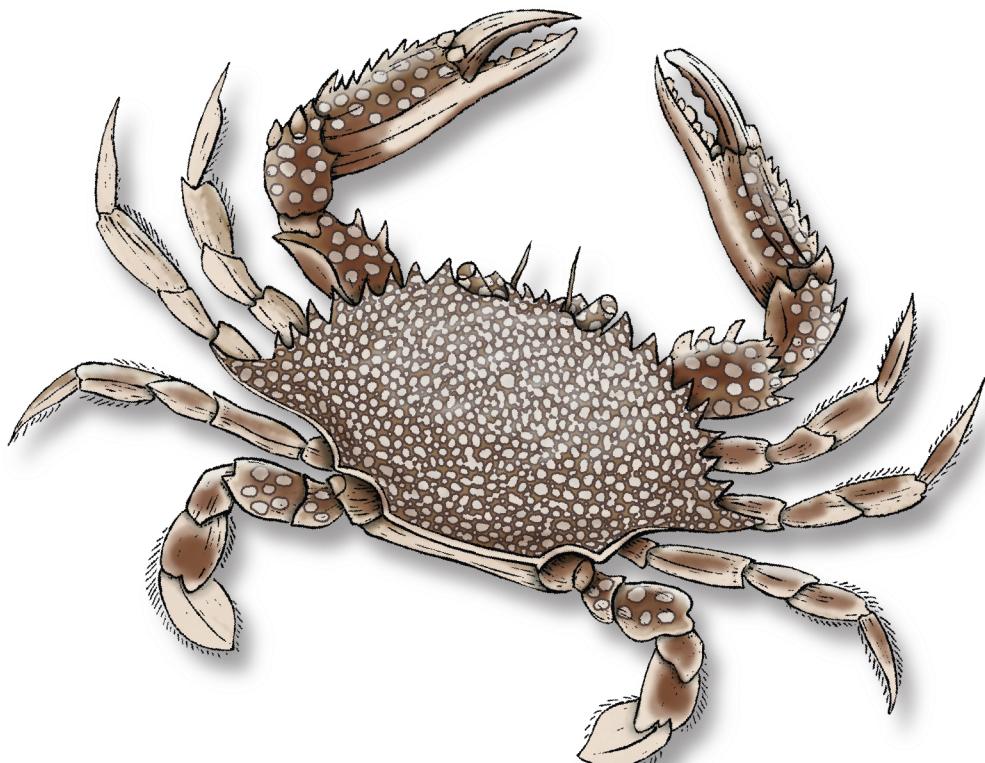


O'REILLY®

Эффективный Rust

35 конкретных способов улучшить код



SPRiNT
book

Дэвид Дрисдейл

Effective Rust

35 Specific Ways to Improve Your Rust Code

David Drysdale

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

Эффективный Rust

35 конкретных способов улучшить код

Дэвид Дрисдейл

SPRiNT
book

2025

ББК 32.973.2-018.1
УДК 004.43
Д74

Дрисдейл Дэвид

Д74 Эффективный Rust. 35 конкретных способов улучшить код. — Астана: «Сprint Бук», 2025. — 304 с.: ил.

ISBN 978-601-08-4868-9

Популярность Rust продолжает расти, в том числе благодаря таким особенностям, как защита памяти, безопасность типов и потокобезопасность. Но они же могут вызвать сложности при изучении Rust даже у опытных программистов. Это практическое руководство научит писать идиоматический код и попутно освоить систему типов, гарантии безопасности и развивающуюся экосистему Rust.

Если у вас есть опыт работы с любым компилируемым языком или вы уже знаете базовый синтаксис Rust и стремитесь получить работающие программы, книга для вас. В ней рассматриваются концептуальные различия между Rust и другими языками и даются конкретные практические рекомендации для программистов. Автор книги Дэвид Дрисдейл быстро научит вас писать код, который выглядит как идиоматический Rust-код, а не как плохой перевод с языка C++.

ББК 32.973.2-018.1
УДК 004.43

16+ (В соответствии с Федеральным законом от 29 декабря 2010 г. № 436-ФЗ.)

Права на издание получены по соглашению с O'Reilly. Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Информация, содержащаяся в данной книге, получена из источников, рассматриваемых издательством как надежные. Тем не менее, имея в виду возможные человеческие или технические ошибки, издательство не может гарантировать абсолютную точность и полноту приводимых сведений и не несет ответственности за возможные ошибки, связанные с использованием книги. Издательство не несет ответственности за доступность материалов, ссылки на которые вы можете найти в этой книге. На момент подготовки книги к изданию все ссылки на интернет-ресурсы были действующими. В книге возможны упоминания организаций, деятельность которых запрещена на территории Российской Федерации, таких как Meta Platforms Inc., Facebook, Instagram и др.

ISBN 978-1098151409 англ.

Authorized Russian translation of the English edition of Effective Rust
ISBN 9781098151409 © 2024 Galloglass Consulting Limited.

This translation is published and sold by permission of O'Reilly Media, Inc., which owns or controls all rights to publish and sell the same.

ISBN 978-601-08-4868-9

© Перевод на русский язык ТОО «Сprint Бук», 2025
© Издание на русском языке, оформление ТОО «Сprint Бук», 2025

Оглавление

Предисловие	11
Для кого эта книга.....	11
Версия Rust	12
Структура издания.....	13
Условные обозначения.....	13
Благодарности.....	14
От издательства	14
Глава 1. Типы.....	15
Рекомендация 1. Используйте систему типов для выражения структур данных	16
Фундаментальные типы.....	16
Составные типы	19
Перечисления	19
Перечисления с полями	22
Распространенные типы перечислений.....	23
Рекомендация 2. Используйте систему типов для выражения типичного поведения.....	24
Функции и методы.....	25
Указатели функций.....	26
Замыкания.....	28
Трейты	32
Рекомендация 3. Предпочитайте преобразования Option и Result вместо явных выражений match.....	35
Запомните	40
Рекомендация 4. Предпочитайте идиоматические типы Error.....	41
Трейт Error	41
Минимальные ошибки	42

Вложенные ошибки	44
Трейт-объекты	47
Библиотеки и приложения	48
Запомните	49
Рекомендация 5. Изучите преобразования типов.....	49
Пользовательские преобразования типов.....	50
Приведения	54
Неявное приведение	55
Рекомендация 6. Используйте паттерн newtype	56
Обход правила сироты для трейтов.....	60
Ограничения newtype	61
Рекомендация 7. Для сложных типов используйте паттерн «Строитель»	62
Рекомендация 8. Познакомьтесь со ссылочными типами и указателями	69
Ссылки в Rust	69
Трейты указателей.....	72
Типы жирных указателей.....	73
Другие трейты указателей	77
Типы умных указателей	78
Рекомендация 9. Рассмотрите использование преобразований итераторов вместо явных циклов	82
Трейты итераторов.....	84
Преобразования итераторов	86
Потребители итераторов	88
Создание коллекций из значений Result	92
Преобразование циклов	93
Когда явный вариант лучше.....	95
Глава 2. Трейты	96
Рекомендация 10. Изучите стандартные трейты	96
Типичные стандартные трейты	97
Стандартные трейты из других рекомендаций.....	105
Перегрузки операторов	106
Обобщение	107
Рекомендация 11. Реализуйте трейт Drop для паттернов RAII.....	110

Рекомендация 12. Уясните компромиссы между обобщениями и трейт-объектами.....	115
Обобщения	115
Трейт-объекты	117
Базовые сравнения	118
Дополнительные границы трейтов	120
Безопасность трейт-объектов	122
Компромиссы.....	123
Рекомендация 13. Минимизируйте число необходимых методов трейтов с помощью предустановленных реализаций	124
Глава 3. Концепции.....	127
Рекомендация 14. Уясните принцип времени жизни.....	128
Знакомство со стеком	128
Эволюция времени жизни.....	130
Продолжительность времени жизни	132
Арифметика времени жизни	135
Правила пропуска времени жизни.....	138
Время жизни 'static.....	139
Время жизни и куча	142
Время жизни в структурах данных	143
Анонимное время жизни	145
Запомните	146
Рекомендация 15. Изучите работу модуля проверки заимствования	146
Контроль доступа	147
Правила заимствования	149
Операции владельца.....	151
Устранение разногласий с модулем проверки заимствования	154
Запомните	166
Рекомендация 16. Страйтесь не писать unsafe-код.....	167
Рекомендация 17. Будьте осторожны при параллельном доступе к общему состоянию.....	170
Гонка данных.....	170
Взаимные блокировки.....	181
Совет.....	185

Рекомендация 18. Не паникуйте	186
Рекомендация 19. Избегайте отражения	189
Повышающее приведение вверх в будущих версиях Rust.....	195
Рекомендация 20. Не поддавайтесь соблазну излишней оптимизации.....	196
Структуры данных и аллокация	196
Кто боится большого и страшного Copy?	200
Ссылки и умные указатели	201
Глава 4. Зависимости.....	203
Рекомендация 21. Разберитесь в принципах семантического версионирования	204
Основные аспекты SemVer	205
SemVer для авторов крейтов	206
SemVer для пользователей крейтов.....	209
Рассуждения.....	210
Рекомендация 22. Минимизация видимости.....	210
Синтаксис для обозначения видимости.....	210
Семантика видимости.....	214
Рекомендация 23. Избегайте импорта с подстановкой.....	215
Рекомендация 24. Повторный экспорт зависимостей, чьи типы присутствуют в вашем API.....	217
Рекомендация 25. Управляйте графиком ваших зависимостей	220
Указание диапазона допустимых версий	223
Решение проблем с помощью инструментов.....	224
От чего зависеть	225
Запомните	226
Рекомендация 26. Остерегайтесь расплазания feature.....	227
Условная компиляция.....	227
Features	228
Запомните	232
Глава 5. Инструменты.....	233
Рекомендация 27. Документируйте публичные интерфейсы.....	233
Инструменты.....	234
Расположение дополнительной документации.....	235

Где публикуется документация крейта	236
Что не нужно документировать	236
Запомните	239
Рекомендация 28. Используйте макросы вдумчиво	239
Декларативные макросы.....	240
Процедурные макросы	244
Когда использовать макросы	249
Недостатки макросов.....	252
Совет.....	253
Рекомендация 29. Прислушивайтесь к Clippy.....	254
Рекомендация 30. Пишите не только модульные тесты.....	258
Модульные тесты.....	258
Интеграционные тесты	260
Тесты документации	260
Примеры	261
Бенчмарки.....	261
Фаззинг-тестирование.....	263
Советы по тестированию	266
Запомните	266
Рекомендация 31. Пользуйтесь преимуществами экосистемы инструментов.....	267
Запомните эти инструменты.....	269
Рекомендация 32. Настройте систему непрерывной интеграции	270
Этапы непрерывной интеграции	270
Принципы CI.....	273
Публичные системы CI.....	275
Глава 6. За пределами стандартного Rust	276
Рекомендация 33. Рассмотрите вариант сделать код библиотеки no_std-совместимым.....	277
core	277
alloc.....	278
Написание кода для no_std.....	279
Выделение с возможностью сбоя	281
Запомните	283

Рекомендация 34. Контролируйте выход за границы FFI.....	283
Вызов функций С из Rust	284
Доступ к данным С из Rust	289
Время жизни.....	290
Вызов Rust из С.....	293
Запомните	296
Рекомендация 35. Используйте bindgen вместо ручного сопоставления встречных компонентов FFI	297
За пределами С	299
Послесловие.....	300
Об авторе	301
Иллюстрация на обложке.....	302

Предисловие

Код выступает скорее в качестве руководства, нежели фактических правил.

Гектор Барбосса (Hector Barbosa)

В густонаселенном ландшафте современных языков программирования Rust стоит особняком. Он сочетает в себе скорость компилируемых языков, эффективность языков без сборки мусора и безопасность типов функциональных языков, а также является уникальным решением проблем с безопасностью памяти. В итоге по результатам опросов Rust регулярно оказывается одним из любимых у программистов (<https://oreil.ly/KKcb6>).

Сила и согласованность системы типов Rust означает, что если программа на этом языке компилируется, то уже наверняка заработает, — этот феномен ранее наблюдался только у более академических и менее доступных языков вроде Haskell. Если программа на Rust компилируется, то она к тому же будет работать *безопасно*.

И эта безопасность — как типов, так и памяти — естественно, возникает не бесплатно. Несмотря на высокое качество базовой документации, Rust прославился своей сложной кривой обучения, когда новичкам приходится раз за разом ритуально сражаться с модулем проверки заимствований, перекраивать свои структуры данных и размышлять над конструкцией «время жизни» (lifetime). Программа на Rust, которая компилируется, с высокой вероятностью заработает с первого раза, но проблема как раз в том, чтобы заставить ее компилироваться — даже с помощью весьма эффективной системы диагностики ошибок.

Для кого эта книга

Книга нацелена на то, чтобы помочь программистам разобраться с названными сложностями, даже если у них уже есть опыт работы с компилируемыми языками вроде C++.

Однако присущая Rust безопасность позволяет несколько иначе взглянуть на приводимые здесь рекомендации. Обычно книги серии «Эффективный язык» содержат *руководства*, а не правила, поскольку у руководств есть

исключения — сопровождая руководства подробным объяснением, автор позволяет читателю самому решить, оправдывает ли конкретный сценарий нарушение правила.

В нашей книге сохранен общий принцип давать советы, объясняя, на чем они основаны. Но поскольку Rust не таит в себе особых подводных камней, приводимые в книге рекомендации больше акцентированы на его принципах. Многие из этих рекомендаций идут под лозунгами «*Поймите...*» и «*Познакомьтесь с...*» и помогают писать гибкий и идиоматический код Rust.

Безопасность Rust обусловливает также полное отсутствие рекомендаций под лозунгом «*Никогда...*». Если вам действительно чего-то никогда не нужно делать, то обычно компилятор вам этого и не позволит.

Версия Rust

Примеры из книги написаны на Rust 2018 года, использующей стабильную цепочку инструментов. Обещанная этим языком обратная совместимость (<https://oreil.ly/husN4>) означает, что любое последующее издание Rust будет поддерживать код, написанный для версии 2018 года, даже если это новое издание будет содержать критические изменения. На сегодня Rust достаточно стабилен для того, чтобы различия между версиями 2018 и 2021 годов были минимальными. Никакие примеры из книги не потребуют изменений для обеспечения совместимости с версией 2021 года (лишь рекомендация 19 содержит одно исключение, в котором более свежая версия Rust делает возможным новое поведение, которое ранее не было доступно).

Рекомендации, приведенные в книге, не раскрывают никаких аспектов функциональности `async` в Rust (<https://oreil.ly/a9r1B>), поскольку это подразумевает более продвинутые концепции и менее стабильную поддержку набора инструментов — по одному только синхронному Rust уже будет предостаточно материала.

Фрагменты кода и сообщения об ошибках создавались в конкретной версии `rustc 1.70`. И если код вряд ли потребует изменений для обеспечения совместимости с будущими версиями языка, то сообщения об ошибках в зависимости от конкретной версии компилятора могут меняться. Сообщения об ошибках, которые встречаются в тексте, были отформатированы, чтобы они вписывались в ширину страницы книги. В остальных случаях они приведены в том виде, в каком их генерирует компилятор.

В тексте содержится множество ссылок к другим статически типизируемым языкам, таким как Java, Go и C++, и сравнений с ними, чтобы помочь сориентироваться читателям, имеющим опыт работы с этими языками. (Ближайшим эквивалентом Rust является, пожалуй, C++, особенно когда дело доходит до семантики перемещения в версии C11.)

Структура издания

Приводимые в книге рекомендации разделены на шесть глав.

Глава 1. Типы

Советы, относящиеся к основной системе типов Rust.

Глава 2. Трейты

Советы по работе с трейтами в Rust.

Глава 3. Концепции

Фундаментальные идеи, которые формируют структуру Rust.

Глава 4. Зависимости

Советы по работе с экосистемой пакетов Rust.

Глава 5. Инструменты

Предложения по улучшению кодовой базы за счет использования не только компилятора Rust.

Глава 6. За пределами стандартного Rust

Рекомендации для случаев, когда вам нужно работать за пределами стандартной безопасной среды Rust.

Несмотря на то что глава «Концепции» наверняка является более фундаментальной, чем главы «Типы» и «Трейты», она специально приводится после них, чтобы читатели, изучающие книгу от начала до конца, сперва могли обрести некоторую уверенность.

Условные обозначения

В этой книге используются следующие шрифтовые обозначения.

Курсив

Отмечает новые термины.

Рубленый шрифт

Им обозначены URL-адреса, адреса электронной почты и элементы интерфейса.

Моноширинный шрифт

Используется для листингов программного кода.

Моноширинный шрифт

Применяется внутри абзацев для ссылки на элементы программы, такие как имена переменных или функций, базы данных, типы данных, переменные окружения, операторы, ключевые слова, а также имена и расширения файлов.

НЕ КОМПИЛИРУЕТСЯ

```
// Содержит некомпилирующиеся фрагменты кода
```

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// Содержит фрагменты кода, демонстрирующие нежелательное поведение
```

Благодарности

Хочу выразить благодарность тем, кто помог воплотить эту книгу в жизнь:

- техническим редакторам Пьетро Альбини (Pietro Albini), Джесс Мэлс (Jess Males), Майку Кэппу (Mike Capp) и особенно Карол Николс (Carol Nichols), которые давали профессиональную и подробную обратную связь по тексту;
- моим редакторам в O'Reilly Джеффу Блейелу (Jeff Bleiel), Брайану Герину (Brian Guerin) и Кэти Тозер (Katie Tozer);
- Тизиано Санторо (Tiziano Santoro), у которого я научился азам Rust;
- Дэнни Эльфенбауму (Danny Elfmanbaum), который оказал столь необходимую техническую помощь с оформлением книги в формате AsciiDoc;
- скрупулезным читателям изначальной веб-версии издания, в частности:
 - Джулиан Россе (Julian Rosse), которая обнаружила десятки опечаток и прочих ошибок;
 - Мартину Дишу (Martin Disch), который подсказал несколько потенциальных доработок и указал на неточности в некоторых рекомендациях;
 - Крису Флитвуду (Chris Fleetwood), Сергею Каунову (Sergey Kaunov), Клиффорду Мэтьюзу (Clifford Matthews), Ремо Сенековичу (Remo Senekowitsch), Кириллу Заборскому (Kirill Zaborsky) и анонимному пользователю Proton Mail, которые указали на ошибки в тексте;
- моей семьи, которая провела много выходных без меня, пока я был занят книгой.

От издательства

Ваши замечания, предложения, вопросы отправляйте по адресу comp@sprintbook.kz (издательство SprintBook, компьютерная редакция).

Мы будем рады узнать ваше мнение!

ГЛАВА 1

Типы

Первая глава книги содержит рекомендации, связанные с системой типов Rust. Его система типов более экспрессивна, чем ее аналоги в других популярных языках, и имеет больше общего с системой типов в академических языках вроде OCaml или Haskell. Одной из основных составляющих этой системы является тип `enum`, который значительно более выразителен, чем типы перечислений в других языках, и делает возможным использование *алгебраических типов данных*.

Рекомендации, приводимые в этой главе, посвящены фундаментальным типам, которые предоставляет Rust, и объясняют, как комбинировать их в структуры данных, точно выраждающие семантику вашей программы. Этот принцип кодирования поведения в систему типов помогает снизить объем проверок и кода, необходимого для выявления ошибок пути, поскольку недействительные состояния отклоняются цепочкой инструментов в процессе компиляции, а не программой в среде выполнения.

В этой главе также описываются некоторые распространенные структуры данных, предоставляемые стандартной библиотекой Rust: `Options`, `Results`, `Error` и `Iterators`. Знакомство с этими стандартными инструментами поможет вам писать идиоматический код Rust, отличающийся эффективностью и компактностью, — в частности, они позволяют использовать оператор вопросительного знака, дающий возможность обрабатывать ошибки необременительным, но при этом типобезопасным способом.

Имейте в виду, что рекомендации в отношении *трейтов* Rust мы рассмотрим в следующей главе, но они будут неизбежно пересекаться с рекомендациями этой главы, поскольку трейты описывают поведение типов.

Рекомендация 1. Используйте систему типов для выражения структур данных

Кто назвал их программистами, а не составителями типов?

@thingskatedid (<https://oreil.ly/hHj5c>)

В этом разделе мы кратко рассмотрим систему типов Rust, начиная с фундаментальных типов, которые предоставляет компилятор, и заканчивая различными способами совмещения значений в структуры данных.

Далее мы переключимся на тип `enum`. И хотя в базовой версии он равнозначен своим аналогам в других языках, возможность комбинировать вариации `enum` с полями данных обеспечивает повышенную гибкость и экспрессивность.

Фундаментальные типы

Основы системы типов Rust будут знакомы всем, кто имеет опыт работы со статически типизируемыми языками, например C++, Go или Java. Здесь мы имеем набор целочисленных типов конкретных размеров — как знаковых (`i8`, `i16`, `i32`, `i64`, `i128`), так и беззнаковых (`u8`, `u16`, `u32`, `u64`, `u128`).

Кроме того, есть знаковые (`isize`) и беззнаковые (`usize`) целые, чей размер соответствует размеру указателя в целевой системе. В Rust вам не потребуется часто заниматься преобразованием между указателями и целыми, так что равнозначность их размера значения не имеет. Тем не менее стандартные коллекции возвращают свой размер в виде `usize` (из `.len()`), поэтому при индексировании коллекции значения `usize` являются типичными, что естественно с точки зрения емкости, поскольку находящаяся в памяти коллекция не может содержать больше элементов, чем имеется адресов памяти в системе.

Целочисленные типы дают нам первое указание на то, что Rust устроен строже, чем C++. В Rust попытка поместить более крупный целочисленный тип (`i32`) в аналогичный тип меньшего размера (`i16`) ведет к ошибке при компиляции:

НЕ КОМПИЛИРУЕТСЯ

```
let x: i32 = 42;
let y: i16 = x;
```

```
error[E0308]: mismatched types
--> src/main.rs:18:18
```

```
18 |     let y: i16 = x;
|         --- ^ expected `i16`, found `i32`
|         |
|         expected due to this
|
help: you can convert an `i32` to an `i16` and panic if the converted value
doesn't fit
18 |     let y: i16 = x.try_into().unwrap();
|           ++++++  
+-----+  
+-----+
```

Такое поведение успокаивает: Rust не будет сидеть молча, пока программист делает что-то рискованное. И хотя мы можем видеть, что значения в этом конкретном преобразовании вполне для него подходят, компилятор должен позволять использовать такие, при которых преобразование будет *недопустимо*.

НЕ КОМПИЛИРУЕТСЯ

```
let x: i32 = 66_000;
let y: i16 = x; // Каким будет это значение?
```

Этот вывод ошибки также на раннем этапе показывает, что, несмотря на повышенную строгость Rust, он содержит полезные сообщения компилятора, которые указывают, как сделать все в соответствии с правилами. Предлагаемое решение поднимает вопрос о том, как быть в ситуациях, когда преобразованию приходится изменять значение, приводя его в соответствующий вид, и позже мы более подробно поговорим об обработке ошибок (см. рекомендацию 4) и использовании `panic!` (см. рекомендацию 18).

Rust также не позволяет выполнять кое-что, что может показаться безопасным, например помещать значение из меньшего целочисленного типа в больший:

НЕ КОМПИЛИРУЕТСЯ

```
let x = 42i32; // Целочисленный литерал с суффиксом типа
let y: i64 = x;
```

```
error[E0308]: mismatched types
--> src/main.rs:36:18
|
36 |     let y: i64 = x;
|         --- ^ expected `i64`, found `i32`
|         |
|         expected due to this
```

```

| help: you can convert an `i32` to an `i64`
| |
36 |     let y: i64 = x.into();
|     ++++++

```

Здесь предлагаемое решение не подразумевает обработку ошибки, но преобразование все равно должно быть явным. Мы более подробно поговорим о преобразованиях в рекомендации 5.

Список стандартных примитивных типов в Rust продолжают `bool`, типы с плавающей запятой (`f32`, `f64`) и единичный тип () (как `void` в C). Но более интересен символьный тип, который содержит значение Юникода (аналогичен типу `rune` в Go). И хотя он сохраняется в виде 4 байт, его также нельзя беспрепятственно преобразовать в 32-битное целое или из него.

Точность в системе типов вынуждает вас явно выражать свои намерения — значение `u32` отличается от `char`, которое отличается от последовательности байтов UTF-8, а она, в свою очередь, отличается от последовательности произвольных байтов, и именно вам нужно точно указывать, что имеется в виду¹. Лучше разобраться в том, какой тип в какой ситуации подходит, вам поможет известная статья Джоэля Спольски (Joel Spolsky) (<https://oreil.ly/wWy7T>).

Естественно, существуют вспомогательные методы, которые позволяют выполнять преобразование между этими типами, но их сигнатуры вынуждают вас обрабатывать (или открыто игнорировать) возможность сбоя. Например, кодовую точку Юникода всегда можно представить 32 битами², значит, '`'a'` as `u32` использовать допустимо, но в обратном направлении все несколько сложнее (есть значения `u32`, которые не являются валидными кодовыми точками Юникода).

`char::from_u32`

Возвращает `Option<char>`, вынуждая вызывающий код обрабатывать сбой.

`char::from_u32_unchecked`

Делает допущение в отношении валидности, но, если оно окажется ложным, это может привести к неопределенному поведению. В результате данная функция обозначается как `unsafe`, заставляя вызывающий код использовать также `unsafe` (см. рекомендацию 16).

¹ Ситуация становится еще более запутанной, если в процесс вовлечена файловая система, поскольку имена файлов на популярных платформах представляют собой нечто среднее между произвольными байтами и последовательностями UTF-8 (читайте документацию `std::ffi::OsString` по адресу <https://doc.rust-lang.org/std/ffi/struct.OsString.html>).

² Технически речь идет о скалярном значении Юникода (https://www.unicode.org/glossary/#unicode_scalar_value), а не о кодовой точке.

Составные типы

Далее идут составные типы. В Rust есть разные способы совмещения связанных значений. Большинство из них представляют собой знакомые эквиваленты механизмов агрегации, доступных в других языках.

Массивы

Содержат множество экземпляров одного типа, количество которых известно на этапе компиляции. Например, `[u32; 4]` — это четыре 4-байтовых целых подряд.

Кортежи

Содержат экземпляры разных типов, когда количество элементов и их типов известно на этапе компиляции, например `(WidgetOffset, WidgetSize, WidgetColor)`. Если типы в кортеже повторяются, например `(i32, i32, &'static str, bool)`, лучше дать каждому элементу имя и использовать структуру.

Структуры

Также содержат экземпляры разных типов, известные на этапе компиляции, но при этом позволяют использовать как общий тип, так и отдельные поля для обращения по имени.

Rust также включает в себя *структуру кортежа*, представляющую смесь `struct` и кортежа: в ней есть имя для общего типа, но нет имен для отдельных полей. Обращение к ним происходит по номеру — `s.0`, `s.1` и т. д.:

```
/// Структура с двумя неименованными полями
struct TextMatch(usize, String);

// Построение путем упорядоченного предоставления содержимого
let m = TextMatch(12, "needle".to_owned());

// Обращение по номеру поля
assert_eq!(m.0, 12);
```

Перечисления

Мы подошли к бриллианту в короне системы типов Rust — `enum`. Рассматривая базовую форму этого типа, сложно понять, что в нем такого восхитительного. Как и в других языках, `enum` позволяет указывать набор взаимно исключающих значений, в том числе с прикреплением числа:

```
enum HttpStatusCode {
    OK = 200,
    NotFound = 404,
```

```
Teapot = 418,  
}  
  
let code = HttpResultCode::NotFound;  
assert_eq!(code as i32, 404);
```

Поскольку каждое определение enum создает отдельный тип, это можно использовать для повышения читаемости и обслуживаемости функций, получающих аргументы bool. Используемая вместо:

```
print_page(/* both_sides= */ true, /* color= */ false);
```

версия с парой enum:

```
pub enum Sides {  
    Both,  
    Single,  
}  
  
pub enum Output {  
    BlackAndWhite,  
    Color,  
}  
  
pub fn print_page(sides: Sides, color: Output) {  
    // ...  
}
```

будет более типобезопасной и лучше читаемой в точке вызова:

```
print_page(Sides::Both, Output::BlackAndWhite);
```

В отличие от версии bool, если пользователь библиотеки случайно изменит порядок аргументов, компилятор тут же начнет ругаться:

```
error[E0308]: arguments to this function are incorrect  
--> src/main.rs:104:9  
|  
104 |     print_page(Output::BlackAndWhite, Sides::Single);  
|     ^^^^^^^^^^ ----- expected `enums::Output`,  
|             |  
|             |  
|             expected `enums::Sides`, found `enums::Output`  
|  
note: function defined here  
--> src/main.rs:145:12  
|  
145 |     pub fn print_page(sides: Sides, color: Output) {  
|     ^^^^^^^^^^ -----  
help: swap these arguments  
|  
104 |     print_page(Sides::Single, Output::BlackAndWhite);  
|     ~~~~~~
```

Используя паттерн newtype (читайте рекомендацию 6) для обертывания `bool`, вы получаете также безопасность типов и обслуживаемость. Лучше всего действовать этот паттерн, если семантика всегда будет логической, и применять `enum`, если есть вероятность возникновения в будущем новой альтернативы, например `Sides::BothAlternateOrientation`.

Безопасность типов при использовании `enum` в Rust сохраняется также за счет выражения `match`:

НЕ КОМПИЛИРУЕТСЯ

```
let msg = match code {
    HttpStatusCode::OK => "OK",
    HttpStatusCode::NotFound => "Not found",
    // Забыл обработать важный код "I'm a teapot"
};
```

```
error[E0004]: non-exhaustive patterns: `HttpStatusCode::Teapot` not covered
--> src/main.rs:44:21
|
44 |     let msg = match code {
|         ^^^^^ pattern `HttpStatusCode::Teapot` not covered
|
note: `HttpStatusCode` defined here
--> src/main.rs:10:5
|
7 | enum HttpStatusCode {
|     -----
...
10 |     Teapot = 418,
|         ^^^^^^ not covered
= note: the matched value is of type `HttpStatusCode`
  help: ensure that all possible cases are being handled by adding
        a match arm with a wildcard pattern or an explicit pattern as shown
|
46 ~         HttpStatusCode::NotFound => "Not found",
47 ~         HttpStatusCode::Teapot => todo!(),
|
```

Компилятор заставляет программиста учесть *все* возможности, предоставленные `enum`¹, даже если результатом будет просто добавление предустановленной

¹ Необходимость учитывать все возможные варианты также означает, что добавление нового варианта в существующее `enum` в библиотеке приведет к *критическому изменению* (см. рекомендацию 21): клиентам библиотеки придется изменять свой код, чтобы он обрабатывал и новый вариант. Если же `enum` является просто C-подобным списком связанных численных значений, такого поведения можно избежать, сделав его `non_exhaustive enum` (подробнее — в рекомендации 21).

ветки `_ => {}`. (Заметьте, что современные компиляторы C++ тоже предупреждают об отсутствии веток `switch` для `enum`.)

Перечисления с полями

Истинная сила `enum` в Rust основывается на том, что каждый содержащийся в нем вариант может иметь сопутствующие данные. Это делает перечисление составным типом, который выступает в качестве *алгебраического типа данных* (algebraic data type, ADT). Эта концепция может быть не знакома программистам, работающим с мейнстримными языками. Например, в системе C/C++ это подобно комбинации `enum` с `union`, только типобезопасной.

Это означает, что инварианты структур данных программы можно закодировать в систему типов Rust. Состояния, которые не соответствуют этим инвариантам, просто не скомпилируются. Грамотно построенное перечисление отчетливо демонстрирует намерение его создателя как для людей, так и для компилятора:

```
use std::collections::{HashMap, HashSet};

pub enum SchedulerState {
    Inert,
    Pending(HashSet<Job>),
    Running(HashMap<CpuId, Vec<Job>>),
}
```

Исходя из определения типа, разумно предположить, что, пока планировщик полностью активен, `Job` добавляются в очередь в состоянии `Pending` и призываются пулу того или иного ядра процессора. Это подчеркивает центральную тему данной рекомендации, которая заключается в использовании системы типов Rust для выражения принципов, связанных с проектированием вашего ПО.

Явным признаком случая, когда этого не происходит, является комментарий, который объясняет, когда определенное поле или параметр валидны.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
pub struct DisplayProps {
    pub x: u32,
    pub y: u32,
    pub monochrome: bool,
    // `fg_color` должен быть (0, 0, 0), если `monochrome` равен true
    pub fg_color: RgbColor,
}
```

Это первый кандидат на замену с помощью `enum`, содержащего данные:

```
pub enum Color {  
    Monochrome,  
    Foreground(RgbColor),  
}  
  
pub struct DisplayProps {  
    pub x: u32,  
    pub y: u32,  
    pub color: Color,  
}
```

Этот небольшой пример иллюстрирует суть рекомендации: *делайте невалидные состояния в ваших типах невыражаемыми*. Типы, которые поддерживают только валидные комбинации значений, свидетельствуют о том, что компилятор будет отвергать целый класс ошибок, обеспечивая более компактный и безопасный код.

Распространенные типы перечислений

Возвращаясь к потенциалу `enum`, следует сказать, что есть два настолько широко распространенных принципа, что стандартная библиотека Rust содержит типы `enum` для их выражения. В коде Rust эти типы встречаются повсеместно.

Option<T>

Первый принцип выражается в типе `Option`, который показывает, что значение определенного типа либо присутствует (`Some(T)`), либо нет (`None`). *Всегда используйте Option для значений, которые могут отсутствовать*. Никогда не прибегайте к использованию сигнальных значений (`-1`, `nullptr`, ...), стараясь выразить тот же принцип в рамках диапазона.

Но здесь нужно учитывать один нюанс. Если вы работаете с *коллекцией*, то следует решить, будет ли присутствие нуля элементов в ней равнозначно ее отсутствию. В большинстве ситуаций различий не будет, и вы вполне можете использовать, скажем, `Vec<Thing>`, когда ноль элементов подразумевает их отсутствие.

Тем не менее есть и другие, редкие сценарии, когда два этих случая нужно разделить с помощью `Option<Vec<Thing>>`, — например, криптографической системе может потребоваться различать полезную нагрузку, передаваемую отдельно (<https://oreil.ly/vuL006Co>), и пустую переданную полезную нагрузку. (Это относится к спорам вокруг маркера `NULL` для столбцов SQL.)

В том же смысле как лучше всего поступить в случае `String`, которая может отсутствовать? Как разумнее указывать на отсутствие значения — с помощью `""` или `None`? Сработает и то и другое, но `Option<String>` отчетливо покажет вероятность того, что это значение может отсутствовать.

Result<T, E>

Второй типичный принцип возникает в контексте обработки ошибок: если функция дает сбой, как об этом нужно сообщать? Традиционно для этого использовались специальные сигнальные значения (например, `-errno`, возвращаемое от системных вызовов в Linux) или глобальные переменные (`errno` в системах POSIX). В последние же годы языки, которые поддерживают множественные или кортежные возвращаемые значения (как в Go) от функций, могут по соглашению возвращать пару (`result, error`), предполагая существование некоего подходящего нулевого значения для `result`, когда `error` ненулевое.

В Rust конкретно для этого есть специальное enum: *всегда кодируйте результат операции, которая может провалиться, в виде Result<T, E>*. Тип `T` содержит успешный результат (в варианте `OK`), а тип `E` — детали об ошибке (в варианте `Err`) в случае сбоя.

Использование стандартного типа отчетливо проясняет замысел кода. Это, в частности, позволяет применять стандартные трансформации (см. рекомендацию 3) и обработку ошибок (см. рекомендацию 4), которые, в свою очередь, дают возможность упрощать обработку ошибок с помощью оператора `?`.

Рекомендация 2. Используйте систему типов для выражения типичного поведения

В рекомендации 1 мы разобрали, как в системе типов выражать структуры данных. Текущая же рекомендация посвящена кодированию в системе типов *поведения* Rust.

Описанные здесь механизмы покажутся знакомыми, поскольку имеют прямые аналоги в других языках.

Функции

Универсальный механизм для ассоциирования фрагмента кода с именем и списком параметров.

Методы

Функции, которые ассоциируются с экземпляром конкретной структуры данных. Методы типичны для языков, разработанных на основе парадигмы объектно-ориентированного программирования.

Указатели функций

Поддерживаются в большинстве языков семейства C, включая C++ и Go, в качестве механизма, позволяющего добавлять еще один уровень косвенности при вызове другого кода.

Замыкания

Изначально характерны преимущественно для семейства Lisp, но были модифицированы для многих популярных языков программирования, включая C++ (начиная с C++ 11) и Java (начиная с Java 8).

Трейты

Описывают коллекцию связанный функциональности, применяемой к одному элементу. Трейты имеют примерные эквиваленты во многих других языках, включая абстрактные классы в C++, а также интерфейсы в Go и Java.

Естественно, все эти механизмы в Rust имеют характерные особенности, которые будут раскрыты в данной рекомендации.

Из всего перечисленного списка трейты играют в книге наиболее важную роль, поскольку описывают значительную часть поведения, предоставляемого компилятором и стандартной библиотекой Rust. Глава 2 будет посвящена рекомендациям по проектированию и реализации трейтов, но то, что они встречаются повсеместно, означает, что они будут всплывать в других рекомендациях и текущей главы.

Функции и методы

Как и в любом другом языке программирования, в Rust *функции* служат для организации кода в виде именованных частей, предполагающих повторное использование и получающих входные данные в виде параметров. Как и в любом статически типизируемом языке, здесь типы параметров и возвращаемое значение указываются явно:

```
// Возвращает `x`, деленное на `y`
fn div(x: f64, y: f64) -> f64 {
    if y == 0.0 {
        // Завершает функцию и возвращает значение
        return f64::NAN;
    }
    // Неявно возвращается последнее выражение в теле функции
    x / y
}

/// Функция вызывается только ради своих побочных эффектов
/// без возвращаемого значения
/// Можно также записать возвращаемое значение как`-> ()` 
fn show(x: f64) {
    println!("x = {}", x);
}
```

Если функция тесно вписана в какую-то структуру данных, она выражается в виде *метода*. Он действует в отношении элемента этого типа, определяемого `self`, и включается в блок `impl DataStructure`. Таким образом связанные данные и код по

анalogии с другими языками инкапсулируются вместе объектно-ориентированным способом. Тем не менее в Rust методы можно добавлять в типы `enum` и `struct`, сохраняя всепроникающую природу `enum` в этом языке (см. рекомендацию 1):

```
enum Shape {
    Rectangle { width: f64, height: f64 },
    Circle { radius: f64 },
}

impl Shape {
    pub fn area(&self) -> f64 {
        match self {
            Shape::Rectangle { width, height } => width * height,
            Shape::Circle { radius } => std::f64::consts::PI * radius * radius,
        }
    }
}
```

Имя метода создает метку кодируемого им поведения, а его сигнатура несет информацию о типе его ввода и вывода. Первым вводом метода будет некий вариант `self`, указывающий, что метод может делать с этой структурой данных.

- Параметр `&self` указывает, что содержимое структуры данных можно считывать, но нельзя изменять.
- Параметр `&mut self` указывает, что метод может изменять содержимое структуры данных.
- Параметр `self` указывает, что метод получает структуру данных.

Указатели функций

В предыдущем разделе было описано, как ассоциировать имя и список параметров с неким кодом. Однако вызов функции всегда приводит к выполнению одного и того же кода. Между вызовами меняются только данные, с которыми работает функция. Это охватывает множество возможных сценариев, но что, если в среде выполнения код должен изменяться?

Простейшей поведенческой абстракцией, которая позволяет это реализовать, является *указатель функции* — указатель только на какой-то код с типом, отражающим сигнатуру его функции:

```
fn sum(x: i32, y: i32) -> i32 {
    x + y
}
// Требуется явное приведение к типу `fn`...
let op: fn(i32, i32) -> i32 = sum;
```

Тип проверяется на этапе компиляции, поэтому к моменту выполнения программы его значение имеет размер указателя. С указателями функций никакие другие данные не ассоциируются, так что их можно рассматривать как значения по-разному:

```
// Типы `fn` реализуют `Copy`
let op1 = op;
let op2 = op;
// Типы `fn` реализуют `Eq`
assert!(op1 == op2);
// `fn` реализует `std::fmt::Pointer`,
/// используемый спецификатором формата `{:p}`.
println!("op = {:p}", op);
// Пример вывода: "op = 0x101e9aeb0"
```

Здесь нужно иметь в виду один технический нюанс — необходимо явное приведение к типу `fn`, поскольку просто использование имени функции *не даст* вам чего-либо с типом `fn`:

НЕ КОМПИЛИРУЕТСЯ

```
let op1 = sum;
let op2 = sum;
// op1 и op2 имеют тип, который в коде пользователя
// прописать невозможно. И этот внутренний тип не реализует `Eq`
assert!(op1 == op2);
```

```
error[E0369]: binary operation `==` cannot be applied to type
  `fn(i32, i32) -> i32 {main::sum}`
--> src/main.rs:102:17
|
102 |     assert!(op1 == op2);
|     --- ^-- fn(i32, i32) -> i32 {main::sum}
|     |
|     fn(i32, i32) -> i32 {main::sum}
|
help: use parentheses to call these
|
102 |     assert!(op1(/* i32 */, /* i32 */) == op2(/* i32 */, /* i32 */));
|     ++++++ ++++++ ++++++ ++++++ ++++++ ++++++
```

Вместо этого выданная компилятором ошибка указывает на то, что здесь у нас тип `fn(i32, i32) -> i32 {main::sum}`, который используется компилятором исключительно внутри (а значит, прописать такой в коде пользователя нельзя) и указывает на конкретную функцию и ее сигнатуру. Иными словами, *тын sum*

в целях оптимизации кодирует сигнатуру функции и ее расположение (<https://oreil.ly/HWxcl>). Этот тип может быть автоматически приведен к типу `fn` (см. рекомендацию 5).

Замыкания

Голые указатели функций весьма ограничены, поскольку единственный возможный ввод для вызываемой функции заключается в явно передаваемых значениях параметров. Рассмотрим в качестве примера код, изменяющий каждый элемент среза с помощью указателя функции:

```
// В реальном коде более подходящим будет метод `Iterator`
pub fn modify_all(data: &mut [u32], mutator: fn(u32) -> u32) {
    for value in data {
        *value = mutator(*value);
    }
}
```

Это сработает для простой мутации среза:

```
fn add2(v: u32) -> u32 {
    v + 2
}
let mut data = vec![1, 2, 3];
modify_all(&mut data, add2);
assert_eq!(data, vec![3, 4, 5]);
```

Тем не менее, если изменение опирается на какое-либо дополнительное состояние, передать его в указатель функции в неявной форме невозможно:

НЕ КОМПИЛИРУЕТСЯ

```
let amount_to_add = 3;
fn add_n(v: u32) -> u32 {
    v + amount_to_add
}
let mut data = vec![1, 2, 3];
modify_all(&mut data, add_n);
assert_eq!(data, vec![3, 4, 5]);
```

```
error[E0434]: can't capture dynamic environment in a fn item
--> src/main.rs:125:13
|
125 |         v + amount_to_add
|             ^^^^^^^^^^^^^^
|
= help: use the `|| { ... }` closure form instead
```

Сообщение об ошибке указывает на подходящий для этой задачи инструмент — *замыкание (closure)*. Так называется фрагмент кода, похожий на определение функции (*лямбда-выражение*), за исключением следующего.

- Его можно создать в рамках выражения, а значит, ему не требуется отдельное имя.
- Входные параметры в нем указываются внутри вертикальных черт `|param1, param2|` (связанные с ними типы обычно выводятся компилятором автоматически).
- Замыкания могут включать часть окружающей их среды:

```
let amount_to_add = 3;
let add_n = |y| {
    // Замыкание, включающее `amount_to_add`
    y + amount_to_add
};
let z = add_n(5);
assert_eq!(z, 8);
```

Чтобы примерно понять принцип их работы, представьте, что компилятор создает одноразовый внутренний тип, содержащий все части среды, упоминаемой в лямбда-выражении. При создании замыкания для хранения нужных значений создается также экземпляр временного типа, а при его вызове этот экземпляр используется в качестве дополнительного контекста:

```
let amount_to_add = 3;
// Примерный эквивалент включающего контекст замыкания
struct InternalContext<'a> {
    // Ссылки на включенные переменные
    amount_to_add: &'a u32,
}
impl<'a> InternalContext<'a> {
    fn internal_op(&self, y: u32) -> u32 {
        // Тело лямбда-выражения
        y + *self.amount_to_add
    }
}
let add_n = InternalContext {
    amount_to_add: &amount_to_add,
};
let z = add_n.internal_op(5);
assert_eq!(z, 8);
```

Значения, хранящиеся в этом гипотетическом контексте, зачастую являются ссылками (см. рекомендацию 8), как в данном случае, но также могут быть мутабельными ссылками на элементы их окружения либо значениями, полностью перемещаемыми из среды путем указания ключевого слова `move` перед входными параметрами.

Возвращаясь к примеру `modify_all`, отмечу, что замыкание нельзя использовать там, где ожидается указатель функции:

```
error[E0308]: mismatched types
--> src/main.rs:199:31
|
199 |     modify_all(&mut data, |y| y + amount_to_add);
|-----|                                         ^^^^^^^^^^ expected fn pointer,
|                                         |                                         found closure
|                                         |
|                                         arguments to this function are incorrect
|
= note: expected fn pointer `fn(u32) -> u32`
      found closure `[closure@src/main.rs:199:31: 199:34]`
note: closures can only be coerced to `fn` types if they do not capture any
      variables
--> src/main.rs:199:39
|
199 |     modify_all(&mut data, |y| y + amount_to_add);
|-----|                                         ^^^^^^ `amount_to_add` captured here
note: function defined here
--> src/main.rs:60:12
|
60  |     pub fn modify_all(data: &mut [u32], mutator: fn(u32) -> u32) {
```

Вместо этого код, который получает замыкание, должен получать экземпляр одного из трейтов Fn*:

```
pub fn modify_all<F>(data: &mut [u32], mut mutator: F)
where
    F: FnMut(u32) -> u32,
{
    for value in data {
        *value = mutator(*value);
    }
}
```

В Rust есть три разных трейта Fn*, показывающих некоторые различия в описанном поведении, связанном с включением элементов окружения.

FnOnce

Описывает замыкание, которое можно вызвать только *один раз*. Если какая-то часть среди `move` перемещается в контекст замыкания, после чего тело замыкания из своего контекста его выносит, то такое перемещение может происходить только раз — не существует другой копии исходного элемента, из которого можно было бы выполнить `move`, поэтому замыкание можно вызвать лишь единожды.

FnMut

Описывает замыкание, которое можно вызывать неоднократно и которое может изменять свою среду, поскольку выполняет из нее заимствование с возможностью *мутации*.

Fn

Описывает замыкание, которое можно вызывать неоднократно и которое заимствует значения из среды только немутабельно.

Компилятор *автоматически* реализует подходящее подмножество этих трейтов Fn* для любого лямбда-выражения в коде. Реализовать какой-либо из этих трейтов вручную невозможно, в отличие от перегрузки `operator()` в C++¹.

Вернемся к рассмотренной модели замыканий. То, какой из трейтов компилятор автоматически реализует, примерно определяется тем, содержит ли включенный контекст окружения следующие элементы.

FnOnce

Любые перемещаемые значения.

FnMut

Любые мутабельные ссылки на значения (&mut T).

Fn

Только формальные ссылки на значения (&T).

Последние два трейта этого списка граничат с предшествующим им, что становится актуальным при рассмотрении элементов, *использующих* замыкания.

- Если что-либо ожидает лишь однократного вызова замыкания (обозначается получением FnOnce), вполне можно передать ему замыкание, допускающее неоднократный вызов (FnMut).
- Если что-либо ожидает повторяющегося вызова замыкания, которое может изменять свое окружение (указывается получением FnMut), допустимо передать ему замыкание, которому свое окружение менять *не нужно* (Fn).

Голый тип указателя функции fn также условно относится к концу этого списка. Любой не-unsafe-тип fn автоматически реализует все трейты Fn*, так как ничего из среды не заимствует.

В результате при написании кода, получающего замыкание, *рекомендуется использовать наиболее обобщенный подходящий трейт Fn**, чтобы получить наибольшую гибкость для вызывающих. Например, примените FnOnce для

¹ На момент написания книги как минимум не в стабильной версии Rust. Экспериментальная функциональность unboxed_closures и fn_traits может в будущем это изменить.

замыканий, которые используются лишь раз. По той же логике рекомендуется *отдавать предпочтение границам трейта Fn* перед голыми указателями функций (fn)*.

Трейты

Трейты Fn* более гибкие по сравнению с голыми указателями функций, но по-прежнему могут описывать поведение лишь одной функции, и то лишь в плане ее сигнатуры. Тем не менее они сами являются примером еще одного механизма описания поведения в системе типов Rust — *трейта*. Трейт определяет набор связанных функций, которые некий внутренний элемент делает публично доступными. Более того, эти функции обычно (но не обязательно) являются *методами*, получая в качестве первого аргумента вариацию `self`.

У каждой функции в трейте есть *имя*, выступающее некой меткой, которая не только помогает компилятору разрешать проблему омонимии в случае совпадения сигнатур функций, но и, что самое главное, позволяет программистам прослеживать смысл функции.

Трейты в Rust в некотором роде аналогичны интерфейсам в Go и Java либо абстрактным классам (полностью виртуальные методы без членов данных) в C++. Реализации трейта должны предоставлять все функции (но имейте в виду, что определение трейта может включать предустановленную реализацию; см. рекомендацию 13), а также иметь связанные данные, которые эта реализация может использовать. Это означает, что код и данные инкапсулируются вместе в общей абстракции *особым* объектно-ориентированным образом.

Код, который получает структуру и вызывает для нее функции, ограничивается работой только с этим конкретным типом. Если существует несколько типов, реализующих общее поведение, тогда более гибким решением будет определить инкапсулирующий его трейт и заставить код использовать функции этого трейта, а не те, которые связаны с конкретной структурой.

Все это ведет к той же рекомендации, которая касается и других опирающихся на объектное ориентирование языков¹: если в будущем потребуется гибкость, то лучше получать типы трейтов, а не конкретные типы.

Порой есть поведение, которое вам нужно выделить в системе типов, но которое нельзя выразить в виде конкретной сигнатуры функции в определении трейта. Рассмотрим, к примеру, трейт `Sort` для упорядочения коллекций. Его реализация может быть *стабильной* (элементы, которые при сравнении оцениваются как одинаковые, будут располагаться в одном порядке до и после упорядочения), но нет способа выразить это в аргументах метода `sort`.

¹ Например, в книге Джошуа Блоха (Joshua Bloch) «Java. Эффективное программирование» (3-е изд.) есть рекомендация 64: «Обращайтесь к объектам по их интерфейсам».

В данном случае для соблюдения этого требования по-прежнему желательно использовать систему типов, задействуя *трейт-маркер*:

```
pub trait Sort {  
    /// Упорядочивает содержимое  
    fn sort(&mut self);  
}  
  
/// Трейт-маркер, указывающий, что [ `Sort` ] стабильно  
/// выполняет упорядочение  
pub trait StableSort: Sort {}
```

Трейт-маркер не содержит функций, но реализация все равно должна объявлять его присутствие. В этом случае он выступает в качестве своеобразного обещания от реализующего: «Клянусь, моя реализация выполняет сортировку стабильно». Тогда код, опирающийся на стабильную сортировку, сможет указать границу трейта `StableSort` с уверенностью в том, что инварианты будут сохранены. *Используйте трейты-маркеры для различения поведения, которое нельзя выразить в сигнатурах функций трейта.*

После того как поведение было инкапсулировано в системе типов Rust в виде трейта, его можно использовать двумя способами:

- в качестве *границы трейта*, определяющей, какие типы будут приемлемы для обобщенного типа данных или функции на этапе компиляции;
- в виде *объекта трейта*, определяющего, какие типы можно сохранять либо передавать в функцию в среде выполнения.

Обе эти возможности описаны в следующем разделе, а в рекомендации 12 более детально раскрыты компромиссы между ними.

Границы трейтов

Граница трейта указывает, что обобщенный код, параметризованный неким типом `T`, можно использовать, только когда этот тип реализует конкретный трейт. Наличие границы трейта означает, что реализация этого обобщения может задействовать функции из данного трейта, зная, что компилятор проследит за присутствием этих функций во всех компилируемых типах `T`. Проверка происходит на этапе компиляции, когда обобщенный код *мономорфизируется*, то есть превращается из обобщенной формы, работающей с произвольным типом `T`, в конкретный код, работающий с конкретным `SomeType` (в C++ это называется *конкретизацией шаблона*).

Это *явное ограничение* целевого типа `T`, оно заложено в границах трейта: реализовывать трейт могут только типы, соответствующие его границам. Такое поведение противоположно аналогичной ситуации в C++, где ограничения

типа T , используемого в `template<typename T>`, неявные¹: шаблонный код в C++ по-прежнему компилируется только при доступности всех функций, на которые есть ссылки, но проверяется это исключительно по имени и сигнатуре функций. (Такая «утинная типизация» может вызывать путаницу: шаблон C++, применяющий `t.pop()`, может компилироваться для параметра типа T , представленного как `Stack` или `Balloon`, что вряд ли окажется желаемым поведением.)

Потребность в явных границах трейтов означает также, что их использует значительная часть обобщений. Чтобы понять, почему так происходит, посмотрим с другой стороны и подумаем, что можно сделать со `struct Thing<T>`, где *нет* никаких границ трейта для T . При их отсутствии `Thing` может выполнять только операции, применимые к *любому* типу T , по сути просто перемещая или отбрасывая значение. Это, в свою очередь, позволяет задействовать обобщенные контейнеры, коллекции и умные указатели, но ничего другого. Для всего, что *использует* тип T , потребуется граница трейта:

```
pub fn dump_sorted<T>(mut collection: T)
where
    T: Sort + IntoIterator,
    T::Item: std::fmt::Debug,
{
    // Следующей строке нужна граница трейта `T: Sort`
    collection.sort();
    // Следующей строке нужна граница трейта `T: IntoIterator`
    for item in collection {
        // Следующей строке нужна граница трейта `T::Item : Debug`
        println!("{}:?", item);
    }
}
```

Итак, советом здесь будет *использовать границы трейтов для выражения требований к используемым в обобщениях типам*. Ему легко следовать, так как компилятор в любом случае вас к этому принудит.

Трейт-объекты

Трейт-объект — это еще один способ задействовать определяемую трейтом инкапсуляцию, но в данном случае возможные реализации трейта выбираются в среде выполнения, а не при компиляции. Такая *динамическая диспетчеризация* схожа с использованием виртуальных функций в C++, и в Rust есть объект `vtable` (таблица диспетчеризации), *примерно* соответствующий его аналогам в C++.

Этот динамический аспект трейт-объектов означает также, что их всегда нужно обрабатывать косвенно посредством некой ссылки (например, `&dyn Trait`) или указателя (например, `Box<dyn Trait>`). Причина заключается в том, что размер

¹ Добавление концепций (<https://oreil.ly/ueIhM>) в C++ 20 сделало возможным явное указание ограничений типов шаблонов, но проверки по-прежнему выполняются, только когда шаблон инстанцируется, а не когда объявляется.

реализующего трейт объекта на этапе компиляции неизвестен — это может быть как гигантская структура, так и миниатюрное перечисление, поэтому под голый трейт-объект невозможно выделить правильное количество места.

Незнание размера конкретного объекта означает также, что используемые в качестве объектов трейты не могут содержать функции, возвращающие тип `Self` или аргументы (за исключением *получателя* — объекта, для которого вызывается метод), которые этот `Self` используют. Причина в том, что компилируемый заранее код, который использует трейт-объект, понятия не имеет, каким может быть размер `Self`.

Трейт, имеющий обобщенную функцию `fn some_fn<T>(t:T)`, делает возможным бесконечное число реализуемых функций для всех типов `T`, которые могут существовать. И это нормально для трейта, применяемого в качестве границы, поскольку бесконечное множество, возможно, вызываемых обобщенных функций становится конечным множеством обобщенных функций, фактически вызываемых на этапе компиляции. Но это не касается трейт-объекта, так как код, доступный при компиляции, должен уметь работать со всеми `T`, которые могут встретиться при выполнении.

Эти два ограничения — не использовать `Self` и обобщенные функции — объединяются в принцип *объектной безопасности* (<https://oreil.ly/gaq4I>). Только объектно-безопасные трейты можно применять в качестве трейт-объектов.

Рекомендация 3. Предпочитайте преобразования Option и Result вместо явных выражений match

В рекомендации 1 мы разобрали достоинства `enum` и показали, что выражения `match` вынуждают программистов учитывать все возможные варианты. Там же были представлены два распространенных `enum`, присутствующих в стандартной библиотеке Rust.

Option<T>

Позволяет показать, что значение, имеющее тип `T`, может либо присутствовать, либо нет.

Result<T, E>

Для случаев, когда операция возвращения значения, имеющего тип `T`, может провалиться и вернуть ошибку, имеющую тип `E`.

Здесь же мы разберем ситуации, когда следует избегать использования явных выражений `match` для этих конкретных `enum`, предпочитая им различные методы преобразования, предоставляемые стандартной библиотекой для данных типов. Применение этих методов преобразования (которые сами обычно реализуются в виде выражений `match`) позволяет создать компактный идиоматический код, смысл которого более отчетлив.

Первый случай, когда `match` необязателен, — это когда важно только значение, а его отсутствие или связанную с этим ошибку можно просто проигнорировать:

```
struct S {
    field: Option<i32>,
}

let s = S { field: Some(42) };
match &s.field {
    Some(i) => println!("field is {}", i),
    None => {}
}
```

В этой ситуации выражение `if let` оказывается на одну строку короче и, что самое главное, яснее:

```
if let Some(i) = &s.field {
    println!("field is {}", i);
}
```

Тем не менее чаще всего программисту нужно предоставить соответствующую ветку `else` — отсутствие значения (`Option::None`), возможно, вместе с сопутствующей ошибкой (`Result::Err(e)`). Проектировать обработку путей сбоя в ПО нелегко, и основная сложность заключается в том, что здесь не поможет никакая синтаксическая поддержка — в частности, в принятии решения о том, что должно произойти, если операция провалится.

В некоторых ситуациях правильным решением будет просто закрыть на это глаза и никак сбой не обрабатывать. Вы не можете *полностью* проигнорировать ветку ошибки, поскольку Rust требует, чтобы код обрабатывал оба варианта `Result`, но при этом можно рассмотреть сбой как фатальный. Выполнение при ошибке `panic!` означает, что программа завершается, но остальной код можно написать так, будто все протекает успешно. Такой маневр в отношении явного `match` окажется излишне громоздким:

```
let result = std::fs::File::open("/etc/passwd");
let f = match result {
    Ok(f) => f,
    Err(_e) => panic!("Failed to open /etc/passwd!"),
};
// Предполагаем, что далее `f` является валидным `std::fs::File`
```

И `Option`, и `Result` предоставляют пару методов `unwrap` и `expect`, которые извлекают свое внутреннее значение и выполняют `panic!`, если оно отсутствует. Второй из них позволяет персонализировать сообщение об ошибке при сбое, но в любом случае итоговый код оказывается короче и проще — обработка ошибок по-прежнему присутствует, хоть и делегируется суффиксу `.unwrap()`:

```
let f = std::fs::File::open("/etc/passwd").unwrap();
```

Тем не менее нужно ясно понимать — эти вспомогательные функции все равно `panic!`, поэтому их выбор равнозначен применению `panic!` (см. рекомендацию 18).

Однако во многих ситуациях обработки ошибки правильным решением будет оставить ее кому-то другому, кто столкнется с этим кодом позднее. Это особенно актуально при написании библиотеки, когда код может использоваться в различных средах, которых автор библиотеки предвидеть не может. Чтобы упростить другому разработчику задачу, рекомендуется показывать ошибки с помощью `Result`, а не `Option`, несмотря на то что это может подразумевать преобразование между различными видами сбоев (см. рекомендацию 4).

Естественно, тогда возникает вопрос: «А что считать ошибкой?» В данном примере ошибкой определено будет сбой при открытии файла, и детали этой ошибки (файл отсутствует? Нет нужного разрешения?) помогут пользователю решить, что делать дальше. В то же время сбой при извлечении `first()` элемента среза из-за того, что срез пуст, ошибкой не будет, поэтому в стандартной библиотеке он отражается в виде возвращаемого типа `Option`. Выбор между двумя этими вариантами требует обдумывания, но если ошибка может сообщать полезную информацию, то лучше применять `Result`.

`Result` также имеет атрибут `#![must_use]`, направляющий пользователей библиотеки в правильную сторону, — если задействующий `Result` код его проигнорирует, компилятор выдаст предупреждение:

```
warning: unused `Result` that must be used
--> src/main.rs:63:5
   |
63 |     f.set_len(0); // Сократите файл
   | ^^^^^^^^^^^^^^
   |
   = note: this `Result` may be an `Err` variant, which should be handled
   = note: `#[warn(unused_must_use)]` on by default
help: use `let _ = ...` to ignore the resulting value
   |
63 |     let _ = f.set_len(0); // Сократите файл
   | +++++++
```

Явное использование `match` позволяет передавать информацию об ошибке вызывающей функции, но требует дополнительного шаблонного кода (напоминает Go):

```
pub fn find_user(username: &str) -> Result<UserId, std::io::Error> {
    let f = match std::fs::File::open("/etc/passwd") {
        Ok(f) => f,
        Err(e) => return Err(From::from(e)),
    };
    // ...
}
```

В данном случае ключевым компонентом для сокращения шаблонного кода в Rust является оператор ?. Этот элемент синтаксического сахара берет на себя обработку ветки Err, при необходимости преобразуя ошибку в нужный вид и создавая выражение `return Err(...)`:

```
pub fn find_user(username: &str) -> Result<UserId, std::io::Error> {
    let f = std::fs::File::open("/etc/passwd")?;
    // ...
}
```

Новичков в Rust это порой дезориентирует: вопросительный знак не так просто заметить, из-за чего возникает неуверенность в том, как код может работать. Тем не менее даже в случае одного символа система типов продолжает делать свое дело, следя за тем, чтобы обрабатывались все возможности, выраженные в соответствующих типах (см. рекомендацию 1). В итоге программист может, не отвлекаясь, сосредоточиться на основном пути кода.

Более того, обычно эти очевидные вызовы методов не требуют никаких затрат — они все являются обобщенными функциями, отмеченными как #[*inline*], в связи с чем генерируемый код, как правило, компилируется в машинный, идентичный ручной версии.

Два этих фактора означают, что вместо явных выражений `match` следует использовать `Option` и `Result`.

В предыдущем примере типы ошибок согласовались — внутренние и внешние методы выражали сбои как `std::io::Error`. Но зачастую это не так: одна функция может аккумулировать ошибки из разных вторичных библиотек, каждая из которых использует свои типы ошибок.

Отображение ошибок в целом рассматривается в рекомендации 4. Сейчас же просто имейте в виду, что ручное отображение:

```
pub fn find_user(username: &str) -> Result<UserId, String> {
    let f = match std::fs::File::open("/etc/passwd") {
        Ok(f) => f,
        Err(e) => {
            return Err(format!("Failed to open password file: {:?}", e))
        }
    };
    // ...
}
```

можно более лаконично и идиоматично выразить с помощью следующей трансформации `.map_err()`:

```
pub fn find_user(username: &str) -> Result<UserId, String> {
    let f = std::fs::File::open("/etc/passwd")
```

```

    .map_err(|e| format!("Failed to open password file: {:?}", e))?;
    // ...
}

```

Но даже это может оказаться необязательным — если тип внешней ошибки может быть создан на основе внутренней путем реализации стандартного трейта `From` (см. рекомендацию 10), то компилятор выполнит это преобразование автоматически, не вызывая `.map_err()`.

Такой вид преобразования действует более широко. Оператор вопросительного знака подобен большому молотку. Поэтому используйте методы преобразования для типов `Option` и `Result`, чтобы разместить их там, где они будут подобны гвоздям. Для этого в стандартной библиотеке есть множество методов преобразования. На рис. 1.1 показаны некоторые наиболее распространенные (в светлых рамках), которые выполняют преобразование между типами (темные рамки). В соответствии с рекомендацией 18 методы, которые могут `panic!`, отмечены звездочкой.

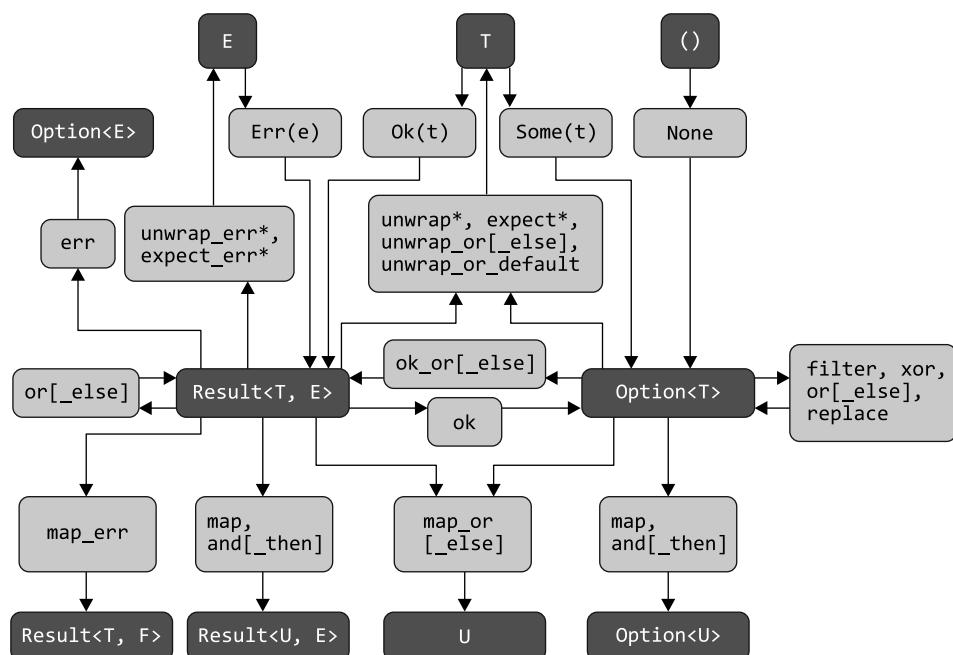


Рис. 1.1. Преобразования Option и Result¹

¹ Онлайн-версия этой схемы (https://oreil.ly/effective_rust_transforms) является кликабельной — каждая рамка в ней ведет к соответствующей документации.

Одна из типичных ситуаций, которую не охватывает эта схема, связана со ссылками. Рассмотрим в качестве примера структуру, которая опционально содержит некие данные:

```
struct InputData {
    payload: Option<Vec<u8>>,
}
```

Метод в этой `struct`, пытающийся передать полезную нагрузку (`payload`) в функцию шифрования с сигнатурой `(&[u8]) -> Vec<u8>`, проваливается, если происходит наивная попытка получить ссылку:

НЕ КОМПИЛИРУЕТСЯ

```
impl InputData {
    pub fn encrypted(&self) -> Vec<u8> {
        encrypt(&self.payload.unwrap_or(vec![]))
    }
}
```

```
error[E0507]: cannot move out of `self.payload` which is behind a shared reference
--> src/main.rs:15:18
   |
15 |     encrypt(&self.payload.unwrap_or(vec![]))
   |     ^^^^^^^^^^ move occurs because `self.payload` has type
   |           `Option<Vec<u8>>`, which does not implement the
   |           `Copy` trait
```

Правильным инструментом для этого будет использование в `Option` метода `as_ref()`¹. Данный метод преобразует ссылку на `Option` в `Option` ссылки:

```
pub fn encrypted(&self) -> Vec<u8> {
    encrypt(self.payload.as_ref().unwrap_or(&vec![]))
}
```

Запомните

- Привыкните преобразовывать `Option` и `Result`, а также отдавайте предпочтение `Result` перед `Option`. Когда преобразования включают ссылки, используйте `.as_ref()`.
- Предпочитайте эти преобразования явным операциям `match` в `Option` и `Result`.
- В частности, используйте эти преобразования для трансформации типов результата в форму, где можно применить оператор `?`.

¹ Заметьте, что этот метод отделен от трейта `AsRef`, несмотря на то что имеет такое же имя.

Рекомендация 4. Предпочитайте идиоматические типы Error

В рекомендации 3 рассказывалось, как использовать предоставляемые стандартной библиотекой преобразования для типов `Option` и `Result`, чтобы лаконично и идиоматично обрабатывать итоговые типы с помощью оператора `?`. В ней мы не стали обсуждать, как лучше всего обрабатывать различные типы ошибок `E`, которые возникают в качестве второго аргумента типа из `Result<T, E>`. Об этом речь пойдет в данной рекомендации.

Эта рекомендация актуальна, только когда в процессе задействовано несколько типов ошибок. Если все ошибки, с которыми может столкнуться функция, уже имеют один тип, она может просто возвращать его. Если же присутствуют ошибки разных типов, то нужно принимать решение о том, должна ли сохраняться информация о типе дополнительной ошибки.

