

#### IN DIESEM KAPITEL

Definieren, was ein Algorithmus ist

Algorithmen verwenden, um mithilfe von Computern Lösungen zu berechnen

Den Unterschied zwischen Aufgaben und Lösungen verstehen

Daten aufbereiten, um Lösungen zu finden

## Kapitel 1

# Grundlegendes über Algorithmen

**W**ie die große Mehrheit aller Menschen sind womöglich auch Sie ein wenig verwirrt, wenn Sie dieses Buch öffnen, um Ihre Abenteuerreise über Algorithmen anzutreten. Die meisten Texte definieren weder, was ein Algorithmus genau ist, noch erklären sie, wofür Algorithmen nützlich sind. Stattdessen nehmen sie an, dass Sie bereits mit Algorithmen vertraut sind und lediglich zur Vertiefung oder Erweiterung Ihres Wissens mehr zum Thema lesen wollen. Viele Bücher enthalten verwirrende Definitionen, die alles andere als Klarheit verschaffen und Algorithmen einem abstrakten, numerischen oder symbolischen Ausdruck gleichsetzen.

Im ersten Teil dieses Kapitels wird erklärt, was der Ausdruck *Algorithmus* eigentlich bedeutet und welchen Nutzen Sie aus dem Gebrauch von Algorithmen ziehen können. Algorithmen sind keine Magie, sondern werden tagtäglich und überall eingesetzt. Mit größter Wahrscheinlichkeit benutzen auch Sie seit Jahren Algorithmen, die Ihren Alltag erleichtern, ohne sich dessen bewusst zu sein. In Wirklichkeit bilden Algorithmen mittlerweile das Rückgrat, das unsere zunehmend komplexe und technische Gesellschaft aufrechterhält und lenkt.

Dieses Kapitel beleuchtet auch, wie Computer mithilfe von Algorithmen Lösungen finden, was der Unterschied zwischen Aufgaben und Lösungen ist und wie sich Datenverarbeitungsmethoden effektiv in der Auffindung von Lösungen einsetzen lassen. Das Ziel dieses Kapitels ist es, den Unterschied zwischen Algorithmen und anderen Methoden, die häufig mit Algorithmen verwechselt werden, zu verdeutlichen. Kurz: Sie werden verstehen, wieso Algorithmen so wesentlich sind und wie sie sich auf Daten anwenden lassen.

## Algorithmen beschreiben

Seit Tausenden von Jahren lösen Menschen Algorithmen von Hand. Je nachdem, wie komplex die zu lösende Aufgabe ist, kann dieser Prozess sehr zeitaufwendig sein und etliche numerische Berechnungen beinhalten. Algorithmen dienen dem Auffinden von Lösungen: je schneller und einfacher, desto besser. Zwischen den mathematischen Algorithmen, die von Genies wie Euklid, Newton oder Gauß entwickelt wurden, und den modernen Algorithmen, die an Universitäten sowie privaten Forschungs- und Testlabors entwickelt werden, besteht jedoch ein signifikanter Unterschied, der hauptsächlich im Gebrauch von Computern begründet ist. Computer können Rechenprozesse erheblich beschleunigen. Der Einsatz immer leistungsfähigerer Computersysteme hat in der Entwicklung neuer Algorithmen bereits einen großen Fortschritt bewirkt. Sie mögen bemerkt haben, dass Lösungen heutzutage immer schneller gefunden werden, was teilweise daran liegt, dass Computerleistung stetig zunimmt und gleichzeitig bezahlbar bleibt. Computer können komplexe Aufgaben mithilfe von Algorithmen lösen – manchmal in Form einer Spezialhardware – und sind mittlerweile allgegenwärtig.

In der Arbeit mit Algorithmen betrachtet man einen Input, den gewünschten Output sowie den Prozess (eine Abfolge von Aktionen), der den gewünschten Output zum gegebenen Input liefert. Wird nicht berücksichtigt, wie Algorithmen in der Praxis funktionieren, werden die wesentlichen Begriffe womöglich falsch verstanden und Algorithmen nicht richtig angewandt. Der dritte Teil des Kapitels widmet sich daher Algorithmen in Bezug auf die reale Welt. Hier werden die Begriffe untersucht, die dem Verständnis von Algorithmen dienen und die verdeutlichen, dass die Realität oft nicht ganz so perfekt ist. Durch die realistische Beschreibung von Algorithmen lassen sich allzu hohe Erwartungen herunterschrauben; so kann genauer überlegt werden, was ein Algorithmus in Wirklichkeit leisten kann.

Algorithmen können sehr vielseitig sein. Dieses Buch will einen Überblick verschaffen, wie Algorithmen unser Leben verändern und bereichern. Daher liegt der Fokus auf Algorithmen, die für die Datenverarbeitung durch einen Computer mit der nötigen Rechenleistung verwendet werden. Für die Algorithmen in diesem Buch muss der Dateninput stets in einer bestimmten Form vorliegen. Dies hat mitunter zur Folge, dass Daten verarbeitet werden müssen, um die Anforderungen des Algorithmus zu erfüllen. Dabei ändert sich der Inhalt nicht, wohl aber die Form und Art ihrer Darstellung. So macht der Algorithmus neue Muster sichtbar, die zuvor nicht erkennbar, jedoch bereits in den Daten vorhanden waren.

