Esercitazione 2

7 e 9 Maggio 2025

Obiettivi

- semplice lettura di file
- uso di array
- iterare sul array e su riferimenti ad array
- gestione caratteri nelle stringhe e operazioni sui caratteri
- lettura dei parametri da command line in modo semplice

Esercizio 1 (disponibile su crownlabs)

Un pangramma è una frase che contiene tutte le lettere dell'alfabeto, indifferentemente maiuscole o minuscole (26 in inglese).
Scrivere un programma che legga come argomento di command line il nome di un file contenente del testo e determini se il testo è un pangramma. Le lettere possono essere ripetute e la frase può contenere altri caratteri di punteggiatura, che devono essere ignorati. Inoltre il programma conti quante volte ciascun carattere si ripete e stampi la statistica.
Il programma deve passare i test forniti a parte, quindi organizzare il codice in funzioni che rispettino i test senza modificare i test stessi.

Esempio di invocazione ed esecuzione:

- creare un file di testo sentence.txt con "`abcdefghijklmnopqrstuvwxyz`"
- da terminale (dentro l'ide) invocare:
  - cargo build
  - cargo run -- sentence.txt
- output:
  "`abcdefghijklmnopqrstuvwxyz" is a pangram`
  `a 1`
  `b 1`
  `c 1`
  `…`
  `z 1`

Suggerimenti per la soluzione:

- copiare il file fornito con i test al posto di main.rs di un nuovo progetto; noterete che dovete implementare le funzioni chiamate nei test che implementano i mattoncini della logica applicativa. quando si è in presenza di IO è sempre buona prassi separare la funzione di IO e quella di logica, in modo da testare la logica con degli stati creati sinteticamente, senza fare io.
- per la lettura semplice degli argomenti da command line guadare il modulo env, in particolare: `let args: Vec<String> = env::args().collect()`
- quando si esegue "cargo run -- qualcosa" tutto quello a destra di -- (doppio hyphen) viene passato come argomento al programma, spezzato secondo gli spazi presenti

- nella soluzione dovrete ciclare sugli elementi di un array, cercare di spiegare la differenza tra
  `fn f(v: &[u32]) {  for el in v {} }`
  e
  `fn f(v: &[u32]) { for &el in v {} }`
  Che tipo è el e perché? Che operazione viene fatta per determinare il tipo di el in &el?

Esercizio 2

Obiettivi:

- gestione stringhe e slice
- differenza tra “caratteri” (char) e “byte” (u8) in Rust
- lettura parametri da command line
- aggiungere dipendenze ad un progetto rust

Con il termine “slug” si intende una stringa convertita in versione semplificata, composta solo dai caratteri [a-z][0-9]-, in modo da poterla usare in url, chiavi, con solo caratteri presenti su tutte le tastiere.
Nella stringa originale i caratteri non ammissibili vengono convertiti seguendo queste regole:

- tutti i caratteri accentati riconosciuti vengono convertiti nell’equivalente non accentato
- tutto viene convertito in minuscolo
- ogni altro carattere rimanente che non sia in [a-z][0-9] viene convertito in “-”
- due “-” consecutivi non sono ammessi, solo il primo viene tenuto
- un “-” finale non è ammesso a meno che non sia l’unico carattere nella stringa

L’obiettivo dell’esercizio è fare un funzione “slugify” che converta una stringa generica in uno slug.

PASSI PER LA SOLUZIONE

1. Creare con cargo un nuovo package chiamato **slugify** e, in main.rs, definire la funzione

   ```
   fn slugify(s: &str) -> String {}
   ```

2. Per convertire le lettere accentate definire una funzione che esegua la conversione e che poi verrà chiamata dentro slugify:

   ```
   fn conv(c: char) -> char {}
   ```

   Conv restituisce il carattere c se è ammesso, la lettera non accentata corrispondente se viene trovata, o “-” negli altri casi.
   Considerare solo c già convertiti a minuscolo in precedenza e, attenzione, la conversione a minuscolo di un char restituisce un iteratore, perché in alcune lingue ad una maiuscola corrispondono due minuscole.