Трейт `Error`

Всегда полезно понимать, какие стандартные трейты используются (см. рекомендацию 10), и в данном случае актуальным трейтом является `std::error::Error`. Параметр `Result` с типом `E` *не обязательно* должен иметь тип, реализующий `Error`, но это обычное соглашение, которое позволяет оберткам показывать фактические границы трейта, поэтому *лучше реализовывать Error для типов ошибок*.

Первым делом нужно заметить, что единственным жестким требованием для типов `Error` являются границы трейта — любой тип, реализующий `Error`, должен также реализовывать следующие трейты:

- `Display`, означающий, что к нему можно применить `format!` с помощью `{}`;
- `Debug`, означающий, что к нему можно применить `format!` с помощью `{:?}`.

Иными словами, он должен уметь отображать типы `Error` как для пользователя, так и для программиста.

В этом трейте присутствует единственный метод `source()`¹, который позволяет типу `Error` выражать внутреннюю, вложенную ошибку. Это необязательный метод — он присутствует в предустановленной реализации (см. рекомендацию 13) и возвращает `None`, указывая на недоступность информации о внутренней ошибке.

¹ Или по крайней мере единственный неустаревший стабильный метод.

И последний нюанс: если вы пишете код для среды `no_std` (см. рекомендацию 33), то можете не иметь возможности реализовать `Error` — трейт `Error` на данный момент реализован в `std`, но не в `core`, в связи с чем недоступен¹.

Минимальные ошибки

Если вложенная информация не требуется, тогда реализации типа `Error` достаточно простой `String` — единственный редкий случай, в котором может подойти переменная, типизированная в виде строки. Однако это должно быть *чуть больше*, чем просто строка, хотя в качестве параметра с типом `E` можно использовать `String`:

```
pub fn find_user(username: &str) -> Result<UserId, String> {
    let f = std::fs::File::open("/etc/passwd")
        .map_err(|e| format!("Failed to open password file: {:?}", e))?;
    // ...
}
```

`String` не реализует `Error`, что для нас даже лучше, поскольку так с `Error` могут работать и другие области кода. Невозможно использовать `impl Error` для `String`, поскольку ни этот трейт, ни тип нам не принадлежат (так называемое *правило сироты*):

НЕ КОМПИЛИРУЕТСЯ

```
impl std::error::Error for String {}
```

```
error[E0117]: only traits defined in the current crate can be implemented
              for types defined outside of the crate
--> src/main.rs:18:5
   |
18 |     impl std::error::Error for String {}
   |     ^^^^^^^^^^^^^^^^^^^^^^^^^^-----|
   |         |
   |             |
   |             `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
```

И *псевдоним типа* здесь тоже не поможет, поскольку он не создает новый тип, а значит, сообщение об ошибке не изменяет:

¹ На момент написания книги `Error` переместили в `core` (<https://oreil.ly/lI0vv>), но пока в стабильной версии Rust он недоступен.

НЕ КОМПИЛИРУЕТСЯ

```
pub type MyError = String;

impl std::error::Error for MyError {}
```

```
error[E0117]: only traits defined in the current crate can be implemented for
              types defined outside of the crate
--> src/main.rs:41:5
41 |     impl std::error::Error for MyError {}
   |     ^^^^^^^^^^^^^^^^^^^^^^-----|
   |     |                         |
   |     |                         `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
```

Как обычно, сообщение компилятора об ошибке дает нам подсказку для решения проблемы. Определение структуры кортежа, которая обертывает тип `String` (паттерн `newtype`; см. рекомендацию 6), позволяет реализовать трейт `Error` при условии, что также будут реализованы `Debug` и `Display`:

```
#[derive(Debug)]
pub struct MyError(String);

impl std::fmt::Display for MyError {
    fn fmt(&self, f: &mut std::fmt::Formatter<'_>) -> std::fmt::Result {
        write!(f, "{}", self.0)
    }
}

impl std::error::Error for MyError {}

pub fn find_user(username: &str) -> Result<UserId, MyError> {
    let f = std::fs::File::open("/etc/passwd").map_err(|e| {
        MyError(format!("Failed to open password file: {:?}", e))
    })?;
    // ...
}
```

Для удобства может оказаться уместной реализация трейта `From<String>`, который позволит преобразовывать строковые значения в экземпляры `MyError` (см. рекомендацию 5):

```
impl From<String> for MyError {
    fn from(msg: String) -> Self {
        Self(msg)
    }
}
```

Когда компилятор встречает вопросительный знак (?), он автоматически применяет любые подходящие реализации трейта `From`, необходимые для получения целевого типа возвращаемой ошибки. Это позволяет дополнительно минимизировать код:

```
pub fn find_user(username: &str) -> Result<UserId, MyError> {
    let f = std::fs::File::open("/etc/passwd")
        .map_err(|e| format!("Failed to open password file: {:?}", e))?;
    // ...
}
```

Путь ошибки здесь включает следующие шаги.

1. `File::open` возвращает ошибку, имеющую тип `std::io::Error`.
2. `format!` преобразует ее в `String` с помощью реализации `Debug` из `std::io::Error`.
3. ? заставляет компилятор искать и использовать реализацию `From`, которая может преобразовать `String` в `MyError`.

Вложенные ошибки

Альтернативным сценарием будет случай, когда содержимое вложенных ошибок достаточно важно для того, чтобы его сохранить и сделать доступным для вызывающего кода.

Рассмотрим библиотечную функцию, которая пытается вернуть первую строку файла в виде строки при условии, что эта строка не слишком длинная. Немного подумав, здесь можно обнаружить не менее трех возможных видов сбоя.

- Файл может не существовать или быть недоступным для чтения.
- Файл может содержать данные не в допустимом формате UTF-8, а значит, их не получится преобразовать в `String`.
- Первая строка файла может оказаться слишком длинной.

В соответствии с рекомендацией 1 здесь уместно использовать систему типов для выражения и включения всех возможных исходов в виде `enum`:

```
#[derive(Debug)]
pub enum MyError {
    Io(std::io::Error),
    Utf8(std::string::FromUtf8Error),
    General(String),
}
```

Это определение `enum` включает `derive(Debug)`, но для удовлетворения требований трейта `Error` необходима также реализация `Display`:

```
impl std::fmt::Display for MyError {
    fn fmt(&self, f: &mut std::fmt::Formatter<'_>) -> std::fmt::Result {
        match self {
            MyError::Io(e) => write!(f, "IO error: {}", e),
            MyError::Utf8(e) => write!(f, "UTF-8 error: {}", e),
            MyError::General(s) => write!(f, "General error: {}", s),
        }
    }
}
```

Кроме того, есть смысл переписать предустановленную реализацию `source()` для упрощения доступа к вложенным ошибкам:

```
use std::error::Error;

impl Error for MyError {
    fn source(&self) -> Option<&(dyn Error + 'static)> {
        match self {
            MyError::Io(e) => Some(e),
            MyError::Utf8(e) => Some(e),
            MyError::General(_) => None,
        }
    }
}
```

Использование `enum` позволяет лаконично обрабатывать ошибки, сохраняя при этом всю информацию типов для разных классов ошибок:

```
use std::io::BufRead; // для `read_until`  
  

/// Максимальная длина строки
const MAX_LEN: usize = 1024;  
  

/// Возврат первой строки указанного файла.
pub fn first_line(filename: &str) -> Result<String, MyError> {
    let file = std::fs::File::open(filename).map_err(MyError::Io)?;
    let mut reader = std::io::BufReader::new(file);

    // (В реальной реализации можно просто
    // использовать `reader.read_line()`)  

    let mut buf = vec![];
    let len = reader.read_until(b'\n', &mut buf).map_err(MyError::Io)?;
    let result = String::from_utf8(buf).map_err(MyError::Utf8)?;
    if result.len() > MAX_LEN {
        return Err(MyError::General(format!("Line too long: {}", len)));
    }
    OK(result)
}
```

Желательно также реализовать трейт `From` для всех типов вложенных ошибок (см. рекомендацию 5):

```
impl From<std::io::Error> for MyError {
    fn from(e: std::io::Error) -> Self {
        Self::Io(e)
    }
}

impl From<std::string::FromUtf8Error> for MyError {
    fn from(e: std::string::FromUtf8Error) -> Self {
        Self::Utf8(e)
    }
}
```

Это избавит пользователей библиотеки от страданий, связанных с самим правилом сироты: они не могут реализовать `From` в `MyError`, поскольку трейт и структура являются для них внешними.

В добавок реализация `From` позволяет добиться еще большей лаконичности, так как оператор `?!` будет автоматически выполнять все необходимые преобразования `From`, исключая потребность в `.map_err()`:

```
use std::io::BufRead; // для `read_until()`

/// Максимальная длина строки
pub const MAX_LEN: usize = 1024;

/// Возврат первой строки указанного файла
pub fn first_line(filename: &str) -> Result<String, MyError> {
    let file = std::fs::File::open(filename)?; // `From<std::io::Error>`
    let mut reader = std::io::BufReader::new(file);
    let mut buf = vec![];
    let len = reader.read_until(b'\n', &mut buf)?; // `From<std::io::Error>`
    let result = String::from_utf8(buf)?; // `From<String::FromUtf8Error>`
    if result.len() > MAX_LEN {
        return Err(MyError::General(format!("Line too long: {}", len)));
    }
    OK(result)
}
```

Написание полноценного типа ошибки может включать большой объем шаблонного кода, что делает его хорошим кандидатом на автоматизацию посредством макросов `derive` (см. рекомендацию 28). Однако не обязательно писать такой макрос самим, так как можно использовать крейт `thiserror` от Дэвида Толная (David Tolnay), который предоставляет его качественную, широко применяемую реализацию. При генерации кода `thiserror` старательно избегает раскрытия любых типов `this error` в генерируемом API. Это означает, что можно забыть о тревогах, связанных с рекомендацией 24.

Трейт-объекты

Первый подход к реализации вложенных ошибок отбрасывал все их подробности, сохраняя лишь строковый вывод (`format!("{}?", err)`). Второй уже сохранял всю информацию типа для всех возможных вложенных ошибок, но требовал полного перечисления всех их возможных типов.

И здесь возникает вопрос. А есть ли некая золотая середина, чтобы сохранялась информация о вложенных ошибках без необходимости вручную включать все их возможные типы?

Кодирование информации о вложенной ошибке в виде *объекта трейта* исключает потребность в создании варианта `enum` для каждого возможного исхода, но исключает и детали конкретных внутренних типов ошибок. Получатель такого объекта будет иметь доступ к методам трейта `Error` и его границам — `source()`, `Display::fmt()` и `Debug::fmt()`, но не будет знать исходный статический тип вложенной ошибки.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
#[derive(Debug)]
pub enum WrappedError {
    Wrapped(Box<dyn Error>),
    General(String),
}

impl std::fmt::Display for WrappedError {
    fn fmt(&self, f: &mut std::fmt::Formatter<'_>) -> std::fmt::Result {
        match self {
            Self::Wrapped(e) => write!(f, "Inner error: {}", e),
            Self::General(s) => write!(f, "{}", s),
        }
    }
}
```

Оказывается, такое поведение возможно, но на удивление труднореализуемо. Отчасти сложность обусловливается ограничениями безопасности объектов трейтов (см. рекомендацию 12), участвуют в этом и *правила согласованности Rust*, которые гласят примерно следующее: для типа может иметься не более одной реализации трейта.

В этом случае будет действовать наивное предположение, что гипотетический тип `WrappedError` реализует оба следующих элемента:

- трейт `Error`, поскольку это и есть сама ошибка;
- трейт `From<Error>`, чтобы упростить обертывание вложенных ошибок.

Это означает, что `WrappedError` можно создать из (`from`) внутреннего `WrappedError`, так как `WrappedError` реализует `Error`, что вступает в конфликт с возвратной (reflexive) blanket-реализацией `From`:

НЕ КОМПИЛИРУЕТСЯ

```
impl Error for WrappedError {}

impl<E: 'static + Error> From<E> for WrappedError {
    fn from(e: E) -> Self {
        Self::Wrapped(Box::new(e))
    }
}
```

```
error[E0119]: conflicting implementations of trait `From<WrappedError>`
for type `WrappedError`
--> src/main.rs:279:5
|
279 |     impl<E: 'static + Error> From<E> for WrappedError {
|     ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
|
= note: conflicting implementation in crate `core`:
- impl<T> From<T> for T;
```

Дэвид Толнай разработал крейт `anyhow`, который уже решил эти проблемы, добавив дополнительный уровень косвенности посредством `Box` (<https://oreil.ly/aWecz>), и привносит другие полезные возможности, такие как трассировка стека. В итоге его использование уже становится стандартной рекомендацией для обработки ошибок, которую даем и мы: *рассматривайте обработку ошибок в приложениях с помощью крейта anyhow*.

Библиотеки и приложения

Последний совет предыдущего раздела включал уточнение «...для обработки ошибок в приложениях». Дело в том, что зачастую есть разница между кодом, написанным для повторного применения, и кодом приложений верхнего уровня¹.

Код, который написан для библиотеки, не может прогнозировать, в какой среде будет использоваться, поэтому в нем лучше избегать конкретной и подробной

¹ Этот раздел основан на статье Ника Гроенена (Nick Groenen) Rust: Structuring and Handling Errors in 2020 (<https://oreil.ly/2K3PH>).

информации об ошибках, оставляя выбор способа ее применения на откуп вызывающему коду. Такой подход склоняет нас к реализации ранее описанных вложенных ошибок в стиле `enum` (плюс исключает зависимость от `anyhow` в публичном API библиотеки; подробнее — в рекомендации 24).

При этом код приложения обычно должен делать больший акцент на выражении ошибки для пользователя. Он также потенциально должен обрабатывать все возможные типы ошибок, генерируемые всеми библиотеками, присутствующими в графе его зависимостей (см. рекомендацию 25). В связи с этим более динамический тип, такой как `anyhow::Error`, не только упрощает обработку ошибок, но и делает ее согласованной по всему приложению.

Запомните

- Стандартный трейт `Error` требует от вас немногого, поэтому стоит использовать для реализации типов ошибок именно его.
- Работая с разнородными типами ошибок, обдумывайте, нужно ли сохранять эти типы:
 - если нет, рассмотрите использование `anyhow` для обертывания вложенных ошибок в код приложения;
 - если да, кодируйте их в `enum` и предоставьте возможность преобразования. В качестве вспомогательного инструмента можете взять `thiserror`.
- Подумайте о применении крейта `anyhow` для удобной идиоматической обработки ошибок в коде приложения.
- Итоговое решение за вами, но, каким бы оно ни оказалось, прописывайте его с помощью системы типов (см. рекомендацию 1).

Рекомендация 5. Изучите преобразования типов

Преобразования типов в Rust разделяются на три категории.

Ручное

Пользовательские преобразования, предоставляемые реализацией трейтов `From` и `Into`.

Полуавтоматические

Явные приведения между значениями с помощью ключевого слова `as`.

Автоматические

Неявное приведение к новому типу.

Основная часть этой рекомендации будет посвящена ручным преобразованиям, поскольку остальные два вида редко применяются к пользовательским типам. Тем не менее есть пара исключений, поэтому далее приводятся разделы, освещдающие приведение и неявное приведение, в том числе применительно к пользовательскому типу.

Обратите внимание, что, в отличие от многих других языков, Rust не выполняет автоматическое преобразование между численными типами, причем это касается даже безопасных операций:

НЕ КОМПИЛИРУЕТСЯ

```
let x: u32 = 2;
let y: u64 = x;
```

```
error[E0308]: mismatched types
--> src/main.rs:70:18
|
70 |     let y: u64 = x;
|          ^ expected `u64`, found `u32`
|          |
|          expected due to this
|
help: you can convert a `u32` to a `u64`
|
70 |     let y: u64 = x.into();
|           ++++++
```

Пользовательские преобразования типов

Как и многое в Rust (см. рекомендацию 10), возможность выполнять преобразования между значениями разных пользовательских типов реализована в виде стандартного трейта, точнее, набора связанных обобщенных трейтов.

Вот четыре трейта, которые позволяют преобразовывать типы значений.

From<T>

Элементы этого типа можно создавать из элементов типа T. Преобразование в этом случае всегда происходит успешно.

TryFrom<T>

Элементы этого типа можно получать из элементов типа T, но в данном случае преобразование может провалиться.

Into<T>

Элементы этого типа можно преобразовывать в элементы типа T со 100%-ным успехом.

TryInto<T>

Элементы этого типа можно преобразовывать в элементы типа T, но преобразование может провалиться.

Учитывая рекомендацию 1, где речь шла о выражении элементов в системе типов, неудивительно обнаружить, что различие между вариантами Try... заключается в том, что этот метод трейта возвращает Result, а не гарантированно новый элемент. Определения трейта Try... требуют также связанного типа, который в случае сбоев будет выдавать тип ошибки E.

Поэтому первым советом, согласующимся с рекомендацией 4, здесь будет такой: *реализовывать (только) трейт Try..., если преобразование может провалиться*. В качестве альтернативы можно проигнорировать вероятность ошибки (например, с помощью .unwrap()), но тогда потребуется сделать целенаправленный выбор, который в большинстве случаев лучше оставить на откуп вызывающему коду.

Трейты для преобразования типов имеют очевидную симметрию: если тип T можно превратить в (into) тип U (посредством Into<U>), не будет ли это равнозначно возможности создать элемент с типом U, выполнив преобразование из (from) элемента с типом T (посредством From<T>)?

Это действительно так, и отсюда вытекает второй совет: *реализуйте для преобразования трейт From*. В стандартной библиотеке Rust нужно было выбрать лишь одну из двух возможностей, чтобы система не зацикливалась на круговом переборе этих вариантов¹. В связи с этим было решено автоматически предоставлять Into из реализации From.

Если вы используете один из этих трейтов в качестве границы трейта своего собственного обобщения, то совет таков: *используйте для границ трейта Into*. Таким образом граница будет удовлетворяться как элементами, непосредственно реализующими Into, так и теми, которые непосредственно реализуют только From.

Документация предполагает это автоматическое преобразование для From и Into, но стоит прочитать и соответствующую часть стандартной библиотеки, которая предоставляет *blanket-реализацию трейта*:

```
impl<T, U> Into<U> for T
where
    U: From<T>,
{
    fn into(self) -> U {
        U::from(self)
    }
}
```

¹ Более широко известно как правило согласованности трейтов.

Перевод спецификации трейта в слова поможет понять более сложные границы трейтов. В этом случае все довольно просто: «Я могу реализовать `Into<U>` для типа `T`, когда `U` уже реализует `From<T>`».

В стандартной библиотеке есть и различные реализации этих трейтов преобразования для используемых в ней типов. Вполне ожидаемо там есть также реализации `From` для целочисленных преобразований, когда целевой тип включает все возможные значения исходного (`From<u32> for u64`), а также реализации `TryFrom`, когда исходное значение может не подходить для целевого типа (`TryFrom<u64> for u32`).

Помимо ранее показанной версии `Into`, есть и другие blanket-реализации трейтов. Они используются преимущественно для типов *умных указателей*, позволяя автоматически создавать такой указатель на основе экземпляра содержащегося в нем типа. Это означает, что обобщенные методы, которые получают параметры умного указателя, также можно вызывать с помощью типичных элементов. Подробнее эта тема будет раскрыта в рекомендации 8.

Трейт `TryFrom` тоже имеет blanket-реализацию для любого типа, который уже реализует трейт `Into` в обратном направлении, что автоматически включает, как было показано ранее, любой тип, реализующий `From` в том же направлении. Иными словами, если вы можете безошибочно преобразовать `T` в `U`, то можете получить и `U` из `T` с вероятностью сбоя. Поскольку это преобразование всегда будет успешным, связанный тип ошибки получил соответствующее имя `Infallible`¹.

Кроме того, существует особая обобщенная форма `From` — *возвратная реализация*:

```
impl<T> From<T> for T {
    fn from(t: T) -> T {
        t
    }
}
```

Если перевести в слова, то здесь говорится: «Имея `T`, я могу получить `T`». Это настолько очевидно, что можно дальше не пытаться понять, в чем тут польза.

Рассмотрим простую newtype-struct (см. рекомендацию 6) и функцию, которая с ней работает (игнорируя тот факт, что эту функцию лучше выразить в виде метода):

```
/// Целочисленное значение из диапазона, предоставленного IANA
#[derive(Clone, Copy, Debug)]
pub struct IanaAllocated(pub u64);
```

¹ На данный момент. В будущей версии Rust его наверняка заменят типом `!` под названием `never`.

```
/// Показывает, зарезервировано ли значение
pub fn is_iana_reserved(s: IanaAllocated) -> bool {
    s.0 == 0 || s.0 == 65535
}
```

Эту функцию можно вызвать с экземплярами `struct`:

```
let s = IanaAllocated(1);
println!("{}:{} reserved? {}", s, is_iana_reserved(s));
// вывод: "IanaAllocated(1) reserved? false"
```

Но, даже если `From<u64>` реализуется для newtype-обертки:

```
impl From<u64> for IanaAllocated {
    fn from(v: u64) -> Self {
        Self(v)
    }
}
```

эту функцию нельзя вызвать непосредственно для значений `u64`:

НЕ КОМПИЛИРУЕТСЯ

```
if is_iana_reserved(42) {
    // ...
}
```

```
error[E0308]: mismatched types
--> src/main.rs:77:25
|
77 |     if is_iana_reserved(42) {
|         ----- ^ expected `IanaAllocated`, found integer
|         |
|         arguments to this function are incorrect
|
note: function defined here
--> src/main.rs:7:8
|
7 | pub fn is_iana_reserved(s: IanaAllocated) -> bool {
|     ^^^^^^^^^^^^^^ -----
help: try wrapping the expression in `IanaAllocated`  

|
77 |     if is_iana_reserved(IanaAllocated(42)) {
|                     +++++++ +
```

Тем не менее обобщенная версия этой функции, которая получает и явно преобразует все, что удовлетворяет условию `Into<ianaAllocated>`:

```
pub fn is_iana_reserved<T>(s: T) -> bool
where
    T: Into<ianaAllocated>,
{
    let s = s.into();
    s.0 == 0 || s.0 == 65535
}
```

позволяет сделать так:

```
if is_iana_reserved(42) {
    // ...
}
```

После установки этой границы возвратная реализация трейта `From<T>` становится более логичной: она означает, что обобщенная функция обрабатывает элементы, которые уже являются экземплярами `ianaAllocated`, не преобразуя их.

Этот паттерн также объясняет, почему и как код Rust иногда *производит* неявные преобразования между типами: комбинация реализаций `From<T>` и границ трейта `Into<T>` ведет к созданию кода, который магически преобразуется в месте вызова, но по-прежнему выполняет безопасное явное преобразование. Данный паттерн становится еще более полезным при совмещении со ссылочными типами и связанными с ними трейтами преобразования. Подробнее об этом — в рекомендации 8.

Приведения

В Rust есть ключевое слово `as`, служащее для выполнения явного преобразования между двумя типами.

Существует небольшое множество пар типов, которые могут быть преобразованы таким образом, и к пользовательским типам среди них относятся лишь C-подобные `enum`, у которых есть только соответствующее целочисленное значение. Сюда входят и общие целочисленные преобразования, предоставляющие альтернативу для `into()`:

```
let x: u32 = 9;
let y = x as u64;
let z: u64 = x.into();
```

Эта версия с `as` позволяет также выполнять преобразования с потерями (lossy conversions)¹:

```
let x: u32 = 9;
let y = x as u16;
```

которые версиями `from/into` будут отклонены:

```
error[E0277]: the trait bound `u16: From<u32>` is not satisfied
--> src/main.rs:136:20
|
136 |     let y: u16 = x.into();
|           ^^^^^ the trait `From<u32>` is not implemented for `u16`
|
|= help: the following other types implement trait `From<T>`:
    <u16 as From<NonZeroU16>>
    <u16 as From<bool>>
    <u16 as From<u8>>
= note: required for `u32` to implement `Into<u16>`
```

В целях обеспечения согласованности и безопасности вам следует *отдавать предпочтение преобразованиям from/into перед явными приведениями с помощью as*, если только вы не нуждаетесь в точной семантике механизма приведения, например, для достижения операционной совместимости с C. Следование этому совету можно подкрепить с помощью Clippy (см. рекомендацию 29), который включает несколько линтов для `as`. Однако по умолчанию эти линты отключены.

Неявное приведение

Явные приведения с помощью `as`, описанные в предыдущем подразделе, являются надмножеством *неявных* (*coercion*), которые компилятор выполняет незаметно: любое неявное приведение можно принудительно выполнить с помощью `as`, но в обратном направлении это не работает. В частности, целочисленные приведения, которые выполнялись в предыдущем разделе, не являются неявными, поэтому всегда будут требовать `as`.

Большинство неявных приведений включают скрытое преобразование типов указателей и ссылочных типов понятным и удобным для программиста способом, например преобразования:

- мутабельной ссылки в немутабельную (чтобы можно было использовать `&mut T` в качестве аргумента для функции, которая получает `&T`);

¹ Возможность выполнять преобразования с потерями в Rust, пожалуй, является ошибкой, и на тему исключения этого поведения даже велись открытые дискуссии (<https://oreil.ly/TpFKB>).

- ссылки в сырой указатель (это не `unsafe` — опасность возникает в момент, когда вы по глупости *разыменовываете* сырой указатель);
- замыкания, не содержащего никаких переменных, в голый указатель функции (см. рекомендацию 2);
- массива в срез;
- конкретного элемента в объект трейта, который этот конкретный элемент реализует;
- времени жизни элемента в более короткое (см. рекомендацию 14)¹.

Есть всего два неявных приведения, чье поведение может повлиять на пользовательские типы. Первое происходит, когда пользовательский тип реализует трейт `Deref` или `DerefMut`. Эти трейты показывают, что пользовательский тип выступает в виде своеобразного *умного указателя* (см. рекомендацию 8), в случае чего компилятор неявно преобразует ссылку на умный указатель в ссылку на элемент с типом, который этот умный указатель содержит (прописан в его `Target`).

Второе неявное приведение пользовательского типа происходит, когда конкретный элемент преобразуется в *объект трейта*. Эта операция создает жирный указатель на элемент. Жирным он является потому, что включает как указатель на расположение элемента в памяти, так и указатель на `vtable` для реализации трейта в конкретном типе (подробнее — в рекомендации 8).

Рекомендация 6. Используйте паттерн `newtype`

В рекомендации 1 описывались *структуры кортежа*, в которых обращение к полям `struct` происходит не по имени, которого у них нет, а по номеру (`self.0`). В этой же рекомендации речь пойдет о структурах, которые содержат единственную запись некоего существующего типа, тем самым создавая новый тип, который может содержать точно такой же диапазон значений, что и включенный в них тип. Это довольно широко распространенный в Rust паттерн, и он заслуживает как своего личного имени — *паттерн newtype*, так и отдельной рекомендации.

Простейшим способом использования этого паттерна выступает указание дополнительной семантики для типа, выходящей за его обычное поведение. Чтобы это понять, представьте проект по запуску спутника на Марс². Это

¹ В Rust такие преобразования называются подтиповацией, но они довольно сильно отличаются от определения подтиповации, используемого в объектно-ориентированных языках.

² В частности, Mars Climate Orbiter.

масштабный проект, поэтому разные группы специалистов создают для него разные детали. Одна из этих групп занимается разработкой кода для двигателей ракеты-носителя:

```
/// Запуск двигателей. Возвращает величину полученного импульса
/// в единицах "фунт-сила-секунда"
pub fn thruster_impulse(direction: Direction) -> f64 {
    // ...
    return 42.0;
}
```

Другая группа разрабатывает инерционную систему наведения:

```
/// Обновляет модель траектории для импульса,
/// указанного в ньютонах в секунду
pub fn update_trajectory(force: f64) {
    // ...
}
```

В конечном счете эти отдельные части будут объединены:

```
let thruster_force: f64 = thruster_impulse(direction);
let new_direction = update_trajectory(thruster_force);
```

Упс¹...

В Rust есть функциональность *псевдонимов типа*, которая позволяет разным группам более отчетливо выражать свои намерения:

```
/// Единицы измерения силы
pub type PoundForceSeconds = f64;

/// Запуск двигателя. Возвращает полученный импульс
pub fn thruster_impulse(direction: Direction) -> PoundForceSeconds {
    // ...
    return 42.0;
}

/// Единицы измерения силы
pub type NewtonSeconds = f64;

/// Обновляет модель траектории импульса
pub fn update_trajectory(force: NewtonSeconds) {
    // ...
}
```

¹ Подробнее о причине сбоя читайте статью Mars Climate Orbiter в «Википедии» (<https://oreilly/AiraF>).

Однако псевдонимы типов, по сути, представляют собой просто документацию. Они более значимы в сравнении с документирующими комментариями предыдущей версии, но ничто не препятствует случайному использованию значения `PoundForceSeconds` там, где ожидается значение `NewtonSeconds`:

```
let thruster_force: PoundForceSeconds = thruster_impulse(direction);
let new_direction = update_trajectory(thruster_force);
```

И снова уп...

Это тот случай, когда поможет паттерн newtype:

```
/// Единицы измерения силы
pub struct PoundForceSeconds(pub f64);

/// Запуск двигателей. Возвращает полученный импульс
pub fn thruster_impulse(direction: Direction) -> PoundForceSeconds {
    // ...
    return PoundForceSeconds(42.0);
}

/// Единицы измерения силы
pub struct NewtonSeconds(pub f64);

/// Обновляет модель траектории импульса
pub fn update_trajectory(force: NewtonSeconds) {
    // ...
}
```

Как и показывает имя, newtype — это новый тип, в связи с чем компилятор возражает, когда замечает несоответствие типов, — в данном случае при попытке передать значение `PoundForceSeconds` туда, где ожидается `NewtonSeconds`:

НЕ КОМПИЛИРУЕТСЯ

```
let thruster_force: PoundForceSeconds = thruster_impulse(direction);
let new_direction = update_trajectory(thruster_force);
```

```
error[E0308]: mismatched types
--> src/main.rs:76:43
|
76 |     let new_direction = update_trajectory(thruster_force);
|           ----- ^^^^^^^^^^^^^^ expected
|                   |           `NewtonSeconds`, found `PoundForceSeconds`
|                   |
|                   arguments to this function are incorrect
```

```
|  
note: function defined here  
--> src/main.rs:66:8  
|  
66 | pub fn update_trajectory(force: NewtonSeconds) {  
|     ^^^^^^^^^^^^^^ -----  
help: call `Into::into` on this expression to convert `PoundForceSeconds`  
      into `NewtonSeconds`  
|  
76 |     let new_direction = update_trajectory(thruster_force.into());  
|           ++++++
```

Как говорилось в рекомендации 5, добавление реализации стандартного трейта `From`:

```
impl From<PoundForceSeconds> for NewtonSeconds {  
    fn from(val: PoundForceSeconds) -> NewtonSeconds {  
        NewtonSeconds(4.448222 * val.0)  
    }  
}
```

позволяет выполнить нужное преобразование единиц и типа с помощью `.into()`:

```
let thruster_force: PoundForceSeconds = thruster_impulse(direction);  
let new_direction = update_trajectory(thruster_force.into());
```

Тот же паттерн использования newtype для обозначения дополнительной семантики единиц в типе помогает сделать чисто логические аргументы менее двусмысленными. Возвращаясь к примеру из рекомендации 1, видим, что применение newtype дополнительно проясняет смысл аргументов:

```
struct DoubleSided(pub bool);  
  
struct ColorOutput(pub bool);  
  
fn print_page(sides: DoubleSided, color: ColorOutput) {  
    // ...  
}  
  
print_page(DoubleSided(true), ColorOutput(false));
```

Если эффективность используемого размера или двоичная совместимость вызывают беспокойство, тогда атрибут `#[repr(transparent)]` обеспечит, чтобы newtype имел в памяти то же представление, что и внутренний тип.

Это пример простого применения newtype, и он явно отражает суть рекомендации 1 — кодировать семантику в системе типов, чтобы компилятор проследил за ее соблюдением.

Обход правила сироты для трейтов

Еще один типичный, но более тонкий сценарий, требующий паттерна newtype, связан с правилом сироты. Грубо говоря, это правило гласит, что крейт может реализовывать трейт для типа, только если соблюдается одно из следующих требований:

- крейт определил трейт;
- крейт определил тип.

Попытка реализовать посторонний трейт для постороннего типа...

НЕ КОМПИЛИРУЕТСЯ

```
use std::fmt;

impl fmt::Display for rand::rngs::StdRng {
    fn fmt(&self, f: &mut fmt::Formatter<'_>) -> Result<(), fmt::Error> {
        write!(f, "<StdRng instance>")
    }
}
```

...вызывает ошибку при компиляции, которая, в свою очередь, указывает обратно на newtype:

```
error[E0117]: only traits defined in the current crate can be implemented for
              types defined outside of the crate
--> src/main.rs:146:1
 |
146 | impl fmt::Display for rand::rngs::StdRng {
| ^^^^^^^^^^^^^^^^^^^^^^-----|
| | | | `StdRng` 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
```

Причина этого ограничения связана с риском внесения двусмыслинности: если два разных крейта в графе зависимостей (см. рекомендацию 25), где оба, например, `impl std::fmt::Display for rand::rngs::StdRng`, тогда компилятор/линковщик не может сделать между ними выбор.

И это часто озадачивает: например, если вы пытаетесь сериализовать данные, которые включают тип из другого крейта, то правило сироты не позволит вам написать `impl serde::Serialize for somecrate::SomeType1`.

Но паттерн newtype означает, что вы определяете *новый* тип, который является частью текущего крейта, а значит, здесь работает вторая часть правила сироты. Теперь реализация постороннего трейта возможна:

```
struct MyRng(rand::rngs::StdRng);

impl fmt::Display for MyRng {
    fn fmt(&self, f: &mut fmt::Formatter<'_>) -> Result<(), fmt::Error> {
        write!(f, "<MyRng instance>")
    }
}
```

Ограничения newtype

Паттерн newtype решает эти два класса проблем, препятствуя преобразованию единиц и обходя правило сироты, но вносит свои странности: каждая операция, которая включает newtype, должна выполнять перенаправление во внутренний тип. В общем смысле это означает, что код должен везде использовать `thing.0`, а не просто `thing`, но это легко, и компилятор сообщит вам, где это необходимо. Более же значимая странность заключается в том, что любые реализации трейта для внутреннего типа утрачиваются: поскольку newtype — это новый тип, существующая внутренняя реализация не работает.

Для выводимых с помощью `derive` трейтов это просто означает, что объявление newtype содержит множество атрибутов `derive`:

```
#[derive(Debug, Copy, Clone, Eq, PartialEq, Ord, PartialOrd)]
pub struct NewType(InnerType);
```

Для более сложных трейтов необходим шаблонный код, восстанавливающий реализацию внутреннего типа, например:

```
use std::fmt;
impl fmt::Display for NewType {
    fn fmt(&self, f: &mut fmt::Formatter<'_>) -> Result<(), fmt::Error> {
        self.0.fmt(f)
    }
}
```

¹ Это настолько распространенная проблема serde, что для этого даже существует вспомогательный механизм (https://oreil.ly/6_6MC).

Рекомендация 7. Для сложных типов используйте паттерн «Строитель»

В данной рекомендации описывается паттерн «Строитель» (Builder), в котором у сложных структур данных есть связанный тип `builder`, упрощающий для пользователей создание экземпляров этой структуры.

Rust требует, чтобы при создании нового экземпляра все поля в `struct` были заполнены. Это обеспечивает безопасность кода, гарантируя отсутствие неинициализированных значений, но из-за этого код получается более громоздким и шаблонным.

Например, любые optionalные поля должны быть явно обозначены как отсутствующие с помощью `None`:

```
/// Номер телефона в формате E164
#[derive(Debug, Clone)]
pub struct PhoneNumberE164(pub String);

#[derive(Debug, Default)]
pub struct Details {
    pub given_name: String,
    pub preferred_name: Option<String>,
    pub middle_name: Option<String>,
    pub family_name: String,
    pub mobile_phone: Option<PhoneNumberE164>,
}

// ...

let dizzy = Details {
    given_name: "Dizzy".to_owned(),
    preferred_name: None,
    middle_name: None,
    family_name: "Mixer".to_owned(),
    mobile_phone: None,
};
```

Кроме того, этот шаблонный код еще и хрупок в том смысле, что будущее изменение, которое добавляет новое поле в `struct`, потребует обновления во всех местах, создающих эту структуру. Но этот код можно значительно сократить, реализовав и используя трейт `Default`, как описано в рекомендации 10:

```
let dizzy = Details {
    given_name: "Dizzy".to_owned(),
    family_name: "Mixer".to_owned(),
    ..Default::default()
};
```

Применение `Default` помогает также сократить объем изменений, необходимых при добавлении нового поля, если только само это поле имеет тип, реализующий

`Default`. И это уже более общая проблема: автоматически выведенная реализация `Default` работает, только если все типы полей реализуют трейт `Default`. Если какое-то из полей от этого отклонится, этап `derive` не сработает:

НЕ КОМПИЛИРУЕТСЯ

```
#[derive(Debug, Default)]
pub struct Details {
    pub given_name: String,
    pub preferred_name: Option<String>,
    pub middle_name: Option<String>,
    pub family_name: String,
    pub mobile_phone: Option<PhoneNumberE164>,
    pub date_of_birth: time::Date,
    pub last_seen: Option<time::OffsetDateTime>,
}
```

```
error[E0277]: the trait bound `Date: Default` is not satisfied
--> src/main.rs:48:9
|
41 |     #[derive(Debug, Default)]
|         ----- in this derive macro expansion
...
48 |         pub date_of_birth: time::Date,
|             ^^^^^^^^^^^^^^^^^^^^^^^^^^^^ the trait `Default` is not
|                             implemented for `Date`
|
= note: this error originates in the derive macro `Default`
```

Этот код не может реализовать `Default` для `time::Date` из-за правила сироты. Но даже если бы мог, это бы не помогло — использование предустановленного значения даты рождения практически во всех случаях окажется ошибкой.

Отсутствие `Default` означает, что все поля необходимо заполнять вручную:

```
let bob = Details {
    given_name: "Robert".to_owned(),
    preferred_name: Some("Bob".to_owned()),
    middle_name: Some("the".to_owned()),
    family_name: "Builder".to_owned(),
    mobile_phone: None,
    date_of_birth: time::Date::from_calendar_date(
        1998,
        time::Month::November,
        28,
    )
    .unwrap(),
    last_seen: None,
};
```

Данную эргономику можно улучшить, если реализовать для сложных структур данных паттерн «Строитель». Простейший его вариант представляет отдельную `struct`, которая содержит информацию, необходимую для построения элемента. В целях упрощения приведенный пример будет содержать экземпляр самого этого элемента:

```
pub struct DetailsBuilder(Details);

impl DetailsBuilder {
    /// Начинает строить новый объект [`Details`]
    pub fn new(
        given_name: &str,
        family_name: &str,
        date_of_birth: time::Date,
    ) -> Self {
        DetailsBuilder(Details {
            given_name: given_name.to_owned(),
            preferred_name: None,
            middle_name: None,
            family_name: family_name.to_owned(),
            mobile_phone: None,
            date_of_birth,
            last_seen: None,
        })
    }
}
```

Далее этот тип «строителя» можно снабдить вспомогательными методами, которые заполнят формирующеся поле элемента. Каждый такой метод получает `self`, генерируя новый `Self`, что позволяет связывать создающие структуру методы в цепочку:

```
/// Установка предпочтительного имени
pub fn preferred_name(mut self, preferred_name: &str) -> Self {
    self.0.preferred_name = Some(preferred_name.to_owned());
    self
}

/// Установка второго имени
pub fn middle_name(mut self, middle_name: &str) -> Self {
    self.0.middle_name = Some(middle_name.to_owned());
    self
}
```

Эти вспомогательные методы могут оказаться полезнее простых сеттеров:

```
/// Обновление поля `last_seen` на текущие дату/время
pub fn just_seen(mut self) -> Self {
    self.0.last_seen = Some(time::OffsetDateTime::now_utc());
    self
}
```

Заключительный метод получает «строителя» и генерирует собираемый элемент:

```
/// Получает объект «строителя» и возвращает готовый объект [ `Details` ]
pub fn build(self) -> Details {
    self.0
}
```

В целом это обеспечивает для клиентов «Строителя» более эргономичный режим сборки:

```
let also_bob = DetailsBuilder::new(
    "Robert",
    "Builder",
    time::Date::from_calendar_date(1998, time::Month::November, 28)
        .unwrap(),
)
.middle_name("the")
.preferred_name("Bob")
.just_seen()
.build();
```

Такая всепоглощающая природа «строителя» ведет к паре загвоздок. Первая состоит в том, что разделение этапов сборки нельзя проделать само по себе:

НЕ КОМПИЛИРУЕТСЯ

```
let builder = DetailsBuilder::new(
    "Robert",
    "Builder",
    time::Date::from_calendar_date(1998, time::Month::November, 28)
        .unwrap(),
);
if informal {
    builder.preferred_name("Bob");
}
let bob = builder.build();
```

```
error[E0382]: use of moved value: `builder`
--> src/main.rs:256:15
|
247 |     let builder = DetailsBuilder::new(
|         ----- move occurs because `builder` has type `DetailsBuilder`,
|             which does not implement the `Copy` trait
...
254 |         builder.preferred_name("Bob");
|             ----- `builder` moved due to this method
|                 call
255 |     }
```

```

256 |     let bob = builder.build();
|         ^^^^^^ value used here after move
|
note: `DetailsBuilder::preferred_name` takes ownership of the receiver `self`,
      which moves `builder`
--> src/main.rs:60:35
|
27 |     pub fn preferred_name(mut self, preferred_name: &str) -> Self {
|         ^

```

Этот нюанс можно обойти, присвоив получаемый объект «строителя» той же переменной:

```

let mut builder = DetailsBuilder::new(
    "Robert",
    "Builder",
    time::Date::from_calendar_date(1998, time::Month::November, 28)
        .unwrap(),
);
if informal {
    builder = builder.preferred_name("Bob");
}
let bob = builder.build();

```

Второй недостаток этого «строителя» в том, что он может собрать только один элемент. Попытка создать несколько экземпляров путем повторного вызова `build()` в одном и том же «строителье» ожидаемо вызовет ошибку при компиляции:

НЕ КОМПИЛИРУЕТСЯ

```

let smithy = DetailsBuilder::new(
    "Agent",
    "Smith",
    time::Date::from_calendar_date(1999, time::Month::June, 11).unwrap(),
);
let clones = vec![smithy.build(), smithy.build(), smithy.build()];

```

```

error[E0382]: use of moved value: `smithy`
--> src/main.rs:159:39
|
154 |     let smithy = DetailsBuilder::new(
|         ----- move occurs because `smithy` has type `base::DetailsBuilder`,
|             which does not implement the `Copy` trait
...
159 |     let clones = vec![smithy.build(), smithy.build(), smithy.build()];
|             ----- ^^^^^^ value used here after move
|             |
|             `smithy` moved due to this method call

```

В качестве альтернативного решения методы «строителя» должны получать `&mut self` и генерировать `&mut Self`:

```
/// Обновляет поле `last_seen` текущими датой/временем
pub fn just_seen(&mut self) -> &mut Self {
    self.0.last_seen = Some(time::OffsetDateTime::now_utc());
    self
}
```

Это исключит потребность в самоприсваивании на отдельных стадиях сборки:

```
let mut builder = DetailsBuilder::new(
    "Robert",
    "Builder",
    time::Date::from_calendar_date(1998, time::Month::November, 28)
        .unwrap(),
);
if informal {
    builder.preferred_name("Bob"); // нет `builder = ...`
}
let bob = builder.build();
```

Тем не менее эта версия делает возможным связывание «строителя» с вызовом его сеттеров:

НЕ КОМПИЛИРУЕТСЯ

```
let builder = DetailsBuilder::new(
    "Robert",
    "Builder",
    time::Date::from_calendar_date(1998, time::Month::November, 28)
        .unwrap(),
)
.middle_name("the")
.just_seen();
let bob = builder.build();
```

```
error[E0716]: temporary value dropped while borrowed
--> src/main.rs:265:19
|
265 |     let builder = DetailsBuilder::new(
|           ^
266 |     |     "Robert",
267 |     |     "Builder",
268 |     |     time::Date::from_calendar_date(1998, time::Month::November, 28)
269 |     |     .unwrap(),
270 |     |
|           ^ creates a temporary value which is freed while still in use
```

```

271 |     .middle_name("the")
272 |     .just_seen();
|         - temporary value is freed at the end of this statement
273 |     let bob = builder.build();
|             ----- borrow later used here
|
|= note: consider using a `let` binding to create a longer lived value

```

Как гласит ошибка компилятора, вы можете обойти эту проблему, присвоив (`let`) собираемому элементу имени:

```

let mut builder = DetailsBuilder::new(
    "Robert",
    "Builder",
    time::Date::from_calendar_date(1998, time::Month::November, 28)
        .unwrap(),
);
builder.middle_name("the").just_seen();
if informal {
    builder.preferred_name("Bob");
}
let bob = builder.build();

```

Такой измененный вариант «строителя», в частности, позволит собирать несколько элементов. Сигнатура метода `build()` не должна получать `self`, а значит, будет выглядеть так:

```

/// Создает готовый объект [ `Details` ].
pub fn build(&self) -> Details {
    // ...
}

```

В этом случае реализация повторяемого метода `build()` должна при каждом вызове создавать свежий элемент. Если внутренний элемент реализует `Clone`, это будет легко — «строитель» может содержать шаблон и клонировать (`clone()`) его для каждой сборки. Если же внутренний элемент не реализует `Clone`, тогда состояние «строитель» должно позволять вручную построить экземпляр внутреннего элемента при каждом вызове `build()`.

Какой бы стиль «строителя» ни использовался, теперь шаблонный код оказывается определен в одном месте — в «строителе» и не требуется в каждой точке, где применяется внутренний тип.

Оставшийся шаблонный код потенциально можно сократить еще больше с помощью макроса (см. рекомендацию 28), но если пойти этим путем, то нужно будет

также проверить, существует ли подходящий крейт (например, `derive_builder`), предоставляющий все необходимое, предполагая, что вас устроит получение от него зависимости (см. рекомендацию 25).

Рекомендация 8. Познакомьтесь со ссылочными типами и указателями

Если брать программирование в целом, то *ссылка* — это способ опосредованно обратиться к некой структуре данных вне зависимости от того, какой переменной она принадлежит. На практике это обычно реализуется в виде *указателя* — числа, значение которого представляет адрес структуры данных в памяти.

Современные процессоры, как правило, накладывают на указатели ряд ограничений — адрес должен находиться в допустимом диапазоне памяти (виртуальной или физической) и может требовать *выравнивания* (например, доступ к четырехбайтовому целому можно получить, только если его адрес кратен 4).

В системах типов высокого уровня языков обычно кодируется больше информации об указателях. В С-подобных языках, включая Rust, указатели имеют тип, который указывает, какая структура данных должна располагаться по соответствующему адресу в памяти. Это позволяет коду интерпретировать содержимое памяти по этому адресу и после него.

Вся базовая информация указателя — предполагаемая область в памяти и ожидаемая схема структуры данных — в Rust имеет вид *сырого указателя*. Однако в безопасном коде Rust сырые указатели не используются, поскольку этот язык предоставляет более широкий спектр ссылочных типов и типов указателей, которые обеспечивают дополнительную безопасность и ограничения. Об этих типах и пойдет речь в текущей рекомендации. А о сырых указателях мы поговорим в рекомендации 16, где будет разбираться `unsafe`-код.

Ссылки в Rust

Самым распространенным видом указателя в Rust является *ссылка*, тип которой для некоего типа `T` записывается как `&T`. И хотя это значение указателя, компилятор обеспечивает, чтобы соблюдались связанные с ним правила: он всегда должен указывать на валидный, корректно выровненный экземпляр соответствующего типа `T`, чей период жизни (см. рекомендацию 14) превышает длительность его использования, и удовлетворять правилам проверки заимствования

(см. рекомендацию 15). В Rust термин «ссылка» всегда подразумевает эти дополнительные ограничения, так что термин «указатель» встречается редко.

Требование Rust, чтобы ссылка указывала на валидный, правильно выровненный элемент, действует и в ссылочных типах C++. Однако в C++ отсутствует механизм времени жизни, из-за чего есть риск наступить на программные «грабли» в виде висячих ссылок¹.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// C++
const int& dangle() {
    int x = 32; // В стеке, переписывается позже
    return x; // Возврат ссылки в переменную стека!
}
```

Имеющиеся в Rust проверки заимствования и времени жизни означают, что равнозначный код даже не скомпилируется:

НЕ КОМПИЛИРУЕТСЯ

```
fn dangle() -> &'static i64 {
    let x: i64 = 32; // В стеке
    &x
}
```

```
error[E0515]: cannot return reference to local variable `x`
--> src/main.rs:477:5
477 |     &x
|     ^^^ returns a reference to data owned by the current function
```

В Rust ссылка `&T` делает возможным доступ к соответствующему элементу только для чтения (примерно как и `const T&` в C++). Мутабельная ссылка, которая также позволяет изменять связанный с ней элемент, записывается как `&mut T` и тоже проверяется на соблюдение правил заимствования, рассматриваемых в рекомендации 15. Этот паттерн именования отражает некоторые различия между ментальными моделями Rust и C++.

- В Rust предустановленный вариант доступен только для чтения, а допускающие запись типы отмечаются особым образом — с помощью `mut`.

¹ Современные компиляторы выдают при этом предупреждение.

- В C++, наоборот, предустановленный вариант запись допускает, а специальным образом (с помощью `const`) отмечаются уже типы только для чтения.

Компилятор преобразует код Rust, использующий ссылки, в машинный код, где применяются уже простые указатели, которые на 64-битной платформе (она подразумевается во всей текущей рекомендации) имеют размер 8 байт. Например, пара локальных переменных вместе со ссылкой на них:

```
pub struct Point {
    pub x: u32,
    pub y: u32,
}

let pt = Point { x: 1, y: 2 };
let x = 0u64;
let ref_x = &x;
let ref_pt = &pt;
```

может расположиться в стеке, как показано на рис. 1.2.

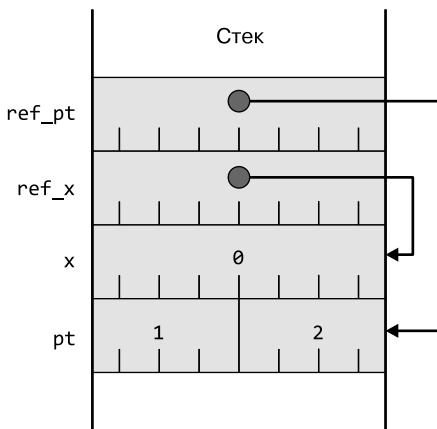


Рис. 1.2. Схема стека с указателями на локальные переменные

Ссылка в Rust может указывать на элементы, расположенные в стеке либо в куче. По умолчанию Rust аллоцирует (выделяет) элементы в стеке, но тип указателя `Box<T>` (примерно равнозначен `std::unique_ptr<T>` в C++) ведет к выделению (allocation) в кучу. Это, в свою очередь, означает, что выделенный элемент может просуществовать дольше области текущего блока. При этом `Box<T>` является простым восьмибайтовым значением указателя:

```
let box_pt = Box::new(Point { x: 10, y: 20 });
```

Это отражено на рис. 1.3.

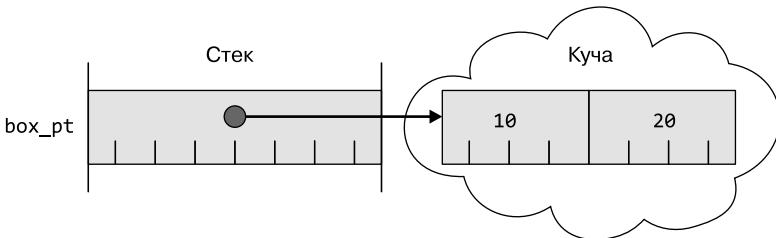


Рис. 1.3. Указатель с Box в стеке на struct в куче

Трейты указателей

Методу, ожидающему аргумент ссылки вроде `&Point`, можно передать также `&Box<Point>`:

```
fn show(pt: &Point) {
    println!("({}, {})", pt.x, pt.y);
}

show(ref_pt);
show(&box_pt);

(1, 2)
(10, 20)
```

Это возможно, поскольку `Box<T>` реализует трейт `Deref`, где `Target = T`. Реализация этого трейта для некоего типа означает, что метод `deref()` трейта можно использовать для создания ссылки на тип `Target`. Существует также равнозначный трейт `DerefMut`, который создает *мутабельную* ссылку на тип `Target`.

Упомянутые трейты `Deref/DerefMut` в некотором смысле особенные, поскольку в ходе работы с реализующими их типами компилятор действует специфическим образом. Встретив разыменовывающее выражение, например `*x`, он ищет места использования реализации одного из этих трейтов в зависимости от того, требует разыменовывание мутабельный доступ или нет. Неявное *приведение Deref* позволяет различным типам умных указателей действовать как типичные ссылки и является одним из немногих механизмов Rust, позволяющих выполнять неявное преобразование (как описывалось в рекомендации 5).

В качестве технического отступления стоит объяснить, почему трейты `Deref` не могут быть обобщенными (`Deref<Target>`) для целевого типа. Если бы могли, то какой-нибудь тип `ConfusedPtr` мог бы реализовывать и `Deref<TypeA>`, и `Deref<TypeB>`, в результате чего компилятор не смог бы вывести единый уникальный тип для выражения вроде `*x`. Поэтому целевой тип кодируется в виде связанного типа с именем `Target`.

Это отступление контрастирует с двумя другими стандартными трейтами указателей, `AsRef` и `AsMut`. Они не вызывают у компилятора особого поведения, но делают возможным преобразование в ссылку или мутабельную ссылку посредством явного вызова к их функциям трейтов `as_ref()` и `as_mut()` соответственно. Целевой тип этих преобразований кодируется в виде параметра типа, например `AsRef<Point>`, то есть один контейнерный тип может поддерживать их несколько.

Например, стандартный тип `String` реализует трейт `Deref` с `Target = str`, то есть выражение вроде `&my_string` можно неявно привести к типу `&str`. Но при этом он также реализует следующее:

- `AsRef<[u8]>` — позволяет неявное приведение к срезу байтов `&[u8]`;
- `AsRef<OsStr>` — позволяет неявное приведение к строке операционной системы;
- `AsRef<Path>` — позволяет неявное приведение к пути файловой системы;
- `AsRef<str>` — позволяет неявное приведение к срезу строк `&str` (как в случае `Deref`).

Типы жирных указателей

В Rust есть два типа *жирных указателей*: срезы и трейт-объекты. Они действуют как указатели, но содержат дополнительную информацию о том, на что указывают.

Срезы

Первым типом жирных указателей является *срез* — ссылка на подмножество некой непрерывной коллекции значений. Он создается из простого указателя (не владеющего значением), сопровождаемого полем длины, что делает его вдвое больше простого указателя (то есть 16 байт на платформе x64). Тип среза записывается как `&[T]` — ссылка на `[T]`, являющийся номинальным типом для непрерывной коллекции значений типа `T`.

Номинальный тип `[T]` инстанцировать нельзя, но есть два типичных контейнера, которые его в себе воплощают. Первый — это *массив* — непрерывная коллекция значений, размер которых известен на этапе компиляции: массив из пяти значений всегда будет иметь пять значений. Следовательно, срез может указывать на подмножество массива (рис. 1.4):

```
let array: [u64; 5] = [0, 1, 2, 3, 4];
let slice = &array[1..3];
```

Еще одним типичным контейнером для непрерывных значений является `Vec<T>`. Он, как и массив, содержит непрерывную коллекцию значений, но, в отличие от него, в `Vec` количество этих значений может расти (например, с помощью `push(value)`) или уменьшаться (например, с помощью `pop()`).

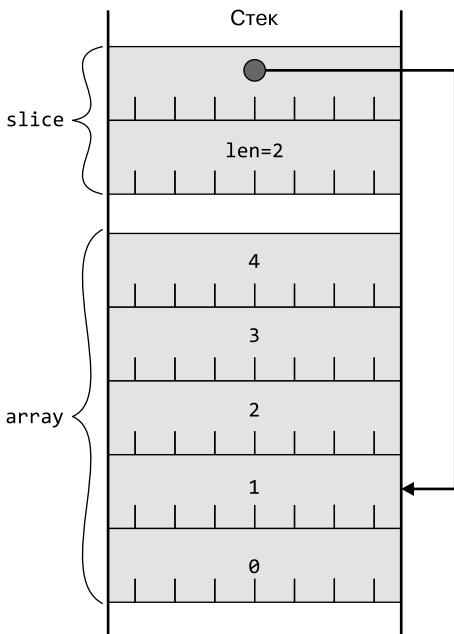


Рис. 1.4. Срез стека, указывающий на массив стека

Содержимое `Vec` хранится в куче подходящего размера, оно всегда непрерывно, а значит, срез может указывать на подмножество вектора (рис. 1.5):

```
let mut vector = Vec::<u64>::with_capacity(8);
for i in 0..5 {
    vector.push(i);
}
let vslice = &vector[1..3];
```

С выражением `&vector[1..3]` происходит многое, поэтому стоит расписать все это по частям.

- Сегмент `1..3` отражает диапазон. Компилятор преобразует его в экземпляр с типом `Range<usize>`, который включает нижнюю границу, но исключает верхнюю.
- Тип `Range` реализует трейт `SliceIndex<T>`, который описывает операции индексирования срезов произвольного типа `T`, поэтому на выходе (`Output`) получается тип `[T]`.
- Сегмент `vector[]` представляет индексирующее выражение. Компилятор преобразует его в вызов метода `index` трейта `Index` для `vector` вместе с разыменовыванием, то есть `*vector.index()`¹.

¹ Равноценным трейтом для мутабельных выражений будет `IndexMut`.

- Следовательно, `vector[1..3]` вызывает `Index<I>` из реализации `Vec<T>`, в которой `I` должна быть экземпляром `SliceIndex<[u64]>`. Это работает, поскольку `Range<usize>` реализует `SliceIndex<[T]>` для любого `T`, включая `u64`.
- `&vector[1..3]` отменяет разыменование, в результате чего итоговым типом выражения оказывается `&[u64]`.

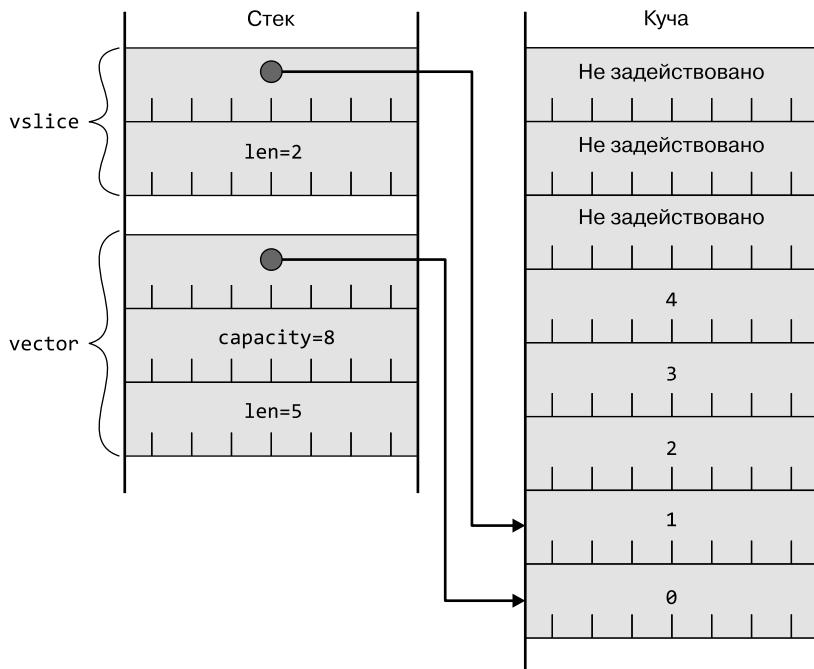


Рис. 1.5. Срез стека указывает на содержимое Vec в куче

Трейт-объекты

Вторым встроенным типом жирных указателей является *трейт-объект* — ссылка на некий элемент, который реализует конкретный трейт. Он состоит из простого указателя на элемент, дополненного внутренним указателем на таблицу *vtable* типа, в результате чего имеет размер 16 байт (на платформе x64). *Vtable* реализованного в типе трейта содержит указатели функции для реализации метода, что делает возможной динамическую диспетчеризацию в среде выполнения (см. рекомендацию 12)¹.

¹ Это упрощенный вариант — полная *vtable* включает также информацию о размере и выравнивании типа вместе с указателем функции `drop()`, чтобы соответствующий объект можно было безопасно отбросить.

Так что простой трейт:

```
trait Calculate {
    fn add(&self, l: u64, r: u64) -> u64;
    fn mul(&self, l: u64, r: u64) -> u64;
}
```

с реализацией его struct:

```
struct Modulo(pub u64);

impl Calculate for Modulo {
    fn add(&self, l: u64, r: u64) -> u64 {
        (l + r) % self.0
    }
    fn mul(&self, l: u64, r: u64) -> u64 {
        (l * r) % self.0
    }
}

let mod3 = Modulo(3);
```

можно преобразовать в трейт-объект с типом `&dyn Trait`. Ключевое слово `dyn` говорит о том, что здесь присутствует динамическая диспетчеризация:

```
// Нужен явный тип для реализации динамической диспетчеризации
let tobj: &dyn Calculate = &mod3;
let result = tobj.add(2, 2);
assert_eq!(result, 1);
```

В памяти все это будет выглядеть так, как показано на рис. 1.6.

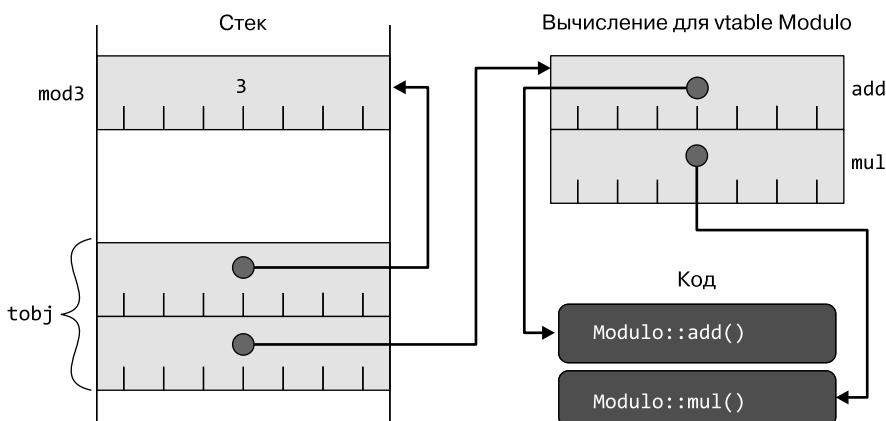


Рис. 1.6. Трейт-объект с указателями на конкретный элемент и vtable

Код, который содержит трейт-объект, может вызывать методы для этого трейта посредством указателей функций из vtable, передавая указатель на элемент в виде параметра `&self`. Дополнительная информация и советы приводятся в рекомендации 12.

Другие трейты указателей

В разделе «Трейты указателей» ранее описывались две пары трейтов, `Deref/DerefMut` и `AsRef/AsMut`, которые используются в работе с типами, легко преобразуемыми в ссылки. Есть и еще несколько стандартных трейтов — как в стандартной библиотеке, так и в пользовательских, которые тоже могут применяться в работе с типами указателей. Самый простой из них — трейт `Pointer`, который форматирует значение указателя для вывода. Это помогает при низкоуровневой отладке, и компилятор автоматически обращается к данному трейту, когда встречает спецификатор формата `{:p}`.

Более интересными являются трейты `Borrow` и `BorrowMut`, каждый из которых имеет один метод — `borrow` и `borrow_mut` соответственно. Сигнатуры этих методов совпадают с сигнатурой равнозначных им методов трейтов `AsRef/AsMut`.

Ключевое различие назначения этих трейтов можно увидеть с помощью предоставляемой стандартной библиотекой blanket-реализации. Для произвольной ссылки `&T` есть blanket-реализация как для `AsRef`, так и для `Borrow`. Аналогично для мутабельной ссылки `&mut T` есть blanket-реализация `AsMut` и `BorrowMut`.

У `Borrow` также есть blanket-реализация для не ссылочных типов:

```
impl<T> Borrow<T> for T
```

Это означает, что метод, получающий трейт `Borrow`, может одинаково успешно обрабатывать как сами экземпляры `T`, так и ссылки на него.

```
fn add_four<T: std::borrow::Borrow<i32>>(v: T) -> i32 {
    v.borrow() + 4
}
assert_eq!(add_four(&2), 6);
assert_eq!(add_four(2), 6);
```

Контейнерные типы стандартной библиотеки используют `Borrow` более практическим способом. Например, `HashMap::get` с его помощью делает возможным удобное извлечение записей, привязанных как по значению, так и по ссылке.

Трейт `ToOwned` основывается на трейте `Borrow`, добавляя метод `to_owned()`, который создает новый присвоенный (`owned`) элемент соответствующего типа. Это является обобщением трейта `Clone`: там, где `Clone` требует ссылку `&T`, трейт

`ToOwned` работает с элементами, которые реализуют `Borrow`. И это дает пару возможностей для единообразной обработки ссылок и перемещаемых элементов.

- Функция, которая работает со ссылками на некий тип, может получать `Borrow`, чтобы ее также можно было вызывать как со ссылками, так и с перемещаемыми элементами.
- Функция, которая работает с присвоенными элементами некоего типа, может получать `ToOwned`, чтобы ее также можно было вызывать как со ссылками на элементы, так и с перемещаемыми элементами. Любые передаваемые ей ссылки будут воспроизводиться в локально присвоенном элементе.

И хотя к типам указателя он не относится, здесь также следует упомянуть тип `Cow`, поскольку он предоставляет альтернативный способ обработки аналогичной ситуации. `Cow` — это `enum`, которое может содержать присвоенные данные или ссылку на заимствованные данные. Его специфическое имя означает `clone-on-write` — «клонирование при записи»: ввод `Cow` может сохраняться в виде заимствованных данных вплоть до момента, когда его нужно будет изменить, но в сам момент изменения он становится присвоенной копией.

Типы умных указателей

Стандартная библиотека Rust содержит различные типы, в той или иной степени действующие как указатели и опосредуемые ранее описанными трейтами. Каждый из этих типов *умных указателей* имеет собственную семантику и обеспечивает некие гарантии. Благодаря этому их правильная комбинация может обеспечивать детальный контроль над поведением указателя, но поначалу пугает своей сложностью (`Rc<RefCell<Vec<T>>` — как вам?).

Первым типом умного указателя является `Rc<T>` — указатель на элемент с подсчетом ссылок (подобен `std::shared_ptr<T>` в C++). Он реализует все связанные с указателями трейты, в связи с чем во многом действует как `Box<T>`.

Этот тип подходит для структур данных, в которых к элементу можно обращаться разными способами, но он нарушает одно из основных правил владения в Rust: «Каждый элемент должен иметь лишь одного владельца». Несоблюдение этого правила означает, что теперь становится возможной утечка данных: если элемент A содержит указатель `Rc` на элемент B, а элемент B содержит указатель `Rc` на элемент A, то эта пара никогда не будет отброшена¹. Иными словами, вам нужно, чтобы `Rc` поддерживал циклические структуры данных, но тогда у вас в структурах данных возникают циклы.

¹ Обратите внимание, что это не влияет на предоставляемые Rust гарантии безопасности памяти — элементы по-прежнему безопасны, просто недоступны.

В некоторых случаях риск утечки можно нивелировать с помощью сопутствующего типа `Weak<T>`, который содержит ссылку на внутренний элемент без владения (подобно `std::weak_ptr<T>` в C++). Присутствие слабой ссылки не мешает отбросить соответствующий элемент, когда все сильные ссылки будут удалены, поэтому использование `Weak<T>` подразумевает повышение до `Rc<T>`, которое может дать сбой.

На момент публикации книги `Rc` реализуется как пара указателей с подсчетом ссылок вместе с целевым элементом. Все это сохраняется в куче (рис. 1.7):

```
use std::rc::Rc;
let rc1: Rc<u64> = Rc::new(42);
let rc2 = rc1.clone();
let wk = Rc::downgrade(&rc1);
```

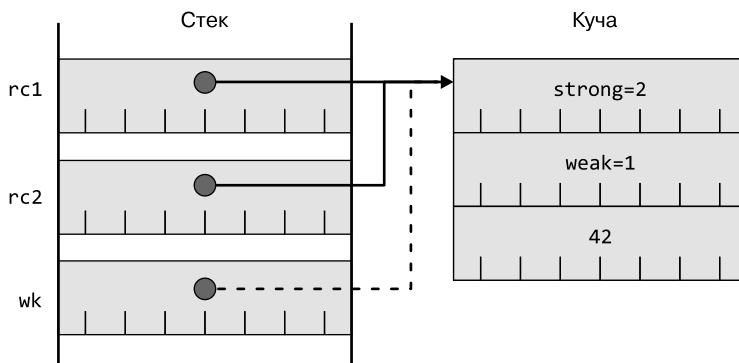


Рис. 1.7. Указатели `Rc` и `Weak`, ссылающиеся на один и тот же элемент кучи

Когда число сильных ссылок сокращается до нуля, внутренний элемент отбрасывается, но сама счетная структура отбрасывается, только когда до нуля сокращается и число слабых ссылок.