In der Literatur werden Algorithmen oftmals entweder zu kompliziert oder schlichtweg falsch und verwirrend dargestellt. Gerne werden Begriffe mit Algorithmen verwechselt, die eigentlich keine sind. Obwohl sich mitunter alternative Definitionen auffinden lassen, definiert dieses Buch die wichtigsten Begriffe wie folgt:

- ✓ **Gleichung:** Zahlen und Symbole, die, als Ganzes betrachtet, für einen bestimmten Wert stehen. Eine Gleichung enthält stets ein Gleichheitszeichen, das verdeutlicht, dass die Zahlen und Symbole den Wert auf der anderen Seite des Gleichheitszeichens ergeben. Allgemein bestehen Gleichungen aus Informationen über Variablen, die in Form von Symbolen ausgedrückt werden, müssen jedoch nicht unbedingt Variablen enthalten.

- ✓ **Formel:** Eine Kombination aus Zahlen und Symbolen, die Informationen oder Ideen ausdrückt. Formeln stehen normalerweise für ein mathematisches oder logisches Konzept wie zum Beispiel die Definition des größten gemeinsamen Teilers (ggT) zweier ganzer Zahlen. Allgemein zeigen sie eine Beziehung zwischen zwei oder mehreren Variablen auf. Für die meisten Menschen ist eine Formel eine besondere Form einer Gleichung.



- ✓ **Algorithmus:** Eine Abfolge von Schritten, die der Lösung eines Problems dient. Dieser Vorgang ist insofern spezifisch, als dass er eine ganz bestimmte Lösung für die Aufgabenstellung liefert. Ein Algorithmus muss nicht mathematischer oder logischer Natur sein, obgleich die Beispiele in diesem Buch oftmals in diese Kategorie fallen, weil Algorithmen in der Regel auf diese Weise verwendet werden. Bei einigen Spezialformeln handelt es sich auch um Algorithmen, beispielsweise bei der Mitternachtsformel zur Lösung von quadratischen Gleichungen. Um als Algorithmus gelten zu können, muss ein Prozess die folgenden Eigenschaften haben:
  - ✓ **Endlichkeit:** Früher oder später muss der Algorithmus die Aufgabe lösen. In diesem Buch werden Aufgabenstellungen behandelt, deren Lösung bekannt ist. So können Sie beurteilen, ob der Algorithmus die Aufgabe richtig gelöst hat.
  - ✓ **Wohldefiniertheit:** Die Abfolge der Schritte muss präzise und leicht verständlich sein. In der Implementierung von Algorithmen kommen Computer zum Einsatz; um einen brauchbaren Algorithmus zu erzeugen, müssen Computer also alle erforderlichen Schritte nachvollziehen können.
  - ✓ **Effektivität:** Ein Algorithmus muss für jeden Fall, den die Aufgabenstellung vorsieht, Ergebnisse berechnen können. Er sollte stets die Aufgabe lösen, für die er entwickelt wurde. Obwohl dabei Fehler auftreten können, sind diese eher eine Seltenheit und tauchen nur in Situationen auf, die im Rahmen des beabsichtigten Einsatzes akzeptabel sind.

Vor diesem Hintergrund sollen die folgenden Abschnitte Algorithmen genauer beleuchten. Ziel ist hier nicht eine genaue Definition dessen, was ein Algorithmus ist. Vielmehr soll genauer dargelegt werden, wie Algorithmen in das Gesamtbild passen, damit Sie ein eigenes Verständnis entwickeln und nachvollziehen können, weshalb Algorithmen so wichtig sind.

## Definitionen zur Anwendung von Algorithmen

Ein Algorithmus besteht stets aus einer Reihenfolge von Schritten zur Lösung einer mathematischen Formel, die jedoch nicht unbedingt ausgeführt werden müssen. Algorithmen können sehr umfangreich sein. Sie berechnen Aufgaben in den Naturwissenschaften, im medizinischen sowie im Finanzbereich, für die Industrieproduktion und -versorgung sowie im Informationsaustausch. Sie unterstützen uns in jedem Bereich unseres täglichen Lebens. Jede endliche, wohldefinierte und effiziente Aktionsabfolge innerhalb unseres Lebens ist ein Algorithmus. Beispielsweise steckt sogar hinter einem eigentlich trivialen und einfachen

## 34 TEIL I Erste Schritte

Toastbrot so etwas wie ein Algorithmus. Tatsächlich findet das Toasten von Brot im Informatikunterricht häufig Erwähnung, da es einen Algorithmus beschreibt:

1. Brot aus dem Schrank nehmen.
2. Brot in den Toaster stecken.
3. Toaster anschalten.
4. Wenn das Toastbrot hochfliegt, herausnehmen und buttern.
5. Essen und genießen.

Eine der wichtigsten Anwendungen von Algorithmen ist das Lösen von Formeln. Zum Beispiel lässt sich der größte gemeinsame Teiler zweier ganzer Zahlen per Hand ermitteln, indem man die Faktoren beider Zahlen aufschreibt und im Anschluss den größten Faktor auswählt, den beide Zahlen gemeinsam haben. So ist  $\text{ggT}(20, 25)$  die Zahl 5, da 5 die größte Zahl ist, die sowohl 20 als auch 25 teilt. Jeden  $\text{ggT}$  per Hand zu finden (auch eine Art von Algorithmus), ist jedoch sehr zeitaufwendig und zudem fehleranfällig. Daher entwickelte der griechische Mathematiker Euklid einen Algorithmus, der die Aufgabe über Division mit Rest effizienter löste.

Eine Formel aus Symbolen und Zahlen, die Informationen oder Ideen ausdrücken, kann mehrere Lösungen haben, von denen jede ein Algorithmus ist. So gibt es im Falle des  $\text{ggT}$  einen weiteren gebräuchlichen Algorithmus von Lehmer, der auf dem Algorithmus von Euklid beruht, ihn aber deutlich beschleunigt. Weil jede Formel auf verschiedene Arten gelöst werden kann, werden Algorithmen oft genau verglichen, um herauszufinden, welcher in einer gegebenen Situation am besten funktioniert.



