

---

# **Zig Basics**

**Systemprogrammierung für das 21.  
Jahrhundert**

**David Pierre Sugar**

# **Zig Basics**

David Pierre Sugar

Copyright © 2024 David Pierre Sugar. All rights reserved.

Munich, Germany 2024: First Edition



Self Publishers Worldwide  
Seattle San Francisco New York  
London Paris Rome Beijing Barcelona

**für Franziska**



# Inhaltsverzeichnis

Zig Crash Course .....	1
Zig installieren .....	1
Funktionen .....	5
Unit Tests .....	7
Comptime .....	8
Kryptographie .....	11
Graphische Applikationen .....	17
Parallelität .....	20
C/C++ Build System .....	20
Bibliothek .....	21
Verwendung der Bibliothek .....	25
Standard Typen .....	27
Ganzzahlen (Integer) .....	27
Fließkommazahlen (Float) .....	29
Arrays und Slices .....	30
Arrays .....	32
Slices .....	33
for-Schleife (Loop) .....	34
Zeiger (Pointer) .....	36
Container .....	37
Struct .....	38
init Pattern .....	39
Anonyme Struct-Literale .....	42
Result-Location .....	42
Tuples .....	44
Enum .....	44
Union .....	45
Speicherverwaltung .....	47
Grundlagen .....	47
Lifetimes .....	50
Static Memory .....	50
Automatic Memory .....	51
Dynamic Memory .....	54
Häufige Fehler .....	57
Speicherzugriffsfehler (Access Errors) .....	57
Buffer Overflow/ Over-Read .....	57
Invalid Page Fault .....	58
Use After Free .....	58
Zusammenfassung .....	59



---

# Vorwort

Zig ist eine Sprache geeignet für die Systemprogrammierung.

Das alleine macht Zig nicht besonders, jedoch verheiratet Zig die Simplizität von C mit vielen modernen Features, was vor allem Neulingen, die eine systemnahe Programmiersprache lernen wollen, zugute kommt.

Zig als Systemprogrammiersprache ist unter anderem geeignet für:

- Kryptographie
- Mikrokontrollerprogrammierung
- Dateisysteme
- Datenbanken
- Betriebssysteme
- Treiber
- Spiele
- Simulationen
- Die Entwicklung von höheren Programmiersprachen

Insbesondere Startups, aber auch große Unternehmen, haben in den letzten Jahren auf Zig als Programmiersprache und Build-System gesetzt. Darunter Uber<sup>1</sup>, Tigerbeetle<sup>2</sup>, und ZML<sup>3</sup>. Dies verwenden Zig in ganz unterschiedlichen Anwendungsbereichen, darunter Datenbanken und maschinellem Lernen.

In der Welt der Systemprogrammiersprachen reiht sich Zig neben C ein und verzichtet auf viele Konzepte die andere Programmiersprachen überkomplex machen, darunter Vererbung. Damit ist Zig erfrischend übersichtlich, was vor allem Einsteigern zu gute kommt, bietet jedoch auch viele Verbesserungen gegenüber C. Ein wichtiger Fokus liegt auf der Lesbarkeit des Codes, d.h. was man sieht wird vom Computer später auch tatsächlich so ausgeführt (mit Abstrichen natürlich). Insbesondere bedeutet das: keine versteckten Allokationen, bei denen die Sprache ohne Zutun des Entwicklers dynamisch Speicher alloziert. Alles was mit der Allokation von dynamischem Speicher zu tun hat ist in Zig explizit!

---

<sup>1</sup><https://www.uber.com/en-DE/blog/bootstrapping-ubers-infrastructure-on-arm64-with-zig/>

<sup>2</sup><https://tigerbeetle.com/>

<sup>3</sup><https://zml.ai/>

Fun-Fact: Während der StackOverflow 2024 Developer Survey<sup>4</sup> gaben 6.2% der Befragten an „umfangreiche Entwicklungsarbeiten“ in Zig getätigt zu haben und 73.8% wollen die Sprache im kommenden Jahr (2025) nutzen. Damit ist Zig trotz seines Alpha-Status eine gern genutzte Programmiersprache und reiht sich von der Zahl der Anwender neben Sprachen wie Swift, Dart, Elixir und Ruby ein.

## Zielgruppe

Falls Sie bereits Erfahrung mit C oder einer anderen systemnahen Programmiersprache haben und mehr über Zig erfahren wollen ist diese Buch für Sie. Wenn Sie Erfahrung mit einer höheren Programmiersprache haben und mehr über Systemprogrammierung und Zig erfahren wollen ist dieses Buch ebenfalls für Sie.

Grundsätzlich empfehle ich Ihnen parallel zum lesen dieses Buches eigene Programmierprojekte zu realisieren um praktische Erfahrung mit der Sprache zu sammeln. Beginnen Sie mit etwas einfachem, vertrauten und steigern Sie sich, sobald Sie ein Gefühl für die Sprache bekommen haben. Sie werden merken, dass die Grundlagen in Zig schnell zu erlernen sind, es gibt jedoch auch nach einiger Zeit viel zu entdecken. Sollten Sie etwas Inspiration benötigen, so kann Ihnen Project Euler<sup>5</sup> eventuell weiterhelfen.

Wichtig zu erwähnen ist, dass Zig derzeit noch nicht die Version 1.0 erreicht hat, d.h. die Sprache und damit auch die Standardbibliothek werden sich in Zukunft noch ändern. Damit kann es sein, dass bestimmte Beispiele mit einer zukünftigen Zig-Compiler-Version nicht mehr compilieren. Sollte das für Sie ein Dealbreaker sein, so empfehle ich Ihnen die Finger von diesem Buch zu lassen und zu warten bis Zig Version 1.0 veröffentlicht wurde.

## Voraussetzungen

Die Zig-Version, die in diesem Buch verwendet wird ist 0.13.0<sup>6</sup>. Je nachdem wann Sie dieses Buch lesen kann es sein, dass diese Version nicht mehr aktuell ist. Bei Abweichungen von der angegebenen Version ist nicht garantiert, dass die in diesem Buch abgebildeten Beispiele compilieren.

Zwar sind die meisten Konzepte und Beispiele in diesem Buch unabhängig von einem bestimmten Betriebssystem und Architektur, jedoch geht das Buch grundsätzlich von einem x86\_64 Linux System aus. Dies wird relevant wenn auf Assembler, Calling-Conventions und ähnliche Konzepte Bezug genommen wird, da diese immer sowohl von der Architektur als auch dem Betriebssystem abhängen. Sollte Ihr Computer eine dieser Anforderungen nicht erfüllen, so empfiehlt es sich ggf. ein virtuelle Maschine zu verwenden<sup>7</sup>.

## Struktur

Die ersten drei Kapitel beschäftigen sich mit den Grundlagen der Programmiersprache Zig. Das erste Kapitel bietet anhand von Beispielen einen Überblick über die Sprache. Im zweiten Kapi-

---

<sup>4</sup><https://survey.stackoverflow.co/2024/>

<sup>5</sup><https://projecteuler.net/about>

<sup>6</sup><https://ziglang.org/download/>

<sup>7</sup><https://ubuntu.com/tutorials/how-to-run-ubuntu-desktop-on-a-virtual-machine-using-virtualbox#1-overview>



tel werden die grundlegenden Datentypen der Programmiersprache näher beleuchtet. In Kapitel drei wird der Leser in grundlegende Konzepte der Speicherverwaltung eingeführt, die für die korrekte und sichere Entwicklung von Anwendungen unabdingbar sind.

Im zweiten Abschnitt des Buches werden wir anhand von Fallbeispielen verschiedene Einsatzszenarios von Zig näher betrachten, darunter:

- Schreiben eines Parsers
- Breakout

Zig bietet für jede Compiler-Version zusätzliche Ressourcen zum Lernen der Sprache und als Referenz<sup>8</sup>, darunter die Language Reference und die Online-Dokumentation der Standardbibliothek. Diese können beim Entwickeln eigener Projekte aber auch beim nachvollziehen der Code-Beispiele eine große Hilfe darstellen.

## Konventionen

Die folgenden Konventionen werden in diesem Buch eingehalten:

*Italic*: Markiert neue Begriffe, URLs, Email-Adressen, Dateinamen und -endungen.

**Konstanter Abstand**: Wird verwendet für Programmbeispiele, sowie zum benennen von Programmbausteinen, wie etwa Variablennamen oder Umgebungsvariablen.

**Konstanter Abstand Fett**: Zeigt Kommandos oder ändern, vom Nutzer zu tippenden, Text.



Ziggy markiert einen Tipp bzw. einen Hinweis.

## Code Beispiele

Die in diesem Buch abgebildeten Code-Beispiele finden sich auf Github unter `todo` zum Download.

Alle Beispiele können von Ihnen ohne Einschränkung verwendet werden. Sie brauchen die Autoren nicht explizit um Genehmigung fragen. Am Schluss geht es darum Ihnen zu helfen und nicht darum Ihnen Steine in den Weg zu legen.

Zitierungen würden uns freuen, sind jedoch keinesfalls notwendig. Ein Zitat umfasst gewöhnlich Titel, Autor, Publizist und ISBN. In diesem Fall wäre dies: „Zig Basics by David Pierre Sugar”.

Sollten Sie Fehler im Buch oder Code finden, die nicht auf unterschiedliche Compiler-Versionen zurückzuführen sind können Sie uns mit einem Verbesserungsvorschlag kontaktieren.

## Fragen, Anmerkungen und Verbesserungen

Ich habe mein Bestes getan dieses Buch so informativ und technisch korrekt wie möglich zu gestalten. Ich bin mir jedoch auch sicher, dass dieses Buch besser sein könnte als es gerade ist.

---

<sup>8</sup><https://ziglang.org/learn/>

Sollten Sie Fehler finden oder generell Feedback zu diesem Buch geben wollen, so können Sie mich unter [david@thesugar.de](mailto:david@thesugar.de) kontaktieren. Dies gibt mir die Möglichkeit dieses Buch über die Zeit zu verbessern. Es ist mir jedoch nicht immer möglich zu antworten. Nehmen Sie es sich deswegen nicht zu Herzen wenn Sie nichts von mir hören.

## **Danksagung**

TDB

## Kapitel 1

---

# Zig Crash Course

In diesem Kapitel schauen wir uns einige kleine Zig Programme an, damit Sie ein Gespür für die Programmiersprache bekommen. Machen Sie sich nicht zu viele Sorgen wenn Sie nicht alles sofort verstehen, in den folgenden Kapiteln werden wir uns mit den hier vorkommenden Konzept noch näher beschäftigen. Wichtig ist, dass Sie diese Kapitel nicht nur lesen sondern die Beispiel auch ausführen, um das meiste aus diesem Kapitel herauszuholen.

## Zig installieren

Um Zig zu installieren besuchen Sie die Seite <https://ziglang.org> und folgen den Instruktionen unter „GET STARTED“<sup>9</sup>.

Die Installation ist unter allen Betriebssystemen relativ einfach durchzuführen. In der Download Sektion<sup>10</sup> finden Sie vorkompilierte Zig-Compiler für die gängigsten Betriebssysteme, darunter Linux, macOS und Windows.

Unter Linux können Sie mit dem Befehl **uname -a** Ihre Architektur bestimmen. In meinem Fall ist dies X86\_64.

```
$ uname -a
Linux ... x86_64 x86_64 x86_64 GNU/Linux
```

Die Beispiele in diesem Buch basieren auf der Zig-Version 0.13.0, d.h. um den entsprechenden Compiler auf meinem Linux system zu installieren würde ich die Datei *zig-linux-x86\_64-0.13.0.tar.xz* aus der Download-Sektion herunterladen.

---

<sup>9</sup><https://ziglang.org/learn/getting-started/>

<sup>10</sup><https://ziglang.org/download/>

**0.13.0**

- 2024-06-07
- [Release Notes](#)
- [Language Reference](#)
- [Standard Library Documentation](#)

OS	Arch	Filename	Signature	Size
Source		<a href="#">zig-0.13.0.tar.xz</a>	minisig	16.4MiB
		<a href="#">zig-bootstrap-0.13.0.tar.xz</a>	minisig	44.3MiB
Windows	x86_64	<a href="#">zig-windows-x86_64-0.13.0.zip</a>	minisig	75.5MiB
	x86	<a href="#">zig-windows-x86-0.13.0.zip</a>	minisig	79.4MiB
	aarch64	<a href="#">zig-windows-aarch64-0.13.0.zip</a>	minisig	71.6MiB
macOS	aarch64	<a href="#">zig-macos-aarch64-0.13.0.tar.xz</a>	minisig	42.8MiB
	x86_64	<a href="#">zig-macos-x86_64-0.13.0.tar.xz</a>	minisig	46.6MiB
Linux	x86_64	<a href="#">zig-linux-x86_64-0.13.0.tar.xz</a>	minisig	44.9MiB
	x86	<a href="#">zig-linux-x86-0.13.0.tar.xz</a>	minisig	49.7MiB
	aarch64	<a href="#">zig-linux-aarch64-0.13.0.tar.xz</a>	minisig	41.1MiB
	armv7a	<a href="#">zig-linux-armv7a-0.13.0.tar.xz</a>	minisig	42.0MiB
	riscv64	<a href="#">zig-linux-riscv64-0.13.0.tar.xz</a>	minisig	43.4MiB
	powerpc64le	<a href="#">zig-linux-powerpc64le-0.13.0.tar.xz</a>	minisig	44.4MiB
FreeBSD	x86_64	<a href="#">zig-freebsd-x86_64-0.13.0.tar.xz</a>	minisig	45.0MiB

Abbildung 1: Download Seite von <https://ziglang.org/download/>

Mit dem `tar` Kommandozeilenwerkzeug kann das heruntergeladene Archiv danach entpackt werden.

```
$ tar -xf zig-linux-x86_64-0.13.0.tar.xz
```

Der entpackte Ordner enthält die Folgenden Dateien.

```
$ ls zig-linux-x86_64-0.13.0
doc lib LICENSE README.md zig
```

- **doc:** Die Referenzdokumentation der Sprache. Diese ist auch online, unter <https://ziglang.org/documentation/0.13.0/>, zu finden und enthält einen Überblick über die gesamte Sprache. Ich empfehle Ihnen ergänzend zu diesem Buch die Dokumentation zu Rate zu ziehen.
- **lib:** Enthält alle benötigten Bibliotheken, inklusive der Standardbibliothek. Die Standardbibliothek enthält viel nützliche Programmbausteine, darunter geläufige Datenstrukturen, einen JSON-Parser, Kompressionsalgorithmen, kryptographische Algorithmen und Protokolle und vieles mehr. Eine Dokumentation der gesamten Standardbibliothek findet sich online unter <https://ziglang.org/documentation/0.13.0/std/>.
- **zig:** Dies ist ein Kommandozeilenwerkzeug mit dem unter anderem Zig-Programme kompiliert werden können.

Um den Zig-Compiler nach dem Entpacken auf einem Linux System zu installieren, können wir diesen nach `/usr/local/bin` verschieben.

```
$ sudo mv zig-linux-x86_64-0.13.0 /usr/local/bin/zig-linux-x86_64-0.13.0
```

Danach erweitern wir die `$PATH` Umgebungsvariable um den Pfad zu unserem Zig-Compiler. Dies können wir in der Datei `~/.profile` oder auch `~/.bashrc` machen<sup>11</sup>.

```
# Sample .bashrc for SuSE Linux
```

```
# ...
```

```
export PATH="$PATH:/usr/local/bin/zig-linux-x86_64-0.13.0"
```

Nach Änderung der Konfigurationsdatei muss diese neu geladen werden. Dies kann entweder durch das öffnen eines neuen Terminalfensters erfolgen oder wir führen im derzeitigen Terminal das Kommando **source .bashrc** in unserem Home-Verzeichnis aus. Danach können wir zum überprüfen, ob alles korrekt installiert wurde, das Zig-Zen auf der Kommandozeile ausgeben lassen. Das Zig-Zen kann als die Kernprinzipien der Sprache und ihrer Community angesehen werden, wobei man dazu sagen muss, dass es nicht „die eine“ Community gibt.

```
$ source ~/.bashrc
```

```
$ zig zen
```

- \* Communicate intent precisely.
- \* Edge cases matter.
- \* Favor reading code over writing code.
- \* Only one obvious way to do things.
- \* Runtime crashes are better than bugs.
- \* Compile errors are better than runtime crashes.
- \* Incremental improvements.
- \* Avoid local maximums.
- \* Reduce the amount one must remember.
- \* Focus on code rather than style.
- \* Resource allocation may fail;  
    resource deallocation must succeed.
- \* Memory is a resource.
- \* Together we serve the users.

Mit dem Kommando **zig help** lässt sich ein Hilfetext auf der Kommandozeile anzeigen, der die zu Verfügung stehenden Kommandos auflistet.

Praktisch ist, dass Zig für uns ein neues Projekt, inklusive Standardkonfiguration, anlegen kann.

```
$ mkdir hello && cd hello
```

```
$ zig init
```

```
info: created build.zig
```

```
info: created build.zig.zon
```

```
info: created src/main.zig
```

```
info: created src/root.zig
```

```
info: see `zig build --help` for a menu of options
```

Das Kommando initialisiert den gegebenen Ordner mit Template-Dateien, durch die sich sowohl eine Executable, als auch eine Bibliothek bauen lassen. Schaut man sich die erzeugten Dateien an so sieht man, dass Zig eine Datei namens *build.zig* erzeugt hat. Bei dieser handelt es sich um

---

<sup>11</sup>Je nach verwendetem Terminal kann die Konfigurationsdatei auch anders heißen.

die Konfigurationsdatei des Projekts. Sie beschreibt aus welchen Dateien eine Executable bzw. Bibliothek gebaut werden soll und welche Abhängigkeiten (zu anderen Bibliotheken) diese besitzen. Ein bemerkenswertes Detail ist dabei, dass *build.zig* selbst ein Zig Programm ist, welches in diesem Fall zur Compile-Zeit ausgeführt wird um die eigentliche Anwendung zu bauen.

Die Datei *build.zig.zon* enthält weitere Informationen über das Projekt, darunter dessen Namen, die Versionsnummer, sowie mögliche Dependencies. Dependencies können dabei lokal vorliegen und über einen relativen Pfad angegeben oder von einer Online-Quelle, wie etwa Github, bezogen werden. Die Endung der Datei steht im übrigen für Zig Object Notation (ZON), eine Art Konfigurationssprache für Zig, die derzeit, genauso wie Zig selbst, noch nicht final ist.

Schauen wir in *src/main.zig*, so sehen wir das Zig für uns ein kleines Programm geschrieben hat.

```
const std = @import("std");

pub fn main() !void {
    std.debug.print("All your {s} are belong to us.\n", .{"codebase"});

    const stdout_file = std.io.getStdOut().writer();
    var bw = std.io.bufferedWriter(stdout_file);
    const stdout = bw.writer();

    try stdout.print("Run `zig build test` to run the tests.\n", .{});

    try bw.flush(); // don't forget to flush!
}
```

Der Code kann auf den ersten Blick überwältigend wirken, schauen wir ihn uns deswegen Stück für Stück an.

```
const std = @import("std");
```

Mit der `@import()` Funktion importieren wir die Standardbibliothek (`std`) und binden diese an eine Konstante mit dem selben Namen. Die Standardbibliothek ist eine Ansammlung von nützlichen Funktionen und Datentypen, die während der Entwicklung von Anwendungen häufiger zum Einsatz kommen und deswegen vom Zig zur Verfügung gestellt werden. Die Funktion `@import()` wird nicht nur zum importieren der Standardbibliothek verwendet, sondern auch um auf Module und andere, zu einem Projekt gehörende, Quelldateien zuzugreifen.