Сам по себе `Rc` дает возможность обращаться к элементу разными способами, но при обращении изменить этот элемент (посредством `get_mut`) можно, только если других способов обратиться к нему нет, то есть не существует других `Rc`- или `Weak`-ссылок на него. И организовать это сложно, поэтому `Rc` зачастую комбинируют с `RefCell`.

Умные указатели следующего типа, `RefCell<T>`, ослабляют правило, гласящее, что изменять элемент может только его владелец или код, содержащий единственную мутабельную ссылку на него (см. рекомендацию 15). Такая *внутренняя мутабельность* обеспечивает повышенную гибкость, например делает возможными реализации трейтов, изменяющие содержимое, даже когда сигнатура

метода допускает только `&self`. Тем не менее она несет сопутствующие затраты: помимо издержек для хранилища (дополнительный `isize` для отслеживания текущих заимствований, как показано на рис. 1.8) стандартные проверки заимствования перемещаются с этапа компиляции в среду выполнения:

```
use std::cell::RefCell;
let rc: RefCell<u64> = RefCell::new(42);
let b1 = rc.borrow();
let b2 = rc.borrow();
```

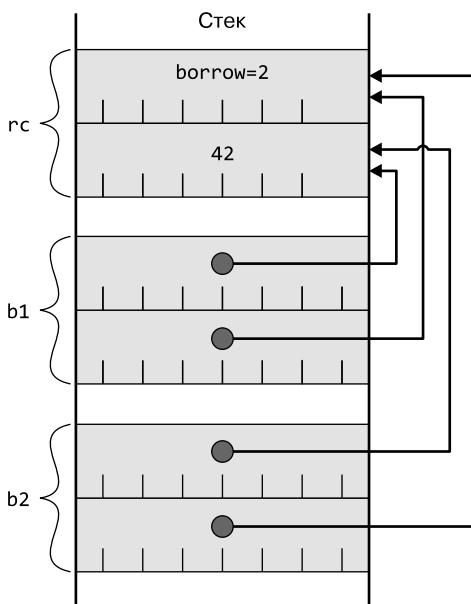


Рис. 1.8. Заимствования `Ref`,
ссылающиеся на контейнер `RefCell`

Привязанность этих проверок к среде выполнения означает, что пользователь `RefCell` должен выбирать из двух неприятных вариантов:

- принять то, что операция заимствования может провалиться, и работать со значениями `Result` из `try_borrow[_mut]`;
- использовать якобы безотказные методы заимствования `borrow[_mut]` и согласиться с риском возникновения `panic!` в среде выполнения, если правила заимствования окажутся нарушены.

В любом случае эта проверка в среде выполнения означает, что сам `RefCell` не реализует ни один из стандартных трейтов указателей. Вместо этого его операции доступа возвращают реализующий их тип умного указателя `Ref<T>` или `RefMut<T>`.

Если внутренний тип T реализует трейт `Copy`, указывающий, что быстрое поразрядное копирование создает валидный элемент (подробнее — в рекомендации 10), тогда тип `Cell<T>` позволяет выполнять внутреннюю мутацию с меньшими издержками — метод `get(&self)` копирует текущее значение, а метод `set(&self, val)` вставляет на его новое место. Реализации `Rc` и `RefCell` используют тип `Cell` для общего отслеживания счетчиков, которые могут изменяться без `&mut self`.

Описанные типы умных указателей подходят только для однопоточного применения. Их реализации предполагают, что к содержимому не будет параллельного доступа. Если это не так, тогда нужны уже умные указатели, включающие дополнительные издержки на синхронизацию.

Потокобезопасным эквивалентом для `Rc<T>` является `Arc<T>`, который использует атомарные счетчики для обеспечения точности числа подсчитанных ссылок. Как и `Rc`, `Arc` реализует всевозможные трейты указателей.

Но сам по себе `Arc` не позволяет осуществлять какой-либо мутабельный доступ к внутреннему элементу. Для этого используется тип `Mutex`, который обеспечивает, чтобы доступ — и мутабельный, и немутабельный — к этому элементу был только у одного потока. Как и в случае с `RefCell`, сам `Mutex` не реализует никакие трейты указателей, но его операция `lock()` возвращает значение с типом, который это делает, — `MutexGuard`, реализующий `Deref[Mut]`.

Если предполагается больше операций чтения, чем записи, то предпочтительнее будет тип `RwLock`, так как он позволяет множеству читающих обращаться к соответствующему элементу параллельно при условии, что в текущий момент нет ни одного записывающего.

В любом случае правила заимствования и обработки потоков в Rust вынуждают применять в многопоточном коде один из этих контейнеров синхронизации (но это защищает лишь от *некоторых* проблем параллельного использования общих данных; подробнее — в рекомендации 17).

Аналогичную стратегию — посмотреть, что компилятор отклоняет и что предлагает взамен, — иногда можно применить и к другим типам умных указателей. Однако быстрее и проще разобраться в том, что предполагает поведение различных умных указателей.

Позаимствуем пример из первой редакции книги по Rust (<https://oreil.ly/DOKi>):

- `Rc<RefCell<Vec<T>>>` содержит вектор (`Vec`) с совместно используемым владением (`Rc`), при котором вектор может изменяться, но только целиком;
- `Rc<Vec<RefCell<T>>>` также содержит вектор с совместно используемым владением, но здесь каждая отдельная запись вектора может изменяться независимо от других.

Используемые типы в точности описывают это поведение.

Рекомендация 9. Рассмотрите использование преобразований итераторов вместо явных циклов

Смиренный цикл проделал длинный путь для повышения удобства и расширения абстракции. В языке В (предшественник С) (<https://oreil.ly/OeiaI>) была только инструкция `while (condition) { ... }`, но с появлением С реализацию типичного сценария перебора индексов массива упростило добавление цикла `for`:

```
// Код C
int i;
for (i = 0; i < len; i++) {
    Item item = collection[i];
    // Тело
}
```

В первых версиях С++ удобство, в том числе определения области действия, было повышенено за счет возможности вложения объявления переменной цикла в инструкцию `for` (эта же функциональность была добавлена в C99):

```
// Код C++98
for (int i = 0; i < len; i++) {
    Item item = collection[i];
    // ...
}
```

Большинство современных языков абстрагируют концепцию цикла еще больше: основная функция цикла зачастую перемещается в следующий элемент или некий контейнер. Отслеживание логистики, необходимой для достижения этого элемента (`index++` или `++it`), по сути, значения не имеет, из чего родились две фундаментальные концепции.

Итераторы

Тип, чьей задачей является повторяющаяся выдача очередного элемента контейнера вплоть до его опустошения¹.

Циклы *for-each*

Компактное выражение цикла для перебора всех элементов контейнера, привязки переменной цикла к *элементу*, а не к деталям доступа к нему.

¹ Фактически итератор может быть более обобщенным — идею выдачи очередных элементов до их исчерпания не нужно ассоциировать с контейнером.

Эти принципы позволяют писать более короткий и, самое главное, понятный код цикла:

```
// Код C++11
for (Item& item : collection) {
    // ...
}
```

Как только эти принципы стали доступными, оказалось, что они необыкновенно мощные, и языки, в которых их не было, оперативно начали их добавлять (например, циклы for-each были добавлены в Java 1.5 и C++11).

В Rust есть итераторы и циклы for-each, но он предлагает и еще один уровень абстракции, позволяя выразить весь цикл в виде *преобразования итератора* (иногда он называется *итератором-адаптером*). Как и в рекомендации 3, посвященной Option и Result, в текущей рекомендации мы постараемся показать, как эти преобразования итераторов можно использовать вместо явных циклов, а также подскажем, когда это делать уместно. В частности, преобразования итераторов часто оказываются эффективнее явного цикла, потому что компилятор иногда пропускает проверку границ, которая в противном случае может быть необходима.

К концу этой рекомендации реализация C-подобного цикла суммирования квадратов первых пяти четных элементов вектора:

```
let values: Vec<u64> = vec![1, 1, 2, 3, 5 /* ... */];

let mut even_sum_squares = 0;
let mut even_count = 0;
for i in 0..values.len() {
    if values[i] % 2 != 0 {
        continue;
    }
    even_sum_squares += values[i] * values[i];
    even_count += 1;
    if even_count == 5 {
        break;
    }
}
```

должна ощущаться вами как более естественная в виде функционального выражения:

```
let even_sum_squares: u64 = values
    .iter()
    .filter(|x| *x % 2 == 0)
    .take(5)
    .map(|x| x * x)
    .sum();
```

Подобные выражения преобразования итераторов можно примерно разделить на три части:

- изначальный итератор из экземпляра типа, реализующего один из трейтов итератора Rust;
- последовательность преобразований итератора;
- конечный метод потребителя, совмещающий результаты итерации в итоговое значение.

Первые две части, по сути, перемещают функциональность из тела цикла в выражение `for`. Последняя полностью исключает потребность в инструкции `for`.

Трейты итераторов

Основной трейт `Iterator` имеет очень простой интерфейс: единственный метод `next`, который выдает некие (`Some`) элементы, пока те не закончатся (`None`). Тип выдаваемых элементов определяется связанным с этим трейтом типом `Item`.

Коллекции, которые позволяют перебирать их содержимое (в других языках они называются *итерируемыми*), реализуют трейт `IntoIterator`. Метод `into_iter` этого трейта получает `Self`, выдавая вместо него `Iterator`. Компилятор автоматически использует этот трейт для выражений следующей формы:

```
for item in collection {
    // Тело
}
```

по сути преобразуя их в код примерно так:

```
let mut iter = collection.into_iter();
loop {
    let item: Thing = match iter.next() {
        Some(item) => item,
        None => break,
    };
    // Тело
}
```

или более лаконично и идиоматично:

```
let mut iter = collection.into_iter();
while let Some(item) = iter.next() {
    // Тело
}
```

Чтобы все выполнялось плавно, для каждого `Iterator` есть реализация `IntoIterator`, которая просто возвращает `self`. В конце концов, несложно преобразовать `Iterator` в `Iterator`.

Эта изначальная форма представляет собой потребляющий итератор, который использует коллекцию по мере ее создания:

```
let collection = vec![Thing(0), Thing(1), Thing(2), Thing(3)];
for item in collection {
    println!("Consumed item {item:?}");
}
```

Любая попытка использовать эту коллекцию после ее перебора провалится:

```
println!("Collection = {collection:?}");

error[E0382]: borrow of moved value: `collection`
--> src/main.rs:171:28
|
163 |     let collection = vec![Thing(0), Thing(1), Thing(2), Thing(3)];
|           ----- move occurs because `collection` has type `Vec<Thing>`,
|           which does not implement the `Copy` trait
164 |     for item in collection {
|           ----- `collection` moved due to this implicit call to
|           ` .into_iter()`
...
171 |     println!("Collection = {collection:?}");
|           ^^^^^^^^^^^^^^^^^ value borrowed here after move
|
note: `into_iter` takes ownership of the receiver `self`, which moves
`collection`
```

И хотя понять это просто, такое всепоглощающее поведение часто оказывается нежелательным. Здесь требуется некий вид *займствования*.

Для более отчетливой демонстрации поведения в приведенных здесь примерах используется тип `Thing`, который *не реализует Copy* (см. рекомендацию 10), так как это скрыло бы вопросы относительно владения (см. рекомендацию 15), поскольку компилятор молча делал бы повсюду копии:

```
// Намеренно без `Copy`
#[derive(Clone, Debug, Eq, PartialEq)]
struct Thing(u64);

let collection = vec![Thing(0), Thing(1), Thing(2), Thing(3)];
```

Если перед итерируемой коллекцией используется префикс `&`:

```
for item in &collection {
    println!("{}", item.0);
}
println!("collection still around {collection:?}");
```

то компилятор Rust будет искать реализацию `IntoIterator` для типа `&Collection`. Правильно присвоенные типы коллекции предоставят такую реализацию.

Она по-прежнему будет получать `Self`, но теперь `Self` является `&Collection`, а не `Collection` и связанный с этой коллекцией тип `Item` будет ссылкой `&Thing`.

В результате коллекция после перебора остается неизменной и равнозначный развернутый код будет выглядеть так:

```
let mut iter = (&collection).into_iter();
while let Some(item) = iter.next() {
    println!("{}", item.0);
}
```

Если есть смысл обеспечить перебор мутабельных ссылок¹, тогда аналогичный паттерн применяется к `for item in &mut collection`: компилятор ищет места использования реализации `IntoIterator` для `&mut Collection`, где каждый `Item` имеет тип `&mut Thing`.

По соглашению стандартные контейнеры предоставляют также метод `iter()`, который возвращает итератор ссылок на соответствующий элемент и, если требуется, эквивалентный метод `iter_mut()` с только что описанным поведением. Эти методы можно задействовать в циклах `for`, но их польза более наглядна, когда они применяются в качестве начала преобразования итератора. Так:

```
let result: u64 = (&collection).into_iter().map(|thing| thing.0).sum();
```

становится:

```
let result: u64 = collection.iter().map(|thing| thing.0).sum();
```

Преобразования итераторов

В трейте `Iterator` есть единственный необходимый метод (`next`), но он предоставляет базовые реализации (см. рекомендацию 13) множества других методов, выполняющих преобразования итератора.

Некоторые из этих преобразований влияют на общий процесс итерации.

`take(n)`

Ограничивает итератор выдачей не более чем n элементов.

`skip(n)`

Пропускает первые n элементов итератора.

¹ Этот метод нельзя предоставить, если мутация элемента может аннулировать внутренние гарантии контейнера. Например, изменение содержимого элемента таким образом, который изменит его значение `Hash`, сделает недействительными внутренние структуры данных `HashMap`.

step_by(n)

Преобразует итератор так, чтобы он выдавал только каждый n -й элемент.

chain(other)

Связывает два итератора для создания единого, который переходит из одного в другой.

cycle()

Преобразует итератор, который превращается в повторяющийся бесконечно, возвращаясь к началу после каждого завершения. (Для этого итератор должен поддерживать `Clone`.)

rev()

Реверсирует направление итератора. (Итератор должен реализовывать трейт `Double EndedIterator`, который содержит дополнительный обязательный метод `next_back`.)

Другие трансформации влияют на природу `Item`, что относится уже к `Iterator`.

map(|item| { ... })

Циклически применяет замыкание для поочередного преобразования каждого элемента. Это наиболее обобщенная версия. Несколько следующих записей списка являются вариантами, которые для удобства можно реализовать и в виде `map`.

cloned()

Создает клоны всех элементов изначального итератора. Это особенно полезно в случае итераторов, перебирающих ссылки `&Item`. (Естественно, для этого нужно, чтобы соответствующий тип `Item` реализовывал `Clone`.)

copied()

Создает копии всех элементов в изначальном итераторе. Это особенно полезно в случае итераторов, перебирающих ссылки `&Item`. (Естественно, для этого нужно, чтобы соответствующий тип `Item` реализовывал `Copy`, но сама операция наверняка будет быстрее `cloned()`.)

enumerate()

Преобразует итератор элементов в итератор пар (`usize, Item`), предоставляя индекс элементов в этом итераторе.

zip(it)

Объединяет итератор с другим итератором, создавая общий, который выдает пары элементов, по одному из каждого исходного итератора, пока не завершится самый короткий.

Еще есть преобразования, которые выполняют фильтрацию `Item`, выдаваемых `Iterator`.

filter(|item| {...})

Применяет к каждой ссылке на элемент замыкание, возвращающее `bool`, чтобы определить, должен ли этот элемент передаваться.

take_while()

Выдает изначальный поддиапазон итератора на основе предиката. Является зеркальным отражением `skip_while`.

skip_while()

Выдает итоговый поддиапазон итератора на основе предиката. Является зеркальным отражением `take_while`.

Метод `flatten()` работает с итератором, элементы которого сами являются итераторами, уплощая результат. Сам по себе он не особо полезен, но пригодится в случаях, когда мы видим, что `Option` и `Result` выступают в виде итераторов: они либо не выдают элементы (для `None`, `Err(e)`), либо выдают один (для `Some(v)`, `Ok(v)`). Это означает, что уплощение (`flatten`) потока значений `Option/Result` обеспечивает простой способ извлечь только валидные из них, игнорируя остальные.

Если брать в целом, то эти методы позволяют преобразовывать итераторы так, чтобы они выдавали последовательность конкретно тех элементов, которые нужны в большинстве ситуаций.

Потребители итераторов

В предыдущих двух разделах описывалось, как получить итератор и как преобразовать его в подходящую для точной итерации форму. Такую итерацию можно реализовать в виде явного цикла `for-each`:

```
let mut even_sum_squares = 0;
for value in values.iter().filter(|x| *x % 2 == 0).take(5) {
    even_sum_squares += value * value;
}
```

Однако большая коллекция методов `Iterator` содержит множество таких, которые позволяют получать итерацию в одном вызове метода, исключая потребность в явном цикле `for`.

Самым общим из этих методов является `for_each(|item| {...})`, который выполняет замыкание для каждого выдаваемого `Iterator` элемента. Этот метод может проделать *большую* часть из того, что может явный цикл `for` (исключения будут описаны в одном из следующих разделов), но его обобщенность делает его не совсем удобным — замыканию для выдачи чего-либо необходимо использовать мутабельные ссылки на внешние состояния:

```
let mut even_sum_squares = 0;
values
    .iter()
    .filter(|x| *x % 2 == 0)
    .take(5)
    .for_each(|value| {
        // Замыканию нужна мутабельная ссылка на внешнее состояние
        even_sum_squares += value * value;
    });
});
```

Тем не менее, если тело цикла `for` соответствует одному из нескольких типичных паттернов, есть конкретные получающие итератор методы — более понятные, сжатые и идиоматичные. К этим паттернам относятся сокращения команд для получения из коллекции одного значения.

`sum()`

Суммирует коллекцию численных значений (целых или с плавающей запятой).

`product()`

Перемножает коллекцию численных значений.

`min()`

Находит минимальное значение коллекции относительно `Ord` реализации элемента (подробнее — в рекомендации 10).

`max()`

Находит максимальное значение коллекции относительно `Ord` реализации элемента (подробнее — в рекомендации 10).

`min_by(f)`

Находит минимальное значение коллекции относительно пользовательской функции сравнения `f`.

`max_by(f)`

Находит максимальное значение коллекции относительно пользовательской функции сравнения `f`.

`reduce(f)`

Создает аккумулированное значение с типом `Item`, выполняя на каждом этапе замыкание, которое получает аккумулированное к этому моменту значение и текущий элемент. Это более обобщенная операция, которая включает в себя предыдущие методы.

`fold(f)`

Создает аккумулированное значение произвольного типа (не только типа `Iterator::Item`), выполняя на каждом шаге замыкание, которое получает аккумулированное к этому моменту значение и текущий элемент. Это более обобщенная форма `reduce`.

scan(init, f)

Создает аккумулированное значение произвольного типа, выполняя на каждом шаге замыкание, которое получает мутабельную ссылку на некое внутреннее состояние и текущий элемент. Это несколько измененное обобщение `reduce`.

Еще есть методы для *выбора* из коллекции одного значения.

find(p)

Находит первый элемент, удовлетворяющий предикату.

position(p)

Также находит первый элемент, удовлетворяющий предикату, но возвращает индекс этого элемента.

nth(n)

Возвращает *n*-й элемент итератора, если таковой имеется.

Существуют методы для тестирования каждого элемента коллекции.

any(p)

Показывает, верен ли (`true`) предикат для *любого* элемента коллекции.

all(p)

Показывает, верен ли (`true`) предикат для *всех* элементов коллекции.

В обоих случаях итерация прекратится, как только будет найден соответствующий контрпример.

Существуют методы, которые допускают возможность сбоя в замыканиях, используемых каждым элементом. В случае, когда замыкание возвращает в отношении элемента сбой, итерация прекращается и вся операция возвращает информацию об этом первом сбое.

try_for_each(f)

Действует как `for_each`, но замыкание может провалиться.

try_fold(f)

Действует как `fold`, но замыкание может провалиться.

try_find(f)

Действует как `find`, но замыкание может провалиться.

Наконец, есть методы, которые собирают все перебираемые элементы в новую коллекцию. Самый значимый из них — `collect()`, его можно использовать для создания нового экземпляра любого типа коллекции, реализующего трейт `FromIterator`.

Трейт `FromIterator` реализуется для всех типов коллекций, имеющихся в стандартной библиотеке (`Vec`, `HashMap`, `BTreeSet` и т. д.), но эта универсальность означает также, что вам придется часто использовать явные типы, поскольку в противном случае компилятор не сможет понять, хотите вы собрать, предположим, `Vec<i32>` или `HashSet<i32>`:

```
use std::collections::HashSet;

// Создает коллекции из четных чисел. Тип должен быть указан,
// поскольку выражение одинаково для любого из них.
let myvec: Vec<i32> = (0..10).into_iter().filter(|x| x % 2 == 0).collect();
let h: HashSet<i32> = (0..10).into_iter().filter(|x| x % 2 == 0).collect();
```

Этот пример демонстрирует также применение выражений диапазона для генерации изначальных данных, которые нужно перебрать.

Вот еще пара методов (менее ясных) для создания коллекций.

`unzip()`

Делит итератор пар на две коллекции.

`partition(p)`

Делит итератор на две коллекции на основе применения к каждому элементу предиката.

В текущей рекомендации был затронут широкий список методов `Iterator`, но это лишь часть доступных. Более подробно о них рассказывается в документации, касающейся итераторов, или в главе 15 книги *Programming Rust*, 2-е издание (O'Reilly), где детально разбираются все возможности.

Эта богатая коллекция преобразований итераторов создана для того, чтобы ею пользовались. С их помощью вы можете создавать более идиоматический, компактный код с ясным назначением.

Выражение циклов в виде преобразований итераторов может помочь писать более эффективный код. Для обеспечения безопасности Rust выполняет *проверку границ* при доступе к непрерывным контейнерам, таким как векторы и срезы. Попытка обратиться к значению за пределами коллекции вместо фактического доступа к недопустимым данным вызывает панику. Старомодный цикл, который обращается к значениям контейнера (например, `values[i]`), может подвергаться этим проверкам среди выполнения, в то время как итератор, создающий значения поочередно, определенно находится в рамках допустимого диапазона.

Есть и еще один факт. Старомодный цикл также *не может* подвергаться дополнительным проверкам границ по сравнению с равнозенным преобразованием

итератора. Компилятор и оптимизатор Rust очень хорошо справляются с анализом кода вокруг среза для определения того, безопасно ли пропускать проверки границ. Связанные с этим тонкости наглядно разбираются в статье Сергея (Shnatsel) Давыдова, опубликованной в 2023 году (<https://oreil.ly/impAX>).

Создание коллекций из значений Result

В предыдущем разделе описывалось использование `collect()` для создания коллекций из итераторов, но `collect()` имеет также отдельную полезную функцию `Result` для работы со значениями.

Рассмотрим попытку преобразовать вектор значений `i64` в байты (`u8`) с оптимистичной надеждой, что они все впишутся.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// В версии Rust 2021 года `TryFrom` находится в начале документа,
// так что инструкция `use` больше не нужна
use std::convert::TryFrom;

let inputs: Vec<i64> = vec![0, 1, 2, 3, 4];
let result: Vec<u8> = inputs
    .into_iter()
    .map(|v| <u8>::try_from(v).unwrap())
    .collect();
```

Такой код будет работать, пока не встретит какой-нибудь неожиданный ввод:

```
let inputs: Vec<i64> = vec![0, 1, 2, 3, 4, 512];
```

который вызовет сбой выполнения:

```
thread 'main' panicked at 'called `Result::unwrap()` on an `Err` value:
TryFromIntError(())', iterators/src/main.rs:266:36
note: run with `RUST_BACKTRACE=1` environment variable to display a backtrace
```

Следуя рекомендации 3, нам нужно сохранить в игре тип `Result` и использовать оператор `?`, чтобы сделать любой сбой проблемой вызывающего кода. Очевидное изменение на генерацию `Result` проблемы не решает:

```
let result: Vec<Result<u8, _>> =
    inputs.into_iter().map(|v| <u8>::try_from(v)).collect();
// И что теперь? Для извлечения результатов и обнаружения ошибок
// нам по-прежнему нужна итерация.
```

Тем не менее есть альтернативная версия `collect()`, которая может собирать не `Vec`, содержащий `Result`, а `Result`, содержащий `Vec`.

Чтобы сделать использование этой версии принудительным, потребуется прием под названием `turbofish (>::<Result<Vec<_>, _>>)`:

```
let result: Vec<u8> = inputs
    .into_iter()
    .map(|v| <u8>::try_from(v))
    .collect::<Result<Vec<_>, _>>()?
```

Совместив этот подход с оператором `?`, мы получим полезное поведение.

- Если во время итерации встретится ошибочное значение, оно будет отправлено вызывающему коду и итерация прекратится.
- Если ошибок не встретится, остаток кода сможет работать с адекватной коллекцией значений правильного типа.

Преобразование циклов

Цель текущей рекомендации — убедить вас в том, что многие явные циклы можно рассматривать как предмет для превращения в преобразования итераторов. Программистам, которые к такому не привыкли, это решение может показаться неестественным, поэтому разберем его пошагово.

Начнем с C-подобного явного цикла, суммирующего квадраты первых пяти четных элементов вектора:

```
let mut even_sum_squares = 0;
let mut even_count = 0;
for i in 0..values.len() {
    if values[i] % 2 != 0 {
        continue;
    }
    even_sum_squares += values[i] * values[i];
    even_count += 1;
    if even_count == 5 {
        break;
    }
}
```

Первым шагом будет замена индексации вектора прямым использованием итератора в цикле `for-each`:

```
let mut even_sum_squares = 0;
let mut even_count = 0;
```

```
for value in values.iter() {
    if value % 2 != 0 {
        continue;
    }
    even_sum_squares += value * value;
    even_count += 1;
    if even_count == 5 {
        break;
    }
}
```

Первая ветка цикла, которая использует `continue` для пропуска некоторых элементов, естественным образом выражается как `filter()`:

```
let mut even_sum_squares = 0;
let mut even_count = 0;
for value in values.iter().filter(|x| *x % 2 == 0) {
    even_sum_squares += value * value;
    even_count += 1;
    if even_count == 5 {
        break;
    }
}
```

Далее ранний выход из цикла после обработки пяти элементов отображается в `take(5)`:

```
let mut even_sum_squares = 0;
for value in values.iter().filter(|x| *x % 2 == 0).take(5) {
    even_sum_squares += value * value;
}
```

Каждая итерация цикла использует только элемент в квадрате в виде комбинации `value * value`, что делает ее идеальной целью для `map()`:

```
let mut even_sum_squares = 0;
for val_sqr in values.iter().filter(|x| *x % 2 == 0).take(5).map(|x| x * x) {
    even_sum_squares += val_sqr;
}
```

В результате рефакторинга изначального цикла его тело становится прекрасным гвоздем для удара молотком метода `sum()`:

```
let even_sum_squares: u64 = values
    .iter()
    .filter(|x| *x % 2 == 0)
    .take(5)
    .map(|x| x * x)
    .sum();
```

Когда явный вариант лучше

Текущая рекомендация подчеркнула преимущества преобразований итераторов, в частности, в плане лаконичности и ясности. А когда эти преобразования будут *неподходящими* или *неidiоматическими*?

- Если тело цикла велико и/или многофункционально, разумнее будет оставить его в явном виде, чем сворачивать в замыкание.
- Если тело цикла содержит условия ошибок, которые ведут к раннему завершению внешней функции, то их обычно лучше оставлять явными — методы `try_..()` помогают здесь не сильно. Тем не менее возможность `collect()` преобразовывать коллекцию значений `Result` в `Result`, содержащий коллекцию значений, зачастую позволяет также обрабатывать условия ошибок с помощью оператора `?`.
- Если производительность крайне важна, преобразование итератора, включающее замыкание, *нужно оптимизировать*, чтобы оно стало таким же быстрым, как равнозначный явный код. Но если так важна производительность внутреннего цикла, *измерьте* различные варианты и внесите нужные корректировки.
 - Обязательно проследите, чтобы ваши измерения отражали реальную производительность, — оптимизатор компилятора может давать слишком оптимистичные результаты на тестовых данных (об этом говорится в рекомендации 30).
 - Проводник компилятора Godbolt (<https://rust.godbolt.org>) — это прекрасный инструмент для изучения его вывода.

Самое главное — не превращайте цикл в преобразование итератора, если это превращение вынужденное или неудобное. Все это определенно дело вкуса, но имейте в виду, что по мере ознакомления с функциональным стилем программирования ваш вкус наверняка изменится.

ГЛАВА 2

Трейты

Вторым столпом системы типов в Rust является использование трейтов, которое позволяет кодировать поведение, стандартное для разных типов. Сам trait можно сравнить с типом интерфейса в других языках, но в Rust трейты привязаны также к *обобщенным типам* (или дженерикам; см. рекомендацию 12), позволяя повторно использовать интерфейс без лишних издержек для среды выполнения.

Рекомендации текущей главы описывают стандартные трейты, которые компилятор и цепочка инструментов Rust делают доступными, а также дают советы относительно того, как проектировать и использовать закодированное в трейтах поведение.

Рекомендация 10. Изучите стандартные трейты

Ключевые поведенческие аспекты системы типов Rust кодирует в самой системе типов посредством коллекции отточенных стандартных трейтов, которые это поведение описывают (подробнее — в рекомендации 2).

Многие из этих трейтов покажутся знакомыми программистам, пришедшим из мира C++, поскольку соответствуют таким механизмам, как конструкторы копий, деструкторы, операторы равенства и присваивания и т. д. Как и в C++, здесь обычно желательно реализовывать многие из трейтов для собственных типов. Компилятор Rust будет выдавать полезные сообщения об ошибках, если какой-то операции для вашего типа потребуется один из этих трейтов, но его не окажется.

Реализация подобной обширной коллекции трейтов может показаться утомительной, но большинство их типичных вариантов можно применять к пользовательским типам автоматически, с помощью макросов `derive`. Эти макросы

генерируют код с очевидной реализацией трейта для данного типа, например сравнение по полям для `Eq` в `struct`. Обычно для этого требуется, чтобы все составные части также реализовывали трейт. Автоматически генерируемая реализация в большинстве случаев окажется подходящим решением, но бывают исключения, о которых мы поговорим далее в разделах, посвященных разным трейтам.

Использование макросов `derive` позволяет создавать определения типов наподобие такого:

```
#[derive(Clone, Copy, Debug, PartialEq, Eq, PartialOrd, Ord, Hash)]
enum MyBooleanOption {
    Off,
    On,
}
```

где автоматическая генерация реализаций активируется для восьми разных трейтов.

Такая детальная спецификация поведения поначалу может озадачивать, но важно познакомиться с самыми распространенными из стандартных трейтов, чтобы с ходу понимать, как может вести себя определение типа.

Типичные стандартные трейты

В этом разделе приводятся самые распространенные стандартные трейты. Вот их краткое описание.

Clone

Элементы этого типа по запросу могут делать копию самих себя, выполняя пользовательский код.

Copy

Если компилятор делает побитовую копию представления этого элемента в памяти (не выполняя пользовательский код), в результате получается новый валидный элемент.

Default

Можно создавать новые экземпляры этого типа с подходящими предустановленными значениями.

PartialEq

Между элементами этого типа существует частичная равнозначность — любые два элемента всегда можно сравнить, но нет гарантии, что их результатом обязательно будет `x==x`.

Eq

Между элементами этого типа существует равнозначность — любые два элемента всегда можно сравнивать, и обязательно окажется `x==x`.

PartialOrd

Некоторые элементы этого типа можно сравнивать и упорядочивать.

Ord

Все элементы этого типа можно сравнивать и упорядочивать.

Hash

Элементы этого типа по запросу могут создавать стабильный хеш своего содержимого.

Debug

Элементы этого типа можно выводить на экран для программистов.

Display

Элементы этого типа можно выводить на экран для пользователей.

Все эти трейты можно выводить (`derive`) для пользовательских типов, за исключением `Display` (был добавлен из-за пересечения с `Debug`). Однако бывают случаи, когда предпочтительнее ручная реализация или ее отсутствие.

В последующих разделах каждый из представленных трейтов разбирается более подробно.

Clone

Трейт `Clone` показывает, что можно сделать копию элемента, вызвав метод `clone()`. Это подобно конструктору копий в C++, но происходит более явно: компилятор никогда не станет молча вызывать этот метод сам по себе (подробнее — в следующем разделе).

`Clone` можно вывести (`derive`) для любого типа, если все поля соответствующего элемента сами реализуют `Clone`. Выведенная таким образом реализация клонирует агрегированный тип, поочередно копируя каждый его член. Опять же это похоже на предустановленный конструктор копий в C++. Такое решение делает трейт используемым по запросу (путем добавления `#[_]derive(Clone)]`), что противоположно поведению C++, где от его применения нужно отказываться (`MyType(const MyType&) = delete;`).

Это настолько типичная и полезная операция, что более интересно изучить ситуации, когда не следует или нельзя реализовать `Clone` или предустановленная реализация `derive` не подходит.

- Вам *не следует* реализовывать `Clone`, если элемент воплощает в себе уникальный доступ к некоему ресурсу (такому как тип RAII; см. рекомендацию 11) или есть другая причина для ограничения создания копий (например, если элемент содержит криптографический материал ключа).
- Вы *не можете* реализовать `Clone`, если некий компонент вашего типа является неклонируемым. Вот примеры:
 - поля, которые являются мутабельными ссылками (`&mut T`), поскольку анализатор заимствований (см. рекомендацию 15) позволяет одновременно использовать только одну мутабельную ссылку;
 - типы стандартной библиотеки, которые относятся к предыдущей категории, такие как `MutexGuard` (реализует уникальный доступ) или `Mutex` (ограничивает создание копий для обеспечения безопасности потоков).

Вам следует *вручную реализовывать* `Clone`, если в вашем элементе есть что-то, что не будет захвачено рекурсивным поочередным копированием полей, или если присутствует дополнительный механизм подсчета, связанный с временем жизни элемента. Рассмотрим в качестве примера тип, который отслеживает количество существующих элементов в среде выполнения для получения метрики. Ручная реализация `Clone` позволит обеспечить сохранение счетчиком точности.

Copy

Объявляется трейт `Copy` весьма тривиально:

```
pub trait Copy: Clone { }
```

В нем нет методов, то есть он является *трейтом-маркером* (описывались в рекомендации 2) — указывает на наличие некоего ограничения типа, которое непосредственно в системе типов не выражено.

Смысль `Copy` в том, что поразрядная копия памяти, содержащей элемент, дает корректный новый элемент. По сути, этот трейт является маркером, который гласит, что тип — это простая структура данных (plain old data, POD).

Это также означает, что граница трейта `Clone` может слегка озадачивать: несмотря на то что тип `Copy` должен реализовывать `Clone`, когда копируется экземпляр типа, метод `clone()` не вызывается — компилятор создает новый элемент без участия пользовательского кода.

В отличие от пользовательских трейтов-маркеров (см. рекомендацию 2) `Copy` особо значим для компилятора (как и несколько других трейтов-маркеров в `std::marker`) — помимо того что он доступен для границ трейта, он также переключает компилятор с *семантикой перемещения* на *семантику копирования*.

В случае семантики перемещения для оператора присваивания мы лишаемся исходного значения:

НЕ КОМПИЛИРУЕТСЯ

```
#[derive(Debug, Clone)]
struct KeyId(u32);

let k = KeyId(42);
let k2 = k; // Значение перемещается из k в k2
println!("k = {k:?}");
```

```
error[E0382]: borrow of moved value: `k`
--> src/main.rs:60:23
   |
58 |     let k = KeyId(42);
   |     - move occurs because `k` has type `main::KeyId`, which does
   |       not implement the `Copy` trait
59 |     let k2 = k; // Значение перемещается из k в k2
   |     - value moved here
60 |     println!("k = {k:?}");
   |             ^^^^^^ value borrowed here after move
   |
= note: this error originates in the macro `$crate::format_args_nl`
help: consider cloning the value if the performance cost is acceptable
   |
59 |     let k2 = k.clone(); // Значение перемещается из k в k2
   |             +++++++
```

В случае же семантики копирования исходный элемент продолжает существовать:

```
#[derive(Debug, Clone, Copy)]
struct KeyId(u32);

let k = KeyId(42);
let k2 = k; // Значение поразрядно копируется из k в k2
println!("k = {k:?}");
```

Это делает `Copy` одним из самых важных трейтов, который нужно иметь в виду: он фундаментально меняет поведение присваивания, включая параметры для вызова методов. Здесь снова наблюдаются пересечения с конструкторами копий C++, но с одним ключевым отличием: в Rust невозможно, чтобы компилятор молча вызвал пользовательский код — либо это происходит явно (вызов к `.clone()`), либо код не пользовательский (поразрядное копирование).

Поскольку `Copy` содержит границу трейта `Clone`, можно выполнить `.clone()` для любого допускающего копирование элемента. Однако делать это нежелательно, так как поразрядное копирование всегда будет быстрее, чем вызов метода трейта. `Clippy` (см. рекомендацию 29) предупредит вас об этом:

НЕ КОМПИЛИРУЕТСЯ

```
let k3 = k.clone();
```

```
warning: using `clone` on type `KeyId` which implements the `Copy` trait
--> src/main.rs:79:14
  |
79 |     let k3 = k.clone();
  |     ^^^^^^^^^ help: try removing the `clone` call: `k`
```

Как и в случае с `Clone`, стоит разобраться, когда вам следует реализовывать `Copy`, а когда — нет.

- Очевидно, что *не нужно реализовывать Copy, если поразрядное копирование не дает валидный элемент*. Это возможно, когда `Clone` нужна ручная реализация, а не выведенная автоматически через `derive`.
- Порой нежелательно реализовывать `Copy` в случае большого типа. Базовая гарантия `Copy` состоит в том, что поразрядная копия будет валидной. Тем не менее это зачастую сопровождается допущением, что копия создается быстро. Если это не так, пропуск `Copy` исключит случайные медленные копии.
- *Нельзя реализовать Copy, если некий компонент вашего типа является некопируемым.*
- Если все компоненты вашего типа допускают применение `Copy`, то обычно `Copy` следует выводить. В компиляторе есть отключенный по умолчанию линт `missing_copy_implementations`, который указывает на доступные для этого возможности.

Default

Трейт `Default` определяет *предустановленный конструктор* посредством метода `default()`. Этот трейт можно выводить (`derive`) для пользовательских типов при условии, что все подтипы тоже будут иметь реализацию `Default`. Если же нет, вам потребуется реализовать трейт вручную. Продолжая сравнение с C++, обратите внимание на то, что предустановленный конструктор нужно активировать явно — компилятор не создает их автоматически.

Трейт `Default` можно выводить также для типов `enum` при наличии атрибута `#[_default]`, который подскажет компилятору, какой вариант является предустановленным:

```
#[derive(Default)]
enum IceCreamFlavor {
    Chocolate,
    Strawberry,
    #[default]
    Vanilla,
}
```

Самый полезный аспект трейта `Default` — это его совмещение с *синтаксисом обновления структур* (<https://oreil.ly/DJW-U>). Этот синтаксис позволяет инициализировать поля `struct` путем копирования или перемещения их содержимого из существующего экземпляра той же `struct` и действует для любых неинициализированных полей. Шаблон, из которого нужно копировать, указывается в конце инициализации после `...`, и трейт `Default` предоставляет его идеальный вариант:

```
#[derive(Default)]
struct Color {
    red: u8,
    green: u8,
    blue: u8,
    alpha: u8,
}

let c = Color {
    red: 128,
    ..Default::default()
};
```

Это существенно упрощает инициализацию структур с множеством полей, лишь часть из которых имеет непредустановленные значения. (Для этого может подойти также паттерн «Строитель», описанный в рекомендации 7.)

PartialEq и Eq

Трейты `PartialEq` и `Eq` позволяют определять равенство в отношении пользовательских типов. Эти трейты имеют особую значимость, так как, если они присутствуют, компилятор автоматически использует их для проверки равенства (`==`) по аналогии с `operator==` в C++. Базовая реализация `derive` выполняет это путем рекурсивного сравнения всех полей.

Версия `Eq` — это просто расширение `PartialEq` в виде трейта-маркера, который добавляет допущение *рефлексивности*: любой тип `T`, который заявляет о поддержке `Eq`, должен обеспечить, чтобы `x == x` для любого `x: T`.

И это довольно странно, поэтому сразу же вызывает вопрос: «А если `x == x` окажется неверно?» Основная стоящая за этим разделением мысль относится к числам с плавающей запятой¹, а конкретно к особому значению `not a number` (`NaN`, *нечисло*, в Rust `f32::NAN` / `f64::NAN`). Спецификация значений с плавающей запятой требует, чтобы ничто не оказывалось равным `NaN`, *включая* само `NaN`. И трейт `PartialEq` является следствием этого.

Для пользовательских типов, у которых нет никаких особенностей, связанных с плавающей запятой, следует *реализовывать Eq при каждой реализации PartialEq*. Полноценный трейт `Eq` необходим и тогда, когда вы хотите использовать этот тип в качестве ключа в `HashMap`, а также в качестве трейта `Hash`.

Следует реализовывать `PartialEq` вручную, если ваш тип содержит поля, которые не влияют на отличительные особенности элемента, такие как внутренний кэш и прочие оптимизации производительности. (Любая ручная реализация будет использована и для `Eq`, если он определен, поскольку `Eq` — это просто трейт-маркер, не имеющий собственных методов.)

PartialOrd и Ord

Трейты упорядочения `PartialOrd` и `Ord` позволяют сравнивать два элемента типа, возвращая `Less`, `Greater` либо `Equal`. Эти трейты требуют реализаций равнозначных трейтов равенства (`PartialOrd` требует `PartialEq`, `Ord` требует `Eq`), причем они должны согласовываться друг с другом (следите за этим, особенно при ручной реализации).

Как и трейты равенства, трейты сравнения имеют особую значимость, поскольку компилятор автоматически использует их для операций сравнения (`<`, `>`, `<=`, `>=`).

Предустановленная реализация, создаваемая `derive`, лексикографически сравнивает поля (или варианты `enum`) в порядке их определения. Так что если это условие не выполняется, вам придется реализовывать трейты вручную или изменять их порядок.

В отличие от `PartialEq` трейт `PartialOrd` соответствует различным реальным ситуациям. Например, его можно использовать для того, чтобы показать связь между коллекцией и ее подмножеством²: `{1, 2}` — это подмножество `{1, 2, 4}`, но `{1, 3}` не подмножество `{2, 4}`, как и наоборот.

¹ Естественно, сравнение значений с плавающей запятой на предмет равенства всегда со-пряжено с риском, поскольку обычно нет гарантий, что округленные вычисления дадут результат, идентичный задуманному вами числу.

² В более общем смысле любая решетчатая структура имеет частичное упорядочение.

Тем не менее, даже если частичный порядок точно моделирует поведение вашего типа, *не следует реализовывать вместо Ord только PartialOrd* (редкий случай, который противоречит совету из рекомендации 2, — кодировать поведение в системе типов), поскольку это может привести к неожиданным результатам.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// Наследует поведение `PartialOrd` от `f32`
#[derive(PartialOrd, PartialEq)]
struct Oddity(f32);

// Входные данные со значениями NaN наверняка дадут
// неожиданные результаты
let x = Oddity(f32::NAN);
let y = Oddity(f32::NAN);

// Сравнение с самим собой вроде бы должно оказываться истинным,
// но так происходит не всегда
if x <= x {
    println!("This line doesn't get executed!");
}

// Программисты редко пишут код, охватывающий все возможные
// ветки сравнения. Если бы задействованные типы реализовывали `Ord`,
// то последние две ветви можно было бы совместить
if x <= y {
    println!("y is bigger"); // Условие не выполняется
} else if y < x {
    println!("x is bigger"); // Условие не выполняется
} else {
    println!("Neither is bigger");
}
```

Hash

Трейт `Hash` используется для создания одного значения, которое с высокой вероятностью будет разным для разных элементов. Это значение служит основой для структур данных на базе хеш-блока вроде `HashMap` и `HashSet`. Поэтому тип ключей в таких структурах должен реализовывать `Hash` и `Eq`.

Если все это развернуть, то необходимо, чтобы одинаковые элементы (согласно `Eq`) всегда давали одинаковый хеш: если `x == y` (посредством `Eq`), тогда всегда должно быть `hash(x) == hash(y)`. *Если у вас есть ручная реализация Eq, посмотрите, нужно ли вам также вручную реализовать Hash, чтобы это требование выполнялось.*

Debug и Display

Трейты `Debug` и `Display` позволяют типу указывать, как его нужно включать в вывод для рутинных нужд (аргумент формата `{}`) или отладки (аргумент формата `:?{}`), что подобно перегрузке `operator<<` для `iostream` в C++.

Однако различаются эти трейты не только требуемым спецификатором формата.

- `Debug` можно выводить автоматически, а `Display` реализуется только вручную.
- Схема вывода `Debug` для разных версий Rust может быть различной. Если вывод в дальнейшем будет парситься другим кодом, используйте `Display`.
- `Debug` предназначен для программистов, а `Display` — для пользователей. Чтобы лучше понять, представьте, что программа локализована на незнакомом авторам языке, — если ее содержимое можно перевести, то подойдет `Display`, если нет, то `Debug`.

В качестве общего правила: если ваши типы не содержат чувствительной информации (личные данные, криптографический материал и т. д.), *добавляйте для них автоматически генерируемую реализацию Debug*. Чтобы было проще следовать этому совету, в компиляторе Rust есть линт `missing_debug_implementations`, который указывает на типы без `Debug`. Этот линт по умолчанию отключен, но его можно включить с помощью одной из следующих инструкций:

```
#![warn(missing_debug_implementations)]  
#![deny(missing_debug_implementations)]
```

Если автоматически генерируемая реализация `Debug` будет выдавать большой объем деталей, тогда лучше включить ее ручную версию, обобщающую содержимое типа.

Реализуйте `Display`, если ваши типы созданы с учетом вывода для конечного пользователя в текстовой форме.

Стандартные трейты из других рекомендаций

Помимо типичных трейтов, описанных в предыдущем разделе, стандартная библиотека включает и менее распространенные. Далее перечислены самые важные из них, но они рассматриваются в других рекомендациях, поэтому здесь приводится лишь их краткое описание.

Fn, FnOnce и FnMut

Элементы, реализующие эти трейты, представляют замыкания, которые можно вызывать. Подробнее — в рекомендации 2.

Error

Элементы, реализующие этот трейт, представляют информацию об ошибках, которую можно выводить для пользователей или программистов и которая может содержать данные о вложенных ошибках. Подробнее — в рекомендации 4.

Drop

Элементы, реализующие этот трейт, выполняют обработку при удалении. Это важное поведение для паттернов RAII. Подробнее — в рекомендации 11.

From и TryFrom

Элементы, реализующие эти трейты, можно автоматически создавать из элементов другого типа, но уже с возможностью сбоя. Подробнее — в рекомендации 5.

Deref и DerefMut

Элементы, реализующие эти трейты, представляют подобные указателям объекты, которые можно разыменовывать для получения доступа к внутреннему элементу. Подробнее — в рекомендации 8.

Iterator и друзья

Элементы, реализующие эти трейты, представляют коллекции, которые можно перебирать. Подробнее — в рекомендации 9.

Send

Элементы, реализующие этот трейт, можно безопасно передавать между несколькими потоками. Подробнее — в рекомендации 17.

Sync

На элементы, реализующие этот трейт, можно безопасно ссылаться из множества потоков. Подробнее — в рекомендации 17.

Ни один из этих трейтов нельзя выводить (`derive`).

Перегрузки операторов

Последняя категория стандартных трейтов относится к перегрузкам операторов, где Rust позволяет перегружать различные встроенные унарные и бинарные операторы для пользовательских типов путем реализации стандартных трейтов из модуля `std::ops`. Эти трейты выводить нельзя, и обычно они нужны только для типов, представляющих алгебраические объекты, в которых присутствует естественная интерпретация этих операторов.

Тем не менее опыт C++ показал, что лучше избегать перегрузки операторов для несвязанных типов, поскольку это часто дает код, который сложно обслуживать,

и неожиданно влияет на производительность (например, `x + y` молча вызывает дорогостоящий метод $O(N)$).

Чтобы соблюсти принцип наименьшего удивления¹ при реализации каких-либо перегрузок операторов, вам следует *реализовывать их связный набор*. Например, если `x + y` содержит перегрузку (`Add`) и `-y` (`Neg`) — тоже, нужно дополнительно реализовать `x - y` (`Sub`), проследив, чтобы это выражение давало тот же ответ, что и `x + (-y)`.

Элементы, передаваемые в трейты перегрузки операторов, перемещаются, то есть по умолчанию будут потреблять типы без копирования (`Copy`). Добавление реализаций для & 'а `MyType` поможет с этим, но потребует больше шаблонного кода, чтобы предусмотреть все варианты (например, для двоичного оператора есть $2 \times 2 = 4$ варианта совмещения ссылочных/не ссылочных аргументов).

Обобщение

В этой рекомендации мы поговорили о многом, поэтому далее приводится несколько таблиц, обобщающих рассмотренные стандартные трейты. Первая из них, табл. 2.1, включает трейты, которые были разобраны детально и которые можно автоматически выводить (`derive`), *за исключением `Display`*.

Таблица 2.1. Типичные стандартные трейты

Трейт	Использование компилятором	Граница	Метод
Clone			<code>clone</code>
Copy	<code>let y = x;</code>	Clone	Трейт-маркер
Default			<code>default</code>
PartialEq	<code>x == y</code>		<code>eq</code>
Eq	<code>x == y</code>	PartialEq	Трейт-маркер
PartialOrd	<code>x < y, x <= y, ...</code>	PartialEq	<code>partial_cmp</code>
Ord	<code>x < y, x <= y, ...</code>	Eq + PartialOrd	<code>cmp</code>
Hash			<code>hash</code>
Debug	<code>format!("{}:?", x)</code>		<code>fmt</code>
Display	<code>format!("{}", x)</code>		<code>fmt</code>

¹ Принцип (правило) наименьшего удивления (англ. principle of least astonishment) в эргономике гласит, что если назначение элемента или сочетания неясно, то его поведение должно быть наиболее ожидаемым со стороны пользователя. — *Примеч. ред.*

В табл. 2.2 приводятся трейты перегрузок операторов. Ни один из них нельзя вывести (`derive`).

Таблица 2.2. Трейты перегрузки операторов¹

Трейт	Использование компилятором	Граница	Метод
Add	<code>x + y</code>		<code>add</code>
AddAssign	<code>x += y</code>		<code>add_assign</code>
BitAnd	<code>x & y</code>		<code>bitand</code>
BitAndAssign	<code>x &= y</code>		<code>bitand_assign</code>
BitOr	<code>x y</code>		<code>bitor</code>
BitOrAssign	<code>x = y</code>		<code>bitor_assign</code>
BitXor	<code>x ^ y</code>		<code>bitxor</code>
BitXorAssign	<code>x ^= y</code>		<code>bitxor_assign</code>
Div	<code>x / y</code>		<code>div</code>
DivAssign	<code>x /= y</code>		<code>div_assign</code>
Mul	<code>x * y</code>		<code>mul</code>
MulAssign	<code>x *= y</code>		<code>mul_assign</code>
Neg	<code>-x</code>		<code>neg</code>
Not	<code>!x</code>		<code>not</code>
Rem	<code>x % y</code>		<code>rem</code>
RemAssign	<code>x %= y</code>		<code>rem_assign</code>
Shl	<code>x << y</code>		<code>shl</code>
ShlAssign	<code>x <<= y</code>		<code>shl_assign</code>
Shr	<code>x >> y</code>		<code>shr</code>
ShrAssign	<code>x >>= y</code>		<code>shr_assign</code>
Sub	<code>x - y</code>		<code>sub</code>
SubAssign	<code>x -= y</code>		<code>sub_assign</code>

Для полноты картины в табл. 2.3 приведены стандартные трейты, разбираемые в других рекомендациях. Ни один из этих трейтов нельзя вывести, но `Send` и `Sync` могут автоматически реализовываться компилятором.

¹ Некоторые из указанных здесь имен могут быть непонятны, например `Rem` для остатка (`remainder`) и `Shl` для сдвига влево (`shift left`), но документация `std::ops` проясняет их назначение.

Таблица 2.3. Стандартные трейты, описываемые в других рекомендациях

Трейт	Использование компилятором	Граница	Метод	Рекомендация
Fn	x(a)	FnMut	call	2
FnMut	x(a)	FnOnce	call_mut	2
FnOnce	x(a)		call_once	2
Error		Display + Debug	[source]	4
From			from	5
TryFrom			try_from	5
Into			into	5
TryInto			try_into	5
AsRef			as_ref	8
AsMut			as_mut	8
Borrow			borrow	8
BorrowMut		Borrow	borrow_mut	8
ToOwned			to_owned	8
Deref	*x, &x		deref	8
DerefMut	*x, &mut x	Deref	deref_mut	8
Index	x[idx]		index	8
IndexMut	x[idx] =...	Index	index_mut	8
Pointer	format("{:p}", x)		fmt	8
Iterator			next	9
IntoIterator	for y in x		into_iter	9
FromIterator			from_iter	9
ExactSizeIterator		Iterator	(size_hint)	9
DoubleEndedIterator		Iterator	next_back	9
Drop	} (конец области видимости)		drop	11
Sized			Трейт-маркер	12
Send	Передача между потоками		Трейт-маркер	17
Sync	Использование несколькими потоками		Трейт-маркер	17

Рекомендация 11. Реализуйте трейт Drop для паттернов RAII

Никогда не поручайте человеку работу машины.

Агент Смит

RAII (Resource Acquisition Is Initialization — получение ресурса есть инициализация) представляет собой идиому программирования, в которой время жизни значения точно привязывается к времени жизни некоего дополнительного ресурса. Идиома RAII была популяризована языком C++ и стала одним из самых значимых его вкладов в сферу программирования в целом.

Корреляция между временем жизни значения и временем жизни ресурса кодируется в типе RAII:

- *конструктор* типа получает доступ к некоторому ресурсу;
- *деструктор* типа освобождает доступ к этому ресурсу.

В результате этого тип RAII имеет *инвариант*: доступ ко внутреннему ресурсу есть, только если этот элемент существует. Поскольку компилятор следит за тем, чтобы при выходе из области локальные переменные уничтожались, то соответствующие ресурсы при выходе из области тоже освобождаются¹.

Это, в частности, хорошо для *обслуживаемости*: если последующее изменение кода изменяет поток управления, время жизни элемента и ресурса все равно остается верным. Чтобы это наблюдать, разберем код, который вручную блокирует и разблокирует мьютекс, не используя паттерн RAII. Этот код написан на C++, поскольку Mutex в Rust запрещает такое уязвимое для ошибок поведение:

```
// Код C++
class ThreadSafeInt {
public:
    ThreadSafeInt(int v) : value_(v) {}

    void add(int delta) {
        mu_.lock();
        // ... еще код
        value_ += delta;
        // ... еще код
        mu_.unlock();
    }
}
```

¹ Это означает, что RAII в качестве техники есть только в языках, где прогнозируется время уничтожения, что исключает большинство языков с функцией сборки мусора (хотя инструкция defer в Go (<https://oreil.ly/2jrtX>) частично дает те же результаты).

Модификация для перехвата состояния ошибки с ранним выходом оставляет мьютекс заблокированным.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// Код C++
void add_with_modification(int delta) {
    mu_.lock();
    // ... еще код
    value_ += delta;
    // Проверка переполнения
    if (value_ > MAX_INT) {
        // Упс, забыли выполнить unlock() перед выходом
        return;
    }
    // ... еще код
    mu_.unlock();
}
```

Тем не менее инкапсуляция блокирующего поведения в классе RAII:

```
// Код C++ (реальный код должен использовать std::lock_guard
// или аналогичный)
class MutexLock {
public:
    MutexLock(Mutex* mu) : mu_(mu) { mu_->lock(); }
    ~MutexLock() { mu_->unlock(); }
private:
    Mutex* mu_;
};
```

означает, что в равнозначном коде это изменение безопасно:

```
// Код C++
void add_with_modification(int delta) {
    MutexLock with_lock(&mu_);
    // ... еще код
    value_ += delta;
    // Проверка на переполнение
    if (value_ > MAX_INT) {
        return; // Безопасно, так как with_lock разблокирует доступ
    }
    // ... еще код.
}
```

В C++ паттерны RAII изначально часто использовались для управления памятью, позволяя обеспечить сохранение синхронности между ручными

операциями аллокации (`(new, malloc())`) и dealлокации¹ (`(delete, free())`). Общая версия такого управления памятью была добавлена в стандартную библиотеку C++11: тип `std::unique_ptr<T>` обеспечивает, чтобы памятью владела только одна область кода, но позволяет заимствовать указатель на эту память для временного использования (`ptr.get()`).

В Rust такое поведение указателей памяти встроено в сам язык (см. рекомендацию 15), но общий принцип RAII по-прежнему полезен для прочих видов ресурсов². *Реализуйте Drop для любых типов, содержащих ресурсы, которые должны быть освобождены*, например следующих.

- Доступ к ресурсам операционной системы. В системах на базе Unix это обычно означает что-то, где есть *дескриптор файла*. Если их не освобождать, системные ресурсы будут удерживаться, и это приведет к тому, что программа упрется в лимит дескрипторов файлов на процесс.
- Доступ к ресурсам синхронизации. Стандартная библиотека уже включает примитивы для синхронизации памяти, но другие ресурсы, например блокировки файлов, блокировки баз данных и т. д., могут потребовать аналогичной инкапсуляции.
- Доступ к неструктурированной памяти (`raw memory`) для `unsafe`-типов, которые связаны с управлением низкоуровневой памятью, например функциональность интерфейса внешней функции (`foreign function interface, FFI`).

В стандартной библиотеке Rust самым очевидным примером RAII является элемент `MutexGuard`, возвращаемый операцией `Mutex::lock()`. Эти операции активно используются в программах, задействующих параллельный доступ к общему состоянию, о чем мы будем говорить в рекомендации 17. Это подобно показанному ранее заключительному примеру C++, но в Rust `MutexGuard`, помимо того что выступает элементом RAII для удерживаемой блокировки, действует и как посредник для защищенных мьютексом данных:

```
use std::sync::Mutex;

struct ThreadSafeInt {
    value: Mutex<i32>,
}

impl ThreadSafeInt {
    fn new(val: i32) -> Self {
```

¹ Аллокация (allocation) — выделение памяти, dealлокация (deallocation) — освобождение памяти. — Примеч. пер.

² RAII по-прежнему полезен для управления памятью в низкоуровневом unsafe-коде, но эта тема выходит за рамки данной книги.