Um in unserer hoch technisierten Welt, die sich immer schneller und schneller dreht, Schritt halten zu können, sind wir auf Algorithmen angewiesen. Wissenschaftliche Errungenschaften wie die Entschlüsselung des menschlichen Genoms wurden erst in unserer Zeit möglich, weil die Wissenschaftler Algorithmen schufen, die die Aufgabe schnell genug bewältigen konnten. Die Abwägung, welcher Algorithmus in einem bestimmten Fall oder in einer durchschnittlichen Gebrauchssituation besser ist, stellt ein komplexes und unter Informatikern häufig diskutiertes Thema dar.

In der Informatik kann ein und derselbe Algorithmus unterschiedliche Formen annehmen. So kann der Euklidische Algorithmus sowohl rekursiv als auch iterativ beschrieben werden. Stark verallgemeinert sind Algorithmen Methoden zur Lösung von Formeln. Es wäre jedoch ein Irrtum anzunehmen, dass es zu jeder Formel nur jeweils einen passenden Algorithmus gibt oder dass jeder Algorithmus nur eine akzeptable Darstellung besitzt. Die Verwendung von Algorithmen zur Berechnung von Lösungen unterschiedlicher Art hat eine lange Geschichte – und wurde nicht erst kürzlich neu erfunden.

Selbst wenn Sie sich auf Informatik, Datenanalyse, künstliche Intelligenz und andere technische Bereiche beschränken, werden Sie viele unterschiedliche Algorithmen antreffen – zu viele für ein einziges Buch. Beispielsweise umfasst *Die Kunst der Computerprogrammierung* von Donald E. Knuth 3.168 Seiten in vier Bänden und bringt es dennoch nicht fertig, das gesamte Thema abzudecken – der Autor wollte sogar noch weitere Bände schreiben. Hier sind einige interessante Anwendungsgebiete von Algorithmen:

- ✓ **Suchen:** Die Suche nach Informationen und die damit einhergehende Überprüfung, ob es sich bei der gefundenen Information tatsächlich um die gewünschte Information handelt, ist eine wichtige Aufgabe. Ohne diese Option wären viele Vorgänge im Internet unmöglich, beispielsweise wenn Sie eine Webseite suchen, die die perfekte Kaffeetasse für das Büro anbietet.
- ✓ **Sortieren:** Heutzutage leiden viele Menschen unter dem allgegenwärtigen Informationsüberfluss. Es ist daher wichtig, eine Reihenfolge festzulegen, in der Informationen dargestellt werden sollen. Durch das Ordnen von Informationen wird das Datenchaos reduziert. Als Kind haben Sie sicherlich gelernt, dass es leichter ist, ein bestimmtes Spielzeug unter Ihren Spielzeugen wiederzufinden, wenn die Spielzeuge ordentlich aufgeräumt und nicht überall verstreut sind. Stellen Sie sich vor, Sie finden bei Amazon Tausende von Kaffeetassen im Angebot, ohne sie nach Preis oder positiven Bewertungen sortieren zu können. Viele komplexe Algorithmen benötigen sogar eine bestimmte Datenordnung, um überhaupt verlässlich angewandt werden zu können. Sortieren ist daher eine wichtige Voraussetzung für den Lösungsfindungsprozess.
- ✓ **Umformungen:** Einen Datentyp in einen anderen umformen zu können, ist wesentlich für das Verständnis und den effektiven Gebrauch von Daten. Zum Beispiel könnte es sein, dass Sie mit dem metrischen Messsystem umgehen können, in Ihren Quellen jedoch das britische System verwendet wird. Umformungen von einem System in das andere erleichtert das Verständnis der Daten. Genauso transformiert die schnelle Fourier-Transformation (FFT) Signale von der Zeitdomäne in die Frequenzdomäne, was beispielsweise in W-LAN-Routern zum Einsatz kommt.
- ✓ **Planung:** Die angemessene Bereitstellung von Ressourcen ist ein weiterer Bereich, aus dem Algorithmen mittlerweile nicht mehr wegzudenken sind. So sind Ampeln an Straßenkreuzungen seit Langem keine simplen Vorrichtungen mehr, die die Sekunden zwischen den Lichtphasen zählen. Moderne Ampeln berücksichtigen viele weitere Aspekte wie Tageszeit, Wetterbedingungen und Verkehrsfluss. Die Planungsmethoden können hierbei unterschiedliche Formen annehmen. Bedenken Sie, dass Ihr Computer gleichzeitig mehrere Aufgaben erledigen kann. Ohne einen Planungsalgorithmus würde das Betriebssystem womöglich alle zur Verfügung stehenden Ressourcen beanspruchen und weitere Programme daran hindern, ausgeführt zu werden.
- ✓ **Kurvenanalyse:** Die Bestimmung der kürzesten Verbindung zwischen zwei Punkten hat etliche Anwendungen. Ohne diesen Algorithmus würde Ihr GPS keine Routenplanung vornehmen können, weil es Sie niemals durch die Straßen der Stadt führen könnte, ohne die kürzeste Strecke zwischen Punkt A und Punkt B auszuwählen.
- ✓ **Kryptografie:** Daten vor Hackerangriffen zu sichern, ist ein ständiger Kampf. Mithilfe von Algorithmen können Daten analysiert, verschlüsselt und am Ende wieder in ihre ursprüngliche Form gebracht werden.
- ✓ **Generieren pseudo-zufälliger Zahlen:** Stellen Sie sich ein Spiel vor, das immer gleich abläuft. Sie fangen jedes Mal am selben Ort an und machen dieselben Schritte auf dieselbe Art und Weise. Ohne die Möglichkeit der Generierung scheinbar zufälliger Zahlen wären viele Computerprozesse unmöglich.

