Bridging the gap between Typestates and Rust in production code

losé Duarte

António Ravara (Advisor)

March 2021

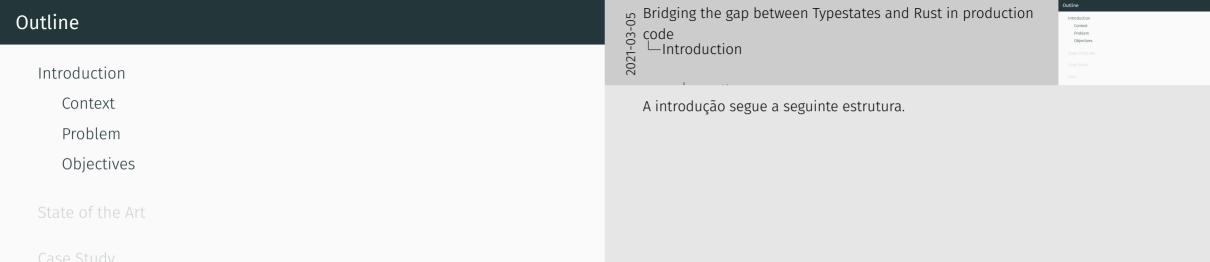
NOVA School of Science and Technology

Bridging the gap between Typestates and Rust in production code

Bridging the gap between Typestates and Rust in production code

Duarte
Inio Rawara (Advisor)
th 2021
School of Science and Technology

Bom dia Professor(es), sou o José Duarte, e o meu tema de tese é "Bridging the gap between typestates and Rust in production code".



Context

Software plays a crucial role in our lives.

From web browsers, to word processors and more!

As software becomes more important, bugs become more expensive.

- · Losing work due to a bug in the save procedure is not nice.
- · A bug in the firmware for a pacemaker may cost a life.

Bridging the gap between Typestates and Rust in production code
—Introduction
—Context

Software plays a crucial role in our lives.

From web browsers, to word processors and more!

As software becomes more important, bugs become more expensive.

Losing work due to a bug in the save procedure is not nice.

A bug in the firmware for a nacemaker may cot at life.

Dado o papel crescente do software nas nossas vidas, os bugs saem cada vez mais caro, tanto às empresas como aos utilizadores.

Perder uma mensagem no Facebook não é grave, mas um dia de trabalho é pior. Imagine-se então um bug num pacemaker, pode causar a morte do utilizador!

Problem

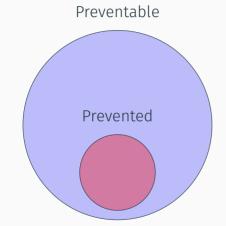


Figure 1: Diagram of preventable bugs and prevented bugs.

Bridging the gap between Typestates and Rust in production code
—Introduction
—Problem



Podemos imaginar todos os tipos de bugs como um conjunto (não mostrado), onde se inserem dois subconjuntos. Os bugs que podemos prevenir e os que, de facto, conseguimos prevenir.

Os que prevenimos são, por exemplo, type mismatch errors; os erros que não conseguimos produzir são, por exemplo, gestão de memória.

Problem — with Rust

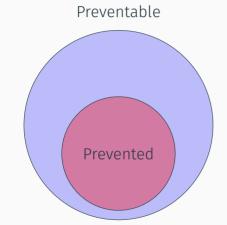


Figure 2: Diagram of preventable bugs and prevented bugs when considering Rust's borrow checker.

Bridging the gap between Typestates and Rust in production

code
Introduction

Problem



Rust permite-nos alargar o conjunto de bugs que conseguimos prevenir, incluindo assim os bugs de memória e de concurrência.

No entanto, continua a não conseguir cobrir todos os bugs.

Problem — Ideal

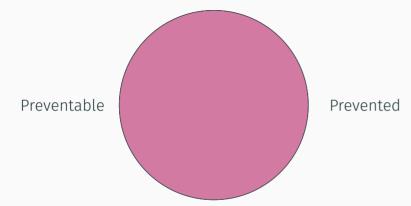


Figure 3: The ideal diagram of preventable bugs and prevented bugs, where all bugs are prevented.

Bridging the gap between Typestates and Rust in production code
Introduction
Problem



O ideal seria uma sobreposição total, como na figura. Os typestates ajudam-nos a dar um passo nesta direção.