   All’interno usare questa tabella di conversione, dove il carattere nella posizione i (come carattere) in SUBS_I corrisponde al carattere nella posizione i in SUBS_O:

   ```
   const SUBS_I : &str =
   "àáâäæãåāăąçćčđďèéêëēėęěğǧḧîïíīįìıİłḿñńňôöòóœøōõőṕŕřßśšşșťțûüùúūǔůűųẃẍÿýžźż";
   const SUBS_O: &str  =
   "aaaaaaaaaacccddeeeeeeeeegghiiiiiiiilmnnnnoooooooooprrsssssttuuuuuuuuuwxyyzzz";
   ```

ATTENZIONE: SUBS_I e SUBS_O essendo degli slice di stringa non possono essere indicizzati direttamente con [pos] (**perché?**), tutto quello che si può assumere è che il carattere corrispondente alla posizione i-esima di SUBS_I è nella stessa posizione in SUBS_O. Conviene quindi convertire le due stringhe in vettori/array di char e fare lookup lì dentro.
(inoltre attenzione che con il copia incolla alcuni accenti potrebbero corrompersi, nel caso farsi a mano una stringa con solo alcune lettere accentate che si hanno su tastiera)

3. Scrivere degli unit test per le funzioni create
(riferimento: https://doc.rust-lang.org/rust-by-example/testing/unit_testing.html)

    a. creare una sezione in main.rs per ospitare i test
    ```
    #[cfg(test)]
    mod tests {
          use super::*;
    }
    ```
    b. i test sono funzioni così definite all'interno:
    ```
    #[test]
    fn my_first_test() {
          // valore = preparazione test
          assert_eq!(valore, valore_atteso)
    }
    ```
    c. definire almeno questi test:
        i. conversione lettera accentata
        ii. conversione lettera non accentata
        iii. conversione lettera non ammessa sconosciuta
        iv. conversione lettera accentata non compresa nella lista (es ῶ)
        v. stringa con più di una parola separata da spazio
        vi. stringa con caratteri accentati
        vii. stringa vuota
        viii. stringa con più spazi consecutivi
        ix. stringa con con più caratteri non validi consecutivi
        x. stringa con solo caratteri non validi
        xi. stringa con spazio alla fine
        xii. stringa con più caratteri non validi consecutivi alla fine
    d. i test possono venire lanciati con

    ```
    cargo test
    ```

4. Rendere la funzione invocabile dalla command line

    Nel main() leggere una sequenza di parole come argomento da command line, invocare la funzione e stampare il risultato.

    Esempio:
    ```
    cargo run -- "Questo che slug    sarà???"
    ```

risultato: “slug: questo-che-slug-sara“

(Attenzione: “--” (doppio hyphen “-”) dopo run, serve per separare i parametri di cargo da quelli del comando; la stringa tra apici doppi viene letta come unico parametro e non come n parametri. Non usate copia incolla dal testo, perché l’editor usato per comporre il doc sostituisce gli apici doppi con altri caratteri)

Per il parsing degli argomenti di command line si suggerisce di inserire nel progetto la libreria clap: https://docs.rs/clap/latest/clap/

Clap è una libreria per leggere e validare i parametri da command line e ha due modalità di funzionamento, “derive” e “build”. La prima permette di definire i parametri da leggere mediante gli attributi di una struct, la seconda è più flessibile ma meno immediata da usare e consente di costruire i parametri da leggere in modo imperativo con una serie di istruzioni.
In questo progetto useremo la modalità “derive” (tutorial completo con esempi a questo indirizzo https://docs.rs/clap/latest/clap/_derive/_tutorial/index.html)

a. aggiungere la libreria cargo, editando il file cargo.toml od eseguendo:`cargo add clap --features derive`
   verrà aggiunto a cargo.toml:
   ```
   [dependencies]
   clap = { version = "4.5.3", features = ["derive"] }
   ```
b. in questo esempio non abbiamo opzioni con nome (es --verbose) e quindi leggiamo tutti i valori passati come una sola stringa, quindi definire una struct Args derivata da clap::Parser in questo modo:
   ```
   #[derive(Parser, Debug)]
   struct Args {
       // input string
       slug_in: String,
   }
   ```
c. la sintassi di clap è semplice: si definisce un attributo (in questo caso slug_in) per ogni parametro che si vuole leggere. Clap cercherà di fare la conversione automatica dei parametri passati al tipo indicato, mentre darà errore nel caso in cui non sia possibile
d. invocando nel main `let args = Args::parse()` si effettua il parsing della command line e restituisce un oggetto di tipo Args con i parametri richiesti; clap aggiunge automaticamente l’opzione di --help e la gestione degli errori
e. provare ad invocare `cargo run -- --help` per verificare che clap funzioni in modo corretto
f. a questo punto l’invocazione `cargo run -- “Questo sarà uno slug?”` dovrebbe inserire in `args.slug_in` tutta la stringa “Questo sarà uno slug!” (notare i doppi apici intorno alla stringa, perché vogliamo che tutte le parole siano interpretate come un unico parametro)
g. per impratichirsi con clap aggiungere due parametri fittizi opzionali con nome --repeat=n e --verbose. Repeat è di tipo intero, verbose boolean.
   Usare l’annotazione #[arg()] come negli esempi al link indicato per impostare il comportamento del parser.