## 36 TEIL I Erste Schritte



Diese Liste stellt lediglich eine stark gekürzte Übersicht dar. Algorithmen werden für viele Zwecke und sehr unterschiedlich verwendet, und Menschen entwickeln ununterbrochen neue Algorithmen, um sowohl bereits vorhandene als auch neue Problemstellungen zu lösen. Der wichtigste Aspekt in der Arbeit mit Algorithmen ist der folgende: Ein bestimmter Input erzeugt einen daraus zu erwartenden Output. Zweitrangige Aspekte sind die Menge der nötigen Ressourcen sowie die Zeit, die für die Lösung benötigt wird. Je nach Problemstellung und gewähltem Algorithmus spielen eventuell auch Aspekte wie Genauigkeit und Konsistenz eine Rolle.

### Algorithmen sind überall

Nicht ohne Grund wurde im vorigen Abschnitt der Toasteralgorithmus genannt. Es handelt sich hierbei vermutlich um den beliebtesten Algorithmus, der jemals erdacht wurde. Bereits Grundschulkinder kennen eine eigene, äquivalente Variante des Toasteralgorithmus, ehe sie die elementarsten Mathematikaufgaben lösen können. Unzählige Varianten des Toasteralgorithmus mit entsprechenden Outputs sind denkbar. Je nach Person und Kreativität können die Ergebnisse ganz unterschiedlich ausfallen. Zusammenfassend lässt sich festhalten, dass Algorithmen in einer großen Vielfalt und an oftmals unerwarteten Stellen auftreten.

Jede mithilfe eines Computers ausgeführte Aufgabe beinhaltet Algorithmen. Einige Algorithmen sind bereits fester Bestandteil der Computerhardware; sie sind darin eingebettet – daher der Begriff *eingebettete Systeme*. Bereits das Einschalten eines Computers erfordert einen Algorithmus. Auch in Betriebssystemen, Anwendungen und anderen Softwarearten sind sie zu finden. Benutzer verlassen sich auf Algorithmen. Anleitungen beschreiben Aufgaben in aufeinanderfolgenden Schritten, jedoch könnten diese Schritte genauso gut Anweisungen oder Nutzungsbedingungen sein.

Auch Tagesroutinen gleichen Algorithmen. Überlegen Sie nur, wie Sie Ihren Tag verbringen. Geht es Ihnen wie den meisten Menschen, so besteht Ihr Alltag in erster Linie aus denselben täglich wiederholten Handlungen, die Sie in derselben Reihenfolge ausführen. Ihr Tag wird so zu einem Algorithmus, der die Aufgabe löst, erfolgreich weiterzuleben, ohne dabei unnötig Energie zu verschwenden. Denn Routine bezweckt dies: Sie macht uns effizient.

Sofortmaßnahmen beruhen häufig auf Algorithmen. Im Flugzeug nehmen Sie etwa die Notfalltafel aus der Rückenlehne vor Ihnen. Darauf abgebildet ist eine Reihe von Piktogrammen, die veranschaulichen, wie die Nottür geöffnet und die Rutschbahn ausgefahren wird. Wahrscheinlich müssen Sie nicht einmal die Sätze lesen, da Sie bereits anhand der Abbildungen verstehen, wie Sie möglichst schnell aus dem Flugzeug gelangen. In diesem Buch finden Sie für jeden Algorithmus immer dieselben drei Punkte:

1. Beschreibung der Aufgabe.
2. Aufstellen einer Abfolge von Schritten, die die Aufgabe lösen (wohldefiniert).
3. Ausführung der Schritte zum Erreichen des gewünschten Ergebnisses (endlich und effektiv).



## Mit Computern Aufgaben lösen

Der Begriff *Computer* mag auf manche Leser zu technisch und etwas überfordernd wirken, doch in Wirklichkeit bewegen wir uns tagtäglich in der Welt der Computer. Die meiste Zeit über tragen Sie mindestens einen Computer mit sich, nämlich Ihr Smartphone. Verwenden Sie ein Spezialgerät wie beispielsweise einen Herzschrittmacher, so befindet sich darin ein weiterer Computer. Ihr Smart-TV enthält mindestens einen Computer, das Gleiche gilt für Ihr Heimnetzwerk. In einem Auto können sich bis zu 30 Computer in Form von eingebetteten Mikroprozessoren befinden, die Benzinverbrauch, Motorverbrennung, Übertragung, Steuerung und Stabilität regulieren und mehr Programmcode benötigen als ein Kampfjet (laut einem Artikel in der *New York Times*, zu finden in englischer Sprache unter <http://www.nytimes.com/2010/02/05/technology/05electronics.html>). Die automatisierten Autos, die nach und nach auf dem Automobilmarkt auftauchen, werden eine noch größere Anzahl an eingebetteten Mikroprozessoren und weit komplexere Algorithmen erfordern. Computer wurden dazu erfunden, Aufgaben schnell und mit geringem Aufwand zu lösen. Es sollte Sie daher nicht überraschen, dass auch in diesem Buch Computer verwendet werden, um genauer zu erklären, was Algorithmen sind.