```

        Self {
            value: Mutex::new(val),
        }
    }

    fn add(&self, delta: i32) {
        let mut v = self.value.lock().unwrap();
        *v += delta;
    }
}

```

Рекомендация 17 советует не использовать удержание блокировок для больших участков кода. Чтобы обеспечить это, *используйте блоки для ограничения зоны действия элементов RAII*. Из-за этого появляются немного странные отступы, но дополнительная безопасность и точность времени жизни того стоят:

```

impl ThreadSafeInt {
    fn add_with_extras(&self, delta: i32) {
        // ... дополнительный код, не нуждающийся в блокировке
        {
            let mut v = self.value.lock().unwrap();
            *v += delta;
        }
        // ... дополнительный код, не нуждающийся в блокировке
    }
}

```

Рекомендуя применение RAII, следует пояснить, как конкретно этот принцип реализовывать. Трейт `Drop` позволяет вам добавлять к уничтожению элемента пользовательское поведение. В этом трейте есть один метод, `drop`, который компилятор выполняет непосредственно перед освобождением памяти, удерживающей блокировку:

```

#[derive(Debug)]
struct MyStruct(i32);

impl Drop for MyStruct {
    fn drop(&mut self) {
        println!("Dropping {self:?}");
        // Код для освобождения ресурсов, принадлежащих элементу
    }
}

```

Метод `drop` специально зарезервирован для компилятора и вручную не вызывается:

НЕ КОМПИЛИРУЕТСЯ

```
x.drop();
```

```
error[E0040]: explicit use of destructor method
--> src/main.rs:70:7
70 |     x.drop();
|     _-^_____
|     |
|     | explicit destructor calls not allowed
| help: consider using `drop` function: `drop(x)`
```

Здесь стоит чуть подробнее разобрать технические детали. Обратите внимание на то, что метод `Drop::drop` имеет сигнатуру `drop(&mut self)`, а не `drop(self)`: он получает мутабельную ссылку на элемент, а не сам этот элемент. Если бы `Drop::drop` действовал как стандартный метод, это бы означало, что элемент остается доступным для использования даже после очистки его внутреннего состояния и освобождения ресурсов.

НЕ КОМПИЛИРУЕТСЯ

```
{
    // Если бы вызов `drop` был разрешен...
    x.drop(); // (не компилируется)

    // `x` останется доступен после
    x.0 += 1;
}
// Кроме того, что произойдет, когда `x` выйдет из рабочей области?
```

Компилятор предположил простую альтернативу, а именно вызов функции `drop()` для ручного отбрасывания элемента. Эта функция получает аргумент перемещения, и реализация `drop(_item: T)` оказывается просто пустым телом `{ }`, поэтому при достижении закрывающей скобки перемещаемый элемент отбрасывается.

Заметьте также, что у сигнатуры метода `drop(&mut self)` отсутствует возвращаемый тип, то есть у него нет способа сообщить о сбое. Если попытка освободить ресурсы может провалиться, то вам наверняка следует задействовать отдельный метод `release`, возвращающий `Result`, чтобы пользователи могли этот сбой обнаружить.

Независимо от технических деталей метод `drop` все равно является ключевым местом для реализации RAII. Его код — идеальная точка для освобождения связанных с элементом ресурсов.

Рекомендация 12. Уясните компромиссы между обобщениями и трейт-объектами

В рекомендации 2 использование трейтов для инкапсуляции поведения в системе типов описывалось в виде коллекции связанных методов. При этом были обозначены два способа применения трейтов: в качестве *границ трейтов* для *обобщений* или в *трейт-объектах*. В текущей же рекомендации речь пойдет о компромиссах между этими двумя вариантами.

В качестве рабочего примера рассмотрим трейт, предоставляющий функциональность для вывода графических объектов:

```
#[derive(Debug, Copy, Clone)]
pub struct Point {
    x: i64,
    y: i64,
}

#[derive(Debug, Copy, Clone)]
pub struct Bounds {
    top_left: Point,
    bottom_right: Point,
}

/// Вычисляет пересечение двух прямоугольников или `None`,
/// если таковое отсутствует
fn overlap(a: Bounds, b: Bounds) -> Option<Bounds> {
    // ...
}

/// Трейт для объектов, которые можно отобразить графически
pub trait Draw {
    /// Возвращает прямоугольник, обрамляющий объект
    fn bounds(&self) -> Bounds;

    // ...
}
```

Обобщения

Обобщения в Rust почти равноценны шаблонам в C++: они позволяют программисту писать код, работающий для некоего произвольного типа T, а конкретное использование этого кода определяется уже при компиляции. В Rust этот

процесс называется *мономорфизацией*, а в C++ – *инстанцированием шаблонов*. В отличие от C++ Rust явно кодирует ожидания от типа T в системе типов в виде границ трейта для обобщения.

Например, обобщенная функция, использующая метод bounds() трейта, имеет явную границу трейта Draw:

```
// Указывает, находится ли объект на экране
pub fn on_screen<T>(draw: &T) -> bool
where
    T: Draw,
{
    overlap(SCREEN_BOUNDS, draw.bounds()).is_some()
}
```

Этот код можно написать и более сжато, установив границу трейта после обобщенного параметра:

```
pub fn on_screen<T: Draw>(draw: &T) -> bool {
    overlap(SCREEN_BOUNDS, draw.bounds()).is_some()
}
```

или используя в качестве типа аргумента `impl Trait`¹:

```
pub fn on_screen(draw: &impl Draw) -> bool {
    overlap(SCREEN_BOUNDS, draw.bounds()).is_some()
}
```

Если тип реализует трейт:

```
#[derive(Clone)] // `Debug` отсутствует
struct Square {
    top_left: Point,
    size: i64,
}

impl Draw for Square {
    fn bounds(&self) -> Bounds {
        Bounds {
            top_left: self.top_left,
            bottom_right: Point {
                x: self.top_left.x + self.size,
                y: self.top_left.y + self.size,
            },
        }
    }
}
```

¹ Использование `impl Trait` в позиции аргумента (https://oreil.ly/k_zrF) не абсолютно равнозначно двум предыдущим версиям, поскольку лишает вызывающий код возможности явно указывать параметр типа в виде чего-то вроде `on_screen::<Circle>(&c)`.

тогда экземпляры этого типа можно передавать в обобщенную функцию, мономорфизируя ее для создания кода, специфичного для одного конкретного типа:

```
let square = Square {
    top_left: Point { x: 1, y: 2 },
    size: 2,
};
// Вызывает `on_screen::<Square>(&square) -> bool`
let visible = on_screen(&square);
```

Если та же обобщенная функция используется с другим типом, который реализует соответствующую границу трейта:

```
#[derive(Clone, Debug)]
struct Circle {
    center: Point,
    radius: i64,
}

impl Draw for Circle {
    fn bounds(&self) -> Bounds {
        // ...
    }
}
```

то задействуется другой мономорфизированный код:

```
let circle = Circle {
    center: Point { x: 3, y: 4 },
    radius: 1,
};
// Вызывает `on_screen::<Circle>(&circle) -> bool`
let visible = on_screen(&circle);
```

Иными словами, программист пишет одну обобщенную функцию, а компилятор выводит ее разные мономорфизированные версии для каждого отдельного типа, с которым эта функция вызывается.

Трейт-объекты

В сравнении с этим трейт-объекты выступают жирными указателями (см. рекомендацию 8), совмещающими указатель на конкретный элемент с указателем на таблицу vtable, которая содержит указатели функции для всех методов в реализации трейта (рис. 2.1):

```
let square = Square {
    top_left: Point { x: 1, y: 2 },
    size: 2,
};
let draw: &dyn Draw = &square;
```

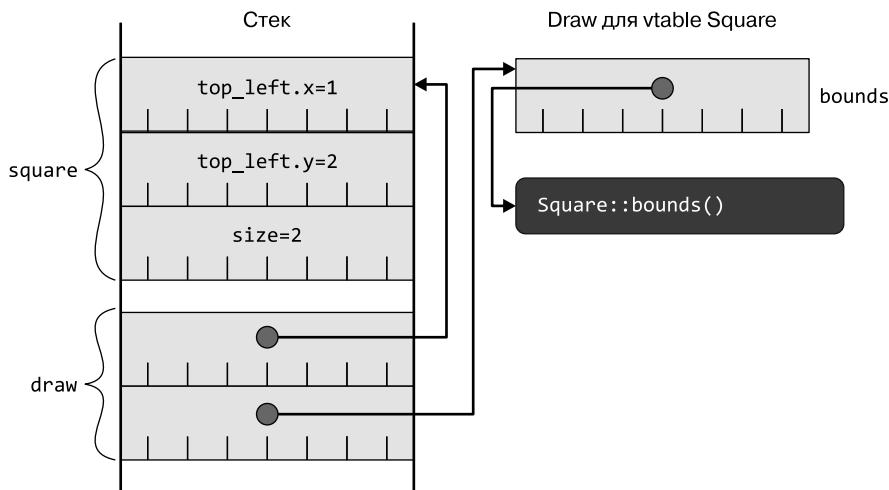


Рис. 2.1. Схема трейт-объекта с указателями на конкретный элемент и vtable

Это означает, что функции, которая получает трейт-объект, не обязательно быть обобщенной и использовать мономорфизацию. Программист пишет функцию, задействуя трейт-объекты, и компилятор выводит только одну ее версию, способную получать трейт-объекты, поступающие от нескольких входных типов:

```
/// Указывает, находится ли объект на экране
pub fn on_screen(draw: &dyn Draw) -> bool {
    overlap(SCREEN_BOUNDS, draw.bounds()).is_some()
}

// Вызывает `on_screen(&dyn Draw) -> bool`
let visible = on_screen(&square);
// Тоже вызывает `on_screen(&dyn Draw) -> bool`
let visible = on_screen(&circle);
```

Базовые сравнения

Эти основные факты уже позволяют с ходу сравнить два описанных варианта.

- Обобщения часто вызывают появление более объемного кода, так как компилятор генерирует его свежую копию (`(on_screen::<T>(&T))`) для каждого типа `T`, использующего обобщенную версию функции `on_screen`. В противоположность этому версии функции с трейт-объектами (`(on_screen(&dyn T))`) достаточно одного экземпляра.
- Вызов метода трейта из обобщения обычно будет лишь ненамного быстрее его вызова из кода, использующего трейт-объект, так как последнему для

нахождения кода нужно выполнить два разыменовывания — трейт-объекта в таблицу vtable и vtable в место реализации.

- Обобщения дольше компилируются, так как компилятор создает больше кода и линковщик проделывает больше работы для сворачивания повторов.

В большинстве ситуаций все эти различия не очень значительны — опираться в решениях на аспекты оптимизации следует, только если вы измерили влияние и выяснили, что оно существенно (обнаружили узкое место производительности или проблемы из-за увеличения загруженности).

Более значительным различием является то, что границы трейтов в обобщениях можно использовать, чтобы делать доступной разную функциональность условия, опираясь на реализацию параметром типа *нескольких* трейтов:

```
// Функция `area` доступна для всех контейнеров,
// содержащих строки, реализующие `Draw`
fn area<T>(draw: &T) -> i64
where
    T: Draw,
{
    let bounds = draw.bounds();
    (bounds.bottom_right.x - bounds.top_left.x)
        * (bounds.bottom_right.y - bounds.top_left.y)
}

// Метод `show` доступен, только если реализован `Debug`
fn show<T>(draw: &T)
where
    T: Debug + Draw,
{
    println!("{} has bounds {}", draw, draw.bounds());
}

let square = Square {
    top_left: Point { x: 1, y: 2 },
    size: 2,
};
let circle = Circle {
    center: Point { x: 3, y: 4 },
    radius: 1,
};

// И `Square`, и `Circle` реализуют `Draw`
println!("area(square) = {}", area(&square));
println!("area(circle) = {}", area(&circle));

// `Circle` реализует `Debug`
show(&circle);

// `Square` не реализует `Debug`, значит, этот код не скомпилируется:
// show(&square);
```

Трейт-объект кодирует vtable реализации только для одного трейта, поэтому проделывать что-то равнозначное будет намного неудобнее. Например, комбинацию трейтов `Debug` и `Draw` можно определить для кейса `show()` вместе с `blanket`-реализацией, чтобы все упростить:

```
trait DebugDraw: Debug + Draw {}

/// Blanket-реализация применяется, когда реализуются отдельные трейты
impl<T: Debug + Draw> DebugDraw for T {}}
```

Тем не менее, если есть несколько комбинаций отдельных трейтов, понятно, что комбинаторика такого подхода быстро станет проблематичной.

Дополнительные границы трейтов

В добавок к использованию границ трейтов для ограничения того, какие параметры типов будут приемлемы для обобщенной функции, их можно применить и к самим определениям трейтов:

```
/// Все, что реализует `Shape`, должно реализовывать и `Draw`
trait Shape: Draw {
    /// Отрисовывает ту часть фигуры, которая попадает в `bounds`
    fn render_in(&self, bounds: Bounds);

    /// Отрисовывает фигуру
    fn render(&self) {
        // Предустановленная реализация отрисовывает ту часть фигуры,
        // которая попадает в область экрана
        if let Some(visible) = overlap(SCREEN_BOUNDS, self.bounds()) {
            self.render_in(visible);
        }
    }
}
```

В этом примере предустановленная реализация метода `render()` (см. рекомендацию 13) использует границу трейта, опираясь на доступность метода `bounds()` из `Draw`.

Программисты, пришедшие из объектно-ориентированных языков, часто путают границы трейтов с наследованием. В данном случае это бы выражалось в ощущении того, что `Shape` — это `Draw`. Но в действительности это не так — связь между этими типами правильнее обозначить как «`Shape` также реализует `Draw`».

У трейт-объектов для трейтов, имеющих границы:

```
let square = Square {
    top_left: Point { x: 1, y: 2 },
```

```

    size: 2,
};

let draw: &dyn Draw = &square;
let shape: &dyn Shape = &square;

```

есть одна совмещенная vtable, которая включает методы верхнего трейта и методы всех его границ. Схематично это показано на рис. 2.2: vtable для Shape включает метод bounds из трейта Draw, а также два метода из самого трейта Shape.

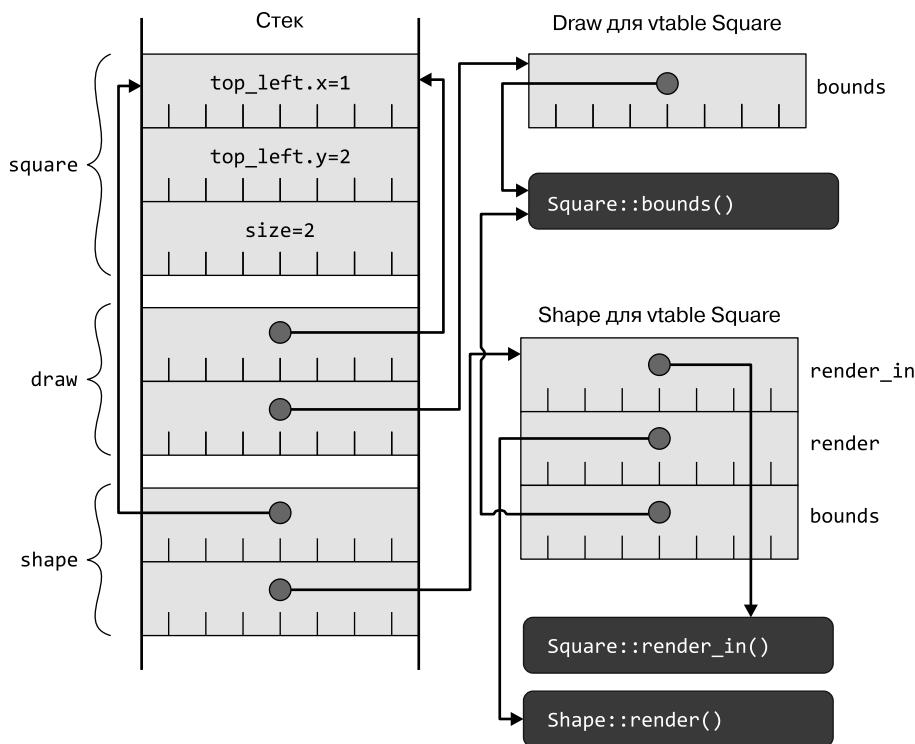


Рис. 2.2. Трейт-объекты для границ трейтов с отдельными vtables для Draw и Shape

На момент написания книги (и для версии Rust 1.70) это означает, что невозможно выполнить приведение из `Shape` в `Draw`, поскольку чистую vtable для `Draw` в среде выполнения восстановить нельзя. Выполнять преобразование между связанными трейт-объектами тоже нельзя, и это означает невозможность использования принципа подстановки Лисков. Однако в последующих версиях Rust это может измениться (подробнее — в рекомендации 19).

Если тот же смысл выразить иными словами, то метод, который получает трейт-объект `Shape`, имеет следующие характеристики.

- Он *может* использовать методы из `Draw`, поскольку `Shape` также реализует `Draw` и в таблице vtable для `Square` есть соответствующие указатели функций.
- Он *не может* (пока) передавать трейт-объект в другой метод, ожидающий трейт-объект `Draw`, потому что `Shape` — это не `Draw` и vtable для `Draw` недоступна.

В противоположность этому обобщенный метод, который получает элементы, реализующие `Shape`, характеризуется так:

- он *может* применять методы из `Draw`;
- он *может* передавать элемент в другой обобщенный метод, который имеет границу трейта `Draw`, поскольку эта граница мономорфизуется на этапе компиляции для использования методов `Draw` конкретного типа.

Безопасность трейт-объектов

Еще одним ограничением трейт-объектов выступает требование *объектной безопасности* (<https://oreil.ly/gaq4l>), согласно которому в качестве трейт-объектов можно использовать только трейты, соответствующие следующим двум правилам:

- методы трейта не должны быть обобщенными;
- методы трейта не должны включать тип, который содержит `Self`, за исключением получателя (объекта, в котором метод вызывается).

Первое ограничение вполне понятно: обобщенный метод `f` — это, по сути, бесконечный набор методов, потенциально включающий `f::<i16>`, `f::<i32>`, `f::<i64>`, `f::<u8>` и т. д. Vtable трейт-объекта, напротив, представляет ограниченную коллекцию указателей функций, а значит, в нее невозможно вписать бесконечный набор мономорфизированных реализаций.

Второе ограничение не столь явное, но на практике встречается чаще. Трейты, которые устанавливают границы `Copy` или `Clone` (см. рекомендацию 10), автоматически подпадают под это правило, поскольку возвращают `Self`. Чтобы понять, почему это запрещено, рассмотрите код, содержащий трейт-объект. Что произойдет, если этот код вызовет, предположим, `let y = x.clone()`? Вызывающему коду нужно зарезервировать в стеке достаточно места для `y`, но он не владеет информацией о размере `y`, так как `Self` имеет произвольный тип. В итоге возвращаемые типы, которые содержат `Self`, ведут к получению небезопасного в объектном смысле трейта¹.

¹ На момент написания книги ограничение методов, которые возвращают `Self`, включает типы наподобие `Box<Self>`, которые можно безопасно сохранить в стеке. В будущем это ограничение вполне может быть ослаблено (<https://oreil.ly/JZH3V>).

Но для второго ограничения есть исключение. Метод, возвращающий связанный с `Self` тип, не влияет на объектную безопасность, если `Self` явно ограничен типами, размер которых известен на этапе компиляции. Это обозначается трейтом-маркером `Sized` в виде границы трейта:

```
/// `Stamp` можно скопировать и отрисовать несколько раз
trait Stamp: Draw {
    fn make_copy(&self) -> Self
    where
        Self: Sized;
}

let square = Square {
    top_left: Point { x: 1, y: 2 },
    size: 2,
};

// `Square` реализует `Stamp`, а значит, он может вызывать `make_copy()`
let copy = square.make_copy();

// Так как возвращающий `Self` метод содержит границу трейта `Sized`,
// создать трейт-объект `Stamp` возможно
let stamp: &dyn Stamp = &square;
```

Эта граница трейта означает, что метод никак нельзя использовать с трейт-объектами, поскольку они связаны с чем-то неизвестного размера (`dyn Trait`), а значит, он объектную безопасность не обеспечивает.

НЕ КОМПИЛИРУЕТСЯ

```
// Тем не менее этот метод нельзя вызвать через трейт-объект
let copy = stamp.make_copy();

error: the `make_copy` method cannot be invoked on a trait object
--> src/main.rs:397:22
   |
353 |     Self: Sized;
   |             ----- this has a `Sized` requirement
...
397 |     let copy = stamp.make_copy();
   |             ^^^^^^^^^^
```

Компромиссы

Баланс указанных факторов предполагает, что *лучше использовать обобщения, а не трейт-объекты*. Но бывают случаи, когда трейт-объекты оказываются подходящим инструментом для решения стоящей задачи.

Первый случай касается практических соображений: если размер генерированного кода на этапе компиляции очень важен, то трейт-объекты окажутся более подходящими (о чём говорилось ранее).

Более теоретический аспект, который ведёт к использованию трейт-объектов, заключается в том, что они неизбежно подразумевают *стирание типа*: информация о конкретном типе в процессе преобразования в трейт-объект утрачивается. И это может обернуться как недостатком (подробнее — в рекомендации 19), так и плюзой, поскольку позволяет задействовать коллекции из разнородных объектов. Объясняется это тем, что код опирается *лишь* на методы трейта — он может вызывать и комбинировать элементы, имеющие разные конкретные типы.

Одним из традиционных объектно-ориентированных примеров сказанного будет отрисовка списка фигур — один и тот же метод `render()` можно использовать для создания квадратов, кругов, эллипсов и звезд в одном цикле:

```
let shapes: Vec<&dyn Shape> = vec![&square, &circle];
for shape in shapes {
    shape.render()
}
```

Намного менее явную пользу трейт-объекты приносят в случаях, когда доступные типы на этапе компиляции неизвестны. Если новый код динамически загружается в среде выполнения, например, посредством `dlopen(3)`, тогда элементы, которые реализуют трейты в новом коде, можно вызывать только через трейт-объект, так как исходный код для мономорфизации отсутствует.

Рекомендация 13. Минимизируйте число необходимых методов трейтов с помощью предустановленных реализаций

Потенциальная аудитория разработчика трейта состоит из программистов, которые будут этот трейт *реализовывать*, и тех, кто станет его использовать. Такая двойственность аудитории вызывает некоторое напряжение при проектировании трейта.

- Для облегчения жизни реализующих желательно, чтобы трейт имел минимум методов для достижения своей цели.
- Для удобства пользователей стоит предоставлять различные варианты методов, которые охватят все типичные способы применения трейта.

Найти баланс между этими двумя приоритетами можно так: добавить множество методов, которые упростят жизнь пользователям, но сопроводить *предустановленными реализациями* те из них, которые могут создаваться из других, более примитивных операций над интерфейсом.

Простым примером будет метод `is_empty()` для `ExactSizeIterator`, представляющего `Iterator`, который знает число перебираемых элементов¹. Этот метод имеет дефолтную реализацию, которая опирается на метод `len()` трейта:

```
fn is_empty(&self) -> bool {
    self.len() == 0
}
```

Но присутствие дефолтной реализации не является ограничением. Если у реализации трейта есть другой способ для определения того, является ли итератор пустым, она может заменить предустановленную `is_empty()` своим вариантом. При использовании этого подхода получаются трейты, имеющие меньше *необходимых методов*, но гораздо больше предустановленных. В итоге при реализации необходимо прописывать только первые, а последние присутствуют по умолчанию.

Этот подход активно реализует стандартная библиотека Rust. Возможно, лучшим примером здесь будет трейт `Iterator`, который содержит один необходимый метод `next()`, но включает множество предопределенных (см. рекомендацию 9) — на момент написания книги более 50.

Методы трейтов могут устанавливать *границы трейтов*, указывая, что метод доступен, только если задействованные типы реализуют конкретные трейты. Трейт `Iterator` также показывает, что это полезно в комбинации с дефолтными реализациями методов. Например, метод перебора `cloned()` содержит границу трейта и предустановленную реализацию:

```
fn cloned<'a, T>(self) -> Cloned<Self>
where
    T: 'a + Clone,
    Self: Sized + Iterator<Item = &'a T>,
{
    Cloned::new(self)
}
```

Иными словами, метод `cloned()` доступен, только если тип внутреннего `Item` реализует `Clone` — при таком условии реализация автоматически оказывается доступной.

¹ Метод `is_empty()` на данный момент является экспериментальной функцией в nightly-версии (<https://oreil.ly/0wZOL>).

Последней особенностью методов трейтов с дефолтными реализациями является то, что новые методы *обычно* можно безопасно добавлять в трейт даже после релиза его первой версии. При этом сохраняется обратная совместимость (подробнее — в рекомендации 21) для пользователей и разработчиков, реализующих трейт, при условии, что имя нового метода не конфликтует с именем метода из другого реализуемого этим типом трейта¹.

Поэтому следуйте примеру стандартной библиотеки и предоставляйте минимальную поверхность API для разработчиков, но удобный и обширный API — для пользователей, добавляя методы с предустановленными реализациями и подобающими границами трейтов.

¹ Если имя нового метода вдруг совпадет с именем другого метода в одном конкретном типе, тогда этот метод, называемый изначальной реализацией, будет использоваться раньше метода трейта. При этом метод трейта можно выбрать явно путем приведения: <Concrete as Trait>::method().

ГЛАВА 3

Концепции

Первые две главы книги были посвящены типам и трейтам в Rust, что позволило изучить терминологию, необходимую для работы с некоторыми *концепциями* написания кода на этом языке, о чем будем говорить в текущей главе.

Анализатор заимствований вместе с проверками времени жизни — это основа, которая делает Rust уникальным. Они же становятся типичным камнем преткновения для новичков, в связи с чем станут предметом обсуждения первых двух рекомендаций этой главы.

В остальных рекомендациях речь пойдет о принципах, которые проще понять, но которые несколько отличаются от принципов написания кода на других языках. К ним относятся:

- советы относительно режима `unsafe` в Rust и того, как его избежать (см. рекомендацию 16);
- хорошие и не очень новости о написании многопоточного кода в Rust (см. рекомендацию 17);
- рассказ о том, как избежать остановки процессов (`abort`) в среде выполнения (см. рекомендацию 18);
- информация о реализуемом в Rust подходе к отражению (см. рекомендацию 19);
- советы по балансированию оптимизации и обслуживания (см. рекомендацию 20).

Хорошо, если вы будете стараться писать код с учетом следствия этих концепций. Можно в некоторой степени воспроизвести поведение C/C++ в Rust, но зачем тогда вообще озадачиваться использованием Rust?

Рекомендация 14. Уясните принцип времени жизни

В текущей рекомендации объясняется принцип *времени жизни* в Rust, который является более точным выражением принципа, существовавшего в более ранних компилируемых языках вроде C и C++ — как в теории, так и на практике. Время жизни (lifetime) — это обязательный ввод от *модуля проверки заимствования*, описываемого в рекомендации 15. Вместе эти две единицы функциональности формируют основу гарантий Rust в отношении безопасности памяти.

Знакомство со стеком

Время жизни фундаментально связано со *стеком*, поэтому следует вкратце напомнить о том, что это за механизм.

В процессе выполнения программы используемая ею память делится на части, иногда называемые *сегментами*. Некоторые из этих частей имеют фиксированный размер, например те, которые содержат код программы или ее глобальные данные. Но размер двух частей — *кучи* и *стека* — в ходе выполнения программы меняется. Чтобы это было возможно, они обычно организуются в противоположных концах области виртуальной памяти программы. В таком случае один сегмент растет вниз, а другой — вверх (по крайней мере, пока программа не исчерпает память или не произойдет сбой). Схематично все это показано на рис. 3.1.

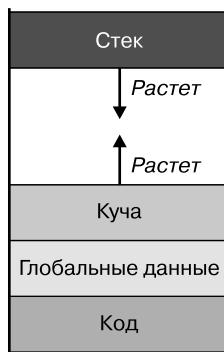


Рис. 3.1. Схема размещения памяти программы, включая растущую вверх кучу и увеличивающийся вниз стек

Стек используется для хранения состояния, связанного с выполняющейся в данный момент функцией. Оно может включать следующие элементы:

- передаваемые в функцию параметры;
- используемые в функции локальные переменные;
- вычисленные в функции временные значения;
- адрес возврата в коде элемента, вызвавшего функцию.

При вызове функции `f()` в стек добавляется новый фрейм, следующий за фреймом вызывающего кода, и процессор обычно обновляет регистр — *указатель стека*, направляя его на этот новый фрейм. Когда внутренняя функция `f()` завершается, указатель стека сбрасывается и начинает указывать туда, куда указывал до вызова, а именно на фрейм вызывающего кода, который остался неизмененным.

Если вызывающий код следом вызывает другую функцию `g()`, процесс повторяется, то есть фрейм стека для `g()` использует ту же область памяти, которую до этого задействовала `f()` (рис. 3.2):

```
fn caller() -> u64 {
    let x = 42u64;
    let y = 19u64;
    f(x) + g(y)
}

fn f(f_param: u64) -> u64 {
    let two = 2u64;
    f_param + two
}

fn g(g_param: u64) -> u64 {
    let arr = [2u64, 3u64];
    g_param + arr[1]
}
```

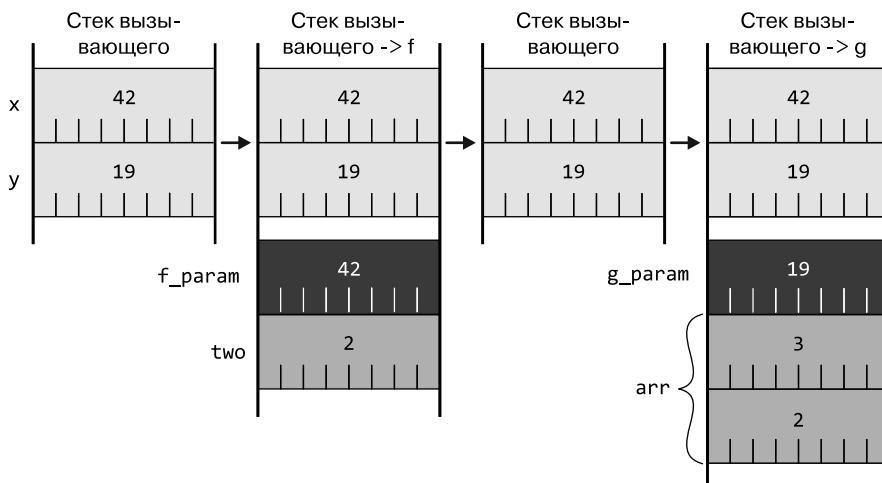


Рис. 3.2. Динамика использования стека в процессе вызова и завершения функций

Естественно, это крайне упрощенная версия того, как все происходит в действительности, — помещение элементов в стек и их извлечение требует времени,

в связи с чем реальные процессоры используют множество оптимизаций. Тем не менее такого упрощенного представления будет достаточно для понимания текущей рекомендации.

Эволюция времени жизни

В предыдущем разделе объяснялось, как параметры и локальные переменные сохраняются в стеке, а также что их значения сохраняются лишь временно. Традиционно эта схема вызывала возникновение опасных программных подводных камней: что произойдет, если сохранить указатель на одно из таких временных значений стека?

Начиная с C, можно было спокойно вернуть указатель на локальную переменную (хотя современные компиляторы в этом случае выдадут предупреждение).

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
/* Код C. */
struct File {
    int fd;
};

struct File* open_bugged() {
    struct File f = { open("README.md", O_RDONLY) };
    return &f; /* возврат адреса объекта! */
}
```

Такой вариант *может* сработать, если вам не повезет и вызывающий код использует возвращенное значение сразу:

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
struct File* f = open_bugged();
printf("in caller: file at %p has fd=%d\n", f, f->fd);
```

in caller: file at 0x7ff7bc019408 has fd=3

А невезение здесь в том, что работает он только *с виду*. Как только произойдет вызов любой другой функции, эта область стека вновь будет использована и память, в которой хранился объект, окажется переписана:

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
investigate_file(f);
```

```
/* Код C. */
void investigate_file(struct File* f) {
    long array[4] = {1, 2, 3, 4}; // Помещает элементы в стек
    printf("in function: file at %p has fd=%d\n", f, f->fd);
}

in function: file at 0x7ff7bc019408 has fd=1592262883
```

Порча содержимого объекта в данном примере вызывает дополнительный негативный эффект: дескриптор открытого файла утрачивается, из-за чего программа теряет ресурс, который содержался в структуре данных.

Впоследствии в C++ проблема потери доступа к ресурсам была решена включением *деструкторов*, позволяющих использовать RAII (подробнее — в рекомендации 11). Теперь содержащиеся в стеке элементы могли прибирать за собой. Если объект содержит некий ресурс, деструктор может его подчистить и компилятор C++ гарантирует, что деструктор объекта в стеке будет вызван в рамках чистки фрейма:

```
// Код C++
File::~File() {
    std::cout << "~File(): close fd " << fd << "\n";
    close(fd);
    fd = -1;
}
```

Теперь вызывающий код получает недействительный указатель на объект, который был уничтожен и ресурсы которого возвращены:

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
File* f = open_bugged();
printf("in caller: file at %p has fd=%d\n", f, f->fd);
```

```
~File(): close fd 3
in caller: file at 0x7ff7b6a7c438 has fd=-1
```

Тем не менее в C++ никак не решалась проблема висячих указателей: в нем все так же можно продолжать хранить указатель на удаленный объект (через вызов деструктора):

```
// Код C++
void investigate_file(File* f) {
    long array[4] = {1, 2, 3, 4}; // Помещает элементы в стек
    std::cout << "in function: file at " << f << " has fd=" << f->fd << "\n";
}

in function: file at 0x7ff7b6a7c438 has fd=-183042004
```

Поэтому программист на C/C++ должен заметить это и проследить, чтобы не произошло разыменование указателя, указывающего на что-то уже удаленное. Если же вы находитесь в роли атакующего и обнаруживаете такой висячий указатель, то наверняка злобно усмехнетесь и с радостью разыменуете его для реализации эксплойта.

Теперь перейдем к Rust. Одним из основных плюсов этого языка является то, что он фундаментально решает проблему висячих указателей, моментально исключая значительную долю проблем с безопасностью¹. Для этого необходимо перенести принцип времени жизни с заднего плана (когда программистам на C/C++ нужно просто не забывать следить за ними, не получая никакой поддержки со стороны языка) на передний: с каждым типом, который содержит амперсанд &, связано определенное время жизни ('а), даже если компилятор в большинстве случаев позволяет вам опустить упоминание о нем.

Продолжительность времени жизни

Время жизни элемента в стеке — это период, в течение которого элемент гарантированно будет находиться в одном и том же месте. Иными словами, это ровно тот период, когда *ссылка* (указатель) на элемент не станет недействительной. Начинается он в момент создания элемента и продолжается до момента, когда он либо *отбрасывается* (в Rust это равнозначно уничтожению в C++), либо *перемещается*.

Повсеместность последнего порой удивляет программистов, приходящих из C/C++: Rust во многих ситуациях перемещает элементы из одной позиции в стеке в другую, или из стека в кучу, или из кучи в стек.

Точное место, в котором элемент автоматически отбрасывается, зависит от того, имеет этот элемент имя или нет. Локальные переменные и параметры функций имеют имена, и время жизни соответствующего элемента начинается в момент его создания, когда ему присваивается имя:

- для локальной переменной — в объявлении `let var = ...;`
- для параметра функции — в рамках установки фрейма выполнения для вызова функции.

Время жизни именованного элемента заканчивается, либо когда он куда-то перемещается, либо когда его имя выходит из области действия:

```
#[derive(Debug, Clone)]  
/// Определение некоего элемента
```

¹ Так, по примерным оценкам проекта Chromium, 70 % багов безопасности связаны с безопасностью памяти (<https://oreil.ly/GJkt0>).

```
pub struct Item {  
    contents: u32,  
}  
  
{  
    let item1 = Item { contents: 1 }; // Здесь создается `item1`  
    let item2 = Item { contents: 2 }; // Здесь создается `item2`  
    println!("item1 = {item1:?}, item2 = {item2:?}");  
    consuming_fn(item2); // `item2` перемещен сюда  
} // Здесь `item1` отбрасывается
```

Также можно создавать элемент на лету в рамках выражения, которое затем передается куда-то еще. Впоследствии эти безымянные временные элементы отбрасываются, когда потребность в них исчезает. Чтобы лучше понять это, можно в упрощенном виде представить, что каждая часть выражения расширяется до границ ее собственного блока и временные переменные вставляются компилятором. Например, выражение:

```
let x = f((a + b) * 2);
```

будет примерно равнозначно следующему:

```
let x = {  
    let temp1 = a + b;  
    {  
        let temp2 = temp1 * 2;  
        f(temp2)  
    } // Здесь `temp2` отбрасывается  
}; // Здесь `temp1` отбрасывается
```

К моменту, когда выполнение достигнет точки с запятой в конце исходной строки, все временные значения будут отброшены.

Для того чтобы увидеть, какое время жизни элемента вычисляет компилятор, можно намеренно добавить ошибку, которую должен будет заметить анализатор заимствований (см. рекомендацию 15). Например, можно сохранить ссылку на элемент вне периода его времени жизни.

НЕ КОМПИЛИРУЕТСЯ

```
let r: &Item;  
{  
    let item = Item { contents: 42 };  
    r = &item;  
}  
println!("r.contents = {}", r.contents);
```

Сообщение об ошибке указывает конкретный момент завершения времени жизни `item`:

```
error[E0597]: `item` does not live long enough
--> src/main.rs:190:13
|
189 |         let item = Item { contents: 42 };
|             ----- binding `item` declared here
190 |         r = &item;
|             ^^^^^^ borrowed value does not live long enough
191 |
| - `item` dropped here while still borrowed
192 | println!("r.contents = {}", r.contents);
|             ----- borrow later used here
```

Аналогично для безымянного временного элемента:

НЕ КОМПИЛИРУЕТСЯ

```
let r: &Item = fn_returning_ref(&mut Item { contents: 42 });
println!("r.contents = {}", r.contents);
```

сообщение об ошибке показывает момент завершения жизни в конце выражения:

```
error[E0716]: temporary value dropped while borrowed
--> src/main.rs:209:46
|
209 |     let r: &Item = fn_returning_ref(&mut Item { contents: 42 });
|             ^^^^^^^^^^^^^^^^^^^^^ - temporary
|                         |           value is freed at the
|                         |           end of this statement
|                         |
|                         creates a temporary value which is
|                         freed while still in use
210 |     println!("r.contents = {}", r.contents);
|             ----- borrow later used here
|
= note: consider using a `let` binding to create a longer lived value
```

И последний нюанс концепции времени жизни *ссылок*: если компилятор может доказать себе, что ссылка станет бесполезной после определенного момента в коде, то в качестве точки завершения ее жизни он рассматривает место последнего использования этой ссылки, а не конец включающей ее области. Эта особенность, известная как нелексическое время жизни (<https://oreil.ly/4uJ73>), позволяет модулю проверки заимствований действовать чуть деликатнее:

```
{  
    // `s` владеет `String`  
    let mut s: String = "Hello, world".to_string();  
  
    // Создает мутабельную ссылку на `String`  
    let greeting = &mut s[..5];  
    greeting.make_ascii_uppercase();  
    // .. после этого момента `greeting` не используется  
  
    // Создание иммутабельной ссылки на `String` допустимо,  
    // несмотря на то что мутабельная ссылка до сих пор находится  
    // в области действия  
    let r: &str = &s;  
    println!("s = '{}'", r); // s = 'HELLO, world'  
} // Мутабельная ссылка `greeting` будет попросту отброшена здесь
```

Арифметика времени жизни

Несмотря на то что в ходе работы со ссылками в Rust время жизни встречается повсеместно, его не нужно прописывать подробно. Здесь нет способа сказать: «Я работаю с временем жизни, которое охватывает период от строки 17 до строки 32 в `ref.rs`». Вместо этого в программе обращение к времени жизни происходит с использованием произвольных имен (традиционно '`a`', '`b`', '`c`...), и компилятор внутренне имеет собственное, недоступное для программиста представление того, чему эти имена соответствуют в исходном коде. (Единственным исключением здесь будет время жизни '`static`', представляющее особый случай, рассматриваемый в одном из следующих подразделов.)

И с этими именами времени жизни ничего особо не сделаешь. Основное, что доступно, — это сравнение двух имен между собой, когда имя повторяется, указывая на то, что два времени жизни совпадают.

Арифметику времени жизни легче всего проиллюстрировать на сигнатурах функций: если ввод и вывод функции связаны со ссылками, то какой будет связь между их временами жизни?

Самым распространенным случаем является функция, которая получает одну ссылку на входе и выдает другую на выходе. Ссылка на выходе должна иметь время жизни, но каким оно может быть? И здесь есть лишь один вариант помимо '`static`' — время жизни входной ссылки. А это означает, что у обеих одно имени, например '`a`'. Добавив это имя в качестве аннотации времени жизни для обоих типов, мы получим:

```
pub fn first<'a>(data: &'a [Item]) -> Option<&'a Item> {  
    // ...  
}
```

Поскольку этот вариант очень распространен и опций для возможного времени жизни на выходе практически нет, в Rust действуют правила *пропуска времени жизни*. Это означает, что в данном случае вам не нужно явно прописывать имена времени жизни. Более идиоматической версией той же сигнатуры функции будет следующая:

```
pub fn first(data: &[Item]) -> Option<&Item> {
    // ...
}
```

Присутствующие здесь ссылки по-прежнему имеют время жизни — правило пропуска означает лишь то, что вам не нужно придумывать произвольное имя и использовать его в обоих местах.

А что, если для отображения во время жизни на выходе на входе есть больше одного его варианта? В этом случае компилятор сделать выбор не сможет:

НЕ КОМПИЛИРУЕТСЯ

```
pub fn find(haystack: &[u8], needle: &[u8]) -> Option<&[u8]> {
    // ...
}
```

```
error[E0106]: missing lifetime specifier
--> src/main.rs:56:55
|
56 | pub fn find(haystack: &[u8], needle: &[u8]) -> Option<&[u8]> {
|           -----      -----      ^ expected named
|           |           |
|           |           lifetime parameter
|
= help: this function's return type contains a borrowed value, but the
      signature does not say whether it is borrowed from `haystack` or
      `needle`
help: consider introducing a named lifetime parameter
|
56 | pub fn find<'a>(haystack: &'a [u8], needle: &'a [u8]) -> Option<&'a [u8]> {
|     +++          ++          ++          ++
|
```

Опираясь на имена функции и параметров, правильно будет предположить, что время жизни вывода должно соответствовать входному `haystack`:

```
pub fn find<'a, 'b>(
    haystack: &'a [u8],
    needle: &'b [u8],
) -> Option<&'a [u8]> {
    // ...
}
```

Интересно, что компилятор предложил другую альтернативу: чтобы оба входных параметра функции использовали *одинаковое* время жизни 'a. К примеру, далее показана функция, где такая комбинация времени жизни может иметь смысл:

```
pub fn smaller<'a>(left: &'a Item, right: &'a Item) -> &'a Item {
    // ...
}
```

Этот вариант *как бы* предполагает, что времена жизни двух входных элементов «одинаково», но пугающие кавычки здесь и ранее подчеркивают, что происходит все не совсем так.

Смысл существования времени жизни в том, чтобы ссылки на элементы не жили дольше самих этих элементов. С учетом этого время жизни 'a выходного элемента просто означает, что входной элемент должен жить дольше вывода.

Когда присутствует два входных элемента с «одинаковым» временем жизни 'a, это лишь означает, что время жизни вывода должно вписываться во время жизни *обоих* этих элементов:

```
{
    let outer = Item { contents: 7 };
    {
        let inner = Item { contents: 8 };
        {
            let min = smaller(&inner, &outer);
            println!("smaller of {inner:?} and {outer:?} is {min:?}");
        } // `min` отброшен
    } // `inner` отброшен
} // `outer` отброшен
```

Иными словами, время жизни вывода должно умещаться в *меньшую* из продолжительностей жизни двух входных элементов.

И напротив, если время жизни вывода не связано с временем жизни одного из входных элементов, то такого требования нет:

```
{
    let haystack = b"123456789"; // Начало времени жизни 'a
    let found = {
        let needle = b"234"; // Начало времени жизни 'b
        find(haystack, needle)
    }; // Конец времени жизни 'b
    // `found` используется в рамках 'a, вне 'b
    println!("found={:?}", found);
} // Конец времени жизни 'a
```

Правила пропуска времени жизни

Помимо правила пропуска «один на входе, один на выходе», описанного в разделе «Арифметика времени жизни» ранее, есть еще два правила, означающих, что имя времени жизни можно опускать.

Первое действует, когда в выводе функции отсутствуют ссылки. В этом случае каждая из входных ссылок автоматически получает собственное время жизни, отличное от времени жизни любого другого входного параметра.

Второе правило работает для методов, которые используют ссылку на `self` (`&self` либо `&mut self`). В этом случае компилятор предполагает, что ссылки на выходе получают время жизни `self`, так как пока что это наиболее типичная ситуация.

Вот краткое описание правил пропуска для функций.

- Один элемент на входе, один или более на выходе. Предположим, что у вывода «такое же» время жизни, как у ввода:

```
fn f(x: &Item) -> (&Item, &Item)
// ... равнозначно ...
fn f<'a>(x: &'a Item) -> ('a Item, &'a Item)
```

- Несколько элементов на входе, вывода нет. Предположим, что у всех входных элементов разное время жизни:

```
fn f(x: &Item, y: &Item, z: &Item) -> i32
// ... равнозначно ...
fn f<'a, 'b, 'c>(x: &'a Item, y: &'b Item, z: &'c Item) -> i32
```

- Несколько входных элементов, включая `&self`, один или несколько элементов на выходе. Предположим, что времена жизни элементов на выходе «такое же», как времена жизни `&self`:

```
fn f(&self, y: &Item, z: &Item) -> &Thing
// ... равнозначно ...
fn f(&'a self, y: &'b Item, z: &'c Item) -> &'a Thing
```

Естественно, если опускаемые имена циклов жизни не совпадают с нужными вам, вы можете явно прописать, какие сроки жизни связаны между собой. На практике к этому наверняка будет подталкивать ошибка компилятора, указывающая, что пропущенные сроки жизни не соответствуют тому, как функция или вызывающий ее компонент используют имеющиеся ссылки.

Время жизни 'static

В предыдущем разделе мы разобрали различные варианты отображения в функции времени жизни входных ссылок во время жизни выходных, но опустили один особый случай. Что происходит, если на входе *нет* элементов с временем жизни, но выходное значение все равно включает ссылку?

НЕ КОМПИЛИРУЕТСЯ

```
pub fn the_answer() -> &Item {  
    // ...  
}
```

```
error[E0106]: missing lifetime specifier  
--> src/main.rs:471:28  
|  
471 |     pub fn the_answer() -> &Item {  
|             ^ expected named lifetime parameter  
|  
= help: this function's return type contains a borrowed value, but there  
      is no value for it to be borrowed from  
help: consider using the `<code>'static` lifetime  
|  
471 |     pub fn the_answer() -> &'static Item {  
|         ++++++
```

Единственным допустимым вариантом будет ситуация, когда время жизни возвращаемой ссылки гарантированно не покинет область действия. Обозначается оно в таком случае с помощью `'static` — единственного специального имени, отличного от его произвольных вариантов, о которых говорилось ранее:

```
pub fn the_answer() -> &'static Item {
```

Самый простой способ получить что-то с временем жизни `'static` — это взять ссылку на глобальную переменную, обозначенную `static`:

```
static ANSWER: Item = Item { contents: 42 };  
  
pub fn the_answer() -> &'static Item {  
    &ANSWER  
}
```

Компилятор Rust гарантирует, что `static`-элемент в ходе выполнения всей программы всегда будет иметь один и тот же адрес и не переместится. В связи с этим ссылка на `static`-элемент имеет время жизни '`static`', что вполне логично.

Во многих случаях время жизни ссылки на элемент `const` тоже будет расширяться до '`static`', но здесь есть пара осложнений. Первое состоит в том, что этого не произойдет, если у причастного к процессу типа есть деструктор или он допускает внутреннее изменение:

НЕ КОМПИЛИРУЕТСЯ

```
pub struct Wrapper(pub i32);

impl Drop for Wrapper {
    fn drop(&mut self) {}
}

const ANSWER: Wrapper = Wrapper(42);

pub fn the_answer() -> &'static Wrapper {
    // `Wrapper` содержит деструктор, значит, время жизни ссылки
    // на константу не расширяется до `static`
    &ANSWER
}
```

```
error[E0515]: cannot return reference to temporary value
--> src/main.rs:520:9
520 |     &ANSWER
|     ^-----
|     |
|     || temporary value created here
|     returns a reference to data owned by the current function
```

Второе возможное осложнение заключается в том, что гарантированно одинаковым везде будет только *значение* `const`. Компилятор может делать любое необходимое число копий там, где эта переменная используется. Если вы реализуете какой-то подлый замысел, который опирается на внутреннее значение указателя за ссылкой '`static`', учитывайте, что это может затрагивать множество областей памяти.

Есть и еще один способ получить что-либо с временем жизни '`static`'. Ключевая гарантия '`static`' заключается в том, что это время жизни должно быть больше времени жизни любого элемента программы. Значение, которое выделено в куче, но *никогда не освобождается*, тоже это требование выполняет.

Для такого случая не подойдет стандартный, выделенный в куче тип `Box<T>`, поскольку нет гарантии, что в ходе выполнения программы элемент не будет отброшен (об этом поговорим в следующем разделе):

НЕ КОМПИЛИРУЕТСЯ

```
{  
    let boxed = Box::new(Item { contents: 12 });  
    let r: &'static Item = &boxed;  
    println!("'static item is {:?}", r);  
}
```

```
error[E0597]: `boxed` does not live long enough  
--> src/main.rs:344:32  
|  
343 |     let boxed = Box::new(Item { contents: 12 });  
|         ----- binding `boxed` declared here  
344 |     let r: &'static Item = &boxed;  
|         ----- ^^^^^^ borrowed value does not live long enough  
|         |  
|             type annotation requires that `boxed` is borrowed for `&'static`  
345 |     println!("'static item is {:?}", r);  
346 | }  
| - `boxed` dropped here while still borrowed
```

Тем не менее функция `Box::leak` преобразует `Box<T>`, которым она владеет, в mutableльную ссылку на `T`. Больше у этого значения владельца нет, значит, отбросить его нельзя — и это уже соответствует требованиям времени жизни `'static'`:

```
{  
    let boxed = Box::new(Item { contents: 12 });  
  
    // `leak()` получает `Box<T>` и возвращает `&mut T`  
    let r: &'static Item = Box::leak(boxed);  
  
    println!("'static item is {:?}", r);  
} // `boxed` здесь не отбрасывается, так как уже был перемещен в `Box::leak()`  
  
// Поскольку `r` теперь находится вне рабочей области,  
// `Item` утрачен навсегда
```

Невозможность отбросить элемент также означает, что память, в которой он содержится, с помощью безопасного кода Rust вернуть не получится, а это может привести к необратимой утечке памяти. (Обратите внимание на то, что утечка памяти не нарушает гарантии Rust в отношении ее безопасности — элемент в памяти, к которому у вас больше нет доступа, по-прежнему безопасен.)

Время жизни и куча

До этого момента речь шла о времени жизни элементов стека, будь то параметры функций, локальные переменные или временные элементы. Но что насчет элементов в куче?

В отношении кучи самое главное — понимать, что у каждого ее элемента есть владелец, за исключением особых случаев вроде сознательно допущенных утечек, о которых говорилось в предыдущем разделе. Например, простой `Box<T>` помещает значение `T` в кучу, притом что его владельцем выступает переменная, содержащая `Box<T>`:

```
{  
    let b: Box<Item> = Box::new(Item { contents: 42 });  
} // Здесь `b` отбрасывается, значит, и `Item` тоже
```

Владеющий `Box<Item>` отбрасывает свое содержимое, когда выходит из области действия, поэтому время жизни `Item` в куче совпадает с временем жизни переменной `Box<Item>` в стеке.

Владелец значения в куче может сам находиться в куче, а не в стеке, но кто тогда владеет этим владельцем?

```
{  
    let b: Box<Item> = Box::new(Item { contents: 42 });  
    let bb: Box<Box<Item>> = Box::new(b); // Здесь `b` перемещается в кучу  
} // Здесь `bb` отбрасывается, значит, и `Box<Item>`  
// тоже отбрасывается, а с ним и `Item`
```

Такая цепочка владения должна где-то завершиться, и здесь есть всего два варианта.

- Цепочка заканчивается локальной переменной или параметром функции, в случае чего время жизни всех ее элементов равняется времени жизни 'а' этой переменной стека. Когда переменная стека покидает область действия, все содержимое цепочки тоже отбрасывается.
- Цепочка завершается на глобальной переменной, отмеченной как `static`, в случае чего время жизни в этой цепочке является '`static`'. Переменная `static` никогда не выходит из области действия программы, а значит, ничто в цепочке автоматически отброшено не будет.

В результате время жизни элементов в куче фундаментально привязывается к времени жизни в стеке.

Время жизни в структурах данных

В одном из предыдущих разделов, посвященном арифметике времени жизни, рассматривались входные данные функций и их вывод, но при хранении ссылок в структурах данных возникают аналогичные вопросы.

Если мы попытаемся внести ссылку в структуру данных, не указывая связанного с ней времени жизни, компилятор будет ругаться:

НЕ КОМПИЛИРУЕТСЯ

```
pub struct ReferenceHolder {  
    pub index: usize,  
    pub item: &Item,  
}
```

```
error[E0106]: missing lifetime specifier  
--> src/main.rs:548:19  
|  
548 |     pub item: &Item,  
|         ^ expected named lifetime parameter  
|  
help: consider introducing a named lifetime parameter  
|  
546 ~     pub struct ReferenceHolder<'a> {  
547 |         pub index: usize,  
548 ~         pub item: &'a Item,  
|
```

Как обычно, в выданной ошибке компилятор подсказывает нам, что делать. Первая часть довольно проста — присвоить типу ссылки явное имя времени жизни `'a`, поскольку при использовании ссылок в структурах данных правила пропуска не действуют.

Вторая часть менее очевидна и имеет более глубокие следствия — в самой структуре данных должен присутствовать параметр времени жизни `<'a>`, который соответствует времени жизни содержащейся в ней ссылки:

```
// В связи с наличием поля со ссылкой необходим параметр времени жизни  
pub struct ReferenceHolder<'a> {  
    pub index: usize,  
    pub item: &'a Item,  
}
```

При этом параметр времени жизни для структуры данных можно назвать — любая содержащая что-либо структура данных, которая использует этот тип, тоже должна получить аналогичный параметр:

```
// В связи с наличием поля, тип которого имеет параметр времени жизни,  
// здесь этот параметр тоже необходим  
pub struct RefHolderHolder<'a> {  
    pub inner: ReferenceHolder<'a>,  
}
```

Потребность в параметре времени жизни возникает также, если структура данных содержит типы срезов, поскольку они опять же являются ссылками на заимствованные данные.

Если структура данных содержит несколько полей, имеющих связанное время жизни, вам нужно выбирать, какая из комбинаций сроков жизни окажется подходящей. Пример кода, который ищет внутри пары строк общие подстроки, — хороший кандидат на использование независимых периодов жизни:

```
/// Места расположения подстроки, которая присутствует в обеих строках пары  
pub struct LargestCommonSubstring<'a, 'b> {  
    pub left: &'a str,  
    pub right: &'b str,  
}  
  
/// Поиск большей подстроки, имеющейся в `left` и `right`  
pub fn find_common<'a, 'b>(  
    left: &'a str,  
    right: &'b str,  
) -> Option<LargestCommonSubstring<'a, 'b>> {  
    // ...  
}
```

При этом структура данных, которая ссылается на несколько мест внутри одной строки, будет иметь общее время жизни:

```
/// Первые два экземпляра подстроки, которая повторяется в строке  
pub struct RepeatedSubstring<'a> {  
    pub first: &'a str,  
    pub second: &'a str,  
}  
  
/// Поиск первой повторяющейся подстроки, присутствующей в `s`  
pub fn find_repeat<'a>(s: &'a str) -> Option<RepeatedSubstring<'a>> {  
    // ...  
}
```

Такое распространение параметров времени жизни имеет смысл: все, что содержит ссылку, вне зависимости от глубины ее вложенности валидно только

в течение времени жизни элемента, к которому ссылка ведет. Если этот элемент перемещается или отбрасывается, то вся цепочка структур данных перестает действовать.

Это означает также, что структуры данных, включающие ссылки, сложнее использовать — владелец структуры данных должен обеспечить, чтобы все сроки жизни были согласованы. В итоге по возможности нужно отдавать предпочтение структурам данных, которые владеют своим содержимым, особенно если их код не нужно сильно оптимизировать (см. рекомендацию 20). Если же такое невозможно, разрешить связанные с временем жизни ограничения помогут умные указатели (например, `Rc`), описанные в рекомендации 8.

Анонимное время жизни

Когда нет возможности использовать структуры данных, которые владеют своим содержимым, любую другую структуру потребуется сопроводить параметром времени жизни, о чём говорилось в предыдущем разделе. Это может несколько усложнить работу с правилами пропуска времени жизни, описанными ранее.

Рассмотрим в качестве примера функцию, которая возвращает структуру данных с параметром времени жизни. Абсолютно явная сигнатура этой функции отчетливо проясняет все имеющиеся сроки жизни элементов:

```
pub fn find_one_item<'a>(items: &'a [Item]) -> ReferenceHolder<'a> {
    // ...
}
```

Но та же сигнатура без указания времени жизни может вводить в заблуждение:

```
pub fn find_one_item(items: &[Item]) -> ReferenceHolder {
    // ...
}
```

Поскольку параметр времени жизни возвращаемого типа опущен, при чтении кода программист не поймет, что здесь нужно учитывать время жизни.

Анонимное время жизни `'_` позволяет отмечать опущенные сроки жизни как присутствующие, не требуя полностью воспроизводить *все* их имена:

```
pub fn find_one_item(items: &[Item]) -> ReferenceHolder<'_> {
    // ...
}
```

Грубо говоря, маркер `'_` просит компилятор придумать уникальное имя времени жизни, которое можно использовать в ситуациях, когда это имя больше нигде не пригодится.

Это решение полезно и для других сценариев с пропуском времени жизни. Например, в объявлении метода `fmt` трейта `Debug` анонимное время жизни используется для указания того, что время жизни экземпляра `Formatter` отличается от времени жизни `&self`, но его имя значения не имеет:

```
pub trait Debug {  
    fn fmt(&self, f: &mut Formatter<'_>) -> Result<(), Error>;  
}
```

Запомните

- У всех ссылок в Rust есть связанное с ними время жизни, обозначаемое специальной меткой, например '`a`'. Такие метки для параметров и возвращаемых значений функций в некоторых распространенных случаях можно пропускать (но они никуда не исчезают).
- Любая структура данных, которая транзитивно включает ссылку, имеет сопутствующий параметр времени жизни. В результате зачастую проще работать со структурами данных, которые владеют своим содержимым.
- Время жизни '`static`' используется для ссылок на элементы, которые гарантированно никогда не покинут область действия, например на глобальные данные или элементы в куче, доступ к которым был сознательно утрачен.
- Метки времени жизни можно использовать только для указания того, что сроки жизни «совпадают», то есть время жизни вывода умещается во время жизни ввода.
- Метку анонимного времени жизни '`_`' можно использовать там, где нет нужды в конкретной метке срока жизни.

Рекомендация 15. Изучите работу модуля проверки заимствования

В Rust у значений есть владелец и он может одалживать значения другим участникам кода. Такой механизм *заимствования* подразумевает создание и использование *ссылок* в соответствии с правилами, отслеживаемыми *модулем проверки заимствования*, о котором и пойдет речь в текущей рекомендации.

Ссылки в Rust используют такие же значения *указателей* (см. рекомендацию 8), которые распространены в коде C/C++, но здесь, чтобы избежать известных проблем этих языков, они дополнительно обвешаны различными правилами и ограничениями. Вот краткое сравнение.

- Как и указатель в C/C++, ссылка в Rust создается с амперсандом — `&value`.
- И в C++, и в Rust ссылка никогда не может быть `nullptr`.
- Аналогично ссылкам в C/C++ в Rust ссылку после ее создания можно изменять, направляя на что-либо другое.
- В отличие от C++ создание ссылки из значения всегда включает явное (`&`) преобразование. Если вы увидите код вроде `f(value)`, то поймете, что `f` получает владение над `value`. (Однако это может быть владение *копией* элемента, если тип `value` реализует `Copy`. Подробнее — в рекомендации 10.)
- В отличие от C/C++ в Rust возможность изменения созданной ссылки всегда явная (`&mut`). Если вы видите код вроде `f(&value)`, то знайте, что `value` изменяться не будет (то есть в терминологии C/C++ является `const`). Изменять содержимое `value` могут только выражения вроде `f(&mut value)`¹.

Самым важным различием между указателем C/C++ и ссылкой в Rust является то, что ссылка обозначается термином «*заимствование*»: вы можете взять ссылку (указатель) на элемент, *но не можете сохранить ее навсегда*. В частности, не можете хранить ее дольше чем время жизни связанного с ней элемента, что отслеживается компилятором (и рассматривается в рекомендации 14).

Ограничения по использованию ссылок позволяют Rust обеспечивать гарантии безопасности, но также создают для вас дополнительную когнитивную нагрузку по освоению правил заимствования и влияют на то, как вы проектируете ПО, — в частности, на структуры данных.

Эта рекомендация начинается с описания возможностей ссылок в Rust, а также правил их использования, за которыми следует анализатор заимствований. Остальная ее часть посвящена работе с последствиями этих правил: как рефакторить, дорабатывать и перепроектировать код так, чтобы анализатор не возмущался.

Контроль доступа

Получить доступ к содержимому элемента в Rust можно тремя способами: через его *владельца* (`item`), *ссылку* (`&item`) или *мутабельную ссылку* (`&mut item`). Каждый из этих способов дает определенную власть над элементом. Если взглянуть на это в контексте модели CRUD (create/read/update/delete — «создание/чтение/изменение/удаление») в отношении хранения (используя

¹ Имейте в виду, что в случае использования выражений вроде `m!(value)`, которые включают макрос (см. рекомендацию 28), все становится непредсказуемым, поскольку он может расширяться на произвольный код.

вместо термина «удаление» термин Rust «отбрасывание»), то выяснится следующее.

- Владелец элемента может его *создать, прочитать, изменить и отбросить*.
- Мутабельные ссылки можно использовать для *чтения* из связанного с ними элемента и его *изменения*.
- Стандартную ссылку можно применять только для *чтения* из связанного элемента.

В Rust у этих правил доступа есть одна особенность: *перемещать* элемент может только его владелец. Это разумно, если рассматривать перемещение как комбинацию *создания* (в новом расположении) и *отбрасывания* памяти элемента (в старом расположении).

И эта особенность может привести к странностям в коде, имеющем мутабельную ссылку на элемент. Например, вполне можно переписать `Option`:

```
/// Некая структура данных, используемая кодом
#[derive(Debug)]
pub struct Item {
    pub contents: i64,
}

/// Замена содержимого `item` на `val`
pub fn replace(item: &mut Option<Item>, val: Item) {
    *item = Some(val);
}
```

Но изменение, подразумевающее возвращение прежнего значения, нарушает связанное с перемещением ограничение¹:

НЕ КОМПИЛИРУЕТСЯ

```
/// Замена содержимого `item` на `val` с возвратом прежнего содержимого
pub fn replace(item: &mut Option<Item>, val: Item) -> Option<Item> {
    let previous = *item; // Перемещение
    *item = Some(val); // Замена
    previous
}
```

```
error[E0507]: cannot move out of `*item` which is behind a mutable reference
--> src/main.rs:34:24
|
```

¹ Рекомендация компилятора здесь не поможет, поскольку `item` необходим на следующей строке.

```
34 |         let previous = *item; // Перемещение
|             ^^^^^^ move occurs because `*item` has type
|                     `Option<inner::Item>`, which does not
|                     implement the `Copy` trait
|
| help: consider removing the dereference here
|
34 -         let previous = *item; // Перемещение
34 +         let previous = item; // Перемещение
|
```

Несмотря на то что из мутабельной ссылки допустимо *читать*, этот код пытается *переместить* значение непосредственно перед его заменой новым, стремясь избежать создания копии исходного значения. Анализатор в данном случае должен быть строг и замечать, что между этими двумя строками есть момент, когда мутабельная ссылка не указывает на валидное значение.

Мы же, как люди, можем видеть, что эта комбинированная операция — извлечение старого значения и его замена новым — безопасна и полезна, поэтому в стандартной библиотеке для ее выполнения есть функция `std::mem::replace`. Для реализации этой перестановки за один шаг `replace` использует `unsafe` (как описано в рекомендации 16):

```
// Замена содержимого `item` на `val` с возвратом прежнего содержимого
pub fn replace(item: &mut Option<Item>, val: Item) -> Option<Item> {
    std::mem::replace(item, Some(val)) // Возвращает прежнее значение
}
```

В частности, для типов `Option` это настолько распространенный паттерн, что в самом `Option` есть специальный метод `replace`:

```
// Замена содержимого `item` на `val` с возвратом прежнего значения
pub fn replace(item: &mut Option<Item>, val: Item) -> Option<Item> {
    item.replace(val) // Возвращает прежнее значение
}
```

Правила заимствования

При заимствовании ссылок в Rust нужно помнить два ключевых правила.

Первое заключается в том, что область действия любой ссылки должна быть меньше времени жизни элемента, на который она указывает. Тема времени жизни подробно разбирается в рекомендации 14, но стоит отметить, что в отношении времени жизни ссылок компилятор действует особым образом. Функциональность *нелексического времени жизни* позволяет сокращать сроки жизни переменных, чтобы они заканчивались в точке последнего использования, а не по завершении окружающего блока кода.

Второе правило заимствования ссылок состоит в том, что в добавок к владельцу элемента могут иметься:

- любое количество иммутабельных ссылок на этот элемент;
- одна мутабельная ссылка на него.

Однако и то и другое одновременно не может присутствовать в одном месте кода. Поэтому в функцию, которая получает несколько иммутабельных ссылок, можно передавать ссылки на один и тот же элемент:

```
/// Указывает, являются ли оба аргумента нулевыми
fn both_zero(left: &Item, right: &Item) -> bool {
    left.contents == 0 && right.contents == 0
}

let item = Item { contents: 0 };
assert!(both_zero(&item, &item));
```

Если же функция получает *мутабельные* ссылки, то этого делать нельзя:

НЕ КОМПИЛИРУЕТСЯ

```
/// Обнуление содержимого обоих аргументов
fn zero_both(left: &mut Item, right: &mut Item) {
    left.contents = 0;
    right.contents = 0;
}

let mut item = Item { contents: 42 };
zero_both(&mut item, &mut item);
```

```
error[E0499]: cannot borrow `item` as mutable more than once at a time
--> src/main.rs:131:26
|
131 |     zero_both(&mut item, &mut item);
|     ----- ^^^^^^^^^^ second mutable borrow occurs here
|     |
|     |         first mutable borrow occurs here
|     first borrow later used by call
```

То же ограничение верно для функции, которая использует комбинацию из мутабельных и иммутабельных ссылок:

НЕ КОМПИЛИРУЕТСЯ

```
/// Устанавливает содержимое `left` на содержимое `right`
fn copy_contents(left: &mut Item, right: &Item) {
    left.contents = right.contents;
}

let mut item = Item { contents: 42 };
copy_contents(&mut item, &item);
```

```
error[E0502]: cannot borrow `item` as immutable because it is also
            borrowed as mutable
--> src/main.rs:159:30
159 |     copy_contents(&mut item, &item);
      |     ----- ^^^^^ immutable borrow occurs here
      |     |
      |     |         |
      |     |         mutable borrow occurs here
      |     |
      |     mutable borrow later used by call
```

Эти правила заимствования позволяют компилятору принимать более объективные решения в отношении *псевдонимов*, то есть отслеживать, когда два разных указателя могут или не могут ссылаться на один элемент в памяти. Если компилятор может убедиться (как в Rust), что область памяти, на которую указывает коллекция иммутабельных ссылок, не может быть изменена через связанную *мутабельную* ссылку, то он может генерировать код, обладающий следующими преимуществами.

Лучшая оптимизация

Значения можно, например, кэшировать в регистрах, будучи уверенными, что внутреннее содержимое памяти за это время не изменится.

Повышенная безопасность

Гонки данных, возникающие в результате асинхронного доступа к памяти разных потоков (см. рекомендацию 17), оказываются невозможными.

Операции владельца

Одно из важных следствий правил существования ссылок состоит в том, что они также определяют, какие операции может выполнять владелец элемента. Чтобы лучше понять это, можно представить, что операции с участием владельца выполняются путем создания и использования ссылок.

Например, попытка обновить элемент через его владельца равнозначна созданию временной мутабельной ссылки и последующему обновлению этого элемента с ее помощью. Если уже существует другая ссылка, то создать гипотетическую мутабельную вторую нельзя:

НЕ КОМПИЛИРУЕТСЯ

```
let mut item = Item { contents: 42 };
let r = &item;
item.contents = 0;
// ^^^ Изменение элемента примерно равнозначно:
//      (&mut item).contents = 0;
println!("reference to item is {:?}", r);
```

```
error[E0506]: cannot assign to `item.contents` because it is borrowed
--> src/main.rs:200:5
|
199 |     let r = &item;
|         ----- `item.contents` is borrowed here
200 |     item.contents = 0;
|     ^^^^^^^^^^^^^^^^^ `item.contents` is assigned to here but it was
|                         already borrowed
...
203 |     println!("reference to item is {:?}", r);
|                                     - borrow later used here
```

В то же время, поскольку допустимо создание нескольких *иммутабельных* ссылок, владелец вполне может выполнять чтение из элемента, пока они существуют:

```
let item = Item { contents: 42 };
let r = &item;
let contents = item.contents;
// ^^^ Чтение из элемента примерно равнозначно:
//      let contents = (&item).contents;
println!("reference to item is {:?}", r);
```

но не в случае, когда присутствует *мутабельная* ссылка:

НЕ КОМПИЛИРУЕТСЯ

```
let mut item = Item { contents: 42 };
let r = &mut item;
let contents = item.contents; // i64 реализует `Copy`
r.contents = 0;
```

```
error[E003]: cannot use `item.contents` because it was mutably borrowed
--> src/main.rs:231:20
|
230 |     let r = &mut item;
|             ----- `item` is borrowed here
231 |     let contents = item.contents; // i64 реализует `Copy`
|             ^^^^^^^^^^^^^^ use of borrowed `item`
232 |     r.contents = 0;
|             ----- borrow later used here
```

Наконец, существование любой активной ссылки не позволяет владельцу элемента перемещать или отбрасывать его, так как в результате этого ссылка начнет указывать на невалидный элемент:

НЕ КОМПИЛИРУЕТСЯ

```
let item = Item { contents: 42 };
let r = &item;
let new_item = item; // Перемещение
println!("reference to item is {:?}", r);
```

```
error[E0505]: cannot move out of `item` because it is borrowed
--> src/main.rs:170:20
|
168 |     let item = Item { contents: 42 };
|             ---- binding `item` declared here
169 |     let r = &item;
|             ----- borrow of `item` occurs here
170 |     let new_item = item; // Перемещение
|             ^^^^ move out of `item` occurs here
171 |     println!("reference to item is {:?}", r);
|                         - borrow later used here
```

Это сценарий, где описанная в рекомендации 14 функциональность нелексического времени жизни оказывается особенно полезной, поскольку, грубо говоря, завершает жизнь ссылки в месте ее последнего использования, а не в конце окружающей области видимости. Переместив точку последнего применения ссылки в место, предшествующее ее перемещению, мы устраним ошибку компиляции:

```
let item = Item { contents: 42 };
let r = &item;
println!("reference to item is {:?}", r);

// Ссылка `r` по-прежнему находится в области видимости,
// но больше не используется, как будто была отброшена
let new_item = item; // Перемещение выполняется без проблем
```

Устранение разногласий с модулем проверки заимствования

У новичков в Rust (да и у опытных специалистов) нередко возникает ощущение, что они много времени тратят на сражения с модулем проверки заимствований. Какие же средства помогут победить в этих сражениях?

Локальный рефакторинг кода

Первая тактика — это внимательное изучение выдаваемых компилятором ошибок, так как создатели Rust постарались сделать их максимально полезными:

НЕ КОМПИЛИРУЕТСЯ

```
/// Если `needle` присутствует в `haystack`,  
/// вернуть содержащий ее срез  
pub fn find<'a, 'b>(haystack: &'a str, needle: &'b str) -> Option<&'a str> {  
    haystack  
        .find(needle)  
        .map(|i| &haystack[i..i + needle.len()])  
}  
// ...  
  
let found = find(&format!("{} to search", "Text"), "ex");  
if let Some(text) = found {  
    println!("Found '{text}'!");  
}
```

```
error[E0716]: temporary value dropped while borrowed  
--> src/main.rs:353:23  
|  
353 |     let found = find(&format!("{} to search", "Text"), "ex");  
|           ^^^^^^^^^^^^^^^^^^^^^^^^^^ - temporary value  
|           |           is freed at the end of this statement  
|           |           creates a temporary value which is freed while still in  
|           |           use  
354 |     if let Some(text) = found {  
|           |           ----- borrow later used here  
|           |  
= note: consider using a `let` binding to create a longer lived value
```

Первая часть сообщения об ошибке самая важная, так как описывает, какое правило заимствования вы, по мнению компилятора, нарушили и почему. Встретив довольно много таких ошибок, вы выработаете чутье в отношении проверки заимствования, которое будет соответствовать более теоретическому аспекту, заключенному в ранее описанных правилах.

Вторая часть сообщения об ошибке включает предложения компилятора по ее исправлению, что в данном случае несложно:

```
let haystack = format!("{} to search", "Text");
let found = find(&haystack, "ex");
if let Some(text) = found {
    println!("Found '{text}'!");
}
// Теперь `found` ссылается на `haystack`, которая живет дольше нее
```

Это пример одного из двух простых приемов в коде, который помогает умаслить анализатор заимствований.

Расширение времени жизни

Преобразование с помощью `let` временной переменной, чей срок жизни продолжается только до конца выражения, в новую локальную переменную с именем, срок жизни которой будет выходить за пределы блока кода.

Сокращение времени жизни

Добавление блока `{ ... }` в окрестности использования ссылки, чтобы ее время жизни заканчивалось в конце него.

Второй вариант менее широко распространен из-за существования нелексического времени жизни: компилятор зачастую может определить, что ссылка перестает использоваться, до момента ее отбрасывания в конце блока. Тем не менее, если вы вдруг увлечетесь регулярным добавлением искусственных блоков кода в окрестности его небольших фрагментов, подумайте о том, не стоит ли заключить этот код в собственный метод.

Предлагаемые компилятором исправления помогают при несложных проблемах, но в случае запутанного кода они наверняка окажутся бесполезными и объяснение нарушенного правила заимствования станет менее понятным:

НЕ КОМПИЛИРУЕТСЯ

```
let x = Some(Rc::new(RefCell::new(Item { contents: 42 })));

// Вызов функции с сигнатурой: `check_item(item: Option<&Item>)`
check_item(x.as_ref().map(|r| r.borrow().deref()));
```

```
error[E0515]: cannot return reference to temporary value
--> src/main.rs:293:35
|
293 |     check_item(x.as_ref().map(|r| r.borrow().deref())));
      |-----^ ^ ^ ^ ^ ^ ^
```

```

|           returns a reference to data owned by the
|           current function
|           temporary value created here
|

```

В этой ситуации может оказаться полезным временное добавление серии локальных переменных — по одной для каждого шага усложняемого преобразования в сопровождении явной аннотации типа:

НЕ КОМПИЛИРУЕТСЯ

```

let x: Option<Rc<RefCell<Item>>> =
    Some(Rc::new(RefCell::new(Item { contents: 42 })));
let x1: Option<&Rc<RefCell<Item>>> = x.as_ref();
let x2: Option<std::cell::Ref<Item>> = x1.map(|r| r.borrow());
let x3: Option<&Item> = x2.map(|r| r.deref());
check_item(x3);

```

```

error[E0515]: cannot return reference to function parameter `r`
--> src/main.rs:305:40
|
305 |     let x3: Option<&Item> = x2.map(|r| r.deref());
|                           ^^^^^^^^^ returns a reference to
|                           data owned by the current function

```

Это сводится к точному преобразованию, на которое жалуется компилятор и которое, в свою очередь, позволяет изменить структуру кода:

```

let x: Option<Rc<RefCell<Item>>> =
    Some(Rc::new(RefCell::new(Item { contents: 42 })));

let x1: Option<&Rc<RefCell<Item>>> = x.as_ref();
let x2: Option<std::cell::Ref<Item>> = x1.map(|r| r.borrow());
match x2 {
    None => check_item(None),
    Some(r) => {
        let x3: &Item = r.deref();
        check_item(Some(x3));
    }
}

```

Разобравшись в возникшей проблеме и устранив ее, вы сможете воссоединить локальные переменные и представить, будто проблем не возникало:

```

let x = Some(Rc::new(RefCell::new(Item { contents: 42 })));

match x.as_ref().map(|r| r.borrow()) {
    None => check_item(None),
    Some(r) => check_item(Some(r.deref())),
};

```

Построение структур данных

Следующая тактика в сражениях против модуля проверки заимствований заключается в проектировании структур данных с учетом этого модуля. Панцеей здесь станет ситуация, когда создаваемые структуры будут владеть всеми используемыми ими данными, избегая применения ссылок и последующего распространения аннотаций времени жизни, описанного в рекомендации 14.

Однако в отношении реальных структур данных это не всегда оказывается возможным. Всякий раз, когда внутренние связи такой структуры формируют граф, который оказывается более взаимосвязанным, чем дерево (где `Root` владеет множеством `Branches`, каждая из которых владеет множеством `Leaf`, и т. д.), простая схема с единым владением становится невозможной.

В качестве примера представим простой реестр информации о гостях, записанной в порядке их прибытия:

```
#[derive(Clone, Debug)]
pub struct Guest {
    name: String,
    address: String,
    // ... множество других полей
}

/// Тип локальной ошибки, используется позже
#[derive(Clone, Debug)]
pub struct Error(String);

/// Реестр гостей, составленный в порядке их прибытия
#[derive(Default, Debug)]
pub struct GuestRegister(Vec<Guest>);

impl GuestRegister {
    pub fn register(&mut self, guest: Guest) {
        self.0.push(guest)
    }
    pub fn nth(&self, idx: usize) -> Option<&Guest> {
        self.0.get(idx)
    }
}
```

Если этому коду нужна *также* возможность эффективного поиска гостей по времени их прибытия и по имени, то это потребует двух фундаментально разных структур данных, лишь одна из которых сможет владеть данными.

Если в процессе участвуют иммутабельные данные небольшого объема, то быстрым решением может стать простое их клонирование:

```
mod cloned {
    use super::Guest;
```

```

#[derive(Default, Debug)]
pub struct GuestRegister {
    by_arrival: Vec<Guest>,
    by_name: std::collections::BTreeMap<String, Guest>,
}

impl GuestRegister {
    pub fn register(&mut self, guest: Guest) {
        // Требует, чтобы `Guest` был `Clone`
        self.by_arrival.push(guest.clone());
        // Для сокращения примера проверка на дубликаты отсутствует
        self.by_name.insert(guest.name.clone(), guest);
    }
    pub fn named(&self, name: &str) -> Option<&Guest> {
        self.by_name.get(name)
    }
    pub fn nth(&self, idx: usize) -> Option<&Guest> {
        self.by_arrival.get(idx)
    }
}
}

```

Но если данные могут быть изменены, подход с клонированием не годится. Например, если адрес `Guest` нужно обновить, вам потребуется находить обе версии и следить за сохранением их согласованности.

Еще один вариант — добавить дополнительный уровень косвенности, когда `Vec<Guest>` начинает рассматриваться как владелец и поиск по имени в этом векторе происходит по индексу:

```

mod indexed {
    use super::Guest;

    #[derive(Default)]
    pub struct GuestRegister {
        by_arrival: Vec<Guest>,
        // Сопоставление имени гостя с индексом в `by_arrival`
        by_name: std::collections::BTreeMap<String, usize>,
    }

    impl GuestRegister {
        pub fn register(&mut self, guest: Guest) {
            // Для сокращения примера проверка на дубликаты отсутствует
            self.by_name
                .insert(guest.name.clone(), self.by_arrival.len());
            self.by_arrival.push(guest);
        }
        pub fn named(&self, name: &str) -> Option<&Guest> {
            let idx = *self.by_name.get(name)?;

```

```
        self.nth(idx)
    }
pub fn named_mut(&mut self, name: &str) -> Option<&mut Guest> {
    let idx = *self.by_name.get(name)?;
    self.nth_mut(idx)
}
pub fn nth(&self, idx: usize) -> Option<&Guest> {
    self.by_arrival.get(idx)
}
pub fn nth_mut(&mut self, idx: usize) -> Option<&mut Guest> {
    self.by_arrival.get_mut(idx)
}
}
```

{}

В этом подходе каждый гость представлен одним элементом `Guest`, что позволяет методу `named_mut()` возвращать мутабельную ссылку на этот элемент. В свою очередь, это подразумевает беспроблемное изменение адреса гостя — `Guest` принадлежит `Vec`, и внутренне обращение к нему всегда будет происходить именно таким образом:

```
let new_address = "123 Bigger House St";
// Реальный код не станет предполагать, что "Bob" существует...
ledger.named_mut("Bob").unwrap().address = new_address.to_string();

assert_eq!(ledger.named("Bob").unwrap().address, new_address);
```

Но если гости могут снимать свою регистрацию, то можно по неосторожности внести ошибку.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// Снятие регистрации `Guest` в позиции `idx`
// с перемещением всех последующих гостей
pub fn deregister(&mut self, idx: usize) -> Result<(), super::Error> {
    if idx >= self.by_arrival.len() {
        return Err(super::Error::new("out of bounds"));
    }
    self.by_arrival.remove(idx);

    // Упс, забыли обновить `by_name`
    OK(())
}
```

Теперь, когда `Vec` можно перемешивать, находящиеся в нем индексы `by_name`, по сути, выступают как указатели и мы вернулись к ситуации, когда может

возникнуть ошибка, при которой они будут указывать в никуда (за пределы `Vec`) или на некорректные данные.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
ledger.register(alice);
ledger.register(bob);
ledger.register(charlie);
println!("Register starts as: {ledger:?}");

ledger.deregister(0).unwrap();
println!("Register after deregister(0): {ledger:?}");

let also_alice = ledger.named("Alice");
// Alice по-прежнему связана с индексом 0, за которым теперь закреплен Bob
println!("Alice is {also_alice:?}");

let also_bob = ledger.named("Bob");
// Bob по-прежнему связан с индексом 1, который теперь принадлежит Charlie
println!("Bob is {also_bob:?}");

let also_charlie = ledger.named("Charlie");
// Charlie по-прежнему закреплен за индексом 2, который теперь
// выходит за пределы Vec
println!("Charlie is {also_charlie:?}");
```

С целью сокращения вывода в этом коде используется персонализированная реализация `Debug` (не показана). Вот итоговый вывод:

```
Register starts as: {
    by_arrival: [{n: 'Alice', ...}, {n: 'Bob', ...}, {n: 'Charlie', ...}]
    by_name: {"Alice": 0, "Bob": 1, "Charlie": 2}
}
Register after deregister(0): {
    by_arrival: [{n: 'Bob', ...}, {n: 'Charlie', ...}]
    by_name: {"Alice": 0, "Bob": 1, "Charlie": 2}
}
Alice is Some(Guest { name: "Bob", address: "234 Bobton" })
Bob is Some(Guest { name: "Charlie", address: "345 Charlieland" })
Charlie is None
```

В предыдущем примере была показана ошибка в коде `deregister`, но даже после ее исправления ничто не мешает вызывающему коду сохранить значение индекса и использовать его с `nth()`, получая неожиданные или недействительные результаты.

Основная проблема заключается в необходимости обеспечения согласованности между этими двумя структурами данных. Более эффективно обработать этот случай можно с помощью умных указателей Rust (см. рекомендацию 8). Переход

к комбинации `Rc` и `RefCell` позволяет избежать проблем с недействительностью, когда индексы используются в качестве псевдоуказателей. Обновив пример — пока с сохранением ошибки, — мы получим следующее:

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
mod rc {
    use super::{Error, Guest};
    use std::cell::RefCell, rc::Rc;
}

#[derive(Default)]
pub struct GuestRegister {
    by_arrival: Vec<Rc<RefCell<Guest>>>,
    by_name: std::collections::BTreeMap<String, Rc<RefCell<Guest>>>,
}

impl GuestRegister {
    pub fn register(&mut self, guest: Guest) {
        let name = guest.name.clone();
        let guest = Rc::new(RefCell::new(guest));
        self.by_arrival.push(guest.clone());
        self.by_name.insert(name, guest);
    }
    pub fn deregister(&mut self, idx: usize) -> Result<(), Error> {
        if idx >= self.by_arrival.len() {
            return Err(Error::new("out of bounds"));
        }
        self.by_arrival.remove(idx);

        // Упс, по-прежнему забыли обновить `by_name`!
        OK(())
    }
    // ...
}
```

```
Bob is Some(RefCell { value: Guest { name: "Bob", address: "234 Bobton" } })
Charlie is Some(RefCell { value: Guest { name: "Charlie",
                                         address: "345 Charlieland" } })
```

В выводе больше нет несовпадающих имен, но зависшая запись про Элис сохраняется до тех пор, пока мы не исправим ошибку, обеспечив сохранение согласованности этих двух коллекций:

```
pub fn deregister(&mut self, idx: usize) -> Result<(), Error> {
    if idx >= self.by_arrival.len() {
        return Err(Error::new("out of bounds"));
    }
    let guest: Rc<RefCell<Guest>> = self.by_arrival.remove(idx);
    self.by_name.remove(&guest.borrow().name);
    Ok(())
}

Register after deregister(0):
by_arrival: [{n: 'Bob', ...}, {n: 'Charlie', ...}]
by_name: [("Bob", {n: 'Bob', ...}), ("Charlie", {n: 'Charlie', ...})]
}
Alice is None
Bob is Some(RefCell { value: Guest { name: "Bob", address: "234 Bobton" } })
Charlie is Some(RefCell { value: Guest { name: "Charlie",
                                         address: "345 Charlieland" } })
```

Умные указатели

Предыдущий раздел закончился примером более общего подхода, заключающегося в *использовании в структурах взаимосвязанных данных умных указателей*.