Inoltre verificare, invocando il comando, che le conversioni, vengano lette in modo corretto, in particolare quella ad intero di --repeat

h. come potrei modificare il tipo di `slug_in` per leggere tutte le parole della stringa da convertire in un vettore di stringhe anziché in un'unica stringa?

Esercizio 3

Obiettivi

- Utilizzo di array
- parsing di stringhe
- mutabilità
- utilizzo di struct ed enum
- gestire i valori di ritorno e gli errori
- lettura/scrittura da file

ESERCIZI Propedeutici

Dopo avere creato un nuovo progetto rust provare questi tre brevi task (in funzioni dedicate) prima di passare alla soluzione dell'esercizio descritto in seguito.

1. Aprire, leggere e salvare un file: leggere un file "test.txt" con dentro del testo e salvare il testo ripetuto 10 volte nello stesso file
   Usare le funzioni read_to_string e write definite in std::fs (https://doc.rust-lang.org/std/fs/)
   a. Testare cosa capita quando il file o il path non esiste, gestire gli errori
   b. che differenza c'è tra read e read_to_string? Provare a leggere un file con delle lettere accentate con read e stampare in esadecimale l'array di byte letto con read, allineato con il testo sulla riga sopra.
      Esempio:
      c i a o \n
      63 69 61 6f 0a
      Cosa notate se nel file è scritto "così\n" al posto di "ciao\n"?

2. Enum con "valore"
   A differenza del C le enum sono tipi che possono ospitare all'interno valore associato agli elementi della enum.
   Ogni elemento di una enum può essere di un tipo diverso e con `match` si può estrarre il contenuto della enum in una variabile.
   Definire quindi una `enum Error` con dentro due valori: `Simple(SystemTime)` e `Complex(SystemTime, String)` e fare una funzione `print_error(e: Error)` che stampi il tipo di errore e le informazioni contenute (senza usare {:?} debug, ma gestendo i valori della enum in modo esplicito)

3. Funzioni che possono restituire errori.
   In rust una funzione per segnalare un errore può usare la enum Result, che è definita in questo modo, analogamente a Option

```
pub enum Result<T, E> {
      Ok(T),
      Err(E)
}
```

   T è il tipo del valore restituito come risultato, mentre E quello dell'errore, che può essere una enum qualsiasi (T e E nel codice sono da sostituire con i tipi veri, come

nell'esempio seguente)
Implementare questa funzione:

```
pub mul(a: i32, b: i32) -> Result<u32, MulErr> {}
```

dove MulErr è

```
pub enum MulErr {Overflow, NegativeNumber};
```

mul quindi deve restituire:

- il risultato se possibile
- MulErr::NegativeNumber se uno tra a e b è negativo (non entrambi)
- MulErr::Overflow se il risultato è più alto del valore massimo di u32

Hint: usare la funzione **checked_mul** per controllare se l'operazione dà overflow
E' possibile vedere su godbolt (https://godbolt.org/) come avviene il check dell'overflow. Confrontare come come viene compilata una funzione mul(int num) che restituisca num*num in C e in Rust. Cosa sono le istruzioni **seto** e **jo**?

4. Uso di self, &self e &mut self
   Nei metodi di una struct si può utilizzare self per avere un riferimento all'oggetto su cui si sta operando: che differenza di comportamento c'è tra self, &self e &mut self?

   Ipotizziamo di avere la struttura

```
struct Node {
	name: String,
	size: u32,
	count: u32,
}
impl Node {
	pub fn new(name: &str) -> Node  {
		Node {name: name.to_str(), size: 0, count: 0}
	}
}
```

   Aggiungere due metodi size(self, n: u32) e count(self, c: u32) in modo che questo codice

```
let node = Node(String::new("nodo")).size(10).count(5);
```

   crei il nodo {"nodo", 10, 5}

   Quante struct di tipo Node costruisco in tutto? Ci sono penalità dal punto di vista dell'efficienza?