Nach der Definition der Konstante `std` beginnt die `main` Funktion:

```
pub fn main() !void {
```

Unsere `main` Funktion beginnt, wie alle Funktionen, mit `fn` und dem Namen der Funktion. Sie gibt keinen Wert zurück, aus diesem Grund folgt auf die leere Parameterliste `()` der Rückgabetypp `void`. Das Ausrufezeichen `!` weist darauf hin, das die Funktion einen Fehler zurückgeben kann. Fehler in Zig sind eigenständige Werte, die von einer Funktion zurückgegeben werden können und sich semantisch vom eigentlichen Rückgabewert unterscheiden.

```
std.debug.print("All your {s} are belong to us.\n", .{"codebase"});
```

Als erstes gibt die `main` Funktion einen String über die Debugausgabe auf der Kommandozeile aus. Die Funktion `print` erwartet dabei einen Format-String, der mit Platzhaltern (z.B. `{s}`) versehen werden kann, sowie eine Liste an Ausdrücken (z.B. `.{ "codebase" }`) deren Werte in den String eingefügt werden sollen. Der Platzhalter `{s}` gibt z.B. an, dass an der gegebenen Stelle ein String eingefügt werden soll. Neben `s` gibt es unter anderem noch `d` für Ganzzahlen und `any` für beliebige Werte.

```
const stdout_file = std.io.getStdOut().writer();
var bw = std.io.bufferedWriter(stdout_file);
const stdout = bw.writer();
```

Via `std.io` können wir mit `getStdIn()`, `getStdOut()` und `getStdErr()` auf `stdin`, `stdout` und `stderr` zugreifen. Alle drei Funktionen geben jeweils ein Objekt vom Typ `File` zurück. Die Funktion `writer()` welche auf der `stdout`-Datei aufgerufen wird, gibt einen `Writer` zurück. Ein `Writer` ist ein Wrapper um ein beliebiges Datenobjekt (z.B. eine offene Datei, ein Array, ...) und stellt eine standardisierte Interface zur Verfügung um Daten zu serialisieren. In unserem Fall wird der `stdout_file` `Writer` wiederum in einen `BufferedWriter` gewrapped, welcher nicht bei jedem einzelnen Schreibvorgang auf die Datei `stdout` zugreift, sondern erst wenn genug Daten geschrieben wurden bzw. wenn die Funktion `flush()` aufgerufen wird. Die Konstante `stdout` ist also ein `Writer` der einen `Writer` umschließt, der eine Datei umschließt, in die schlussendlich geschrieben werden soll.

```
try stdout.print("Run `zig build test` to run the tests.\n", .{});
```

Der `BufferedWriter` (`stdout`) wird verwendet um (indirekt) den String „Run zig build test to run the tests.“ nach `stdout` (standardmäßig die Kommandozeile) zu schreiben. Da diese Schreiboperation fehlschlagen kann wird vor den Ausdruck ein `try` gestellt. Damit wird ein potenzieller Fehler „nach oben“ propagiert, was im gegebenen Fall zu einem Programmabsturz führen würde, da `main` keine Funktion über sich besitzt. Als Alternative könnte mit einem `catch` Block der Fehler explizit abgefangen werden.

```
try bw.flush();
```

Um sicher zu gehen, dass auch alle Daten aus dem `BufferedWriter` tatsächlich geschrieben wurden, muss schlussendlich `flush()` aufgerufen werden.

Das von Zig vorbereitete „Hello, World“-Programm kann mit **zig build run**, von einem beliebigen Ordner innerhalb des Zig-Projekts, ausgeführt werden.

```
$ zig build run
All your codebase are belong to us.
Run `zig build test` to run the tests.
```

Im gegebenen Beispiel wurden zwei Schritte ausgeführt. Zuerst wurde der Zig-Compiler aufgerufen um das Programm in `src/main.zig` zu kompilieren und im zweiten Schritt wurde das Programm ausgeführt. Zig platziert dabei seine Kompilierten Anwendungen in `zig-out/bin` und Bibliotheken in `zig-out/lib`.

## Funktionen

Zig's Grammatik ist sehr überschaubar und damit leicht zu erlernen. Diejenigen mit Erfahrung in anderen C ähnlichen Programmiersprachen wie C, C++, Java oder Rust sollten sich direkt Zuhause fühlen. Die unterhalb abgebildete Funktion berechnet den größten gemeinsamen Teiler (greatest common divisor) zweier Zahlen.

```
// chapter01/gcd.zig
fn gcd(n: u64, m: u64) u64 {
    return if (n == 0)
        m
    else if (m == 0)
        n
    else if (n < m)
        gcd(m, n)
    else
        gcd(m, n % m);
}
```

Das `fn` Schlüsselwort markiert den Beginn einer Funktion. Im gegebenen Beispiel definieren wir eine Funktion mit dem Name `gcd`, welche zwei Argumente `m` und `n`, jeweils vom Typ `u64`, erwartet. Nach der Liste an Argumenten in runden Klammern folgt der Typ des erwarteten Rückgabewertes. Da die Funktion den größten gemeinsamen Teiler zweier `u64` Ganzzahlen berechnet ist auch der Rückgabewert vom Typ `u64`. Der Körper der Funktion wird in geschweifte Klammern gefasst.

Zig unterscheidet zwischen zwei Variablen-Typen, Variablen und Konstanten. Konstanten können nach ihrer Initialisierung nicht mehr verändert werden, während Variablen neu zugewiesen werden können. Funktionsargumente zählen grundsätzlich zu Konstanten, d.h. sie können nicht verändert werden. Der Zig-Compiler erzwingt die Nutzung von Konstanten, sollte eine Variable nach ihrer Initialisierung nicht mehr verändert werden. Dies ist eine durchaus kontroverse Designentscheidung, welche aber auf das Zig-Zen zurückgeführt werden kann, das besagt: „Favor reading code over writing code“. Sollten Sie also eine Variable in fremden Code sehen so können Sie sicher sein, dass diese an einer anderen Stelle manipuliert bzw. neu zugewiesen wird.

Eine Besonderheit, die Zig von anderen Sprachen unterscheidet ist, dass Integer mit beliebiger Präzision unterstützt werden. Im obigen Beispiel handelt es sich bei `u64` um eine vorzeichenlose Ganzzahl (unsigned integer) mit 64 Bits, d.h. es können alle Zahlen zwischen 0 und  $2^{64} - 1$  dargestellt werden. Zig unterstützt jedoch nicht nur `u8`, `u16`, `u32` oder `u128` sondern alle unsigned Typen zwischen `u0` und `u65535`.



Alle Zig-Basistypen sind Teil des selben union: `std.builtin.Type`. Das union beinhaltet den `Int` Typ welcher ein struct mit zwei Feldern ist, `signedness` und `bits`, wobei `bits` vom Typ `u16` ist, d.h. es können alle Integer-Typen zwischen 0 und  $2^{16} - 1$  Bits verwendet werden. Ja Sie hören richtig, der Zig-Compiler ist seit Version 0.10.0 selbst in Zig geschrieben, d.h. er ist self-hosted.

Innerhalb des Funktionskörpers werden mittels `if` verschiedene Bedingungen abgefragt. Sollte eine beider Zahlen 0 sein, so wird jeweils die andere zurückgegeben, ansonsten wird `gcd` rekursiv



aufgerufen bis für eine der beiden Zahlen die Abbruchbedingung (0) erreicht ist. Wie auch bei C muss die Bedingung in runde Klammern gefasst werden. Bei Einzeilern können die geschweiften Klammern um einen Bedingungsblock weggelassen werden. In diesem Fall wird der Rückgabewert der Bedingung an den umschließenden Block gereicht.

Mittels eines `return` Statements kann von einer Funktion in die aufrufende Funktion zurückgekehrt werden. Das Statement nimmt bei Bedarf zusätzlich einen Wert der an die aufrufende Funktion zurückgegeben werden soll. Im obigen Beispiel gibt `gcd` mittels `return` den Wert des ausgeführten If-Else-Asudruck zurück.

Das vollständige Programm finden Sie im zugehörigen Github-Repository. Mittels **`zig build-exe chapter01/gcd.zig`** kann das Beispiel kompiliert werden.

## Unit Tests

Wie von einer modernen Programmiersprache zu erwarten bietet Zig von Haus aus Unterstützung für Tests. Tests beginnen mit dem Schlüsselwort `test`, gefolgt von einem String der den Test bezeichnet. In geschweiften Klammern folgt der Test-Block.

```
// chapter01/gcd.zig
test "assert that the gcd of 21 and 4 is 1" {
    try std.testing.expectEqual(@as(u64, 1), gcd(21, 4));
}
```

Die Standardbibliothek bietet unter `std.testing` eine ganze Reihe an Testfunktionen für verschiedene Datentypen und Situationen. Im obigen Beispiel verwenden wir `ExpectEqual`, welche als erstes Argument den erwarteten Wert erhält und als zweites Argument das Resultat eines Aufrufs von `gcd`. Die Funktion überprüft beide Werte auf ihre Gleichheit und gibt im Fehlerfall einen `error` zurück. Dieser Fehler kann mittels `try` propagiert werden, wodurch der Testrunner im obigen Beispiel erkennt, dass der Test fehlgeschlagen ist.

```
$ zig test chapter01/gcd.zig
All 1 tests passed.
```

Innerhalb einer Datei sind Definitionen auf oberster Ebene (top-level definitions) unabhängig von ihrer Reihenfolge, was die Definition von Tests mit einschließt. Damit können Tests an einer beliebigen Stelle definiert werden, darunter direkt neben der zu testenden Funktion oder am Ende einer Datei. Der Zig-Test-Runner sammelt automatisch alle definierten Tests und führt dies beim Aufruf von **`zig test`** aus. Worauf Sie jedoch achten müssen ist, dass Sie ausgehend von der Wurzel-Datei (in den meisten Fällen `src/root.zig`), die konzeptionell den Eintrittspunkt für den Compiler in ihr Programm oder Ihre Bibliothek darstellt, Zig mitteilen müssen in welchen Dateien zusätzlich nach Tests gesucht werden soll. Dies bewerkstelligen Sie, indem Sie die entsprechende Datei innerhalb eines Tests importieren.

```
const foo = @import("foo.zig");

test "main tests" {
    _ = foo; // Tell test runner to also look in foo for tests
}
```

# Comptime

Die meisten Sprachen erlauben eine Form von Metaprogrammierung, d.h. das Schreiben von Code der wiederum Code generiert. In C können die gefürchteten Makros mit dem Präprozessor verwendet werden und Rust bietet sogar zwei verschiedene Typen von Makros, jeweils mit einer eigenen Syntax. Zig bietet mit `comptime` seine eigene Form der Metaprogrammierung. Was Zig von anderen kompilierten Sprachen unterscheidet ist, dass die Metaprogrammierung in der Sprache selber erfolgt, d.h., wer Zig programmieren kann, der hat alles nötige Handwerkszeug um auch Metaprogrammierung in Zig zu betreiben.

Ein Aufgabe für die Metaprogrammierung sehr gut geeignet ist, ist die Implementierung von Container-Typen wie etwa `std.ArrayList`. Eine `ArrayList` ist ein Liste von Elementen eines beliebigen Typen, die eine Menge an Standardfunktionen bereitstellt um die Liste zu manipulieren. Nun wäre es sehr aufwändig die `ArrayList` für jeden Typen einzeln implementieren zu müssen. Aus diesem Grund ist `ArrayList` als Funktion implementiert, welche zur Compilezeit einen beliebigen Typen übergeben bekommt auf Basis dessen einen `ArrayList`-Typ generiert.

```
var list = std.ArrayList(u8).init(allocator);
try list.append(0x00);
```

Der Funktionsaufruf `ArrayList(u8)` wird zur Compilezeit ausgewertet und gibt einen neuen Listen-Typen zurück, mit dem sich eine Liste an `u8` Objekten managen lassen. Auf diesem Typ wird `init()` aufgerufen um eine neu Instanz des Listen-Typs zu erzeugen. Mit der Funktion `append()` kann z.B., ein Element an das Ende der Liste angehängt werden. Eine stark simplifizierte Version von `ArrayList` könnte wie folgt aussehen.

```
// chapter01/my-arraylit.zig
const std = @import("std");

// Die Funktion erwartet als Compilezeitargument einen Typen `T`
// und gibt ein Struct zurück, dass einen Wrapper um einen Slice
// des Type `T` darstellt.
//
// Der Wrapper implementiert Funktionen zum managen des Slices
// und unterstützt unter anderem:
// - das Hinzufügen neuer Elemente
pub fn MyArrayList(comptime T: type) type {
    return struct {
        items: []T,
        allocator: std.mem.Allocator,

        // Erzeuge eine neue Instanz von MyArrayList(T).
        // Der übergebene Allocator wird von dieser Instanz gemanaged.
        pub fn init(allocator: std.mem.Allocator) @This() {
            return .{
                .items = &[_]T{},
                .allocator = allocator,
            };
        }
    }
}
```

```

pub fn deinit(self: *@This()) void {
    self.allocator.free(self.items);
}

// Füge da Element `e` vom Typ `T` ans ende der Liste.
pub fn append(self: *@This(), e: T) !void {
    // `realloc()` kopiert die Daten bei Bedarf in den neuen
    // Speicherbereich aber die Allokation kann auch
    // fehlschlagen. An dieser Stelle verbleiben wir der
    // Einfachheit halber bei einem `try`.
    self.items = try self.allocator.realloc(self.items, self.items.len +
1);
    self.items[self.items.len - 1] = e;
}
};
}

pub fn main() !void {
    var gpa = std.heap.GeneralPurposeAllocator(.{}){};
    const allocator = gpa.allocator();

    var list = MyArrayList(u8).init(allocator);
    defer list.deinit();

    try list.append(0xAF);
    try list.append(0xFE);

    std.log.info("{s}", .{std.fmt.fmtSliceHexLower(list.items[0..])});
}

```

Mit dem `comptime` Keyword sagen wir dem Compiler, dass das Argument `T` zur Compilezeit erwartet wird. Beim Aufruf von `MyArrayList(u8)` wertet der Compiler die Funktion aus und generiert dabei einen neuen Typen. Das praktische ist, dass wir `MyArrayList` nur einmal implementieren müssen und diese im Anschluss mit einem beliebigen Typen verwenden können.

Der `comptime` Typ `T` kann innerhalb und auch außerhalb des von der Funktion `MyArrayList` zurückgegebenen Structs, anstelle eines expliziten Typs, verwendet werden.

Structs die mit `init()` initialisiert und mit `deinit()` deinitialisiert werden sind ein wiederkehrendes Muster in Zig. Dabei erwartet `init()` meist einen `std.mem.Allocator` der von der erzeugten Instanz verwaltet wird.

Ein weiterer Anwendungsfall bei dem Comptime zum Einsatz kommen kann ist die Implementierung von Parsern. Ein Beispiel hierfür ist der Json-Parser der Standardbibliothek (`std.json`), welcher dazu verwendet werden kann um Zig-Typen als Json zu serialisieren und umgekehrt<sup>12</sup>.

---

<sup>12</sup>Die JavaScript Object Notation (JSON) ist eines der gängigsten Datenformate und wird unter anderem zur Übermittlung von Daten im Web verwendet (<https://en.wikipedia.org/wiki/JSON>).

```
// chapter01/reflection.zig
const std = @import("std");

const MyStruct = struct {
    a: u32 = 12345,
    b: []const u8 = "Hello, World",
    c: bool = false,
};

fn isStruct(obj: anytype) bool {
    const T = @TypeOf(obj);
    const TInf = @typeInfo(T);

    return switch (TInf) {
        .Struct => |S| blk: {
            inline for (S.fields) |field| {
                std.log.info("{s}: {any}", .{ field.name, @field(obj, field.name) });
            }

            break :blk true;
        },
        else => return false,
    };
}

pub fn main() void {
    const s = MyStruct{};

    std.debug.print("{s}", .{if (isStruct(s)) "is a struct!" else "is not a struct!"});
}
```

Anstelle eines Typen kann `anytype` für Parameter verwendet werden. In diesem Fall wird der Typ des Parameters, beim Aufruf der Funktion, abgeleitet. Zig erlaubt Reflexion (type reflection). Unter anderem erlaubt Zig die Abfrage von (Typ-)Informationen über ein Objekt. Funktionen denen ein `@` vorangestellt sind heißen Builtin-Function (eingebaute Funktion) und werden direkt vom Compiler bereitgestellt, d.h., sie können überall in Programmen, ohne Einbindung der Standardbibliothek, verwendet werden.

Die Funktion `@TypeOf()` ist insofern speziell, als dass sie eine beliebige Anzahl an Ausdrücken als Argument annimmt und als Rückgabewert den Typ des Resultats zurückliefert. Die Ausdrücke werden dementsprechend evaluiert. Im obigen Beispiel wird `@TypeOf()` genutzt um den Typen des übergebenen Objekts zu bestimmen, da `isStruct()` aufgrund von `anytype` mit einem Objekt beliebigen Typs aufgerufen werden kann.

Die eigentliche Reflexion kann mithilfe der Funktion `@typeInfo()` durchgeführt werden, die zusätzliche Informationen über einen Typ zurückliefert. Felder sowie Deklarationen von structs, unions, enums und error Sets kommen dabei in der selben Reihenfolge vor, wie sie auch im Source Code zu sehen sind. Im obigen Beispiel testen wir mittels eines switch Statements ob es

sich um ein `struct` handelt oder nicht und geben dementsprechend entweder `true` oder `false` zurück. Sollte es sich um ein `struct` handeln, so iterieren wir zusätzlich über dessen Felder und geben den Namen des Felds, sowie dessen Wert aus. Den Wert des jeweiligen Felds erhalten wir, indem wir mittels `@field()` darauf zugreifen. Die Funktion `@field()` erwartet als erstes Argument ein Objekt (ein `Struct`) und als zweites Argument einen zu Compile-Zeit bekannten String, der den Namen des Felds darstellt, auf das zugegriffen werden soll. Damit ist `@field(s, "b")` das Äquivalent zu `s.b`.

Für jeden Typen, mit dem `isStruct()` aufgerufen wird, wird eine eigene Kopie der Funktion (zur Compile Zeit) erstellt, die an den jeweiligen Typen angepasst ist. Das Iterieren über die einzelnen Felder eines `structs` muss zur Compile Zeit erfolgen, aus diesem Grund nutzt die obige Funktion `inline` um die For-Schleife zu entrollen, d.h., aus der Schleife eine lineare Abfolge von Instruktionen zu machen.

```
$ zig build-exe chapter01/reflection.zig
$ ./reflection
info: a: 12345
info: b: { 72, 101, 108, 108, 111, 44, 32, 87, 111, 114, 108, 100 }
info: c: false
```

Reflexion kann in vielen Situationen äußerst nützlich sein, darunter der Implementierung von Parsern für Formate wie JSON oder CBOR<sup>13</sup>, da im Endeffekt nur zwei Funktionen implementiert werden müssen, eine zum Serialisieren der Daten und eine zum Deserialisieren. Mithilfe von Reflexion kann dann, vom Compiler, für jeden zu serialisierenden Datentyp eine Kopie der Funktionen erzeugt werden, die auf den jeweiligen Typen zugeschnitten ist.

## Kryptographie

Ein Großteil der Anwendungen, die Sie wahrscheinlich täglich verwenden, benutzt in irgend einer Form Kryptographie. Dabei handelt es sich grob gesagt um mathematische Algorithmen, mit denen vorwiegend die Vertraulichkeit (Confidentiality), Integrität (Integrity) und Authentizität (Authenticity) von Daten gewährleistet werden kann. Typische Anwendungsbereiche die Kryptographie verwenden sind Messenger, Video Chats, Networking (TLS), Passwortmanager und Smart Cards. Zig bietet in seiner Standardbibliothek bereits jetzt eine Vielzahl an kryptographischen Algorithmen und Protokollen, wobei ein Großteil davon von Frank Denis<sup>14</sup>, Online auch bekannt als jedisct1, beigetragen wurde. Ohne groß ein Autoritätsargument aufmachen zu wollen, ist Frank der Maintainer von `libsodium`<sup>15</sup> und `libhydrogen`<sup>16</sup>, zwei viel genutzte, kryptographische Bibliotheken.