В рекомендации 8 описывались самые распространенные виды умных указателей, предоставляемые стандартной библиотекой Rust.

- `Rc` позволяет реализовывать общее владение, когда на один элемент ссылаются несколько компонентов. Этот вид указателей часто совмещается с `RefCell`.
- `RefCell` делает возможной внутреннюю мутабельность, когда внутреннее состояние можно изменять, не используя мутабельную ссылку. Но это происходит ценой перемещения проверок заимствования с этапа компиляции в среду выполнения.
- `Arc` — это многопоточный эквивалент `Rc`.
- `Mutex` (и `RwLock`) обеспечивает внутреннюю мутабельность в многопоточной среде, в некотором смысле являясь эквивалентом `RefCell`.
- `Cell` обеспечивает внутреннюю мутабельность для типов `Copy`.

Для программистов, которые переходят на Rust с C++, самым популярным инструментом становится `Rc<T>` (и его потокобезопасный собрат `Arc<T>`), который часто совмещается с `RefCell` (или его потокобезопасной альтернативой `Mutex`).

Наивный перевод общих указателей (или даже `std::shared_ptr`) в экземпляры `Rc<RefCell<T>` обычно дает результат, который работает в Rust без особых претензий со стороны анализатора заимствований.

Однако этот подход подразумевает утрату ряда предоставляемых Rust защит. В частности, ситуация, когда один элемент мутабельно заимствуется (через `borrow_mut()`), притом что существует другая ссылка, ведет к возникновению `panic!` в среде выполнения, а не к ошибке этапа компиляции.

Например, одним из паттернов, нарушающих односторонний поток владения в древоподобных структурах данных, является случай, когда присутствует указатель «владения», направленный назад от элемента на его владельца (рис. 3.3). Такие ссылки `owner` полезны для перемещения по структуре данных. К примеру, процесс добавления в `Leaf` нового сиблинга должен включать владеющую `Branch`.

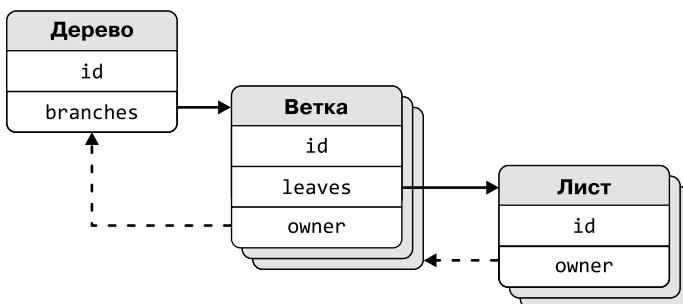


Рис. 3.3. Схема древовидной структуры данных

Реализация этого паттерна в Rust может строиться вокруг более интуитивного шаблона `Rc<T>` — `Weak<T>`:

```

use std::{
    cell::RefCell,
    rc::{Rc, Weak},
};

// Использование "newtype" для каждого типа идентификатора
struct TreeId(String);
struct BranchId(String);
struct LeafId(String);

struct Tree {
    id: TreeId,
    branches: Vec<Rc<RefCell<Branch>>,
}

struct Branch {
    id: BranchId,
    leaves: Vec<Rc<RefCell<Leaf>>,
}
  
```

```

    owner: Option<Weak<RefCell<Tree>>>,
}

struct Leaf {
    id: LeafId,
    owner: Option<Weak<RefCell<Branch>>>,
}

```

Ссылка `Weak` не инкрементирует основное число ссылок, а значит, должна явно проверять, не исчез ли соответствующий элемент:

```

impl Branch {
    fn add_leaf(branch: Rc<RefCell<Branch>>, mut leaf: Leaf) {
        leaf.owner = Some(Rc::downgrade(&branch));
        branch.borrow_mut().leaves.push(Rc::new(RefCell::new(leaf)));
    }
    fn location(&self) -> String {
        match &self.owner {
            None => format!("<unowned>.{})", self.id.0),
            Some(owner) => {
                // Апгрейд слабого указателя владения
                let tree = owner.upgrade().expect("owner gone!");
                format!("{}.{})", tree.borrow().id.0, self.id.0)
            }
        }
    }
}

```

Если умные указатели Rust не позволяют реализовать все необходимое для ваших структур данных, всегда есть возможность прибегнуть к написанию `unsafe`-кода с использованием сырых (и намеренно не умных) указателей. Тем не менее, как говорилось в рекомендации 16, это должно быть последним средством — кто-то другой уже мог реализовать нужную вам семантику внутри безопасного интерфейса, и если вы поищете в стандартной библиотеке или на [crates.io](#), то вполне можете найти подходящие инструменты.

Представьте, что у вас есть функция, которая иногда возвращает ссылку на один из входных элементов, но в некоторых случаях должна вернуть только что аллокированные данные. В соответствии с рекомендацией 1 естественным способом выражения такого поведения в системе типов будет `enum`, которое кодирует эти две возможности, и вы можете реализовать различные трейты указателей, описанные в рекомендации 8. Но вам это делать не обязательно: в стандартной библиотеке уже есть тип `std::borrow::Cow`, который подходит конкретно под описанный сценарий¹.

¹ Cow означает clone-on-write (клонирование при записи) — копия соответствующих данных делается, только если в них нужно внести изменение (произвести запись).

Самореферентные структуры данных

Программисты, начинающие изучать Rust после других языков, часто сталкиваются с проблемой проверки заимствования при попытке создания самореферентных структур данных, которые содержат смесь присвоенных (owned) данных и ссылок, ведущих внутрь этих данных.

НЕ КОМПИЛИРУЕТСЯ

```
struct SelfRef {  
    text: String,  
    // Срез `text`, который содержит текст title  
    title: Option<&str>,  
}
```

На уровне синтаксиса этот код не скомпилируется, поскольку не соответствует правилам времени жизни, описанным в рекомендации 14: ссылке необходима аннотация времени жизни, и это означает, что содержащей ее структуре данных тоже потребуется параметр времени жизни. Но время жизни будет относиться к чему-то внешнему для этой структуры `SelfRef`, что не соответствует замыслу, согласно которому данные по ссылке должны находиться внутри структуры.

Стоит подумать о причине этого ограничения на семантическом уровне. Структуры данных в Rust могут *перемещаться*: из стека в кучу, из кучи в стек и из одного места в другое. Если такое происходит, то внутренний указатель `title` становится недействительным и сохранить его рабочим никак не получится.

Простой альтернативой для этого случая будет использование ранее рассмотренного подхода с индексированием: диапазон смещений в `text` не утрачивает своей валидности в результате перемещения и для анализатора заимствований остается невидимым, так как не задействует ссылки:

```
struct SelfRefIdx {  
    text: String,  
    // Индексы в `text`, где находится текст title  
    title: Option<std::ops::Range<usize>>,  
}
```

Однако подход с индексированием работает только для простых примеров и имеет недостатки, о которых говорилось ранее: сам индекс становится псевдоуказателем, который может утратить связь со своими данными или даже начать указывать на уже не существующие диапазоны `text`.

Более общая версия проблемы самореферентности возникает, когда компилятор обрабатывает `async`-код¹. Грубо говоря, он объединяет кучу ожидающего `async` кода в замыкание, которое содержит как сам этот код, так и все захваченные им части среды, с которыми он работает (как описывалось в рекомендации 2). Эта захваченная часть среды может включать как значения, так и ссылки на них. По своей сути такая структура самореферентна, в связи с чем поддержка `async` стала основной причиной добавления в стандартную библиотеку типа `Pin`. Этот тип указателя прикрепляет свое значение к месту, вынуждая его оставаться в одной области памяти и обеспечивая тем самым сохранение валидности внутренних самореферентных связей.

Итак, для самореферентных типов можно использовать `Pin`, но сделать это правильно нелегко, поэтому рекомендуем почитать официальную документацию.

По возможности *избегайте применения самореферентных структур данных* или старайтесь найти библиотечные крейты, которые упрощают эти сложности, например `ouroborous`.

Запомните

- Ссылки в Rust являются *заимствуемыми*, а значит, не могут удерживаться бесконечно.
- Анализатор заимствований позволяет использовать множество иммутабельных ссылок на элемент или одну мутабельную, но не то и другое одновременно. Благодаря концепции нелексического времени жизни существование ссылки обрывается в точке ее последнего применения, а не в конце окружающего ее блока кода.
- Ошибки, выдаваемые анализатором заимствований, можно исключить разными способами:
 - добавлением области `{ ... }` для сокращения продолжительности времени жизни значения;
 - присваиванием значения именованной локальной переменной для расширения его жизни до конца области видимости;
 - временным добавлением нескольких локальных переменных для точной локализации предмета недовольства анализатора.
- Умные указатели в Rust помогают обойти правила проверки заимствования, в связи с чем пригождаются для структур взаимосвязанных данных.
- Тем не менее работать в Rust с самореферентными структурами по-прежнему не очень удобно.

¹ Работа с `async`-кодом выходит за рамки тематики этой книги. Чтобы лучше понять его потребность в самореферентных структурах данных, прочитайте главу 8 книги Йона Гьенсета (Jon Gjengset) *Rust for Rustaceans* (No Starch Press).

Рекомендация 16. Страйтесь не писать unsafe-код

Гарантии безопасности памяти — без издержек для среды выполнения — являются уникальным маркетинговым преимуществом. Именно этой функциональности нет ни в одном из популярных языков программирования. Но эти гарантии имеют свою цену — написание кода Rust требует реорганизовывать его в соответствии с требованиями анализатора заимствований (см. рекомендацию 15) и точно указывать типы используемых ссылок (см. рекомендацию 8).

Unsafe-код Rust — это надмножество языка, в котором ослаблены некоторые из этих ограничений, а с ними и соответствующие гарантии. Предваряя блок кода ключевым словом `unsafe`, мы переключаем его в небезопасный режим, который позволяет делать то, что в стандартном Rust недопустимо. В частности, это дает возможность использовать *сырые указатели*, которые работают подобно старым добрым указателям C. На эти указатели не накладываются правила заимствования, и здесь уже программист должен следить за тем, чтобы в случае разыменования они продолжали указывать на валидную область памяти.

Так что на внешнем уровне рекомендация текущего раздела будет тривиальной: зачем переходить на Rust, если вы планируете писать на нем код в стиле C? Тем не менее бывает, что `unsafe`-код просто необходим: для низкоуровневого кода библиотеки или когда вашему коду Rust нужно взаимодействовать с кодом других языков (см. рекомендацию 34).

Формулировка текущей рекомендации достаточно точна: *страйтесь не писать небезопасный код*. И акцент здесь делается на написании, потому что `unsafe`-код, который вам может понадобиться, наверняка уже написан.

Стандартные библиотеки Rust содержат много `unsafe`-кода. Короткий поиск показывает около 1000 случаев использования `unsafe` в библиотеке `alloc`, 1500 — в `core` и 2000 — в `std`. Этот код написан экспертами и прошел проверку боем в тысячах баз кода Rust.

Часть этого `unsafe`-кода реализуется в функциональности стандартной библиотеки, которую мы уже разобрали.

- В типах умных указателей — `Rc`, `RefCell`, `Arc` и пр., описанных в рекомендации 8, применяют `unsafe`-код (зачастую сырье указатели), чтобы иметь возможность предоставлять свою семантику пользователям.
- Примитивы синхронизации — `Mutex`, `RwLock` и связанные с ними защиты (см. рекомендацию 17) задействуют специфичный для ОС `unsafe`-код. Если вы хотите разобраться в нюансах работы с этими примитивами, рекомендую почитать книгу Мары Боса (Mara Bos) *Rust Atomics and Locks* (O'Reilly).

В стандартной библиотеке есть и другая функциональность, которая охватывает более продвинутые возможности, реализованные с помощью `unsafe`-кода¹.

- `std::pin::Pin` не дает элементу менять расположение в памяти (см. рекомендацию 15). Позволяет создавать самореферентные структуры данных, которые зачастую вызывают сложности (<https://oreil.ly/JBnWU>) у новичков в Rust.
- `std::borrow::Cow` предоставляет умный указатель с клонированием при записи — один и тот же указатель можно использовать для чтения/записи, когда клонирование связанных с ним данных происходит только в случае записи.
- Различные функции (`take`, `swap`, `replace`) в `std::mem` позволяют манипулировать элементами в памяти, не подвергаясь нападкам со стороны анализатора заимствований.

Для корректного использования всех этих возможностей по-прежнему требуется осторожность, но `unsafe`-код был инкапсулирован в них таким образом, чтобы исключить целые классы проблем.

Если выйти за пределы стандартной библиотеки, то экосистема `crates.io` содержит множество крейтов, включающих `unsafe`-код для предоставления востребованных возможностей.

`once_cell`

Позволяет использовать такие элементы, как глобальные переменные, инициализируемые только раз.

`rand`

Позволяет генерировать случайные числа, задействуя низкоуровневую функциональность операционной системы и процессора.

`byteorder`

Позволяет преобразовывать сырье байты данных в числа и наоборот.

`cxx`

Делает возможным взаимодействие кода C++ и Rust (также упоминается в рекомендации 35).

Есть и много других примеров, но мы надеемся, что общая идея вам ясна. Если вы хотите проделывать нечто, что явно не вписывается в ограничения Rust,

¹ На практике большая часть этой функциональности `std` предоставляется `core`, в связи с чем доступна для `no_std`-кода, как описано в рекомендации 33.

особенно рассмотренные в рекомендациях 14 и 15, поищите в стандартной библиотеке функциональность, которая может отвечать вашему запросу. Если найти подходящее решение не удастся, обратитесь к базе `crates.io`. Все же редко встречается настолько уникальная задача, чтобы никто ее еще не реализовывал.

Естественно, всегда будут случаи, когда без `unsafe` не обойтись, — например, когда вам нужно наладить взаимодействие с кодом, написанным на других языках, через интерфейс внешних функций (foreign function interface, FFI), о чем пойдет речь в рекомендации 34. Но, когда возникает такая потребность, *подумайте о написании оберточного слоя, который будет содержать весь unsafe-код*, необходимый для того, чтобы другие программисты могли последовать совету, данному в текущей рекомендации. Это позволит также локализовать возможные проблемы: когда что-то идет не так, `unsafe`-обертка может стать первым подозреваемым.

Кроме того, если вы вынуждены писать `unsafe`-код, имейте в виду ограничение, подразумеваемое самим этим ключевым словом: *hic sunt dracones* («тут обитают драконы»).

- Добавляйте комментарии безопасности (<https://oreil.ly/MHQvh>), документирующие предварительные условия и инварианты, на которые опирается `unsafe`-код. В Clipper (см. рекомендацию 29) есть предупреждение, которое об этом напоминает.
- Минимизируйте количество кода, содержащегося в блоке `unsafe`, чтобы ограничить возможный радиус действия ошибки. Подумайте о включении линта `unsafe_op_in_unsafe_fn`, чтобы при выполнении `unsafe`-операций требовалось использовать явные блоки `unsafe`, даже когда эти операции выполняются в функции, которая сама является `unsafe`.
- Пишите больше тестов (см. рекомендацию 30), чем обычно.
- Проверяйте код дополнительными инструментами диагностики (см. рекомендацию 31). В частности, *рекомендую анализировать unsafe-код с помощью Miri* — Miri интерпретирует промежуточный вывод компилятора, который позволяет этому инструменту обнаруживать классы ошибок, невидимые для компилятора Rust.
- Тщательно продумывайте применение многопоточности, особенно если в коде есть совместно используемое состояние (см. рекомендацию 17).

Добавление маркера `unsafe` не означает, что все правила отменяются, — это значит лишь то, что теперь за сохранение гарантий безопасности Rust отвечает *вы*, программист, а не компилятор.

Рекомендация 17. Будьте осторожны при параллельном доступе к общему состоянию

В Rust даже самые смелые формы совместного использования ресурсов гарантированно безопасны.

*Аарон Туран (Aaron Turon)
(<https://oreil.ly/wKFxX>)*

В официальной документации языка Rust описывается как допускающий смелое применение многопоточности (<https://oreil.ly/R7eq9>), но в текущей рекомендации мы разберем, почему, к сожалению, даже в нем все же остаются причины ее бояться.

Эта рекомендация посвящена конкретно многопоточной обработке при использовании *общего состояния* — ситуации, когда разные потоки выполнения взаимодействуют друг с другом, используя общую память. Совместное использование памяти потоками в любом языке обычно сопровождают две серьезные проблемы.

Гонка данных (data race)

Может вызвать повреждение данных.

Взаимная блокировка (deadlock)

Может привести к зависанию программы.

Обе эти проблемы ужасны — они вызывают или склонны вызывать ошибки, потому что на практике их бывает очень трудно отлаживать: сбои происходят несистематически и зачастую именно под нагрузкой, то есть в модульных, интеграционных и прочих тестах не проявляются (см. рекомендацию 30), зато проявляются на производстве.

Rust стал значительным прорывом, поскольку полностью решает одну из этих двух проблем, но вторая, как вы далее увидите, все же сохраняется.

Гонка данных

Начнем с хорошего — с гонки данных в Rust, а именно ее отсутствия. Точное определение этого состояния от языка к языку различается, но в целом его основные компоненты можно описать так. Гонка данных происходит, когда два потока пытаются получить доступ к одной области памяти при следующих условиях:

- как минимум один из них производит запись;
- отсутствует механизм синхронизации, который упорядочивал бы операции доступа.

Гонки данных в C++

Суть этого понятия лучше всего продемонстрировать на примере. Рассмотрим структуру данных, отслеживающую состояние банковского счета.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// Код C++
class BankAccount {
public:
    BankAccount() : balance_(0) {}

    int64_t balance() const {
        if (balance_ < 0) {
            std::cerr << "** Упс, произошел перерасход: " << balance_ << "! **\n";
            std::abort();
        }
        return balance_;
    }
    void deposit(uint32_t amount) {
        balance_ += amount;
    }
    bool withdraw(uint32_t amount) {
        if (balance_ < amount) {
            return false;
        }
        // А если в этот момент `balance_` изменит другой поток?
        std::this_thread::sleep_for(std::chrono::milliseconds(500));

        balance_ -= amount;
        return true;
    }

private:
    int64_t balance_;
};
```

Этот пример специально приведен на C++, а не на Rust, и скоро станет понятно почему. Тот же общий принцип применим во многих других языках (не Rust) – Java, Go, Python и пр.

Показанный ранее класс отлично работает в однопоточном сеттинге, но давайте взглянем на многопоточный:

```
BankAccount account;
account.deposit(1000);

// Запуск потока, который отслеживает опустошение счета и пополняет его
std::thread payer(pay_in, &account);
```

```
// Запуск трех потоков, каждый из которых циклически пытается
// снять со счета деньги
std::thread taker(take_out, &account);
std::thread taker2(take_out, &account);
std::thread taker3(take_out, &account);
```

Здесь несколько потоков циклически пытаются снять средства со счета, и есть дополнительный поток, который пополняет его при достижении балансом нижнего порога:

```
// Непрерывно мониторит баланс `account` и пополняет его в случае снижения
void pay_in(BankAccount* account) {
    while (true) {
        if (account->balance() < 200) {
            log("[A] Balance running low, deposit 400");
            account->deposit(400);
        }
        // (Бесконечный цикл с приостановками приводится
        // исключительно в целях демонстрации)
        std::this_thread::sleep_for(std::chrono::milliseconds(5));
    }
}

// Повторяющаяся попытка снять средства с `account`
void take_out(BankAccount* account) {
    while (true) {
        if (account->withdraw(100)) {
            log("[B] Withdrew 100, balance now " +
                std::to_string(account->balance()));
        } else {
            log("[B] Failed to withdraw 100");
        }
        std::this_thread::sleep_for(std::chrono::milliseconds(20));
    }
}
```

Рано или поздно в этой схеме произойдет промах:

```
** Oh no, gone overdrawn: -100! **
```

Проблему обнаружить несложно, в частности, благодаря комментарию к методу `withdraw()`: когда в процессе участвуют несколько потоков, между проверками может происходить несколько изменений баланса. Но в реальных сценариях подобные ошибки обнаружить намного сложнее, особенно если компилятору разрешено выполнять различные оптимизации и переупорядочение (как в случае с C++).

Здесь вызовы `sleep` добавлены для того, чтобы искусственно повысить вероятность возникновения ошибки и ее перехвата на раннем этапе. Подобные

проблемы в реальной рабочей среде обычно проявляются редко и бессистемно, что сильно затрудняет их отладку.

Класс `BankAccount` является *потокосовместимым*, то есть его можно применять в многопоточной среде при условии, что пользователи обеспечат некий механизм синхронизации для регулирования доступа к нему.

Этот класс можно преобразовать в *потокобезопасный*, добавив внутренние операции синхронизации, и тогда его смогут спокойно использовать несколько потоков¹:

```
// Код C++
class BankAccount {
public:
    BankAccount() : balance_(0) {}

    int64_t balance() const {
        // Блокировка mu_ на протяжении всей этой области видимости
        const std::lock_guard<std::mutex> with_lock(mu_);
        if (balance_ < 0) {
            std::cerr << "*** Oh no, gone overdrawn: " << balance_ << " **!\n";
            std::abort();
        }
        return balance_;
    }

    void deposit(uint32_t amount) {
        const std::lock_guard<std::mutex> with_lock(mu_);
        balance_ += amount;
    }

    bool withdraw(uint32_t amount) {
        const std::lock_guard<std::mutex> with_lock(mu_);
        if (balance_ < amount) {
            return false;
        }
        balance_ -= amount;
        return true;
    }

private:
    mutable std::mutex mu_; // Защищает balance_
    int64_t balance_;
};
```

Теперь внутреннее поле `balance_` защищено *мьютексом* `mu_` — синхронизирующим объектом, который обеспечивает, чтобы этот мьютекс одновременно мог удерживаться только одним потоком. Вызывающий код получает мьютекс

¹ Третья категория поведения — это потокоопасный код, который несет угрозу многопоточной среде, даже если все операции доступа к нему внешне синхронизированы.

через вызов `std::mutex::lock()`, далее второй и последующие вызывающие `std::mutex::lock()` будут приостанавливаться до тех пор, пока изначальный не вызовет `std::mutex::unlock()`, после чего *один* из приостановленных потоков активируется и получит запрошенный им `std::mutex::lock()`.

Теперь все операции доступа к балансу протекают через получение мьютекса, в связи с чем его значение между изменением и проверкой остается согласованным. Обратите также внимание на `std::lock_guard` — это класс RAII (см. рекомендацию 11), который вызывает `lock()` при создании и `unlock()` при уничтожении. Этот механизм обеспечивает, чтобы при выходе из области видимости мьютекс освобождался, снижая вероятность внесения ошибки при согласовании ручных вызовов `lock()` и `unlock()`.

Тем не менее безопасность потоков здесь по-прежнему под угрозой. Для ее нарушения достаточно одного опрометчивого изменения класса:

```
// Добавили метод C...
void pay_interest(int32_t percent) {
    // ...но забыли о mu_
    int64_t interest = (balance_ * percent) / 100;
    balance_ += interest;
}
```

В результате чего безопасность потоков оказалась утрачена¹.

Гонки данных в Rust

Для книги по Rust в этой рекомендации было приведено довольно много материала из C++, поэтому разберем простой перевод этого класса на Rust:

```
pub struct BankAccount {
    balance: i64,
}

impl BankAccount {
    pub fn new() -> Self {
        BankAccount { balance: 0 }
    }
    pub fn balance(&self) -> i64 {
        if self.balance < 0 {
```

¹ Компилятор Clang из C++ содержит опцию `-Wthread-safety`, иногда называемую аннотацией (от комбинации аннотации с анализом), которая позволяет аннотировать данные информацией о том, какие мьютексы и какую их часть защищают, а функции — информацией о получаемых ими блокировках. В результате, когда эти инварианты нарушаются, в C++, как и в Rust, возникают ошибки этапа компиляции. Однако здесь ничто не заставляет использовать эти аннотации, например, когда библиотека с поддержкой потоков применяется в многопоточной среде впервые.

```

        panic!("** Oh no, gone overdrawn: {}", self.balance);
    }
    self.balance
}
pub fn deposit(&mut self, amount: i64) {
    self.balance += amount
}
pub fn withdraw(&mut self, amount: i64) -> bool {
    if self.balance < amount {
        return false;
    }
    self.balance -= amount;
    true
}
}

```

вместе с функциями, которые бесконечно пытаются внести или снять деньги со счета:

```

pub fn pay_in(account: &mut BankAccount) {
    loop {
        if account.balance() < 200 {
            println!("[A] Running low, deposit 400");
            account.deposit(400);
        }
        std::thread::sleep(std::time::Duration::from_millis(5));
    }
}

pub fn take_out(account: &mut BankAccount) {
    loop {
        if account.withdraw(100) {
            println!("[B] Withdrew 100, balance now {}", account.balance());
        } else {
            println!("[B] Failed to withdraw 100");
        }
        std::thread::sleep(std::time::Duration::from_millis(20));
    }
}

```

В однопоточном контексте этот код работает прекрасно, даже если его поток не является основным:

```

{
    let mut account = BankAccount::new();
    let _payer = std::thread::spawn(move || pay_in(&mut account));
    // В конце области видимости поток `_payer` отсоединяется
    // и оказывается единственным владельцем `BankAccount`
}

```

Но наивная попытка использовать `BankAccount` множеством потоков:

```

НЕ КОМПИЛИРУЕТСЯ

{
    let mut account = BankAccount::new();
    let _taker = std::thread::spawn(move || take_out(&mut account));
    let _payer = std::thread::spawn(move || pay_in(&mut account));
}
```

тут же вызывает ошибку на этапе компиляции:

```
error[E0382]: use of moved value: `account`
--> src/main.rs:102:41
|
100 |     let mut account = BankAccount::new();
|           ----- move occurs because `account` has type
|           `broken::BankAccount`, which does not implement the
|           `Copy` trait
101 |     let _taker = std::thread::spawn(move || take_out(&mut account));
|           ----- variable
|           |
|           |
|           value moved into closure here
102 |     let _payer = std::thread::spawn(move || pay_in(&mut account));
|           ^^^^^^ ----- use occurs due
|           |           to use in closure
|           |
|           value used here after move
```

Правила анализатора заимствований (см. рекомендацию 15) отчетливо проясняют проблему: здесь у нас две мутабельные ссылки на один элемент, то есть на одну больше, чем допустимо. Согласно правилам, на элемент могут указывать либо одна мутабельная ссылка, либо несколько иммутабельных, но не то и другое одновременно.

И это интересно перекликается с состоянием гонки данных, описанным в начале рекомендации: требование, чтобы присутствовал либо один пишущий, либо несколько читающих, но не то и другое, позволит избежать состояния гонки. Обеспечивая безопасность памяти, Rust бонусом получает безопасность потоков (<https://oreil.ly/wKFxX>).

Как и в C++, чтобы сделать эту `struct` потокобезопасной, потребуется определенная синхронизация. Наиболее распространенный механизм для этого также называется `Mutex`, но его версия в Rust обертывает защищаемые данные, а не выступает отдельным объектом, как в C++:

```
pub struct BankAccount {
    balance: std::sync::Mutex<i64>,
}
```

Метод `lock()` в этом дженерике `Mutex`, подобно `std::lock_guard` в C++, возвращает объект `MutexGuard` с поведением RAII: мьютекс автоматически освобождается в конце области действия, когда защита отбрасывается (`drop`). (В отличие от C++ в Rust `Mutex` не содержит методов, требующих ручного освобождения мьютекса, так как они создали бы риск, при котором разработчик может не уследить за сохранением парности этих вызовов.)

Говоря точнее, `lock()` фактически возвращает `Result`, который содержит `MutexGuard`, чтобы исключить возможность *отравления Mutex*. Отравление происходит, если в удерживающем блокировку потоке возникает сбой, так как это может означать, что ни на какие защищенные мьютексом инварианты больше опираться нельзя. На практике отправление блокировки происходит достаточно редко для того, чтобы просто применять `.unwrap()` к `Result` (вопреки совету из рекомендации 18). И желательно, чтобы программа в этом случае завершалась.

За счет реализации трейтов `Deref` и `DerefMut` (см. рекомендацию 8) объект `MutexGuard` выступает также в роли посредника для данных, охватываемых `Mutex`, что позволяет использовать его как для чтения:

```
impl BankAccount {
    pub fn balance(&self) -> i64 {
        let balance = *self.balance.lock().unwrap();
        if balance < 0 {
            panic!("** Oh no, gone overdrawn: {}", balance);
        }
        balance
    }
}
```

так и для записи:

```
impl BankAccount {
    // Примечание: больше не требует `&mut self`
    pub fn deposit(&self, amount: i64) {
        *self.balance.lock().unwrap() += amount
    }
    pub fn withdraw(&self, amount: i64) -> bool {
        let mut balance = self.balance.lock().unwrap();
        if *balance < amount {
            return false;
        }
        *balance -= amount;
        true
    }
}
```

В сигнтурах этих методов кроется интересная деталь: несмотря на то что они изменяют баланс `BankAccount`, они теперь получают `&self`, а не `&mut self`. И это неизбежно: если потоки будут удерживать ссылки на один и тот же `BankAccount`, то в соответствии с правилами анализатора заимствований этим ссылкам лучше не быть мутабельными. Кроме того, это еще один случай паттерна *внутренней мутабельности*, описанного в рекомендации 8: проверки заимствований, по сути, переносятся с этапа компиляции в среду выполнения, но теперь попутно с синхронизацией потоков. Если мутабельная ссылка уже существует, то попытка получить вторую станет блокироваться до тех пор, пока первая не будет уничтожена.

Обертывание совместно используемого состояния в `Mutex` успокаивает анализатор заимствований, но еще нужно исправить некоторые проблемы с временем жизни (см. рекомендацию 14):

НЕ КОМПИЛИРУЕТСЯ

```
{
    let account = BankAccount::new();
    let taker = std::thread::spawn(|| take_out(&account));
    let payer = std::thread::spawn(|| pay_in(&account));
    // В конце области видимости `account` отбрасывается,
    // но потоки `_taker` и `_payer` отсоединяются и продолжают
    // содержать иммутабельные ссылки на `account`
}
```

```
error[E0373]: closure may outlive the current function, but it borrows `account`
    which is owned by the current function
--> src/main.rs:206:40
|
206 |     let taker = std::thread::spawn(|| take_out(&account));
      |             ^^^          ----- `account` is
      |             |           borrowed here
      |
      |             may outlive borrowed value `account`

note: function requires argument type to outlive ``static``
--> src/main.rs:206:21
|
206 |     let taker = std::thread::spawn(|| take_out(&account));
      |             ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^

help: to force the closure to take ownership of `account` (and any other
referenced variables), use the `move` keyword
|
```

```
206 |     let taker = std::thread::spawn(move || take_out(&account));
|           +++
error[E0373]: closure may outlive the current function, but it borrows `account`  

|           which is owned by the current function  

--> src/main.rs:207:40
|  
207 |     let payer = std::thread::spawn(|| pay_in(&account));
|           ^          ----- `account` is  

|           |          borrowed here  

|           |  
|           may outlive borrowed value `account`  

|  
note: function requires argument type to outlive ``static``  

--> src/main.rs:207:21
|  
207 |     let payer = std::thread::spawn(|| pay_in(&account));
|           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^  
help: to force the closure to take ownership of `account` (and any other  
referenced variables), use the `move` keyword  

|  
207 |     let payer = std::thread::spawn(move || pay_in(&account));
|           +++
```

Сообщение об ошибке проясняет проблему: `BankAccount` в конце блока будет отброшен (`drop`), но есть два потока, которые содержат ссылку на него и могут продолжить выполнение после этого. (Предложенное компилятором решение данной проблемы не особо помогает — если элемент `BankAccount` переместить в первое замыкание, он станет недоступным для второго замыкания, которому нужно получить на него ссылку.)

Стандартный инструмент, позволяющий обеспечить сохранение активности объекта, пока на него существуют ссылки, — это указатель с подсчетом ссылок, и в Rust его многопоточным вариантом выступает `std::sync::Arc`:

```
let account = std::sync::Arc::new(BankAccount::new());
account.deposit(1000);

let account2 = account.clone();
let _taker = std::thread::spawn(move || take_out(&account2));

let account3 = account.clone();
let _payer = std::thread::spawn(move || pay_in(&account3));
```

Каждый поток получает собственную копию указателя с подсчетом ссылок, который помещается в замыкание, и соответствующий `BankAccount` будет отброшен (`drop`), только когда количество ссылок упадет до нуля. Эта комбинация

`Arc<Mutex<T>>` часто применяется в программах Rust, использующих многопоточную обработку общего состояния.

Если отвлечься от технических деталей, то можно увидеть, что Rust полностью избежал проблемы с гонкой данных, которая создает много сложностей с многопоточным программированием в других языках. Естественно, это касается только *безопасного* кода Rust — `unsafe`-код (см. рекомендацию 16) и границы FFI в частности (см. рекомендацию 14) не всегда могут избежать гонки данных. Тем не менее эта особенность является большим достоинством.

Стандартные трейты-маркеры

Существует два стандартных трейта, влияющих на использование объектов Rust между потоками. Оба они являются *трейтами-маркерами* (см. рекомендацию 10), которые не имеют никаких методов, зато особо значимы для компилятора в многопоточных сценариях.

- Трейт `Send` указывает, что элементы некоего типа можно безопасно передавать между потоками. Владение элементом этого типа можно передавать из одного потока в другой.
- Трейт `Sync` указывает, что к элементам некоего типа можно безопасно обращаться из нескольких потоков в соответствии с правилами заимствования.

Иными словами, `Send` означает, что между потоками можно передавать `T`, а `Sync` — что между ними можно передавать `&T`.

Оба этих трейта являются также *автоматретами* (<https://oreil.ly/7eeE7>): компилятор автоматически выводит их для новых типов при условии, что составляющие тип части тоже реализуют `Send`/`Sync`.

`Send` и `Sync` реализуются очень многими безопасными типами — их так много, что проще выяснить, какие типы их *не реализуют* (записываются эти трейты в форме `impl !Sync for Type`).

Тип, который не реализует `Send`, можно использовать только в одиночном потоке. Каноническим примером является асинхронный указатель с подсчетом ссылок `Rc<T>` (см. рекомендацию 8). Реализация этого типа явно предполагает однопоточное использование (ради скорости). Здесь отсутствует необходимость синхронизировать внутренний подсчет ссылок, как при многопоточном применении. В связи с этим передача `Rc<T>` между потоками не допускается. Задействуйте для этого случая `Arc<T>`, предлагающий дополнительные издержки, связанные с синхронизацией.

Тип, который не реализует `Sync`, небезопасно использовать из нескольких потоков посредством `ne-mut`-ссылок, поскольку анализатор заимствований будет

следить, чтобы никогда не возникало нескольких `mut`-ссылок. Традиционными примерами этого являются типы, которые обеспечивают *внутреннюю мутабельность* асинхронным способом, такие как `Cell<T>` и `RefCell<T>`. Для получения мутабельности в многопоточной среде используйте `Mutex<T>` или `RwLock<T>`.

Типы сырых указателей вроде `*const T` и `*mut T` также не реализуют ни `Send`, ни `Sync`. Подробнее — в рекомендациях 16 и 24.

Взаимные блокировки

А теперь о плохом. Несмотря на то что Rust решил проблему с гонкой данных, о чём говорилось ранее, в нем сохранилась *вторая* ужасная проблема многопоточного кода с общим состоянием — *взаимные блокировки*.

Рассмотрим упрощенный сервер мультиплерной игры, реализованный в виде многопоточного приложения для одновременного обслуживания большого числа игроков. Две его основные структуры данных будут выступать набором игроков, индексируемых по имени пользователя, и набором активных игровых партий, индексируемых по уникальному идентификатору:

```
struct GameServer {
    // Сопоставление имени игрока с информацией о нем
    players: Mutex<HashMap<String, Player>>,
    // Текущие игры, индексируемые по уникальному игровому ID
    games: Mutex<HashMap<GameId, Game>>,
}
```

Обе эти структуры данных защищены `Mutex`, а значит, гонкам данных не подвержены. Однако в коде, который управляет ими, могут возникнуть другие проблемы. Единичное взаимодействие между ними может работать гладко:

```
impl GameServer {
    // Добавление игрока для присоединения к текущей игре
    fn add_and_join(&self, username: &str, info: Player) -> Option<GameId> {
        // Добавление игрока
        let mut players = self.players.lock().unwrap();
        players.insert(username.to_owned(), info);

        // Поиск игры с доступным местом для нового игрока
        let mut games = self.games.lock().unwrap();
        for (id, game) in games.iter_mut() {
            if game.add_player(username) {
                return Some(id.clone());
            }
        }
    }
    None
}
```

При повторном взаимодействии между этими двумя независимо блокируемыми структурами данных возникают проблемы:

```
impl GameServer {
    // Бан игрока `username` и удаление его из всех текущих игр
    fn ban_player(&self, username: &str) {
        // Поиск всех игр, где присутствует этот пользователь,
        // и его удаление из них
        let mut games = self.games.lock().unwrap();
        games
            .iter_mut()
            .filter(|(_id, g)| g.has_player(username))
            .for_each(|(_id, g)| g.remove_player(username));

        // Удаление пользователя из общего списка
        let mut players = self.players.lock().unwrap();
        players.remove(username);
    }
}
```

Чтобы понять проблему, представьте два отдельных потока, использующих эти два метода, когда их выполнение происходит в порядке, показанном в табл. 3.1.

Таблица 3.1. Последовательность взаимной блокировки потоков

Поток 1	Поток 2
Входит в <code>add_and_join()</code> и сразу получает блокировку <code>players</code>	
	Входит в <code>ban_player()</code> и сразу получает блокировку <code>games</code>
Пытается получить блокировку <code>games</code> , которая удерживается потоком 2, в связи с чем поток 1 входит в ожидание	
	Пытается получить блокировку <code>players</code> , которая удерживается потоком 1, поэтому поток 2 входит в ожидание

В этот момент программа попадает в состояние *взаимной блокировки*: ни один из рассматриваемых потоков не может продолжать выполнение, как и никакой другой поток, взаимодействующий с любой из защищенных `Mutex` структур данных.

Основной причиной этого является *инверсия блокировки*: одна функция получает блокировки в порядке `players`, затем `games`, в то время как другая делает это в противоположном порядке — `games`, затем `players`.

Это простой пример более общей проблемы. Аналогичная ситуация может возникнуть в более длинных цепочках вложенных блокировок (поток 1 получает блокировку A, затем B, после чего пытается получить C, при этом поток 2 получает C, а затем пытается получить A) и между большим числом потоков (поток 1 получает A, потом B, поток 2 получает B, потом C, поток 3 получает C, потом A).

Простейшей попыткой решить эту проблему будет сокращение области действия блокировок, чтобы исключить момент, когда обе блокировки удерживаются одновременно:

```
/// Добавление нового игрока и его присоединение к текущей игре
fn add_and_join(&self, username: &str, info: Player) -> Option<GameId> {
    // Добавление нового игрока
    {
        let mut players = self.players.lock().unwrap();
        players.insert(username.to_owned(), info);
    }

    // Поиск игры с доступным местом для присоединения нового игрока
    {
        let mut games = self.games.lock().unwrap();
        for (id, game) in games.iter_mut() {
            if game.add_player(username) {
                return Some(id.clone());
            }
        }
    }
    None
}

/// Бан игрока `username` и его удаление из всех текущих игр
fn ban_player(&self, username: &str) {
    // Поиск всех игр, где присутствует этот пользователь,
    // и его удаление из них
    {
        let mut games = self.games.lock().unwrap();
        games
            .iter_mut()
            .filter(|(_id, g)| g.has_player(username))
            .for_each(|(_id, g)| g.remove_player(username));
    }

    // Удаление пользователя из общего списка
    {
        let mut players = self.players.lock().unwrap();
        players.remove(username);
    }
}
```

(Более удачной версией этого решения станет заключение управления структурой `players` во вспомогательные методы `add_player()` и `remove_player()`). Это позволит снизить вероятность того, что разработчик забудет закрыть область действия.)

Предложенный подход решает проблему взаимной блокировки, но оставляет неизменной проблему согласованности данных: учитывая последовательность выполнения, показанную в табл. 3.2, структуры `players` и `games` могут рассинхронизироваться.

Таблица 3.2. Последовательность, ведущая к несогласованности состояний

Поток 1	Поток 2
Входит в <code>add_and_join("Alice")</code> и добавляет Alice в структуру данных <code>players</code> , затем освобождает блокировку <code>players</code>	
	Входит в <code>ban_player("Alice")</code> и удаляет Alice из всех игр, затем освобождает блокировку <code>games</code>
	Удаляет Alice из структуры данных <code>players</code> . Поток 1 уже освободил блокировку, значит, в ожидание не входит
Продолжает выполнение и получает блокировку <code>games</code> , уже освобожденную потоком 2. Удерживая ее, добавляет Alice в активную игру	

В этот момент есть игра, где присутствует игрок, которого, если верить структуре данных `players`, не существует.

Суть проблемы в том, что есть две структуры данных, между которыми необходимо поддерживать согласованность. И лучше всего для этого использовать синхронизирующий примитив, включающий обе:

```
struct GameState {
    players: HashMap<String, Player>,
    games: HashMap<GameId, Game>,
}

struct GameServer {
    state: Mutex<GameState>,
    // ...
}
```

Совет

Самым очевидным советом, который поможет избежать проблем, возникающих при многопоточной обработке с общим состоянием, будет простое избегание такой обработки. В книге по Rust приводится цитата из документации к языку Go (<https://oreil.ly/HiKmp>): «Не взаимодействуйте через обмен памятью, наоборот, обменивайтесь памятью через взаимодействие».

В языке Go есть *каналы*, которые подходят для реализации этого приема (<https://oreil.ly/uBPhZ>). В Rust аналогичная функциональность присутствует в модуле `std::sync::mpsc` стандартной библиотеки: функция `channel()` возвращает пару (`Sender`, `Receiver`), которая позволяет передавать между потоками значения конкретного типа.

Если многопоточности с использованием общей памяти не избежать, можно воспользоваться одним из способов снизить вероятность написания уязвимого для взаимных блокировок кода.

- *Помещать структуры данных, которые требуют взаимной согласованности, под одну блокировку.*
- *Делать области действия блокировок короткими и очевидными.* По возможности использовать вспомогательные методы, которые получают код и размещают его под соответствующей блокировкой.
- *Избегать вызова замыканий, удерживающих блокировки.* Иначе надежность кода станет зависеть от замыкания, которое могут добавить в будущем.
- *Избегать возвращения вызывающему коду MutexGuard* — с позиции взаимной блокировки это подобно передаче заряженного пистолета.
- *Добавить в свою систему непрерывной интеграции инструменты для обнаружения взаимных блокировок* (см. рекомендацию 32), такие как `no_deadlocks`, `ThreadSanitizer` или `parking_lot::deadlock`.
- В качестве крайней меры проектируйте, документируйте, тестируйте и организуйте *иерархию блокировки*, которая будет описывать допустимый/требуемый порядок получения блокировок. Это решение должно становиться последним средством, поскольку любая стратегия, опирающаяся на то, что инженер никогда не допустит ошибки, в долгосрочной перспективе обречена на провал.

Если говорить более абстрактно, то многопоточный код — это идеальное место, чтобы последовать общему совету: «Стремитесь к настолько простому коду, что его правильность станет очевидной, избегая настолько сложного, что будет непросто разглядеть его ошибочность».

Рекомендация 18. Не паникуйте

Книга казалась безумно сложной, и это стало одной из причин, по которой на ее пластиковой обертке крупным разборчивым шрифтом напечатали: «DON'T PANIC».

Дуглас Адамс (Douglas Adams)

Более точной формулировкой для текущей рекомендации будет такая: *вместо использования panic! старайтесь возвращать Result* (но «не паникуйте» звучит более интригующе).

Механизм паники в Rust создан в первую очередь на случай возникновения в программе ошибок, после которых невозможно восстановить работу, и *по умолчанию* он завершает работу потока, который выполняет panic!. Тем не менее этому базовому поведению есть альтернативы.

В частности, новички в Rust, которые пришли из языков, где есть система исключений (например, Java или C++), порой прибегают к использованию std::panic::catch_unwind в качестве средства для их симуляции, поскольку этот механизм позволяет перехватывать паники в более высокой точке стека вызовов.

Рассмотрим функцию, которая паникует при недействительном вводе:

```
fn divide(a: i64, b: i64) -> i64 {
    if b == 0 {
        panic!("Cowardly refusing to divide by zero!");
    }
    a / b
}
```

Попытка вызвать ее с недопустимым вводом ожидаемо приведет к сбою:

```
// Попытка выяснить, чему равно 0/0...
let result = divide(0, 0);

thread 'main' panicked at 'Cowardly refusing to divide by zero!', main.rs:11:9
note: run with `RUST_BACKTRACE=1` environment variable to display a backtrace
```

Обертка, которая использует catch_unwind для перехвата паники:

```
fn divide_recover(a: i64, b: i64, default: i64) -> i64 {
    let result = std::panic::catch_unwind(|| divide(a, b));
    match result {
        Ok(x) => x,
        Err(_) => default,
    }
}
```

казалось бы, работает и симулирует `catch`:

```
let result = divide_recover(0, 0, 42);
println!("result = {result}");

result = 42
```

Но видимость может быть обманчивой. Первая проблема этого подхода в том, что паники не всегда вскрываются. В компиляторе есть опция (доступна также через настройку профиля `Cargo.toml`), изменяющая поведение паники таким образом, что она сразу же завершает процесс:

```
thread 'main' panicked at 'Cowardly refusing to divide by zero!', main.rs:11:9
note: run with `RUST_BACKTRACE=1` environment variable to display a backtrace
/bin/sh: line 1: 29100 Abort trap: 6 cargo run --release
```

В таком случае возможность симуляции исключения может быть сохранена только благодаря более глобальным настройкам проекта. Кроме того, некоторые целевые платформы, например WebAssembly, в случае паники *всегда* отменяют процесс независимо от настроек компилятора или проекта.

Более тонкой проблемой, которая обнаруживается при обработке паники, является *безопасность исключений*: если паника произойдет посреди операции над структурой данных, она исключит все гарантии сохранения согласованности этих данных. Известно, что с 1990-х годов¹ сохранить внутренние инварианты в случае исключения становится крайне сложно. И это одна из главных причин, по которой Google запрещает использование исключений в коде на C++ (https://oreil.ly/Bc-_z).

Наконец, паники могут преодолевать (<https://oreil.ly/fIAtD>) границы FFI (foreign function interface — «интерфейс внешних функций»), речь о котором пойдет в рекомендации 34. *Используйте `catch_unwind`, чтобы не позволить паникам в коде Rust распространяться на вызывающий код другого языка* через границу FFI.

И что же применять для обработки ошибок вместо `panic!`? Для кода библиотек лучшей альтернативой будет переложить этот вопрос на кого-то другого, вернув `Result` с правильным типом ошибки (см. рекомендацию 4). Это позволит пользователям библиотеки принимать собственные взвешенные решения о последующих действиях, которые могут включать в себя передачу проблемы следующему вызывающему по цепочке через оператор `?`.

Но передача этой эстафеты должна где-то прекратиться, и в качестве общего правила нужно помнить, что вполне допустимо использовать `panic!` (или `unwrap()`,

¹ Том Каргилл (Tom Cargill) в своей статье 1994 года из журнала *C++ Report* (<https://oreil.ly/J9hes>) разбирает проблемы с безопасностью исключений в шаблонном коде на C++. О том же пишет Херб Саттер (Herb Sutter) на своем ресурсе *Guru of the Week* в разделе 8 (<https://oreil.ly/521d9>).

`expect()` и т. д.), если вы контролируете `main`. На этом этапе уже нет последующего вызывающего кода, которому можно было бы делегировать решение.

Еще одним чувствительным случаем применения `panic!`, даже в коде библиотеки, является ситуация, когда ошибки встречаются крайне редко и вы не хотите, чтобы пользователи засоряли свой код вызовами `.unwrap()`.

Если ошибка происходит только из-за, скажем, повреждения внутренних данных, а не в результате недопустимого ввода, тогда вызвать `panic!` можно. Иногда даже полезно разрешить ее возникновение при недопустимом вводе, но только там, где подобный ввод нетипичен. Такое решение лучше всего работает, когда соответствующие точки входа идут парами в виде:

- «безотказной» версии, чья сигнатура подразумевает, что она всегда завершится успешно (и которая паникует в случае провала);
- версии, допускающей сбой и возвращающей `Result`.

В отношении первых руководство по API (<https://oreil.ly/vXxDV>) предполагает, что `panic!` должна быть описана в конкретном разделе встроенной документации (см. рекомендацию 27).

Примером такой пары являются точки входа `String::from_utf8_unchecked` и `String::from_utf8` в стандартной библиотеке, хотя в этом случае паника фактически откладывается до момента использования `String`, созданной из недопустимого ввода.

Следуя советам этой рекомендации, вам нужно кое-что помнить. Во-первых, паника может проявляться в разных формах и то, что вы избежали `panic!`, подразумевает, что избежали также:

- `unwrap()` и `unwrap_err()`;
- `expect()` и `expect_err()`;
- `unreachable!()`.

Становится труднее обнаружить следующие моменты:

- `slice[index]`, когда индекс выходит за границы диапазона;
- `x / y`, когда у нулевой.

Второй нюанс при избегании паники заключается в следующем: никогда не нужно полагаться на то, что бдительность человека окажется безупречной.

А вот постоянная бдительность машин — дело другое: добавление в систему непрерывной интеграции (см. рекомендацию 32) проверки, которая будет обнаруживать новый потенциально паникующий код, окажется более надежным подходом. В качестве простого варианта вполне сойдет выполнение `grep`

в отношении наиболее часто паникующих точек входа, как показано ранее. Более тщательная проверка может включать дополнительные инструменты из экосистемы Rust (см. рекомендацию 31), такие как настройка варианта сборки с подтягиванием крейта `no_panic`.

Рекомендация 19. Избегайте отражения

Программисты, приходящие в Rust из других языков, часто имеют привычку использовать в качестве инструмента отражение. Они зачастую тратят немало времени, пытаясь реализовать на его основе те или иные проекты, но в итоге такое решение либо оказывается неэффективным, либо не работает вообще. Задача текущей рекомендации — попытаться застраховать вас от трясины лишнего времени на подобные тупиковые решения. Для этого мы разберем, что Rust конкретно может, а чего не может делать в плане отражения и какие доступны альтернативы.

Отражение — это способность программы проанализировать себя в процессе выполнения. Получая в ходе работы некий элемент, она отвечает для себя на следующие вопросы.

- Что можно выяснить о типе этого элемента?
- Что можно сделать с этой информацией?

Языки программирования с полноценной поддержкой отражения дают развернутые ответы на эти вопросы. Опираясь на полученную из отражения информацию, они способны предоставить некоторые из следующих возможностей или их все:

- определить тип элемента;
- изучить его содержимое;
- изменить его поля;
- вызвать его методы.

Языки, имеющие подобный уровень отражения, также *обычно* являются динамически типизируемыми (например, Python, Ruby), но существуют и статически типизируемые языки, которые тоже поддерживают отражение, в частности Java и Go.

Rust не поддерживает такой уровень отражения, так что рекомендации *избегать его* следовать легко. И для программистов, приходящих из языков с полноценной поддержкой отражения, подобный ее дефицит поначалу может казаться значительным пробелом, но в Rust есть другие возможности, которые предоставляют альтернативные способы решения многих аналогичных задач.

В C++ реализована более ограниченная поддержка отражения, известная как *идентификация типов в среде выполнения* (RTTI). В нем оператор `typeid` возвращает для объектов *полиморфных типов* (грубо говоря, классов с виртуальными функциями) уникальный идентификатор каждого типа.

`typeid`

Может выводить конкретный класс объекта, на который идет ссылка базового класса.

`dynamic_cast<T>`

Когда это безопасно и уместно, позволяет преобразовывать ссылки базовых классов в производные классы.

Rust не поддерживает и подобный RTTI-стиль отражения, так что и в этом случае следовать рекомендации легко.

В Rust есть возможности, предоставляющие *похожую* функциональность с помощью модуля `std::any`, но они весьма ограничены (позднее будет сказано, как именно), поэтому при наличии альтернатив их лучше избегать.

Поначалу *кажется* магией первая подобная отражению возможность `std::any` — определение имени типа элемента. В примере далее задействуется пользовательская функция `tname()`:

```
let x = 42u32;
let y = vec![3, 4, 2];
println!("x: {} = {}", tname(&x), x);
println!("y: {} = {:?}", tname(&y), y);
```

Она выводит типы вместе со значениями:

```
x: u32 = 42
y: alloc::vec::Vec<i32> = [3, 4, 2]
```

Реализация `tname()` показывает, что нам готовит компилятор, — это обобщенная функция (как описывалось в рекомендации 12), а значит, каждый ее вызов является другой функцией (`tname::<u32>` или `tname::<Vec<i32>>`):

```
fn tname<T: ?Sized>(_v: &T) -> &'static str {
    std::any::type_name::<T>()
}
```

Показанная реализация предоставляется библиотечной функцией `std::any::type_name<T>`, тоже обобщенной. Эта функция имеет доступ только к информации *этапа компиляции* — в ней нет кода, который бы определял тип в среде выполнения. Если вернуться к типам трейт-объектов, использованных в рекомендации 12, можно отметить, что она демонстрирует следующее:

```
let square = Square::new(1, 2, 2);
let draw: &dyn Draw = &square;
let shape: &dyn Shape = &square;

println!("square: {}", tname(&square));
println!("shape: {}", tname(&shape));
println!("draw: {}", tname(&draw));
```

Доступны только типы трейт-объектов, а не тип (`Square`) конкретного внутреннего элемента:

```
square: reflection::Square
shape: &dyn reflection::Shape
draw: &dyn reflection::Draw
```

Строка, возвращаемая `type_name`, подходит только для диагностики — это явно лучший помощник, чье содержимое может меняться и не является уникальным, поэтому *не пытайтесь спарсить результаты* `type_name`. Если вам нужен глобально уникальный идентификатор, используйте `TypeId`:

```
use std::any::TypeId;

fn type_id<T: 'static + ?Sized>(_v: &T) -> TypeId {
    TypeId::of::<T>()
}

println!("x has {:?}", type_id(&x));
println!("y has {:?}", type_id(&y));

x has TypeId { t: 18349839772473174998 }
y has TypeId { t: 2366424454607613595 }
```

Для людей этот вывод не особо полезен, но гарантия уникальности означает, что его можно использовать в коде. Тем не менее обычно лучше не брать непосредственно `TypeId`, а использовать трейт `std::any::Any`, поскольку в стандартной библиотеке есть дополнительная функциональность для работы с экземплярами `Any` (описаны далее).

Трейт `Any` содержит единственный метод `type_id()`, который возвращает значение `TypeId` для типа, этот трейт реализующего. Тем не менее самостоятельно его реализовать нельзя, так как `Any` уже идет с blanket-реализацией для большинства произвольных типов `T`:

```
impl<T: 'static + ?Sized> Any for T {
    fn type_id(&self) -> TypeId {
        TypeId::of::<T>()
    }
}
```

Blanket-реализация не охватывает *каждый* тип T: *граница времени жизни* 'static означает, что если T содержит любые ссылки, имеющие не 'static время жизни, то typeId для T не реализуется. Это специально сделанное ограничение (<https://oreil.ly/BjglR>), причина которого в том, что сроки жизни не являются полностью частью типа: typeId::of::&'a T будет тем же, что и typeId::of::&'b T, несмотря на различное время жизни. В связи с этим повышается вероятность путаницы и появления ошибок в коде.

Напомню (см. рекомендацию 8), что трейт-объект — это жирный указатель, который содержит ссылку на соответствующий элемент, а также ссылку на vtable реализации трейта. Как показано на рис. 3.4, vtable для Any содержит одну запись, связанную с методом type_id(), возвращающим тип элемента:

```
let x_any: Box<dyn Any> = Box::new(42u64);
let y_any: Box<dyn Any> = Box::new(Square::new(3, 4, 3));
```

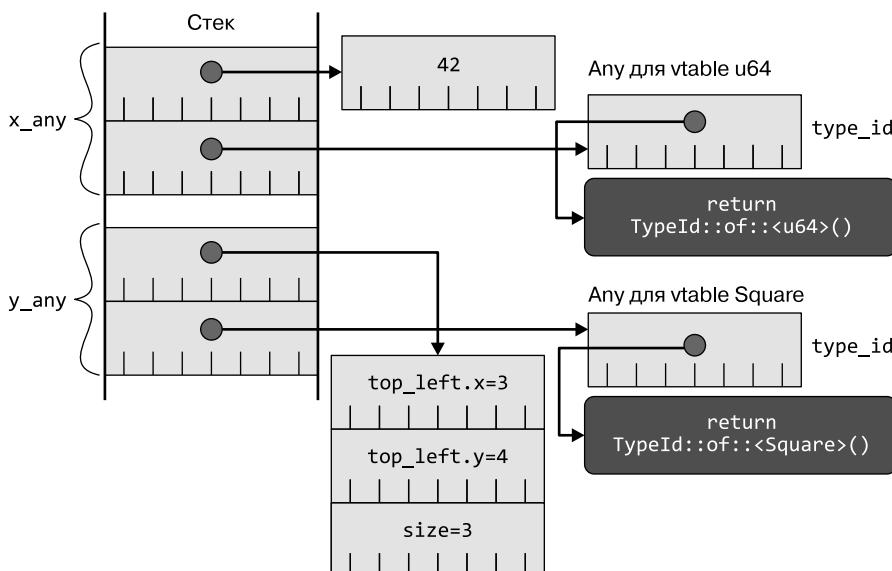


Рис. 3.4. Трейт-объекты Any, каждый с указателями на конкретные элементы и vtable

Помимо пары косвенных адресаций, трейт-объект `dyn Any`, по сути, является комбинацией сырого указателя и идентификатора типа. Это означает, что стандартная библиотека может предложить дополнительные обобщенные методы, определенные для трейт-объекта `dyn Any`. Эти методы являются обобщенными относительно некоего дополнительного типа T.

`is::<T>()`

Указывает, равен ли тип трейт-объекта некоему другому типу T.

downcast_ref::<T>()

Возвращает ссылку на конкретный тип T при условии, что тип трейт-объекта соответствует T .

downcast_mut::<T>()

Возвращает мутабельную ссылку на конкретный тип T при условии, что тип трейт-объекта соответствует T .

Заметьте, что трейт `Any` только аппроксимирует функциональность отражения — именно программист выбирает (на этапе компиляции), что нужно явно собрать что-либо (`&dyn Any`) отслеживающее тип элемента при компиляции, а также его расположение. Возможность, скажем, выполнить пониждающее приведение к изначальному типу доступна, только если трейт-объект `Any` уже создан.

Существует сравнительно немного сценариев, где в Rust на этапе компиляции и в среде выполнения с элементом ассоциируются разные типы. И доминируют среди них *трейт-объекты*: элемент конкретного типа `Square` можно неявно привести к трейт-объекту `dyn Shape` для трейта, который этот тип реализует. Такое приведение создает жирный указатель (объект + vtable) из простого указателя (объект/элемент).

Напомню также (см. рекомендацию 12), что трейт-объекты на деле не являются объектно-ориентированными. То есть `Square` не является `Shape`, он лишь реализует интерфейс `Shape`. То же касается границ трейтов: граница трейта `Shape: Draw` не означает связь по типу «является» — эта запись значит «*также реализует*», поскольку vtable для `Shape` включает записи для методов `Draw`.

Для некоторых простых границ трейтов:

```
trait Draw: Debug {
    fn bounds(&self) -> Bounds;
}

trait Shape: Draw {
    fn render_in(&self, bounds: Bounds);
    fn render(&self) {
        self.render_in(overlap(SCREEN_BOUNDS, self.bounds()));
    }
}
```

равнозначные трейт-объекты:

```
let square = Square::new(1, 2, 2);
let draw: &dyn Draw = &square;
let shape: &dyn Shape = &square;
```

имеют схему со стрелками (рис. 3.5; взята из рекомендации 12), которая проясняет проблему: при наличии объекта `dyn Shape` невозможно напрямую создать трейт-объект `dyn Draw`, поскольку нет возможности вернуться к vtable,

относящейся к `impl Draw` для `Square`, несмотря на то что соответствующую часть его содержимого (адрес метода `Square::bounds()`) теоретически восстановить можно. (В последующих версиях Rust это наверняка изменится — читайте последний раздел этой рекомендации.)

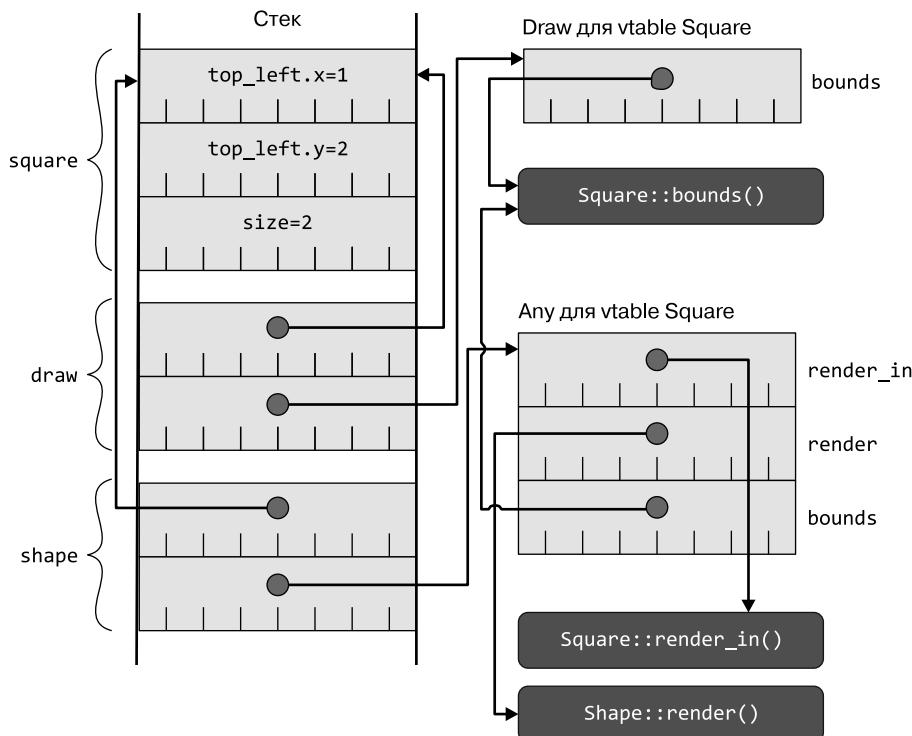


Рис. 3.5. Трейт-объекты для границ трейтов с отдельными vtable для Draw и Shape

При сравнении этой схемы с предыдущей (см. рис. 3.4) становится ясно, что явно построенный трейт-объект `&dyn Any` не поможет. `Any` позволяет восстановить изначальный конкретный тип соответствующего элемента, но в среде выполнения невозможно увидеть, какой трейт он реализует, или получить доступ к связанной vtable, которая дала бы возможность создать трейт-объект.

Что же нам доступно вместо этого?

Основным инструментом здесь станут определения трейтов, что согласуется с рекомендацией для других языков — рекомендация 65 из книги *Effective Java* советует «использовать вместо отражения интерфейсы». Если коду нужно опираться на доступность для элемента определенного поведения, прописывайте последнее в виде трейта (см. рекомендацию 2). Даже если нужное поведение

нельзя выразить через набор сигнатур методов, применявшиеся трейты-маркеры, чтобы обозначить совместимость с этим поведением, — так будет безопаснее и эффективнее, чем, например, просматривать имя класса для выяснения конкретного префикса.

Код, ожидающий трейт-объекты, можно использовать также с объектами, имеющими вспомогательный код, который не был доступен на этапе линковки программы, так как динамически подгружался в среде выполнения (посредством `dlopen(3)` или чего-то аналогичного), что означает невозможность мономорификации обобщения (см. рекомендацию 12).

Кроме того, отражение иногда используется и в других языках, чтобы сделать возможной одновременную загрузку в программу нескольких несовместимых версий одной библиотечной зависимости, обходя ограничения линковки, утверждающие, что «может присутствовать только одна версия». В Rust это не обязательно, так как здесь Cargo вполне справляется с обработкой нескольких версий одной библиотеки (см. рекомендацию 25).

Наконец, в качестве более эффективной и типобезопасной альтернативы коду, который парсит содержимое элемента в среде выполнения, в Rust можно использовать макросы, особенно `derive`, для автоматической генерации вспомогательного кода, понимающего тип элемента на этапе компиляции. Подробнее система макросов Rust рассматривается в рекомендации 28.

Повышающее приведение вверх в будущих версиях Rust

Текст этой рекомендации был составлен в 2021 году и оставался актуальным вплоть до момента подготовки книги к печати в 2024-м, когда в Rust должна была появиться новая функциональность, меняющая некоторые сопутствующие детали.

Эта новая функциональность, названная *повышающим приведением трейтов* (<https://oreil.ly/gWJUW>), позволяет приводить трейт-объект `dyn T` к трейт-объекту `dyn U`, когда `U` выступает одним из супертрейтов `T` (`trait T: U { ... }`). Эта функциональность была включена в `#![feature(trait_upcasting)]` еще до своего официального релиза, который должен состояться в Rust 1.76.

Для предыдущего примера это означает, что трейт-объект `&dyn Shape` теперь можно преобразовать в трейт-объект `&dyn Draw`, приблизившись к связи по типу «является», отражающей принцип подстановки Лисков. В результате подобное преобразование может вызывать цепную реакцию для внутренних деталей реализации `vtable`, которая наверняка станет сложнее версии, показанной на рис. 3.5.

Тем не менее основные пункты описанной рекомендации это не затрагивает — трейт `Any` не имеет родительских трейтов (supertraits), поэтому возможность повышающего приведения ничего к его функциональности не добавляет.

Рекомендация 20. Не поддавайтесь соблазну излишней оптимизации

То, что Rust позволяет вам безопасно писать суперкрутые алгоритмы без аллокации и копирования, еще не значит, что таковым будет каждый создаваемый алгоритм.

trentj (<https://oreil.ly/fQMfu>)

Большинство рекомендаций книги нацелены на то, чтобы помочь программистам познакомиться с Rust и его идиомами. Но текущая посвящена проблеме, которая может возникнуть, когда программист слишком отклоняется от этого направления и стремится задействовать потенциал Rust для максимизации эффективности программы в ущерб удобству ее использования и обслуживаемости.

Структуры данных и аллокация

Подобно указателям в других языках, ссылки в Rust позволяют повторно использовать данные, не создавая их копий. Причем Rust отличается тем, что его правила времени жизни и заимствования ссылок позволяют делать это *безопасно*. Тем не менее следование правилам заимствования (см. рекомендацию 15) порой усложняет итоговый код.

В особенности это касается структур данных, когда вы можете выбирать — аллокировать свежую копию чего-либо или включить ссылку на уже существующую копию.

В качестве примера рассмотрим код, который парсит потоки байтов, извлекая данные, закодированные в виде структур «тип — длина — значение» (type — length — value, TLV) в таком формате:

- сначала идет байт, который описывает тип значения (здесь хранится в поле `type_code`)¹;
- далее следует байт, описывающий длину значения в байтах (здесь используется для создания среза установленной длины);
- последним указывается заданное количество байтов значения, сохраняемое в поле `value`:

¹ Данное поле нельзя назвать `type`, так как в Rust это зарезервированное ключевое слово. Можно это ограничение обойти, используя префикс идентификатора `r#` (<https://oreil.ly/oC8VO>) (получится поле `r#type: u8`), но обычно проще переименовать само поле.