   Questo modo di creare l'oggetto, un pezzo per volta a partire dai default, viene definito "**builder pattern**". Ciò permette di ovviare al problema della mancanza di valori default e del polimorfismo nei metodi rust: nel caso di molti parametri opzionali posso usare il pattern per creare una versione "base" dell'oggetto con i valori di default e modificare solo gli attributi che hanno un valore diverso.

   Aggiungere un metodo to_string() che lo trasformi in stringa "name:node size:10 count:5". Come deve essere definito self in questo caso? Con o senza "&"?
   Infine aggiungere due metodi grow() e inc() che aumentano rispettivamente la size e

il count di 1 senza creare un nuovo oggetto. Qui self come va definito?

TESTO ESERCIZIO

Un programma deve gestire la costruzione di uno schema di battaglia navale 20x20 salvato su file.
La costruzione dello schema è a passi. Alla prima invocazione si costruisce una board vuota, poi ad ogni successiva invocazione si aggiunge una nave nella posizione indicata e si salva lo schema aggiornato.
Il formato del file è il seguente (21 righe):

- **LINEA 1**: N1 N2 N3 N4, 4 interi separati da spazio che indicano il numero di navi rispettivamente di lunghezza 1, 2 , 3 e 4, che si possono ancora aggiungere alla board
- **LINEE 2..21**, 20 righe di 20 caratteri con " " (spazio) per le caselle vuote e "B" per quelle con navi

La costruzione della board avviene per passi, invocando il programma con dei parametri

- **cargo run -- board.txt new 4,3,2,1**
  questo crea una nuova board vuota nel file board.txt e può ospitare 4 navi da 1, 3 navi da 2, 2 navi da 3, e una da 4.
- **cargo run -- board.txt add_boat V,3,10,10**
  legge la board in board.txt, aggiunge una nave da 3 caselle in verticale, partendo dalla casella (10,10) e andando giù 3 caselle, fino a (12,10). Possibili direzioni: H e V, il verso è sempre da alto a basso e da sinistra a destra.
  Aggiunta la nave, salva il risultato in board.txt, aggiornando anche le navi disponibili nella prima linea.
  Gli indici iniziano da 1 (in alto a sinistra) fino a 20

L'operazione di add_boat deve essere "safe" e stampare errore, senza panic e senza modificare il file nel caso in cui non si possa aggiungere la nave. Casi in cui la nave non si può aggiungere:

- sono già state aggiunte tutte le navi possibili
- una nave si sovrappone o ha almeno un casella che "tocca" una nave esistente (anche di angolo)
- la nave va fuori dallo schema

(per la gestione della command line utilizzare sempre clap, come nell'esercizio precedente, notare gli spazi che separano file, dimensione boat e posizione start: provare a leggere tutti gli args in un vettore di stringhe, poi il primo valore avrà il nome del file, il secondo il comando, il terzo i parametri del comando, che saranno da "interpretare")

Per gestire la board utilizzare la seguente struttura ed implementare i metodi indicati **senza modificare la signature dei metodi presenti**:

```
const bsize: usize = 20;
pub struct Board {
```

```
    boats: [u8; 4],
    data: [[u8; bsize]; bsize],
}

pub enum Error {
    Overlap,
    OutOfBounds,
    BoatCount,
}

pub enum Boat {
    Vertical(usize),
    Horizontal(usize)
}

impl Board {
      /** crea una board vuota con una disponibilità di navi */
      pub fn new(boats: &[u8]) -> Board {}

      /* crea una board a partire da una stringa che rappresenta tutto
      il contenuto del file board.txt  */
      pub fn from(s: String)->Board {}

      /* aggiunge la nave alla board, restituendo la nuova board se
      possibile */
      /* bonus: provare a *non copiare* data quando si crea e restituisce
      una nuova board con la barca, come si può fare?  */
      pub fn add_boat(self, boat: Boat, pos: (usize, usize))
            -> Result<Board, Error> {}

      /* converte la board in una stringa salvabile su file */
      pub fn to_string(&self) -> String
}
```

Bonus

1. identificare e scrivere i test per struct Board
2. modificare Board in modo che add_boat alteri la struttura Board anziché crearne una nuova; per fare questo prestare attenzione al parametro self, come va modificato?
3. modificare il parse di clap in modo che gestisca i parametri in modo più espressivo e provare ad utilizzare la modalità builder anziché derive (https://docs.rs/clap/latest/clap/_tutorial/chapter_0/index.html):
   cargo run -- add --file=board.txt --boat=3V --start=10,10