Apesar da programação defensiva tentar maximizar os erros prevenidos, existe sempre o limite do erro humano.

Problem

Error happens at runtime, possibly crashing the program.

```
fn main() {
    let builder = XYBuilder::new();
   builder.setX(0.0):
   builder.build(): // runtime error: missing `v`
```

Our tools should work for us, not make us work for them.

Bridging the gap between Typestates and Rust in production code -Introduction └─ Problem



Queremos diminuir os bugs. No entanto, os bugs que queremos endereçar são principalmente maus usos das APIs.

Como na figura, o programa iria lançar um panico. Enquanto programadores, queremos que a máquina trabalhe por nós, não ao contrário.

Minimizando assim o tempo de debug, e a complexidade associada ao trabalho.

2021-

What are typestates?

Typestates are an approach to behavioral types that allow the developer to cleanly and concisely express API constraints.

Bridging the gap between Typestates and Rust in production

code
Introduction

Objectives

es are an approach to behavioral types that allow the developer y and concisely express API constraints.

Typestates são uma abordagem a tipos comportamentais. Estes visam descrever o estado em runtime do programa através do sistema de tipos.

Podem ser modelados como máquinas de estados e permitem ao programador expressar de forma clara e consisa restrições de uso de uma API.

Objectives

A library which brings *practical* typestates to Rust.

- · Minimal learning overhead.
- · Scalable to large projects.

Bridging the gap between Typestates and Rust in production code
—Introduction
—Objectives

ibrary which brings practical typestates to Ru

- Minimal learning overhead.

- Scalable to large projects.

Objectives

Esta tese visa então providenciar typestates práticos na linguagem Rust. Os mesmos:

- · Devem ser fáceis de aprender e usar.
- · Devem escalar a projetos grandes.



State of the Art

Session Types

Typestates

10

code

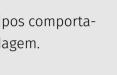
-State of the Art

Devido a restrições de tempo vou cingir-me a implementações de tipos comportamentais para linguagens de alto nível em linha com a minha abordagem.

Bridging the gap between Typestates and Rust in production

State of the Art

Outline



Session Types — StMungo

StMungo is a tool that converts Scribble protocols into a Mungo typestate specification and Java skeleton.



```
module example;

type <xsd> "..."
    from "..." as Hello;

global protocol P(role A, role B) {
    connect(Hello) from A to B;
}
```



Bridging the gap between Typestates and Rust in production code
State of the Art
Session Types



O StMungo (ou Scribble to Mungo) é uma ferramenta que converte protocolos de Scribble para especificações de typestates em Mungo e esqueletos de API em Java. A implementação Java é depois verificada pelo Mungo.

O Scribble permite ainda a conversão de multiparty session types para binary session types.

Session Types — Session Types & Rust

	Kind	Compile-Time	Ext. Tool
Jespersen et al.	Binary	✓	
Lagaillardie et al.	Multiparty	✓	\checkmark

Dridging the gan between Typestates and Dust in produc	Session Types — Session Types & Rust
Bridging the gap between Typestates and Rust in produc	LIOII
code code	Kind Comp
⊆ ∟State of the Art	Jespersen et al. Binary ✓ Lagaillardie et al. Multiparty ✓
Session Types	

(Tanto quanto sei) O trabalho de Jespersen et al. foi o primeiro a introduzir session types em Rust, entre esse e o seu principal sucessor as diferenças residem no suporte de multiparty session types e o uso ferramentas externas. Ambos válidos para o seu sucessor Lagaillardie.

Typestates — Plaid

Plaid is a typestate-oriented language.

It also supports aliasing control through keyword usage:

	Aliasing	Mutation
unique		√
immutable	√	
shared	√	√

Typestates — Plaid			
Plaid is a typestate-oriented l	anguage.		
It also supports aliasing conti	rol through key	word usage:	
	Aliasing	Mutation	
unique		√	
unique immutal	ole 🗸	√	

O Plaid é uma linguagem orientada aos typestates, fazendo uso dos mesmos como cidadãos de primeira classe.

A outra caracteristica mais relevante da linguagem é o seu controlo de aliasing, efetuado através de palavras-chave, como descrito na tabela.