```

/// Тип – длина – значение из потока данных
#[derive(Clone, Debug)]
pub struct Tlv<'a> {
    pub type_code: u8,
    pub value: &'a [u8],
}

pub type Error = &'static str; // Какая-то локальная ошибка типа
/// Извлечение следующего TLV из `input`
/// и возвращение оставшихся необработанных данных
pub fn get_next_tlv(input: &[u8]) -> Result<(Tlv, &[u8]), Error> {
    if input.len() < 2 {
        return Err("too short for a TLV");
    }
    // В TLV размер каждого из элементов Т и Л составляет 1 байт
    let type_code = input[0];
    let len = input[1] as usize;
    if 2 + len > input.len() {
        return Err("TLV longer than remaining data");
    }

    let tlv = Tlv {
        type_code,
        // Ссылка на соответствующий фрагмент входных данных
        value: &input[2..2 + len],
    };
    Ok((tlv, &input[2 + len..]))
}

```

Эффективность этой структуры данных `Tlv` объясняется тем, что она содержит ссылку на соответствующий кусок входных данных без их копирования и предоставляемые Rust гарантии безопасности памяти обеспечивают сохранение валидности этой ссылки. В некоторых сценариях это оказывается прекрасным решением, но в случаях, когда какому-то элементу нужно сохранять неизменную связь с экземпляром структуры данных, возникают сложности (о чем говорилось в рекомендации 15).

Рассмотрим в качестве примера сетевой сервер, который получает сообщения в форме TLV. Полученные данные затем можно спарсить в экземпляры `Tlv`, но срок их жизни будет соответствовать сроку жизни входящего сообщения, которое может быть временным `Vec<u8>` в куче или расположенным где-то буфером, который повторно используется для нескольких сообщений.

Все это создает проблему, если код сервера хочет сохранить входящее сообщение, чтобы иметь возможность обратиться к нему позднее:

```

pub struct NetworkServer<'a> {
    // ...
    /// Последнее сообщение максимального размера
    max_size: Option<Tlv<'a>>,
}

```

```

/// Код типа сообщения set-maximum-size
const SET_MAX_SIZE: u8 = 0x01;

impl<'a> NetworkServer<'a> {
    pub fn process(&mut self, mut data: &'a [u8]) -> Result<(), Error> {
        while !data.is_empty() {
            let (tlv, rest) = get_next_tlv(data)?;
            match tlv.type_code {
                SET_MAX_SIZE => {
                    // Сохранение последнего сообщения `SET_MAX_SIZE`
                    self.max_size = Some(tlv);
                }
                // (Обработка других типов сообщений)
                // ...
                _ => return Err("unknown message type"),
            }
            data = rest; // Обработка оставшихся данных
                         // на следующей итерации
        }
        Ok(())
    }
}

```

Этот код компилируется, но использовать его невозможно: время жизни `NetworkServer` должно быть меньше времени жизни любых данных, поступающих в его метод `process()`. Это означает, что простой цикл обработки...

НЕ КОМПИЛИРУЕТСЯ

```

let mut server = NetworkServer::default();
while !server.done() {
    // Считывание данных в новый вектор
    let data: Vec<u8> = read_data_from_socket();
    if let Err(e) = server.process(&data) {
        log::error!("Failed to process data: {:?}", e);
    }
}

```

...не компилируется, так как время жизни скоротечных данных привязывается к более долгоживущему серверу:

```

error[E0597]: `data` does not live long enough
--> src/main.rs:375:40
|
372 |     while !server.done() {
|         ----- borrow later used here
373 |         // Считывание данных в новый вектор
|

```

```

374 |     let data: Vec<u8> = read_data_from_socket();
|         ---- binding `data` declared here
375 |     if let Err(e) = server.process(&data) {
|             ^^^^^^ borrowed value does not live
|                     long enough
...
378 |     }
|         - `data` dropped here while still borrowed

```

Перестройка кода таким образом, чтобы он повторно использовал более долгоживущий буфер, тоже не помогает.

НЕ КОМПИЛИРУЕТСЯ

```

let mut perma_buffer = [0u8; 256];
// Время жизни внутри `perma_buffer`
let mut server = NetworkServer::default();

while !server.done() {
    // Повторное использование того же буфера
    // для очередной загрузки данных
    read_data_into_buffer(&mut perma_buffer);
    if let Err(e) = server.process(&perma_buffer) {
        log::error!("Failed to process data: {:?}", e);
    }
}

```

На сей раз компилятор жалуется, что код пытается зафиксировать иммутабельную ссылку, при этом выдавая мутабельную на тот же буфер:

```

error[E0502]: cannot borrow `perma_buffer` as mutable because it is also
              borrowed as immutable
--> src/main.rs:353:31
|
353 |     read_data_into_buffer(&mut perma_buffer);
|           ^^^^^^^^^^^^^^^^^^^^^ mutable borrow occurs here
|
354 |     if let Err(e) = server.process(&perma_buffer) {
|             -----
|             |           |
|             |               immutable borrow occurs here
|             |
|             immutable borrow later used here
|

```

Основная проблема заключается в том, что структура `Tlv` ссылается на временные данные — и это вполне нормально для временной обработки, но несовместимо с сохранением состояния для последующего использования.

Тем не менее, если структуру `Tlv` преобразовать так, чтобы она владела собственным содержимым:

```
#[derive(Clone, Debug)]
pub struct Tlv {
    pub type_code: u8,
    pub value: Vec<u8>, // Данные, которыми владеет куча
}
```

и соответствующим образом скорректировать код `get_next_tlv()`, добавив дополнительный вызов к `.to_vec()`:

```
// ...
let tlv = Tlv {
    type_code,
    // Копирование соответствующего фрагмента данных в кучу
    // Поле длины в TLV представляет один `u8`, значит,
    // копируется не более 256 байт
    value: input[2..2 + len].to_vec(),
};
```

то работа кода сервера значительно упростится. Владеющая данными структура `Tlv` не имеет параметра времени жизни, а значит, серверной структуре данных он тоже не нужен и оба варианта цикла прекрасно работают.

Кто боится большого и страшного `Copy`?

Одна из причин, по которой программисты так любят сокращать копирование, состоит в том, что Rust обычно делает копирование и аллоцирование явными. Видимый вызов метода наподобие `.to_vec()` или `.clone()` либо функции вроде `Box::new()` открыто показывает, что происходит копирование или аллокация. Это резко отличается от поведения C++, где можно опрометчиво написать код, который будет беспечно выполнять аллокацию за кадром, в частности, с помощью конструктора копий или оператора присваивания.

Явное выражение операции аллокации или копирования не является объективной причиной оптимизировать ее, особенно ценой удобства использования кода. Во многих ситуациях разумнее будет сначала сосредоточиться именно на удобстве и *вносить корректировки для повышения эффективности, только если производительность приоритетна* — и если бенчмарки (см. рекомендацию 30) показывают, что сокращение количества копий окажет ощутимое влияние.

Кроме того, эффективность вашего кода обычно важна, только если необходимо его масштабирование для широкого применения. Если же по факту допущенные в коде компромиссы себя не оправдывают и при аудитории в миллионы пользователей он перестает справляться, то это становится серьезной проблемой.

Здесь нужно помнить пару важных моментов. Первый был скрыт за неоднозначным словом «*как правило*», когда говорилось, что копии, как правило, видимы. Большим исключением здесь выступают типы `Copy`, когда компилятор волей-неволей молча создает копии, переходя от семантики перемещения к семантике копирования. Так что здесь будет нeliшним повторить совет из рекомендации 10: «Не реализуйте `Copy`, если только побитовое копирование не является допустимым и быстрым». Но верно и обратное: «Подумайте о реализации `Copy`, если побитовое копирование допустимо и выполняется быстро». Например, типы `enum`, которые не несут дополнительных данных, обычно проще использовать, если они производят `Copy`.

Второй актуальный момент — это возможные компромиссы при использовании `no_std`. В рекомендации 33 говорится, что зачастую можно писать код, который будет совместим с `no_std` при минимуме изменений, и код, полностью исключающий аллокацию, упрощает эту задачу. Тем не менее ориентация на `no_std`-среду с поддержкой аллокации в куче (с помощью библиотеки `alloc`, также описанной в рекомендации 33) может обеспечивать оптимальный баланс удобства и поддержки `no_std`.

Ссылки и умные указатели

Совсем недавно я сознательно попробовал эксперимент — перестал заботиться о написании гипотетически идеального кода. Вместо этого я при необходимости вызываю `clone()` и использую Arc для более плавного задействования локальных объектов в потоках и промисах.

Работает прекрасно.

Джош Триплетт (*Josh Triplett*)
(<https://oreil.ly/1ViCT>)

Проектирование структуры данных таким образом, чтобы она владела своим содержимым, определенно может повысить эргономику программы, но все равно остаются потенциальные проблемы, если некоторым структурам данных нужно использовать одну информацию. Если эти данные иммутабельны, тогда подход с наличием у каждого своей копии прекрасно работает. Если же информация может меняться (что бывает очень часто), тогда наличие нескольких копий подразумевает необходимость обновления нескольких областей с сохранением их согласованности.

Решить эту проблему помогают типы умных указателей, позволяющие переключаться с модели одного владельца на модель совместного владения. Умные

указатели `Rc` (в однопоточном коде) и `Arc` (в многопоточном) обеспечивают подсчет ссылок, который поддерживает эту модель. Если оттолкнуться от предположения, что нам нужна мутабельность, то они обычно совмещаются с внутренним типом, который делает ее возможной вне зависимости от проверки правил заимствования Rust.

RefCell

Для обеспечения внутренней мутабельности в однопоточном коде в виде типичной комбинации `Rc<RefCell<T>>`.

Mutex

Для обеспечения внутренней мутабельности в многопоточном коде (см. рекомендацию 17) в виде типичной комбинации `Arc<Mutex<T>>`.

Более подробно этот переход описывается в примере с `GuestRegister` в рекомендации 15, но суть в том, что вам не нужно рассматривать умные указатели Rust лишь в качестве крайней меры. Не станет признанием поражения то, что в вашем дизайне вместо сложной сети взаимосвязанных сроков жизни ссылок будут использоваться умные указатели — *они порой позволяют написать более простой, легко обслуживаемый и удобный в работе код*.

ГЛАВА 4

Зависимости

Когда боги желают нас наказать, они отвечают на наши молитвы.

Оскар Уайльд

На протяжении десятилетий идея повторного использования кода была только мечтой. Мысль о том, что код можно написать один раз, упаковать в библиотеку и задействовать во множестве разных приложений, была неким идеалом, реализованным лишь в нескольких стандартных библиотеках и для внутренних корпоративных инструментов.

Однако стремительное развитие Интернета и ПО с открытым исходным кодом изменило эту ситуацию. Первым доступным хранилищем с обширной коллекцией полезных библиотек, инструментов и вспомогательных утилит, упакованных для упрощения использования, стал CPAN (Comprehensive Perl Archive Network) — обширный сетевой архив ресурсов для языка Perl, запущенный в 1995 году. Сегодня же практически у каждого современного языка есть обширная коллекция библиотек с открытым исходным кодом, расположенных в хранилище пакетов, упрощающем и ускоряющем процесс добавления новых зависимостей¹.

Однако вместе с простотой, удобством и скоростью пришли новые проблемы. *Обычно* по-прежнему быстрее повторно использовать существующий код, нежели писать свой, но создание зависимостей от чьего-то стороннего кода несет с собой возможные подвохи и риски. Это и стало стимулом для написания данной главы книги, которая поможет вам лучше их понимать.

Акцент здесь сделан конкретно на Rust и сопутствующем использовании инструмента Cargo, но многие рассматриваемые вопросы, темы и проблемы в той же степени касаются и других наборов инструментов, в том числе других языков.

¹ За ярким исключением в виде C и C++, где управление пакетами остается в определенной степени фрагментированным.

Рекомендация 21. Разберитесь в принципах семантического версионирования

Если мы признаем, что SemVer — это только приблизительная оценка, которая представляет лишь часть возможного спектра изменений, то начнем воспринимать его как несовершенный инструмент.

Титус Винтерс (Titus Winters), автор книги Software Engineering at Google (O'Reilly)

Cargo, пакетный менеджер Rust, позволяет автоматически выбирать зависимости (см. рекомендацию 25) для кода в соответствии с *семантическим версионированием* (SemVer). Стока в `Cargo.toml` вроде:

```
[dependencies]
serde = "1.4"
```

сообщает Cargo, какой диапазон версий приемлем для этой зависимости. В официальной документации (<https://oreil.ly/fchXS>) подробно описывается, как можно указывать точные диапазоны допустимых версий, и самыми популярными вариантами являются следующие.

"1.2.3"

Указывает, что допустима любая версия, которая в соответствии с SemVer совместима с 1.2.3.

"^1.2.3"

Иной способ указать то же самое более явно.

"=1.2.3"

Закрепляет одну конкретную версию, не допуская никаких альтернатив.

"~1.2.3"

Допускает версии, которые согласно SemVer совместимы с 1.2.3, но только те, в которых меняется последний номер (то есть 1.2.4 допустима, а 1.3.0 — нет).

"1.2.*"

Допускает любую версию, соответствующую подставленному значению.

В табл. 4.1 показан пример диапазонов, которые обозначают эти спецификаторы.

При выборе версий зависимостей Cargo, как правило, выбирает максимальную, вписывающуюся в комбинацию всех диапазонов SemVer.

И поскольку семантическое версионирование лежит в основе процесса разрешения зависимостей в Cargo, текущая рекомендация посвящена изучению дополнительных деталей этой системы.

Таблица 4.1. Обозначение допустимых версий зависимостей в Cargo

Спецификатор	1.2.2	1.2.3	1.2.4	1.3.0	2.0.0
"1.2.3"	Нет	Да	Да	Да	Нет
"^1.2.3"	Нет	Да	Да	Да	Нет
"=1.2.3"	Нет	Да	Нет	Нет	Нет
"~1.2.3"	Нет	Да	Да	Нет	Нет
"1.2.*"	Да	Да	Да	Нет	Нет
"1.*"	Да	Да	Да	Да	Нет
"*"	Да	Да	Да	Да	Да

Основные аспекты SemVer

Основные составляющие семантического версионирования перечислены в документации (<https://oreil.ly/sBrbZ>) и приведены далее.

На основе номера версии в формате «Старшая.Младшая.Патч» инкрементировать:

- старшую версию при внесении несовместимых изменений API;
- младшую при добавлении функциональности обратно совместимым образом;
- версию патча при внесении обратно совместимых исправлений ошибок.

И один важный момент кроется в деталях.

1. После выпуска версионированного пакета содержимое этой его версии изменяться **не должно**. Любые последующие изменения **необходимо** выпускать в рамках новой версии.

Иными словами:

- изменение *чего-либо* требует указания новой версии патча;
- *добавление* чего-либо в API таким образом, что существующие пользователи крейта продолжат компилироваться и работать, требует повышения младшей версии;
- *удаление* или *изменение* чего-либо в API требует повышения уже старшей версии.

Есть и еще одно важное дополнение к правилам SemVer.

2. Нулевая старшая версия (0.y.z) означает начальную стадию разработки. В ней в любой момент **может** измениться что угодно, и ее публичный API **не следует** считать стабильным.

Cargo слегка адаптирует последнее правило, «сдвигая влево» прежний порядок так, чтобы изменение крайнего левого ненулевого номера отражало несовместимые модификации кода. Это означает, что версии с 0.2.3 по 0.3.0, как и версии с 0.0.4 по 0.0.5, могут включать несовместимые изменения API.

SemVer для авторов крейтов

Теоретически теория — это то же самое, что и практика. Но практически это не так.

Будучи автором крейта, следовать первому из перечисленных правил теоретически легко: если вы хоть что-то меняете, нужно делать новый релиз. В этом помогут *теги Git* — по умолчанию тег закреплен за конкретным коммитом, и его можно перемещать только вручную с помощью опции `--force`. В отношении крейтов, публикуемых на `crates.io`, автоматически работает та же политика, поскольку реестр отклоняет вторую попытку опубликовать одну и ту же версию. Основная опасность несоответствия возникает в случае, когда вы замечаете ошибку *сразу после релиза* и вам нужно устоять перед соблазном просто внести исправление.

В спецификации SemVer описывается совместимость API, так что если вы вносите в поведение незначительное изменение, которое не затрагивает API, тогда будет достаточно простого обновления версии патча. (Однако если ваш крейт является популярной зависимостью, то на практике вам может потребоваться учесть закон Хайрума (https://oreil.ly/7lQQ_): насколько бы незначительное изменение вы ни вносили в код, кто-нибудь из его потребителей наверняка окажется зависим от его старого поведения, даже если API останется неизменным.)

Сложность для авторов крейтов представляют последние правила, которые требуют точного определения того, является ли изменение обратно совместимым. Некоторые изменения очевидно несовместимы — удаление публичных точек входа или типов, корректировка сигнатур методов, в то время как другие, наоборот, совместимы (например, добавление нового метода в `struct` или добавление новой константы), но между ними лежит большая зона неопределенности.

Чтобы прояснить этот момент, в книге по Cargo (<https://oreil.ly/Y6ZLJ>) подробно рассказывается, что является, а что не является обратно совместимым. В большинстве случаев все вполне очевидно, но есть моменты, которые стоит проговорить.

- Добавление новых элементов *обычно* безопасно, но может вызывать конфликты, если в использующем крейт коде уже задействуется что-то с таким же именем, как у нового элемента.
 - Это особенно опасно, если пользователь выполняет из крейта импорт с подстановкой, поскольку тогда все элементы этого крейта автоматически находятся в основном пространстве имен пользователя. В рекомендации 23 советуется не делать этого.

- И даже без импорта с подстановкой имени нового метода трейта (с предустановленной реализацией; см. рекомендацию 13) или нового наследуемого метода может конфликтовать с существующим именем.
- Стремление Rust охватить все варианты означает, что изменение их доступного набора может стать критическим.
 - Выполнение `match` для `enum` должно охватывать все возможные варианты, поэтому, если крейт добавляет новую вариацию `enum`, это становится критическим изменением (если только это `enum` уже не отмечено как `non_exhaustive` — добавление `non_exhaustive` тоже является критическим изменением).
 - Явное создание экземпляра `struct` требует изначального значения для всех полей, поэтому добавление поля в структуру, которая может быть инстанцирована публично, — это критическое изменение. Структуры, которые имеют закрытые поля, опасности не представляют, так как пользователи крейта все равно не могут создавать их явно. При этом `struct` можно отметить и как `non_exhaustive`, чтобы не позволить внешним пользователям создавать ее явно.
- Изменение трейта, после которого он утрачивает *объектную безопасность* (см. рекомендацию 12), является критическим. Любые пользователи, которые создают для трейта объекты, перестанут компилироваться.
- Добавление новой `blanket`-реализации для трейта — это критическое изменение. Любые пользователи, которые уже реализуют этот трейт, в итоге окажутся с двумя конфликтующими реализациями.
- Изменение *лицензии* крейта с открытым исходным кодом является несовместимым изменением: оно может нарушить работу пользователей вашего крейта, имеющих строгие ограничения в отношении допустимых лицензий. *Подумайте о том, чтобы сделать лицензию частью своего API.*
- Изменение предустановленной функциональности (см. рекомендацию 26) крейта потенциально критическое. Удаление базовой возможности наверняка приведет к нарушению работы (если только она уже не фиктивная). Добавление базовой возможности может нарушить работу в зависимости от того, что она привносит. *Подумайте о том, чтобы сделать набор базовых функций крейта частью своего API.*
- Изменение кода библиотеки для применения новой возможности Rust может стать несовместимым изменением, так как нарушит работу пользователей крейта, которые еще не обновили свой компилятор до версии, включающей эту возможность. Тем не менее большинство крейтов Rust воспринимают повышение минимальной поддерживаемой версии Rust (`minimum supported Rust version, MSRV`) как *некритическое* (<https://oreil.ly/Tadgz>), так что *подумайте, не составляет ли MSRV часть вашего API.*

Из всех этих правил вытекает очевидное следствие: чем меньше у крейта публичных элементов, тем меньше вероятность внести критическое изменение (см. рекомендацию 22).

Однако сравнение всех публичных элементов API двух последовательных релизов на совместимость неизбежно требует много времени и в лучшем случае дает лишь *примерную* оценку уровня изменений (старшая/младшая/патч). Учитывая, что это сравнение в некотором смысле механический процесс, можно надеяться, что для его упрощения будет разработан соответствующий инструментарий (см. рекомендацию 31)¹.

Если вам нужно внести несовместимое изменение, обновляющее старшую версию, то желательно упростить жизнь пользователям, обеспечив итоговую доступность всей прежней функциональности, даже если API претерпел кардинальную модификацию. Для удобства пользователей желательно по возможности придерживаться такой последовательности действий.

1. Выpusкать обновление младшей версии, включающее новый вариант API и отмечающее старый вариант как `deprecated`, с указанием инструкций по переходу.
2. Выпускать обновление старшей версии, удаляющее устаревшие части API.

Более тонкий момент заключается в том, чтобы *делать критические изменения действительно критическими*. Если ваш крейт меняет свое поведение таким образом, что становится несовместимым для имеющихся пользователей, но при этом *может* продолжать применять тот же API — не делайте так. Делайте обязательным изменение типов (и повышение старшей версии), чтобы пользователи точно не могли по оплошности задействовать новую версию некорректно.

Что касается менее осозаемых элементов вашего API, таких как MSRV (<https://oreil.ly/Gast->) или лицензии, — подумайте об использовании в системе непрерывной интеграции (continuous integration, CI) инструмента для проверки (см. рекомендацию 32) наличия изменений (например, `cargo-deny`; подробнее — в рекомендации 25).

Наконец, не бойтесь версии 1.0.0, так как она станет гарантией того, что ваш API теперь стабилен. Многие крейты попадают в ловушку, бесконечно оставаясь в версии 0.x, но это снижает и без того ограниченную экспрессивность версионирования из трех категорий (старшая/младшая/патч) всего до двух (действующая старшая/действующая младшая).

¹ Например, есть инструмент `cargo-semver-checks`, нацеленный как раз на это.

SemVer для пользователей крейтов

С позиции пользователя крейта *теоретическими* ожиданиями от новой версии зависимости будут следующие.

- Новая версия патча зависимости Should Just Work™ (должна просто работать).
- Обновленные части API могут требовать изучения на предмет новых и более эффективных способов использования крейта. Однако если вы эти части применяете, то не сможете откатить зависимость обратно к старой версии.
- От новой старшей версии можно ожидать чего угодно. В этом случае ваш код наверняка перестанет компилироваться и вам потребуется переписывать его фрагменты, чтобы они соответствовали новому API. И даже если код продолжит компилироваться, вам следует *убедиться, по-прежнему ли вы используете API валидным образом после изменения старшей версии*, так как ограничения и предварительные условия библиотеки могли измениться.

На практике в силу закона Хайрума даже эти первые два типа изменений *могут* неожиданным образом сказываться на поведении, в том числе в коде, который без проблем компилируется.

В результате описанных ожиданий ваши спецификации зависимостей обычно будут принимать форму вроде "1.4.3" или "0.7", которая включает последовательные совместимые версии. *Старайтесь не применять целиком подстановочную форму версии зависимости вроде "*" или "0.*".* Такая запись говорит, что ваш крейт может использовать *любую* версию зависимости с *любым* API, что вряд ли будет соответствовать тому, что вы хотите показать. Отсутствие возможности подстановки также является требованием для публикации на crates.io. Заявки с символами "*" этим порталом отклоняются (<https://oreil.ly/i6ELc>).

Тем не менее в более длительной перспективе простое игнорирование изменений старшей версии зависимостей становится небезопасным. Когда у библиотеки меняется старшая версия, высока вероятность того, что для предыдущей перестанут предоставляться исправления ошибок и, что самое главное, обновления безопасности. Тогда спецификация версии, например "1.4", по мере появления новых релизов 2.x будет все больше отставать, сохраняя неустранимые проблемы с безопасностью.

В итоге вам нужно либо принять риски, сопряженные с продолжением использования старой версии, либо *в конечном итоге обновлять старшие версии своих зависимостей*. Такие инструменты, как Cargo update или Dependabot (см. рекомендацию 31), могут уведомлять вас о появлении обновлений, позволяя запланировать их применение на удобное время.

Рассуждения

Семантическое версионирование имеет свою цену: каждое изменение в крейте необходимо анализировать, чтобы определить подходящий для него тип повышения версии. Кроме того, семантическое версионирование является инструментом ограниченного применения, так как в лучшем случае отражает догадку владельца крейта относительно того, к какой из трех категорий относится текущий релиз. Не каждый понимает все это правильно, да и слово «правильно» здесь не имеет точного определения. И даже если вы все уясните верно, всегда остается вероятность столкнуться с законом Хайрума.

Как бы то ни было, SemVer — это единственный инструмент для тех, кому недоступна роскошь работать в среде наподобие тщательно протестированного внутреннего монорепозитория Google (<https://oreil.ly/i-WpN>). Поэтому для управления зависимостями понимание принципов работы этой системы и ее ограничений является обязательным.

Рекомендация 22. Минимизация видимости

Rust позволяет скрывать либо раскрывать элементы кода для других частей кодовой базы. Текущая рекомендация познакомит вас с предоставляемыми для этого механизмами, а также посоветует, где и когда их следует использовать.

Синтаксис для обозначения видимости

Базовой единицей видимости в Rust выступает модуль. По умолчанию элементы модуля (типы, методы, константы) являются *закрытыми* и доступными только для кода в этом модуле и его подмодулях.

Код, который должен быть доступен более широко, отмечается ключевым словом `pub`, что делает его открытым для некой другой области видимости. В большинстве синтаксических возможностей Rust обозначение кода как `pub` не раскрывает автоматически его содержимое — типы и функции в `pub mod`, как и поля в `pub struct`, не являются публичными. Тем не менее есть пара исключений, когда раскрытие видимости содержимого оказывается кстати.

- Обозначение `enum` как публичного автоматически делает открытыми и варианты этого типа (вместе со всеми полями, которые могут в них присутствовать).
- Обозначение `trait` как публичного автоматически делает публичными и его методы.

Поэтому коллекция типов в модуле:

```
pub mod somemodule {
    // Обозначение `struct` как публичной не делает публичными ее поля
    #[derive(Debug, Default)]
    pub struct AStruct {
        // По умолчанию поля недоступны
        count: i32,
        // Чтобы быть видимыми, поля должны быть явно обозначены как `pub`
        pub name: String,
    }

    // Аналогично методы в структуре требуют индивидуальных маркеров `pub`
    impl AStruct {
        // По умолчанию методы недоступны
        fn canonical_name(&self) -> String {
            self.name.to_lowercase()
        }
        // Чтобы сделать методы видимыми, их нужно явно обозначить как `pub`
        pub fn id(&self) -> String {
            format!("{}-{}", self.canonical_name(), self.count)
        }
    }

    // Обозначение `enum` как публичного делает публичными и все его варианты
    #[derive(Debug)]
    pub enum AnEnum {
        VariantOne,
        // Поля в вариантах также становятся публичными
        VariantTwo(u32),
        VariantThree { name: String, value: String },
    }

    // Обозначение `trait` как публичного делает публичными и все его методы
    pub trait DoSomething {
        fn do_something(&self, arg: i32);
    }
}
```

открывает доступ к публичным (`pub`) элементам и ранее обозначенным исключениям:

```
use somemodule::*;

let mut s = AStruct::default();
s.name = "Miles".to_string();
println!("s = {:?}", s, name='{}', id={}, s, s.name, s.id());
```

```
let e = AnEnum::VariantTwo(42);
println!("e = {e:?}");

#[derive(Default)]
pub struct DoesSomething;
impl DoesSomething for DoesSomething {
    fn do_something(&self, _arg: i32) {}
}

let d = DoesSomething::default();
d.do_something(42);
```

Однако непубличные элементы, как правило, остаются недоступными:

```
let mut s = AStruct::default();
s.name = "Miles".to_string();
println!("(inaccessible) s.count={}", s.count);
println!("(inaccessible) s.canonical_name()={}", s.canonical_name());

error[E0616]: field `count` of struct `somemodule::AStruct` is private
--> src/main.rs:230:45
|
230 |     println!("(inaccessible) s.count={}", s.count);
|                         ^^^^^^ private field
error[E0624]: method `canonical_name` is private
--> src/main.rs:231:56
|
86 |     fn canonical_name(&self) -> String {
|           ----- private method defined here
...
231 |     println!("(inaccessible) s.canonical_name()={}", s.canonical_name());
|                         ^^^^^^^^^^^^^^^^^ private method ^^^^^^^^^^^^^^
Some errors have detailed explanations: E0616, E0624.
For more information about an error, try `rustc --explain E0616`.
```

Самым распространенным маркером видимости выступает ключевое слово `pub`, которое делает элемент доступным для всего, что видит модуль, в котором он находится. И эта последняя деталь очень важна: если модуль `somemodule::somemodule` невидим для другого кода, то и любые его `pub`-элементы тоже будут невидимыми.

Есть и более специфичные варианты `pub`, которые позволяют ограничивать его область видимости. Перечислю их в нисходящем порядке по степени полезности.

`pub(crate)`

Доступен везде внутри крейта. Этот вариант особенно полезен для действующих в пределах крейта внутренних вспомогательных функций, которые не должны раскрываться для его внешних пользователей.

pub(super)

Доступен для модуля, являющегося родителем текущего, и его подмодулей. Такой вариант иногда пригождается для выборочного расширения видимости в крейте с глубокой структурой модулей. Это также фактический уровень видимости для модулей: простой `mod mymodule` видим для своего родительского модуля либо крейта и соответствующих подмодулей.

pub(in <path>)

Доступен для кода в `<path>`, который должен быть описанием некоего модуля, являющегося предком текущего. Иногда пригождается для организации исходного кода, так как позволяет перемещать часть функциональности в подмодули, не обязательно видимые в публичном API. Например, стандартная библиотека Rust объединяет все итераторные адаптеры во внутреннем подмодуле и содержит:

- маркер видимости `pub(in crate::iter)` для всех необходимых методов адаптера в подмодулях, таких как `std::iter::adapters::Map::new` (<https://oreil.ly/XbQGP>);
- публичное использование (`pub use`) всех типов `adapters::` во внешнем модуле (<https://oreil.ly/uLZ-d>).

pub(self)

Эквивалент `pub(in self)`, равнозначный отсутствию свойства `pub`. Примеры его использования очень туманны — например, сокращение числа особых случаев, необходимых в макросах генерации кода.

Компилятор Rust предупредит вас в случае наличия элемента кода, который приватен для модуля, но внутри него и его подмодулей не применяется:

```
pub mod anothermodule {
    // Приватная функция, не используемая внутри ее модуля
    fn inaccessible_fn(x: i32) -> i32 {
        x + 3
    }
}
```

И хотя предупреждение указывает, что код не используется в модуле, к которому принадлежит, на практике оно зачастую говорит о том, что код ввиду ограничений видимости *не может* быть задействован извне этого модуля:

```
warning: function `inaccessible_fn` is never used
--> src/main.rs:56:8
  |
56 |     fn inaccessible_fn(x: i32) -> i32 {
  |     ^^^^^^^^^^^^^^
  |
  = note: #[warn(dead_code)] on by default
```

Семантика видимости

Отдельно от вопроса о том, *как* расширить видимость, стоит вопрос о том, *когда* это делать. Общепринятым ответом на него будет: *как можно реже*, по крайней мере в случае кода, который может повторно применяться в будущем.

Первая причина появления такой рекомендации в том, что изменение видимости бывает трудно отменить. Как только элемент крейта становится публичным, его нельзя снова сделать приватным, не нарушив работу кода, который этот крейт использует. В связи с этим такие изменения ведут к повышению старшей версии (см. рекомендацию 21). Но в обратном направлении все иначе: смена статуса элемента с приватного на публичный обычно требует повышения лишь младшей версии, не влияя на работоспособность пользователей крейта. Почитайте руководства по совместимости Rust API (<https://oreil.ly/CkFWN>) и обратите внимание на то, сколько категорий версии затрагивается, когда в процессе участвуют `pub`-элементы.

Более важной, но менее явной причиной для предпочтения закрытости является то, что она оставляет для вас доступные возможности. Чем больше элементов раскрыто, тем больше всего должно оставаться фиксированным в будущем (ради отсутствия несовместимых изменений). Если вы раскроете внутренние детали реализации структуры данных, то впоследствии гипотетический переход на более эффективный алгоритм станет критическим изменением. Если вы раскроете внутренние вспомогательные функции, какой-то внешний код обязательно окажется зависимым от каких-то деталей этих функций.

Естественно, это касается только кода библиотек, которые намерены существовать долго и потенциально имеют множество пользователей. Но нет ничего более постоянного, чем временное решение, так что этот подход станет хорошей привычкой.

Также стоит отметить, что совет ограничивать видимость касается не только текущей рекомендации или Rust.

- Руководство по Rust API (<https://oreil.ly/nSrkD>) рекомендует: структуры должны содержать приватные поля.
- Книга *Effective Java*, 3-е издание (Addison-Wesley Professional)¹, содержит следующие рекомендации:
 - рекомендация 15 — минимизируйте доступность классов и их членов;
 - рекомендация 16 — в публичных классах используйте методы доступа, а не публичные поля.

¹ Блох Дж. Java, эффективное программирование. 3-е изд. — М., 2022.

- Второе издание книги *Effective C++* Скотта Мейерса (Scott Meyers) (Addison-Wesley Professional)¹ содержит такие рекомендации:
 - рекомендация 18 – старайтесь создавать полноценные и *минималистичные* интерфейсы классов;
 - рекомендация 20 – избегайте использования в публичном интерфейсе членов данных.

Рекомендация 23. Избегайте импорта с подстановкой

Инструкция `use` в Rust подтягивает указанный элемент из другого крейта или модуля и делает его имя доступным для использования в коде локального модуля без проверки на конфликты. *Импорт с подстановкой* (или *глобальный импорт*) в виде `use somecrate::module::*` гласит, что в локальное пространство имен должны быть добавлены *все* публичные символы из этого модуля.

Как говорилось в рекомендации 21, внешний крейт может добавлять в свой API новые элементы в рамках обновления младшей версии. Такое изменение считается обратно совместимым. Совместно эти два нюанса вызывают обеспокоенность тем, что некритическое изменение в зависимости может сломать код: «Что произойдет, если зависимость добавит новый символ, конфликтующий с уже используемым именем?»

На самом простом уровне это не создает проблемы: имена элементов, импортируемых с помощью подстановки, рассматриваются как имеющие меньший приоритет, поэтому, если в коде есть такие же имена, то предпочтение отдается им:

```
use bytes::*;

// Локальный тип `Bytes` не конфликтует с `bytes::Bytes`
struct Bytes(Vec<u8>);
```

К сожалению, в некоторых случаях конфликты все же могут происходить. Рассмотрим, к примеру, ситуацию, когда зависимость вносит новый трейт и реализует его для некоего типа:

```
trait BytesLeft {
    // Имя конфликтует с методом `remaining` в импортированном
    // с подстановкой трейте `bytes::Buf`
    fn remaining(&self) -> usize;
}
```

¹ Майерс С. Эффективное использование C++. 50 рекомендаций по улучшению ваших программ и проектов. — СПб., Питер, 2006.

```
impl BytesLeft for &[u8] {
    // Реализация конфликтует с `impl bytes::Buf for &[u8]`
    fn remaining(&self) -> usize {
        self.len()
    }
}
```

Если имена каких-либо методов из нового трейта конфликтуют с именами существующих методов, которые применяются к этому типу, то компилятор не может однозначно определить, какой из них предполагается.

НЕ КОМПИЛИРУЕТСЯ

```
let arr = [1u8, 2u8, 3u8];
let v = &arr[1..];

assert_eq!(v.remaining(), 2);
```

Об этом сообщает ошибка этапа компиляции:

```
error[E0034]: multiple applicable items in scope
--> src/main.rs:40:18
 |
40 |     assert_eq!(v.remaining(), 2);
|          ^^^^^^^^^^ multiple `remaining` found
|
note: candidate #1 is defined in an impl of the trait `BytesLeft` for the
      type `&[u8]`
--> src/main.rs:18:5
|
18 |     fn remaining(&self) -> usize {
|          ^^^^^^^^^^^^^^^^^^^^^^^^^^
= note: candidate #2 is defined in an impl of the trait `bytes::Buf` for the
      type `&[u8]`
help: disambiguate the method for candidate #1
|
40 |     assert_eq!(BytesLeft::remaining(&v), 2);
|          ~~~~~~
help: disambiguate the method for candidate #2
|
40 |     assert_eq!(bytes::Buf::remaining(&v), 2);
|          ~~~~~~
```

В результате следует избегать выполнения импорта с подстановкой из крейтов, которые вы не контролируете.

Если же вы контролируете источник импорта, то упомянутые соображения утрачивают актуальность. Например, модуль `test` нередко выполняет `use`

`super::*`;.

Это может касаться и крейтов, которые используют модули преимущественно для деления кода, чтобы выполнить импорт с подстановкой из внутреннего модуля:

```
mod thing;
pub use thing::*;


```

Есть и еще одно распространенное исключение, когда импорт с подстановкой имеет смысл. Некоторые крейты соблюдают соглашение о том, что типичные элементы из крейта повторно экспортируются из вводных модулей (*prelude module*), которые изначально задуманы для импорта с подстановкой:

```
use thing::prelude::*;


```

И хотя теоретически в этом случае актуальны те же проблемы, на практике такой модуль наверняка будет тщательно организован и повышенное удобство может перевесить небольшой риск возникновения проблем в будущем.

Наконец, если вы не будете следовать предложенным здесь советам, *подумайте о прикреплении зависимостей, которые импортируете с подстановкой, к конкретной версии* (подробнее — в рекомендации 21), чтобы повышение младшей версии зависимости не было автоматически разрешено.

Рекомендация 24. Повторный экспорт зависимостей, чьи типы присутствуют в вашем API

У этой рекомендации получилось слегка мудреное название, но проработка примеров все прояснит¹.

В рекомендации 25 описывается, как в Cargo реализована прозрачная поддержка различных версий одного библиотечного крейта, связанных в единый двоичный файл. Разберем такой файл, который использует крейт `rand`, в частности его версию 0.8:

```
# Файл Cargo.toml для двоичного файла верхнего уровня
[dependencies]
# Двоичный файл зависит от крейта `rand` из crates.io
rand = "=0.8.5"
# Он зависит и от другого крейта, (`dep-lib`)
dep-lib = "0.1.0"
```

¹ Этот пример, да и вся рекомендация вдохновлены подходом, используемым в крейтах RustCrypto (<https://oreil.ly/7w1iF>).

```
// Исходный код:  
let mut rng = rand::thread_rng(); // rand 0.8  
let max: usize = rng.gen_range(5..10);  
let choice = dep_lib::pick_number(max);
```

В последней строке кода в качестве еще одной зависимости используется условный крейт `dep-lib`. Он тоже может относиться к `crates.io` или же быть локальным и обнаруживаться с помощью механизма Cargo path.

Этот крейт `dep-lib` использует версию 0.7 крейта `rand`:

```
# Файл Cargo.toml для библиотечного крейта `dep-lib`  
[dependencies]  
# Эта библиотека зависит от крейта `rand` из crates.io  
rand = "=0.7.3"  
  
// Исходный код:  
/// Крейт `dep-lib` предоставляет функциональность выбора чисел  
use rand::Rng;  
  
/// Выбирает номер между 0 и n, не включая крайние значения  
pub fn pick_number(n: usize) -> usize {  
    rand::thread_rng().gen_range(0, n)  
}
```

Внимательный читатель может заметить различие между этими двумя примерами кода.

- В версии `rand` 0.7.x, используемой библиотечным крейтом `dep-lib`, метод `rand::gen_range()` получает два параметра, `low` и `high`.
- В версии `rand` 0.8.x, используемой двоичным крейтом, метод `rand::gen_range()` получает один параметр `range`.

Это не обратно совместимое изменение, а значит, у `rand` в соответствии с правилами семантического версионирования (см. рекомендацию 21) повышается крайний левый номер версии. Однако двоичный файл, совмещающий в себе две несовместимые версии, прекрасно работает — Cargo все улаживает.

При этом ситуация существенно усложняется, если API библиотечного крейта `dep-lib` раскрывает тип из своей зависимости, делая ее *публичной* (<https://oreil.ly/yAm3Y>).

Предположим, к примеру, что точка входа `dep-lib` содержит элемент `Rng` — конкретно его версию 0.7:

```
/// Выбирает номер между 0 и n (без крайних значений),  
/// используя предоставленный экземпляр `Rng`  
pub fn pick_number_with<R: Rng>(rng: &mut R, n: usize) -> usize {  
    rng.gen_range(0, n) // Метод из Rng версии 0.7.x  
}
```

Причем нужно хорошо подумать, прежде чем использовать в своем API типы из другого крейта, так как это тесно привяжет ваш крейт к крейту зависимости. Например, повышение старшей версии зависимости (см. рекомендацию 21) будет автоматически требовать повышения старшей версии вашего крейта.

В данном случае `rand` — это полустандартный крейт, который широко используется и подтягивает лишь ограниченное число собственных зависимостей (см. рекомендацию 25), так что включение его типов в API крейта с позиции баланса станет вполне допустимым.

Вернемся к примеру. Попытка использовать эту точку входа из двоичного файла верхнего уровня проваливается.

НЕ КОМПИЛИРУЕТСЯ

```
let mut rng = rand::thread_rng();
let max: usize = rng.gen_range(5..10);
let choice = dep_lib::pick_number_with(&mut rng, max);
```

И сообщение компилятора здесь не особо помогает, что для Rust довольно необычно:

```
error[E0277]: the trait bound `ThreadRng: rand_core::RngCore` is not satisfied
--> src/main.rs:22:44
22 |     let choice = dep_lib::pick_number_with(&mut rng, max);
   |     ----- ^^^^^^^^ the trait
   |         | `rand_core::RngCore` is not
   |         | implemented for `ThreadRng`
   |
   |         required by a bound introduced by this call
|
= help: the following other types implement trait `rand_core::RngCore`: &'a mut R
```

Анализ причастных типов вызывает путаницу, так как необходимые трейты *вроде бы* реализованы, но вызывающий код автоматически реализует условный `RngCore_v0_8_5`, а библиотека ожидает `RngCore_v0_7_3`.

Расшифровав наконец сообщение об ошибке, вы поймете, что внутренней причиной является конфликт версий, но как его исправить?¹ Нужно понять главное: хотя бинарный файл и не может *напрямую* использовать две разные

¹ Такая ошибка может появиться, даже когда в графе зависимостей присутствуют две альтернативы для крейта с одной версией. В этом случае какой-то элемент графа сборки вместо адреса crates.io использует поле path для указания локального каталога.

версии одного крейта, он может делать это *косвенно* (как в изначальном примере ранее).

С точки зрения автора этого файла, проблему можно обойти, добавив промежуточный оберточный крейт, скрывающий прямое использование типов `rand v0.7`. Оберточный крейт отличается от бинарного, а значит, может зависеть от `rand v0.7`, не пересекаясь с зависимостью двоичного крейта от `rand v0.8`.

Получается нескладно, и автору этого библиотечного крейта доступен гораздо более эффективный подход. Он может упростить жизнь пользователям, явным образом повторно экспортав что-то одно:

- типы, причастные к API;
- весь крейт зависимости.

Для этого примера больше подойдет второй вариант — он делает доступными не только версию `Rng 0.7` и типы `RngCore`, но и методы наподобие `thread_rng()`, которые создают экземпляры этого типа:

```
// Повторный экспорт версии `rand`, используемой в API этого крейта
pub use rand;
```

Теперь у вызывающего кода есть другой способ напрямую обращаться к `rand v0.7`, используя `dep_lib::rand`:

```
let mut prev_rng = dep_lib::rand::thread_rng(); // экземпляр v0.7 Rng
let choice = dep_lib::pick_number_with(&mut prev_rng, max);
```

После ознакомления с этим примером название текущей рекомендации должно проясниться: *повторно экспортавте зависимости, чьи типы присутствуют в вашем API*.

Рекомендация 25. Управляйте графиком ваших зависимостей

Как и почти все современные языки программирования, Rust упрощает подтягивание внешних библиотек в виде *крайтов*. Большинство нетипичных программ Rust используют внешние крейты, которые сами имеют дополнительные зависимости, формирующие *граф* для всей программы.

По умолчанию Cargo скачивает все крейты, указанные в разделе `[dependencies]` файла `Cargo.toml` с `crates.io`, отыскивая такие их версии, которые отвечают требованиям, установленным в том же файле.

Но этот, казалось бы, простой процесс скрывает в себе некоторые тонкости. Первая состоит в том, что имена крейтов из `crates.io` формируют единое плоское

пространство имен — и это глобальное пространство пересекается с именами функциональности в крейте (подробнее — в рекомендации 26)¹.

Если вы планируете публиковать крейт на `crates.io`, то имейте в виду, что имена обычно выделяются по принципу «первым пришел, первым получил». То есть вы можете обнаружить, что предпочтительное имя для вашего публичного крейта уже занято. При этом сквоттинг имен — резервирование имени крейта путем предварительной регистрации его пустой версии — не приветствуется (<https://oreil.ly/ArljK>), если только вы не планируете в ближайшем будущем выпустить этот крейт с кодом.

Есть здесь и некоторые нюансы. Первый связан с небольшим различием между тем, что допустимо использовать в качестве имени крейта в пространстве имен, а что допустимо применять в качестве соответствующего идентификатора в коде: крейт можно назвать `some-crate`, но в коде он будет представлен как `some_crate` (с нижним подчеркиванием). Иными словами, если вы видите в коде `some_crate`, то именем соответствующего крейта может быть `some-crate` либо `some_crate`.

Второй важный нюанс касается того, что Cargo позволяет присутствовать в сборке нескольким версиям одного крейта, соответствующим правилам SemVer. Это немного удивляет, поскольку каждый файл `Cargo.toml` может иметь лишь одну версию любой конкретной зависимости, но такая ситуация часто возникает в случае косвенных зависимостей, когда ваш крейт зависит от `some-crate` версии 3.x, а также от `older-crate`, который, в свою очередь, зависит от `some-crate` версии 1.x.

И это может вызывать путаницу, если зависимость раскрывается не только для внутреннего использования, но и как-то иначе (см. рекомендацию 24) — компилятор будет рассматривать эти две версии как отдельные крейты, но его сообщения об ошибках не обязательно это прояснят.

Вероятность присутствия нескольких версий крейта также может создать проблемы, если крейт включает код C/C++, проанализированный механизмами FFI (см. рекомендацию 34). Набор инструментов Rust может распознать отдельные версии кода Rust, но любой включенный код C/C++ подвергается *правилу одного определения*: «Может присутствовать лишь одна версия любой функции, константы или глобальной переменной».

Поддержка в Cargo нескольких версий имеет некоторые ограничения. Cargo не позволяет использовать несколько версий одного крейта в рамках допустимого в соответствии с SemVer диапазона (см. рекомендацию 21):

- `some-crate` 1.2 и `some-crate` 3.1 могут сосуществовать;
- `some-crate` 1.2 и `some-crate` 1.3 — не могут.

¹ Можно также настроить альтернативный реестр крейтов, например внутренний корпоративный реестр. Тогда в каждой записи о зависимости в `Cargo.toml` можно будет использовать ключ `registry` для указания того, из какого реестра нужно получать зависимость.

Cargo также расширяет правила семантического версионирования для крейтов, не достигших версии 1.0, чтобы первая ненулевая подверсия считалась старшей, в результате чего названное ограничение работает и в их отношении:

- `other-crate` 0.1.2 и `other-crate` 0.2.0 могут сосуществовать;
- `other-crate` 0.1.2 и `other-crate` 0.1.4 — не могут.

Заложенный в Cargo алгоритм выбора версий (<https://oreil.ly/qIxXh>) определяет, какие версии нужно включить. Каждая строка зависимости в `Cargo.toml` указывает допустимый диапазон версий в соответствии с правилами версионирования, и Cargo учитывает это, когда один и тот же крейт встречается в нескольких частях графа зависимостей. Если допустимые диапазоны накладываются и отвечают требованиям SemVer, тогда Cargo по умолчанию будет выбирать последнюю версию крейта из наложившегося диапазона. Если же отвечающего требованиям SemVer наложения нет, то Cargo создаст несколько копий зависимости с разными версиями.

После того как Cargo выбрал допустимые версии для всех зависимостей, его выбор фиксируется в файле `Cargo.lock`. Теперь все последующие сборки будут использовать этот прописанный в `Cargo.lock` выбор, чтобы получаться стабильными и не требовать дополнительного скачивания.

Это оставляет вам выбор: отправлять файлы `Cargo.lock` в систему контроля версий или нет. Разработчики Cargo советуют следующее (<https://oreil.ly/pppkQ>).

- Компоненты, создающие итоговый продукт, а именно приложения и двоичные файлы, должны коммитить `Cargo.lock` для обеспечения детерминированной сборки.
- Библиотечные крейты *не* должны коммитить файл `Cargo.lock`, так как для любых нижестоящих потребителей библиотеки он неважен — у них будет собственный файл `Cargo.lock`. *Имейте в виду, что файл `Cargo.lock` для библиотечного крейта пользователями библиотеки игнорируется.*

Даже библиотечному крейту бывает полезно иметь загруженный в систему контроля версий файл `Cargo.lock`, чтобы регулярные сборки и система непрерывной интеграции (см. рекомендацию 32) работали со стабильной, а не изменчивой версией. И хотя гарантии системы семантического версионирования (см. рекомендацию 21) в теории должны предотвращать сбои, на практике ошибки случаются, и очень неприятно получать сборки, которые не работают из-за того, что кто-то изменил зависимость зависимости.

Тем не менее, *если вы версионируете `Cargo.lock`, настройте процесс для обработки его обновлений*, например, предоставляемый GitHub Dependabot. Если вы этого не сделаете, ваши зависимости останутся закрепленными за версиями, которые стареют, утрачивая свою актуальность и потенциально безопасность.

Закрепление версий с загруженным файлом `Cargo.lock` не исключает боли с обработкой обновлений зависимостей, но это означает, что вы сможете

обрабатывать их в удобное для вас время, а не сразу после изменения выше-стоящего крейта. Есть еще некоторые проблемы с обновлением зависимостей, которые исчезают сами по себе: крейт, который выпускается с ошибкой, зачастую сопровождается второй, исправленной версией, которая выходит чуть позже, и пакетный процесс апгрейда версий может увидеть только последнюю.

Третий нюанс при разрешении версий в Cargo заключается в *унификации функциональности*: функциональность, которая активируется для зависимого крейта, является *объединением* функциональности, выбранный разными участками в графе зависимостей. Подробнее читайте в рекомендации 26.

Указание диапазона допустимых версий

Текущий раздел посвящен определению диапазона допустимых версий в соответствии с правилами, прописанными в книге по Cargo (<https://oreil.ly/9YHm->).

Избегайте слишком конкретной зависимости от версии

Придерживаться одной конкретной версии ("=1.2.3") *обычно* нежелательно: в этом случае вы не видите более новые версии, потенциально включающие исправления безопасности, и существенно сужаете возможный диапазон пересечения с другими крейтами графа, опирающимися на ту же зависимость (напомню, что Cargo допускает использование лишь одной версии крейта в рамках диапазона, соответствующего SemVer). Если вы хотите сделать так, чтобы ваши сборки задействовали согласованный набор зависимостей, реализовать это поможет файл `Cargo.lock`.

Избегайте слишком большой зависимости от слишком обобщенной версии

Можно указывать спецификатор ("*"), который означает возможность использования *любой* зависимости, но это тоже плохая идея. Если выйдет новая старшая версия крейта зависимости, кардинально меняющая все аспекты его API, ваш код вряд ли продолжит работать после того, как `cargo update` подтянет эту новую версию.

Самый распространенный оптимальный вариант — не слишком точный и не слишком обобщенный — допустить использование SemVer-совместимых версий крейта ("1"), возможно, с конкретной минимальной версией, включающей функциональность или исправление, которое вам нужно ("1.4.23"). Оба этих варианта указания допустимых версий реализуются с помощью стандартного поведения Cargo, которое выражается в разрешении версий, совместимых с указанной в соответствии с SemVer. Это можно сделать более явным, добавив знак вставки.

- Версия "1" равнозначна "^1", то есть допускает все версии 1.x, а значит, равноцenna записи "1.*".
- Версия "1.4.23" равнозначна "^1.4.23", то есть допускает любые версии 1.x старше 1.4.23.

Решение проблем с помощью инструментов

Рекомендация 31 советует использовать набор инструментов, имеющийся в экосистеме Rust. В текущем разделе описываются проблемы графа зависимостей, при решении которых эти инструменты способны помочь.

Компилятор не мешкая сообщит вам, если вы задействуете в коде зависимость, которую не включили в `Cargo.toml`. Но что насчет обратного случая? Если в `Cargo.toml` есть зависимость, которую вы в своем коде *не используете* или, скорее, *больше* не используете, тогда Cargo вмешиваться не станет. Для решения такой проблемы существует инструмент `cargo-udeps`: он предупреждает, когда в файле `Cargo.toml` присутствует неиспользуемая зависимость (`udep`).

Есть и более гибкий инструмент `cargo-deny`, который анализирует ваш граф зависимостей на предмет различных проблем в переходных зависимостях:

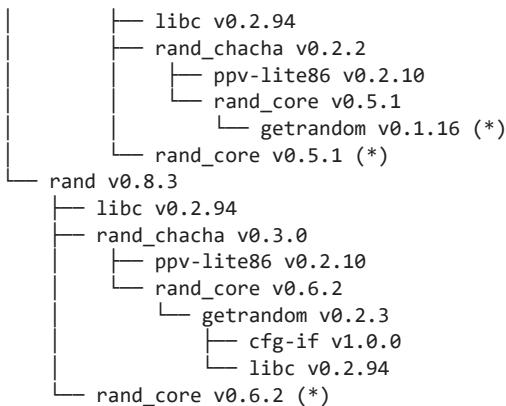
- имеющих известные проблемы безопасности во включенной версии;
- находящихся под неприемлемой лицензией;
- которые сами по себе неприемлемы;
- которые включены в несколько версий древа зависимостей.

Каждый из этих аспектов можно настроить, указав исключения. Механизм исключений обычно необходим для более крупных проектов, в частности, для предупреждения о присутствии нескольких версий: по мере роста графа зависимостей увеличивается и вероятность появления переходной зависимости от разных версий одного крейта. Есть смысл попытаться уменьшить число таких повторов, где это возможно, как минимум по причинам двоичного размера и этапа компиляции, но порой нет такой комбинации версий зависимостей, которая исключала бы повтор.

Эти инструменты можно использовать как одноразовые, но лучше наладить их регулярное и надежное выполнение, включив в систему непрерывной интеграции (см. рекомендацию 32). Это поможет перехватывать возникающие в выше-стоящей зависимости проблемы, включая такие, которые могли быть внесены извне кода (например, недавно обнародованную уязвимость).

Если один из этих инструментов сообщит о проблеме, может быть сложно выяснить, где конкретно в графе зависимостей она возникает. И здесь помогает команда `cargo tree`, которая показывает этот график в виде древовидной структуры:

```
dep-graph v0.1.0
└── dep-lib v0.1.0
    └── rand v0.7.3
        ├── getrandom v0.1.16
        │   ├── cfg-if v1.0.0
        │   └── libc v0.2.94
```



Для решения конкретных задач в `cargo tree` есть несколько опций, например:

`--invert`

Показывает, что зависит *от* конкретного пакета, помогая сосредоточиться на конкретной проблемной зависимости.

`--edges features`

Показывает, какие функции крейта активируются зависимой связью. Помогает определить, что происходит с унификацией функциональности (см. рекомендацию 26).

`--duplicates`

Показывает крейты, имеющие в графе зависимостей несколько версий.

От чего зависеть

В предыдущих разделах освещалась преимущественно механическая сторона работы с зависимостями, но есть и более философский, а значит, и более сложный вопрос: «Когда следует включать зависимость?»

Чаще всего рассуждать особо не приходится: если вам нужна функциональность крейта, то придется от него зависеть либо самому писать эту функциональность¹.

Однако любая зависимость имеет свою цену, отчасти в плане увеличения времени сборки и размера двоичных файлов, но в основном в плане усилий разработчика, необходимых для исправления возникающих проблем.

Чем шире график ваших зависимостей, тем выше вероятность того, что у вас появятся подобные проблемы. Экосистема крейтов Rust так же уязвима для

¹ Если вы нацелены на среду `no_std`, этот выбор может быть сделан за вас: многие крейты несовместимы с `no_std`, в частности, если также недоступна `alloc` (см. рекомендацию 33).

неожиданных проблем с зависимостями, как и другие экосистемы пакетов. Как показывает история, удаление пакета одним разработчиком (<https://oreil.ly/8lTZ8>) или исправление командой проблем с лицензией пакета (<https://oreil.ly/7HjSi>) может дать масштабный обратный эффект.

Но еще большее беспокойство вызывают атаки на цепочки поставок, когда злоумышленник намеренно пытается саботировать популярные зависимости путем тайпсквоттинга, захвата аккаунта специалиста по обслуживанию (<https://oreil.ly/b8D-f>) или с помощью других более сложных атак.

Подобная атака не просто влияет на скомпилированный вами код — имейте в виду, что зависимость может выполнять произвольный код на этапе *сборки* посредством скриптов `build.rs` или процедурных макросов (см. рекомендацию 28). Это означает, что в случае взлома в нее могут внедрить, к примеру, майнер криптовалют, который будет работать в рамках вашей системы непрерывного развертывания. Так что при желании добавить более «косметические» зависимости иногда стоит оценить оправданность этого действия.

Обычно ответом будет «да». В конечном итоге время, затраченное на решение проблем с зависимостями, оказывается намного меньше времени, которое потребуется для написания аналогичной функциональности с нуля.

Запомните

- Имена крейтов на `crates.io` формируют единое одноуровневое пространство имен, общее с названиями элементов функциональности.
- Имена крейтов могут включать дефис, но в коде будут отображаться с нижним подчеркиванием.
- Cargo поддерживает несколько версий одного крейта в графе зависимостей, но только если они относятся к разным не-SemVer-совместимым версиям. В случае крейтов, включающих код FFI, это может создать проблемы.
- Страйтесь допускать SemVer-совместимые версии зависимостей ("1" или "1.4.23", чтобы включать минимальную версию).
- Используйте файлы `Cargo.lock`, чтобы обеспечить повторяемость своих сборок, но помните, что файл `Cargo.lock` с публикуемым крейтом не предоставляется.
- Применяйте инструменты (`cargo tree`, `cargo deny`, `cargo udep` и т. п.), чтобы находить и исправлять проблемы с зависимостями.
- Поймите, что подтягивание зависимостей сокращает объем написания кода, но не дается даром.

Рекомендация 26. Остерегайтесь расположения feature

Rust позволяет одной и той же базе кода поддерживать различные конфигурации посредством механизма *feature*, созданного поверх базового механизма для условной компиляции. Механизм *feature* имеет некоторые тонкости, которые стоит иметь в виду. О них и пойдет речь в текущей рекомендации.

Условная компиляция

Rust включает поддержку условной компиляции, регулируемой атрибутами `cfg` и `cfg_attr`. Эти атрибуты управляют тем, будет ли элемент, к которому они прикрепляются, — функция, строка, блок кода или что-то иное — включен в компилируемый исходный код (что отличается от механизма препроцессора в C/C++, основанного на обработке строк). Такое условное включение управляется опциями конфигурации, которые представляют либо простые имена (например, `test`), либо имена и значения (например, `panic = "abort"`).

Заметьте, что опции конфигурации, отражающие имя/значение, допускают использование более одного значения для одного имени:

```
// Сборка с установкой `RUSTFLAGS` на:
//   '--cfg myname="a" --cfg myname="b"'
#[cfg(myname = "a")]
println!("cfg(myname = 'a') is set"); #
#[cfg(myname = "b")]
println!("cfg(myname = 'b') is set");

cfg(myname = 'a') is set
cfg(myname = 'b') is set
```

Помимо значений *feature*, описываемых в этом разделе, чаще всего в конфигурации используются такие, которые инструменты устанавливают автоматически, описывая целевую среду сборки. К ним относятся операционная система (`target_os`), архитектура процессора (`target_arch`), размер указателя (`target_pointer_width`) и порядок байтов (`target_endian`). Это обеспечивает портируемость кода, когда функциональность, специфичная для конкретной цели, включается в компиляцию только при сборке для этой цели.

Стандартная опция `target_has_atomic` также предоставляет пример использования множества значений в опциях конфигурации: и `[cfg(target_has_atomic = "32")]`, и `[cfg(target_has_atomic = "64")]` будут установлены для целей, которые поддерживают 32- и 64-битные атомарные операции. Подробнее об атомарных операциях читайте в главе 2 книги Мары Боса *Rust Atomics and Locks* (O'Reilly).

Features

Пакетный менеджер Cargo строится на основе этого механизма обработки имен/значений `cfg`, предоставляя `features`: именованные выборочные элементы функциональности крейта, которые можно включать при его создании. Cargo обеспечивает, чтобы опция `feature` заполнялась всеми установленными значениями для каждого компилируемого крейта. При этом устанавливаемые значения действуют в отношении конкретного крейта.

Вот функциональность Cargo: для компилятора Rust `feature` — это просто еще одна опция конфигурации.

На момент написания книги самый надежный способ определения доступных функций крейта — это проверка его файла манифеста `Cargo.toml`. Например, приведенный далее фрагмент такого файла включает *шесть* функций:

```
[features]
default = ["featureA"]
featureA = []
featureB = []
# Включение `featureAB` ведет к активации `featureA` и `featureB`
featureAB = ["featureA", "featureB"]
schema = []

[dependencies]
rand = { version = "^0.8", optional = true }
hex = "^0.4"
```

Учитывая, что в блоке `[features]` приведено всего пять записей, здесь определенно есть пара нюансов, на которые нужно обратить внимание.

Первый заключается в том, что строка `default` в блоке `[features]` представляет особое имя функциональности, сообщающее Cargo, какие функции должны быть включены по умолчанию. Эти функции также можно отключить, передав в команду сборки флаг `--no-default-features`, и потребитель крейта может закодировать их в своем файле `Cargo.toml` так:

```
[dependencies]
somecrate = { version = "^0.3", default-features = false }
```

Тем не менее `default` все равно считается именем функциональности, которое можно протестировать в коде:

```
#[cfg(feature = "default")]
println!("This crate was built with the \"default\" feature enabled.");
#[cfg(not(feature = "default"))]
println!("This crate was built with the \"default\" feature disabled.");
```

Второй нюанс определений функциональности скрыт в разделе `[dependencies]` изначального примера `Cargo.toml`: крейт `rand` — это зависимость, которая

отмечена как `optional = true`, что фактически вносит "rand" в имя функциональности¹. Если крейт компилируется с использованием `--features rand`, тогда эта зависимость активируется:

```
#[cfg(feature = "rand")]
pub fn pick_a_number() -> u8 {
    rand::random::<u8>()
}

#[cfg(not(feature = "rand"))]
pub fn pick_a_number() -> u8 {
    4 // Выбирается случайно
}
```

Это также означает, что *имена крейтов и функций используют общее пространство имен*, несмотря на то что оно обычно глобальное (к тому же обычно управляет ресурсом `crates.io`) и локально для соответствующего крейта. Поэтому нужно *осторожно выбирать названия для функций*, чтобы избежать конфликта с именами крейтов, которые могут оказаться в числе потенциальных зависимостей. Конфликта можно избежать, поскольку в Cargo есть механизм, который позволяет переименовывать импортируемые крейты (ключ `package`), но проще будет, если этого делать не придется.

Так что вы можете *определять функциональность крейта, просматривая [features] и optional [dependencies]* в файле крейта `Cargo.toml`. Чтобы включить нужную функцию зависимости, добавьте опцию `features` на соответствующую строку в блоке `[dependencies]` файла манифеста:

```
[dependencies]
somecrate = { version = "^0.3", features = ["featureA", "rand"] }
```

Эта строка обеспечит создание `somecrate` с включением `featureA` и функции `rand`. Тем не менее вследствие принципа под названием «*унификация функциональности*», помимо этих функций, могут оказаться включены и другие. Это означает, что крейт будет собран с *объединением* всех функций, запрошенных всеми компонентами графа зависимостей. Иными словами, если некая иная зависимость в графе опирается также на `somecrate`, но в нем включена только `feature`, тогда этот крейт будет собран с включением `featureA`, `featureB` и `rand`, удовлетворяя все запросы². Тот же принцип относится и к предустановленным функциям: если ваш крейт устанавливает для зависимости `default-features = false`, но некий другой компонент графа сборки оставляет включенными предустановленные функции, то они таковыми и останутся.

¹ Это предустановленное поведение можно отключить, используя ссылку "dep:<crate>" в другом месте блока `features`. Подробнее читайте в документации (<https://oreil.ly/HnLOJ>).

² Команда `cargo tree --edges features` поможет определить, какие функции для каких крейтов включены и почему.

Унификация функциональности означает, что *функции должны быть аддитивными*. Нежелательно применять взаимно несовместимые, поскольку ничто не мешает разным пользователям включить такие функции одновременно.

К примеру, если крейт делает `struct` и ее поля публичными, нежелательно делать эти поля функционально зависимыми.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
/// Структура, чье содержимое является публичным,
/// чтобы внешние пользователи могли создавать ее экземпляры
#[derive(Debug)]
pub struct ExposedStruct {
    pub data: Vec<u8>,

    /// Дополнительные данные, необходимые,
    /// только когда включена функция `schema`
    #[cfg(feature = "schema")]
    pub schema: String,
}
```

Пользователь крейта, который пытается создать экземпляр `struct`, оказывается в затруднении: нужно ли ему заполнять поле `schema`? Для ответа на этот вопрос можно *попробовать* добавить соответствующую функцию в файл `Cargo.toml` пользователя:

```
[features]
# Здесь функция `use-schema` включает функцию `schema` в `somecrate`
# (в этом примере для ясности используются разные имена функций.
# В реальном коде одно имя наверняка будет повторно задействоваться
# в обоих местах)
use-schema = ["somecrate/schema"]
```

Это сделает создание `struct` зависимым от данной функции.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
let s = somecrate::ExposedStruct {
    data: vec![0x82, 0x01, 0x01],

    // Заполняем поле, только если запрошена
    // активация `somecrate/schema`
    #[cfg(feature = "use_schema")]
    schema: "[int int]",
};
```

Тем не менее это не охватывает все конечные исходы: код не сможет скомпилироваться, если этот его фрагмент не активирует `somewhere/schema`, но некая другая переходная зависимость это сделает. Суть проблемы в том, что проверка статуса функции доступна только для крейта, который ее содержит. Пользователь крейта никак не сможет определить, включил Cargo `somewhere/schema` или нет. Так что следует *избегать привязывания (feature gating) публичных полей структур к функциональности*.

То же касается публичных трейтов, предназначенных для применения вне крейта, в котором они определены. Рассмотрим трейт, имеющий функциональную привязку к одному из своих методов.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
/// Трейт для элементов, поддерживающих сериализацию в формат CBOR
pub trait AsCbor: Sized {
    /// Преобразование элемента в данные CBOR
    fn serialize(&self) -> Result<Vec<u8>, Error>;

    /// Создание экземпляра элемента из сериализованных данных
    fn deserialize(data: &[u8]) -> Result<Self, Error>;

    /// Возврат схемы, соответствующей этому элементу
    #[cfg(feature = "schema")]
    fn cddl(&self) -> String;
}
```

Те, кто будет реализовывать этот трейт вовне, снова столкнутся с дилеммой: «Нужно ли реализовывать метод `cddl(&self)?`» Внешний код, который пытается реализовать этот трейт, не знает и не может определить, нужно ли ему воссоздавать функционально обусловленный метод.