Wir werden uns in einem späteren Kapitel noch genauer mit Kryptographie auseinandersetzen, machen Sie sich deshalb keine Sorgen, wenn Sie nicht alles in diesem Abschnitt auf Anhieb verstehen. Fürs erste schauen wir uns einen gängigen Anwendungsfall von Kryptographie an, die Verschlüsselung einer Datei. Angenommen wir haben eine Datei deren Inhalt geheim bleiben

<sup>13</sup><https://github.com/r4gus/zbor>

<sup>14</sup><https://github.com/jedisct1>

<sup>15</sup><https://github.com/jedisct1/libsodium>

<sup>16</sup><https://github.com/jedisct1/libhydrogen>

soll und wir wollen des weiteren überprüfen können, dass der Inhalt der Datei nicht verändert wurde. In solch einem Fall bietet sich die Verwendung eines AEAD (Authenticated Encryption with Associated Data) Ciphers an. Zig bietet unter `std.crypto.aead` verschiedene AEAD Cipher an. Die Unterschiede zwischen den Ciphern ist für dieses Beispiel Out-of-Scope. Sie müssen sich fürs erste damit begnügen mir zu glauben, dass `XChaCha20Poly1305`<sup>17</sup> für diese Art von Problem eine gute Wahl ist. Der Name `XChaCha20Poly1305` enthält dabei zwei Informationen, die uns Aufschluss über die Zusammensetzung des Ciphers geben:

- `XChaCha20`: Zur Verschlüsselung der Daten wird die „Nonce-eXtended“ Version der `ChaCha20` Stromchiffre verwendet. `XChaCha20` erwartet einen Schlüssel und eine Nonce (Number used once: Eine Byte-Sequenz die nur einmal für eine Verschlüsselung verwendet werden darf) und leitet daraus eine Schlüsselsequenz ab, die mit dem Klartext XORed wird. Die eXtended Version verwendet dabei eine 192-Bit Nonce anstelle einer 96-Bit Nonce, was es deutlich sicherer macht diese zufällig mittels eines (kryptographisch sicheren) Zufallszahlengenerators zu erzeugen. Dieser Teil des Algorithmus ist für die Vertraulichkeit der Daten verantwortlich.
- `Poly1305`: `Poly1305` ist ein Hash, der zur Erzeugung von (one-time) Message Authentication Codes (MAC) verwendet werden kann. MACs sind sogenannte Keyed-Hashfunktionen, bei denen in einen Hash (keine Sorge, wir werden uns noch näher damit beschäftigen) ein geheimer Schlüssel integriert wird. Die Hashsumme wird dabei in unserem Beispiel über den Ciphertext, d.h. den Verschlüsselten Text, gebildet<sup>18</sup>. Durch den Einbezug eines Schlüssels kann nicht nur überprüft werden, dass die Integrität der Datei nicht verletzt wurde (sie wurde nicht verändert), sondern es kann auch sichergestellt werden, dass die MAC von Ihnen generiert wurde, da nur Sie als Nutzer der Anwendung den geheimen Schlüssel kennen.

```
// chapter01/encrypt.zig
const std = @import("std");

const argon2 = std.crypto.pwhash.argon2;
const XChaCha20Poly1305 = std.crypto.aead.chacha_poly.XChaCha20Poly1305;

var gpa = std.heap.GeneralPurposeAllocator(.{}){};
const allocator = gpa.allocator();

const Mode = enum {
    encrypt,
    decrypt,
};

pub fn main() !void {
    var password: ?[]const u8 = null;
    var mode: ?Mode = null;

    // Als erstes parsen wir die übergebenen Kommandozeilenargumente.
    // Diese bestimmen zum einen mit welchem Passwort die Daten
```

---

<sup>17</sup><https://datatracker.ietf.org/doc/html/rfc7539>

<sup>18</sup>Dies wird als Encrypt-then-Mac bezeichnet.

```

// verschlüsselt werden sollen und zum anderen den Modus, d.h.
// ob ver- bzw. entschlüsselt werden soll.
var ai = try std.process.argsWithAllocator(allocator);
defer ai.deinit();

while (ai.next()) |arg| {
    // `std.mem.eql` kann dazu verwendet werden zwei Strings mit einander zu
    // vergleichen...
    if (arg.len > 11 and std.mem.eql(u8, "--password=", arg[0..11])) {
        password = arg[11..];
    } else if (arg.len >= 9 and std.mem.eql(u8, "--encrypt", arg[0..9])) {
        mode = .encrypt;
    } else if (arg.len >= 9 and std.mem.eql(u8, "--decrypt", arg[0..9])) {
        mode = .decrypt;
    }
}

// Sollten nicht alle benötigten Argumente übergeben worden sein, so beenden
// wir den Prozess.
if (password == null or mode == null) {
    std.log.err("usage: ./encrypt --password=<password> [--encrypt|--
decrypt]", .{});
    return;
}

// Als nächstes lesen wir die übergebenen Daten von `stdin` ein.
const stdin = std.io.getStdIn();
const data = try stdin.readToEndAlloc(allocator, 64_000);
defer {
    // Wir überschreiben die Daten bevor wir den Speicher wieder freigeben.
    @memset(data, 0);
    allocator.free(data);
}

if (mode == .encrypt) {
    // Bei der Verschlüsselung müssen wir eine Reihe an (öffentlichen)
    // Parametern festlegen, die bei der Entschlüsselung wiederverwendet
    // werden müssen.

    // Als erstes müssen wir ein Schlüssel von unserem Passwort ableiten.
    // Hierfür verwenden wir die Argon2id Key-Derivation-Function (KDF).
    var salt: [32]u8 = undefined;
    std.crypto.random.bytes(&salt);

    var key: [XChaCha20Poly1305.key_length]u8 = undefined;
    try argon2.kdf(allocator, &key, password.?, &salt, .{
        // Die Parameter bestimmen wie aufwendig die Brechnung des Schlüssels
        // `key` ist.
        // Damit wird verhindert, diesen durch "Brute-Forcing" brechen zu

```

können.

```

        .t = 3,
        .m = 4096,
        .p = 1,
    }, .argon2id);

// Nun können wir die Daten ver-/ bzw. entschlüsseln.

// Der TAG wird von der encrypt() Funktion erzeugt und später
// von decrypt() überprüft.
var tag: [XChaCha20Poly1305.tag_length]u8 = undefined;

// Für jede Verschlüsselung muss eine neue, einzigartige Nonce
// verwendet werden. Da wir die eXtended Version von ChaCha20
// verwenden, kann diese durch einen kryptographisch sicheren
// Zufallszahlengenerator festgelegt werden.
var nonce: [XChaCha20Poly1305.nonce_length]u8 = undefined;
std.crypto.random.bytes(&nonce);

XChaCha20Poly1305.encrypt(data, &tag, data, "", nonce, key);

// Der Salt, Nonce und Tag müssen mit den verschlüsselten Daten serialisiert
werden,
// da wir diese später zur Entschlüsselung benötigen.
const stdout = std.io.getStdOut();
try std.fmt.format(stdout.writer(), "{s}:{s}:{s}:{s}", .{
    // Wir serialisieren die Binärdaten in Hexadezimal.
    std.fmt.fmtSliceHexLower(salt[0..]),
    std.fmt.fmtSliceHexLower(nonce[0..]),
    std.fmt.fmtSliceHexLower(tag[0..]),
    std.fmt.fmtSliceHexLower(data),
});
} else {
    // Da wir die Daten in Hexadezimal serialisiert haben, müssen wir diese
    // wieder voneinander trennen und in Binärdaten umwandeln.
    var si = std.mem.split(u8, data, ":");

    const salt = si.next();
    if (salt == null or salt.?.len != 32 * 2) {
        std.log.err("invalid data (missing salt)", .{});
        return;
    }
    var salt_: [32]u8 = undefined;
    _ = try std.fmt.hexToBytes(&salt_, salt.?);

    const nonce = si.next();
    if (nonce == null or nonce.?.len != XChaCha20Poly1305.nonce_length * 2) {
        std.log.err("invalid data (missing nonce)", .{});
        return;
    }
}

```



```

    }
    var nonce_: [XChaCha20Poly1305.nonce_length]u8 = undefined;
    _ = try std.fmt.hexToBytes(&nonce_, nonce.?);

    const tag = si.next();
    if (tag == null or tag?.len != XChaCha20Poly1305.tag_length * 2) {
        std.log.err("invalid data (missing tag)", .{});
        return;
    }
    var tag_: [XChaCha20Poly1305.tag_length]u8 = undefined;
    _ = try std.fmt.hexToBytes(&tag_, tag.?);

    const ct = si.next();
    if (ct == null) {
        std.log.err("invalid data (missing cipher text)", .{});
        return;
    }

    const pt = try allocator.alloc(u8, ct?.len / 2);
    defer {
        @memset(pt, 0);
        allocator.free(pt);
    }

    _ = try std.fmt.hexToBytes(pt, ct.?);

    // Danach können wir die deserialisierten Daten verwenden um
    // den Ciphertext zu entschlüsseln.
    var key: [XChaCha20Poly1305.key_length]u8 = undefined;
    try argon2.kdf(allocator, &key, password.?, &salt_, .{
        .t = 3,
        .m = 4096,
        .p = 1,
    }, .argon2id);

    try XChaCha20Poly1305.decrypt(pt, pt, tag_, "", nonce_, key);

    const stdout = std.io.getStdOut();
    try std.fmt.format(stdout.writer(), "{s}", .{pt});
}
}

```

In diesem Beispiel laufen eine Vielzahl von Konzepten zusammen, die sie im Lauf der Zeit noch häufiger antreffen werden. Unsere Anwendung erwartet Daten, z.B. den Inhalt einer Datei, über stdin, sowie zwei Kommandozeilenargumente: `--password` und `--encrypt` bzw. `--decrypt`. Basierend auf diesen Argumenten werden die übergebenen Daten entweder verschlüsselt oder entschlüsselt und nach stdout geschrieben.

Wir beginnen mit einigen Top-Level-Deklarationen, damit wir den Pfad zu Datenstrukturen, wie etwa `XChaCha20Poly1305`, nicht immer ausschreiben müssen. Weiterhin definieren wir ein Enum `Mode` welches zwei operationelle Zustände ausdrücken kann, Verschlüsselung (`encrypt`) und Entschlüsselung (`decrypt`).

Innerhalb von `main` parsen wir zuerst die übergebenen Argumente, indem wir durch die Funktion `argsWithAllocator()` einen Iterator über die Kommandozeilenargumente beziehen und mithilfe dessen über die einzelnen Argumente iterieren. Iteratoren sind ein häufig wiederzufindendes Konzept und lassen sich hervorragend mit `while` Schleifen kombinieren. Solange `ai.next()` ein Element zurückliefert, wird diese an `arg` gebunden und die Schleife wird fortgeführt. Liefert `next()` den Wert `null` zurück, so wird automatisch aus der Schleife ausgebrochen.

Danach stellen wir sicher, dass sowohl ein Passwort als auch ein Modus vom Nutzer spezifiziert wurden. Sollte eines der beiden Argumente fehlen, so wird ein entsprechender Fehler gelogged und der Prozess vorzeitig beendet.

Als nächstes wird eine über `stdin` übergebene Datei eingelesen und an die Konstante `datagebunden`. Da der für die Datei benötigte Speicher dynamisch alloziert wird muss dieser wieder freigegeben werden. Hierfür wird eine `defer`-Block verwendet, der vor Beendigung der Anwendung ausgeführt wird. Innerhalb dieses Blocks wird zusätzlich der Speicherinhalt mittels `@memset` überschrieben.



Der Umgang mit Sicherheitsrelevanten Daten ist durchaus herausfordernd. Grundsätzlich muss darauf geachtet werden, dass sensible Daten nicht zu lange im Speicher verweilen. Voraussetzung hierfür ist, dass Sie überhaupt wissen wo überall sensible Daten abgespeichert werden. Zum einen können Sie Daten, nachdem diese nicht mehr benötigt werden, überschreiben. Sie sollten jedoch auch weniger offensichtlich Angriffsvektoren, wie das Swappen von Hauptspeicher, im Hinterkopf behalten.

Sowohl für die Ver- als auch Entschlüsselung muss zuerst ein geheimer Schlüssel vom übergebenen Passwort, mittels einer Key-Derivation-Funktion, abgeleitet werden. Für diese Beispiel wird *Argon2id*<sup>19</sup>, der Gewinner der 2015 Password Hashing Competition, verwendet. Die Berechnung eines Schlüssels durch Argon2 hängt von den Folgenden (öffentlichen) Parametern ab:

- Salt: eine zufällige Sequenz die in die Schlüsselberechnung einfließt.
- Time: Die Anzahl an Iterationen für die Berechnung.
- Memory: Die Speicher-Kosten für die Berechnung.
- Parallelismus: Die Anzahl an parallelen Berechnungen.

Time, Memory und Parallelismus bestimmen wie aufwändig die Ableitung eines Schlüssels ist. Grundsätzlich gilt: je aufwendiger desto besser, jedoch schlägt sich dies auch in einer längeren Wartezeit nieder (spielen Sie deshalb gerne mit den Parametern). Alle Parameter werden bei der Verschlüsselung festgelegt und müssen mit dem Ciphertext zusammen gespeichert werden, da

---

<sup>19</sup><https://en.wikipedia.org/wiki/Argon2>

bei der Entschlüsselung die selben Parameter wieder in die KDF einfließen müssen um den Selben Schlüssel vom Passwort abzuleiten.

Zur Verschlüsselung wird eine zufällige Nonce generiert welche zusammen mit den zu verschlüsselnden Daten, einem Zeiger auf ein Array für den Tag, zusätzliche Daten (in diesem Fall der leere String "") und dem abgeleiteten Schlüssel an `encrypt` übergeben werden. Die Funktion verschlüsselt daraufhin die Daten. Danach wird der Salt, die Nonce, der Tag, sowie die verschlüsselten Daten, getrennt durch ein `:`, in die Standardausgabe `stdout` geschrieben.

Für die Entschlüsselung wird dieser String anhand der `:`, mittels `split`, aufgeteilt. Sollten die eingelesenen Daten nicht im erwarteten Format vorliegen, das heißt Salt, Nonce, Tag oder Ciphertext fehlen, so wird ein Fehler ausgegeben und die Anwendung beendet. Andernfalls, werden die eingelesenen Parameter verwendet um den Ciphertext, mittels `decrypt`, wieder zu entschlüsseln.

Das kleine Verschlüsselungsprogramm kann wie folgt verwendet werden:

```
$ cat hello.txt
Hello, World!
$ cat hello.txt | ./encrypt --password=supersecret --encrypt > secret.txt
$ cat secret.txt
828dfa14efa4b1f8242a8258a411301bd79bc4b7528294500305a4e9baaecbba:
85e4593786697e4e49212131a8e6e6bb68d25f43613dd870:ec666e95ebel1fa4c
53a1183379ae0dbd:80a7fe0475834364229c15dfb96d
$ cat secret.txt | ./encrypt --password=supersecret --decrypt
Hello, World!
```

## Graphische Applikationen

Ein weiterer Anwendungsfall für Zig ist die Entwicklung graphischer Applikationen. Hierfür existiert eine Vielzahl an Bibliotheken, darunter GTK und QT. Was beide Bibliotheken gemeinsam haben ist, dass sie in C beziehungsweise C++ geschrieben sind. Normalerweise würde das die Entwicklung von Bindings voraussetzen, um die Bibliotheken in anderen Sprachen nutzen zu können. Zig integriert jedoch direkt C, wodurch C-Bibliotheken direkt verwendet werden können<sup>20</sup>.

In diesem Abschnitt zeige ich Ihnen, wie sie eine simple GUI-Applikation mit GTK4 und Zig schreiben können. Hierfür müssen Sie zuerst einen neuen Projektordner anlegen.

```
$ mkdir gui
$ cd gui
$ zig init
info: created build.zig
info: created build.zig.zon
info: created src/main.zig
info: created src/root.zig
info: see `zig build --help` for a menu of options
```

---

<sup>20</sup>Mit wenigen Einschränkungen. Zig scheitert zurzeit noch an der Übersetzung einer Makros.

Danach fügen Sie gtk4 als Bibliothek zu Ihrer Anwendung hinzu. Hierfür öffnen Sie `build.zig` mit einem Texteditor und erweitern die Datei um die folgenden Zeilen:

```
// chapter01/gui/build.zig
//...
const exe = b.addExecutable(.{
    //...
});
// Fügen Sie die folgenden beiden Zeilen hinzu
exe.linkLibC();
exe.linkSystemLibrary("gtk4");
//...
```

Stellen Sie sicher, dass Sie die Developer-Bibliothek von GTK4 auf Ihrem System installiert haben. Unter Debian/Ubuntu können Sie diese über den APT-Paket-Manager installieren.

```
sudo apt install libgtk-4-dev
```

Führen Sie danach **zig build** aus um zu überprüfen, dass Zig die benötigte Bibliothek auf Ihrem System findet. Erzeugen Sie als nächstes die Datei `src/gtk.zig` und fügen Sie den Folgenden Code hinzu:

```
// chapter01/gui/src/gtk.zig
pub usingnamespace @cImport({
    @cInclude("gtk/gtk.h");
});

const c = @cImport({
    @cInclude("gtk/gtk.h");
});

/// g_signal_connect re-implementieren
pub fn z_signal_connect(
    instance: c.gpointer,
    detailed_signal: [*c]const c.gchar,
    c_handler: c.GCallback,
    data: c.gpointer,
) c.gulong {
    var zero: u32 = 0;
    const flags: *c.GConnectFlags = @as(*c.GConnectFlags, @ptrCast(&zero));
    return c.g_signal_connect_data(
        instance,
        detailed_signal,
        c_handler,
        data,
        null,
        flags.*,
    );
}
```

Zig ist zwar ziemlich gut darin mit C zu integrieren, jedoch werden Sie von Zeit zu Zeit noch auf Probleme stoßen. In den meisten Fällen lässt sich dies jedoch relativ einfach lösen. Innerhalb von

src/gtk.zig inkludieren wir zuerst die GTK4 Header-Datei gtk.h. Wie Ihnen vielleicht aufgefallen ist, haben wir an keiner Stelle innerhalb von build.zig auf diese Datei verwiesen. Zig reicht es in den aller meisten Fällen aus, wenn Sie die Bibliothek benennen die Sie einbinden möchten und fügt die benötigten Pfade automatisch hinzu.

Das Schlüsselwort `usingnamespace` sorgt dafür, dass wir auf alle in `gtk.h` deklarierten Objekte, über `gtk.zig`, direkt zugreifen können.

Eine in `gtk.h` deklarierte Funktion, die wir später noch benötigen, ist `g_signal_connect`. Diese lässt sich leider nicht ohne weiteres direkt verwenden (einer der seltenen Fälle bei denen Zig derzeit noch versagt). Aus diesem Grund implementieren wir die Funktion selber und nennen unsere Implementierung `z_signal_connect`.

Nun haben wir alles vorbereitet und können uns um die eigentliche Anwendung kümmern. Ersetzen Sie den Code in `src/main.zig` mit dem folgenden Programm:

```
// chapter01/gui/src/main.zig
const std = @import("std");
const gtk = @import("gtk.zig");

fn onActivate(app: *gtk.GtkApplication) void {
    const window: *gtk.GtkWidget = gtk.gtk_application_window_new(app);

    gtk.gtk_window_set_title(
        @as(*gtk.GtkWindow, @ptrCast(window)),
        "Zig Basics",
    );
    gtk.gtk_window_set_default_size(
        @as(*gtk.GtkWindow, @ptrCast(window)),
        920,
        640,
    );

    gtk.gtk_window_present(@as(*gtk.GtkWindow, @ptrCast(window)));
}

pub fn main() !void {
    const application = gtk.gtk_application_new(
        "de.zig.basics",
        gtk.G_APPLICATION_FLAGS_NONE,
    );
    _ = gtk.z_signal_connect(
        application,
        "activate",
        @as(gtk.GCallback, @ptrCast(&onActivate)),
        null,
    );
    _ = gtk.g_application_run(
        @as(*gtk.GApplication, @ptrCast(application)),
        0,
    );
}
```

```

    null,
);
}

```

Ganz oben importieren wir die Standardbibliothek, als auch die Datei `gtk.zig` unter dem Namen `gtk`. Danach folgt die Funktion `onActivate`, welche verwendet wird um ein GTK-Fenster zu erzeugen. Schauen wir uns aber zuerst die `main` Funktion an.