Computer können sich stark voneinander unterscheiden. Der Computer in Ihrer Armbanduhr ist sehr klein, während der auf Ihrem Schreibtisch recht groß ist. Supercomputer sind gigantisch und enthalten viele weitere kleine Computer, die alle darauf abgestimmt sind, gemeinsam komplexe Themen wie die Wettervorhersage für morgen zu bearbeiten. Die komplexesten Algorithmen verwenden besondere Computerfunktionen, um Lösungen für Aufgaben zu finden, die die Anwender ihnen vorgeben. In der Tat könnten hier erhebliche Ressourcen gespart werden, doch würde dies längere Wartezeiten oder ungenauere Ergebnisse bedeuten. In manchen Fällen ist die Wartezeit so endlos, dass das Ergebnis am Ende nicht mehr interessiert. Unter Berücksichtigung von Geschwindigkeit und Genauigkeit beleuchten die folgenden Abschnitte einige Besonderheiten von Computern, die Auswirkungen auf Algorithmen haben können.

## Moderne CPUs und GPUs wirksam einsetzen

Universalprozessoren (englisch: *CPUs*) dienten ursprünglich der Durchführung von Berechnungen mithilfe von Algorithmen. Aufgrund ihrer allgemeinen Natur können sie jedoch eine Vielzahl anderer Aufgaben erledigen, wie zum Beispiel Daten verschieben oder mit externen Apparaten interagieren. Prozessoren können zwar alle Schritte innerhalb eines Algorithmus ausführen, doch nicht unbedingt schnell. Die Besitzer der ersten Universalprozessoren stateten ihr System daher gerne mit mathematischen Koprozessoren aus, um die Rechenarbeit zu beschleunigen. Heutzutage sind mathematische Koprozessoren bereits in allen Universalprozessoren enthalten, sodass Sie beim Kauf eines Intel-i7-Prozessors eigentlich mehrere Prozessoren in einem erwerben.



Interessanterweise vermarktet Intel immer noch Erweiterungen für Spezialprozessoren wie den Xeon-Phi-Prozessor, der mit Xeon-Chips arbeitet. Der Xeon-Phi-Chip wird zusammen mit einem Xeon-Chip verwendet, wenn rechenintensive Prozesse wie maschinelles Lernen durchgeführt werden sollen (um zu verstehen, wie man im Rahmen des maschinellen Lernens Algorithmen einsetzt, die anhand von Daten Prognosen erstellen und Informationen sinnvoll organisieren, siehe *Maschinelles Lernen für Dummies* von John Mueller und Luca Massaron).

## 38 TEIL I Erste Schritte

Sie fragen sich womöglich, wieso in diesem Teil Grafikprozessoren (*Graphics Processing Units*, kurz: *GPUs*) erwähnt werden. GPUs haben die Aufgabe, Daten einzuspeisen, zu bearbeiten und im Anschluss auf dem Bildschirm abzubilden. Doch jede Computerhardware kann mehrere Zwecke erfüllen. So sind GPUs besonders geeignet für Datenumformungen – ein wesentlicher Bestandteil vieler Algorithmen. Ein GPU ist ein Prozessor für Spezialzwecke, der sich für die schnelle Ausführung von Algorithmen eignet. Es sollte nicht verwundern, dass Menschen, die Algorithmen entwickeln, viel Zeit mit Querdenkerei verbringen und dabei häufig unkonventionelle Lösungsansätze finden.

In der Tat sind CPUs und GPUs die am häufigsten verwendeten Chips zur Ausführung algorithmenbasierter Aufgaben. CPUs handhaben dabei meist Universalaufgaben besonders gut, während GPUs auf anspruchsvollere mathematische Aufgaben und insbesondere Datenumformungen spezialisiert sind. Der Gebrauch mehrerer Kerne ermöglicht paralleles Rechnen, also die Ausführung mehrerer Schritte zur gleichen Zeit. Werden weitere Chips hinzugefügt, erhöht dies die Anzahl der verfügbaren Kerne. Mehr Kerne erhöhen wiederum die Rechenleistung, doch aus verschiedenen Gründen steigt die Schnelligkeit hierbei nicht erheblich. Zwei i7-Chips rechnen nicht doppelt so schnell wie ein i7-Chip.

### Arbeiten mit Spezialchips

Mathematische Koprozessoren und GPUs sind Beispiele für die gängigsten Spezialchips, die jedoch nicht für Zwecke wie etwa den Systemstart eingesetzt werden. Algorithmen hingegen erfordern oftmals andere Spezialchips, um Aufgaben zu lösen. Obwohl dies kein Hardwarebuch ist, lohnt es sich dennoch, andere interessante Chips zu betrachten. So entwickelt IBM aktuell neue künstliche Neuronen. Ein solcher Speicher, der das menschliche Gehirn simulieren kann, würde neue, interessante Algorithmenprozesse ermöglichen, die es heute in dieser Form noch nicht gibt.

Neuronale Netze simulieren das menschliche Denken und ermöglichen tiefergehende Lern-techniken des maschinellen Lernens. Auch diese Technologie profitiert vom Gebrauch von Spezialchips wie dem Tesla P100 von NVidia. Diese Chips können algorithmische Prozesse nicht nur extrem schnell ausführen, sondern lernen währenddessen dazu und werden hierdurch wiederum bei jeder Iteration beschleunigt. Lernende Computer werden alsbald Roboter betreiben, die sich (so gut es geht) eigenständig bewegen können, so wie die Roboter im Film *I, Robot*. Auch gibt es Spezialchips mit Funktionen wie visueller Objekterkennung.



Ungeachtet ihrer genauen Funktionsweise werden Spezialprozessoren letztendlich eine Vielfalt von Algorithmen betreiben, und dies nicht ohne Konsequenzen für unsere reale Welt. Bereits heute finden sich einige dieser realen Anwendungen in recht einfacher Form. Stellen Sie sich beispielsweise die Schritte vor, die ein Pizza backender Roboter ausführen müsste – mitsamt der Variablen, die dabei in Echtzeit bedacht werden müssen. Dieser Robotertyp existiert bereits. Es handelt sich hier nur um ein weiteres Beispiel der vielen Industrieroboter, die mithilfe von Algorithmen unterschiedliche Produkte herstellen; auch hier bilden Algorithmen die Handlungsgrundlage und Spezialchips die Voraussetzung dafür, dass die Aktionen rasch ausgeführt werden.