Typestates — Fugue

```
[WithProtocol("open", "closed")]
class OuterSocket {
    [InState("connected",
       WhenEnclosingState="open"),
     NotAliased(WhenEnclosingState="open")]
    [Unavailable(WhenEnclosingState="closed")]
   private Socket innerSocket;
```

Bridging the gap between Typestates and Rust in production 2021-03code -State of the Art —Typestates

Seguindo então para os typestates, começo com o Fugue. Um projeto vindo da Microsoft, onde o meu trabalho se inspira.

O Fugue permite a descrição de estados e transições através de anotações.

Permite ainda a ligação entre sub-estados para garantir o bom uso das APIs.

Typestates — Mungo

```
typestate Socket {
    Init = { Status open(): <OK: Open, ERROR: end> }
    Open = { void read(): <OK: Open> }
    Close = { void close(): end }
}
```

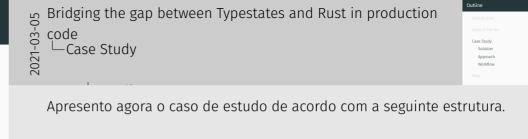
Bridging the gap between Typestates and Rust in production code
State of the Art
Typestates



Por fim temos o Mungo, já falado antes mas não apresentado formalmente. O Mungo é um sistema de tipos / um verificador de typestates.

O seu uso em Java consiste no processamento de uma anotação que liga a descrição do typestate de uma classe, após o processamento da anotação, o Mungo verifica todos os usos da classe validando os mesmos contra a especificação.

Outline



Outline

Case Study Solution Approach

Workflow

fn main() {

Bridging the gap between Typestates and Rust in production code -Case Study

Para relembrar, o problema é o compilador não avisar que aquela chamada está fora de ordem.

Solution

Ideally, we want to catch the error at compile-time.

```
fn main() {
    let builder = XYBuilder::new();
    builder.setX(0.0);
    builder.build();
    // ^^^^^
    // error: you cannot build without setting `y`
}
```

How?

Bridging the gap between Typestates and Rust in production

code

Case Study

Solution

Nós queremos que o compilador nos diga "hey, não podes chamar este método nesta altura". Como é que o podemos fazer?

Solution

Ideally, we want to catch the error at compile-time.

```
fn main() {
    let builder = XYBuilder::new();
    builder.setX(0.0);
    builder.build();
    // ^^^^^
    // | error: you cannot build without setting `y`
}
```

How? Declare the protocol.

Bridging the gap between Typestates and Rust in production

code
Case Study

Solution

Solution

Ideally we want to catch the error at complie-time.

In main() {
 let builder = XYMulider::new();
 but leter = xtd(a, a);
 but leter = xtd(a, a);
 but leter = xtd(a, a);
 let leter = xtd(a, a);
 leter = xtd(a, a);

Nós queremos que o compilador nos diga "hey, não podes chamar este método nesta altura". Como é que o podemos fazer?

Annroach — Overview

We can exploit the Rust typesystem to emulate typestates.

Usando o sistema de tipos de Rust, podemos emular typestates. No entanto, esta abordagem requer muito boilerplate.

Esta abordagem não requer um esforço especial por parte do utilizador, visto que os macros são um elemento comum da linguagem.

Usando macros, podemos gerar o código por nós. Podemos ainda validar o código contra outras propriedades e lançar erros durante a compilação.

2021-03-

-Case Study └─Approach However this approach requires boilerplate.

Bridging the gap between Typestates and Rust in production 2021-03code Case Study —Approach

We can exploit the Rust typesystem to emulate typestate

Annroach — Overview

Usando o sistema de tipos de Rust, podemos emular typestates. No entanto, esta abordagem requer muito boilerplate.

Esta abordagem não requer um esforço especial por parte do utilizador, visto que os macros são um elemento comum da linguagem.

Usando macros, podemos gerar o código por nós. Podemos ainda validar o código contra outras propriedades e lançar erros durante a compilação.

Approach — Overview

We can exploit the Rust typesystem to emulate typestates.

However this approach requires boilerplate.

Use macros! Part of the language, requiring no new experience.

Bridging the gap between Typestates and Rust in production

code

Case Study

Approach

We can exploit the Rust typesystem to emulate typestates.

However this approach requires boilerplate.

Use macros! Part of the language, requiring no new experience.