Innerhalb von `main` wird als erstes, mithilfe von `gtk_application_new`, ein Anwendungsobjekt erzeugt, welches an die Konstante `application` gebunden wird. Als nächstes wird ein Callback registriert, der durch das Signal `activate` aufgerufen wird. Als Callback nutzen wir die Funktion `onActivate`. Nachdem die Anwendung mittels `g_application_run` die GTK-Anwendung gestartet hat, wird das `activate` Signal ausgelöst, wodurch `onActivate` aufgerufen wird.

Die Funktion `onActivate` erzeugt als erstes ein neues Fenster für die Anwendung und weist dem Fenster, mithilfe von `get_window_set_title`, den Titel *Zig Basics* zu. Danach wird eine Fenstergröße von `920 x 640` Pixeln festgelegt, bevor das Fenster mit `gtk_window_present` angezeigt wird.

Innerhalb des Root-Verzeichnisses des Projekts können Sie mit **zig build run** die Anwendung starten. Nach dem Starten des Programms sollten Sie ein leeres Fenster sehen.



Abbildung 2: Leeres GKT4-Fenster

## Parallelität

### C/C++ Build System

Wie wir anfangs gesehen haben bietet `zig` nicht nur einen Compiler, sondern ein vollständiges Build-System. Die Beschreibung, wie ein Projekt gebaut werden soll, erfolgt dabei direkt in `Zig` und nicht in einer eigenen Sprache wie Sie es vielleicht von `Make` oder `CMake` gewohnt sind. Das besondere ist, dass `zig` nicht nur ein Build-System für die Sprache `Zig` bereitstellt, sondern auch als Build-System für `C` und `C++` Projekte verwendet werden kann und damit u.a. eine Alternative zu `Make` und `CMake` darstellt. Dies unterstreicht die enge Beziehung zwischen `Zig` und `C` bzw. `C++`.

Um zu demonstrieren wie Sie Zig als Build-System für ein C Projekt verwenden können, werden wir für [\*https://github.com/libusb/hidapi\*](https://github.com/libusb/hidapi)<sup>21</sup> ein kleines, unvollständiges Build-Skript schreiben und im Anschluss für eine kleine Beispielanwendung verwenden. Zuerst müssen wir mittels **git clone** [\*\*https://github.com/libusb/hidapi.git\*\*](https://github.com/libusb/hidapi) das Projekt beziehen. Danach wechseln wir in den Ordner *hidapi* und erstellen eine neue Datei mit dem Namen *build.zig*.

```
git clone https://github.com/libusb/hidapi.git
cd hidapi
touch build.zig
```

## Bibliothek

Wird von einem C bzw. C++ Projekt bereits ein Build-System verwendet, so lohnt es sich die zugehörigen Dateien zu analysieren um ein Verständniss davon zu bekommen, wie das Projekt derzeit gebaut wird. Im gegebenen Fall ist dies CMake, d.h., wir schauen uns zuerst die Datei *CMakeLists.txt* an. Ein Großteil des Inhalts der Datei ist für unsere Absichten erst einaml irrelevant und hat größtenteils mit der Konfiguration des Projekts zu tun. Zu diesem Zeitpunkt möchten wir *hidapi* jedoch lediglich für Linux, ohne besondere Optionen, compilieren. Aus diesem Grund schauen wir nach weiteren Unterordnern, die von *CMakeLists.txt* eingebunden werden. Einer davon ist *src*.

```
add_subdirectory(src)
```

Zwei Dinge die wir von *src/CMakeLists.txt* entnehmen können ist, dass das Projekt genau eine Header-Datei enthält (*hidapi/hidapi.h*) und dass für jedes Target (Linux, BSD, macOS und Windows) ein eigener Unterordner existiert. Die Datei *linux/CMakeLists.txt* wird uns für diese Beispiel als Vorlage dienen.

```
cmake_minimum_required(VERSION 3.6.3...3.25 FATAL_ERROR)

add_library(hidapi_hidraw
    ${HIDAPI_PUBLIC_HEADERS}
    hid.c
)
target_link_libraries(hidapi_hidraw PUBLIC hidapi_include)

find_package(Threads REQUIRED)

include(FindPkgConfig)
pkg_check_modules(libudev REQUIRED IMPORTED_TARGET libudev)

target_link_libraries(hidapi_hidraw PRIVATE PkgConfig::libudev Threads::Threads)

set_target_properties(hidapi_hidraw
    PROPERTIES
        EXPORT_NAME "hidraw"
        OUTPUT_NAME "hidapi-hidraw"
        VERSION ${PROJECT_VERSION})
```

---

<sup>21</sup><https://github.com/libusb/hidapi>

```

        SOVERSION ${PROJECT_VERSION_MAJOR}
        PUBLIC_HEADER "${HIDAPI_PUBLIC_HEADERS}"
    )

# compatibility with find_package()
add_library(hidapi::hidraw ALIAS hidapi_hidraw)
# compatibility with raw library link
add_library(hidapi-hidraw ALIAS hidapi_hidraw)

if(HIDAPI_INSTALL_TARGETS)
    install(TARGETS hidapi_hidraw EXPORT hidapi
        LIBRARY DESTINATION "${CMAKE_INSTALL_LIBDIR}"
        ARCHIVE DESTINATION "${CMAKE_INSTALL_LIBDIR}"
        PUBLIC_HEADER DESTINATION "${CMAKE_INSTALL_INCLUDEDIR}/hidapi"
    )
endif()

```

```
hidapi_configure_pc("${PROJECT_ROOT}/pc/hidapi-hidraw.pc.in")
```

Als erstes definieren wir eine neue Bibliothek und geben dieser einen Namen, sowie Informationen über das Target (z.B. X86\_64 Linux) sowie Informationen zu gewünschten Optimierungen. Außerdem definieren wir eine Option, um Anwendern die Möglichkeit zu geben zwischen einer statischen oder dynamischen Bibliothek zu wählen. Als Default wählen wir dabei `true`, d.h. standardmäßig wird eine statische Bibliothek gebaut.

```

// hidapi/build.zig
const std = @import("std");

pub fn build(b: *std.Build) void {
    const static = b.option(bool, "static", "Make a static library") orelse true;

    const target = b.standardTargetOptions(.{});
    const optimize = b.standardOptimizeOption(.{});

    const lib = if (static) blk: {
        break :blk b.addStaticLibrary(.{
            .name = "hidapi",
            .target = target,
            .optimize = optimize,
        });
    } else blk: {
        break :blk b.addSharedLibrary(.{
            .name = "hidapi",
            .target = target,
            .optimize = optimize,
        });
    };

    b.installArtifact(lib);
}

```



Nachdem Sie den obigen Code zu *build.zig* hinzugefügt haben können Sie mittels **build zig --help** sich einen Hilfetext ausgeben lassen, der u.a. unsere Option beinhalten sollte.

```
$ build zig --help
Usage: /usr/local/bin/zig-linux-x86_64-0.13.0/zig build [steps] [options]
...
Project-Specific Options:
  -Dstatic=[bool]           Make a static library
```

Zwar können wir uns einen Hilfetext anzeigen lassen, jedoch schlägt das Compilieren beim Ausführen von **build zig** fehl. Was fehlt ist Quellcode, aus der die Bibliothek gebaut werden soll. Um es einfach zu halten unterstützen wir in diesem Beispiel nur Linux, es steht Ihnen jedoch Frei auch Unterstützung für andere Betriebssysteme hinzuzufügen. Über die *target* Variable können wir u.a. das Betriebssystem, für welches die Bibliothek gebaut werden soll, bestimmen.



Zig unterstützt Cross-Compilation, d.h. Architektur und Betriebssystem auf dem eine Anwendung oder Bibliothek compiliert wird kann sich von der Architektur bzw. dem Betriebssystem unterscheiden, für welches Compiliert wird. Die Variable *target* enthält Informationen über das Zielsystem.

Für unser Beispiel machen wir eine einfache Fallunterscheidung. Sollte das Zielsystem *.linux* sein, so fügen wir die benötigten *.c* Dateien zu *lib* hinzu und linken zusätzlich die benötigte Bibliothek *libudev*. Andernfalls kehren wir einfach frühzeitig von der *build* Funktion zurück.

```
// hidapi/build.zig
// ...

if (target.result.os.tag == .linux) {
    lib.addSourceFiles(.{
        .files = &{"linux/hid.c"},
        .flags = &{"-std=gnu11"},
    });

    // Manche Linux-Distros (z.B. OpenSuse) besitzen keine Developer-Package
    // von libudev, d.h. es fehlt die Datei `libudev.h`. In diesem Fall kann
    // die Datei manuell bezogen
    // https://github.com/mcatalancid/libudev/blob/1.8.2/src/libudev.h
    // und in das Projekt integriert werden. In diesem Fall einfach die
    // folgende Zeile einfügen:
    // lib.addSystemIncludePath(b.path("./"));

    // Abhängig von der Linux-Distor muss `udev` evtl. durch `libudev` ersetzt
    // werden.
    lib.linkSystemLibrary("udev");
} else {
    // An dieser Stelle wäre eine bessere Fehlerkommunikation angebracht.
    return;
}
```

```
// Der Unterordner ./hidapi enthält die `hidapi.h` Header-Datei
lib.addIncludePath(b.path("hidapi"));
lib.linkLibC();

// ...

b.installArtifact(lib);
```

Beim linken von Bibliotheken mittels `linkSystemLibrary` sucht Zig nach den Bibliotheken in den gängigsten Ordnern, darunter `/usr/lib`. Außerdem werden die zugehörigen Header automatisch dem Include-Pfad hinzugefügt. Oft wird jedoch das Developer-Paket einer Bibliothek benötigt um auch an die Header-Dateien auf seinem System zu gelangen. Im Fall von `libudev` ist ein solches Developer-Paket jedoch nicht unter allen gängigen Linux Distributionen vorhanden, z.B. gibt es unter Ubuntu `libudev-dev`, jedoch nicht unter OpenSuse<sup>22</sup>.

Da unsere `hidapi` Bibliothek selber auch `hidapi.h` zum compilieren benötigt, geben wir dessen relativen Pfad mittels der Funktion `addIncludePath()` an. Außerdem müssen wir bei C bzw. C++ Projekten in den meisten Fällen, mittels `linkLibC()`, die C-Standard-Bibliothek verlinken.

Damit ist unser Build-Script vollständig und wir können mit **build zig** unsere Bibliothek (unter Linux bzw. für Linux) bauen. Nach dem Compilieren findet sich diese unter `zig-out/lib` im Root-Verzeichnis von `hidapi`.

Um die `hidapi` Bibliothek, welche durch `build.zig` beschrieben wird, in anderen Zig Projekten einfach verwenden zu können, müssen wir noch zwei Dinge hinzufügen. Zum einen sollten wir die für `hidapi` benötigten Header (`hidapi/hidapi.h`) exportieren.

```
// hidapi/build.zig
// ...
lib.installHeader(b.path("hidapi/hidapi.h"), "hidapi.h");
// ...
```

Damit wird beim Linken von `hidapi` durch ein anderes Zig-Projekt, diesem auch gleichzeitig der Pfad zur Header-Datei mitgeteilt.

Zum anderen wird eine `build.zig.zon` Datei benötigt, die weitere Informationen zum `hidapi` Projekt bereitstellt.

```
// hidapi/zig.build.zon
.{
    .name = "hidapi",
    .version = "0.1.0",
    .dependencies = .{},
    .paths = .{
        "build.zig",
        "build.zig.zon",
        ".builds",
        "android/jni",
    },
}
```

---

<sup>22</sup>Die Beschaffung benötigter Bibliotheken ist in der Regel „Out-of-Scope“, d.h. es ist die Aufgabe des Nutzers die benötigten Bibliotheken auf seinem System zu installieren. Für das gegebene Beispiel habe ich trotzdem ein Kommentar diesbezüglich eingefügt, mit einem Link zur benötigten Header-Datei.

```

    "dist",
    "documentation",
    "doxygen",
    "hidapi",
    "hidtest",
    "libusb",
    "linux",
    "m4",
    "mac",
    "pc",
    "src",
    "subprojects",
    "testgui",
    "udev",
    "windows",
    ".appveyor.yml",
    ".cirrus.yml",
    ".gitattributes",
    ".gitignore",
    "AUTHORS.txt",
    "HACKING.txt",
    "LICENSE.txt",
    "README.md",
    "VERSION",
    "bootstrap",
    "configure.ac",
    "meson.build",
  },
}

```

Die meisten Felder sind dabei selbsterklärend. Das Feld `.paths` gibt an welche Dateien zum gegebenen Projekt gehören, woraus sich u.a. der Paket-Hash berechnet.

## Verwendung der Bibliothek



# Standard Typen

Zig ist eine kompilierte Sprache, d.h. sie wird, bevor der Programmcode ausgeführt werden kann, in eine Sprache übersetzt die vom Prozessor verstanden wird. Die Übersetzungsarbeit übernimmt dabei der Compiler.

TBD

Typ	Beschreibung	Beispielwerte
i8, u65	Vorzeichen(un)behaftete Ganzzahlen mit der angegebenen Bitbreite (von 0 bis $2^{16} - 1$ ).	0x32, -1
usize, isize	Vorzeichen(un)behaftete Ganzzahlen deren Bitbreite mit der der Architektur übereinstimmt, d.h. auf x86_64 wäre usize gleichbedeutend mit u64.	0xcafe_babe

## Ganzzahlen (Integer)

Zig unterstützt Ganzzahlen mit einer beliebigen Bitbreite. Der Bezeichner eines jeden Integer-Typen beginnt mit einem Buchstaben i (vorzeichenbehaftet; signed) oder u (vorzeichenunbehaftet; unsigned) gefolgt von einer oder mehreren Ziffern, welche die Bitbreite in Dezimal darstellen. Als Beispiel, i7 ist eine vorzeichenbehaftete Ganzzahl der sieben Bit zur Kodierung der Zahl zur Verfügung stehen. Die Aussage, dass die Bitbreite beliebig ist entspricht dabei nicht ganz der Wahrheit. Die maximal erlaubte Bitbreite beträgt  $2^{16} - 1 = 65535$ .

Typ	Wertebereich
i7	$-2^6$ bis $2^6 - 1$
i32	$-2^{31}$ bis $2^{31} - 1$
u8	0 bis $2^8 - 1$
u64	0 bis $2^{64} - 1$

Vorzeichenbehaftete Ganzzahlen werden im Zweierkomplement dargestellt<sup>23</sup>. In Assembler wird nicht zwischen vorzeichenbehafteten und vorzeichenunbehafteten Zahlen unterschieden. Alle mathematischen Operationen werden von der CPU auf Registern, mit einer festen Bitbreite (meist 64 Bit auf modernen Computern), ausgeführt. Dabei entspricht jede, vom Computer ausgeführte, arithmetische Operationen effektiv einem „Rechnen mit Rest“, auch bekannt als modulare Arithmetik<sup>24</sup>. Die Bitbreite  $m$  der Register (z.B. 64) repräsentiert dabei den Modulo  $2^m$ . Damit entspricht ein 64 Bit Register dem Restklassenring  $\mathbb{Z}_{2^{64}} = \{0, 1, 2, \dots, 2^{64} - 1\}$  und jegliche Addition zweier Register resultiert in einem Wert der ebenfalls in  $\mathbb{Z}_{2^{64}}$  liegt, d.h. auf x86\_64 wäre die Instruktion `add rax, rbx` äquivalent zu  $rax = rax + rbx \bmod 2^{64}$ . Dieses Verhalten überträgt sich analog auf Ganzzahlen in Zig.

Das Zweierkomplement einer Zahl  $a \in \mathbb{Z}_m$  ist das additive Inverse  $a'$  dieser Zahl, d.h.  $a + a' \equiv 0$ . Dieses kann mit  $a' = m - a$  berechnet werden. Für i8 wäre das additive Inverse zu  $a = 4$  die Zahl  $a' = 2^8 - 4 = 256 - 4 = 252$ . Addiert man beide Zahlen modulo 256, so erhält man wiederum das neutrale Element 0,  $a + a' \bmod 256 = 4 + 252 \bmod 256 = 256 \bmod 256 = 0$ . Das Zweierkomplement hat seinen Namen jedoch nicht von der Subtraktion, sondern von der speziellen Weise wie das additive Inverse einer Zahl bestimmt wird. Dieser Vorgang kann wie folgt beschrieben werden:

1. Gegeben eine Zahl in Binärdarstellung, invertiere jedes Bit, d.h. jede 1 wird zu einer 0 und umgekehrt.
2. Addiere 1 auf das Resultat und ignoriere mögliche Überläufe.

Für das obige Beispiel mit der Zahl 4 vom Typ i8 sieht dies wie folgt aus:

$$\begin{aligned} 00000100_2 &= 4_{16} && \text{invertiere alle Bits der Zahl 4} \\ 11111011_2 &= 251_{16} && \text{addiere 1 auf die Zahl 251} \\ 11111100_2 &= 252_{16} \end{aligned}$$

Zur Compile-Zeit bekannte Literale vom Typ `comptime_int` haben kein Limit was ihre Größe (in Bezug auf die Bitbreite) und konvertieren zu anderen Integertypen, solange das Literal im Wertebereich des Typen liegt.

```
// Variable `i` vom Typ `comptime_int`
var i = 0;
```

Um die Variable zur Laufzeit modifizieren zu können, muss ihr eine expliziter Type mit fester Bitbreite zugewiesen werden. Dies kann auf zwei Weisen erfolgen.

1. Deklaration der Variable `i` mit explizitem Typ, z.B. `var i: usize = 0`.
2. Verwendung der Funktion `@as()`, von welcher der Compiler den Type der Variable `i` ableiten kann, z.B. `var i = @as(usize, 0)`.

Ein häufiger Fehler, der aber schnell behoben ist, ist die Verwendung einer Variable vom Typ `comptime_int` in einer Schleife.

<sup>23</sup>[https://en.wikipedia.org/wiki/Two's\\_complement](https://en.wikipedia.org/wiki/Two's_complement)

<sup>24</sup>[https://de.wikipedia.org/wiki/Modulare\\_Arithmetik](https://de.wikipedia.org/wiki/Modulare_Arithmetik)

```
var i = 0;
while (i < 100) : (i += 1) {}
```

Was zu einem entsprechenden Fehler zur Compilezeit führt.

```
$ zig build-exe chapter02/integer.zig
error: variable of type 'comptime_int' must be const or comptime
    var i = 0;
    ^
```

**note:** to modify this variable at runtime, it must be given an explicit fixed-size number type

Der Zig-Compiler ist dabei hilfreich, indem er neben dem Fehler auch einen Lösungsansatz bietet. Nachdem der Variable `i` ein expliziter Typ zugewiesen wird (`var i: usize`) compiliert das Programm ohne weitere Fehler.

Optional können die Prefixe `0x`, `0o` und `0b` an ein Literal angehängt werden um Literale in Hexadezimal, Octal oder Binär anzugeben, z.B. `0xcafebab`.

Um größere Zahlen besser lesbar zu machen, kann ein Literal mit Hilfe von Unterstrichen aufgeteilt werden, z.B. `0xcafe_babe`.