То есть суть в том, что вам следует *избегать функциональной привязки методов в публичных трейтах*. Частичным исключением здесь может стать метод трейта с предустановленной реализацией (см. рекомендацию 13), но только если для внешнего кода не будет смысла переопределять эту предустановку в будущем.

Унификация функциональности также означает, что если ваш крейт имеет N независимых функций¹, то на практике могут проявиться все 2^N возможных комбинаций. Для исключения нежелательных сюрпризов нужно обеспечить, чтобы ваша система непрерывной интеграции (см. рекомендацию 32) охватывала все эти 2^N комбинаций во всех возможных вариантах тестов (см. рекомендацию 30).

¹ Функции могут активировать другие функции. В изначальном примере функция `featureAB` приводит к включению `featureA` и `featureB`.

Тем не менее использование optionalных функций очень помогает контролировать доступность для расширенного графа зависимостей (см. рекомендацию 25). Это особенно полезно в низкоуровневых крейтах, которые можно задействовать в среде `no_std` (см. рекомендацию 33), — часто бывает так, что присутствует компонент `std` или `alloc`, который активирует функциональность, опирающуюся на эти библиотеки.

Запомните

- Имена функций пересекаются с именами зависимостей.
- Имена функций нужно выбирать осторожно, чтобы они не конфликтовали с именами потенциальных зависимостей.
- Функции должны быть аддитивными.
- Избегайте функциональной привязки полей в `struct` или методов трейтов.
- Наличие множества независимых функций способно вызвать комбинаторный взрыв числа различных конфигураций сборки.

ГЛАВА 5

Инструменты

Титус Уинтерс (Titus Winters), ведущий специалист по библиотеке C++ в Google, описывает программную инженерию как программирование, обобщаемое с течением времени, а иногда как программирование, обобщаемое с течением времени и с текучкой людей. По прошествии длительных промежутков времени и если работает большая команда, кодовая база несет в себе уже не просто код.

Современные языки, включая Rust, об этом знают и сопровождаются экосистемой инструментов, которая выходит за пределы простого конвертирования программы в исполняемый двоичный код (компиляции).

Текущая глава посвящена изучению инструментария Rust, и главной рекомендацией здесь будет такая: используйте всю возможную инфраструктуру. Очевидно, что при этом не стоит впадать в крайности — настройка непрерывной интеграции (continuous integration, CI), сборка документации и шесть видов тестов станут перебором для проходной программы, которая будет запущена всего дважды. Но для большинства вещей, описанных в этой главе, можно добиться существенной отдачи: небольшое вложение в интеграцию инструментов окупится преимуществами, которые она несет.

Рекомендация 27. Документируйте публичные интерфейсы

Если ваш крейт будет использоваться другими программистами, то хорошо бы добавить к его содержимому документацию, в частности, для его публичного API. Если ваш крейт представляет собой не просто некий временный код, то к этим другим программистам относитесь и вы в будущем, когда забудете детали своего нынешнего кода.

Эта рекомендация касается не только Rust, как не является и чем-то новым: например, в книге «Эффективный Java», 2-е издание, есть рекомендация 44: «Пишите комментарии для всех публичных элементов API».

Какие особенности имеет написание комментариев в Rust? Они пишутся в формате Markdown и отделяются от кода символами `///` или `//`, о которых говорится в документации к Rust (https://oreil.ly/_WGEv). Вот пример:

```
/// Вычисляет [ `BoundingBox` ], которая включает ровно
/// пару объектов [ `BoundingBox` ]
pub fn union(a: &BoundingBox, b: &BoundingBox) -> BoundingBox {
    // ...
}
```

Этот формат написания имеет ряд деталей, которые стоит обозначить.

Используйте для кода соответствующий шрифт

Все, что будет вводиться в исходный код как есть, заключайте в обратные кавычки, чтобы итоговая документация имела фиксированную ширину, отчетливо демонстрируя различие между кодом и текстом.

Обильно добавляйте перекрестные ссылки

Добавляйте ссылку Markdown для всего, что может предоставлять контекст читателю документации. В частности, давайте *перекрестные ссылки на идентификаторы* с помощью удобного синтаксиса [`SomeThing`]: если в области видимости есть `SomeThing`, тогда в итоговой документации будет присутствовать гиперссылка на правильное место.

Подумайте о добавлении примеров кода

Если способ использования точки входа неочевиден, нeliшним будет добавить раздел `# Примеры` с образцом кода. Имейте в виду, что образец кода в комментариях документации компилируется и исполняется при выполнении `cargo test` (подробнее — в рекомендации 30), что помогает ему сохранять согласованность с демонстрируемым им кодом.

Документируйте места возникновения паники и ограничения для unsafe-кода

Если есть входные данные, которые вызывают панику функции, документируйте (в разделе `# Паники`) предварительные условия, которые необходимы для того, чтобы избежать `panic!`. Аналогичным образом документируйте (в разделе `# Безопасность`) требования для `unsafe`-кода.

Документация для стандартной библиотеки Rust — прекрасный пример, на глядно демонстрирующий все эти детали.

Инструменты

Формат Markdown, используемый для документирования, обеспечивает элегантный вид на выходе, но также подразумевает явный этап преобразования (`cargo doc`). Это, в свою очередь, создает вероятность возникновения попутных неполадок.

Простейшим советом в данном случае будет просто читать отображаемую документацию после ее написания, выполнив `cargo doc --open` (или `cargo doc --no-deps --open` для ограничения генерируемой документации текущим крейтом).

Также можно проверить валидность всех сгенерированных гиперссылок. Но эта задача больше подходит для машины и выполняется посредством атрибута крейта `broken_intra_doc_links`¹.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
#![deny(broken_intra_doc_links)]  
  
/// Ограничивающая рамка для [ `Polygone` ]  
#[derive(Clone, Debug)]  
pub struct BoundingBox {  
    // ...  
}
```

После активации этого атрибута `cargo doc` начнет обнаруживать невалидные ссылки:

```
error: unresolved link to `Polygone`  
--> docs/src/main.rs:4:30  
|  
4 |     /// Ограничивающая рамка для [ `Polygone` ]  
|           ^^^^^^^^^ no item named `Polygone` in scope  
|
```

Кроме того, вы можете требовать документацию, включив для крейта атрибут `#![warn(missing_docs)]`. После его включения компилятор будет выдавать предупреждение для каждого незадокументированного публичного элемента. Но при этом есть риск того, что включение данной опции вызовет появление некачественных комментариев, написанных наспех, только чтобы успокоить компилятор (подробнее об этом чуть позже).

Как обычно, любой инструмент, который обнаруживает потенциальные проблемы, должен быть частью вашей системы CI (см. рекомендацию 32) для перехвата любых недочетов, которые могут прокрасться в код.

Расположение дополнительной документации

Вывод `cargo doc` — основное, но не единственное место документирования вашего крейта. Как работать с кодом, пользователям помогут понять и другие части проекта Cargo.

¹ Традиционно эта опция называлась `intra_doc_link_resolution_failure`.

Подкаталог `examples/` этого проекта может содержать код для отдельных двоичных файлов, которые задействуют ваш крейт. Эти программы создаются и выполняются во многом аналогично интеграционным тестам (см. рекомендацию 30), но предназначены конкретно для хранения образцов кода, иллюстрирующих правильное использование интерфейса крейта.

Кроме того, имейте в виду, что интеграционные тесты в подкаталоге `tests/` также могут служить примерами для испытывающего трудности пользователя, несмотря на то что их основная цель — тестировать внешний интерфейс крейта.

Где публикуется документация крейта

Если вы хотите опубликовать крейт на `crates.io`, то сопутствующая документация будет доступна на `docs.rs` — это официальный проект Rust по сборке и размещению документации для публикуемых крейтов.

Имейте в виду, что `crates.io` и `docs.rs` предназначены для немного разной аудитории: первый нацелен на людей, которые выбирают, какой крейт им использовать, а второй — для тех, кто изучает уже добавленный крейт (хотя эти аудитории, очевидно, будут преимущественно совпадать).

В итоге на домашней странице крейта в каждой области свое содержимое.

`docs.rs`

Показывает страницу верхнего уровня из вывода `cargo doc`, сгенерированную на основе комментариев `//!` в файле `src/lib.rs` верхнего уровня.

`crates.io`

Показывает содержимое любого файла `README.md` верхнего уровня, добавленного в репозиторий проекта¹.

Что не нужно документировать

Когда проект *требует* включения документации для всех публичных элементов (о чем говорилось в первом разделе), по недосмотру можно запросто создать такую, которая только впустую потратит ценные пиксели. Сообщения компилятора об отсутствии документирующих комментариев лишь подразумевают, что вам реально нужна полезная документация, и программисты наверняка предпочтут их просто заглушить.

Хорошие комментарии — это благо, которое помогает пользователям понять код, с которым они работают. Плохие же создают проблемы с поддержкой и в случае

¹ Предустановленное поведение автоматического добавления `README.md` можно переопределить через поле `readme` в файле `Cargo.toml`.

расхождения с кодом могут человека запутать. Как же отличить плохие комментарии от хороших?

Главный совет: *стараться не повторять в их тексте что-то очевидное из самого кода.* Рекомендация 1 призывала максимально выражать семантику кода через систему типов Rust. Как только вы это сделаете, позвольте системе типов задокументировать эту семантику. Предположим, читатель знаком с Rust — возможно, потому, что прочитал полезную серию рекомендаций по его эффективному использованию, — и не описывает повторно то, что и так ясно из сигнатур и присутствующих в коде типов.

Например, излишне перегруженный документирующий комментарий может выглядеть так.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
/// Возвращает новый объект `BoundingBox`, который включает
/// ровно два объекта `BoundingBox`
///
/// Параметры:
/// - `a`: иммутабельная ссылка на `BoundingBox`
/// - `b`: иммутабельная ссылка на `BoundingBox`
/// Возвращает: новый объект `BoundingBox`
pub fn union(a: &BoundingBox, b: &BoundingBox) -> BoundingBox {
```

Этот комментарий без какой-либо пользы повторяет многие детали, очевидные из сигнатуры функции.

Хуже того, подумайте, что может произойти, если код после очередного рефакторинга будет сохранять результат в одном из своих изначальных аргументов (это станет критическим изменением; подробнее — в рекомендации 21). Никакой компилятор или инструмент не пожалуется, что комментарий не был обновлен, поэтому он легко может утратить актуальность.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
/// Возвращает новый объект `BoundingBox`, который включает
/// ровно два объекта `BoundingBox`
///
/// Параметры:
/// - `a`: иммутабельная ссылка на `BoundingBox`
/// - `b`: иммутабельная ссылка на `BoundingBox`
/// Возвращает: новый объект `BoundingBox`
pub fn union(a: &mut BoundingBox, b: &BoundingBox) {
```

В результате изначальный комментарий переживает рефакторинг без изменений, так как его текст описывает поведение, а не синтаксические детали:

```
/// Вычисляет `BoundingBox`, которая включает ровно
/// два объекта `BoundingBox`
pub fn union(a: &mut BoundingBox, b: &BoundingBox) {
```

Помогает улучшить документацию и зеркальная форма предыдущего совета: *добавляйте в текст все, что из кода неочевидно*. Сюда относятся предварительные условия, инварианты, паники, состояния ошибки и все остальное, что может удивить пользователя. Если не можете обеспечить соблюдение кодом принципа наименьшего удивления, проследите, чтобы все его сюрпризы были задокументированы. В этом случае вы хотя бы сможете сказать: «Я вас предупреждал».

Еще один типичный пример ошибки — документирующие комментарии, описывающие то, как некий другой код использует метод, а не то, что конкретно этот метод делает:

```
/// Определяет область пересечения двух объектов `BoundingBox`,
/// возвращая `None`, если пересечения нет. Код определения коллизии
/// в `hits.rs` использует эту информацию для изначальной проверки
/// возможного наложения объектов, прежде чем выполнять более
/// дорогостоящую попиксельную проверку в функции `objects_overlap`
pub fn intersection(
    a: &BoundingBox,
    b: &BoundingBox,
) -> Option<BoundingBox> {
```

Подобные комментарии почти всегда будут утрачивать актуальность: при изменении использующего кода (здесь это `hits.rs`) описывающий поведение комментарий окажется не в тему.

Если перефразировать подобный комментарий так, чтобы он больше акцентировался на описании «*почему*», то он станет более устойчивым к будущим изменениям:

```
/// Определяет область пересечения двух объектов `BoundingBox`,
/// возвращая `None`, если пересечения нет. Заметьте, что факт
/// пересечения ограничивающих рамок необходим, но не достаточен
/// для коллизии объектов, то есть дополнительная попиксельная
/// проверка по-прежнему нужна
pub fn intersection(
    a: &BoundingBox,
    b: &BoundingBox,
) -> Option<BoundingBox> {
```

Хорошим ориентиром при написании ПО будет девиз «Программируйте в будущем времени»¹, то есть структурируйте код с учетом возможных изменений

¹ Скотт Мейерс (Scott Meyers), «Эффективное использование C++», см. рекомендацию 32.

в дальнейшем. То же касается документации: ставя акцент на семантике «*почему*» и «*почему нет*», мы получаем текст, который с большей вероятностью долго останется актуальным.

Запомните

- Добавляйте комментарии для публичных элементов API.
- Описывайте аспекты кода, которые из него самого неочевидны, такие как паника и критерии безопасности.
- *Не описывайте* то, что понятно из кода.
- Упрощайте навигацию, давая перекрестные ссылки и выделяя идентификаторы.

Рекомендация 28. Используйте макросы вдумчиво

В некоторых случаях можно легко склониться к написанию макроса вместо функции, так как только макрос может выполнить то, что нужно.

*Пол Грэхэм (Paul Graham),
On Lisp (Prentice Hall)*

Система макросов Rust позволяет выполнять *метапрограммирование*, то есть писать код, генерирующий другой код в вашем проекте. Чаще всего это пригождается, когда есть фрагменты шаблонного кода, которые детерминированы и повторяются и которые в противном случае пришлось бы согласовывать вручную.

Программисты, переходящие на Rust, уже знакомы с макросами, предоставляемыми препроцессором C/C++, которые выполняют подстановку токенов входного текста. Но в Rust механизм макросов устроен иначе, так как они работают не просто на текстовом содержимом программы, а на ее спарсенных токенах либо *абстрактном синтаксическом дереве* (abstract syntax tree, AST).

Это означает, что макросы в Rust могут понимать структуру кода и, как следствие, избегать целых категорий сопутствующих подводных камней. В частности, следующий раздел покажет, что декларативные макросы в этом языке *гигиеничны*, то есть не могут случайно обращаться к локальным переменным в окружающем коде и захватывать их.

Макросы можно представить как очередной уровень абстракции кода. Простая форма абстракции — это функция: она абстрагирует различия между разными значениями одного *типа*, реализуя код, который может использовать любую

функциональность и методы этого типа вне зависимости от обрабатываемого в данный момент значения. Еще один вид абстракции — обобщения: они абстрагируют различия между разными *типами*, которые удовлетворяют границам трейта, реализуя код, который может применять любые методы, предоставляемые этими границами, вне зависимости от текущего обрабатываемого типа.

Макрос абстрагирует различия между разными фрагментами программы, которые играют одну роль (тип, идентификатор, выражение и т. д.). Его реализация может включать любой код, использующий эти фрагменты в одной роли.

В Rust есть два способа определения макросов.

- Декларативные макросы, также известные как макросы на основе образца, позволяют вставлять в программу произвольный код Rust на основе входных параметров макроса, которые делятся на категории согласно их роли в AST.
- Процедурные макросы. Позволяют вставлять в программу произвольный код Rust, исходя из спарсенных токенов исходного кода. Чаще всего используются для макросов `derive`, которые могут генерировать код на основе содержимого определений структур данных.

Декларативные макросы

И хотя текущая рекомендация не место для воспроизведения документации декларативных макросов (<https://oreil.ly/Vm7AZ>), некоторые ее важные детали все же обозначить нужно.

Во-первых, имейте в виду, что правила видимости при использовании декларативных макросов отличаются от аналогичных правил для других элементов Rust. Если декларативный макрос определен в файле исходного кода, то задействовать его может только код *после* определения макроса:

НЕ КОМПИЛИРУЕТСЯ

```
fn before() {
    println!("[before] square {} is {}", 2, square!(2));
}

/// Макрос, возводящий свой аргумент в квадрат
macro_rules! square {
    { $e:expr } => { $e * $e }
}

fn after() {
    println!("[after] square {} is {}", 2, square!(2));
}
```

```
error: cannot find macro `square` in this scope
--> src/main.rs:4:45
 |
4 |     println!("[before] square {} is {}", 2, square!(2));
 |                                         ^^^^^^
 |
= help: have you added the `#[macro_use]` on the module/import?
```

Атрибут `#[macro_export]` расширяет область видимости макроса, но из-за этого возникает странность — он отображается в верхней части крейта, даже если определен в модуле:

```
mod submod {
    #[macro_export]
    macro_rules! cube {
        { $e:expr } => { $e * $e * $e }
    }
}

mod user {
    pub fn use_macro() {
        // Примечание: *не* `crate::submod::cube!`
        let cubed = crate::cube!(3);
        println!("cube {} is {}", 3, cubed);
    }
}
```

Декларативные макросы в Rust известны своей *гигиеничностью*: расширенный код в их теле не может использовать локально присвоенные переменные. Например, макрос, предполагающий, что существует некая переменная `x`:

```
// Создает макрос, предполагающий существование локальной `x`
macro_rules! increment_x {
    {} => { x += 1; };
}
```

на этапе компиляции в момент использования вызовет сбой:

НЕ КОМПИЛИРУЕТСЯ

```
let mut x = 2;
increment_x!();
println!("x = {}", x);
```

```
error[E0425]: cannot find value `x` in this scope
--> src/main.rs:55:13
 |
55 |     {} => { x += 1; };
 |           ^ not found in this scope
```

```
...
314 |     increment_x!();
|     ----- in this macro invocation
|
= note: this error originates in the macro `increment_x`
```

Эта гигиеничность означает, что макросы в Rust безопаснее макросов препроцессора в C. Тем не менее при их использовании все же есть пара подвохов.

Первый — то, что хотя вызов макроса и *выглядит* как вызов функции, он им не является. Макрос генерирует в точке вызова код, который может выполнять манипуляции с его аргументами:

```
macro_rules! inc_item {
    { $x:ident } => { $x.contents += 1; }
}
```

То есть интуитивное понимание того, передаются параметры или же используются через &, не работает:

```
let mut x = Item { contents: 42 }; // Тип не `Copy`
// Элемент *не* перемещается, несмотря на синтаксис (x),
// но тело макроса *может* изменять `x`
inc_item!(x);

println!("x is {x:?}");
x is Item { contents: 43 }
```

Это становится понятным, если вспомнить, что макрос вставляет код в точке вызова, в данном случае добавляя строку, инкрементирующую `x.contents`. Инструмент `cargo-expand` показывает код, который компилятор видит после раскрытия макроса:

```
let mut x = Item { contents: 42 };
x.contents += 1;
{
    ::std::io::_print(format_args!("x is {0:?}\n", x));
};
```

Раскрытый код включает изменение на месте через владельца элемента, а не ссылку. (Также интересно просмотреть раскрытую версию `println!`, которая опирается на макрос `format_args!`, о котором мы вскоре поговорим¹.)

Итак, восклицательный знак служит предупреждением: раскрытый код макроса может выполнять произвольные действия со своими аргументами.

¹ Внимательный читатель заметит, что `format_args!` выглядит как вызов макроса даже после его расширения. Дело в том, что это специальный макрос, встроенный в компилятор.

Кроме того, раскрытий код может включать операции потока управления, которые в вызывающем коде невидимы, будь то циклы, условные конструкции, инструкции `return` или использование оператора `?`. Очевидно, что это наверняка нарушит принцип наименьшего удивления, поэтому по возможности *используйте такие макросы, поведение которых соответствует нормальному поведению Rust*. (В то же время, если цель макроса — допустить странности в потоке управления, так и поступайте. Только упростите жизнь пользователям, отчетливо задокументировав это поведение потока управления.)

В качестве примера рассмотрим макрос для проверки HTTP-кодов состояния, который молча включает в свое тело `return`:

```
/// Проверяет, соответствует ли HTTP-статус успеху.
/// Если нет — функция завершается
macro_rules! check_successful {
    { $e:expr } => {
        if $e.group() != Group::Successful {
            return Err(MyError("HTTP operation failed"));
        }
    }
}
```

Код, который использует этот макрос для проверки результата некой HTTP-операции, способен внести неоднозначность в поток управления:

```
let rc = perform_http_operation();
check_successful!(rc); // Может молча завершить функцию
// ...
```

Альтернативная версия макроса для генерации кода, создающего `Result`:

```
/// Преобразует HTTP-статус в `Result<(), MyError>`, указывающий на успех
macro_rules! check_success {
    { $e:expr } => {
        match $e.group() {
            Group::Successful => Ok(()),
            _ => Err(MyError("HTTP operation failed")),
        }
    }
}
```

дает более понятный код:

```
let rc = perform_http_operation(); check_success!(rc)?;
// Ошибка в потоке видима из-за `?`
// ...
```

Второй нюанс декларативных макросов, который нужно иметь в виду, общий для Rust и препроцессора C: если аргумент для макроса — это выражение с побочными эффектами, остерегайтесь его повторного применения в этом макросе. Определенный ранее макрос `square!` получает в качестве аргумента произвольное выражение, после чего использует его дважды, что может вызвать сюрпризы.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
let mut x = 1;
let y = square!({
    x += 1;
    x
});
println!("x = {x}, y = {y}");
// Вывод: x = 3, y = 6
```

Если такое поведение не предполагалось, исправить его можно, просто вычислив выражение один раз и присвоив результат локальной переменной:

```
macro_rules! square_once {
    {$e:expr} => {
        {
            let x = $e;
            x*x // Примечание: здесь есть нюанс, который будет разъяснен далее...
        }
    }
}
// Вывод сейчас: x = 2, y = 4
```

Альтернативным решением будет не позволять использовать в качестве ввода для макроса произвольные выражения. Если спецификатор фрагмента синтаксиса `expr` (<https://oreil.ly/8u0NJ>) заменить спецификатором `ident`, то макрос будет получать в качестве ввода только идентификаторы и попытка передать ему произвольное выражение приведет к ошибке компиляции.

Процедурные макросы

Rust поддерживает также *процедурные макросы*. Подобно декларативным, они могут вставлять произвольный код Rust в исходный код программы. Однако здесь входные данные макроса являются уже не просто конкретными передаваемыми в него аргументами. Процедурный макрос имеет доступ к спарсенным токенам, соответствующим определенному фрагменту оригинального исходного кода. И это обеспечивает такую степень экспрессивности, которая приближается к гибкости динамических языков вроде Lisp, но с сохранением гарантий этапа компиляции. Это помогает обойти ограничения отражения Rust, о чём говорилось в рекомендации 19.

ФОРМАТИРОВАНИЕ ЗНАЧЕНИЙ

Один из распространенных стилей создания декларативных макросов подразумевает составление сообщения, включающего различные значения из текущего состояния кода. Например, стандартная библиотека включает `format!` для сборки `String`, `println!` — для вывода в `stdout`, `eprintln!` — для стандартного вывода ошибок и т. д. Синтаксис этих директив форматирования описан в документации и примерно аналогичен инструкции `printf` в C. Тем не менее аргументы форматирования проверяются на этапе компиляции, и в реализациях макроса для форматирования отдельных значений используются трейты `Display` и `Debug`, описанные в рекомендации 10¹.

Вы можете (и вам следует) применять одинаковый синтаксис форматирования для всех своих макросов, выполняющих похожие функции. Например, макросы журналирования, предоставляемые крейтом `log`, задействуют тот же синтаксис, что и `format!`. Для этого *в макросах, форматирующих аргументы, используйте `format_args!`*, не пытаясь изобрести колесо:

```
// Журналирует ошибку, включая ее место в коде,
// с помощью аргументов в стиле `format!`
// В реальном коде наверняка будет использоваться крейт `log`
macro_rules! my_log {
    { $($arg:tt)* } => {
        eprintln!("{}:{}: {}", file!(), line!(), format_args!($($arg)*));
    }
}

let x = 10u8;
// Спецификаторы формата:
// - `x` означает "выводить в шестнадцатеричном виде"
// - `#` означает "добавлять префикс '0x'"
// - `04` означает "добавлять ведущие нули, чтобы ширина была не менее 4"
//   (сюда относится префикс '0x')
my_log!("x = {:#04x}", x);

src/main.rs:331: x = 0x0a
```

Процедурные макросы необходимо определять в отдельном крейите (с типом `proc-macro`), откуда они в дальнейшем берутся. И данный крейт почти всегда будет зависеть от вспомогательной библиотеки `proc-macro`, предоставляемой стандартными инструментами, или `proc-macro2`, разработанной Дэвидом Толнаем (David Tolnay) и позволяющей работать с входными токенами.

¹ Модуль `std::fmt` включает также другие трейты, используемые при отображении данных в конкретных форматах. Например, `LowerHex` применяется, когда спецификатор формата `x` указывает, что вывод должен быть в шестнадцатеричной форме и нижнем регистре.

Процедурные макросы бывают трех типов.

Функциональные макросы

Вызываются с аргументом.

Макросы-атрибуты

Прикрепляются к фрагменту синтаксиса в программе.

Derive-макросы

Прикрепляются к определению структуры данных.

Функциональные макросы

Функциональные макросы вызываются с аргументом. Их код использует считанные токены, которые этот аргумент составляют, и в качестве результата генерирует другие токены. Обратите внимание на то, что в предыдущем предложении слово «аргумент» используется в единственном числе — даже если функциональный макрос вызывается с чем-то похожим на несколько аргументов:

```
my_func_macro!(15, x + y, f32::consts::PI);
```

сам он получает лишь один, являющийся потоком считанных токенов. Реализация макроса, которая просто выводит на этапе компиляции содержимое этого потока:

```
use proc_macro::TokenStream;

// Функциональный макрос, который просто выводит свой поток ввода
// (на этапе компиляции)
#[proc_macro]
pub fn my_func_macro(args: TokenStream) -> TokenStream {
    println!("Input TokenStream is:");
    for tt in args {
        println!(" {tt:?}");
    }
    // Возвращает пустой поток токенов, заменяя им аргумент
    // вызова макроса
    TokenStream::new()
}
```

показывает поток, соответствующий переданному на вводе:

```
Input TokenStream is:
Literal { kind: Integer, symbol: "15", suffix: None, span: #0
          bytes(10976..10978) }
Punct { ch: ',', spacing: Alone, span: #0 bytes(10978..10979) }
Ident { ident: "x", span: #0 bytes(10980..10981) }
Punct { ch: '+', spacing: Alone, span: #0 bytes(10982..10983) }
Ident { ident: "y", span: #0 bytes(10984..10985) }
```

```
Punct { ch: ',', spacing: Alone, span: #0 bytes(10985..10986) }
Ident { ident: "f32", span: #0 bytes(10987..10990) }
Punct { ch: ':', spacing: Joint, span: #0 bytes(10990..10991) }
Punct { ch: ':', spacing: Alone, span: #0 bytes(10991..10992) }
Ident { ident: "consts", span: #0 bytes(10992..10998) }
Punct { ch: ':', spacing: Joint, span: #0 bytes(10998..10999) }
Punct { ch: ':', spacing: Alone, span: #0 bytes(10999..11000) }
Ident { ident: "PI", span: #0 bytes(11000..11002) }
```

Низкоуровневая природа этого ввода означает, что реализация макроса должна выполнять собственный парсинг. Например, выделение того, что выглядит как отдельные аргументы макроса, подразумевает поиск токенов `TokenTree::Punct`, которые содержат запятые, эти аргументы разделяющие. В крейте `syn` Дэвида Толна есть библиотека парсинга, которая в этом помогает, что описывается в пункте «*Derive-макросы*» далее.

Из-за этого обычно проще использовать декларативный макрос, нежели функциональный, так как тогда ожидаемую структуру входных данных можно выразить в соответствующем паттерне.

Недостатком потребности в ручной обработке является то, что функциональные макросы достаточно гибкие для получения ввода, который *не парсится* как нормальный код Rust. Требуется такое нечасто (и не всегда уместно), поэтому подобные макросы встречаются сравнительно редко.

Макросы-атрибуты

Для вызова макросы-атрибуты размещаются перед нужным элементом программы, используя в качестве ввода считанные для этого элемента токены. Такие макросы тоже могут генерировать на выходе произвольный поток токенов, но в данном случае вывод обычно представляет преобразованную форму ввода.

Например, макрос-атрибут можно использовать для обертывания тела функции:

```
#[log_invocation]
fn add_three(x: u32) -> u32 {
    x + 3
}
```

чтобы вызов этой функции журналировался в консоль:

```
let x = 2;
let y = add_three(x);
println!("add_three({x}) = {y}");

log: calling function 'add_three'
log: called function 'add_three' => 5
add_three(2) = 5
```

Реализация этого макроса слишком велика для того, чтобы привести ее здесь, так как коду нужно проверить структуру входных токенов и создать новые, но в этом нам снова помогает крейт `syn`.

Derive-макросы

Это последний тип процедурных макросов. Они позволяют автоматически прикреплять генерируемый код к определению структуры данных (`struct`, `enum` или `union`). Эти макросы похожи на макросы-атрибуты, но у них есть несколько важных аспектов, связанных конкретно с `derive`.

Первый состоит в том, что `derive`-макросы *добавляют* результат к входным токенам, а не заменяют их. Это означает, что определение структуры данных остается нетронутым, но макрос может добавить к нему сопутствующий код.

Второй момент заключается в том, что `derive`-макрос способен объявлять вспомогательные атрибуты, которые затем можно использовать для обозначения деталей структуры данных, требующих особой обработки. Например, макрос `Deserialize` из модели данных `serde` содержит вспомогательный атрибут `serde`, который может предоставлять метаданные для координации процесса десериализации:

```
fn generate_value() -> String {
    "unknown".to_string()
}

#[derive(Debug, Deserialize)]
struct MyData {
    // Если `value` при десериализации отсутствует,
    // поле заполняется с помощью вызова `generate_value()`
    #[serde(default = "generate_value")]
    value: String,
}
```

Последним значимым нюансом `derive`-макросов является то, что крейт `syn` может взять на себя значительную нагрузку, связанную с парсингом входных токенов в равноценные узлы AST. Макрос `syn::parse_macro_input!` преобразует токены в структуру данных `syn::DeriveInput`, которая описывает содержимое элемента, а обработать `Derive Input` намного проще, чем сырой поток токенов.

На практике `derive`-макросы являются самым распространенным видом процедурных макросов — возможность генерировать реализации «поле за полем» (для `struct`) или «вариант за вариантом» (для `enum`) позволяет обеспечить богатую функциональность при минимуме усилий, например добавив одну строку вроде `#[derive(Debug, Clone, PartialEq, Eq)]`.

Поскольку выводимые (`derive`) реализации генерируются автоматически, это также означает, что они сохраняют согласованность с определением структуры

данных. Например, если вы добавите в `struct` новое поле, прописанную реализацию `Debug` потребуется обновить вручную при том, что автоматически выведенная версия отобразит новое поле без дополнительных усилий (или не скомпилируется, если это окажется невозможным).

Когда использовать макросы

Основная причина применения макросов — стремление избежать повторения кода, особенно такого, который иначе придется вручную синхронизировать с другими частями программы. В этом смысле написание макроса — это просто расширенная форма процесса обобщения, являющегося частью программирования.

- Если вы повторяете одинаковый код для нескольких значений конкретного типа, инкапсулируйте его в общую функцию и вызывайте ее из всех мест повторов.
- Если вы повторяете одинаковый код для нескольких типов, инкапсулируйте его в обобщение с границей трейта и используйте это обобщение из всех мест повторов.
- Если вы повторяете одинаковый код в нескольких местах, инкапсулируйте его в макрос и используйте последний из всех мест повторов.

Например, избежать повторов кода, работающего в различных вариантах `enum`, можно только с помощью макроса:

```
enum Multi {
    Byte(u8),
    Int(i32),
    Str(String),
}

/// Извлекает копии всех значений конкретного варианта enum
#[macro_export]
macro_rules! values_of_type {
    { $values:expr, $variant:ident } => {
        {
            let mut result = Vec::new();
            for val in $values {
                if let Multi::$variant(v) = val {
                    result.push(v.clone());
                }
            }
            result
        }
    }
}
```

```
fn main() {
    let values = vec![
        Multi::Byte(1),
        Multi::Int(1000),
        Multi::Str("a string".to_string()),
        Multi::Byte(2),
    ];
    let ints = values_of_type!(&values, Int);
    println!("Integer values: {ints:?}");

    let bytes = values_of_type!(&values, Byte);
    println!("Byte values: {bytes:?}");

    // Вывод:
    //     Целочисленные значения: [1000]
    //     Значения байтов: [1, 2]
}
```

Еще один сценарий, при реализации которого макрос поможет избежать ручного повторения кода, — это ситуация, когда в противном случае информация о наборе значений данных будет разбросана по разным частям кода.

Рассмотрим в качестве примера структуру данных, которая кодирует информацию о HTTP-кодах состояния. Здесь макрос поможет сохранить всю сопутствующую информацию в одном месте:

```
// Модуль http.rs

#[derive(Debug, PartialEq, Eq, Clone, Copy)]
pub enum Group {
    Informational, // 1xx
    Successful, // 2xx
    Redirection, // 3xx
    ClientError, // 4xx
    ServerError, // 5xx
}

// Информация об HTTP-кодах ответа
http_codes! {
    Continue          => (100, Informational, "Continue"),
    SwitchingProtocols => (101, Informational, "Switching Protocols"),
    // ...
    OK              => (200, Successful, "OK"),
    Created          => (201, Successful, "Created"),
    // ...
}
```

Вызов этого макроса содержит всю сопутствующую информацию — численное значение, группу, описание — для каждого HTTP-кода состояния, выступая

чем-то наподобие предметно-ориентированного языка (domain-specific language, DSL), содержащего источник истины для этих данных.

Определение макроса описывает сгенерированный код. Каждая строка в форме `$(...)+` в генерируемом коде раскрывается в несколько строк — по одной для каждого аргумента макроса:

```
macro_rules! http_codes {
    { $( $name:ident => ($val:literal, $group:ident, $text:literal), )+ } => {
        #[derive(Debug, Clone, Copy, PartialEq, Eq, Hash)]
        #[repr(i32)]
        enum Status {
            $($name = $val, )+
        }
        impl Status {
            fn group(&self) -> Group {
                match self {
                    $( Self::$name => Group::$group, )+
                }
            }
            fn text(&self) -> &'static str {
                match self {
                    $( Self::$name => $text, )+
                }
            }
        }
        impl core::convert::TryFrom<i32> for Status {
            type Error = ();
            fn try_from(v: i32) -> Result<Self, Self::Error> {
                match v {
                    $( $val => Ok(Self::$name), )+
                    _ => Err(())
                }
            }
        }
    }
}
```

В результате на выходе макрос генерирует весь код, который получается из указанных в источнике истины значений:

- определение `enum`, содержащее все варианты;
- определение метода `group()`, указывающее, к какой группе принадлежит HTTP-код;
- определение метода `text()`, которое отображает код состояния в его текстовое описание;
- реализацию `TryFrom<i32>` для преобразования чисел в `enum`-значения кодов состояний.

Если позднее потребуется добавить значения, достаточно будет включить одну дополнительную строку:

```
ImATeapot => (418, ClientError, "I'm a teapot"),
```

Без макроса все эти четыре разбросанные области пришлось бы обновлять вручную. Компилятор укажет на некоторые из них (так как выражения `match` должны охватывать все случаи), но не на все — `TryFrom<i32>` запросто можно упустить.

Поскольку макросы раскрываются на месте вызова в коде, их также можно использовать для автоматического создания дополнительной диагностической информации — в частности, используя макросы стандартной библиотеки `file!()` и `line!()`, которые выдают информацию о расположении исходного кода:

```
macro_rules! log_failure {
    { $e:expr } => {
        {
            let result = $e;
            if let Err(err) = &result {
                eprintln!("{}:{}: operation '{}' failed: {:?}",,
                         file!(),
                         line!(),
                         stringify!($e),
                         err);
            }
            result
        }
    }
}
```

Тогда при сбое в файл журнала будут автоматически записываться подробности о том, что вызвало сбой и где:

```
use std::convert::TryInto;

let x: Result<u8, _> = log_failure!(512.try_into()); // too big for `u8`
let y = log_failure!(std::str::from_utf8(b"\xc3\x28")); // invalid UTF-8

src/main.rs:340: operation '512.try_into()' failed: TryFromIntError(())
src/main.rs:341: operation 'std::str::from_utf8(b"\xc3\x28")' failed:
          Utf8Error { valid_up_to: 0, error_len: Some(1) }
```

Недостатки макросов

Основным недостатком макроса является влияние на читаемость и обслуживаемость кода. В подразделе «Декларативные макросы» ранее говорилось, что макросы позволяют создавать DSL для лаконичного выражения ключевых возможностей кода и данных. Но это означает также, что любому, кто читает или обслуживает код, дополнительно к пониманию Rust потребуется понимать и этот

DSL, включая его реализацию в определении макроса. Например, `http_codes!` из предыдущего раздела создает `enum` с именем `Status`, но в DSL, используемом для вызова макроса, это перечисление невидимо.

И, поскольку такой непрозрачный код больше не соблюдает синтаксические соглашения Rust, это коснется не только инженеров, но и инструментов, работающих с ним. Один тривиальный пример такого случая содержал показанный ранее макрос `square_once!`: тело макроса не было отформатировано в соответствии со стандартными правилами `rustfmt`:

```
{  
    let x = $e;  
    // Инструмент `rustfmt` не сможет обработать код макроса,  
    // поэтому он был переформатирован в `x * x`  
    x*x  
}
```

Еще одним примером является ранее приведенный макрос `http_codes!`, в котором DSL использует имена вариантов перечисления `Group` без префикса `Group::` или инструкции `use`, что может сбить с толку некоторые инструменты навигации по коду.

И даже компилятор здесь не особо помогает: его сообщения об ошибках не всегда прослеживают всю цепочку определения и применения макроса. (Однако в экосистеме инструментов есть такие, которые помогают решить эту проблему, например использованный ранее `cargo-expand` Дэвида Толна. Подробнее — в рекомендации 31.)

Еще одним недостатком макросов является вероятность появления раздутого кода — одна строка вызова макроса может дать сотни строк генерированного кода, который при беглом просмотре программы окажется незаметным. Такая проблема менее вероятна, когда код только создается, поскольку на тот момент он нужен и избавляет человека от необходимости писать его самому. Тем не менее, если впоследствии необходимость в этом коде отпадает, будет сложно понять, что в нем есть большие фрагменты, которые можно удалить.

Совет

Несмотря на то что у макросов есть недостатки, описанные в предыдущем разделе, они все равно остаются полезными для случаев, когда есть разные фрагменты кода, которые нужно поддерживать согласованными, чего нельзя сделать иным способом: *используйте макрос, когда это единственный способ обеспечить согласованность разнородного кода*.

Макросы подходят и для ситуаций, когда нужно сократить объем шаблонного кода: *применяйте макрос для повторяющегося рутинного кода*, который нельзя объединить в функцию или обобщение.

Чтобы сократить влияние на читаемость, старайтесь избегать в макросах синтаксиса, конфликтующего со стандартными правилами Rust. Делайте вызов макроса либо похожим на нормальный код, либо явно *отличающимся*, чтобы никто их не путал. Вот конкретные рекомендации.

- По возможности *избегайте макросов, которые при раскрытии вставляют ссылки*, — например, вызов макроса `my_macro!(&list)` будет лучше сочетаться с нормальным кодом Rust, чем вызов `my_macro!(list)`.
- *Старайтесь не использовать в макросе нелокальные операции потока управления*, чтобы при чтении кода можно было понять его поток, не вникая в детали самого макроса.

Это стремление к читаемости в стиле Rust иногда влияет на выбор между декларативным и процедурным макросом. Если вам нужно генерировать код для каждого поля структуры или каждого варианта перечисления, *отдавайте предпочтение derive-макросу, а не такому, который генерирует тип* (несмотря на пример из подраздела «Когда использовать макросы»), — так вы сделаете код более идиоматичным и понятным.

Однако если добавляете `derive`-макрос с функциональностью, не характерной для вашего проекта, поищите подходящий внешний крейт (см. рекомендацию 25). Например, проблема преобразования целочисленных значений в подходящий вариант C-подобного `enum` хорошо проработана: все `enumn::N`, `num_enum::TryFromPrimitive`, `num_derive::FromPrimitive` и `strum::FromRepr` решают тот или иной ее аспект.

Рекомендация 29. Прислушивайтесь к Clippy

Похоже, вы пишете письмо. Вам нужна помощь?

Microsoft Clippit

В рекомендации 31 описывается экосистема полезных инструментов для Rust, но один из них, Clippy, настолько важен и эффективен, что достоин собственного раздела.

Clippy — это дополнительный компонент Cargo (`cargo clippy`), который по разным категориям генерирует предупреждения относительно использования Rust.

Корректность

Предупреждает о распространенных ошибках программирования.

Идиоматичность

Предупреждает о конструкциях кода, написанных в нестандартном для Rust стиле.

Лаконичность

Подсказывает более компактные варианты кода.

Производительность

Предлагает альтернативы, избегающие лишней обработки или аллокации.

Читаемость

Описывает изменения в коде, которые сделают его более понятным и читаемым.

К примеру, следующий код прекрасно скомпилируется.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
pub fn circle_area(radius: f64) -> f64 {
    let pi = 3.14;
    pi * radius * radius
}
```

Но Clippy укажет, что локальная аппроксимация `pi` необязательна и неточна:

```
error: approximate value of `f{32, 64}::consts::PI` found
--> src/main.rs:5:18
  |
5 |     let pi = 3.14;
  |     ^^^^
  |
= help: consider using the constant directly
= help: for further information visit
  https://rust-lang.github.io/rust-clippy/master/index.html#approx\_constant
= note: `#[deny(clippy::approx_constant)]` on by default
```

Указанная в коде веб-страница объясняет проблему и предоставляет вариант исправления кода:

```
pub fn circle_area(radius: f64) -> f64 {
    std::f64::consts::PI * radius * radius
}
```

Как уже говорилось, каждое предупреждение Clippy сопровождается ссылкой на веб-страницу с описанием ошибки, где объясняется, *почему* этот код считается плохим. И это важно, так как позволяет решить, применимы ли указанные причины к вашему коду, или определить, что результат проверки линтера здесь неактуален. В некоторых случаях текст сообщений описывает также известные проблемы оформления, что может прояснить непонятное ложное срабатывание.

Если вы решите, что предупреждение линтера для вашего кода неактуально, можете отключить его для конкретного элемента (`#[allow(clippy::some_lint)]`)

или всего крейта (#! [allow(clippy::some_lint)]), используя дополнительный ! на верхнем уровне. Тем не менее обычно лучше выполнить небольшой рефакторинг кода, чем тратить время и силы, доказывая, что предупреждение действительно ложное.

Независимо от того, выберете ли вы отключение предупреждений Clippy или внесение исправлений, нужно *обеспечить их отсутствие в своем коде*.

Таким образом, при появлении нового предупреждения — хоть в результате изменения кода, хоть из-за добавления в Clippy новых проверок — оно будет очевидно. Clippy также нужно включать в систему CI (см. рекомендацию 32).

Предупреждения Clippy особенно полезны при изучении Rust, так как вскрывают подвохи, которые вы могли упустить, и помогают лучше усвоить идиому Rust.

Во многих рекомендациях данной книги предупреждения Clippy присутствуют, когда есть возможность механически проверить указанную проблему.

- В рекомендации 1 предлагается использовать более экспрессивные типы вместо простых `bool`. И Clippy тоже укажет на применение нескольких `bool` в параметрах функции и структурах.
- В рекомендации 3 описываются манипуляции с типами `Option` и `Result`, и Clippy указывает на несколько возможных недочетов:
 - необязательное преобразование `Result` в `Option`;
 - возможность использовать `unwrap_or_default`.
- Рекомендация 3 также предлагает по возможности возвращать ошибки вызывающему коду, и Clippy отмечает отсутствие некоторых возможностей для этого.
- Рекомендация 5 предлагает реализовывать `From` вместо `Into`, о чем также говорит Clippy.
- Помимо этого, в рекомендации 5 описывается приведение, и Clippy выдает предупреждения для следующего (по умолчанию отключено):
 - приведений `as`, которые могли быть `from`;
 - приведений `as`, которые можно обрезать;
 - приведений `as`, которые можно обернуть;
 - приведений `as`, которые утрачивают точность;
 - приведений `as`, которые могут преобразовывать знаковые отрицательные числа в большие положительные числа;
 - **любого** использования `as`.

- Рекомендация 8 описывает типы жирных указателей, и различные линты Clippy указывают на случаи, когда присутствует лишняя косвенная адресация, которая:
 - содержит аллоцированную в куче коллекцию в `Box`;
 - содержит аллоцированную в куче коллекцию элементов `Box`;
 - получает ссылку на `Box`.
- В рекомендации 9 описывается множество способов манипулирования экземплярами `Iterator`. Clippy включает поистине поразительное число линтов, указывающих на комбинации методов итератора, которые можно упростить.
- В рекомендации 10 описываются стандартные трейты Rust и приводятся некоторые требования к реализации, которые Clippy проверяет:
 - `Ord` должен согласовываться с `PartialOrd`;
 - `PartialEq::ne` не должен требовать нестандартной реализации (подробнее — в рекомендации 13);
 - `Hash` и `Eq` должны быть последовательными;
 - `Clone` для типов `Copy` должны совпадать.
- Рекомендация 18 предлагает ограничить использование `panic!` и сопутствующих методов вроде `expect`. Clippy все это также обнаруживает.
- В рекомендации 21 отмечается, что импорт подстановочной версии крейта не имеет смысла. Clippy с этим согласен.
- Рекомендация 23 предлагает избегать импорта с подстановкой. Clippy тоже.
- В рекомендациях 24 и 25 затрагивается тот факт, что в графе зависимостей может присутствовать несколько версий одного крейта. Clippy тоже можно настроить на оповещение о подобных случаях.
- В рекомендации 26 объясняется аддитивная природа функциональности Cargo, и в Clippy есть предупреждение о случаях ее «отрицательных» имен (например, `"no_std"`), которые наверняка говорят о том, что функциональность этому не соответствует.
- Рекомендация 26 также объясняет, что необязательные зависимости крейта являются частью его функциональности, и Clippy предупреждает, если существуют имена явных функций (например, `"use-crate-x"`), которые могут по ошибке использовать их.
- В рекомендации 27 описываются соглашения для составления документирующих комментариев, и Clippy также будет указывать на следующее:
 - отсутствующие описания `panic!`;
 - отсутствующие описания требований для `unsafe`-кода.

Размер этого списка говорит о том, что в целях обучения будет очень полезно ознакомиться с перечнем предупреждений *Clippy* (<https://oreil.ly/Nt5zE>), включая проверки, которые по умолчанию отключены из-за своей излишней педантичности или частых ложных срабатываний. И хотя вы вряд ли захотите включать эти предупреждения для своего кода, понимание причин их появления поможет вам лучше вникнуть в принципы Rust и его идиому.

Рекомендация 30. Пишите не только модульные тесты

У всех компаний есть среда для тестирования.

И счастливы те, у которых она отделена от про-
дакшен-среды.

@FearlessSon (<https://oreil.ly/UzBRq>)

Как и большинство современных языков, Rust предоставляет возможности, которые упрощают написание тестов, существующих параллельно с кодом и дающих уверенность в корректности его работы.

Нет, я не буду разглагольствовать на тему важности тестов. Достаточно сказать, что код, который не тестировался, вряд ли будет работать так, как от него ожидают. Поэтому текущая рекомендация предполагает, что вы уже запланировали *писать для своего кода тесты*.

Ключевыми формами проверки выступают модульные и интеграционные тесты, которые будут рассмотрены в следующих двух подразделах. Но инструментарий Rust наряду с сопутствующими расширениями позволяет использовать и другие виды тестов, о целесообразности применения и механизмах работы которых также будет сказано далее.

Модульные тесты

Самый популярный вид тестов для кода Rust — модульные тесты. Вот пример:

```
// ... (код, определяющий функции `nat_subtract`  
// для вычитания натуральных чисел)  
  
#[cfg(test)]  
mod tests {  
    use super::*;

    #[test]  
    fn test_nat_subtract() {  
        assert_eq!(nat_subtract(4, 3).unwrap(), 1);  
    }
}
```

```
    assert_eq!(nat_subtract(4, 5), None);
}

#[should_panic]
#[test]
fn test_something_that_panics() {
    nat_subtract_unchecked(4, 5);
}
}
```

В каждом модульном teste будут встречаться некоторые аспекты этого примера:

- набор функций для модульного teste;
- обозначение каждой такой функции атрибутом `#[test]`;
- модуль, содержащий функции для тестирования, который аннотируется атрибутом `#[cfg(test)]`, чтобы его код собирался только в тестовых конфигурациях.

Прочие аспекты этого примера представляют необязательные детали, актуальные только для определенных тестов.

- Код teste здесь помещен в отдельный модуль, по соглашению именуемый `tests` или `test`. Этот модуль может быть встроенным (как здесь) или находиться в отдельном файле `tests.rs`. Использование для тестового модуля отдельного файла помогает понять, является ли задействующий функцию код тестовым или реальным.
- Тестовый модуль может содержать символ подстановки `use super::*`, чтобы подтягивать все содержимое родительского тестируемого модуля. Это упрощает добавление тестов и является исключением для рекомендации 23 «Избегайте импорта с подстановкой».
- Согласно стандартным правилам видимости модулей модульный teste может использовать все содержимое родительского модуля независимо от того, является оно `pub` или нет. Это делает возможным открытое тестирование кода, когда модульные teste проверяют внутреннюю функциональность, невидимую для обычных пользователей.
- Ожидаемые результаты в коде teste указываются с помощью выражения `expect()` или `unwrap()`. Совет из рекомендации 18 неактуален для кода в режиме тестирования, когда `panic!` используется в качестве сигнала о провале проверки. Аналогичным образом тестовый код проверяет ожидаемые результаты с помощью `assert_eq!`, вызывающей в случае сбоя панику.
- Тестируемый код включает функцию, которая паникует при получении определенного недопустимого ввода. Для реализации этого используется функция, обозначаемая атрибутом `#[should_panic]`. Она может потребоваться при тестировании внутренней функции, которая ожидает, что остальной

код будет учитывать ее инварианты и предварительные условия. Или же это может быть публичная функция, у которой есть причина игнорировать рекомендацию 18. (Документирующий комментарий к такой функции должен содержать раздел «Паники», оформленный в соответствии с рекомендацией 27.)

В рекомендации 27 предлагается *не документировать* то, что уже выражено в системе типов. По этой же логике не нужно тестировать то, что гарантируется системой типов. Если ваши типы `enum` будут содержать значения, которых нет в списке допустимых вариантов, вы можете столкнуться с гораздо большими проблемами, нежели провал модульного теста.

Однако если ваш код опирается на конкретную функциональность зависимостей, будет нeliшним включить базовые тесты этой функциональности. Цель здесь не в том, чтобы повторить тестирование, которое такая зависимость уже выполнила, а наладить систему раннего обнаружения, определяющую, не изменилось ли нужное вам поведение этой зависимости. Происходить это должно отдельно от проверки изменения публичной сигнатуры API, которое отражается в виде семантического номера версии (см. рекомендацию 21).

Интеграционные тесты

Еще одним типичным видом проверки, добавляемым в проект Rust, являются *интеграционные тесты*, размещаемые в каталоге `tests/`. Каждый файл в этом каталоге выполняется в отдельной тестовой программе, которая исполняет все функции с атрибутом `##[test]`.

Интеграционные тесты *не имеют* доступа к содержимому крейта, а значит, выступают в роли тестов поведения, которые могут проверять только его публичный API.

Тесты документации

В рекомендации 27 говорилось о добавлении в документирующие комментарии коротких фрагментов кода для иллюстрации использования конкретного элемента публичного API. Каждый такой фрагмент кода заключается в неявную `fn main() { ... }` и выполняется в рамках `cargo test`, по сути становясь для вашего кода дополнительным тестовым кейсом, называемым *тестом документации*. С помощью `cargo test --doc <item-name>` можно также выборочно выполнять отдельные тесты.

Регулярный прогон тестов в рамках системы CI (см. рекомендацию 32) позволит избежать серьезных отклонений фрагментов кода от текущей реальности вашего API.

Примеры

В рекомендации 27 описывалась возможность предоставления примеров программ, проверяющих ваш публичный API. Каждый файл кода Rust в каталоге `examples/` или каждом его подкаталоге, включающем `main.rs`, можно выполнить как самостоятельный двоичный файл командой `cargo run --example <name>` или `cargo test --example <name>`.

Эти программы имеют доступ только к публичному API вашего крейта и служат для иллюстрации использования этого API в целом. Примеры не обозначаются специально как тестовый код (не имеют `#[test]` или `#[cfg(test)]`) и являются неудачным местом для размещения кода, проверяющего вызывающие сомнения тонкости вашего крейта, в частности, потому, что по умолчанию `cargo test` примеры *не выполняет*.

Тем не менее желательно обеспечить, чтобы ваша система CI (см. рекомендацию 32) собирала и выполняла все связанные примеры для крейта (с помощью `cargo test -examples`), так как она может выступать в качестве системы раннего оповещения о недочетах, способных повлиять на большое число пользователей. Как уже говорилось, если ваши примеры демонстрируют основное применение вашего API, тогда сбой в них говорит о неисправности чего-то серьезного.

- Если это откровенная ошибка, то она наверняка повлияет на многих пользователей — код примера по своей природе подразумевает частое копирование, вставку и адаптирование.
- Если это преднамеренное изменение API, тогда примеры нужно соответствующим образом обновить. Изменение API подразумевает также обратную совместимость, поэтому если крейт опубликован, то его семантическую версию необходимо должным образом обновить (см. рекомендацию 21).

Поскольку высока вероятность того, что пользователи будут копировать примеры кода, их стиль должен отличаться от стиля тестируемого кода. В соответствии с рекомендацией 18 следует составлять для пользователей хороший пример кода, избегая вызовов `unwrap()` для `Result`. Вместо этого делайте так, чтобы функция `main` в каждом примере возвращала что-то вроде `Result<(), Box<dyn Error>>`, после чего повсеместно используйте оператор вопросительного знака (см. рекомендацию 3).

Бенчмарки

Рекомендация 20 утверждает, что абсолютная оптимизация кода не всегда обязательна. Однако бывают случаи, когда крайне важна именно производительность, и если это так, то будет важно измерять ее и отслеживать. Наличие регулярно выполняющихся *бенчмарков* (например, в рамках системы CI;

см. рекомендацию 32) позволяет обнаруживать, когда изменения кода или набора инструментов негативно сказываются на быстродействии.

Команда `cargo bench` выполняет особые тестовые случаи, которые циклически прогоняют определенную операцию, и сообщает среднее время обработки этой операции. На момент написания книги поддержка бенчмарков еще нестабильна, поэтому точной формой этой команды может быть `cargo +nightly bench`. (Нестабильные возможности Rust, включая используемую здесь `test`, описаны в книге *The Unstable Book*; <https://oreil.ly/tDaYl>.)

Тем не менее есть риск того, что оптимизации компилятора могут давать сбивающие с толку результаты, в частности, если вы ограничите выполняемую операцию небольшой частью реального кода. Разберем простую арифметическую функцию:

```
pub fn factorial(n: u128) -> u128 {
    match n {
        0 => 1,
        n => n * factorial(n - 1),
    }
}
```

Наивный бенчмарк для этого кода:

```
#![feature(test)]
extern crate test;

#[bench]
fn bench_factorial(b: &mut test::Bencher) {
    b.iter(|| {
        let result = factorial(15);
        assert_eq!(result, 1_307_674_368_000);
    });
}
```

дает невероятно позитивные результаты:

test bench_factorial	... bench:	0 ns/iter (+/- 0)
----------------------	------------	-------------------

Оперируя фиксированными входными данными и небольшим объемом тестируемого кода, компилятор может в рамках оптимизации исключить итерацию и сразу выдать обманчиво оптимистичный результат.

Здесь поможет функция `std::hint::black_box`. Это функция тождества, реализация которой «подталкивает, но не заставляет» (<https://oreil.ly/UEpFk>) компилятор дать пессимистичную оценку.

Бенчмарк, дополненный этим приемом:

```
#![bench]
fn bench_factorial(b: &mut test::Bencher) {
    b.iter(|| {
        let result = factorial(std::hint::black_box(15));
        assert_eq!(result, 1_307_674_368_000);
    });
}
```

дает более реалистичный результат:

```
test blackboxed::bench_factorial ... bench:           16 ns/iter (+/- 3)
```

В этом случае поможет также инструмент Godbolt Compiler Explorer (<https://rust.godbolt.org>), который покажет фактический машинный код, генерируемый компилятором. Это сделает очевидными моменты выполнения оптимизаций, неуместных для кода, выполняющего реальный сценарий.

Наконец, если вы добавляете бенчмарки для своего кода на Rust, крейт `criterion` может предоставить более удобную и полноценную альтернативу стандартной функциональности `test::bench::Bencher` с поддержкой статистики и графов.

Фаззинг-тестирование

Фаззинг-тестирование (fuzz testing, также нечеткое тестирование) — это процесс выполнения кода с рандомизированным вводом с целью обнаружить ошибки, в частности сбои, возникающие из-за определенного ввода. И хотя эта техника полезна в целом, она обретает гораздо большую важность, когда ваш код проверяется на потенциальных входных данных, которые могли быть переданы злоумышленником для атаки. Поэтому фаззинг *следует выполнять, если ваш код является потенциальной целью для взлома*.

Основная часть дефектов в коде C/C++, раскрываемых фаззерами, была связана с безопасностью памяти. Выявлялись эти дефекты обычно совмещением фаззинга с инструментами среды выполнения, например AddressSanitizer или ThreadSanitizer, либо паттернами доступа к памяти.

Rust лишен некоторых из этих проблем безопасности памяти (не всех), особенно когда присутствует `unsafe`-код (см. рекомендацию 16). Тем не менее этот язык не исключает ошибок в целом, и путь кода, активирующий `panic!` (см. рекомендацию 18), по-прежнему может стать инструментом для атаки типа «отказ в обслуживании» (denial of service, DoS), влияющей на всю кодовую базу.

Самые эффективные формы фаззинга *регулируются по охвату*: инфраструктура теста отслеживает, какие части кода выполняются, и отдает предпочтение произвольным мутациям ввода, которые проходят новыми путями кода. Чемпионом в этой технике изначально являлась программа American fuzzy lop (AFL), но в последние годы аналогичная функциональность была добавлена в набор инструментов LLVM в виде **libFuzzer**.

Компилятор Rust построен на базе LLVM, за предоставление функциональности **libFuzzer** в нем отвечает подкоманда **cargo-fuzz** (хотя работает это не для всех платформ).

Основным требованием для нечеткого тестирования является выявление в коде точки входа, которая получает ввод произвольных байтов данных (или может быть скорректирована для их получения).

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
/// Определяет, начинается ли ввод с "FUZZ"
pub fn is_fuzz(data: &[u8]) -> bool {
    if data.len() >= 3 /* oops */
    && data[0] == b'F'
    && data[1] == b'U'
    && data[2] == b'Z'
    && data[3] == b'Z'
    {
        true
    } else {
        false
    }
}
```

В книге *Rust Fuzz Book* (<https://oreil.ly/xFOEx>) описывается, как после определения целевой точки входа выстроить подпроект фаззинга. В основе этого лежит небольшой драйвер, который подключает эту точку входа к инфраструктуре фаззинга:

```
// Файл fuzz/fuzz_targets/target1.rs
#![no_main]
use libfuzzer_sys::fuzz_target;

fuzz_target!(<|data: &[u8]| {
    let _ = somecrate::is_fuzz(data);
});
```

Запуск `cargo +nightly fuzz run target1` приводит к непрерывному выполнению цели фаззинга со случайными данными, которое прекращается только при обнаружении сбоя. В нашем случае сбой обнаруживается почти сразу:

```
INFO: Running with entropic power schedule (0xFF, 100).
INFO: Seed: 1607525774
INFO: Loaded 1 modules: 1624 [0x108219fa0, 0x10821a5f8),
INFO: Loaded 1 PC tables (1624 PCs): 1624 [0x10821a5f8,0x108220b78),
INFO: 9 files found in fuzz/corpus/target1
INFO: seed corpus: files: 9 min: 1b max: 8b total: 46b rss: 38Mb
#10 INITED cov: 26 ft: 26 corp: 6/22b exec/s: 0 rss: 39Mb
thread panicked at 'index out of bounds: the len is 3 but the index is 3',
    testing/src/lib.rs:77:12
stack backtrace:
 0: rust_begin_unwind
    at /rustc/f77bfb7336f2/library/std/src/panicking.rs:579:5
 1: core::panicking::panic_fmt
    at /rustc/f77bfb7336f2/library/core/src/panicking.rs:64:14
 2: core::panicking::panic_bounds_check
    at /rustc/f77bfb7336f2/library/core/src/panicking.rs:159:5
 3: somecrate::is_fuzz
 4: _rust_fuzzer_test_input
 5: __rust_try
 6: _LLVMFuzzerTestOneInput
 7: __ZN6fuzzer6Fuzzer15ExecuteCallbackEPKhm
 8: __ZN6fuzzer6Fuzzer6RunOneEPKhmbPNS_9InputInfoEbPb
 9: __ZN6fuzzer6Fuzzer16MutateAndTestOneEv
10: __ZN6fuzzer6Fuzzer4LoopERNSt3_16vectorINS_9SizedFileENS_
    16fuzzer_allocatorIS3_EEEE
11: __ZN6fuzzer12FuzzerDriverEPiPPPcPFiPKhmE
12: _main
```

И в консоль выводится ввод, который к этому сбою привел.

Как правило, фаззинг не находит сбои так быстро, а значит, *нет смысла* выполнять подобные тесты в рамках системы CI. Такая временная неопределенность фаззинга и вытекающие из этого вычислительные затраты означают, что вам следует подумать о том, как и когда его лучше выполнять — возможно, делать это только для новых релизов или критических изменений либо ограничить его по времени¹.

Вы также можете повысить эффективность последующих выполнений фаззинга, сохраняя и повторно используя *корпус* предыдущих входных данных, которые фаззер протестировал, чтобы изучать новые пути кода. Это позволит ему при последующих выполнениях исследовать новые области, не повторяя проверку уже пройденных путей.

¹ Если ваш код является популярным крейтом с открытым исходным кодом, программа Google OSS-Fuzz может заинтересоваться выполнением его фаззинга за вас.

Советы по тестированию

Рекомендация по тестированию не будет полноценной, если не повторить некоторые типичные советы (большинство из них касаются не только Rust).

- Как неоднократно повторялось в этой рекомендации, *выполните все тесты в рамках CI при каждом изменении* (кроме фаззинга).
- При исправлении ошибки *сначала напишите тест, который ее демонстрирует, и только потом приступайте к делу*. Таким образом вы сможете убедиться, что ошибка устранена и случайно не возникнет в будущем.
- Если в вашем крейте есть `features` (функции; см. рекомендацию 26), выполните тесты для всех возможных комбинаций доступных функций.
- В более общем смысле, если в вашем крейте есть какой-либо код конфигурации, например `#[cfg(tar get_os = "windows")]`, *выполните тесты для всех платформ*, которые этот код содержит.

В текущей рекомендации было рассмотрено несколько видов тестирования, и дальше уже вам решать, какое их количество актуально и целесообразно использовать для своего проекта.

Если у вас в программе много тестируемого кода и вы публикуете крейт на crates.io, стоит подумать о том, какие тесты есть смысл в него включить. По умолчанию Cargo будет включать модульные и интеграционные тесты, бенчмарки и примеры (но не фаззинг, так как инструменты `cargo-fuzz` сохраняют их в виде отдельного крейта в подкаталоге), чего может оказаться слишком много для конечного пользователя. Если так и есть, рекомендую либо исключить (`exclude`) некоторые файлы (<https://oreil.ly/SCuug>), либо перенести проверки крейта в отдельный тестовый крейт (этот прием годится для тестов поведения).

Запомните

- Пишите модульные тесты для всеобъемлющей проверки, включающей тестирование кода, работающего только скрыто. Выполните их с помощью `cargo test`.
- Пишите тесты для проверки своих публичных API. Выполните их также с помощью `cargo test`.
- Пишите тесты документации, которые демонстрируют использование отдельных элементов вашего публичного API. Выполните их с помощью `cargo test`.
- Пишите примеры программ, которые показывают, как использовать ваш публичный API в целом. Выполните их с помощью `cargo test --examples` или `cargo run --example <name>`.