Früher oder später könnte es sogar möglich sein, das menschliche Gehirn wie einen Prozessor zu verwenden und Informationen über eine besondere Schnittstelle auszugeben. Einige Firmen experimentieren derzeit bereits mit Prozessoren, die direkt in das menschliche Gehirn integriert werden und dabei die Kapazität, Informationen zu verarbeiten, steigern. Stellen Sie sich eine Welt vor, in der Menschen mit der gleichen Geschwindigkeit wie Computer Algorithmen lösen können, jedoch mit dem kreativen »Was wäre wenn«-Potenzial, das den kleinen Unterschied im menschlichen Denken ausmacht.

## Netzwerke wirksam einsetzen

Sogar mit Spezialchips lassen sich einige Algorithmen nicht ohne unbegrenzte Geldmittel einsetzen. Ein Ausweg ist hier die Vernetzung mehrerer Computer. Mithilfe von Spezialsoftware kann ein Computer, der sogenannte Master, durch einen *Agenten* (eine Art Speicher-Hintergrundprogramm, die den Prozessor verfügbar macht) auf alle Prozessoren der Slavecomputer zugreifen. Hierdurch lassen sich komplexe Aufgaben durch die Aufteilung auf weitere Slavecomputer lösen. Hat ein Netzwerkcomputer seine Aufgabe gelöst, schickt er die Ergebnisse an den Master zurück, der die Informationen wiederum zusammenfügt und in eine kompakte Form bringt; diese Technik nennt man *Computerclustern*.

Halten Sie dies nicht für Science-Fiction-Fantasien. Computerclustertechniken werden bereits sehr vielfältig eingesetzt. Der Artikel unter <http://www.zdnet.com/article/build-your-own-supercomputer-out-of-raspberry-pi-boards/> beschreibt beispielsweise sehr detailliert, wie Sie Ihren eigenen Supercomputer aus der Kombination mehrerer Raspberry-Pi-Platten zu einem einzigen Cluster zusammenbauen können.

*Verteilte Systeme* sind eine weitere Version von Computerclustern (jedoch lockerer organisiert) und auch sehr beliebt. Unter <http://www.distributedcomputing.info/projects.html> findet sich eine Liste über Projekte, die mit verteilten Systemen arbeiten. Die Projektliste beinhaltet auch große Vorhaben wie die Suche nach extraterrestrischer Intelligenz (*Search for Extraterrestrial Intelligence, SETI*). Es besteht auch die Möglichkeit, die ungenutzte Rechenleistung Ihres Computers der Krebsforschung zur Verfügung zu stellen. Die Liste der möglichen Projekte ist erstaunlich. Durch Netzwerke kann auch die Rechenleistung externer Geräte unabhängig genutzt werden. So können Sie beispielsweise über die Computer von Amazon Web Services (AWS) und anderen Anbietern Aufgaben erledigen lassen. Durch den Fernzugriff über eine Netzwerkverbindung werden die externen Computer Teil Ihres Netzes. Netzwerke können also mit einer Vielzahl unterschiedlicher Algorithmen arbeiten und Aufgaben lösen, die mit einem Computer schier unlösbar wären.

## Daten effektiv nutzen

Ein Großteil der Arbeit mit Algorithmen hat rein gar nichts zu tun mit Rechenleistung, kreativem und unkonventionellem Denken oder Hardware. Um Lösungen zu finden, sind auch verlässliche Daten nötig, auf deren Grundlage das Endergebnis berechnet wird. Beispielsweise muss für den Toasteralgorithmus bekannt sein, dass Brot, Toaster, Strom für den Toaster und so weiter vorhanden sind, ehe die eigentliche Aufgabe des Toastens ausgeführt werden

## 40 TEIL I Erste Schritte

kann. Daten spielen eine große Rolle, da der Algorithmus nicht ausgeführt und keine Lösung gefunden werden kann, wenn auch nur ein Element fehlt. Eventuell sind auch zusätzliche Inputdaten notwendig. So könnte die Person, die das Brot toastet, beispielsweise keinen Roggen mögen. Ist dies der Fall, wird das Vorhandensein von Brot trotzdem nicht zum Erfolg führen, wenn es sich um Roggenbrot handelt.

Daten entstammen unterschiedlichen Quellen und existieren in vielen verschiedenen Formen. Man kann Daten über eine Quelle, zum Beispiel einen Echtzeitmonitor, ausgeben lassen, auf eine öffentliche Datenbank zugreifen, sich auf private Datenquellen stützen, Daten von Webseiten sammeln oder auf etliche andere Weisen erhalten, die wir hier nicht alle aufzählen können. Daten können statisch (unveränderlich) oder dynamisch (sich ständig verändernd) sein, vollständig oder mangelhaft. Es kann vorkommen, dass Daten nicht in der richtigen Form vorliegen (zum Beispiel in britischen statt metrischen Gewichtseinheiten) oder dass sie in Tabellenform sind, obwohl Sie sie in einer anderen Form bräuchten. Eventuell sind sie nicht strukturiert (etwa in einer NoSQL-Datenbank oder in mehreren Dateien), obwohl eine formale Formatierung wie in einer relationalen Datenbank angebracht wäre. Kurz: Es müssen alle möglichen Details über die Daten Ihres Algorithmus bekannt sein, ehe damit Aufgaben gelöst werden können.