Operatoren wie `+` (Addition), `-` (Subtraktion), `*` (Multiplikation) und `/` (Division) führen bei einem Überlauf zu undefiniertem Verhalten (engl. undefined behavior). Aus diesem Grund stellt Zig spezielle Versionen dieser Operatoren zur Verfügung, darunter:

- Operatoren für Sättigungsarithmetik: Alle Operationen laufen in einem festen Intervall zwischen einem Minimum und einem Maximum ab welches nicht unter- bzw. überschritten werden kann.
  - Addition (`+|`): `@as(u8, 255) +| 1 == @as(u8, 255)`
  - Subtraktion (`-|`): `@as(u32, 0) -| 1 == 0`
  - Multiplikation (`*|`): `@as(u8, 200) *| 2 == 255`
- Wrapping-Arithmetik: Dies ist äquivalent zu modularer Arithmetik.
  - Addition (`+%`): `@as(u32, 0xffffffff) +% 1 == 0`
  - Subtraktion (`-%`): `@as(u8, 0) -% 1 == 255`
  - Multiplikation (`*%`): `@as(u8, 200) *% 2 == 144`

## Fließkommazahlen (Float)

Im Gegensatz zu Integern erlaubt Zig keine beliebige Bitbreite für Fließkommazahlen. Zur Verfügung stehen:

Typ	Repräsentation
f16	IEEE-754-2008 binary16
f32	IEEE-754-2008 binary32
f64	IEEE-754-2008 binary64

Typ	Repräsentation
f80	IEEE-754-2008 80-bit extended precision
f128	IEEE-754-2008 binary128

Literale sind immer vom Typ `comptime_float`, welcher äquivalent zum größtmöglichen Fließkommatypen (f128) ist, und können zu jedem beliebigen Fließkommatypen konvertiert werden. Enthält ein Literal keinen Bruchteil, so ist eine Konvertierung zu einem Integertyp ebenfalls möglich.

```
const fp = 123.0E+77;
const hfp = 0x103.70p-5;
```

Der Typ f32 entspricht dem Typ `float` (single precision) in C, während f64 dem Typ `double` (double precision) entspricht. Je nach Prozessortyp stehen dedizierte Maschineninstruktionen für zumindest einen Teil der Typen zur Verfügung, was eine effizientere Verwendung ermöglicht. Auf *x86\_64* Prozessor stehen z.B. Instruktionen für single und double Precision zur Verfügung.

Die interne Darstellung einer Fließkommazahl besteht für das Format *IEEE-754* aus einem Vorzeichenbit, gefolgt von einem Exponenten und einem Bruch. Wie viele Bits jeweils für Exponent und Bruch zur Verfügung stehen ist abhängig von der Bitbreite der Fließkommazahl. Für *IEEE-754 binary32* sieht dies wie folgt aus:

31	30	...	23	22	...	0
s	exponent (e)			fraction (f)		

Diese Darstellung entspricht der Gleichung  $(-1)^s * 1.f * 2^{e-127}$ . Der Bruch  $f$  entspricht einer normalisierten, binär kodierten Fließkommazahl, d.h. die Zahl wird um eine entsprechende Anzahl an Stelle verschoben, sodass genau eine führende Eins vor dem Komma steht. Als Beispiel entspricht die Fließkommazahl 3.25 in binär der Zahl 11.01 oder anders ausgedrückt  $11.01 * 2^0$ . Um die Zahl zu normalisieren wird diese nun um eine Stelle nach rechts verschoben  $1.101 * 2^1$ . Die Zahl nach der führenden Eins (101) entspricht  $f$  und der Exponent  $e$  ist die Summe des Exponenten der normalisierten Darstellung und einem Bias (im Fall von f32 ist dieser 127), d.h.  $e = 1 + 127 = 128_{10} = 10000000_2$ . Damit wird 3.25 wie folgt kodiert:

31	30	...	23	22	...	0
0	10000000 <sub>2</sub>			101000000000000000000000 <sub>2</sub>		

Aufgrund der Darstellung von Fließkommazahlen kann sich die Ausführung bestimmter Operationen, wie ein Tests auf Gleichheit (`==`), als trickreich herausstellen. Ein Beispiel ist die wiederholte Addition der Fließkommazahl 0.1. Die Summe  $\sum_{k=1}^{10} 0.1$  ist erwartungsgemäß 1.0, je nach Präzision der Fließkommazahl gilt jedoch  $\sum_{k=1}^{10} 0.1 \neq 1.0$ .

## Arrays und Slices



Zig besitzt eine Vielzahl an Datentypen um eine (lineare) Sequenz an Werten im Speicher darzustellen, darunter:

- Der Typ `[N]T` repräsentiert ein Array vom Typ `T` bestehend aus `N` Werten. Die Größe eines Arrays ist zur Compilezeit bekannt und Arrays werden grundsätzlich auf dem Stack alloziert. Damit kann ein Array weder erweitert noch verkleinert werden.
- Der Typ `[]T` bzw. `[]const T` repräsentiert ein Slice vom Typ `T`, bestehend aus einem Zeiger und einer Länge. Die Länge eines Slices ist zur Laufzeit bekannt. Slices referenzieren eine Sequenz von Werten. Dies kann z.B. ein Array sein oder auch eine auf dem Heap gespeicherte Sequenz. Die von einem konstanten Slice `[]const T` referenzierten Werte können gelesen, jedoch nicht verändert werden, während die Werte eines Slices `[]T` sowohl gelesen als auch verändert werden können.

Sowohl Arrays als auch Slices erlauben den Zugriff auf deren Länge durch den Ausdruck `.len`.

```
// chapter02/slices.zig
var a = [_]u8{ 1, 2, 3, 4 };
std.log.info("length of a is {d}", .{a.len});
const s = &a;
std.log.info("length of a is still {d}", .{s.len});
```

Mit dem Address-Of Operator `&` kann ein Slice für ein Array erzeugt werden. Alternativ kann auch der Ausdruck `a[0..]` verwendet werden, der einen Bereich innerhalb des Arrays beschreibt. Grundsätzlich liegt das erste Element einer Sequenz immer an Index 0 und es kann mit `a[0]` auf dieses zugegriffen werden. Das letzte Element liegt immer an der Stelle `a.len - 1` und es kann mit `a[a.len - 1]` darauf zugegriffen werden. Der Index muss dabei immer ein Integer vom Typ `usize` oder ein Literal sein, das zu diesem Typ konvertiert werden kann. Die Verwendung anderer Typen als Index führt zu einem Fehler zur Compilezeit.

Auf den Zeiger eines Slices kann mit `.ptr` zugegriffen werden, z.B. `s.ptr`.

Zig überprüft bei dem Zugriff auf eine Array oder Slice zur Laufzeit, dass der Index innerhalb des Speicherbereichs der Sequenz liegt. Läßt eine Anwendung über die Grenzen der Sequenz, so führt dies zu einem Fehler zur Laufzeit der den Prozess beendet. Dies verhindert typische Speicherfehler wie Buffer-Overflows and Buffer-Overreads die in Sprachen wie C weit verbreitet sind und in der Vergangenheit zu Hauf von Angreifern ausgenutzt wurden um Anwendungen zu exploiten.

```
// chapter02/slices.zig
var i: usize = 0;
while (true) : (i += 1) {
    a[i] += 1;
}

$ zig build-exe slices.zig -Doptimize=ReleaseFast
$ ./slices
info: length of a is 4
info: length of a is still 4
thread 1232 panic: index out of bounds: index 4, len 4
slices.zig:14:10: 0x103544c in main (slices)
```

```

    a[i] += 1;
    ^
start.zig:514:22: 0x1034c99 in posixCallMainAndExit (slices)
    root.main();
    ^
start.zig:266:5: 0x1034801 in _start (slices)
    asm volatile (switch (native_arch) {
    ^
???:?:?: 0x0 in ??? (???)
Aborted (core dumped)

```

## Arrays

Es gibt eine Vielzahl von Möglichkeiten um Arrays in Zig zu definieren. Die einfachste Möglichkeit ist, eine Sequenz von Werten in geschweiften Klammern anzugeben.

```

const prime: [5]u8 = {2, 3, 5, 7, 11};
const names = [3][]const u8{"David", "Franziska", "Sarah"};

```

Für den Fall, dass initial keine Werte bekannt sind kann ein Array mit undefined initialisiert werden. In diesem Fall ist der Inhalt des Speichers undefiniert.

```

const some: [1000]u8 = undefined;

```

Arrays können aber auch mit einem bestimmten Wert initialisiert werden. Im unteren Beispiel wird das gesamte Array mit 0 Werten initialisiert.

```

const some: [1000]u8 = {0} ** 1000;

```

Die Länge eines Arrays muss immer zur Compilezeit bekannt sein. Dementsprechend können keine Variablen zur Angabe der Länge verwendet werden, außer die Variable ist vom Typ `comptime_int`. Sollte ein Array benötigt werden, dessen Länge nur zur Laufzeit bekannt ist, so muss der Speicher entweder manuell alloziert oder auf einen Kontainertypen wie `ArrayList` aus der Standardbibliothek zurückgegriffen werden<sup>25</sup>.

Viel Funktionen die über Sequenzen arbeiten erwarten ein Slice und kein Array. Zig konvertiert dabei nicht automatisch Arrays zu Slices, d.h. bei einem Aufruf muss explizit der Address-Of Operator & auf das Array angewandt werden oder alternativ ein Slice mit dem [] Operator festgelegt werden.

```

// chapter02/coersion.zig
const std = @import("std");

pub fn main() void {
    const a: [5]u8 = { 1, 2, 3, 4, 5 };

    foo(&a);
    foo(a[1..]);
}

fn foo(s: []const u8) void {

```

<sup>25</sup>Mehr dazu in folgenden Kapiteln.

```

    for (s) |e| {
        std.log.info("{d}", .{e});
    }
}

$ ./coersion
info: 1
info: 2
info: 3
info: 4
info: 5
info: 2
info: 3
info: 4
info: 5

```

## Slices

Slices `[]T` werden ohne Angabe einer Länge geschrieben und repräsentieren eine lineare Sequenz an Werten. Konzeptionell ist ein Slice eine Zeiger vom Typ `std.builtin.Type.Pointer`. Schaut man sich die Definition von `Slice` in `zig/src/mutable_value.zig`<sup>26</sup> an, so sieht man, dass ein Slice durch einen Zeiger (`ptr`) auf den Beginn des referenzierten Speicherbereichs und eine Länge (`len`) beschrieben wird.

```

// github.com/ziglang/zig/src/mutable_value.zig
pub const Slice = struct {
    ty: InternPool.Index, // wir ignorieren dieses Feld :)
    /// Must have the appropriate many-ptr type.
    ptr: *MutableValue,
    /// Must be of type `usize`.
    len: *MutableValue,
};

```

Je nach Typ einer Variable bzw. eines Parameters konvertiert Zig die Referenz zu einem Struct automatisch in ein Slice.

```

const b: [3][]const u8 = .{ "David", "Franziska", "Sarah" };

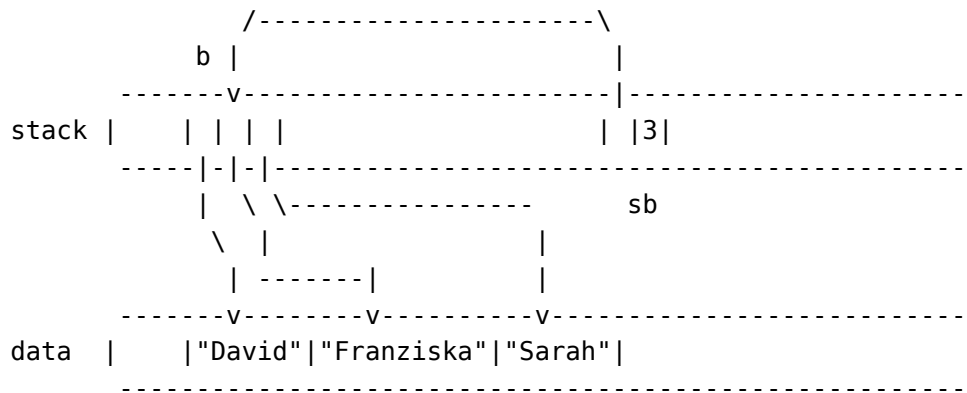
// Zig konvertiert die Referenz automatisch zu einem Slice.
const sb: []const []const u8 = &b;
_ = sb;

// `rb` ist ein Pointer zu einem Array.
const rb: *const [3][]const u8 = &b;
_ = rb;

```

Da `b` eine Konstante ist, muss auch das Slice `sb` (`[]const T`), sowie der Pointer `rb` auf das Array (`*const [N]T`) konstant sein. Wäre `b` eine Variable, so wäre auch das `const`, in Bezug auf das Slice bzw. den Pointer, optional, je nachdem ob das Array durch die jeweilige Referenz verändert werden soll oder nicht.

<sup>26</sup>[https://github.com/ziglang/zig/blob/624fa8523a2c4158ddc9fce231181a9e8583a633/src/mutable\\_value.zig](https://github.com/ziglang/zig/blob/624fa8523a2c4158ddc9fce231181a9e8583a633/src/mutable_value.zig)



Mithilfe des `[]` Operators können Slices für einen bestehenden Speicherbereich angegeben werden. Innerhalb der eckigen Klammern muss dafür ein Bereich spezifiziert werden, der durch das Slice eingegrenzt werden soll:

- `[0.. ]` : Der gesamte Bereich, vom ersten bis zum letzten Element.
- `[N..M]` : Ein Bereich beginnend ab Index N (eingeschlossen) und endend bei Index M (ausgeschlossen).

```
const name = "David";  
// Die ersten drei Buchstaben  
std.log.info("{s}", .{name[0..3]});  
// Die letzten zwei Buchstaben  
std.log.info("{s}", .{name[3..]});  
// Die mittleren drei Buchstaben  
std.log.info("{s}", .{name[1..4]});
```

Um Buffer-Overreads vorzubeugen überprüft Zig, dass die angegeben Indices valide sind. Sind die Indices zur Compilezeit bekannt, so führt ein invalider Index zu einem Compile-Fehler, andernfalls zu einer Panic zur Laufzeit.

```
// chapter02/slice_error.zig
const a = "this won't work";
// ...
const n: usize = 20;
std.log.info("{s}", .{a[1..n]});
```

Versucht man den obigen Code mit **zig build-exe chapter02/slice\_error.zig** zu Compilieren so erhält man den folgenden Fehler:

```
error: end index 20 out of bounds for array of length 15 +1 (sentinel)
std.log.info("{s}", .{a[1..n]});
```

## for-Schleife (Loop)

Es kann äußerst nützlich sein über den Inhalt eines Arrays oder Slices zu iterieren. Eine Möglichkeit dies zu tun ist mit Hilfe einer for-Schleife.

```
// chapter02/loop.zig
const names = [_][]const u8{ "David", "Franziska", "Sarah" };

for (names) |name| {
```

```
std.log.info("{s}", .{name});
}
```

Eine for-Schleife beginnt mit dem Schlüsselwort `for`, gefolgt von einer Sequenz, über die iteriert werden soll, in runden Klammern. Danach wird ein Bezeichner zwischen zwei `|` angegeben. Dem Bezeichner wird für jede Iteration der aktuelle Wert zugewiesen, d.h. für das obige Beispiel wird im ersten Schleifendurchlauf `name` der Wert "David" zugewiesen, im zweiten Durchlauf "Franziska" und so weiter. Nachdem über alle Elemente iteriert wurde, wird automatisch aus der Schleife ausgebrochen.

Eine Besonderheit von Zig ist, dass innerhalb einer for-Schleife simultan über mehrere Sequenzen iteriert werden kann.

```
for (names, 0..) |name, i| {
    std.log.info("{s} ({d})", .{ name, i });
}
```

Die Sequenzen werden, getrennt durch ein Komma, innerhalb der runden Klammer angegeben. Selbes gilt für die Bezeichner, an die die einzelnen Werte der Sequenzen gebunden werden. Im obigen Beispiel wird als zweite Sequenz `0..` angegeben, d.h. eine Sequenz von Ganzzahlen beginnend bei 0. Zig sorgt dabei automatisch dafür, dass `names` und `0..` über die selbe Länge verfügen, indem das Ende von `0..` automatisch bestimmt wird, d.h. für das gegebene Beispiel ist `0..` äquivalent zu `0..3`.

Sollten Sie über mehrere Arrays bzw. Slices gleichzeitig iterieren, so müssen sie sicherstellen, dass alle die selbe Länge besitzen!

```
const dishes = [_][]const u8{ "Apfelstrudel", "Pasta", "Quiche" };

for (names, dishes) |name, dish| {
    std.log.info("{s} likes {s}", .{ name, dish });
}
```

Mit dem Schlüsselwort `break` kann aus einer umschließenden Schleife ausgebrochen werden, d.h. das Programm wird unter der Schleife fortgeführt.

```
for (1..5) |i| {
    std.log.info("{d}", .{i});
    if (i == 2) break;
}
```

Mit dem Schlüsselwort `continue` können sie den restlichen Körper der Schleife überspringen und mit der nächsten Iteration beginnen. Sollten `continue` in der letzten Iteration der Schleife ausgeführt werden, so wird aus dieser ausgebrochen.

```
for (1..5) |i| {
    if (i == 2) continue;
    std.log.info("{d}", .{i});
}
```

Schleifen können auch geschachtelt werden. Wenn Sie innerhalb einer der inneren Schleifen, aus einer der Äußeren ausbrechen wollen, müssen Sie sogenannte Label verwenden, mit der sie einer bestimmten Schleife einen Namen geben können. Labels kommen vor dem `for` Schlüssel-

wort und enden immer mit einem `::`. Sie können sowohl mit `break` als auch `continue` verwendet werden.

```
outer: for (names) |name| {
    for (dishes) |dish| {
        std.log.info("{s}, {s}", .{ name, dish });
        // Da wir an dieser stelle aus der äußeren Schleife ausbrechen
        // ist nur eine Ausgabe auf der Kommandozeile zu sehen.
        break :outer;
    }
}
```

Zig erlaubt auch die Verwendung von `for`-Loops in Ausdrücken.

```
const pname = outer: for (names) |name| {
    if (name.len > 0 and (name[0] == 'p' or name[0] == 'P'))
        break :outer name;
} else blk: {
    break :blk "no name starts with p!";
};
std.log.info("found: {s}", .{pname});
```

In diesem Beispiel suchen wir nach einem Namen der mit dem Buchstaben P bzw. p beginnt. Sollte aus der Schleife mit `break` ausgebrochen werden, so wird der `else` Block nicht ausgeführt. Da `names` keinen solchen Namen beinhaltet wird der `else` Block aufgerufen und der String "no name starts with p!" der Konstanten `pname` zugewiesen.

Neben Schleifen können auch `if/else` Blöcken Label zugewiesen werden. Dies erlaubt es, mittels `break`, Werte aus dem Block heraus zu reichen, wie oben zu sehen ist.

Sie können das Beispiel mit **zig build-exe chapter02/loop.zig && ./loop** compilieren und ausführen.

## Zeiger (Pointer)

Zig unterscheidet zwischen zwei Arten von Zeigern, *single-item* und *many-item* Pointer.

Ein *single-item* Pointer `*T` zeigt auf exakt einen Wert im Speicher und kann mit der Syntax `ptr.*` dereferenziert werden. Mit Hilfe des Address-of-Operators `&` kann ein *single-item* Pointer bezogen werden.

```
// Definiere eine Variable vom Typ u8
var v: u8 = 128;
// Beziehe einen Zeiger auf `v`
const v_ptr = &v;
// Dereferenziere den Zeiger `v_ptr` und addiere 1 zu `v`
v.* += 1;
```

Ein *multi-item* Pointer `[*]T` zeigt auf eine lineare Sequenz an Werten im Speicher mit unbekannter Länge. Der Zeiger eines Slice (`.ptr`) ist ein *multi-item* Pointer. Allgemein teilen Slices und *multi-item* Pointer die selbe Index- und Slice-Syntax.

- `ptr[i]`

- `ptr[start..end]`
- `ptr[start..]`

Genau wie C erlaubt auch Zig Zeigerarithmetik auf multi-item Pointer.

```
// chapter02/pointer.zig
var array = [_]i32{ 1, 2, 3, 4 };

var array_ptr = array[0..].ptr;

std.log.info("{d}", .{array_ptr[0]});
array_ptr += 1;
std.log.info("{d}", .{array_ptr[0]});
```

Nach dem Compilieren mit **zig build-exe chapter02/pointer.zig** können wir die Beispiel Anwendung ausführen und sehen, dass die ersten beiden Zahlen von `array` ausgegeben werden, obwohl wir den selben Index für `array_ptr` verwenden. Grund dafür ist, dass wir den Zeiger selbst, zwischen dem ersten und zweiten Aufruf von `std.log.info()`, inkrementiert haben.

```
$ ./pointer
info: 1
info: 2
```

Ein weit verbreitetes Konzept in C sind NULL-terminierte Strings, d.h. ein 0-Byte wird hinter den letzten Character eines Strings geschrieben und markiert so dessen Ende. Zig bietet etwas sehr ähnliches, nämlich sentinel-terminated Pointer.