- Пишите бенчмарки, если к производительности вашего кода предъявляются высокие требования. Выполняйте их с помощью `cargo bench`.
- Пишите фаззинг-тесты, если ваш код предполагает получение недоверенного ввода. Выполняйте их регулярно с помощью `cargo fuzz`.

Рекомендация 31. Пользуйтесь преимуществами экосистемы инструментов

Экосистема Rust содержит великое множество дополнительных инструментов, которые предоставляют функциональность, выходящую за рамки базовой задачи преобразования кода Rust в машинный код.

При настройке среды разработки вам наверняка понадобится большинство следующих базовых инструментов¹:

- Cargo — для организации зависимостей (см. рекомендацию 25) и управления компилятором;
- `rustup` — для управления установленными наборами инструментов Rust;
- IDE с поддержкой Rust или плагин IDE/редактора вроде `rust-analyzer`, который позволит быстро перемещаться по кодовой базе Rust и обеспечит поддержку автозаполнения при написании кода;
- песочница Rust (<https://play.rust-lang.org>) для обособленного изучения синтаксиса Rust и обмена результатами с коллегами;
- закладка со ссылкой на документацию стандартной библиотеки Rust.

Помимо этих фундаментальных компонентов, Rust содержит множество инструментов, которые помогают решать более широкие задачи обслуживания кодовой базы и повышения ее качества. Официальный набор инструментов Cargo охватывает ряд важнейших задач, выходящих за `cargo build`, `cargo test` и `cargo run`, например следующие.

`cargo fmt`

Переформатирует код Rust в соответствии с требованиями стандарта.

`cargo check`

Проверяет этап компиляции без генерации машинного кода, что пригождается для быстрой проверки синтаксиса.

¹ В некоторых средах этот список может сократиться. Например, разработка на Rust для Android (<https://oreil.ly/pntNC>) подразумевает централизованное управление набором инструментов (то есть никакого `rustup`) и интеграцию с системой сборки Android Soong (то есть никакого Cargo).

cargo clippy

Проверяет стилистику, попутно выявляя неэффективный и неidiоматический код (см. рекомендацию 29).

cargo doc

Генерирует документацию (см. рекомендацию 27).

cargo bench

Выполняет тесты бенчмарка (см. рекомендацию 30).

cargo update

Обновляет зависимости до последних версий, выбирая такие, которые по умолчанию соответствуют принципам семантического версионирования (см. рекомендацию 21).

cargo tree

Выводит граф зависимостей (см. рекомендацию 25).

cargo metadata

Выводит метаданные пакетов, присутствующих в рабочем пространстве, а также их зависимостей.

Последний инструмент особенно полезен, хоть и косвенно: поскольку есть команда, которая выводит информацию о крейтах в четко обозначенном формате, человеку гораздо проще создать другие инструменты, использующие ее (обычно через крейт `cargo_metadata`, который предоставляет набор типов Rust для хранения метаданных). В рекомендации 25 описывались некоторые из подобных инструментов, появившихся благодаря доступности метаданных, например `cargo-udeps` (позволяет обнаруживать неиспользуемые зависимости) или `cargo-deny` (позволяет проверять множество аспектов, включая дубликаты зависимостей, допустимые лицензии и рекомендации по безопасности).

Такая расширяемость набора инструментов Rust не ограничивается одними лишь метаданными пакетов. В качестве основы для разработки инструментов можно использовать абстрактное синтаксическое дерево компилятора, зачастую при помощи крейта `syn`. Именно эта информация делает процедурные макросы (см. рекомендацию 28) столь полезными, лежит она в основе и многих других инструментов.

cargo-expand

Показывает весь исходный код, создаваемый при раскрытии макроса, что может быть важно для отладки каверзных определений макросов.

cargo-tarpaulin

Поддерживает генерацию и отслеживание информации о покрытии кода.

Любой список конкретных инструментов всегда будет субъективным, утратившим актуальность и неполным. Более универсальным советом здесь будет *изучать доступные инструменты*. Например, поиск инструмента `cargo-` выдаст десятки результатов. Некоторые окажутся неподходящими, другие уже не обслуживаются, но какие-то могут быть именно тем, что вам нужно.

Существуют также различные проекты применения к коду Rust формальной верификации (<https://oreil.ly/51DfU>), что может пригодиться, когда требуется большая уверенность в корректности кода.

Ну и напоследок напомню: если инструмент пригождается неоднократно, следует *интегрировать его в систему CI* (об этом говорится в рекомендации 32). Если инструмент быстр и не дает ложноположительных результатов, может иметь смысл *интегрировать его также в редактор или IDE*. На странице Rust Tools (<https://rust-lang.org/tools>) приведены ссылки на соответствующую документацию.

Запомните эти инструменты

Помимо инструментов, которые необходимо настроить для регулярного и автоматического применения к кодовой базе (см. рекомендацию 32), на страницах книги говорилось и о других, которые в качестве напоминания перечислены далее. Но важно помнить, что экосистема инструментов намного богаче.

- В рекомендации 16 дается совет использовать Miri при написании `unsafe`-кода.
- В рекомендациях 21 и 25 упоминается Dependabot — средство для управления обновлениями.
- В рекомендации 21 также говорится о `cargo-semver-checks` как о возможном решении для проверки корректности семантического версионирования.
- В рекомендации 28 поясняется, что `cargo-expand` может пригодиться при отладке проблем с макросами.
- Рекомендация 29 полностью посвящена использованию Clippy.
- Godbolt Compiler Explorer (<https://rust.godbolt.org>) позволяет изучать машинный код в сопоставлении с его исходной версией, о чем говорилось в рекомендации 30.
- В рекомендации 30 упоминаются также дополнительные инструменты тестирования, такие как `cargo-fuzz` для фаззинга и `criterion` для бенчмарков.
- В рекомендации 35 описывается использование `bindgen` для автоматической генерации оберток Rust FFI на основе кода C.

Рекомендация 32. Настройте систему непрерывной интеграции

Система непрерывной интеграции (continuous integration, CI) — это механизм автоматического применения инструментов к базе кода, который активируется в случае фактического или предполагаемого изменения в ней.

Настройка системы CI не эксклюзивный аспект Rust, поэтому далее будут приведены общие рекомендации по ее выполнению, дополненные советами по использованию конкретных инструментов Rust.

Этапы непрерывной интеграции

Из каких шагов должна строиться непрерывная интеграция? Очевидными фундаментальными кандидатами будут:

- сборка кода;
- выполнение тестов.

В каждом случае отдельный этап CI должен выполняться чисто, быстро, детерминированно и без ложноположительных результатов. Подробнее об этом — в следующем подразделе.

Требование детерминированности ведет к совету для этапа сборки: прописывайте фиксированную версию набора инструментов в файле `rust-toolchain.toml` вашей сборки CI. Этот файл указывает, какую версию Rust нужно использовать для сборки кода — конкретную (например, 1.70) либо «канал» (`stable`, `beta` или `nightly`), возможно дополненный датой (например, `nightly-2023-09-19`)¹. Выбор для канала плавающего значения привел бы к изменению результатов CI при появлении новых версий набора инструментов. Фиксированное значение более детерминированное и позволяет работать с обновлением набора инструментов отдельно.

В различных рекомендациях из этой книги предлагались инструменты и техники, способные помочь улучшить кодовую базу. По возможности их желательно включать в систему CI. Например, два ранее упомянутых фундаментальных этапа CI можно улучшить.

- Сборка кода.
 - В рекомендации 26 описывается использование `features` (функций) для условного включения различных частей кода. Если у вашего крейта есть

¹ Если ваш код опирается на конкретные функции, доступные только в `nightly`-версии компилятора, файл `rust-toolchain.toml` также прояснит эту зависимость от набора инструментов.

функции, *компилируйте все допустимые их комбинации в CI* (имейте в виду, что это могут быть 2^N различных вариантов, чем объясняется рекомендация избегать расплаззания функциональности).

- В рекомендации 33 предлагается по возможности делать код библиотеки совместимым с `no_std`-кодом. Вы можете быть уверены, что ваш код полностью `no_std`-совместим, только если *протестируете эту совместимость в CI*. Одно из решений — использовать возможности кросс-компиляции и делать сборку для явной `no_std`-цели (например, `thumbv6m-none-eabi`).
- В рекомендации 21 затрагивается тема объявления минимальной поддерживаемой версии Rust (minimum supported Rust version, MSRV) для вашего кода. Если вы так и сделали, то *проверяйте MSRV в CI*, добавив этап тестирования кода в отношении этой конкретной версии.
- Тестирование кода.
 - В рекомендации 30 описываются различные виды тестирования. *Выполните все виды тестов в CI*. Некоторые виды автоматически включаются в `cargo test` (модульные тесты, интеграционные и тесты документации), но другие (такие как примеры программ) могут потребовать явного включения.

Есть также дополнительные инструменты и рекомендации для повышения качества кодовой базы.

- Рекомендация 29 расписывает плюсы использования Clippy для проверки кода. *Выполните Clippy в рамках CI*. Чтобы гарантировать указание на недочеты, установите опцию `-Dwarnings`, например, командой `cargo clippy -- -Dwarnings`.
- В рекомендации 27 предлагается документировать публичный API. Проверяйте корректность сгенерированной документации и разрешения гиперссылок с помощью инструмента `cargo doc`.
- В рекомендации 25 упоминаются такие инструменты, как `cargo-udeps` и `cargo-deny`, которые помогают управлять графиком зависимостей. Применение этих инструментов в качестве этапа CI исключит возможные недочеты.
- В рекомендации 31 обсуждается экосистема инструментов Rust. Подумайте, какие из этих инструментов оправдывают регулярное применение в отношении базы кода. Например, выполнение `rustfmt/cargo fmt` в рамках CI позволяет обнаруживать код, не соответствующий руководству по стилю для вашего проекта. Чтобы гарантировать обнаружение подобных сбоев, установите опцию `--check`.

Также можете включить этапы CI, оценивающие конкретные аспекты вашего кода.

- Генерируйте статистику покрытия кода (например, с помощью `cargo-tarpaulin`), чтобы видеть, какую часть кодовой базы покрывают тесты.
- Выполняйте бенчмарки (например, с помощью `cargo-bench`; см. рекомендацию 30), чтобы оценивать производительность кода в ключевых сценариях. Но имейте в виду, что большинство систем CI работают в совместно используемых средах, где на результат могут повлиять внешние факторы. Для получения более надежных данных бенчмарка наверняка потребуется обособленная среда.

Реализовать предложенные методы анализа непросто, так как вывод этапа оценки более полезен по сравнению с предыдущими результатами. В идеале система CI должна обнаруживать, когда изменение кода протестировано не полностью или негативно сказывается на производительности. Это обычно подразумевает совмещение с внешней системой отслеживания.

Вот еще несколько советов относительно этапов CI, которые могут оказаться актуальными для вашей базы кода.

- Если вы создаете библиотеку, то помните, что любой файл `Cargo.lock`, отправленный вместе с ней в главную ветку, пользователями библиотеки будет проигнорирован (см. рекомендацию 25). Теоретически требования семантического версионирования (см. рекомендацию 21) в `Cargo.toml` должны означать, что все будет работать корректно в любом случае. На практике же подумайте о том, чтобы включить в CI этап, который будет выполнять сборку без локального `Cargo.toml` с целью убедиться в корректности работы текущих версий зависимостей.
- Если ваш проект содержит какие-либо машинно-генерируемые ресурсы, подверженные версионированию (например, код, который `prost` генерирует из сообщений буфера протокола), то включите в CI этап, восстанавливающий эти ресурсы и проверяющий отсутствие отличий от версии, отправленной в главную ветку.
- Если ваша кодовая база содержит код, специфичный для конкретной платформы, например `#[cfg(target_arch = "arm")]`, выполните в рамках CI этапы, которые будут подтверждать, что такой код компилируется и в идеале работает на этой платформе. (Первое реализовать проще, так как в наборе инструментов Rust есть поддержка кросс-компиляции.)
- Если ваш проект работает с секретными значениями, такими как токены доступа или криптографические ключи, подумайте о включении этапа CI,

который будет искать в базе кода секреты, случайно отправленные в главную ветку. Это особенно важно, если проект публичный, в случае чего есть смысл переместить проверку из системы CI на этап проверки версий перед коммитом кода.

Проверки CI не всегда нужно интегрировать с Cargo и набором инструментов Rust. Иногда простой скрипт оболочки может оказаться эффективнее, особенно когда у кодовой базы есть локальное соглашение, которое не выполняется по-всеместно. Например, это может быть требование того, чтобы любое вызывающее панику выполнение метода (см. рекомендацию 18) сопровождалось комментарием или чтобы каждый комментарий `TODO`: имел владельца — человека или ID для отслеживания. Для проверки всего этого идеально подойдет именно скрипт оболочки.

Наконец, рекомендую изучить системы CI публичных проектов Rust, чтобы лучше понять, какие дополнительные этапы непрерывной интеграции могут оказаться полезными для вашего проекта. Например, много интересных идей можно почерпнуть из системы CI, используемой Cargo (<https://oreil.ly/DOqCc>).

Принципы CI

Теперь перейдем от частного к общему. Существует ряд принципов, которые должны определять детали вашей системы CI. Основной звучит так: *не тратить впустую время людей*. Если система CI необоснованно тратит время человека, он начнет искать способ обойти это.

Больше всего в этом смысле инженеров раздражают *нестабильные тесты* (flaky test), которые иногда проходят, а иногда — нет, даже при одинаковой базе кода и конфигурации. Страйтесь избавляться от ненадежных тестов: отслеживайте их и не жалейте времени, чтобы найти и устраниТЬ причину нестабильности — это обязательно оправдается в долгосрочной перспективе.

Еще одна типичная трата времени возникает из-за систем CI, которые долго выполняются и делают это только *после* запроса ревью кода. В такой ситуации есть риск впустую потратить время двух людей: автора и ревьюера кода. Последний в этом случае будет искать в коде проблемы, часть которых можно было обнаружить с помощью специальных программ CI.

Чтобы этого избежать, страйтесь реализовать простое выполнение проверок в рамках CI независимо от автоматизированной системы. Это позволит инженерам выработать привычку проводить их регулярно, чтобы ревьюерам кода не встречались проблемы, которые могла обозначить система CI. Еще лучше сделать интеграцию более непрерывной, добавив в редактор или IDE инструменты,

чтобы, например, плохо отформатированный код не попадал даже на диск. Для этого может потребоваться разделить проверки, если среди них будут длительные тесты, которые редко находят проблемы, но служат защитой от сбоя в сомнительных сценариях.

В более общем смысле крупный проект может потребовать разделения проверок CI в соответствии с порядком их выполнения на следующие виды.

- Проверки, интегрированные в среду разработки каждого инженера (например, `rustfmt`).
- Проверки, которые выполняются при каждом запросе код-ревью (например, `cargo build`, `cargo clippy`) и легко производятся вручную.
- Проверки, которые выполняются при каждом изменении, которое попадает в главную ветку проекта (например, полноценная `cargo test` во всех поддерживаемых средах).
- Проверки, которые выполняются по заданному графику, допустим, ежедневно или раз в неделю, и могут перехватывать редкие недочеты после их возникновения (например, длительные интеграционные тесты и сравнительные бенчмарки).
- Проверки, которые выполняются в отношении текущего кода всегда (например, фаззинг).

Важно, чтобы система CI была интегрирована с любой системой код-ревью вашего проекта. Тогда ревьюеры смогут четко видеть зеленые отметки тестов и будут уверены, что анализируют важные аспекты кода, а не тривиальные детали.

Подобная потребность в «зеленой» сборке означает, что ни одной проверкой из включенных в систему CI нельзя пренебречь. И это оправданно, даже если вам приходится находить обходное решение в случае периодического ложного срабатывания одного из инструментов. Если в вашей системе CI возникает допустимый провал даже одной проверки («Да все знают, что этот тест никогда не проходит»), становится намного труднее отслеживать новые недочеты.

В рекомендации 30 приводился распространенный совет добавлять тесты для воспроизведения ошибки перед ее исправлением. Тот же принцип относится и к системе CI: *прежде чем исправлять обнаруженную в процессе проблему, добавьте этап CI, который будет ее выявлять*. Например, если вы выясните, что какой-то автоматически генерируемый код утратил синхронность со своим источником, добавьте в систему CI соответствующую проверку. Сначала она будет проваливаться, но после устранения проблемы получит зеленую отметку, тем самым обеспечив вам уверенность в том, что подобных ошибок процесса в будущем не возникнет.

Публичные системы CI

Если ваша кодовая база имеет открытые исходники и является публично доступной, в отношении системы CI нужно учесть еще некоторые нюансы.

Начнем с хорошего. Существует множество надежных бесплатных вариантов создания системы CI для открытого исходного кода. На момент написания книги лучшим решением, пожалуй, является GitHub Actions, но это далеко не единственный вариант, и постоянно появляются новые.

Второй важный момент — то, что система CI должна выступать в качестве руководства для конфигурирования всех необходимых кодовой базе предварительных требований. И это касается не только чистых крейтов Rust. Если вашей базе кода нужны внешние зависимости — базы данных, альтернативные наборы инструментов для кода FFI, конфигурация и т. д., тогда скрипты CI станут доказательством того, что все это будет работать в свежей системе. Заключение этих этапов настройки в скрипты позволит людям и программам получать рабочую систему в результате простых действий.

Наконец, есть и плохие новости, которые касаются видимых всем крейтов, — существует вероятность нецелевого использования и атак. Их спектр может простираться от попыток внедрить в вашу систему CI майнера криптовалют до кражи токенов доступа базы кода (<https://oreil.ly/yP-8m>), атак на цепочки поставок и даже хуже. Чтобы снизить подобные риски, ориентируйтесь на следующие советы.

- Ограничьте доступ, чтобы скрипты CI выполнялись автоматически только для известных участников и требовали ручного выполнения для новых.
- Фиксируйте для внешних скриптов конкретные версии, а еще лучше закрепляйте их за конкретными известными хешами.
- Внимательно отслеживайте все этапы интеграции, которые требуют не только права чтения базы кода.

ГЛАВА 6

За пределами стандартного Rust

Набор инструментов Rust поддерживает не только прикладной код, написанный на этом языке, но и множество других рабочих сред, выполняющихся в пространстве пользователя.

- Он поддерживает *кросс-компиляцию*, когда система, в которой применяется набор инструментов (*хост*), отличается от системы, в которой будет выполняться скомпилированный код (*цель*). Это упрощает сборку кода для встраиваемых систем.
- Он поддерживает линковку с кодом, скомпилированным из других языков, посредством встроенных возможностей FFI.
- Он поддерживает конфигурации без полноценной стандартной библиотеки `std`, позволяя компилировать код для систем, не имеющих полноценной операционной системы, например таких, где нет файловой системы или сетевых возможностей.
- Он даже поддерживает конфигурации, лишенные возможности аллокации в куче и имеющие только стек (исключая использование стандартной библиотеки `alloc`).

В этих нестандартных средах Rust бывает сложно работать, и они менее безопасны, а иногда и откровенно `unsafe`, но зато предоставляют больше возможностей для реализации нужных задач.

В текущей главе речь пойдет всего о нескольких основных принципах работы в этих средах. Помимо освоения основ, вам также потребуется ознакомиться с предметной документацией, например *Rustonomicon*.

Рекомендация 33. Рассмотрите вариант сделать код библиотеки no_std-совместимым

У Rust есть стандартная библиотека `std`, которая содержит код для широкого спектра типичных задач — от стандартных структур данных до сетевых взаимодействий, от поддержки многопоточности до файлового ввода/вывода. Для удобства некоторые элементы `std` автоматически импортируются в вашу программу через вводный модуль (*prelude module*) — набор распространенных инструкций `use`, которые делают доступными наиболее популярные типы, не используя их полные имена, например `Vec` вместо `std::vec::Vec`.

Rust также поддерживает сборку кода для сред, где нет возможности предоставить для работы всю стандартную библиотеку, таких как загрузчики, прошивки или встраиваемые платформы в целом. О том, что крейт нужно собирать именно таким образом, сообщает атрибут `#![no_std]` в начале файла `src/lib.rs`.

В текущей рекомендации мы разберем, что конкретно утрачивается при сборке для `no_std` и на какие библиотечные функции вы по-прежнему можете полагаться — их, оказывается, довольно много. Тем не менее эта рекомендация посвящена конкретно поддержке `no_std` в коде *библиотек*. Сложности создания *двоичного* файла `no_std` останутся за кадром¹, акцент будет сделан именно на том, как обеспечить доступность библиотечного кода для тех несчастных, кому приходится работать в подобных ограниченных средах.

core

Даже при сборке для самых ограниченных платформ нам по-прежнему остаются доступны многие фундаментальные типы из стандартной библиотеки, например `Option` и `Result`, хоть и под другими именами. То же касается различных видов `Iterator`.

Другие имена для этих базовых типов начинаются с `core::`, указывая на то, что они взяты из библиотеки `core` — стандартной библиотеки, доступной даже в самых нестандартных средах. Эти типы `core::` действуют точно так же, как их `std::`-аналоги, поскольку фактически являются теми же типами — во всех случаях `std::`-версия представляет собой просто реэкспорт соответствующего типа `core::`.

¹ Подробнее о том, из чего строится процесс создания двоичного файла `no_std`, читайте в книге *The Embedonomicon* (<https://oreil.ly/x74WK>) или в старой статье Филиппа Оппермана (Philipp Oppermann) (<https://oreil.ly/WTn-j>).

Это означает, что есть быстрый и грубый способ понять, доступен ли `std`:- элемент в `no_std`-среде, — найти на сайте doc.rust-lang.org интересующий `std`- элемент и перейти по ссылке `source` (в верхнем правом углу)¹. Если она перенесет вас в `src/core/...`, значит, элемент доступен в `no_std` через `core::`.

Все программы Rust автоматически получают доступ к типам из `core`. Тем не менее в `no_std`-среде их обычно нужно использовать (`use`) явно, так как вводный `std` отсутствует.

На практике расчет исключительно на `core` во многих средах, даже `no_std`, окажется ограничивающим фактором. Основным ограничением библиотеки `core` является отсутствие возможности *аллокации в куче*.

И хотя Rust превосходно справляется с размещением элементов в стеке и безопасным отслеживанием соответствующих сроков жизни (см. рекомендацию 14), это ограничение по-прежнему означает, что стандартные структуры данных — векторы, карты, множества — предоставить нельзя, так как пространство под их содержимое должно выделяться в куче. Это существенно снижает количество доступных крейтов, работающих в данной среде.

alloc

Однако если `no_std`-среда *поддерживает* выделение в куче, тогда поддержка многих стандартных структур данных из `std` сохранится. Эти структуры данных вместе с другой функциональностью, использующей выделение памяти в куче, сгруппированы в библиотеке Rust `alloc`.

Как и в случае с `core`, эти варианты `alloc` фактически являются теми же типами. Например, реальное имя `std::vec::Vec` на деле выглядит как `alloc::vec::Vec`.

`no_std`-крайт Rust необходимо явно настраивать на использование `alloc`, добавляя в `src/lib.rs` объявление `extern crate alloc`²:

```
//! Мой `no_std`-совместимый крайт
#![no_std]

// Требует `alloc`
extern crate alloc;
```

¹ Имейте в виду, что иногда могут возникнуть осечки. Например, на момент написания книги крайт `Error` определен в `core::`, но обозначен как нестабильный. Стабильна только его `std::`-версия.

² До Rust 2018 для подтягивания зависимостей использовались объявления `extern crate`. Сейчас это полностью обрабатывается через файл `Cargo.toml`, но механизм `extern crate` по-прежнему применяется для подтягивания тех частей стандартной библиотеки Rust — крайтов `sysroot` (<https://oreil.ly/sJzAv>), — которые для `no_std`-сред необязательны.

Подтягивание крейта `alloc` делает доступными множество знакомых полей, которые теперь адресуются по своим истинным именам:

- `alloc::boxed::Box<T>;`
- `alloc::rc::Rc<T>;`
- `alloc::sync::Arc<T>;`
- `alloc::vec::Vec<T>;`
- `alloc::string::String;`
- `alloc::format!;`
- `alloc::collections::BTreeMap<K, V>;`
- `alloc::collections::BTreeSet<T>;`

Доступность всего этого позволяет многим библиотечным крейтам получать поддержку `no_std` — например, если библиотека не подразумевает ввод/вывод или сетевое взаимодействие.

Однако в предоставляемых `alloc` структурах данных есть значительный пробел — коллекции `HashMap` и `HashSet` относятся к `std`, а не к `alloc`. Дело в том, что эти контейнеры для защиты от коллизионных атак опираются на случайные начальные числа, а безопасная случайная генерация требует содействия со стороны операционной системы, существование которой `alloc` прогнозировать не может.

Еще одним значимым пробелом является функциональность синхронизации наподобие `std::sync::Mutex`, необходимой для многопоточного кода (см. рекомендацию 17). Эти типы относятся именно к `std`, так как опираются на соответствующие ОС примитивы синхронизации, которые без ОС недоступны. Если вам нужно написать код, который будет поддерживать и `no_std`, и многопоточность, то решением станут такие крейты, как `spin`.

Написание кода для `no_std`

Ранее мы прояснили, что для *некоторых* библиотечных крейтов обеспечение совместимости с `no_std` подразумевает лишь:

- замену `std::`-типов идентичными им крейтами `core::` или `alloc::`, что ввиду отсутствия вводного `std` требует использования (`use`) полного имени типа;
- переход с `HashMap`/`HashSet` на `BTreeMap`/`BTreeSet`.

Но актуально это, только если все крейты, от которых вы зависите (см. рекомендацию 25), тоже совместимы с `no_std` — нет смысла добиваться `no_std`-совместимости, если какой-то пользователь вашего крейта все равно вынужден применять `std`.

Есть здесь и один подвох: компилятор Rust не сообщит, если ваш `no_std`-крейт зависит от библиотеки, использующей `std`. Это означает, что можно легко перечеркнуть работу по обеспечению `no_std`-совместимости для крейта, просто добавив или обновив зависимость, подтягивающую `std`.

Чтобы от этого застраховаться, *добавьте в систему CI проверку no_std-сборки*, чтобы получать соответствующее предупреждение (см. рекомендацию 32). Набор инструментов Rust базово поддерживает кросс-компиляцию, так что для этого может быть достаточно кросс-компилирования (<https://oreil.ly/DAbwt>) под целевую систему без поддержки `std` (например, `--target thumbv6m-none-eabi`). Тогда любой код, который случайно начнет требовать `std`, для этой цели компилироваться не будет.

Итак, если ваши зависимости поддерживают это и достаточно простых описанных ранее преобразований, *подумайте о том, чтобы сделать код библиотеки no_std-совместимым*. Если такое окажется возможным, то большой работы не потребуется, зато вы получите библиотеку, пригодную для максимально широкого применения.

Если же эти преобразования *не охватывают* весь код вашего крейта, но незатронутые фрагменты составляют лишь его малую и локализованную часть, рассмотрите вариант добавления в крейт функции (см. рекомендацию 26), которая будет включать только эти фрагменты.

Такая функция по соглашению называется либо `std`, если позволяет использовать функциональность, присущую `std`:

```
#![cfg_attr(not(feature = "std"), no_std)]
```

либо `alloc`, если делает доступной функциональность на основе `alloc`:

```
#[cfg(feature = "alloc")]
extern crate alloc;
```

Но здесь есть нюанс, требующий внимательности: не создавайте функцию `no_std`, которая будет *отключать* функциональность, требующую `std` (то же касается функции `no_alloc`). Как говорилось в рекомендации 26, функции должны быть аdditивными, и никак не получится обслужить двух пользователей крейта, из которых один налаживает работу с `no_std`, а другой — нет, так как первый будет инициировать удаление кода, на который опирается второй.

Как и в случае с кодом, привязанным к функциональности, обеспечьте для системы CI (см. рекомендацию 32) возможность собирать все актуальные комбинации, включая сборку с отключением функции `std` для явно `no_std`-платформы.

Выделение с возможностью сбоя

В предыдущих разделах текущей рекомендации были рассмотрены две no_std-среды: полностью встраиваемая среда без выделения памяти в куче (аллокации) (`core`) и более общая среда, где выделение памяти в куче доступно (`core +alloc`).

Однако между двумя этими вариантами есть и другие значимые среды, в частности такие, где выделение памяти в куче возможно, но может проваливаться из-за ее ограниченного объема.

К сожалению, стандартная библиотека Rust `alloc` во всех случаях предполагает, что выделение памяти в куче провалиться не может, но это предположение не всегда верно.

Даже простое использование `alloc::vec::Vec` потенциально может обусловить выделение памяти на каждой строке:

```
let mut v = Vec::new();
v.push(1); // может выделять память
v.push(2); // может выделять память
v.push(3); // может выделять память
v.push(4); // может выделять память
```

Ни одна из этих операций не возвращает `Result`, так что же происходит, если эти выделения памяти проваливаются?

Ответ будет зависеть от набора инструментов, цели и конфигурации (<https://oreil.ly/5kPxH>), но наверняка приведет к `panic!` и завершению программы. И определенно нет ответа, при котором можно было бы обработать сбой выделения памяти на строке 3 так, чтобы программа перешла к строке 4.

Это предположение о безотказном выделении памяти обеспечивает хорошую эргономику для кода, который выполняется в типичном пространстве пользователя, где объем доступной памяти, по сути, неограничен или где ее исчерпание говорит о том, что в компьютере есть другие, более серьезные проблемы.

Как бы то ни было, безотказное выделение памяти никак не подходит для кода, который должен выполняться в средах, где памяти мало и программы вынуждены как-то обходиться ее ограниченным объемом. Это та редкая область, в которой лучшую поддержку обеспечивают более старые и менее безопасные в смысле памяти языки.

- С является достаточно низкоуровневым для того, чтобы выделение памяти в нем производилось вручную, в связи с чем возвращаемое от `malloc` значение можно проверить на `NULL`.

- C++ может использовать свой механизм исключений для перехвата сбоев выделения памяти в виде `std::bad_alloc`¹.

Эта неспособность стандартной библиотеки Rust справляться со сбоями при выделении памяти неоднократно озвучивалась в популярных и значимых контекстах, например в ядре Linux (<https://oreil.ly/hzzCR>), Android и инструменте Curl (<https://oreil.ly/jvfOw>), так что работа по исправлению этого недочета ведется.

Первым шагом стали изменения в выделении памяти коллекций с возможностью сбоя (<https://oreil.ly/gyhpR>), которые добавили допускающие сбой альтернативы для многих API коллекций, включающих выделение памяти. Это, как правило, сопряжено с добавлением варианта `try_<operation>`, который приводит к `Result<_, AllocError>`, например:

- `Vec::try_reserve` доступен в качестве альтернативы для `Vec::reserve`;
- `Box::try_new` доступен (в наборе инструментов версии `nightly`) в качестве альтернативы для `Box::new`.

Собственно, этим допускающие сбой API и ограничиваются. Например, пока что нет допускающего сбой эквивалента для `Vec::push`, поэтому код, который собирает вектор, может потребовать тщательных вычислений, чтобы исключить ошибки при выделении памяти:

```
fn try_build_a_vec() -> Result<Vec<u8>, String> {
    let mut v = Vec::new();

    // Выполняет тщательные вычисления, чтобы определить
    // необходимый объем пространства. Здесь код упрощен до...
    let required_size = 4;

    v.try_reserve(required_size)
        .map_err(|e| format!("Failed to allocate {} items!", required_size))?;

    // Теперь мы знаем, что можно безопасно выполнить:
    v.push(1);
    v.push(2);
    v.push(3);
    v.push(4);

    OK(v)
}
```

¹ Также можно добавить перегрузку `std::nothrow` в вызовы `new` и проверку для возвращаемых значений `nullptr`. Тем не менее есть контейнеры вроде `vector<T>::push_back`, которые выделяют память скрыто, а значит, могут сообщить о ее сбое только через исключение.

Помимо добавления допускающих сбой точек выделения памяти, можно также убрать *безотказные* операции выделения памяти, отключив флаг конфигурации `no_global_oom_handling` (по умолчанию включен). Среды с ограниченным пространством кучи могут явно отключать этот флаг, гарантируя, что в код случайно не просочится ни один эпизод использования безотказного выделения памяти.

Запомните

- Многие элементы в крейте `std` в действительности взяты из `core` или `alloc`.
- В результате сделать код библиотеки совместимым с `no_std` может оказаться проще, чем кажется.
- Убедитесь, что `no_std`-код таковым и остается, проверив его через систему CI.
- Имейте в виду, что работа в среде с ограниченным пространством кучи на данный момент имеет ограниченную поддержку библиотек.

Рекомендация 34. Контролируйте выход за границы FFI

Несмотря на то что Rust располагает обширной стандартной библиотекой и активно развивающейся экосистемой крейтов, другого кода в мире все равно намного больше.

Как и другие молодые языки, Rust помогает решить эту проблему, предлагая механизм *интерфейса внешних функций* (foreign function interface, FFI), который делает возможным взаимодействие с кодом и структурами данных, написанными на других языках, — несмотря на свое название, FFI не ограничивается одними лишь функциями. И это позволяет использовать существующие библиотеки других языков, и не только те, которым повезло стать частью инициативы «перепишите на Rust» (rewrite it in Rust, RiiR).

По умолчанию Rust, как и другие языки, нацелен на интероперабельность с C. Основная причина в широкой распространенности всевозможных библиотек на этом языке, а также в его простоте: C выступает в качестве «наименьшего общего знаменателя» интероперабельности, так как не требует от набора инструментов поддержки более продвинутой функциональности, которая потребовалась бы для совместимости с другими языками (например, сборки мусора для Java или Go, исключений и шаблонов для C++, переопределений функций для Java и C++ и т. д.).

Тем не менее это не значит, что добиться интероперабельности с чистым С легко. При добавлении кода, написанного на другом языке, попадают под удар все предоставляемые Rust гарантии и защиты, особенно касающиеся безопасности памяти.

В итоге код FFI на Rust автоматически становится `unsafe`, и советом из рекомендации 16 приходится пренебречь. В текущей рекомендации я дам альтернативный совет, а в рекомендации 35 речь пойдет об инструментах, помогающих избежать некоторых (не всех) «граблей», возникающих во время работы с FFI. (Глава Rustonomicon, посвященная FFI (<https://oreil.ly/2jHBA>), также содержит полезные советы и сопутствующую информацию.)

Вызов функций C из Rust

Простейшим взаимодействием через FFI является вызов из кода Rust функции C, получающей непосредственные аргументы, которые не содержат указателей, ссылок или адресов в памяти:

```
/* Файл lib.c */
#include "lib.h"

/* Определение функции C */
int add(int x, int y) {
    return x + y;
}
```

Этот код C предоставляет *определение* функции и обычно сопровождается заголовочным файлом, содержащим *объявление* этой функции, позволяющее другому коду C ее использовать:

```
/* Файл lib.h */
#ifndef LIB_H
#define LIB_H

/* Объявление функции C */
int add(int x, int y);

#endif /* LIB_H */
```

Это объявление гласит примерно следующее: «Где-то есть функция `add`, которая получает два целых числа и возвращает другое целое число». Это позволяет коду на C использовать функцию `add` с условием, что фактический код `add` будет предоставлен позднее — в частности, на этапе линковки.

Код на Rust, который хочет использовать `add`, должен содержать аналогичное объявление с той же целью описать сигнатуру функции и указать, что соответствующий код будет доступен позже:

```
use std::os::raw::c_int;
extern "C" {
    pub fn add(x: c_int, y: c_int) -> c_int;
}
```

Это объявление обозначено как `extern "C"`, указывая, что код функции предоставит внешняя библиотека C¹. Маркер `extern "C"` также автоматически отмечает функцию как `no_mangle`, о чем мы будем говорить в пункте «Декорирование имен» далее.

Механизмы линковки

Детали процесса генерации инструментами С внешней библиотеки и ее формата зависят от конкретной среды и выходят за рамки тематики нашей книги. Тем не менее простым вариантом, типичным для Unix-подобных систем, является *статичный файл библиотеки*, который обычно принимает форму `lib<something>.a` (например, `libcffi.a`) и который можно сгенерировать при помощи инструмента `ar`.

В итоге система сборки Rust требует указания того, какая библиотека содержит нужный код С. Сделать это можно в атрибуте `link` (https://oreil.ly/_N-Fv) в коде:

```
#[link(name = "cffi")] // Необходима внешняя библиотека наподобие `libcffi.a`
extern "C" {
    // ...
}
```

или с помощью скрипта сборки, отправляющего инструкцию `cargo:rustc-link-lib` в `cargo`²:

```
// Файл build.rs
fn main() {
    // Необходима внешняя библиотека наподобие `libcffi.a`
    println!("cargo:rustc-link-lib=cffi");
}
```

¹ Если функциональность FFI, которую вы хотите использовать, является частью стандартной библиотеки С, то создавать эти объявления не нужно — их уже предоставляет крейт `libc`.

² Соответствующий ключ `links` в манифесте `Cargo.toml` поможет сделать эту зависимость видимой для `Cargo`.

Второй вариант более гибок, так как скрипт сборки может проверить свою среду и придерживаться поведения, соответствующего результату.

В любом случае системе сборки Rust потребуется сообщить, как найти эту библиотеку C, если только она не расположена в стандартном месте. Указать адрес библиотеки можно через скрипт сборки, отправляющий в Cargo инструкцию `cargo:rustc-link-search` с адресом библиотеки:

```
// Файл build.rs
fn main() {
    // ...

    // Извлекает местоположение `Cargo.toml`
    let dir = std::env::var("CARGO_MANIFEST_DIR").unwrap();
    // Ищет нативные библиотеки в каталоге одним уровнем выше
    println!(
        "cargo:rustc-link-search=native={}",
        std::path::Path::new(&dir).join("..").display()
    );
}
```

Нюансы кода

Возвращаясь к исходному коду, отмечу, что даже у этих простейших примеров есть подвох. В первую очередь использование FFI-функций неизбежно будет `unsafe`:

```
let x = add(1, 1);

error[E0133]: call to unsafe function is unsafe and requires unsafe function
              or block
--> src/main.rs:176:13
   |
176 |     let x = add(1, 1);
   |           ^^^^^^^^^^ call to unsafe function
   |
   = note: consult the function's documentation for information on how to avoid
         undefined behavior
```

а значит, требует заключения в `unsafe { }.`

Следующим нюансом является использование типа C `int`, представленного как `std::os::raw::c_int`. Насколько большой этот `int`? Здесь можно сказать, что два следующих значения, *пожалуй*, будут одинаковыми:

- размер `int` для набора инструментов, компилировавшего библиотеку C;
- размер `std::os::raw::c_int` для набора инструментов Rust.

Но зачем испытывать судьбу? *Старайтесь использовать на границе FFI типы установленного размера*, что для C означает применение типов (например, `uint32_t`), определенных в `<stdint.h>`. Однако если вы работаете с уже существующей базой кода, где задействуется `int/long/size_t`, такой роскоши у вас может не быть.

Последним практическим нюансом является то, что код на C и равноценное ему объявление на Rust должны совпадать. Хуже того, если будет расхождение, инструменты сборки об этом не сообщат, а просто молча сгенерируют некорректный код.

В рекомендации 35 обсуждается решение этой проблемы с помощью инструмента `bindgen`, но все же желательно разобраться в происходящем, чтобы понимать, почему инструменты сборки не могут обнаружить проблему сами. В частности, нужно вникнуть в основы механизма *декорирования имен*.

Декорирование имен

Компилируемые языки обычно поддерживают *отдельную компиляцию*, когда разные части программы преобразуются в машинный код в виде отдельных фрагментов (объектных модулей), которые затем совмещаются в цельную программу *линковщиком*. Это означает, что если изменится лишь одна небольшая часть исходного кода программы, то заново сгенерировать придется всего один объектный модуль. На этапе линковки программа будет пересобрана, чтобы измененный модуль совместился с остальными, которые остались без изменений.

Этап линковки, грубо говоря, является операцией проведения связей (<https://oreil.ly/gzfEl>): некоторые объектные модули предоставляют определения функций и переменных, другие содержат маркеры-заглушки, указывающие на ожидание определения из другого объекта, который на этапе компиляции доступен не был. Компоновщик совмещает эти два вида файлов: он обеспечивает, чтобы все заглушки в скомпилированном коде были заменены ссылкой на соответствующее определение конкретной функции.

Сопоставление между заглушками и определениями компоновщик делает путем простой проверки совпадения названий, то есть для всех сопоставлений используется общее пространство имен. Традиционно такой механизм отлично работал для линковки программ на языке C, где никакое имя не могло быть задействовано повторно где-то еще — в объектном модуле указывалось конкретное имя функции. (В результате типичным соглашением для библиотек C является ручное добавление ко всем символам префикса, чтобы `lib1_process` не конфликтовал с `lib2_process`.)

Появление C++ создало проблему, так как этот язык допускает использование измененных определений с одним именем:

```
// Код на C++
namespace ns1 {
    int32_t add(int32_t a, int32_t b) { return a+b; }
    int64_t add(int64_t a, int64_t b) { return a+b; }
}
namespace ns2 {
    int32_t add(int32_t a, int32_t b) { return a+b; }
}
```

Решением этой проблемы стало *декорирование имен*: компилятор кодирует сигнатуру и информацию о типе переопределяемой функции в имя, записываемое в объектный модуль, а компоновщик продолжает выполнять банальное сопоставление между заглушками и определениями один к одному.

В Unix-подобных системах увидеть, с чем работает компоновщик, можно с помощью инструмента nm:

```
% nm ffi-lib.o | grep add # Что линковщик видит для C
0000000000000000 T _add
```

```
% nm ffi-cpp-lib.o | grep add # Что компоновщик видит для C++
0000000000000000 T ZN3ns13addEii
0000000000000020 T ZN3ns13addExx
0000000000000040 T ZN3ns23addEii
```

В этом случае он показывает три декорированных символа, которые относятся к коду (T указывает на *текстовый раздел* двоичного файла, являющийся стандартным именем для места расположения кода).

Инструмент c++filt помогает перевести это обратно в представление, которое было бы видимо в коде на C++:

```
% nm ffi-cpp-lib.o | grep add | c++filt # Что видит программист
0000000000000000 T ns1::add(int, int)
0000000000000020 T ns1::add(long long, long long)
0000000000000040 T ns2::add(int, int)
```

Так как декорированное имя содержит информацию о типе, компоновщик будет ругаться на любое несоответствие этой информации между заглушкой и определением. Это обеспечивает некую степень безопасности типов: если определение изменится, но использующую его область не обновят, набор инструментов будет ругаться.

Возвращаясь к Rust: внешние функции `extern "C"` неявно обозначаются как `#[no_mangle]`, и этот символ в объектном файле является голым именем, которым он являлся бы и для программы на C. Это означает, что безопасность типов сигнатур функции утрачена: так как компоновщик видит только голые имена

функций, то в случае расхождения между ожидаемыми типами в определении и месте фактического использования он молча продолжит линковку и проблемы возникнут только в среде выполнения.

Доступ к данным C из Rust

В примере с `add` на C в предыдущем разделе между Rust и C передавался простейший тип данных — целое число, которое умещается в машинный регистр. Даже в этом случае есть о чём беспокоиться, поэтому неудивительно, что работа с более сложными структурами данных также имеет свои опасные нюансы.

И в C, и в Rust для совмещения связанных данных в одну структуру используется `struct`. Тем не менее, когда `struct` реализуется в памяти, эти два языка могут решить разместить разные поля в разных местах или даже в разном порядке (*схема*; <https://oreil.ly/cjkXO>). Чтобы избежать расхождений, *используйте для типов Rust, задействованных в FFI, конструкцию `#[repr(C)]`*. Этот атрибут создан для обеспечения интероперабельности с C:

```
/* Определение структуры данных на C */
/* Произведенные здесь изменения должны быть отражены в lib.rs */
typedef struct {
    uint8_t byte;
    uint32_t integer;
} FfiStruct;

// Эквивалентная структура данных на Rust
// Произведенные здесь изменения должны быть отражены в lib.h / lib.c
#[repr(C)]
pub struct FfiStruct {
    pub byte: u8,
    pub integer: u32,
}
```

Эти определения структур содержат комментарий, напоминающий программистам, что два обозначенных фрагмента должны быть согласованными. Упование на неусыпную бдительность человека рано или поздно приведет к проблемам. Что касается сигнатур функций, то лучше автоматизировать эту синхронизацию между двумя языками с помощью инструмента наподобие `bindgen` (см. рекомендацию 35).

Особой внимательности в ходе работы с FFI требует строковый тип данных. Базовые определения строки в Rust и C немного различаются:

- `String` в Rust содержит данные в кодировке UTF-8, возможно включая нулевые байты с явно известной длиной;
- в C строка (`char *`) содержит байтовые значения (знаковые или беззнаковые), длина которых определяется первым нулевым байтом (`\0`) данных.

К счастью, работать в Rust со строками в стиле C довольно легко, так как создатели библиотеки Rust уже все подготовили, предоставив пару типов для их кодирования. Для хранения (присвоенных) строк, которые должны поддерживать интероперабельность с C, используйте тип `CString`, а в работе с заимствованными строковыми значениями — тип `CStr`. Второй тип включает метод `as_ptr()`, который можно задействовать для передачи содержимого строки любой FFI-функции, ожидающей строку C `const char*`. Имейте в виду, что ключевое слово `const` важно — его нельзя использовать для FFI-функции, которая должна изменять содержимое (`char *`) передаваемой ей строки.

Время жизни

Большинство структур данных слишком велики для размещения в регистре, поэтому помещаются в память. В результате доступ к ним происходит с помощью поиска в памяти. В терминологии C это означает использование *указателя* — числа, кодирующего адрес памяти, — без добавления какой-либо другой семантики (см. рекомендацию 8).

В Rust расположение в памяти обычно выражается в виде *ссылки*, и ее численное значение можно извлечь в виде *сырого указателя*, готового для передачи на границу FFI:

```
extern "C" {
    // Функция C, выполняющая операцию с содержимым `FfiStruct`
    pub fn use_struct(v: *const FfiStruct) -> u32;
}

let v = FfiStruct {
    byte: 1,
    integer: 42,
};
let x = unsafe { use_struct(&v as *const FfiStruct) };
```

Тем не менее ссылки в Rust несут дополнительные ограничения, связанные с *временем жизни* соответствующего фрагмента памяти, о чем говорилось в рекомендации 14. В ходе преобразования в сырой указатель эти ограничения исчезают.

В итоге использование сырых указателей по своему характеру является `unsafe`, предупреждая об опасности: код C по ту сторону границы FFI может проделать огромное множество операций, способных разрушить предоставленные Rust гарантии безопасности памяти.

- Он может задержать значение некоего указателя и задействовать его позднее, когда связанная с ним область памяти будет либо освобождена в куче, либо повторно применена уже в стеке (это называется *использованием данных после освобождения памяти*).

- Он может решить отказаться от неизменности (`const`) переданного ему указателя и изменить данные, которые, по ожиданиям Rust, должны быть иммутабельны.
- К коду на С не применяются защиты Rust в виде `Mutex`, что ведет к потенциальному возникновению гонок данных (см. рекомендацию 17).
- Код на С может по ошибке вернуть аллокатору выделенную в куче память, вызвав библиотечную функцию С `free()`, то есть код на Rust после этого может начать выполнять операции использования данных после освобождения памяти (`use-after-free`).

Все эти опасности следует учитывать при анализе соотношения рисков и пользы от применения библиотеки через FFI. Из плюсов вы получаете возможность использовать существующий код, предположительно, в хорошем рабочем состоянии, для чего потребуется лишь написать или сгенерировать соответствующие объявления. Из минусов же — вы теряете защиту памяти, которая во многом является причиной использования Rust в принципе.

В качестве первого шага для снижения вероятности возникновения проблем, связанных с памятью, рекомендую *выделять и освобождать память по одну сторону границы FFI*. Например, следующий код может выглядеть как симметричная пара функций:

```
/* Функции С */
/* Выделение памяти под `FfiStruct` */
FfiStruct* new_struct(uint32_t v);
/* Освобождение памяти, ранее выделенной под `FfiStruct` */
void free_struct(FfiStruct* s);
```

с соответствующими объявлениями FFI на Rust:

```
extern "C" {
    // Код С для выделения памяти под `FfiStruct`
    pub fn new_struct(v: u32) -> *mut FfiStruct;
    // Код С для освобождения памяти, выделенной под `FfiStruct`
    pub fn free_struct(s: *mut FfiStruct);
}
```

Чтобы обеспечить сохранение согласованности между операциями выделения и освобождения памяти, желательно реализовать обертку RAII, которая автоматически исключит утечку памяти, выделенной кодом С (см. рекомендацию 11). Структура этой обертки владеет памятью, выделенной библиотекой С:

```
/// Структура обертки, владеющая памятью, выделенной библиотекой С
struct FfiWrapper {
    // Инвариант: inner не NULL
    inner: *mut FfiStruct,
}
```

а реализация `Drop` возвращает эту память библиотеке C, чтобы исключить возможность ее утечки:

```
// Ручная реализация [`Drop`], обеспечивающая, чтобы память,
// выделенная библиотекой C, была ей возвращена
impl Drop for FfiWrapper {
    fn drop(&mut self) {
        // Безопасность: `inner` не NULL; кроме того, `free_struct()`
        // обрабатывает NULL-указатели
        unsafe { free_struct(self.inner) }
    }
}
```

Тот же принцип применим не только к памяти в куче: *реализуйте Drop, чтобы применять принцип RAII к ресурсам, произведенным FFI*, — открытым файлам, подключениям к базам данных и т. д. (подробнее — в рекомендации 11).

Инкапсуляция взаимодействий с библиотекой C в обертку `struct` позволяет также перехватить ряд других возможных проблем, например преобразовав невидимый в ином случае сбой в `Result`:

```
type Error = String;

impl FfiWrapper {
    pub fn new(val: u32) -> Result<Self, Error> {
        let p: *mut FfiStruct = unsafe { new_struct(val) };
        // Нет гарантии, что сырье указатели будут не NULL
        if p.is_null() {
            Err("Failed to get inner struct!".into())
        } else {
            Ok(Self { inner: p })
        }
    }
}
```

В этом случае оберточная структура сможет предоставить безопасные методы, позволяющие использовать функциональность библиотеки C:

```
impl FfiWrapper {
    pub fn set_byte(&mut self, b: u8) {
        // Безопасность: опирается на инвариант, что `inner` не NULL
        let r: &mut FfiStruct = unsafe { &mut *self.inner };
        r.byte = b;
    }
}
```

В качестве альтернативы если соответствующая структура C имеет эквивалент на Rust и можно безопасно напрямую манипулировать этой структурой данных,

то реализаций трейтов `AsRef` и `AsMut` (описаны в рекомендации 8) обеспечат более прямое использование:

```
impl AsMut<FfiStruct> for FfiWrapper {
    fn as_mut(&mut self) -> &mut FfiStruct {
        // Безопасность: `inner` не NULL
        unsafe { &mut *self.inner }
    }
}

let mut wrapper = FfiWrapper::new(42).expect("real code would check");
// Прямое изменение содержимого структуры данных,
// выделенной библиотекой C
wrapper.as_mut().byte = 12;
```

Этот пример иллюстрирует полезный принцип работы с FFI: *инкапсулируйте доступ к unsafe FFI-библиотеке внутри безопасного кода на Rust*. Тогда остальная часть приложения сможет следовать рекомендации 16, избегая написания `unsafe`-кода. Кроме того, так вы сберегете в одном месте весь опасный код, который затем сможете внимательно изучить и протестировать для выявления возможных проблем, а также рассматривать как наиболее вероятного подозреваемого в случае неполадок.

Вызов Rust из C

Внешний элемент определяется относительно того, где мы находимся: если пишем приложение на C, то это может быть библиотека Rust, используемая через интерфейс внешних функций. В этом случае предоставление коду C доступа к библиотеке Rust осуществляется по тем же основным принципам, что и доступ из Rust к C.

- Функции Rust, которые предполагается использовать из C, необходимо обозначить маркером `extern "C"`, чтобы обеспечить их совместимость с этим языком.
- Символы Rust по умолчанию подвергаются декорированию имен (как и в C++)¹, поэтому определения функций нужно сопроводить также атрибутом `#[no_mangle]`, чтобы можно было обращаться к ним по простому имени. Это, в свою очередь, означает, что имя функции относится к одному глобальному пространству имен и может конфликтовать с любым другим

¹ В Rust эквивалентом инструмента `c++filt` для перевода декорированных имен обратно в имена, видимые программисту, является `rustfilt`, основывающийся на команде `rustc-demangle`.

символом, определенным в программе. В связи с этим, чтобы избежать неоднозначности, желательно дополнять имя раскрываемой функциональности префиксом (`mylib_...`).

- Определения структур данных нужно сопровождать атрибутом `#[repr(C)]`, чтобы схема расположения их содержимого была совместима с эквивалентными структурами данных C.

Кроме того, при использовании кода Rust из C, как и в обратном случае, есть ряд тонких нюансов, связанных с обработкой указателей, ссылок и времени жизни. Указатель C отличается от ссылки Rust, о чём вы, к своему сожалению, забываете:

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
#[no_mangle]
pub extern "C" fn add_contents(p: *const FfiStruct) -> u32 {
    // Преобразование сырого указателя, переданного вызывающим кодом,
    // в ссылку Rust
    let s: &FfiStruct = unsafe { &*p }; // Упс
    s.integer + s.byte as u32
}
```

```
/* Код C, вызывающий Rust */
uint32_t result = add_contents(NULL); // Бах!
```

В ходе работы с сырыми указателями на вас лежит ответственность за обеспечение того, чтобы любое их использование соответствовало допущениям и гарантиям Rust в отношении ссылок:

```
#[no_mangle]
pub extern "C" fn add_contents_safer(p: *const FfiStruct) -> u32 {
    let s = match unsafe { p.as_ref() } {
        Some(r) => r,
        None => return 0, // Коварный код C передал NULL
    };
    s.integer + s.byte as u32
}
```

В этих примерах код C передает коду Rust сырой указатель, который тот преобразует в ссылку, чтобы выполнить операцию со структурой. Но откуда происходит этот указатель? Куда ведет полученная ссылка Rust?

В первом примере из рекомендации 8 было показано, что обеспечиваемая Rust безопасность памяти исключает возвращение ссылок на уже неактивные объекты стека. Такие проблемы возникают, если вы передаете сырой указатель:

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
impl FfiStruct {
    pub fn new(v: u32) -> Self {
        Self {
            byte: 0,
            integer: v,
        }
    }

    // Никаких ошибок при компиляции
#[no_mangle]
pub extern "C" fn new_struct(v: u32) -> *mut FfiStruct {
    let mut s = FfiStruct::new(v);
    &mut s // Возвращает сырой указатель на объект стека,
           // который вскоре утратит актуальность
}
```

Любые указатели, передаваемые обратно из Rust в C, обычно должны ссылаться на память кучи, а не стека. Но наивная попытка поместить объект в кучу посредством `Box` не поможет.

НЕЖЕЛАТЕЛЬНОЕ ПОВЕДЕНИЕ

```
// Здесь тоже нет ошибок при компиляции
#[no_mangle]
pub extern "C" fn new_struct_heap(v: u32) -> *mut FfiStruct {
    let s = FfiStruct::new(v); // Создание `FfiStruct` в стеке
    let mut b = Box::new(s); // Перемещение `FfiStruct` в кучу
    &mut *b // Возврат сырого указателя на объект кучи,
           // который вскоре утратит актуальность
}
```

Владеющий `Box` находится в стеке, поэтому, когда он выходит из области видимости, он освобождает объект кучи и возвращаемый сырой указатель снова окажется недействительным.

Подходящим инструментом здесь будет `Box::into_raw`, который откажется от ответственности за объект кучи, по сути забыв о нем:

```
#[no_mangle]
pub extern "C" fn new_struct_raw(v: u32) -> *mut FfiStruct {
    let s = FfiStruct::new(v); // Создает `FfiStruct` в стеке
    let b = Box::new(s); // Перемещает `FfiStruct` в кучу

    // Поглощает `Box` и берет на себя ответственность за память в куче
    Box::into_raw(b)
}
```

Здесь возникает вопрос: как теперь освободить объект кучи? Предыдущий совет заключался в выполнении выделения и освобождения памяти по одну сторону границы FFI, то есть в данном случае это освобождение нужно возложить на сторону Rust. Подходящим инструментом для этого будет `Box::from_raw`, который создает `Box` из сырого указателя:

```
#[no_mangle]
pub extern "C" fn free_struct_raw(p: *mut FfiStruct) {
    if p.is_null() {
        return; // Коварный код C передал NULL
    }
    let _b = unsafe {
        // Безопасность: известно, что p не NULL
        Box::from_raw(p)
    };
} // `_b` в конце области видимости отбрасывается, освобождая `FfiStruct`
```

Но это все равно оставляет участок кода Rust на милость кода C. Если код C по недосмотру попросит Rust освободить один и тот же указатель дважды, это наверняка приведет аллокатор Rust к ступору и сбою.

Подобная ситуация хорошо иллюстрирует основной посыл текущей рекомендации: использование FFI несет вам риски, которых в стандартном Rust нет. Это может быть вполне оправданным при условии, что вы в курсе сопутствующих опасностей и затрат. Контроль того, что передается через границу FFI, помогает снизить эти риски, но не устраняет их.

При этом контроль границы FFI для кода C, вызывающего Rust, сопряжен еще с одним нюансом: если ваш код Rust игнорирует рекомендацию 18, то вам следует *не допускать проникновения panic! через границу FFI*, так как это всегда будет приводить к неопределенному поведению, причем обязательно плохому (<https://oreil.ly/qmUe0>)¹

Запомните

- При взаимодействии с кодом на других языках в качестве наименьшего общего знаменателя используется C. Это означает, что все символы находятся в одном глобальном пространстве имен.
- Минимизируйте вероятность возникновения проблем на границе FFI с помощью следующих приемов:
 - инкапсулируйте `unsafe`-код FFI в безопасные обертки;
 - выделяйте и освобождайте память согласованно по одну сторону границы;

¹ Имейте в виду, что Rust версии 1.71 включает C-unwind ABI (<https://oreil.ly/VVqVY>), позволяющий раскрывать некоторую функциональность между языками.

- используйте для структур данных совместимый с C порядок составления;
- задействуйте целочисленные типы фиксированного размера;
- используйте вспомогательные возможности для FFI, предоставляемые стандартной библиотекой;
- не допускайте выхода `panic!` за границы Rust.

Рекомендация 35. Используйте bindgen вместо ручного сопоставления встречных компонентов FFI

В рекомендации 34 разбиралась механика вызова кода C из программы Rust. Там говорилось, что для использования через FFI объявлений структур и функций C в коде Rust должны присутствовать их эквиваленты. Эти объявления на C и Rust необходимо сохранять согласованными. В той же рекомендации говорилось, что набор инструментов здесь не поможет — расхождения будут молча игнорироваться, скрывая проблемы, которые возникнут позднее.

Сохранение точного соответствия этих двух встречных компонентов выглядит достойной задачей для автоматизации, и у Rust есть подходящий инструмент — `bindgen`. Основная функция `bindgen` заключается в парсинге заголовочного файла C и генерации соответствующих объявлений на Rust.

Разберем в качестве примера одно из объявлений на C из рекомендации 34:

```
/* Файл lib.h */
#include <stdint.h>

typedef struct {
    uint8_t byte;
    uint32_t integer;
} FfiStruct;

int add(int x, int y);
uint32_t add32(uint32_t x, uint32_t y);
```

Инструмент `bindgen` можно вызывать вручную (или с помощью скрипта сборки `build.rs`) для создания соответствующего файла на Rust:

```
% bindgen --no-layout-tests \
--allowlist-function="add.*" \
--allowlist-type=FfiStruct \
-o src/generated.rs \
lib.h
```

Сгенерированный код Rust идентичен написанным вручную объявлением из рекомендации 34:

```
/* Автоматически сгенерировано инструментом rust-bindgen 0.59.2 */

#[repr(C)]
#[derive(Debug, Copy, Clone)]
pub struct FfiStruct {
    pub byte: u8,
    pub integer: u32,
}
extern "C" {
    pub fn add(
        x: ::std::os::raw::c_int,
        y: ::std::os::raw::c_int,
    ) -> ::std::os::raw::c_int;
}
extern "C" {
    pub fn add32(x: u32, y: u32) -> u32;
}
```

И его можно подтянуть в код Rust с помощью макрона `include!`:

```
// Включает сгенерированные объявления на Rust
include!("generated.rs");
```

Используйте bindgen для генерации кода Rust, соответствующего коду C, во всех случаях, кроме самых тривиальных объявлений FFI. Это та самая область, где обширный машинно-генерируемый код определено предпочтительнее ручных объявлений. В случае изменения определения функции С компилятор С начнет ругаться, если ее объявление перестанет этому определению соответствовать. Если же вручную прописанное объявление на Rust перестанет соответствовать объявлению на С, то никто ругаться не станет. В связи с этим желательно использовать автоматическую генерацию объявлений на Rust из объявлений на С, чтобы гарантировать сохранение их согласованности.

Это также означает, что этап `bindgen` отлично подходит для включения в CI (см. рекомендацию 32). Если генерируемый код вносится в систему контроля версий, механизм CI может выдавать ошибку в случае несоответствия созданного файла той версии, которая была загружена в главную ветку.

Инструмент `bindgen` служит самостоятельным помощником в работе с существующей кодовой базой С, имеющей обширный API. Написание эквивалентов для заголовочного файла `lib_api.h` происходит вручную, оно очень утомительно, из-за чего возрастает риск ошибок — а многие категории ошибок из-за рас согласованности, как уже говорилось, набор инструментов просто не заметит. Кроме того, `bindgen` имеет множество опций (<https://oreil.ly/WSC8t>), позволяющих

выполнять генерацию для конкретных наборов API, например ранее показанные опции `--allowlist-function` и `--allowlist-type`¹.

Это делает возможным двухуровневый подход к использованию существующей библиотеки С из Rust. В соответствии с общепринятым соглашением для обертывания некой библиотеки `xyzzy` нужно следующее:

- крейт `xyzzy-sys`, где находится только сгенерированный `bindgen` код, использование которого неизбежно `unsafe`;
- крейт `xyzzy`, инкапсулирующий `unsafe`-код и предоставляющий для Rust безопасный доступ к соответствующей функциональности.

Таким образом, весь `unsafe`-код концентрируется на одном уровне, позволяя остальной программе следовать рекомендации 16.

За пределами С

Инструмент `bindgen` способен обрабатывать конструкции C++ (<https://oreil.ly/vn8Hf>), но лишь некоторые и ограниченно. Для лучшей, но слегка неполноценной интеграции рекомендую реализовывать взаимодействие C++ с Rust при помощи крейта `cxx`. Этот инструмент вместо генерации кода Rust из объявлений C++ и тот и другой код генерирует на основе общей схемы, обеспечивая более тесную интеграцию.

¹ Для упрощения вывода в примере использовалась также опция `--no-layout-tests`. По умолчанию сгенерированный код будет включать код `#[test]`, проверяющий, чтобы структуры были выстроены корректно.

Послесловие

Надеюсь, что представленные в книге советы, предложения и информация помогут вам стать гибким и продуктивным программистом. Как говорилось в предисловии, данный учебник пригодится на втором этапе этого процесса, когда вы уже освоили основы из фундаментального пособия по Rust. Теперь перед вами открывается немало дальнейших шагов и направлений для изучения.

- В этой книге не затрагивался *асинхронный код Rust*, хотя он наверняка потребуется для создания эффективных многопоточных серверных приложений. Вы можете познакомиться с `async` в онлайн-документации (<https://oreil.ly/a9r1B>), а также почитать книгу *Async Rust* Максвелла Флиттона (Maxwell Flitton) и Кэролайн Мортон (Caroline Morton) (O'Reilly, 2024).
- В качестве альтернативного направления для ваших интересов и требований может подойти концепция *Bare-metal Rust*. Она выходит за пределы знакомства с `no_std`, о котором говорилось в рекомендации 33, перенося вас в мир, где нет операционной системы и аллокации. В качестве хорошего ознакомительного материала по этой теме рекомендую соответствующий раздел онлайн-курса *Comprehensive Rust* (https://oreil.ly/h_cfR).
- Независимо от того, ориентированы вы на низкоуровневую или же высокоуровневую разработку, вам следует познакомиться с экосистемой сторонних крейтов с открытым исходным кодом `crates.io` и стать ее участником. Разобраться в огромном спектре предоставляемых ею возможностей вам помогут грамотно организованные тематические таблицы рекомендаций `blessed.rs` или `lib.rs`.
- Помощь в работе с самим языком вы можете найти на тематических форумах, таких как непосредственно форум Rust (<https://users.rust-lang.org>) и его канал Reddit `r/rust`. Кроме того, они позволяют выполнять поиск по уже заданным ранее вопросам (тем, на которые даны ответы).
- Если вы окажетесь в ситуации зависимости от библиотеки, которая на Rust не реализована (о чем говорилось в рекомендации 34), то можете *переписать ее на Rust* (RiiR). Только не стоит недооценивать усилия, необходимые (<https://oreil.ly/iKRzI>) для воссоздания зрелой и проверенной временем кодовой базы.
- Когда ваши навыки работы с Rust окрепнут, хорошим пособием для освоения продвинутых аспектов этого языка станет книга Йона Гъенсета (Jon Gjengset) *Rust for Rustaceans* (No Starch, 2022).

Удачи!

Об авторе

Дэвид Дрисдейл (David Drysdale) занимает должность ведущего специалиста по программному обеспечению в Google и с 2019 года работает преимущественно с Rust. Он портировал на Rust криптографическую библиотеку Tink, а также руководил проектом по замене аппаратной библиотеки криптографии для Android (KeyMint) ее аналогом на Rust. Дэвид имеет богатый опыт работы с C/C++ и Go, а также участвовал в разработке различных проектов, включая ядро Linux, ПО для плоскостей управления в сетевых архитектурах и мобильные приложения для видеосвязи.

Иллюстрация на обложке

Животное на обложке — краб-плавунец (*Arenaeus cibrarius*). Для него характерен ночной одиночный образ жизни, в случае угрозы он может демонстрировать агрессивное поведение. Этот вид встречается в нескольких местах вдоль Атлантического побережья, начиная от Массачусетса в США и заканчивая Аргентиной.

Верхняя часть панциря краба-плавунца может иметь светло-коричневый, светло-бордовый или оливковый окрас с белыми или темно-рыжими пятнами. Каждая сторона панциря состоит из девяти боковых зубцов, последний из которых выходит наружу. Между глазницами шесть частично сросшихся передних зубцов. Ширина тела этого морского животного может достигать 15 см и вдвое превышает его длину.

Окраска позволяет крабу маскироваться в окружающей среде во время охоты. Его основная пища — детрит (органические остатки), но он также питается рыбой, моллюсками и другими ракообразными.

Охранный статус этого вида крабов не оценивался, но на них ведется коммерческая охота в Бразилии, что может повлиять на общую численность популяции. Многие из животных с обложек книг издательства O'Reilly являются вымирающими. Все они крайне важны для нашего мира.

Дэвид Дрисдейл

Эффективный Rust.

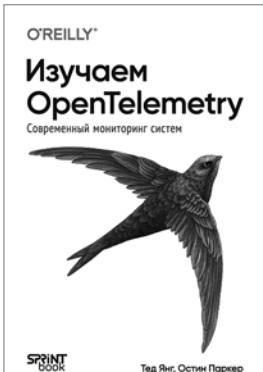
35 конкретных способов улучшить код

Перевел с английского Д. Брайт

Изготовлено в России. Изготовитель: ТОО «Спринт Бук». Место нахождения и фактический адрес: 010000, Казахстан, город Астана, район Алматы, проспект Ракымжан Кошкарбаев, д. 10/1, н. п. 18.

Дата изготовления: 03.2025. Наименование: книжная продукция. Срок годности: не ограничен.

Подписано в печать 26.02.25. Формат 70×100/16. Бумага офсетная. Усл. п. л. 24,510. Тираж 700. Заказ 0000.



Тед Янг, Остин Паркер

ИЗУЧАЕМ OPENTELEMETRY: СОВРЕМЕННЫЙ МОНИТОРИНГ СИСТЕМ

Появление OpenTelemetry произвело революцию в сфере наблюдаемости. Вместо того чтобы использовать несколько разрозненных систем, OpenTelemetry интегрирует трассировки, метрики и журналы в общий поток данных, предоставляя возможность оценить недоступные ранее взаимосвязи между ними. В этом практическом руководстве показано, как настраивать, использовать и диагностировать систему наблюдаемости OpenTelemetry.

Авторы Тед Янг и Остин Паркер, руководители, основатели и участники проекта OpenTelemetry, представляют все компоненты OpenTelemetry, а также лучшие практики наблюдаемости для многих популярных облачных сервисов, платформ и сервисов данных, таких как Kubernetes и AWS Lambda. Вы узнаете, как OpenTelemetry дает возможность сервисам и библиотекам OSS создавать собственное нативное инструментирование — впервые в отрасли.