Usando o sistema de tipos de Rust, podemos emular typestates. No entanto, esta abordagem requer muito boilerplate.

Esta abordagem não requer um esforço especial por parte do utilizador, visto que os macros são um elemento comum da linguagem.

Usando macros, podemos gerar o código por nós. Podemos ainda validar o código contra outras propriedades e lançar erros durante a compilação.

Approach — Overview

We can exploit the Rust typesystem to emulate typestates.

However this approach requires boilerplate.

Use macros! Part of the language, requiring no new experience.

· Rewrite the annotated code, generating boilerplate for the user.

Bridging the gap between Typestates and Rust in production code

Case Study

Approach

We can applied the Rust typesystem to emulate typestates.

However this approach requires bolleriplate.

Les marcord Fart of the language, requiring on new experience.

Rewrite the annotated code, generating bolleriplate for the user.

Annroach — Overview

Usando o sistema de tipos de Rust, podemos emular typestates. No entanto, esta abordagem requer muito boilerplate.

Esta abordagem não requer um esforço especial por parte do utilizador, visto que os macros são um elemento comum da linguagem.

Usando macros, podemos gerar o código por nós. Podemos ainda validar o código contra outras propriedades e lançar erros durante a compilação.

Approach — Overview

We can exploit the Rust typesystem to emulate typestates.

However this approach requires boilerplate.

Use macros! Part of the language, requiring no new experience.

- · Rewrite the annotated code, generating boilerplate for the user.
- Throw errors during compile-time.



We can exploit the Rust typesystem to emulate typestates.

However this approach requires boilerplate.

Use macrosl Part of the language, requiring no new experience.

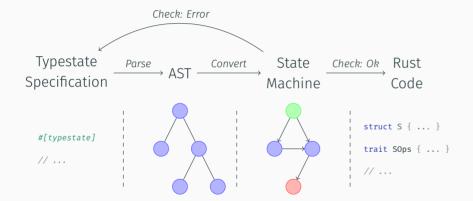
Annroach — Overview

Rewrite the annotated code, generating boilerplate for the use
 Throw errors during compile-time.

Usando o sistema de tipos de Rust, podemos emular typestates. No entanto, esta abordagem requer muito boilerplate.

Esta abordagem não requer um esforço especial por parte do utilizador, visto que os macros são um elemento comum da linguagem.

Usando macros, podemos gerar o código por nós. Podemos ainda validar o código contra outras propriedades e lançar erros durante a compilação.

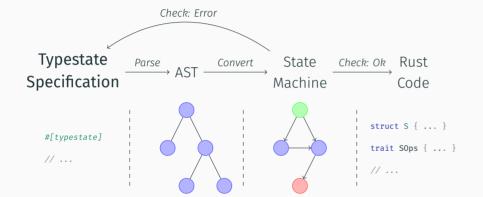


Bridging the gap between Typestates and Rust in production code
Case Study
Approach



Aprofundando, para fazer com que tal aconteça a arquitetura do macro será como na figura. A especificação é código Rust normal, recorrendo a macros.

Daí, o sistema de macros extrai a AST, a qual precisamos apenas de processar para extrair uma máquina de estados e gerar o código final.



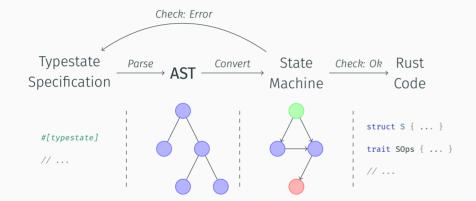
Bridging the gap between Typestates and Rust in production code —Case Study

Approach



Aprofundando, para fazer com que tal aconteça a arquitetura do macro será como na figura. A especificação é código Rust normal, recorrendo a macros.

Daí, o sistema de macros extrai a AST, a qual precisamos apenas de processar para extrair uma máquina de estados e gerar o código final.

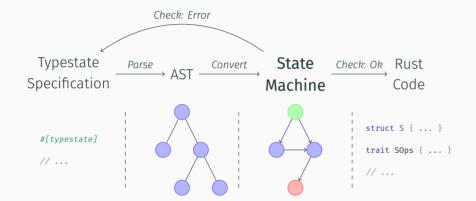


Bridging the gap between Typestates and Rust in production code
Case Study
Approach



Aprofundando, para fazer com que tal aconteça a arquitetura do macro será como na figura. A especificação é código Rust normal, recorrendo a macros.