Ein sentinel-terminated Pointer wird durch einen Typ `[*:x]T` beschrieben, wobei `x` ein Wert vom Typ `T` ist und den Sentinel darstellt, der das Ende einer Sequenz markiert. Analog zu einem NULL-terminierten String vom Typ `char*` in C, schreibt man in Zig `[:0]u8`.



Im Allgemeinen werden in Zig Slices, gegenüber sentinel-terminated Pointern, präferiert. Der Grund hierfür ist, dass Slices über Bounds-Checking verfügen und so gängige Speicherfehler abgefangen werden können. Es gibt jedoch auch Situationen, in denen many-item Pointer bzw. sentinel-terminated Pointer explizit benötigt werden, z.B. beim Arbeiten mit C Code. Auf die Interoperabilität zwischen Zig und C wird in einem späteren Kapitel noch näher eingegangen.

## Container

Jedes syntaktische Konstrukt in Zig, welches als Namensraum dient und Variablen- oder Funktionsdeklarationen umschließt, wird als Container bezeichnet. Weiterhin können Container selbst Typdeklarationen sein, welche instantiiert werden können. Dazu zählen `structs`, `enums`, `unions` und sogar Sourcdateien mit der Dateiendung `.zig`.

Ein Merkmal welches Container von Blöcken unterscheidet ist, dass Container keine Ausdrücke enthalten, obwohl sowohl Container als auch Blöcke, mit der Ausnahme von Sourcdateien, in geschweifte Klammern (`{}`) gefasst werden.



In Zig ist die Definition von Structs, Enums und Unions ein Ausdruck, d.h. Definitionen müssen mit einem Semikolon ; abgeschlossen werden (z.B. struct {};

## Struct

Ein Struct erlaubt die Definition eines neuen Datentyps der eine Menge an Werten, von einem bestimmten Typ, zusammenfasst. Structs werden mit dem struct Schlüsselwort deklariert. Der Inhalt eines Structs wird dabei in geschweifte Klammern gefasst. Innerhalb der geschweiften Klammern wird jeder Wert, den ein Struct umschließt, durch einen Bezeichner bzw. Namen und einen Typen, getrennt durch ein :, deklariert. Diese Kombination aus Name und Typ wird als Feld (engl. field) bezeichnet. Nach jedem Feld folgt ein Komma (,), welches das Feld vom danach folgenden Feld trennt. Neben Feldern können Structs auch Methoden, Funktionen, Konstanten und Variablen enthalten.

```
// chapter02/color.zig

// Das Struct wird der Konstante `RgbColor` zugewiesen.
const RgbColor = struct {
    // Felder mit Standardwert `0`
    r: u8 = 0,
    g: u8 = 0,
    b: u8 = 0,

    // Constanten für die drei Grundfarben.
    // Mit `@This()` kann auf den umschließenden Container
    // zugegriffen werden.
    const RED = @This(){ .r = 255 };
    const GREEN = @This(){ .g = 255 };
    const BLUE = @This(){ .b = 255 };

    // Eine Methode ist eine Funktion die direkt auf einem
    // Objekt aufgerufen werden kann. Ihr erster Parameter
    // ist ein Instanz oder Referenz auf eine Instanz des Typen.
    pub fn add(self: @This(), other: @This()) @This() {
        // ...
    }
};
```

Mit Hilfe der Funktion @This() kann auf den umschließenden Kontext, im obigen Beispiel das Struct, welches an RgbColor<sup>27</sup> gebunden wird, zugegriffen werden.

Funktionen im allgemeinen Sinn, die innerhalb eines Structs definiert sind, werden in Methoden und (Struct-)Funktionen unterteilt. Der Unterschied zwischen beiden ist dabei subtil. Der erste Parameter einer Methode besitzt als Typ immer das Struct selbst, z.B. @This(), \*@This() oder

---

<sup>27</sup>Die gängige Konvention ist, dass Typbezeichner Camel-Case verwenden, d.h. ein zusammengesprochenes Wort beginnend mit einem Großbuchstaben.



`*const @This()`. Alternativ zu `@This()` kann auch direkt der Name verwendet werden, z.B. `pub fn getRed(self: *const RgbColor) u8 { ... }`. Methoden können direkt auf einer Instanz aufgerufen werden. Funktionen auf der anderen Seite haben als ersten Parameter nicht den Typ des Structs und werden über den Namen der Funktion aufgerufen, z.B. `RgbColor.foo()`.

Konstanten innerhalb von Structs können dazu verwendet werden um Werte, wie etwa die Länge eines kryptografischen Schlüssels oder wie oben zu sehen, gängige Farben, die im Bezug zu dem gegebenen Struct stehen im selben Scope zu deklarieren.

Um ein Struct nach seiner Definition zu verwenden, muss dieses instanziiert werden, indem für jedes Feld ein konkreter Wert angegeben wird. Das Instanziiieren erfolgt indem der Name des Struct, gefolgt von geschweiften Klammern, angegeben wird. Innerhalb der Geschweiften Klammern wird jedem Feld ein Wert zugewiesen. Alternativ kann, wie oben zu sehen ist, bei der Definition eines Structs jedem Feld ein Standardwert zugewiesen werden, der automatisch übernommen wird, sollte beim Instanziiieren des Structs kein Wert für das Feld angegeben werden.

```
// Die Zuweisung der Felder muss nicht in der selben Reihenfolge
// erfolgen, in der die Felder deklariert wurden.
const red = RgbColor{ .r = 255, .b = 0, .g = 0 };
// Angabe des Grün-Werts. Für die restlichen Felder wird der
// Standardwert `0` übernommen.
var green = RgbColor{ .g = 255 };
// Zugriff auf die Konstante `BLUE` definiert in `RgbColor`
const blue = RgbColor.BLUE;
```

Um auf ein bestimmtes Feld zuzugreifen wird Punktnotation verwendet. Um zum Beispiel auf den Rot-Wert der Konstante `red` zuzugreifen wird `red.r` verwendet. Auf die selbe Weise kann auch auf Methoden zugegriffen werden (z.B. `red.add(green)`). Im Fall von Variablen erfolgt eine (Neu-)Zuweisung von Feldern ebenfalls über die Punktnotation (z.B. `green.g = 128;`).

```
// Wir addieren die Werte zweier Farben.
const new_color = red.add(green);
```

In Zig sind alle Structs anonym. Ihr Name ist dementsprechend abhängig von ihrer Umgebung:



- Ist das Struct der initialisierende Ausdruck einer variable, so wird es nach der Variable benannt.
- Ist das Struct Teil eines return Ausdrucks, so wird es nach der Funktion benannt, von welcher es zurückgegeben wird.
- Andernfalls wird dem Struct ein Name nach dem Muster `filename.funcname.__struct_ID` zugewiesen.

## init Pattern

Sobald Sie mit Feldern arbeiten, deren Werte dynamisch alloziert werden, stellt sich schnell die Frage wie und von wem der allozierte Speicher verwaltet werden soll. Verwalten kann dabei mehrere Dinge bedeuten:

- Das initiale Allozieren von Speicher, für bestimmte Felder, beim instanziiieren eines Structs.

- Die Freigabe des Speichers, sobald das Struct nicht mehr benötigt wird.
- Die Neuuzuweisung eines Feldes vom Typ Single-Item-/Multi-Item-Pointer.

Zwar können all diese Aufgaben dem Nutzer des entsprechenden Struct-Datentyps auferlegt werden<sup>28</sup>, je nach Anzahl der Zeiger-Felder kann dies jedoch extrem mühsam werden. Außerdem laufen Nutzer Gefahr, den Speicher nicht korrekt zu managen und so Speicherfehler, wie etwa Memory-Leaks bei denen Speicher nicht mehr freigegeben wird, in ihren Code einzubauen.

Aus diesem Grund hat sich in Zig ein Konzept durchgesetzt, bei dem der dynamisch allozierte Speicher von einer Struct-Instanz selbst verwaltet wird. Structs die den Speicher ihrer Felder managen, verfügen meist über eine `init()` oder `new()` Funktion, die ein Objekt vom Typ `std.mem.Allocator` übergeben bekommt und das Struct initialisiert bzw. instanziiert, sowie eine `deinit()` Funktion, die den verwalteten Speicher wieder freigibt. Optional kann ein Struct über Setter-Funktionen für Felder verfügen, die vor der Neuuzuweisung das alte Objekt deallozieren.

```
// chapter02/managed.zig
const std = @import("std");

const String = struct {
    s: ?[]u8 = null,
    allocator: std.mem.Allocator,

    /// Erzeuge eine neue Instanz von `String` die den
    /// Speicher des Strings mit Hilfe von `allocator` verwaltet.
    pub fn init(allocator: std.mem.Allocator) @This() {
        // Wir geben an dieser Stelle ein anonymes Struct-Literal zurück, dessen
        // Typ (`String`) vom Rückgabewert der Funktion abgeleitet wird.
        return .{
            // Der Standardwert für `s` ist null, daher müssen
            // wir `s` nicht explizit initialisieren.
            .allocator = allocator,
        };
    }

    /// Deinitialisiere den referenzierten String.
    pub fn deinit(self: *@This()) void {
        // Da die Freigabe von Speicher immer erfolgreich sein muss,
        // ist der Rückgabewert void, d.h. innerhalb der Funktion
        // kann kein Fehler passieren.
        if (self.s == null) return;
        // Many-Item-Pointer werden mit `free` deinitialisiert.
        self.allocator.free(self.s?);
        // Wir weisen an dieser Stelle `s` den `null`-Wert zu um klar
        // zu machen, dass `s` kein valider Slice ist.
        self.s = null;
    }
}
```

---

<sup>28</sup>In den meisten Fällen sind das wohl Sie selbst.

```

    /// Weise dem referenzierten `String` den Wert `str` zu.
    /// Der Wert von `str` wird kopiert, d.h. der Caller behält
    /// die Ownership über `str`.
    ///
    /// Ein Aufruf dieser Funktion kann fehlschlagen, z.B. weil
    /// kein Speicher mehr zur Verfügung steht.
    pub fn set(self: *@This(), str: []const u8) !void {
        // Entweder `s` ist `null` oder es wurde bereits ein Wert gemanaged.
        if (self.s) |s| {
            // Wir reallozieren Speicher für `s`.
            const s_ = try self.allocator.realloc(s, str.len);
            @memcpy(s_, str);
            self.s = s_;
        } else {
            // Wir kopieren `str`.
            const s_ = try self.allocator.dupe(u8, str);
            self.s = s_;
        }
    }

    /// Beziehe den von `self` gemanageden String.
    pub fn get(self: *const @This()) ?[]const u8 {
        // An dieser Stelle geben wir entweder den Wert des Strings zurück oder,
        // falls dieser nicht existiert, `null`.
        return if (self.s) |s| s else null;
    }
};

```

Der Test für den obige Code kann mit **zig test chapter02/managed.zig** ausgeführt werden.

```

const allocator = std.testing.allocator;

var s = String.init(allocator);
// Sie können die untere Zeile auskommentieren um zu sehen, wie
// Sie einen Memory-Leak provozieren.
defer s.deinit();

try s.set("Hello, World!");
try std.testing.expectEqualStrings("Hello, World!", s.get().?);

try s.set("Ich liebe Kryptografie");
try std.testing.expectEqualStrings("Ich liebe Kryptografie", s.get().?);

```

Innerhalb des Tests weisen wir die mit `init` erzeugte String-Instanz der Variable `s` zu. Direkt danach platzieren wir einen `defer` Ausdruck der dafür sorgt, dass `deinit` am Ende des Blocks aufgerufen wird.



Wenn Sie wissen, dass sie ein Objekt im selben Block deinitialisieren wollen, sollten Sie sich grundsätzlich angewöhnen dies mit einem `defer`, direkt nach der Instanziierung des Objekts, zu machen. So vergessen Sie nicht, ihre Objekte auch wieder freizugeben.

Danach weisen wir `s` mittels der Setter-Funktion `set()` den String "Hello, World!" zu und überprüfen im Anschluss mit `expectEqualStrings()`, unter Verwendung des Getters `get()`, ob der String auch korrekt zugewiesen wurde. Dies wiederholen wir mit einem zweiten String "Ich liebe Kryptografie" um zu überprüfen, dass die Deallokation des alten, von `s` gemanagten, Strings statt findet, bevor der neue String zugewiesen wird.

Eine Besonderheit des `std.testing allocator` ist, dass beim ausführen eines Tests, nicht freigegebener Speicher automatisch, als Fehler, an den Aufrufer kommuniziert wird. Dadurch testen wir nicht nur das `set()` den gemanagten String neu zuweist, sondern auch den alten String freigibt.

### Anonyme Struct-Literale

Zig erlaubt es den Struct-Typ eines Literals wegzulassen. In diesem Fall wird der Typ des Structs abgeleitet. Bei der Konvertierung zu einem anderen Typen, Instanziert das Literal direkt die *Result-Location*.

```
const Point = struct { x: i32, y: i32 };
const pt: Point = .{
    .x = 16,
    .y = 16,
};
```

### Result-Location

Ein Konzept, das nicht nur Structs betrifft, auf welches ich an dieser Stelle trotzdem eingehen weil es die Instanziierung von Structs betrifft, ist das Konzept von *Result-Locations*.

Bestimmten Ausdrücken in Zig wird eine sogenannte Result-Location zugewiesen, d.h. ein Zeiger auf einen Speicherbereich, in welchen das Ergebnis des Ausdrucks direkt geschrieben werden muss. Dies verhindert die Erzeugung von Kopien des Ergebnisses während der Initialisierung von Datenstrukturen. In vielen Fällen hat dies keine praktischen Auswirkungen.

Anders sieht es u.a. bei der Instanziierung von Structs mithilfe eines Literals aus. Angenommen der Ausdruck `{ .a = x, .b = y }` hat die Result-Location `ptr`. In diesem Fall hätte der Ausdruck `x` die Result-Location `&ptr.a` und `y` die Result-Location `&ptr.b`. Ohne das Konzept von Result-Locations würde der Ausdruck ein temporäres Struct auf dem Stack anlegen, um dieses im Anschluss an die Zieladresse zu kopieren. Anders ausgedrückt, Zig zerlegt den Ausdruck `foo = { .a = x, .b = y }` in zwei gesonderte Ausdrücke `foo.a = x;` und `foo.b = y;`.

Ein klassisches Beispiel bei dem dies zum Verhängnis werden kann, ist beim Tauschen von Feldern eines Structs oder Arrays.

```
var arr: [2]u32 = { 1, 2 };
arr = { arr[1], arr[0] };
```

Da keine temporären Wert für `arr[0]` und `arr[1]` gespeichert werden sieht der obige Ausdruck erst einem O.K. aus, führt jedoch zu einem unerwarteten Resultate, sollte man sich dem Konzept von Result-Location nicht bewusst sein. Die zweite Zeile ist nämlich äquivalent zu folgendem:

```
arr[0] = arr[1];
arr[1] = arr[0];
```

Zuerst wird das erste Element von `arr` mit dem Wert 2 überschrieben, welcher sich an Index 1 befindet. Danach wird das zweite Element mit dem Wert des Ersten überschrieben. Schlussendlich ist der Inhalt von Array äquivalent zu `{ 2, 2 }`.

Die Folgende Tabelle listet Ausdrücke, für die das Konzept von Result-Locations zutrifft.

Ausdruck	Result Location	Result Location für Teilausdruck
<code>const val: T = x</code>	keine	x hat die Result-Location &val
<code>var val: T = x</code>	keine	x hat die Result-Location &val
<code>val = x</code>	keine	x hat die Result-Location &val
<code>@as(T, x)</code>	ptr	keine
<code>&amp;x</code>	ptr	keine
<code>f(x)</code>	ptr	keine
<code>.{x}</code>	ptr	x hat die Result-Location &ptr[0]
<code>.{ .a = x }</code>	ptr	x hat die Result-Location &ptr.a
<code>T{x}</code>	ptr	keine (Typinitialisierer propagieren keine Result-Locations!)
<code>T{ .a = x }</code>	ptr	keine (Typinitialisierer propagieren keine Result-Locations!)
<code>@Type(x)</code>	ptr	keine
<code>@TypeInfo(x)</code>	ptr	keine
<code>x &lt;&lt; y</code>	ptr	keine

Wie aus der Tabelle zu entnehmen ist, macht es in Bezug auf Structs, sowie Arrays, einen Unterschied ob während der Initialisierung ein anonymes Struct-Literal `.{ .a = x }` angegeben wird oder nicht `T{ .a = x }`, da nur bei anonymen Struct-Literalen die Result-Location zu den Teilausdrücken propagiert.

## Tuples

Anonyme Struct-Literale ohne Feldnamen werden als *Tupel* bezeichnet. Die Felder eines Tupel werden automatisch durchnummeriert, wobei dem ersten Feld der Index 0 zugewiesen wird, dem zweiten Feld der Index 1 und so weiter.

```
const v = .{
    @as(u16, 0xcafe),
    "dave",
    true,
};
```

Es gibt zwei Möglichkeiten auf die Felder eines Tupel zuzugreifen:

1. Mithilfe der Dot-Notation `.`, wobei der Feldname in eine `@"` gefasst werden muss<sup>29</sup>, z.B. `v["@1"]`.
2. Alternativ kann auch ein Index in `[]` angegeben werden (z.B. `v[1]`), wobei der Index zu Compilezeit bekannt sein muss.

## Enum

Ein weiterer Typ der in keiner Programmiersprache fehlen darf sind Enums. Enums erlauben die Definition einer Menge an Werten, die intern durch einen numerischen Wert gedeckt sind.

```
const IPType = enum {
    IPv4,
    IPv6,
};
```

Anders ausgedrückt, Enums erlauben es einer Menge an Zahlen einen Namen zuzuordnen, die in einem bestimmten Kontext Sinn ergeben.

Um einer Variable einen Enum-Wert zuzuweisen kann Punktnotation verwendet werden.

```
const ipv4 = IPType.IPv4;
```

Normalerweise wird der, zu jedem Enum-Wert gehörende, numerische Wert, sowie dessen Typ, von Zig festgelegt. Sollte ein bestimmter Zahlentyp benötigt werden, kann dieser in runden Klammern, hinter dem enum Schlüsselwort, angegeben werden.

```
// Ein Enum gedeckt durch ein `u8`
const IPType = enum(u8) {
    IPv4,
    IPv6,
};
```

Weiterhin kann jedem Enum-Wert explizit ein numerischer Typ zugewiesen werden.

```
const IPType = enum(u8) {
    IPv4 = 4,
    IPv6 = 6,
};
```

---

<sup>29</sup>Namen innerhalb von `@"` können immer als Identifier verwendet werden, was Zahlen und Strings mit Leerzeichen einschließt!

Dabei muss nicht jedem Wert explizit ein numerischer Wert zugewiesen werden. Für nicht zugewiesene Werte legt Zig automatisch einen numerischen Wert fest.

Genau wie Stucts können auch Enums Methoden enthalten. Dabei gelten die selben Regeln, d.h. der erste Parameter muss den Typ des umschließenden (e.g. `IPType`) enthalten.

```
const IPType = enum(u8) {
    IPv4 = 4,
    IPv6 = 6,

    pub fn isIPv4(self: @This()) bool {
        // Der Typ muss nicht immer explizit angegeben werden. In diesem Beispiel
        // reicht es einen Punkt (`. `) gefolgt vom Namen des Enum-Werts anzugeben.
        // Dies wird als Enum-Literal bezeichnet.
        return self == .IPv4;
    }
};
```

Enums können in switch Statements verwendet werden. Wie auch bei anderen Typen muss dabei darauf geachtet werden, dass alle Fälle abgedeckt werden oder alternativ ein else Zweig verwendet wird.

```
const ip = IPType.IPv4;
const desc = switch (ip) {
    IPType.IPv4 => "a IPv4 address",
    IPType.IPv6 => "a IPv6 address",
};
std.log.info("{s}", .{desc});
```

Mit `@intFromEnum` kann ein Enum-Wert in seine numerische Repräsentation umgewandelt werden. Diese Operation kann nicht fehlschlagen, da jeder Enum-Wert von einer Zahl gedeckt wird.

```
std.debug.assert(@intFromEnum(IPType.IPv4) == 4);
```

Die Inverse Funktion zu `@intFromEnum` ist `@enumFromInt`. Mit ihr kann ein Integer in einen Enum-Wert umgewandelt werden. Da nicht jede Zahl mit einem Enum-Wert in Beziehung stehen muss kann diese Operation fehlschlagen.

```
std.debug.assert(@as(IPType, @enumFromInt(4)) == IPType.IPv4);
```

Enums haben einen besonderen Bezug zu Unions, welche wir uns als nächstes genauer anschauen werden.

