new subtype for i32
				to implement Iterator

May return to discussion of extending data types and available functions later time permitting, hopefully discuss the **Expression Problem**

Java Interfaces Are Similar

	Interf	Class	Ok?	Notes
Α	New	New	Yes	Can implement MyInterf for MyClass
В	Exists	New	Yes	Can implement Iterator for MyClass
C	New	Exists	Sub	Must subclass Object implement MyInterf
D	Exists	Exists	Sub	Must subclass Object implement Iterator

Trait Usage

Rust Type system allows for the use of Traits to constrain Generic (polymorphic) functions to accept only data with certain methods

```
1 fn show_it<T:Display>(thing: T){ // accept any type with Display trait
     println!("The thing is {}",thing);
 3
 5 fn main(){
                                    // show use of generic function
     show it("Hello"):
     show it(1.234);
     show it(Omelet{cook time: 1.0, is cooked: false,
8
                    ingredients: "ham cheese".to string()}):
 9
10 }
11
   // Alternative syntaxes for constraining types
   fn show it2(thing: impl Display){
     println!("The thing is {}".thing):
14
15 }
16
   fn show it3<T>(thing: T)
   where T: Display
19
   {
     println!("The thing is {}",thing);
20
21 }
```

NOTE: 3 separate syntaxes available to dictate data must have a certain trait

Ref in Returns create Lifetime Problems

► A common example of ownership difficulties is the following function from Ch 10.3 of The Rust Programming Language

```
fn longest(x: &str, y: &str) -> &str {
   if x.len() > y.len() {
      x
   } else {
      y
   }
}
```

- Because a &str (string slice / ref) is returned, the borrow checker is nervous that its underlying owned data could be dropped prior to use.
- rustc needs a little re-assurance and is kind enough to ask:

Lifetimes as Generic Types

```
1 // generic on datatype T: constrains v / z to have a type relation
 2 fn add front<T>(v: Vec<T>), z: T){
     v.push(z);
 3
     v.rotate right(1);
 5
 6
   // generic on lifetime 'a: constrains x / y / return to have a lifetime relation
8 fn longest<'a>(x: &'a str, y: &'a str) -> &'a str {
 9
     if x.len() > v.len() {
10
       x
    } else {
11
12
13
14 }
```

- 'a is a lifetime, <'a> indicates generic func on lifetimes
- x: &'a str indicates x's ref'd data has a lifetime at least as long as 'a
- y and the returned &str do as well
- Borrow checker assuaged in that returned type will have lifetime at least as long as parameters
 - Obvious in this simple case to smart humans
 - One day the compiler may get that smart too

'a Notation as Minimum Lifetimes

- Note that 'a lifetimes are best interpreted as a "minimum" lifetime, NOT "equal" lifetimes²
- Example below shows shorter + longer lifetime data passed to longest()

```
fn longest<'a>(x: &'a str, y: &'a str) -> &'a str { }

// a static string: program-length lifetime
static STAT_STRING: &'static str = "Who wants to live forever?";
fn different_lifetimes() {
    // a string with function-length lifetime
    let string1 = String::from("abcd");

    // compiles and runs fine
    let result = longest(string1.as_str(), STAT_STRING);
    println!("The longest string is {}", result);
}
```

 $^{^2}$ Rust made an unfortunate notation choice as folks coming from OCaml have seen 'a and will instinctively assume two things that are both 'a are equal in type

Linking Structures

Challenge: Define a Singly Linked List with the following "methods"

- ▶ new() to create an empty list
- push(data) to add a new node with given data at the head
- pop() to remove the head node and return its data
- Works with any data type (generic / polymorphic)

You should feel confident of doing so in

Java, Python, OCaml

How about in Rust?

Rust Singly Linked Lists

- box_linked_list.rs shows one possibility for a Rust singly-linked list which leans towards mutability
- Uses combination of enum for Nodes (a la OCaml) and struct for list itself
- ▶ Requires use of Box, a "smart pointer" which allows heap allocation of arbitrary data
- Requires use of the mem::replace(dest,src) function to deal with tricky ownership issues
- ➤ Singly linked lists which add/remove from the head are as simple as it gets and require "Chapter 15" functionality
- Doubly Linked Lists, Trees, and more intense data structures require more advanced techniques beyond our scope here

Difficulty with creating Linked Data Structures is the price of the Safety that Rust guarantees

Highlights of box_linked_list.rs

31 }

```
1 enum Node {
                                                  // type for nodes in list
     Cons(i32, Box<Node>),
     Nil,
 3
   }
 4
 5
6 struct List {
                                                  // type for list
     head: Node,
     len: u32.
 8
9 }
10
   impl List {
11
12
     fn push(&mut self, data: i32){
                                                  // add element at head of list
13
       let stolen =
                                                  // adjust ownership of head to stolen,
14
         mem::replace(&mut self.head, Nil);
                                                  // head becomes Nil momentarily
15
       self.head = Cons(data, Box::new(stolen)); // assign head to newly allocated node
16
       self.len += 1:
                                                  // update length
17
18
     fn pop(&mut self) -> Option<i32> {
                                                  // remove head element and return it or None
19
20
       let stolen =
                                                  // steal the head making it Nil for a moment
21
         mem::replace(&mut self.head, Nil);
                                                  // so that it can be owned in match
22
       match stolen {
                                                  // match head node
23
         Nil => return None,
                                                  // head was Nil so remains so
         Cons(data, box_next) => {
24
                                                // head was Cons
           self.head = *box next:
25
                                                 // set head to next node, deref needed
26
           self.len -= 1:
                                                 // decrement length
                                                  // return stolen data
27
           return Some(data);
28
29
```

Further Study

- Rust is a BIG language and environment
- Immediate next steps in study would be to examine
 - cargo the build tool and package manager
 - Functional language features and Closures
- More advanced features of note include
 - "Fearless" Concurrency with Threads and Smart Pointers
 - ► Macro implementation like with vec! [] and println!()
 - Unsafe layer which is used to implement much of the standard library, which is very safe. We promise.

Or you could wait 10 years to see if Rust stabilizes, perhaps adds a full garbage collector, or collapses under its own complexity giving rise to a new language that Gen-Next rabidly admires and believes is the one true way... at which point you'll be old enough to snort derisively at it and talk about what you had to go through to learn programming

```
(quote "Prediction is very difficult, especially about the future!"
    "-- attributed to Niels Bohr, physicist and Nobel laureate")
(quote "Fashion changes, but style endures." 'Coco 'Chanel)
(quote "God had a deadline. So he wrote it all in Lisp." 'Eternal 'Flame)
```

Sundries

Arithmetic Overflow Checks overflow.rs

- Default rustc options augment every + * operation to check for overflow
- Can be disabled with rustc -C overflow-checks=off, increases speed of arithmetic by about 2X on simple benchmarks
- Opposite of C which defaults to no checks but can be enabled with compiler options gcc -ftrapv

Sorted Print sorted_print.rs

- Attempted to provide a more complex example of Trait usage in sorted_print.rs
- Attempts to create a generic function which accepts an iterable data structure with elements that implements several traits
- Sort the data and print it
- Lots of strange syntax, borrowing problems abound
- Gave up after seeing a 3-year old Stack Overflow post with the same errors and no answers
- Student contributions welcome