Daí, o sistema de macros extrai a AST, a qual precisamos apenas de processar para extrair uma máquina de estados e gerar o código final.

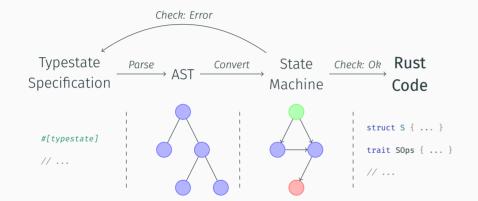


Bridging the gap between Typestates and Rust in production code
Case Study
Approach



Aprofundando, para fazer com que tal aconteça a arquitetura do macro será como na figura. A especificação é código Rust normal, recorrendo a macros.

Daí, o sistema de macros extrai a AST, a qual precisamos apenas de processar para extrair uma máquina de estados e gerar o código final.



Bridging the gap between Typestates and Rust in production code
Case Study
Approach

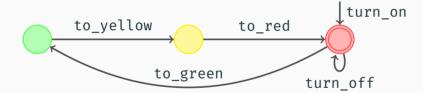


Aprofundando, para fazer com que tal aconteça a arquitetura do macro será como na figura. A especificação é código Rust normal, recorrendo a macros.

Daí, o sistema de macros extrai a AST, a qual precisamos apenas de processar para extrair uma máquina de estados e gerar o código final.

Workflow — Design the state machine

Consider a traffic light as a state machine.



Bridging the gap between Typestates and Rust in production code
Case Study
Workflow

Workflow — Design the state machine

Consider a traffic light as a state machine.

to_verted turn_on to_verd turn_off

Consideremos um semáforo. Ele percorre um ciclo, verde, amarelo, vermelho e o mesmo pode ser ligado e desligado.

Vamos então modelá-lo com a biblioteca.

Workflow — Declaring state in Rust

Using the DSL we first declare the module, main automata and the states:

```
#[typestate] mod traffic light {
   #[automata] struct TrafficLight;
   #[state] struct Green:
   #[state] struct Yellow;
   #[state] struct Red;
   // ...
```

We still need transitions!

Bridging the gap between Typestates and Rust in production Workflow — Declaring state in Rust 2021-03code -Case Study └─Workflow We still need transitions!

Começamos por declarar o autómato e os estados do mesmo.

Workflow — Declaring transitions in Rust

All transition functions take ownership of the current state and return the new state.

```
// code from the previous slide ...
fn to yellow(self: Green) -> Yellow;
fn to red(self: Yellow) -> Red;
fn to green(self: Red) -> Green;
// ...
```



De seguida, as transições entre estados.

Note-se o consumo do estado atual e o retorno do novo estado.

2021-03-

Workflow — Declaring start states in Rust

Functions that do not use **self** and *return* a valid state are inferred as the *start* state.

```
// code from the previous slides ...
fn turn_on() -> Red;
```

Bridging the gap between Typestates and Rust in production

code

Case Study

Workflow

Finctions the seri raise.

Finctions the seri raise.

Finctions the seri raise.

Finctions the seri raise.

	_
uncti	ns that do not use self and return a valid state are inferred as the
LUFE .	ate.

Por fim, o estado inicial e final, que neste caso é o mesmo.

Funções que não consomem o estado atual e retornam um estado são consideradas declarações de estados iniciais.

Funções que consomem o estado atual e não retornam um estado são consideradas declarações de estados finais.

Workflow — Declaring *end* states in Rust

Functions that take **self** and do *not return* a valid state are inferred as the *end* state.

```
fn turn_off(self: Red);
}
```

Our traffic light is ready!



Por fim, o estado inicial e final, que neste caso é o mesmo.

Funções que não consomem o estado atual e retornam um estado são consideradas declarações de estados iniciais.

Funções que consomem o estado atual e não retornam um estado são consideradas declarações de estados finais.

Workflow — The other features

For maintenance purposes, consider that our traffic light is now required to count every GYR cycle.

The **TrafficLight** structure is now declared as follows:

#[automata] struct TrafficLight { cycles: u64 }

Bridging the gap between Typestates and Rust in production code -Case Study └─Workflow

Workflow — The other features

Existem ainda mais alguns detalhes a apresentar.