## Union

Unions sind nutzerdefinierte Typen die in sich mehrere verschiedene Typen vereinen können. Die verschiedenen Typen, die ein Union in sich vereint, werden als Liste an Feldern definiert. Zu einem bestimmten Zeitpunkt kann für eine Instanz immer nur ein Feld aktiv sein.

```
const IPAddr = union {
    IPv4: [4]u8,
    IPv6: [8]u16,
};
```

```
// Die Variable `ipv4` bindet einen Wert vom Typ `IPAddr` wobei
// das Feld `IPv4` des Unions aktiv ist.
const ipv4 = IPAddr{ .IPv4 = .{127, 0, 0, 1} };
```

Der Speicher, den ein Union benötigt, ist abhängig von dem Union-Feld mit dem größten Speicherbedarf. Im obigen Beispiel benötigt das IPv4-Feld  $4 * 8 \text{ Bit} = 32 \text{ Bit} = 4 \text{ Byte}$  und das IPv6-Feld benötigt  $8 * 16 \text{ Bit} = 128 \text{ Bit} = 16 \text{ Byte}$ . Demnach benötigt jede Instanz von IPAddr immer 16 Byte an Speicher, unabhängig davon welches Feld aktiv ist.

Um Unions in switch-Statements verwenden zu können, müssen sogenannte Tagged-Unions verwendet werden. Diese können definiert werden, indem nach dem union Schlüsselwort, in runden Klammern, ein Enum angegeben wird, dessen Felder sich mit den Feldern des Unions überschneiden.

```
const IPAddr = union(IPType) {
    IPv4: [4]u8,
    IPv6: [8]u16,
};

const ipv4 = IPAddr{ .IPv4 = .{127, 0, 0, 1} };

switch (ipv4) {
    .IPv4 => |v| std.log.info("{d}.{d}.{d}.{d}", .{v[0], v[1], v[2], v[3]}),
    .IPv6 => |_| std.log.info("a IPv6 address", .{});
}
```

Innerhalb eines switch-Statements kann nach dem => eine Variable innerhalb von | | angegeben werden, an welche der Wert der Union-Instanz gebunden werden soll. Wird der Wert nicht benötigt, so kann anstelle einer Variable auch ein \_ angegeben werden.

Soll der Wert eines Unions innerhalb eines switch-Statements modifiziert werden, so muss der Variable, an welche der Wert des Unions gebunden werden soll, ein \* vorangestellt werden.

```
switch (ipv4) {
    .IPv4 => |*v| v.* = .{ 192, 168, 13, 128 },
    .IPv6 => {}; // Eine weitere Möglichkeit diesen Zweig zu ignorieren
}
```

Weiterhin können Unions, genau wie Structs und Enums, über Methoden verfügen.

```
const IPAddr = union(IPType) {
    IPv4: [4]u8,
    IPv6: [8]u16,

    pub fn isIPv4(self: @This()) bool {
        return switch (self) {
            .IPv4 => true,
            else => false,
        };
    }
};
```



---

# Speicherverwaltung

Im Vergleich zu anderen Sprachen, wie etwa Java oder Python, muss der Speicher in Zig manuell verwaltet werden. Dies bringt einige Vorteile mit sich, birgt aber auch Risiken, die bei Nichtbeachtung zu Schwachstellen in den eigenen Anwendungen führen können. Was Zig von anderen Sprachen mit manueller Speicherverwaltung hervorhebt ist die explizite Verwendung und Verwaltung von Allokatoren, in der Programmiersprache repräsentiert durch den `Allocator` Typ. Dies kann von anderen Programmiersprachen kommenden Entwicklern anfangs ungewohnt vorkommen, bietet jedoch ein hohes Maß an Flexibilität, da Speicher zur Laufzeit dynamisch von verschiedenen Speicherquellen alloziert werden kann.

## Grundlagen

In den meisten Fällen kann ein Programm von zwei verschiedenen Quellen Speicher allozieren, dem Stack und dem Heap. Wird eine Funktion aufgerufen, so alloziert diese Speicher auf dem Stack der von den lokalen Variablen und Parametern zur Speicherung der zugehörigen Werte verwendet wird. Dieser, von einer Funktion allozierte, Speicherbereich wird als Stack-Frame bezeichnet. Die Allokation eines Stack-Frames wird durchgeführt, indem der Wert eines speziellen CPU-Register, der sog. Stack-Pointer welcher auf das Ende des Stacks zeigt, verringert wird. Die Anzahl an Bytes um die der Stack-Pointer verringert werden muss um alle lokalen Variablen halten zu können wird vom Compiler zur Compilezeit berechnet und in entsprechende Assemblerinstruktionen übersetzt.

Durch die Einschränkung, dass die Größe eines Stack-Frames zur Compilezeit bekannt sein muss, lassen sich bestimmte Aufgaben schwer lösen. Angenommen Sie wollen eine Zeichenkette unbekannter Länge von Ihrem Programm einlesen lassen, um diese später zu verarbeiten. Eine Möglichkeit um die Zeichenkette zu speichern wäre innerhalb der `main` Funktion eine Variable vom Typ `Array` mit fester Länge zu deklarieren, jedoch ist dieser Ansatz sehr unflexibel da Sie in dem gegebenen Szenario die Länge der zu erwartenden Zeichenkette nicht kennen. Bei besonders kurzen Zeichenketten verschwenden Sie ggf. Speicher während sich besonders lange Zeichenketten nicht einlesen lassen, da nicht genügend Speicher auf dem Stack alloziert wurde. Um Probleme solcher Art besser lösen zu können, kann Speicher dynamisch zur Laufzeit eines Programms alloziert werden. Der Heap kann als linearer Speicherbereich betrachtet werden, der von einem Allokator verwaltet wird. Wird Speicher zur Laufzeit benötigt, so kann der Allokator durch einen Funktionsaufruf angewiesen werden eine bestimmte Menge an Bytes zu allozieren.

Der Allokator sucht ein Stück Speicher mit der passenden Länge heraus, markiert dieses als alloziert und gibt einen Zeiger auf den Beginn des Speicherbereichs zurück. Wird der Speicher nicht mehr benötigt, so kann der Allokator durch einen weiteren Funktionsaufruf aufgefordert werden den Speicher wieder frei zu geben. In C und C++ verwenden Sie i.d.R. `malloc` und `free` um Speicher zu allozieren bzw. freizugeben, in den wenigsten Fällen müssen Sie sich jedoch Gedanken um den zu verwendenden Allokator machen. Im Gegensatz dazu verwenden Sie in Zig immer explizit einen Allokator.

In vielen Fällen, vor allem als Neuling, ist die Unterscheidung zwischen den vielen verschiedenen Arten von Allokatoren, welche die Zig Standardbibliothek bereitstellt, weniger interessant. Wird ein Standard-Allokator, im Sinne von `malloc` und `free`, benötigt, so kann auf den `GeneralPurposeAllocator` zurückgegriffen werden.

```
const Gpa = std.heap.GeneralPurposeAllocator(.{});  
var gpa = Gpa{};  
const allocator = gpa.allocator();
```

Die Funktion `GeneralPurposeAllocator` erwartet ein Konfigurationsstruct als Argument zur Compilezeit und gibt einen neuen `GeneralPurposeAllocator`-Typ zurück der der Konstante `Gpa` zugewiesen wird. In den meisten Fällen kann durch Verwendung von `.{}` als Argument die Standardkonfiguration übernommen werden. Danach kann der `Gpa` Allokator-Typ verwendet werden um ein neues Allokator-Object zu erzeugen und an die Variable `gpa` zu binden. Durch Aufruf der `allocator()` Funktion auf dem Objekt kann schlussendlich auf den eigentlichen Allokator zugegriffen werden. Dies mag auf den ersten Blick kompliziert wirken, vor allem im Vergleich zu anderen Sprachen wo Funktionen wie `malloc()` scheinbar immer zur Verfügung stehen, in den meisten Fällen reicht es aber aus, den Allokator einmal am Anfang der Anwendung zu instanziiieren. Danach kann dieser zur Allokation von Speicher verwendet werden. Der `Allocator`-Typ erlaubt es verschiedene Allokatoren durch das selbe, standardisierte Interface zu verwenden. Das bedeutet, dass Entwickler von Bibliotheken bzw. Modulen das gesamte dynamische Speichermanagement durch einen Typen (`Allocator`) handhaben können, während die Verwender von besagten Bibliotheken die freie Wahl bezüglich des dahinter liegenden Allokators besitzen.

Beim Allozieren von Speicher wird in Zig grundsätzlich zwischen der Allokation von exakt einem Objekt und der Allokation mehrerer Objekte unterschieden. Soll Speicher für genau ein Objekt alloziert werden, so muss `create()` zum allozieren und `destroy()` zum Freigeben des Speichers verwendet werden. Andernfalls können die Funktionen `alloc()` und `free()` verwendet werden. Die Funktion `create()` erwartet einen Typen (`type`) als Argument und alloziert daraufhin Speicher für exakt eine Instanz dieses Typen. Eine Allokation kann jedoch fehlschlagen, z.B. weil kein ausreichender Speicher auf dem Heap vorhanden ist. Aus diesem Grund gibt `create()` nicht direkt einen Zeiger auf den allozierten Speicher zurück, sondern einen Fehler-Typ. Damit werden Entwickler gezwungen sich bewusst zu machen, dass eine Allokation fehlschlagen kann. Dies spiegelt sich auch im Zig-Zen wieder, in welchem es u.a. heißt: „Resource allocation may fail; resource deallocation must succeed“ (auf Deutsch: Die Allokation von Ressourcen kann fehlschlagen; die deallokation von Ressourcen muss gelingen).

```
// chapter03/hello_world.zig  
const std = @import("std");
```

```

const Gpa = std.heap.GeneralPurposeAllocator(.{});
var gpa = Gpa{};
const allocator = gpa.allocator();

pub fn main() !void {
    const T = u8;
    const L = "Hello, World".len;

    // Hier allozieren wir Speicher für L Elemente vom Typ `u8` .
    const hello_world = allocator.alloc(T, L) catch {
        // Im Fall, dass der Speicher nicht alloziert werden kann,
        // geben wir eine Fehlermeldung aus und beenden den
        // Prozess ordnungsgemäß.
        std.log.err("We ran out of memory!", .{});
        return;
    };
    // Defer wird vor dem Verlassen der Funktion ausgeführt.
    // Es ist 'good practice' Speicher mittels `defer` zu
    // deallozieren.
    defer allocator.free(hello_world);

    // Wir kopieren "Hello, World" in den allozierten Speicher.
    @memcpy(hello_world, "Hello, World");

    // Nun geben wir den String auf der Kommandozeile aus.
    const stdout_file = std.io.getStdOut().writer();
    var bw = std.io.bufferedWriter(stdout_file);
    const stdout = bw.writer();

    try stdout.print("{s}\n", .{hello_world});
    try bw.flush();
}

```

Das obige Programm gibt „Hello, World” auf der Kommandozeile aus, jedoch allozieren wir vor der Ausgabe, zur Veranschaulichung, Speicher für den auszugebenden String auf dem Heap. Die Funktion `alloc()` erwartet als Argument den Typ, für den Speicher alloziert werden soll (`u8`), sowie die Anzahl an Elementen. Aus dem Typ `T` und der Anzahl `L` berechnet sich die Anzahl an Bytes die benötigt werden um `L` mal den Typ `T` im Speicher zu halten (`@sizeof(T) * L`). Wie bereits erwähnt kann die Speicherallokation fehlschlagen, aus diesem Grund müssen wir denn Fehlerfall berücksichtigen bevor wir auf den Rückgabewert von `alloc()` zugreifen können<sup>30</sup>. Da `alloc()` Speicher für mehr als ein Objekt alloziert, gibt die Funktion anstelle eines Zeigers auf den Typ `T` einen Slice vom Typ `T` zurück. Ein Slice ist ein Wrapper um einen Zeiger, der zusätzlich die Länge des referenzierten Speicherbereichs kodiert. Nach außen verhält sich ein Slice wie ein Zeiger in C, d.h. mit dem Square-Bracket-Operator `[]` kann auf einzelne Element zugegriffen werden, jedoch wird vor jedem Zugriff überprüft, ob der angegebene Index innerhalb des allo-

<sup>30</sup>Anstelle eines catch Blocks hätten wir an dieser Stelle auch try verwenden können.

zierten Bereichs liegt um Out-of-Bounds-Reads zu vorzubeugen. Slices ersetzen in vielen Fällen null-terminierte Strings, was dabei hilft Speicherfehlern vorzubeugen.

Ein wichtiger Punkt der zu jeder Allokation gehört ist die Deallokation des allozierten Speichers. In Zig kann diese direkt nach dem Aufruf von `alloc()` bzw. `create()` platziert werden, indem dem Aufruf von `free()` der `defer` Operator vorangestellt wird. `Defer` sorgt dafür, dass vor dem Verlassen eines Blocks, im obigen Beispiel ist dies der Funktionsblock von `main`, alle `defer` Blöcke ausgeführt werden und zwar in umgekehrter Reihenfolge in der sie deklariert werden. Dies ist vor allem zum Aufräumen von Ressourcen sehr hilfreich.



Sehen Sie beim Lesen von Zig-Code keinen `defer` Block zur Bereinigung von Speicher direkt nach einer Allokation sollten Sie erst einmal stutzig werden. Es gibt aber auch Situationen, z.B. bei der Verwendung eines `ArenaAllocators`, in denen nicht jede einzelne Allokation manuell bereinigt werden muss. In solchen Fällen ist es aber durchaus nützlich für Leser Ihres Quellcodes, wenn Sie durch ein Kommentar ersichtlich machen, dass das Fehlen einer Deallokation beabsichtigt ist.

## Lifetimes

Bei der Verwendung von Programmiersprachen mit manuellem Speichermanagement ist die Berücksichtigung der Lifetime (Lebenszeit) von Objekten essenziell um Speicherfehler zu vermeiden. Die Lifetime eines Objekts beschreibt ein abstraktes Zeitintervall zur Laufzeit, in welchem ein bestimmtes Objekt oder eine Sequenz von Objekten im Speicher existieren und auf diese zugegriffen werden darf. Die Art wie bzw. wo Speicher für ein Objekt alloziert wird hat dabei großen Einfluss auf dessen Lebenszeit. Im Allgemeinen beginnt die Lifetime eines Objekts mit dessen Erzeugung und endet wenn der Speicher des Objekt wieder freigegeben wird. Bezogen auf die Art der Allokation kann grob zwischen den folgenden Fällen unterschieden werden:

- Statische Allokation
- Automatische Allokation
- Dynamische Allokation

## Static Memory

In Zig, wie auch in C, befinden sich statische Variablen und Konstanten, die im globalen Scope, bzw. im Fall von Zig in einem Container<sup>31</sup>, einer Anwendung deklariert werden, in der `.data` oder `.bss` Section eines Programms. Speicher für diese Sektionen wird beim Start eines Prozesses gemapped und er bleibt bis zur Terminierung des Prozesses valide. Variablen die dies betrifft haben eine statische Lifetime, d.h. sie sind vom Start eines Prozesses bis zu dessen Beendigung valide. Selbes gilt für statische, lokale Variablen.

```
const std = @import("std");
```

---

<sup>31</sup>Ein Container in Zig ist jedes Konstrukt, das als Namensraum (engl. namespace) dient. Dazu zählen u.a. Structs aber auch Source Dateien.

```
const hello = "Hello, World";

pub fn main() void {
    const local_context = struct {
        var x: u8 = 128;
    };

    std.log.info("{s}, {d}", .{ hello, local_context.x });
}
```

Die Konstante `hello` wird im umschließenden Container, dargestellt durch die Quelldatei, deklariert und ist damit zum einen statisch, zum anderen ist sie aufgrund des `const` Modifiers zur Compilezeit bekannt. Selbes gilt für die lokale, statische Variable `x`. Im Gegensatz zu C werden statische, lokalen Variablen nicht mit dem `static` Keyword deklariert sondern innerhalb eines lokalen Structs welches ebenfalls einen Container darstellt. Lokale, statische Variablen können nützlich sein um z.B. einen gemeinsamen „Shared-State“ zwischen Aufrufen der selben Funktion zu verwirklichen.



In Zig wird jede Translationunit, d.h. jede Datei in der Quellcode liegt, als Struct und damit als Container betrachtet. Dementsprechend gibt es eigentlich keine global deklarierten Variablen wie man sie aus C kennt, sondern nur statische Variablen die in Containern deklariert werden.

## Automatic Memory

Objekte die durch Deklaration bzw. Definition innerhalb eines (Funktions-)Blocks erzeugt werden erben ihre Lifetime von dem umschließenden Block. Für Variablen und Parameter von Funktionen bedeutet dies, dass sich ihre Lifetime an der Lifetime eines Stack-Frames orientiert. Bei jedem Funktionsaufruf wird für den Aufruf ein Stück zusammenhängender Speicher auf dem Stack alloziert (der Stack-Frame) welcher groß genug ist um alle lokalen Variablen und Parameter zu halten. Der Frame wird dabei durch zwei Spezialregister der CPU, dem Stack-Pointer (SP) und dem Base-Pointer (BP), eingegrenzt. Der Stack-Pointer zeigt dabei auf das Ende vom Stack.



Es gibt verschiedene Arten von Stacks, jedoch ist die wohl häufigst auftretende Form der Full-Descending-Stack. Das bedeutet, dass der Stack nach unten „wächst“ (Descending), d.h. von höheren zu niedrigeren Speicheradressen, und der Stack-Pointer auf das erste valide Element des Stacks zeigt (Full).

```
// chapter03/stack_01.zig
const std = @import("std");

pub fn main() !void {
    var i: usize = 0; // Begin der Lifetime von 'i' --/
                        //                               /
    try foo(&i); // foo referenziert 'i' /
} // Ende der Lifetime von 'i' -----/
```

```

pub fn foo(a: *u64) !void { // Begin der Lifetime von 'a' --/
    a.* += 1;                //                               /
} // Ende der Lifetime von 'a' -----/

```

Das obige Programm übergibt eine Referenz auf die Variable `i` als Argument an die Funktion `foo()`, welche `i` inkrementiert. Die Lifetime der Variable startet mit ihrer Definition und endet mit dem Funktionsblock von `main`. Innerhalb der Lifetime darf `i` von anderen Programmteilen, in diesem Fall der Funktion `foo()`, referenziert und ggf. modifiziert werden.

Die Lifetime der Referenz `a` zu `i` beginnt mit dem Funktionsblock von `foo()` und endet mit dem Ende des Funktionsblocks. Wichtig ist, dass die Lifetime einer Referenz immer innerhalb der Lifetime des referenzierten Objekts liegen muss. Überschreitet die Lifetime einer Referenz die Lifetime des referenzierten Objekts so spricht man von einem `dangling Pointer` (auf Deutsch hängender Zeiger). Die Verwendung solcher `dangling Pointer` können zu schwerwiegenden Programmfehlern führen, da der referenzierte Speicher als undefiniert gilt.