Daten haben zahlreiche Formen und Verarbeitungsmöglichkeiten. Daher ist dieses Buch ausführlich diesem Thema gewidmet. Das 6. Kapitel befasst sich damit, wie Daten strukturiert sind. Im 7. Kapitel werden wir anfangen, Daten nach gewünschten Informationen zu durchsuchen. Kapitel 12 bis 14 erläutern, wie man mit Big Data arbeitet. In der Tat findet sich in jedem Kapitel dieses Buches die eine oder andere datenspezifische Information, da kein Algorithmus ohne Daten auskommt.

## Zwischen Aufgaben und Lösungen unterscheiden

In diesem Buch werden zwei Teile der algorithmischen Sicht auf die Welt vorgestellt. Einerseits gibt es *Aufgaben*, die gelöst werden müssen. Eine Aufgabe kann den gewünschten Output eines Algorithmus beschreiben oder eine Hürde, die zunächst zu überwinden ist, um den gewünschten Output zu erhalten. *Lösungen* hingegen sind die Methoden oder Schritte, die implementiert werden, um die Aufgabe zu behandeln. Eine Lösung kann sich auf einen oder auf viele Schritte innerhalb des Algorithmus beziehen. Der Output eines Algorithmus, also die Antwort auf den letzten Schritt, ist in der Tat eine Lösung. Im folgenden Teil werden einige wichtige Aspekte von Aufgaben und Lösungen erklärt.

## Richtigkeit und Effizienz

Alle Algorithmen kreisen um die Frage nach einer akzeptablen Antwort. Man sucht nach einer *akzeptablen* Antwort, da einige Algorithmen mehrere Lösungen finden, etwa bei ungenauem Dateninput. Denn in Wirklichkeit sind präzise Antworten selten. Obwohl das Ziel stets eine präzise Antwort ist, erhält man oftmals nur eine akzeptable Antwort.

Die Suche nach der exaktesten Lösung kann sehr zeitaufwendig sein. Erhalten Sie endlich eine genaue Lösung, mag es bereits zu spät sein. Die Information ist mittlerweile unbrauchbar; Sie haben Zeit verschwendet. Die Wahl zwischen zwei Algorithmen, die beide dieselbe Fragestellung behandeln, bedeutet oft eine Entscheidung zwischen Geschwindigkeit und Präzision. Ein schneller Algorithmus erzeugt womöglich keine allzu genaue Lösung, die jedoch trotzdem als brauchbarer Output dienen kann.

Auch fehlerhafte Lösungen stellen ein Problem dar. Schnell gefundene, aber falsche Ergebnisse sind mindestens genauso fatal wie penibel genaue Lösungen, die langsam berechnet wurden. Der Fokus dieses Buches liegt auf dem Mittelweg zwischen schnell und langsam sowie zwischen ungenau und allzu genau. Obwohl Ihr Mathematiklehrer mit Sicherheit stets betonte, dass nur die richtige Antwort zählt, spielt sich die Mathematik der wirklichen Welt oftmals in Abwägungen von Alternativen und Mittelwegen ab, deren Auswirkungen Sie oftmals nicht abschätzen können.

## Die Erkenntnis, dass nichts umsonst ist

Es ist ein allgemein verbreiteter Mythos, dass jeder nur erdenkliche Computeroutput erzeugen kann, ohne dass allzu viel Mühe in die Herleitung der Lösung gesteckt werden muss. Bedauerlicherweise gibt es keine absolut gültigen Lösungen auf Fragestellungen, und bessere Ergebnisse sind oftmals kostspielig. Im Zuge der Arbeit mit Algorithmen werden Sie schnell feststellen, dass für schnelle, genaue Ergebnisse meist zusätzliche Ressourcen nötig sind. Größe und Komplexität der verwendeten Datenquellen können die Lösungsmenge stark beeinflussen. Mit zunehmender Größe und Komplexität steigt auch der Bedarf an Ressourcen.

## Die Strategie an die Aufgabe anpassen

Teil V dieses Buches untersucht mögliche Strategien, um Kosten zu reduzieren, wenn Sie mit Algorithmen arbeiten. Die besten Mathematiker verwenden Tricks, um mit weniger Rechenarbeit mehr Output zu erzeugen. Beispielsweise können Sie einen einzigen Algorithmus zum Lösen einer Aufgabe erstellen oder eine Menge simpler Algorithmen einsetzen, um dieselbe Aufgabe zu lösen, jedoch unter Verwendung mehrerer Prozessoren. Viele einfache Algorithmen sind normalerweise schneller und besser als ein einzelner, komplexer Algorithmus, obwohl dieser Ansatz der Intuition widersprechen mag.

## Algorithmen in einer Lingua franca beschreiben

Algorithmen bilden eine Basis für menschliche Kommunikation – auch wenn die Individuen, die sie verwenden, unterschiedliche Ansichten haben und verschiedene Sprachen sprechen. So ist zum Beispiel der Satz von Bayes (die Wahrscheinlichkeit eines Ereignisses unter bestimmten Umständen)

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

## 42 TEIL I Erste Schritte

immer der gleiche, ob Sie ihn nun auf Englisch, Spanisch, Chinesisch, Deutsch, Französisch oder in einer anderen Sprache lesen. Der Algorithmus sieht gleich aus und liefert bei gleichem Dateninput die gleichen Ergebnisse. Algorithmen dienen der Überwindung möglicher Hindernisse, die Menschen voneinander trennen, indem sie Ideen in einer Form ausdrücken, die jeder nachvollziehen und beweisen kann. Beim Lesen werden Sie die Schönheit und magische Kraft bemerken, die Algorithmen im Ausdruck selbst der subtilsten Gedanken eigen sind.