Imaginemos que queremos saber o número de ciclos atravessados pelo semáforo, para questões de manutenção.

Adicionamos então um campo novo ao autómato, tornando-o disponível em todos os estados.

2021-03-

Workflow — The other features

How can we check if the light requires maintenance?

We add a **pure** function:

fn requires maintenance(self: &TrafficLight) -> bool;

This function is not able to perform mutations due to the immutable reference (&).

Bridging the gap between Typestates and Rust in production code

Case Study

Workflow

Workflow

Workflow

Para confirmar se o semáforo precisa de manutenção criamos um predicado. Neste caso, a função só tem acesso imutável a uma referência de qualquer estado. Sendo assim uma função pura.

Workflow — The other features

After maintenance, how can we reset the counter?

fn reset counter(self: &mut TrafficLight);

We add an **impure** function:

The function is able to perform mutations over fields, due to usage of a mutable reference (&mut), but it is unable to transition between states.

Bridging the gap between Typestates and Rust in production

code

Case Study

Workflow—The other features

After maintenance, box can we reset the counter

We add an impure function:

In reset, counter(self), flows Trafficially

The function is able to perform mutations over femulations over femulations over femulations over femulations

with the function of the function is able to perform mutations over femulations over femulations over femulations over femulations.

Após a manutenção, é prático poder dar reset ao semáforo. Para tal usamos uma função impura, capaz de mutar o estado atual, mas não de fazer transições.

Workflow — Final steps

After the code is checked and expanded, all that is left is for the user to implement the state transitions.

```
impl GreenState for TrafficLight<Green> {
    fn to_yellow(self) -> TrafficLight<Yellow> {
        /* ... */
    }
}
```

Bridging the gap between Typestates and Rust in production code
Case Study
Workflow



É possível adicionar a implementação diretamente às declarações das funções, no entanto, tal poderia tornar o código díficil de ler devido às transformações necessárias ao código.

The final user API would look like the following:

```
let tl = TrafficLight<Red>::turn_on();
let green = tl.to_green();
if green.requires_maintenance() {
    // ...
    green.reset();
}
green.to_red(); // compile-time error
```

Bridging the gap between Typestates and Rust in production code
Case Study
Workflow



A API apresentada ao utilizador final seria a seguinte.

The final user API would look like the following:

```
let tl = TrafficLight<Red>::turn_on();
let green = tl.to_green();
if green.requires_maintenance() {
    // ...
    green.reset();
}
green.to_red(); // compile-time error
```

Bridging the gap between Typestates and Rust in production code
Case Study
Workflow



A API apresentada ao utilizador final seria a seguinte.

The final user API would look like the following:

```
let tl = TrafficLight<Red>::turn_on();
let green = tl.to_green();
if green.requires_maintenance() {
    // ...
    green.reset();
}
green.to_red(); // compile-time error
```

Bridging the gap between Typestates and Rust in production code
Case Study
Workflow



A API apresentada ao utilizador final seria a seguinte.

The final user API would look like the following:

```
let tl = TrafficLight<Red>::turn_on();
let green = tl.to_green();
if green.requires_maintenance() {
    // ...
    green.reset();
}
green.to_red(); // compile-time error
```

Bridging the gap between Typestates and Rust in production code
Case Study
Workflow

```
Orkflow — Usage
The final user APT would look like the following:

[set il - Trafficisiphicades/strum_mm();
let green - title (green);
if green, requires, maintenance() {
// ...
green.reset();
green.te_red() // compile-time error
```

A API apresentada ao utilizador final seria a seguinte.





Outline

Plan

Plan Overview

																_															
	March			Ap	ril			1	May Ju					June			July				Aug	ust			September						
	1	2	2 3	4	1	2	-	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4		
Macro Development									T								П														
- Macro Parsing																															
- Code Expansion																															
- Testing																															
State Machine Verification Development																															
- Implementation of the algorithms																															
- Testing																															
Usability Checking					I																										
- Use case implementation																															
- User survey																															
- User feedback review and implementation																															
Thesis Writing																															

Figure 4: Work plan Gantt chart

Bridging the gap between Typestates and Rust in production code — Plan



Por fim, apresento o plano de trabalho. (Apresentar e explicar a tabela.)