Um das Verhalten der Anwendung besser nachvollziehen zu können, besteht die Möglichkeit mithilfe des Programms `objdump` die kompilierte Anwendung zu disassemblieren: **`objdump -d -M intel stack_01`**. Der Eintrittspunkt einer jeden Anwendung ist dabei die `main` Funktion.

```

00000000010349b0 <stack_01.main>:
10349b0: push    rbp           ; Begin Prolog ---|
10349b1: mov     rbp, rsp       ;                |
10349b4: sub     rsp, 0x10      ; End Prolog ----|
10349b8: mov     QWORD PTR [rbp-0x8], 0x0 ; i = 0
10349bf:         00
10349c0: lea     rdi, [rbp-0x8] ; &i
10349c4: call    10348e0 <stack_01.foo>
10349cb: add     rsp, 0x10      ; Begin Epilog --|
10349cf: pop     rbp           ; End Epilog ----|
10349d0: ret

```

Jede Funktion besitzt ein Symbol (im Fall von `main` ist dies `stack_01.main`) welches repräsentativ für die Adresse der ersten Instruktion steht. Beim Funktionsaufruf wird diese Adresse in das Instruktions-Zeiger-Register (Instruction Pointer - IP) geschrieben, welcher immer auf die nächste auszuführende Instruktion zeigt. Jede Funktion beginnt mit dem sogenannten Funktions-Prolog, welcher einen neuen Stack-Frame für den Funktionsaufruf erzeugt, und endet mit dem Funktions-Epilog, welcher den Stack-Frame wieder entfernt, d.h. den Stack in den Zustand vor dem Funktionsaufruf zurückversetzt.

Im Prolog wird zuerst der Zustand des Base Pointers (BP), welcher auf den oberen Teil des derzeitigen Stack-Frames zeigt, auf den Stack gepushed, um diesen im Epilog wieder herstellen zu können. Danach wird der BP mit dem Wert des SP überschrieben, d.h. BP und SP zeigen beide auf den alten BP auf dem Stack. Danach werden 16 Bytes (`0x10`) für die Variablen und Parameter von `main` auf dem Stack alloziert, indem der SP um die entsprechende Anzahl an Bytes verringert wird. Innerhalb des allozierten Speicherbereichs wird die Variable `i` mit dem Wert `0` initialisiert.

Die Adresse der Variable `i` (`BP - 8`) wird an die Funktion `foo()` mittels des `RDI` Registers übergeben<sup>32</sup>.



In einer kompilierten Anwendung existieren Variablen nur implizit, d.h. es gibt keine Symbole oder ähnliches mit denen z.B. die Variable `i` klar identifiziert werden kann. In Assembler ist eine Variable lediglich ein Bereich im (Haupt-)Speicher in welchem der Wert der Variable gespeichert ist.

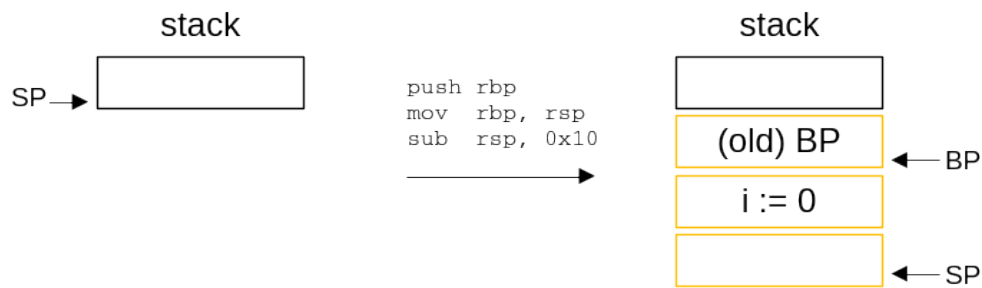


Abbildung 3: Stack-Frame von `main()`

```
00000000010348e0 <stack_01.foo>:
10348e0: push    rbp
10348e1: mov     rbp, rsp
10348e4: sub     rsp, 0x20
10348e8: mov     QWORD PTR [rbp-0x18], rdi
10348ec: mov     QWORD PTR [rbp-0x8], rdi
10348f0: mov     rax, QWORD PTR [rdi]
10348f3: add     rax, 0x1
10348f7: mov     QWORD PTR [rbp-0x10], rax
10348fb: setb    al
10348fe: jb      1034902 <stack_01.foo+0x22>
1034900: jmp     1034924 <stack_01.foo+0x44>
1034902: movabs  rdi, 0x101ed99
1034909:         00 00 00
103490c: mov     esi, 0x10
1034911: xor     eax, eax
1034913: mov     edx, eax
1034915: movabs  rcx, 0x101dfb0
103491c:         00 00 00
103491f: call    1034940 <builtin.default_panic>
1034924: mov     rax, QWORD PTR [rbp-0x18]
1034928: mov     rcx, QWORD PTR [rbp-0x10]
103492c: mov     QWORD PTR [rax], rcx
103492f: xor     eax, eax
1034931: add     rsp, 0x20
```

<sup>32</sup>Wer mehr über Assembler-Programmierung lernen möchte, dem empfehle ich das Buch „x86-64 Assembly Language Programming with Ubuntu“ von Ed Jorgensen. Diese ist öffentlich zugänglich und bietet einen sehr guten und verständlichen Einstieg.

```

1034935: pop    rbp
1034936: ret

```

Nach dem Aufruf von `foo()` wird zuerst ein neuer Stack-Frame für den Funktionsaufruf erzeugt. Innerhalb dieses Stack-Frames wird der Parameter `a` mit der Adressen von `i` initialisiert. Was auffällt ist, dass die in RDI gespeicherte Adresse gleich mehrmals auf den Stack geschrieben wird und zusätzlich direkt dereferenziert wird um den Wert von `i` in das Register RAX zu laden. Schaut man sich jedoch den gesamten Funktionskörper an so sieht man, dass von der Speicherstelle `BP - 24` die Adresse von `i` zum Zurückschreiben des inkrementierten Werts geladen wird. Damit ist `BP - 24` in diesem Fall der Parameter `a`. Dementsprechend sieht der Stack nach aufruf von `foo()` wie folgt aus.

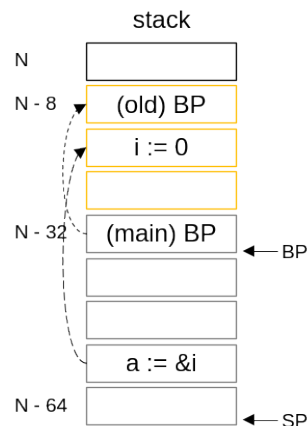


Abbildung 4: Stack-Frame von `foo()`

Diese im Funktionsprolog vollzogenen Schritte werden vor dem Verlassen der Funktion, im Epilog, umgekehrt, d.h. beim ausführen der `ret` Instruktion befindet sich der Stack, bezogen auf sein Layout, im selben Zustand wie vor dem Funktionsaufruf. Was sich natürlich geändert hat ist der Wert der Variable `i`.

## Dynamic Memory

Wir haben uns die Allokation von dynamischem Speicher anhand des `GeneralPurposeAllocator` am Anfang dieses Kapitels schon etwas angeschaut. Die Lifetime von dynamisch allozierten Objekten ist etwas tückischer als die von statisch oder automatisch allozierten. Der Grund ist, dass bei komplexeren Programmen sowohl die Allokation als auch die Deallokation eines Objekts an verschiedenen Stellen im Code passieren kann, z.B. abhängig von einer Bedingung.

Ein Beispiel hierfür ist eine verkettete Liste, bei der alle Element dynamisch auf dem Heap alloziert werden. Bezogen auf die Allokation würde es in diesem Szenario mindestens eine Stelle geben und zwar der Bereich des Codes, in dem ein neues Listenelement erzeugt wird. Bei der Deallokation eines Elements muss zumindest unterschieden werden, ob ein Element aus der Liste entfernt wird oder ob die gesamte Liste, zum Ende des Programms, dealloziert werden soll, wobei letzteres als ein Sonderfall angesehen werden kann. Ein weiterer Aspekt auf den geachtet werden muss ist, dass nach dem Löschen eines Elements der Liste, alle Referenzen auf dieses Element invalide sind, d.h. es darf nicht mehr auf den Speicher zugegriffen werden.



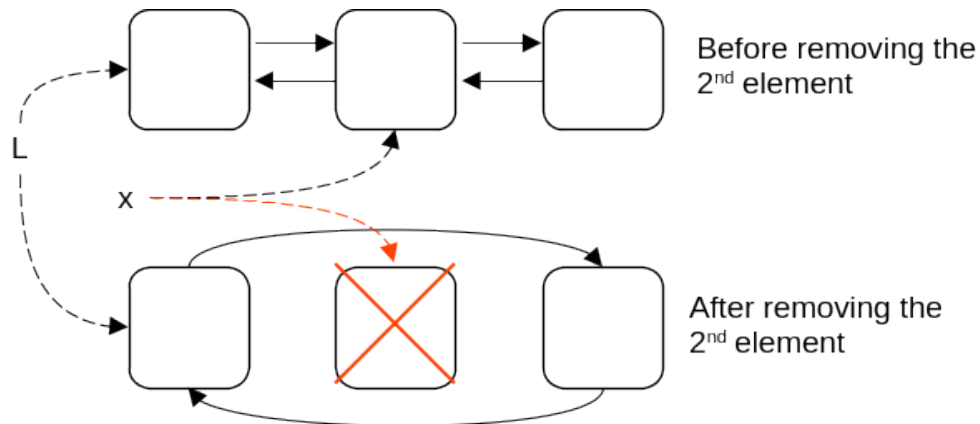


Abbildung 5: Beispiel einer verketteten Liste

```
// chapter03/linked-list.zig
const std = @import("std");

// Element einer verketteten Liste mit zwei (optionalen)
// Zeigern auf das nächste und vorherige Element.
const Elem = struct {
    prev: ?*Elem = null,
    next: ?*Elem = null,
    i: u32,

    pub fn new(i: u32, allocator: std.mem.Allocator) !*@This() {
        var self = try allocator.create(@This());
        self.i = i;
        return self;
    }
};

pub fn main() !void {
    // Hier fassen wir die Erzeugung eines neuen Allokator-Typen
    // und dessen Instanziierung in einen Ausdruck zusammen...
    var gpa = std.heap.GeneralPurposeAllocator(.{}){};
    // ... und binden dann den Allokator an eine Konstante.
    const allocator = gpa.allocator();

    // Als nächstes erzeugen wir (manuell) eine verkettete
    // Liste mit drei Elementen.
    var lhs = try Elem.new(1, allocator);
    defer allocator.destroy(lhs);
    var middle = try Elem.new(2, allocator);
    var rhs = try Elem.new(3, allocator);
    defer allocator.destroy(rhs);

    lhs.next = middle;
    middle.prev = lhs;

    middle.next = rhs;
```

```

rhs.prev = middle;

// Die Konstante L referenziert das erste Element aus
// der Liste (`lhs`).
const L = lhs;
// Ausgehend vom ersten Element geben wir alle Werte der
// Liste nacheinander aus.
std.log.info("Wert von lhs: {d}", .{L.i});
std.log.info("Wert von middle: {d}", .{L.next?.i});
std.log.info("Wert von rhs: {d}", .{L.next?.next?.i});

// Die Konstante `x` referenziert das mittlere Element...
const x = middle;
std.log.info(
  "Wert von Elem referenziert von x vor deallokation: {d}",
  .{x.i});

// ... welches als nächstes aus der Liste (manuell) entfernt wird.
lhs.next = middle.next;
rhs.prev = middle.prev;
allocator.destroy(middle);
// ... ab diesem Zeitpunkt ist `x` ein dangling Pointer und
// darf nicht mehr dereferenziert werden...

// ... wir machen es trotzdem aber der Wert des referenzierten
// Objekts ist ab diesem Zeitpunkt undefiniert.
std.log.info(
  "Wert von Elem referenziert von x NACH deallokation: {d}",
  .{x.i});
}

```

In folgendem Beispiel erzeugen wir eine verkettete Liste mit drei Elementen. Als nächstes definieren wir eine Konstante L, die das erste Element der Liste lhs referenziert und geben nach und nach, durch Dereferenzierung, die Werte aller drei Elemente aus. Da weder lhs, middle noch rhs bis zum Zeitpunkt der Ausgabe dealloziert wurden, ist die Dereferenzierung erlaubt. Wie Sie vielleicht gesehen haben ist die dritte Ausgabe, die letzte Stelle an der L dereferenziert wird. Damit überschreitet die Lifetime von L zwar theoretisch die Lifetime von lhs, middle und rhs, in der Praxis spielt dies für die Korrektheit der Anwendung jedoch keine Rolle. Anders sieht es mit der Konstanten x aus. Zwischen der ersten und zweiten Dereferenzierung von x wird middle aus der Liste entfernt und dealloziert. Damit ist x ab der Deallokation von middle ein dangling Pointer was zum Problem wird, da x später noch einmal dereferenziert wird. Auch nach der Deallokation zeigt x weiterhin auf eine existierende Speicherstelle, jedoch wurde diese durch die Deallokation freigegeben, d.h. die Daten an dieser Stelle sind undefiniert. Das hält Zig jedoch nicht davon ab den referenzierten Speicher als Instanz von Elem zu interpretieren, was sich auch in der Kommandozeilenausgabe widerspiegelt.

```

$ ./linked-list
info: Wert von lhs: 1

```

```

info: Wert von middle: 2
info: Wert von rhs: 3
info: Wert von Elem ... von x vor deallokation: 2
info: Wert von Elem ... von x NACH deallokation: 2863311530

```

Diese Art von Speicherfehler wird als Use-After-Free bezeichnet und kann unter den richtigen Bedingungen von Angreifern genutzt werden, den Kontrollfluss des Programms zu übernehmen, sollte es für den Angreifer möglich sein die Speicherstelle zu kontrollieren.

Etwas das Sie sich grundsätzlich Angewöhnen sollten ist, Referenzen die Sie nicht mehr benötigen zu invalidieren. Eine Möglichkeit dies zu tun ist anstelle eines Pointers einen optionalen Pointer zu verwenden.

```

var x: ?*Elem = middle;
// ...
lhs.next = middle.next;
rhs.prev = middle.prev;
allocator.destroy(middle);
x = null; // wir invalidieren x direkt nach der Deallokation

```

## Häufige Fehler

Im Gegensatz zu speichersicheren (engl. memory safe) Sprachen wie etwa Rust, bietet Zig einige Fallstricke, die das Leben als Entwickler schwer machen können, aber nicht müssen! In diesem Abschnitt werden wir uns einige davon näher anschauen und ich zeige Ihnen, wie Zig Ihnen dabei hilft sicheren Code zu schreiben.

### Speicherzugriffsfehler (Access Errors)

Speicherzugriffsfehler sind eine typische Fehlerquelle und haben in der Vergangenheit schon zu einigen Exploits geführt. Allgemein handelt es sich dabei um einen Oberbegriff für Programmierfehler, durch die unzulässig auf eine Speicherstelle zugegriffen wird. Zu den Speicherzugriffsfehlern gehören der Buffer-Overflow, Buffer-Over-Read, Invalid-Page-Fault und Use-After-Free. Den Use-After-Free haben wir uns im Kontext von Lifetimes schon angeschaut, an dieser Stelle möchte ich Ihnen die verbleibenden Fehler etwas näher bringen.

#### Buffer Overflow/ Over-Read

Der Buffer-Overflow und Buffer-Over-Read sind nah miteinander verwandt, kommen jedoch jeweils mit ihren eigenen Problemen. Beim Buffer-Overflow wird Speicher außerhalb eines validen Bereichs beschrieben. Dies ist meist das Resultat der unzureichenden Überprüfung der Grenzen eines Objekts, z.B. eines Arrays. Ein klassisches Beispiel ist ein Array, dass an einer Stelle indiziert wird die außerhalb der Grenzen des Arrays liegt.

```

var x: [10]u8 = .{0} ** 10;
x[10] = 1; // Index 10 ist out-of-bounds -> buffer overread!

```

Diese Art von Fehlern können genutzt werden um Daten von naheliegenden Objekten oder sogar Adressen zu überschreiben. In der Vergangenheit wurde diese Art von Fehler von Angreifern genutzt um Schadcode in Anwendungen einzuschleusen, die Rücksprungsadresse zu überschreiben und so die Kontrolle über den Prozess zu übernehmen. Moderne Compiler injizieren deswegen

sogenannte Stack-Canaries, einen randomisierten Wert der von einem Angreifer nicht erraten werden kann und der vor der Rückkehr in die aufrufende Funktion überprüft wird, in Stack-Frames die potenziell von einem Buffer-Overflow betroffen sein könnten. Ist ein Stack-Frame von einem Buffer-Overflow betroffen und wurde die Rücksprungadresse überschrieben, so bedeutet dies, dass auch der Canary überschrieben wurde. In diesem Fall wird der Prozess zur Sicherheit beendet. Wie das Zig-Zen so schön sagt: „Laufzeit-Crashes sind besser als Bugs“ (engl. „Runtime crashes are better than bugs“).

Im obigen Fall wird der Buffer-Overflow schon zur Compile-Zeit erkannt, da Arrays eine zur Compile-Zeit bekannte Länge besitzen (Zig-Zen: „Compile errors are better than runtime crashes“).

```
error: index 10 outside array of length 10
  x[10] = 1;
```

Allozieren wir den Speicher jedoch dynamisch so kann der Compiler uns nicht mehr vor unserem Fehler bewahren.

```
var x = try allocator.alloc(u8, 10);
x[10] = 1;
```

Da wir in Zig jedoch in den meisten Fällen mit Arrays oder Slices arbeiten und nicht mit rohen Zeigern und Zig für beide Datenstrukturen die Grenzen bei einem Speicherzugriff überprüft, wird der Buffer-Overflow zumindest zur Laufzeit erkannt und der Prozess beendet (Zig-Zen: „Runtime crashes are better than bugs“).

```
thread 7940 panic: index out of bounds: index 10, len 10
buffer-overflow.zig:12:6: 0x1037424 in main (buffer-overflow)
  x[10] = 1;
```

## Invalid Page Fault

Moderne Betriebssysteme schirmen den Hauptspeicher ab und stellen jedem Prozess stattdessen virtuellen Speicher zur Verfügung, der aus Sicht des Prozesses nahezu unbegrenzt ist. Bei Bedarf werden Teile des Hauptspeichers, sogenannte Pages, in den virtuellen Adressraum eines Prozesses gemapped, wodurch dieser den Speicher lesend, so wie schreibend, nutzen kann.

Ein Invalid-Page-Fault ist ein Fehler bei dem ein Prozess versucht auf einen Speicherbereich zuzugreifen, der nicht durch eine Page gedeckt wird. Ein klassisches Beispiel hierfür ist die NULL-Pointer-Exception, bei dem innerhalb eines Programms versucht wird einen NULL-Zeiger zu dereferenzieren. Grund hierfür ist, dass die meisten Betriebssysteme keine physische Page auf den Speicherbereich mappen, der die Adresse 0x0000000000000000 mit einschließt. Wird eine Page-Fault durch das Betriebssystem erkannt, so erzeugt der Kernel einen Segmentation-Fault (Schutzverletzung) für den betroffenen Prozess, welcher diesen abnormal beendet. Jeder der schon einmal in C programmiert hat wird den entsprechenden Fehler "sigsegv segmentation fault" schon einmal auf der Kommandozeile gesehen haben.

## Use After Free

Bei einem Use-After-Free wird Speicher, nach dem Ende dessen Lifetime, verwendet. Grundsätzlich hält Zig sie nicht davon ab, durch einen dangling Pointer, auf Speicher zuzugreifen, nachdem

dieser wieder freigegeben wurde. Eine Möglichkeit das Risiko für Use-After-Free Bugs in Ihrem Code zu reduzieren, ist durch die Verwendung von optionalen Zeigern `?*T` bzw. `?*const T`.

## Zusammenfassung

Bei dem Arbeiten mit Referenzen bzw. Slices sind zwei Fragen von essenzieller Bedeutung: Umschließt die Lifetime des referenzierten Objekts die der Referenz und wenn nein, habe ich dafür gesorgt, dass nach dem Ende der Lifetime des Objekts nicht mehr versucht wird auf dieses zuzugreifen. Fall Sie diese Fragen nicht beantworten können besteht eine hohe Wahrscheinlichkeit, dass sich Speicherfehler in Ihren Code einschleichen.