Abgesehen von universellen mathematischen Bezeichnungen werden in der Arbeit mit Algorithmen Programmiersprachen eingesetzt, die die angewandten Formeln erklären und beschreiben. Sämtliche Algorithmen finden sich in C, C++, Java, Fortran, Python (wie in diesem Buch) sowie in anderen Sprachen. Einige Autoren benutzen Pseudocode für die Beschreibung von Algorithmen, sollte dem Leser die Programmiersprache unbekannt sein. Durch *Pseudocode* können Computeroperationen in einer Sprache wie Deutsch ausgedrückt werden.

## Schwierige Aufgaben angehen

Sie können Algorithmen benutzen, um Aufgaben unterschiedlichster Komplexität zu lösen. Ein Algorithmus denkt nicht nach; er besitzt weder Gefühle noch eine Vorstellung, wie er verwendet werden will. Algorithmen lassen sich auf beliebige Art und Weise gebrauchen, solange dies der Lösung der Aufgabe zuträglich ist. Beispielsweise könnte ein Algorithmus, der aus Sicherheitsgründen anstatt einer Passworteingabe auf Gesichtserkennung spezialisiert ist, ebenso Terroristen aufspüren, die sich in einem Flughafen aufhalten, wie ein vermisstes Kind auf der Straße erkennen. Somit kann ein und derselbe Algorithmus verschiedene Anwendungen haben; allein der Benutzer entscheidet, wie er ihn einsetzen will. Es lohnt sich daher, dieses Buch sorgfältig durchzulesen; nur so können auch schwierige Fragestellungen mit oftmals überraschend simplen Algorithmen gelöst werden.

## Daten für Lösungen strukturieren

Menschen haben eine ungenaue Vorstellung von Daten; sie wenden auf dieselben Daten unterschiedliche Regeln an und ziehen daraus Schlüsse, die ein Computer niemals sehen würde. Computer haben eine strukturierte, einfache, kompromisslose und höchst unkreative Sicht auf Daten. Nachdem Menschen Daten für die Auswertung durch einen Computer aufbereitet haben, geben Algorithmen mitunter unerwartete Ergebnisse aus und erzeugen unerwünschten Output. Dies liegt daran, dass es Menschen schwerfällt, Daten so beschränkt zu betrachten, wie dies ein Computer tut. Im Folgenden werden zwei Aspekte beschrieben, die auch in den nächsten Kapiteln wiederkehren.

## Die Sichtweise eines Computers

Computer haben eine vereinfachte Sicht auf die Dinge. Trotzdem ist dies eine Sichtweise, die Menschen in der Regel fremd ist. Computer kennen ausschließlich Zahlen, weil sie nicht für die Arbeit mit anderen Datentypen konzipiert wurden. Wenn Menschen Buchstaben auf dem

Computerbildschirm sehen, nehmen sie an, dass der Computer nach dem gleichen Schema mit den Daten umgeht. Der Computer versteht aber weder die Daten noch deren Bedeutung. Der Buchstabe A ist für den Computer schlichtweg die Zahl 65, und eigentlich sogar weniger als das. Für den Computer ist lediglich eine Reihe elektrischer Impulse sichtbar, die die Binärzahl 0100 0001 ergeben.

Genauso machen Computer keinen Unterschied zwischen großen und kleinen Buchstaben. Für einen Menschen ist das kleine *a* eine andere Form des großgeschriebenen A, doch für einen Computer handelt es sich hierbei um zwei völlig unterschiedliche Buchstaben. Das *a* ist für den Computer die Zahl 97 (die Binärzahl 0110 0001).

Wenn bereits diese einfachen Buchstabenvergleiche Verwirrung zwischen Menschen und Computern stiften, kann man sich leicht ausmalen, was geschieht, wenn Menschen Annahmen über andere Datentypen treffen. Beispielsweise haben Computer kein Gehör für Musik. Dennoch können Computerlautsprecher Musik produzieren. Das Gleiche gilt für Grafiken. Computer sehen eine Reihe von 0 und 1 – kein Bild, auf dem eine schöne Landschaft abgebildet ist.



Im Umgang mit Algorithmen ist es wichtig, Daten aus der Sicht eines Computers zu betrachten. Der Computer sieht ausschließlich die Zahlen 0 und 1. Auch Sie sollten Daten aus dieser Perspektive betrachten, wenn Sie mit der Arbeit am Algorithmus beginnen. Sie mögen es sogar vorteilhaft finden, dass manche Lösungen durch die Perspektive des Computers leichter auffindbar sind. Über diese Kuriosität werden Sie im Laufe des Buches mehr erfahren.

## Datenordnung muss sein

Für Computer sind Daten mit einer strengen Form und Struktur versehen. Anfangs werden Sie sicherlich feststellen, dass die Arbeit mit Algorithmen zu einem großen Teil darin besteht, Daten in eine für den Computer brauchbare Form zu bringen, sodass der Algorithmus darüber laufen kann. Obwohl der menschliche Verstand Muster in ungeordneten Daten erkennen kann, brauchen Computer präzisere Informationen, um dieselben Muster zu sehen. Diese Genauigkeit kann bislang unbekannte Muster sichtbar machen, was eine der Hauptmotivationen für die Verwendung von Algorithmen mit Computern ist. Neue Muster können aufgespürt und anschließend verwendet werden, um weitere Aufgaben auszuführen. So kann zum Beispiel das Kaufverhalten eines Kunden analysiert und die gewonnenen Informationen verwendet werden, um automatisch neue Käufe zu generieren.

